European Journal of Educational Sciences, EJES

June 2019

European Scientific Institute, ESI

The content is peer-reviewed

June 2019 Edition Vol. 6, No. 2

ISSN 1857- 6036

About The Journal

The European Journal of Educational Sciences (EJES) is a peer-reviewed international journal which accepts high quality research articles. It is a quarterly journal published at the end of March, June, September and December, and it is available to all researchers who are interested in publishing their scientific achievements. We welcome submissions focusing on theories, methods, and applications in educational sciences.

In the past few years, academicians from over 50 countries from Europe and around the globe have published their papers in the European Journal of Educational Sciences.

Authors can publish their articles after a review by our editorial board. Our mission is to provide greater and faster flow of the newest scientific thought. EJES's role is to serve as a kind of bridge between researchers around the world. EJES is open to all researchers, regardless of their geographical origin, race, nationality, religion or gender, as long as they have an adequate scientific paper in the field of educational sciences.

EJES fully supports the open access and open science concept. The full content of the papers is available on EJES website and is free for usage.

Dejan Marolov

Managing Editor

Table of Contents:

English Language Teachers' View on Assessment and their Reflection on Teaching Practices: A Mongolian Case Study

Jargaltuya Ragchaa, PhD Candidate
University of Szeged, Hungary
Dornod University, Mongolia

Doi: 10.19044/ejes.v6no2a1 URL:http://dx.doi.org/10.19044/ejes.v6no2a1

Abstract

In Mongolia, teachers' attitudes towards large-scale assessments have been a largely uncovered area. Therefore, this paper focuses on examining the general overview of English language teachers' belief about state-level assessments, and their reflection on teaching practices. Participants of this study were 307 teachers of primary and secondary school teachers, and 36 of them were English language teachers of Dornod province in Mongolia. Independent sample t-test was used to explore how English teachers change their instructions in teaching English compared to other subjects teachers. Result showed that they usually search for more effective teaching methods, take less liberty on how they design their lessons, reduce instructional content, and focus more on Educational standards. As a result of a correlation analysis, English language teachers' assessment view is significantly related to the content of the assessment that is designed by teachers in a class. Based on the results, it can be concluded that teachers focus more on the assessment content that they design for progress and final exams. They prefer to prepare students for this assessment by making them practice the test items that are similar to the school achievement test items during classes. Understanding the reasons for ineffective instructions can help policy makers and teachers to change the assessment content and its accountability, and would also help to improve their classroom instructions to have better learning outcome.

Keywords: Large-scale assessments, test-based accountability, instructional change, English language teaching.

Introduction

A lot of children are learning English language in different schools around the world. English is increasingly perceived as a basic competence to succeed in life. Mongolia has adopted English language as a second language, and schools are offering English as a main mandatory subject. English is included as one of the main subjects in school achievement tests. Majority of parents often search for schools that offers English Language and that have better quality programs for their children.

Given the recognized importance of English, language education and its assessment are changing from developing students' academic skills to the use of English in real life. Nikolov (2016) noted that one of the best programs of English, documented by recent interest, is content and language integrated learning. Johnstone (2009) and Rixon (2013, 2016) remark that this new development poses new opportunities and challenges for assessment. Nikolov (2016) added that this shift towards assessment and accountability are not limited to foreign language programs. However, there is an international trend in educational assessment for accountability in public education policies in all subjects and competencies. Assessment and its accountability have become inseparable parts of education and, based on the assessment, program accountability calls for the quality of education to be continually improved. However, recent studies indicate that in most cases, assessment is administrated to see that the implementation of standards and curriculum are being met. Based on the results of the study, assessment can be used for ranking the schools, teachers, and students in a bid to improve the teaching and learning process (Nikolov, 2016). The aim of this paper is to specify and understand what English language teachers think of state-level assessments and their usefulness, and how their instructions and test preparation strategies are changed due to their perceptions of state-level assessments.

Literature Review

Early research indicates different directions of the impacts of high stakes tests. Thus, they have both negative (anxiety and fear) and positive (changes of teaching instruction and test taking strategy) effects on learning and teaching practices. External pressure can lead teachers to critically revise their practices and adapt effective teaching strategies (Terhart, 2013). In contrast, Hamilton et al. (2002) argued that test-based accountability can also lead to negative reallocation of instructional time to focus on tested aspects of the standards to the exclusion of untested aspects of the standards. English language instructors are encouraged (Baker & Westrup, 2000) to use many methods to teach receptive skills in pre-stages and post-stages. On the other hand, Alkaff (2013) noted that students concentrate more on terminology and that they are usually tested with multiple choice questions because of limited practice on everyday interactions in the classroom.

Tran (2012) highlighted the importance of validity, reliability, practicality, equivalency, authenticity, and wash back of second language assessment. He explained that test validity needs to measure the test takers' real ability based on empirical and theoretical research. Bachman and Palmer (1996 cited in Tran, 2012) say that reliability refers to similar results when the test is administered on different occasions. Practicality refers to the relationship between the resources (human and materials, time, and location) and the use of the test. Equivalency and authenticity indicate whether or not the test is directly based on curriculum standards or instructional activities. Brown and Hudson (1998 cited in Tran, 2012) pointed out that a wash back is the reflection of testing and assessment on the language teaching curriculum and instruction. These studies show that including all of these criteria for writing tests is really important to assess students' actual skills and their learning outcome. Second language testing assesses learners' progress and their specific skills. Therefore, language instructors need to design tests to measure the learners' functional use of language, not a specific linguistic point.

Consequently, the most important thing test makers need to consider in language assessment is to understand the roles of abilities and contexts, the interactions between them, and the influence of ability and context on the performance of language assessment tasks (Fox et al., 2007). Powers (2010) observed that language receptive (reading and listening) and productive skills (speaking and writing) are assessed differently. Receptive skills are usually assessed through computer-based and paper-pencil with multiple choice items, while productive skills are assessed with performance-based tests. Language testing experts and language researchers such as Hakuta and Beatty (2000), Bailey and Butler (2003), and Garcia, McKoon and August (2006) have criticized previous English language assessments used for ESL students. This is because those assessments do not measure up with the development of the academic English language skills that students need to become successful in a school settings. Language educators noted that an interactional approach was becoming more important in language teaching and assessment. For example, Bachman (2007) and Chapelle (1998) noted that the English language program includes skill-based, trait-based, task-based, and interactional approaches in a given context. Chapelle (1998) remarks that an interactional approach to language learning improves communicative language abilities. Chalhoub-Deville (2003) noted that language competence is a process involving improvement over time in combining knowledge and context with language performance.

Across the world, English teachers have different assessment views. Language assessments can be different or similar in different countries. Rixon (2013) found that, at the end of primary school years, English language assessments were different in some countries. For instance, in France, at the end of primary school years, teachers complete an evaluation which covers five skills areas : (1) listening comprehension, (2) oral interaction, (3) individual speaking with no interaction (e.g. reproducing a model, a song, a rhyme, a phrase, reading aloud, giving a short presentation), (4) reading comprehension, and (5) writing. In Taiwan, instructors are now developing their own English proficiency tests (Rixon, 2013) at the primary school level. The purpose of their proficiency test is to assess the effectiveness of English instruction and to identify those in need of remedial teaching. In Finland, many primary schools use a voluntary “national” test of English designed by the English teachers’ association of Finland to guide their final grading of students and to get some information for them about how they are doing against the average of other schools (Rixon, 2013). A New National Curriculum in Cyprus was implemented in September 2011. It introduced English at the primary level, emphasized the role of portfolio assessment, and introduced content and language integrated learning (Rixon, 2013).

Teachers’ Instructional Change based on Assessment and Accountability

Researchers differently indicate that high stakes tests contribute to negative (anxiety and fear) and positive (changes of teaching instruction and test taking strategy) effects in teaching practices. A good dose of pressure can force teachers to adapt effective teaching strategies (Terhart, 2013). Tóth and Csapó (2011) explored how Hungarian teachers in elementary schools felt pressured by different stakeholders than their counterparts in upper secondary schools. However, they claimed that greater incentives and heightened external pressure were needed to induce school agents to raise educational quality. Hamilton et al. (2005) noted that the integration of mechanisms of educational accountability system can positively affect the quality of education. As they reported, the mechanisms—incentives, information, and assistance—are likely to affect student achievement primarily by altering what occurs in the classroom: Incentives are intended to motivate teachers to focus on the goals embodied in the standards, information from tests should provide data to guide instructional decision making, and assistance should help them improve their practice” (Hamilton et al., 2005, p.3).

According to Hamilton et al. (2002), test-based accountability can lead educators to work harder and to adopt better curricula or more-effective teaching methods. It can lead to coaching students to perform better by focusing on aspects of the test that are relevant to the domain the test is intended to represent. Due to a test-based accountability system, teachers may pay more attention to test-taking strategies. Often, multiple-choice state school achievement tests differ widely from the format used in classroom tests. Pederson and Yager (2014 in Ngang, Hong & Chanya, 2014, p.536) remarked that becoming a highly qualified teacher in today's educational system is dependent on how well teachers work together with their principals and colleagues. Through collective work, teachers explore the potential to practice more effective decision making as a skill for supporting acquisition of additional professional knowledge and skills.

A number of other studies have shown that test-based accountability programs have had a positive impact on students' test scores (e.g. Carnoy & Loeb, 2002; Jacob, 2005; Linn & Dunbar, 1990; Nichols, Glass, & Berliner, 2012).

In contrast, Boyd et al. (2008 in Fuller & Ladd, 2012, p.13) noted that teachers avoid high stakes tests that may induce teachers' anxiety of unwanted inquisition, loss of flexibility in classroom practices, a feeling of coercion to teach based on the test, and fear for their jobs. Tóth and Csapó (2011) found that in Hungary, teacher beliefs' about changes in their teaching are rather similar in elementary and lower secondary schools. However, the level of agreement in the case of many of the statements differs between elementary and upper secondary school teachers. In Hungary, teachers typically refuse to narrow down the curriculum due to the national assessment system; they focus their efforts more on students with poor results in the state tests by giving extra-curricular tutoring.

Koretz et al. (2001) found that test-based accountability has no effect, or even a negative effect on students' knowledge and skills. Hamilton et al. (2002) points out that test-based accountability can also lead to negative reallocation of instructional time to focus on tested aspects of the standards to the exclusion of untested aspects of the standards. In addition, high-stakes testing may become a barrier to the development of intrinsic motivation as its implementation is generally accompanied by a high amount of pressure on students and teachers (Moore & Waltman, 2007). Thus, the various studies reviewed above show the usefulness of test-based accountability systems. Herman and Golan (1991) noted that high-stakes testing leads to a narrowing of curricula and instruction, and such testing appears to influence teaching and learning within schools. Teachers spend most of their time and attention to increase students' test scores rather than focus on student learning. Thus, state test results, under conditions of accountability pressure, remain a critical issue to understand when designing and implementing accountability measures. Meaningful learning requires a critical approach based on the productive use of assessment in stimulating educational reform.

The NBETPP (National Board on Educational Testing and Public Policy) (2003) reports that teachers often spend more time on subjects that are tested with high stakes, and less time on non-tested subjects. Therefore, students have limited time to practice with fine arts, physical education, foreign language, and other extra-curricular activities. Similarly, Abrams, Pedulla and Madaus (2003) and Abrams (2004) conducted a survey among Florida teachers and the result showed that teachers had reallocated instructional schedules, allowing for more time to be spent on tested content while reducing the time for the material that would not appear on the large scale assessment. Hence, they reduce the time spent on fostering activities in order to prepare students for the state test.

Hadley (2010) carried out a survey on 12 school principals from eight different district schools in the state of Utah to explore their opinions about how high-stakes testing impacts teaching and learning. The findings showed that principals were concerned that teachers should teach a curriculum that would result in improved test scores. Additionally, the principals encouraged teachers to use the results of large scale assessments to guide their instruction to produce high test scores. Eslami-Rasekh and Valizadeh (2008) conducted a survey on Iranian young EFL teachers. They responded that they felt more successful in applying instructional strategies than in managing an EFL class. They also reported that their ability to motivate and engage students to learn English was not as high as their ability to use instructional strategies.

Teachers' Test Preparation Strategies based on Assessment and Accountability

Clearly, educational researchers should pay attention to teachers' test preparation strategies caused by large-scale assessment and accountability systems. In the NBETPP report (2003), teachers responded to some questions related to preparing their students for the state-level test such as test preparation methods and amount of time spent on test preparation. They stated that more time is spent due to high-stakes tests with intense preparation using materials that closely resemble the test. Also, they try to motivate their students to do well in the state test.

In addition, majority of teachers changed their assessment practices by modeling their own classroom tests following the format of the state test. Abrams et al. (2003) report that teachers from high-stakes states spend more time than do their counterparts in low-stakes states preparing students for the state test. Abrams (2004) also found that in Florida, many teachers and schools are highly stressed by the pressure to improve student test performance. Sixty-three percent of teachers indicated that the pressure was so much that they had little time to teach anything that would not appear on the test. Furthermore, majority of them reported that they found ways to raise test scores without improving learning. Hadley (2010) remarks that test subjects and test preparation activities restrict the amount of time spent on a particular subject, and the tests dictated the kind of teaching strategies used, resulting in fewer activities, less creativity, and less benefit to the students.

Methodology

Participants

The participants were 262 different subject teachers and 36 English language teachers from 19 schools in Dornod province. Dornod province lies at the eastern part of Mongolia and includes a major city, Choibalsan. Those 19 schools were in Choibalsan and in nearby villages (soums) in the surrounding metropolitan area. The subjects were 100% female with a mean age of 33.8 and a mean teaching experience of 9.4 years. 75% of them were Bachelor of Arts holders and 25% of them were MA degree holders.

Instruments

The teachers' view on educational assessment and accountability questionnaire was created based on numerous international questionnaires (Hamilton, Berend, & Stecher, 2005; Moore & Waltman, 2007). This is the questionnaire of the IPEA (International Project for the Study of Educational Accountability Systems) project. Tóth and Csapó (2011) adapted this questionnaire to fit into the Hungarian context. In adapting this version into Mongolian context, some questions related to International assessment were discarded because Mongolia is not included in some International studies such as PISA and TIMMS. The questionnaire consisted of seven blocks of questions (61 items). Each block represented a particular assessment or accountability procedure: (1) Teachers' background questions consisting of five items, (2) views on large-scale assessments consisting of 10 items, (3) test preparation strategies consisting of nine items, (4) perceived pressure for different types of assessment consisting of seven items, (5) amount of test practice consisting of three categorical items, (6) changes in instructional practice including 20 statements, and (7) perceived pressure from different stakeholders consisting of seven items. Teachers' opinions were assessed on a four point Likert scale (1=disagree; 4=agree).

Results

An independent sample t-test was used based on the acceptance of the large scale assessments to compare the means of selected 36 English teachers' opinions with other teachers' opinions. Table 1 shows that English language teachers in Dornod province have similar ideas compared to other teachers. Specifically, they think that school achievement tests should be conducted on a regular basis (M=3.6, SD=.60), that tests contribute to an increased effort in schools (M=3.4, SD=.93), that tests provide an objective basis to evaluate schools (M=3.3, SD=.79), and that these tests are important for work in schools (M=3.2, SD=.98). They also, like teachers of other subjects, somewhat disagree with the view that school achievement tests support the debate about the concept of competence (M=2.2, SD=.96), and they provide a basis for discussion among colleagues (M=2.3, SD=.96). Other teachers, however, had different opinions about the statements “School achievement tests are important for work in schools” (t=-1.16, p<.05) and “School achievement tests are not useful for my job as a teacher” (t=-1.46, p<.05).

Table 1. English language teachers’ view on school achievement tests

	School achievement tests	Groups	N	M	SD	t	P
1	Should be conducted on a regular basis	ENG	36	3.6	.60	-.60	n.s
		Other	262	3.6	.65		
2	Contribute to an increased effort in schools	ENG	34	3.4	.93	.24	n.s
		Other	253	3.3	.85		
3	Provide an objective basis to evaluate schools	ENG	35	3.3	.79	1.70	n.s
		Other	257	3.0	1.0		
4	Are important for work in schools	ENG	36	3.2	.98	-1.16	p<.05
		Other	255	3.3	.79		
5	Create more problems than solutions	ENG	35	2.7	.93	-.82	n.s
		Other	257	2.8	.92		
6	Provide a basis for discussion among colleagues	ENG	34	2.3	1.0	-1.79	n.s
		Other	252	2.6	1.1		
7	Support the debate about the concept of competence	ENG	36	2.2	.96	-.69	n.s
		Other	253	2.4	1.0		
8	Are barely applicable for individual student evaluations	ENG	34	2.1	1.1	-.29	n.s
		Other	257	2.6	1.0		
9	Only cause trouble in schools	ENG	36	2.0	.96	5.50	n.s
		Other	253	2.0	1.0		
10	Are not useful for my job as a teacher	ENG	36	1.5	.80	-1.46	p<.05
		Other	257	1.8	1.0		

Note: N= number of participants, M= mean value of participants, SD= standard deviation, t= t-value (the size of the difference between means), p= p-value (significance level), n.s=not significant.

An independent sample t-test was also used to explore which instructional changes were mostly made by English teachers in teaching English in comparison to other teachers. The results in Table 2 below show that English language teachers and other subjects teachers usually search for more effective teaching methods (M=3.8, SD=.35), take less liberty on how they design their lessons (M=3.7, SD=.62), reduce instructional content (M=3.6, SD=.60), and focus more on Educational standards (M=3.5, SD=.56). English teachers have different opinions on the statement “I search for more effective teaching methods” (t=1.32, p<.05) and “I have narrowed down the curricular content of my instruction” (t=-2.02, p<.05) compared to teachers of other subjects in general. The results suggest that English language teachers are less willing to narrow down their curricular content than teachers in other fields, and that they spend more effort searching for effective teaching methods.

Table 2. English language teachers' instructional changes

	School achievement tests	Groups	N	M	SD	T	P
1	I search for more effective teaching methods	ENG	35	3.8	.35	1.32	p<.05
		Other	246	3.7	.47		
2	I take less liberties on how I design my lessons	ENG	36	3.6	.62	.33	n.s
		Other	244	3.6	.57		
3	I reduce instructional content	ENG	36	3.5	.60	1.34	n.s
		Other	244	3.3	.74		
4	I focus more on Educational standards	ENG	36	3.5	.56	.11	n.s
		Other	247	3.5	.67		
5	My instruction focuses more strongly on competences rather than content	ENG	35	3.5	.78	1.61	n.s
		Other	247	3.5	.62		
6	I focus more strongly on multiple choice tests	ENG	36	3.4	.69	1.69	n.s
		Other	246	3.1	.83		
7	I focus more strongly on overarching competences (writing and reading in mathematics instruction)	ENG	36	3.3	.75	-1.22	n.s
		Other	239	3.4	.70		
8	I have narrowed down the curricular content of my instruction	ENG	36	3.0	.58	-2.02	p<.05
		Other	245	3.2	.72		

Note: N= number of participants, M= mean value of participants, SD= standard deviation, t= t-value (the size of the difference between means), p= p-value (significance level), n.s=not significant.

Based on confirmatory factor analysis, the following factors of changes in teachers' instructional practice were identified: (1) giving homework, (2) teaching methods, (3) content of the instruction, (4) testing strategy, and (5) teachers' attention to special students. A correlation analysis was done to identify how English teachers' view on school achievement tests (SATs) are related to their teaching practices. As a result, English language teachers' assessment view was found to be significantly related to the content of the instruction (r=.357, p<.05). The analysis also shows that teaching methods and testing strategy are correlated (r=.356, p<.05), and the content of the instruction is correlated with testing strategy (r=.334, p<.05) and attention to special students (r=.478, p<.01). These results are presented in Table 3 which follows.

Table 3. Relationship between assessment attitude and teachers' instructional changes

	Giving homework	Teaching methods	Content of instruction	the Testing strategy	Attention to special students
View on assessments					
Giving homework	.027				
Teaching methods	.137	-.133			
Content of the instruction	.357*	.058	.254		
Testing strategy	.031	-.074	.356*	.334*	
Attention to special students	.189	.037	.175	.478**	-.008

Note. * p<.05, ** p<.01.

English language teachers often use tasks in regular instruction that are similar to those in school achievement tests (M=3.6, SD=.47), discuss general task-taking strategies with students (M=3.6, SD=.58), have students practice test formats that are used in school achievement tests (M=3.5, SD=.66), and seek to improve students' motivation to do well on SATs (M=3.4, SD=.73). One significantly different statement compared to other teachers' answers was "I discuss general task-taking strategies with students" (t=1.49, p<.05). Other teachers, however, see more coherence between instructional content and tasks in SATs (t=-1.40, p<.05) and try to improve students' test taking skills (practice on public release tasks that are used in SATs) (t=-.90, p<.05) as summarised in Table 4 below.

Table 4. English language teachers' test preparation strategies

	Test preparation strategies	Groups	N	M	SD	T	P
1	I more often use tasks in regular instruction that are similar to those in school achievement test	ENG	34	3.6	.47	.21	n.s
		Other	253	3.6	.64		
2	I discuss general task-taking strategies with students	ENG	34	3.6	.58	1.49	p<.05
		Other	256	3.5	.81		
3	I practice test formats that are used in school achievement test	ENG	35	3.5	.66	-.55	n.s
		Other	255	3.5	.64		
4	I seek to improve students' motivation to do well on SATs	ENG	36	3.3	.72	-.52	n.s
		Other	258	3.4	.73		
5	I see to it that coherence between instructional content and the tasks of the SAT is increased	ENG	35	3.2	1.0	-1.40	p<.05
		Other	252	3.5	.74		
6	I try to improve students' test taking skills (practice on public release tasks that are used in SATs)	ENG	35	3.2	1.1	-.90	p<.05
		Other	260	3.4	.79		
7		ENG	35	2.1	1.0	-1.15	n.s

I set aside or put less emphasis, in regular instruction, on content that will not be tested	Other	259	2.3	1.0

Note: N= number of participants, M= mean value of participants, SD= standard deviation, t= t-value (the size of the difference between means), p= p-value (significance level), n.s=not significant.

Conclusion

Since English language is an important subject included in school achievement tests, it is important that English language teachers should believe that state-level assessments are important for their work and that the results of the assessment are linked with the school and teachers' efforts. The state-level assessments cause English teachers to focus more on the assessment content and influence their design of progress and final exams. English teachers also prefer to prepare students for the assessments by practicing test items that are similar to the school achievement test items during class. The results of this study will help give insights into the issues behind the teaching and learning process of English language education in Mongolia. It is important to explore the reasons behind ineffective teaching and learning strategies and their effect on learning achievement, and how English language instruction has been changing due to the educational assessment and accountability system in Mongolia. An independent sample t-test was used for exploring the frequencies and differences between the perceptions of assessment and accountability, and their instructional changes. The main results indicated that English language teachers think state-level assessments are important for improving the quality of language education since English language is included in school achievement tests. They also think it is better to conduct these assessments regularly.

However, they believe that these assessments are aimed only at evaluating schools, not for developing individuals' learning outcomes. English language teachers try to use more effective teaching methodologies even though they already do not have enough time to prepare their lessons due to their work load and the different types of assessments. Therefore, they reduce their instructional content and focus more on preparing students for exams. In addition, their view on the importance of large scale assessments influences the content of the assessment that they design for progress and final tests in their classes. Thus, this may be the reason why English teachers prefer to ask students to practice on the test formats that are used in the school achievement tests during class.

References:

1. Abrams, L. M., Pedulla, J. J., & Madaus, G. F. (2003). Views from the classroom: Teachers' opinions of statewide testing programs. *Theory into Practice, 42*(1), 8–29.
2. Abrams, L. M. (2004). Teachers' Views on High-Stakes Testing: Implications for the Classroom, Arizona State University Education Policy Studies Laboratory, Working Paper EPSL0401-104-EPRU.
3. Alkaff, A.A. (2013). Students' Attitudes and Perceptions towards Learning English, *AWEJ*, Vol.4. pp.106-121.
4. Bachman, L. F., & Palmer, A. S. (1996). Language testing in practice: Designing and developing useful language tests, Oxford England: Oxford University Press.
5. Bachman, L. F. (2007). What is the construct? The dialectic of abilities and contexts in defining constructs in language assessments. In J. Fox, M. Wesche, D. Bayliss, L. Cheng, C. E.Turner, & C. Doe (Eds.), *Language testing reconsidered* (pp. 41–71). Ottawa, Canada: Ottawa University Press.
6. Bailey, A. L., & Butler, F. A. (2003). *An evidentiary framework for operationalizing academic language for broad application to K–12 education*: A design document (CSE Technical Report No. 611). Los Angeles: University of California, National Center for Research on Evaluation, Standards, and Student Testing (CRESST). *classes with few resources*. London: Continuum.
7. Baker, J. & Westrup, H. (2000). *The English language teacher's handbook: How to teach large classes with few resources*. London: Continuum.
8. Boyd, D., Grossman, P., Hammerness, K., Lankford, H., Loeb, S., McDonald, M., Reininger, M., Ronfeldt, M., & Wyckoff, J. (2008). Surveying the landscape of teacher education in New York city: Constrained variation and the challenge of innovation. *Education Evaluation and Policy Analysis*, 30(4), 319-343.
9. Brown, J. D., & Hudson, T. (1998). The alternatives in language assessment. TESOL Quarterly, 32 (4), 653-675.
10. Carnoy, M., & Loeb, S. (2002). Does external accountability affect student outcomes? A cross-state analysis.*Educational Evaluation and Policy Analysis*, *24*(4), 305–331.
11. Chalhoub-Deville, M. (2003). Second language interaction: Current perspectives and future trends, *Language Testing*, *20,* 369–383.
12. Chapelle, C. A. (1998). Construct definition and validity inquiry in SLA research. In L. F. Bachman & A. D. Cohen (Eds.), *Second language acquisition and language testing interfaces* (pp. 32–70). Cambridge, England: Cambridge University Press.

13. Eslami-Rasekh, Z. & Valizadeh, K. (2008). Teachers' sense of self-efficacy, English proficiency, and instructional strategies: A study of nonnative EFL teachers in Iran. TESL EJ, 11(4), 1-19. Retrieved October 9, 2008 from http://tesl-ej.org/ej44/a1.pdf.
14. Fox, J., Wesche, M., Bayliss, D., Cheng, L., Turner, C.E., & Doe, C. (2007). *Language testing reconsidered*, University of Ottawa Press, Ottawa.
15. Fuller, S. C. & Ladd, H. F. (2012).*School Based Accountability and the Distribution of Teacher Quality Among Grades in Elementary Schools.* Panel Session for the Association for Education Finance and Policy, Boston, MA, March.
16. Garcia, G. E., McKoon, G., & August, D. (2006). Synthesis: Language and literacy assessment. In D. August & T. Shanahan (Eds.), *Developing literacy in second-language learners: Report of the National Literacy Panel on language minority children and youth* (pp. 269–318). Mahwah, NJ: Lawrence Erlbaum.
17. Hamilton, L. S., Berends, M., & Stecher, B. (2005). Teachers' Responses to Standards-Based Accountability. Working Paper. RAND Corporation, Santa Monica.
18. Hamilton, L, S., Stecher, B. M., & Klein, S. P. (2002). *Making Sense of Test-Based Accountability in Education.* RAND Corporation, Santa Monica.
19. Hakuta, K., & Beatty, A. (2000).Testing English-language learners in U.S. schools. Washington, DC: National Academies Press.
20. Herman, J., & Golan, S. (1991). Effects of Standardized Testing on Teachers and Learning—Another Look. Los Angeles, CA: National Center for Research on Evaluation, Standards and Student Testing, UCLA.
21. Jacob, B. A. (2005). Accountability incentives and behavior: The impact of high stakes testing in the Chicago public schools. *Journal of Public Economics*, *89*(5–6), 297-327.
22. Johnstone, R. (2009). An early start: What are the key conditions for generalized success? In Koda, K. (1996). *L2 word recognition research: A critical review. Modern Language Journal*, 80, 14, 450–460.
23. Koretz, D., Mccaffrey, D., & Hamilton, L. (2001). *Toward a Framework for Validating GainsUnder High-Stakes Conditions*. CSE Technical Report.No. 551. Los Angeles, CA: Center for the Study of Evaluation, Universityof California.
24. Linn, R. L., & Dunbar, S. B. (1990). The Nation's report card goes home: Good news and bad about trends in achievement. *Phi Delta Kappan, 72* (2), October, 127–133.

25. Moore, J. L. & Waltman, K. (2007).*Pressure to Increase Test Scores in Reaction to NCLB: An Investigation of Related Factors*, Annual Meeting for the American Educational Research and Evaluation Association.
26. NBETPP Report (2003). Perceived effects of state-mandated testing programs on teaching and learning, Joseph J. Pedulla, Lisa M. Abrams, George F. Madaus, Michael K. Russell, Miguel A. Ramos, and Jing Miao Lynch, School of Education, Boston College.
27. Ngang, T. K., Hong, C. S., & Chanya, A. (2014). Collective work of novice teachers in changing teaching practices, *Procedia-Social and Behavioral Sciences*, 116 (21), pp. 536-540.
28. Nichols, S., Glass, G., & Berliner, D. (2012). High-stakes testing and student achievement: Updated analyses with NAEP data. *Education Policy Analysis Archives, 20* (20), 1–35.
29. Nikolov, M. (2016). A framework for young EFL learners' diagnostic assessment: Can dostatements and task types. In M. Nikolov (Ed.), *Assessing young learners' of English: Global and local perspectives* (pp. 65-92). New York: Springer. 15
30. Powers, D. E. (2010). The case for a comprehensive, four skills assessment of English language proficiency.*TOEIC Compendium study*,12.pp. *2-11*.
31. Rixon, S. (2013).*British Council Survey of policy and practice in primary English languag teaching worldwide*. London: British Council.
32. Rixon, S. (2016). Do developments in assessment represent the 'coming of age' of young learners English language teaching initiatives? The international picture. In M. Nikolov (Ed.), *Assessing young learners' of English: Global and local perspectives* (pp. 19-41). New York: Springer.
33. Terhart, E. (2013). Teacher resistance against school reform: reflecting an inconvenient truth, *School leadership management*, 33(5), 486-500.
34. Tóth, E. & Csapó, B. (2011). Comparing primary and high-school teachers' attitudes towards testing and accountability. Paper presented at the 14th European Conference for the Research on Learning and Instruction. Exeter, United Kingdom, August 30- September 3, 2011. 57.
35. Tran, T. H. (2012). Second Language Assessment for Classroom Teachers , Paper presented at MIDTESOL-2012, Ames, Iowa Rolla, Missouri, USA

Validation of an Instrument to Assess Beliefs About Nature of Science and Scientific Inquiry in Namibia

Simson Ndadaleka Shaakumeni
Doctoral School of Education, University of Szeged, Hungary

Doi: 10.19044/ejes.v6no2a2 URL:http://dx.doi.org/10.19044/ejes.v6no2a2

Abstract

The aim of this study was to validate a new 5-point Likert scale questionnaire that is being developed to assess secondary school science students' beliefs about the nature of scientific knowledge and scientific inquiry. Following a review of literature, it was found that out of the many existing instruments, most of them were open-ended and none were found to have been developed in the cultural context of Namibia. The beliefs about nature of science (BANOS) questionnaire is currently being developed to break the ground for future cross-sectional research on the nature of science and scientific inquiry in Namibia. The BANOS questionnaire taps from aspects of nature of science and scientific inquiry as a theoretical framework. The 35-item BANOS questionnaire was administered to a sample of 124 (39% male and 61% female) secondary school science teachers. The analysis of the instrument showed that the average inter-item correlation was reasonable at $r = .40$, the mean item-total correlations were acceptable at $r = .63$ and the reliability was very high at $\alpha = .96$. Science teachers response pattern showed several variations indicating mixed beliefs about the nature of science and scientific inquiry. At non-parametric level, it was established that gender, science teaching and the type of science subject taught did not influence teachers' beliefs about nature of science and scientific inquiry. Further piloting of the questionnaire on adequate sample of student population is needed to enable more robust statistical analyses to assess the psychometric properties of the instrument.

Keywords: Nature of science; beliefs; validity/reliability, scientific inquiry.

Introduction

The National Curriculum for Basic Education (NCBE) in Namibia, which is the broad curriculum, demands that students should develop into scientific literate citizens (Ministry of Education, 2010). According to the NCBE, one of the components of scientific literacy is the understanding of the nature of scientific knowledge. The nature of science entails what makes science different from other disciplines. In other words, it characterises scientific knowledge that is derived from how the knowledge is developed (Lederman, et al., 2014). However, the assessment of science knowledge in Namibian schools does not include these aspects of scientific literacy. All assessments mainly focus on the subject content knowledge. They hardly include assessing students' abilities to carry out inquiry and whether they acquire fundamental understanding of the characteristics of nature of science and scientific inquiry. One way to ascertain students' understanding of the nature of scientific knowledge and knowing is to assess their science epistemic beliefs. Advancing students' beliefs about the nature of scientific knowledge and knowing has featured prominently in recent research in science education (Conley, Pintrich, Vekiri, & Harrison, 2004; Tsai, Ho, Liang, & Lin, 2011; Chen, 2012; Chen, Metcalf, & Tutwiler, 2014). However, none of such studies appear to have been conducted in Namibia.

Theoretical Background

Nature of Science

One important goal of science education is to foster students' scientific literacy (Nowak, Tiemann, & Upmeier zu Belzen, 2013; Peters-Burton, 2016). Scientific literacy consists of different components, namely; content knowledge, nature of science, and scientific inquiry. The concept nature of science has been commonly used to refer to "the epistemology of science; science as a way of knowing or the values and beliefs inherent to the development of scientific knowledge" (Lederman, 1992, p. 331; 2007). This definition of nature of science is rather general and till date, there is still disagreement among philosophers of science, historians of science, scientists, and science educators on the specific definition of the concept (Abd-El-Khalick, 1998)

Students should develop certain habits of the mind such as believing that scientific knowledge (1) can change over time (tentative), (2) empirically-based (based on and/or derived from observations of the natural world), (3) there is no one way of doing science called "the Scientific Method", (4) subjective (theory-laden) - partially based on human inference, (5) influenced by imagination and creativity, (6) socially and culturally embedded, (7) observation and inference are different, and (8) theories and laws are distinct kinds of scientific knowledge (Abd-El-Khalick & Lederman, 2000; Lederman, 2007; McComas, 2008; Osborne, Collins, Ratcliffe, & Duschl, 2003; Niaz, 2009; Chen, 2012; Abd-El-Khalick, et al., 2017).

Scientific Inquiry

Scientific inquiry as a component of scientific literacy has been defined in various ways. Schmidt, Smyth, and Kowalski (2014) defined scientific inquiry as "a powerful way of understanding science content. Students learn how to ask questions and use evidence to answer them. In the process of learning the strategies of scientific inquiry, students learn to conduct an investigation and collect evidence from a variety of sources, develop an explanation from the data and communicate and defend their conclusions" (p. 4).

According to Lederman et al. (2014), scientific inquiry refers to "the combination of general science process skills with traditional science content, creativity, and critical thinking to develop scientific knowledge" (p. 65). The aspects of scientific inquiry that students in secondary level should be able to understand and articulate is that: (1) all scientific investigations begin with a question but do not necessarily test a hypothesis, (2) there is no single set or sequence of steps followed in all investigations, (3) inquiry procedures are guided by the question(s) asked, (4) all scientists performing the same procedures may not get the same results, (5) inquiry procedures can influence results, (6) research conclusions must be consistent with the data collected, (7) scientific data are not the same as scientific evidence, and (8) explanations are developed from a combination of collected data and what is already known (Lederman et al., 2017).

Scientific inquiry, though closely related, is not necessarily a component of nature of science. It refers to how scientists carry out their work and how the resultant knowledge is generated and accepted. Moreover, nature of science constitutes what makes science different from other disciplines such as Sociology or History. Hence, it refers to the characteristics of scientific knowledge that are necessarily derived from how the knowledge is developed (Lederman et al., 2014). Developing a tool for assessing beliefs about the preceding aspects that characterise the nature of science and scientific inquiry is the focus of this study.

Assessing Beliefs about Nature of Science

Numerous instruments for assessing beliefs about the nature of science exist (Lederman, Wade, & Bell, 1998; Summers & Abd-El-Khalick, 2017). However, none were found to have been developed in the cultural context of Namibia. Moreover, the widely-used instrument in this area of research in recent years is the Views of the Nature of Science (VNOS) questionnaire developed by Lederman, Abd-El-Khalick, Bell, and Schwartz (2002).

The VNOS is an open-ended questionnaire and items that often do not directly specify the targeted aspects of nature of science. Subsequently, it becomes too optimistic to assume that participants would be able to adequately articulate the details sought within a limited time frame. Thus, it becomes difficult for a researcher to obtain the intended information from every participant without follow-up interviews (Chen, 2006). Although Lederman et al. (2002) advised that a large proportion of respondents should be interviewed after taking the questionnaire to validate responses, this has not been ideal for this study at the moment due to limited access to state schools in Namibia.

Another argument against the use of open-ended questionnaire to gather views is that it requires a substantial commitment of time and mental grit from each participant to articulate their views (Osborne, Collins, Ratcliffe, Millar & Duschl, 2003). As a result, only few respondents may endeavour to articulate their views. This renders open-ended questionnaires inappropriate for large scale empirical studies, attributable to the smaller samples involved as well as limited possibilities of applying statistical analysis to qualitative data. The forgoing arguments form the basis to justify the choice of the type of instrument being developed in this study.

Research Aims and Objectives

The aim of this study is to validate a new instrument that was developed to assess secondary school science students' beliefs about the nature of scientific knowledge and scientific inquiry in Namibia. Hence, this paper reports on the first pilot study that was carried out in Namibia using a sample of science teachers as one of the preliminary steps in the validation process. The following questions were answered:

1. How do items function in the instrument?
2. What is the reliability of the instrument?
3. What are the science teachers' overall beliefs about nature of science and scientific inquiry?
4. Is there any difference in beliefs among teachers based on gender, teaching experience or type of science subject taught?

Methods

Participants

It should be noted that the questionnaire was developed to assess students' beliefs about nature of science and scientific inquiry. However, as part of the preliminary stages of the validation process, the questionnaire was administered to a sample of 124 (39% male and 61% female) science teachers in three regions in Namibia. The criterion for inclusion was that teachers must have been teaching either of the two science subjects that make up the Natural Sciences in the Namibian curriculum. This means they should either be teaching Biology or Physical Science at secondary level. The respondents were grouped into ranges of years of experience in science teaching from 0-3 years to over 10 years.

Instrument Development

The theoretical framework for the development of the instrument for assessing beliefs about the nature of science is based on the general and symbiotic aspects of nature of science and scientific inquiry as proposed by Lederman and other scholars (Lederman, Abd-El-Khalick, Bell, & Schwartz, 2002; Lederman & Abd-El-Khalick, 1998; McComas, Almazroa, & Clough, 1998, Lederman et al.,2014). These aspects of nature of science, though criticised by some science educators as being too general, over-simplified, prescriptive and narrow (Irzik & Nola, 2011; Mathews, 2012; Dagher & Erduran, 2016; Grandy & Duschl, 2008), they are considered to be a useful guiding framework for this study. This framework was adopted based on the clarification provided by proponents of the general aspects of nature of science, who in response to criticisms stated that the suggested general aspects of nature of science and scientific inquiry are by no means "a definitive or universal definition of the construct" (Lederman, Antink, & Bartos, 2014, p. 286).

In the present study, a new questionnaire termed "Beliefs about Nature of Science" (BANOS) has been developed. However, ideas for possible items were obtained from existing scales in the literature. The first version of the BANOS questionnaire comprised of thirty-five items. The items were declarative statements describing particular aspects of nature of scientific knowledge including scientific inquiry. Respondents gave their personal level of belief or agreement with the five-point Likert scale (Cohen, Manion, & Morrison, 2007), namely; 1 = strongly disagree, 2 = disagree, 3 = not sure, 4 = agree, and 5 = strongly agree.

The declarative statements are organised according to the general aspects of nature of science and scientific inquiry. All statements were in a form of sophisticated/informed views of respondents about nature of science and scientific inquiry obtained from the literature (Vhurumuku, 2010; Dogan & Abd-El-Khalick, 2008; Chen, 2006; Khishfe & Abd-El-Khalick, 2002; Summers & Abd-El-Khalick, 2017). All statements were positively worded so that a high score indicate more sophisticated beliefs about the nature of science and scientific inquiry. The BANOS questionnaire maximum total raw score is 175 and a minimum of 35.

Procedure

An assessment of validity and reliability was conducted on the questionnaire. To ascertain face and content validity, a review of over seventy recent researches on nature of science and scientific inquiry was done. This served to ensure the accuracy of each declarative statement (item) that was used in the questionnaire. To check for comprehension and readability, the paper-and-pencil questionnaire was administered to a sample of 124 science teachers that were purposefully sampled from three regions in Namibia. The involvement of teachers is also to some extent expected to provide some validation as teachers' beliefs are expected to be different from that of students.

An English language expert was also engaged to read through the statements and modifications made were applicable. The reliability of the instrument was demonstrated by confirming the internal consistency of the construct using Cronbach's alpha.

Ordinal scales were analysed as if they were interval scales (Glynn, Brickman, Armstrong & Taasoobshirazi, 2011). In this case, items are assumed to be generally parallel indicators of the underlying latent variable (DeVellis, 2003).

Results and Discussions

Inter-item Correlations

In order to produce a reliable scale of measurement, it is imperative to analyse the extent to which responses on one item are related to responses on all other items in a scale. For this reason, inter-item correlations are an essential element in the initial analysis of a set of items.

This analysis provides information about the extent to which items on a scale are assessing the same construct (Cohen & Swerdlik, 2005). The average inter-item correlation for a set of items should preferably be between r = .20 and r = .40 (Piedmont, 2014). This implies that although items should be reasonably similar in some way, it is pointless to have items on a scale that measure the construct in exactly the same way. When correlation values are below r = .20, it suggests that items do not relate to each other very well. Hence, it may not be suitable for measuring a single construct. Moreover, if the values are above r = .40, the items tend to be very similar to each other, almost to the point that they become redundant (Piedmont, 2014).

Table 1. Summary of item statistics

	Mean	No. of Items
Item Means	2.95	35
Inter-Item Correlations	**.403**	35

The average inter-item correlation for the items (N = 35) in the first version of Beliefs about Nature of Science (BANOS) instrument was r = .40 (Table 1). As indicated above, correlations between r = .20 and r = .40 are preferable. This indicates that items are related to each other fairly well. Therefore, in the context of this study, they may be suitable for measuring beliefs about the nature of science and scientific inquiry in Namibia.

Item-total Correlations

Items were further examined to ascertain whether there is an individual item whose score is not related to the summated score of all other items in the scale and such items are revised or discarded. This is essentially performing item-total correlations. Item-total correlation values between r = .20 and r = .40 are desirable and a small correlation suggests that the item is not measuring the same construct measured by the other items in the scale (Field, 2005; Everitt, 2002). The mean item-total correlation (N = 35) was r = .63. One item had correlation values less than r = .20. This shows that the item was not measuring the same construct measured by the other items in the instrument. This item was examined to ascertain whether it can be improved or be discarded. Three other items had correlations less than r = .40 but greater than r = .20. These items displayed very weak correlations and were reviewed. Since this is a validation process, the intention was not to discard items yet, but rather to modify them and pilot them for the second time. Subsequently, items were reviewed and none of the items were discarded based on these results.

Reliability

Reliability is a measure of how well the items in a scale measure the same construct over time (Streiner, 2003). This measure is commonly estimated using Cronbach's alpha reliability coefficient, which normally ranges between 0 and 1. The closer Cronbach's alpha coefficient is to 1, the greater the internal consistency of the items in the scale (Gliem & Gliem, 2003). George and Mallery (2003) suggest that Cronbach's alpha of α = .70 is acceptable. The Cronbach's alpha computed on all the 35 items (N = 124) using SPSS version 23 was α = .96. This shows that the instrument overall has very high reliability. Further assessment of reliability analysis reveals that there was no need to delete any item in order to improve the overall reliability value of the instrument.

The eight general aspects of nature of science and scientific inquiry made up the nine subscales in the instrument. Participants' responses on the items for each subscale were computed to determine the median (Elliot & McGregor, 2001). Despite that the overall reliability of the instrument as a whole was very high, two subscales namely the empirical nature of scientific knowledge and the scientific method did not show very good reliability. The reliability of the two subscales were α = .51 and α = .67 respectively (Table 2). Deleting one item from the empirical subscale improved the alpha coefficient value slightly to α = .54. However, this was still very low. Moreover, by deleting one item from the scientific method subscale improved the alpha coefficient to α = .73. This alpha coefficient value is now reasonable, However, the items in this subscale were reviewed and will be piloted for the second time before any permanent dropping of the items is implemented.

Table 2. Reliability of the subscales

NOS aspects	No. of items	Cronbach's alpha (α)
Tentative nature of science	5	.81
Empirical nature of science	4	.51 (.54)*
The scientific method	3	.67 (.73)*
Subjective nature of science	3	.78
Imaginative and Creativity	4	.75
Socio-cultural influence	3	.88
Observations and inferences	3	.90
Theories & laws	5	.83
Scientific inquiry	5	.91
Total items	35	.96

* = α if 1 item deleted

The number of items in a scale partially influences the Cronbach's alpha values. These values could be increased by increasing the number of items in the scale or deleting individual items whose values are higher than the overall Cronbach's alpha value of the scale. The very high alpha values indicate that items were measuring the same construct very well. Notwithstanding the above, Gliem and Gliem (2003) cautioned that although a higher Cronbach's alpha indicates a good internal consistency of the items in the scale, it does not mean that the scale is uni-dimensional. Hence, factor analysis is still required to determine the dimensionality of the scale. However, for this first pilot study, factor analysis was not appropriate as the sample was inadequate to satisfy all the assumptions for factor analysis procedure. This will be done during the second pilot study.

Subscale Correlations

The intertwined aspects of nature of science and scientific inquiry formed subscales that reflect the core constructs of nature of science and scientific inquiry. When separate constructs are combined to form one scale, there is a need to justify that they are closely related (Summers & Abd-El-Khalick, 2017). One way to show the relationship between constructs is by computing correlations. Without assuming causation (Carver & Nash, 2012), the analysis showed that there was an overall significant positive relationship between subscales at $p < .001$ (Table 3) level. The weakest correlation was between tentativeness and empirical subscales ($r = .27$). The empirical nature of science subscale showed poor reliability values ($\alpha = .54$, Table 3). Thus, this may explain this weak relationship. The strongest correlation was between observations and inferences and scientific inquiry ($r = .88$). It should be noted that scientific inquiry though closely related, is not necessarily a component of nature of science because it entails the process of how scientists perform their work and how the resultant knowledge is generated and accepted (Lederman, et al., 2014). However, beliefs about it was considered important for this study.

Table 3. Subscales correlations

Subscales	1	2	3	4	5	6	7	8	9
1. Tentative nature of science	-								
2. Empirical nature of science	.27**	-							
3. The scientific method	.42**	.51**	-						
4. Subjective nature of science	.59**	.70**	.50**	-					
5. Imaginative and creativity	.44**	.77**	.56**	.78**	-				
6. Sociocultural	.38**	.44**	.43**	.43**	.57**	-			
7. Observations and inferences	.64**	.54**	.56**	.69**	.66**	.66**	-		
8. Theories and laws	.53**	.64**	.71**	.59**	.69**	.65**	.77**	-	
9. Scientific inquiry	.74**	.55**	.69**	.70**	.68**	.53**	.88**	.86**	-

**. Correlation is significant at p < .001 level (2-tailed).

It should be noted, however, that very high significant correlations between subscales point to strong similarity. As indicated in preceding section, the overall Cronbach's alpha value was very high (α = .96) and most likely indicate unnecessary redundancy of items (and subsequently the subscales) rather than a desirable level of internal consistency (Streiner, 2003).

Teachers' Response Patterns on Subscales

Firstly, the median of responses was calculated for each subscale using the Statistical Package for Social Sciences (SPSS) version 23. The median as a measure of central tendency was deemed the most appropriate indicator of respondents' likeliest beliefs about each subscale. Secondly, the interquartile range (IQR) which is a measure of dispersion was also computed to indicate whether responses are clustered together or scattered across the range of possible responses. This is shown by the boxplot in Figure 1. With reference to the tentative nature of science, as shown in Figure 1, most respondents were more likely to choose option 4 (Agree) (Mdn = 4, IQR = 3). This aspect of nature of science is concerned with the idea that scientific knowledge is never absolute or certain but is subject to change (Abd-El-Khalick, et al., 2017; Lederman, 2007). However, the length of the boxplot indicates variability in opinion about this aspect of nature of science with more variability among the lower quartile.

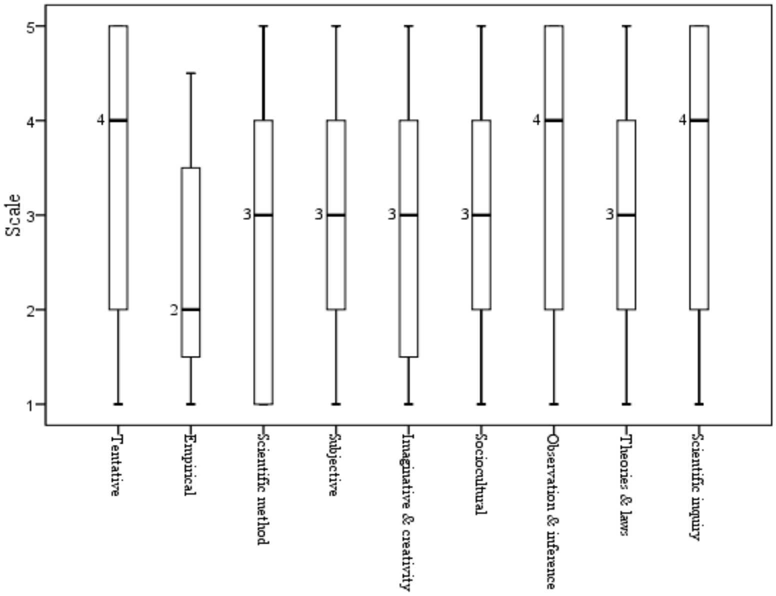

Figure 1. Boxplot of responses to subscales showing the median

Four items that made up the empirical nature of science subscale are: (1) science does not depend on experiments only to get evidence about the natural world, (2) experiments cannot prove a scientific theory true, (3) science cannot prove anything but is still valuable, and (4) scientific evidence can be obtained from observations of the natural world. Respondents indicated disagreement with this notion (Mdn = 2, IQR = 2). As shown in Figure 1, most respondents were likely to choose option 2 (disagree). This indicates that teachers' beliefs about this aspect of nature of science were not accurate or they may have interpreted the items differently than expected. Experiments are useful tools in science but are not the only means to generate scientific knowledge (McComas, 1996). Scientific knowledge is also derived from the observation of the natural world (Lederman, 2007; Lederman, et al., 2014).

With regards to the scientific method, respondents appear to be unsure about this aspect of the nature of science (Mdn = 3, IQR = 3). Respondents were most likely to choose option 3 (Not sure) and the variability of responses spread more within the lower quartile. This indicates that those who did not choose option 3 (Not sure) mostly disagreed or strongly disagreed. There is a commonly held misconception about science that there exists a single procedure which all scientists follow (Lederman, Abd-El-Khalick, Bell, & Schwartz, 2002). This was inaccurate as there is no single scientific method that all scientists follow, but scientists use various methods in search of scientific knowledge (Abd-El-Khalick, Waters, & Le, 2008; Abd-El-Khalick, et al., 2017).

The work of scientists is influenced by their theoretical and disciplinary commitments, beliefs, prior work knowledge, training, and expectations (Abd-El-Khalick, et al., 2017). This suggests that scientific knowledge is subjective in nature as it is affected by scientists' backgrounds. Respondents' beliefs about this aspect of nature of science appear to be divided. About 48% of respondents disagreed but roughly equal proportion amounting to 45% of respondents indicated that they agreed (Mdn = 3, IQR = 2). This is also evident in Figure 1; the upper and lower quartiles appear roughly equal.

Respondents' beliefs about the imaginative and creative nature of science appear to show more uncertainty as well with more variability in the lower quartile (Mdn = 3, IQR = 2.5). Scientific knowledge production involves human creativity in terms of scientists inventing explanations and theoretical models, and this requires a great deal of creativity by scientists (Abd-El-Khalick, Waters, & Le, 2008). Creativity and imagination are vital at all stages of a scientific endeavour; from planning and designing through data collection to data interpretation though with variable extent between stages (Wong & Hodson, 2009).

Scientific knowledge affects and is affected by various cultural elements and spheres including social fabric, trends, prestige, power structures, religion, political and economic factors (McComas, 2008). Respondents' beliefs about this aspect of nature of science seem divided. About 32% of respondents agreed with this notion, but a sizeable proportion amounting to 32% of respondents disagreed (Mdn = 3, IQR = 2).

With reference to the difference between observations and inferences, most respondents were again more likely to choose option 4 (Agree) (Mdn = 4, IQR = 3). This aspect of nature of science deals with the fact that there is a crucial distinction between observations and inferences, although both are scientific processes skills. Observations are descriptions of the natural world that are accessible to the human senses whereby several observers could easily reach an agreement whilst inferences are interpretations or explanations of observations gathered (Lederman, Antink, & Bartos, 2014; Schwartz, Lederman, & Crawford, 2004). The response pattern in this subscale is also showing variability in responses about this aspect of nature of science with more variability visible in the lower quartile as evident in Figure 1.

About 35% of respondents disagreed with the notion that theories and laws are different, but a sizeable proportion amounting to 48% of respondents indicated that they agree (Mdn = 3, IQR = 2). This is also evident in Figure 1; the upper and lower quartiles appear roughly similar. There are common misconceptions that there is a simplistic and hierarchical relationship between observations, hypotheses, theories and laws of science; and the belief that laws have a higher status than theories within a scientific endeavour (Lederman, Abd-El-Khalick, Bell, & Schwartz, 2002). Theories and laws are related but are distinct kinds of scientific knowledge and for this reason they serve different roles in the scientific enterprise. Hence, theories do not in any way become laws, even with additional evidence (Niaz, 2009; McComas, 2008; Lederman, 2007).

Scientific inquiry, though closely related, is not necessarily a component of nature of science as it entails the process of how scientists carry out their work and how the resultant knowledge is generated. This subscale was made up of five items. Respondents were more likely to choose option 4 (Agree) (Mdn = 4, IQR = 3). Similar to some of the subscales discussed, the length of the boxplot indicates variability in opinion about this component of scientific literacy with more variability among the lower quartile (Figure 1).

The response pattern revealed an interesting trend in the science teachers' beliefs about the nature of science and scientific inquiry. It shows that teachers' beliefs about the tentative nature of science, observations and inferences, and scientific inquiry is quite similar (Mdn = 4). This means teachers' beliefs about these subscales are quite informed because option 4 in the scale represents 'agree'. However, the response pattern for five of the subscales, namely; the scientific method, the subjective method, imaginative and creative method, socio-cultural nature of science and scientific theories and laws, show substantive amount of uncertainty in the teachers' beliefs (Mdn = 3). This suggests that they were not sure whether to agree or disagree with the statements. Furthermore, teachers' beliefs about the empirical nature of science was more confounding because they generally disagreed with this notion (Mdn = 2). It was not possible to ascertain the reasons for this disagreement, whether or not it was a misinterpretation of the items or they simply had different views regarding the empirical nature of science. For this reason, it was imperative to conduct a follow-up interview with participants to get clarity on their responses.

Beliefs Across Gender, Years of Experience and Science Subject Taught

The assessment of normality of the data was conducted and the Shapiro-Wilk test was significant ($p < .001$). This suggests that the data was not normally distributed and non-parametric tests would be appropriate to examine the difference in beliefs (Field, 2009). The Mann-Whitney U test was used to examine the difference in beliefs about nature of science and scientific inquiry based on gender. The results showed that there was no significant difference ($U = 1496$, $p = .092$). It can be concluded that gender does not influence science teachers' beliefs about nature of science and scientific inquiry. Similarly, the Mann-Whitney U test showed no significant difference in beliefs based on the type of science subject (Biology or Physical Science) teachers taught ($U = 1896$, $p = .712$).

Kruskal-Wallis test was used to examine the difference in beliefs based on ranges of years of science teaching experience. The results once again showed that there was no significant difference in beliefs between the different ranges of years of science teaching experience, $\chi^2(3) = 6.717$, $p = .081$, with the mean rank belief score of 83.50 for 0-3 years, 66.21 for 3-5 years, 53.33 for 5-10 years, and 66.70 for 10+ years of science teaching experience. This suggests that teaching experience did not influence teachers' beliefs about nature of science and scientific inquiry.

Conclusion

This pilot study was carried out in Namibia as part of the initial stages of the validation process in the development of questionnaire for assessing beliefs about nature of science and scientific inquiry. The initial analyses of this instrument indicated that it had potential to be a reliable instrument for assessing beliefs about the nature of scientific knowledge and scientific inquiry in Namibia, with respect to its preliminary very high reliability. This indicates that the items were pulling in one direction in terms of assessing the same construct. However, very high reliability values may indicate unnecessary redundancy of items in the scales. The correlations between subscales that constituted the instrument were generally significantly high. A phenomenon that may indicate that subscales were too similar and therefore the instrument may have poor discriminant validity. This study therefore revealed that science teachers in Namibia have mixed beliefs about the nature of science and scientific inquiry. The study further established that at non-parametric level, gender, science teaching experience, and the type of science subject taught did not influence teachers' beliefs about nature of science and scientific inquiry. Some methodological limitations may have influenced the findings of this study. First, respondents were not interviewed to ascertain accurate interpretation of the questionnaire items or to dig deeper and uncover possible explanations for the variability in their responses (Lederman, Abd-El-Khalick, Bell, & Schwartz, 2002). It was assumed they interpreted the items as expected. Secondly, it was not possible at this juncture to apply more advanced statistical analyses such as factor analysis to the data due to inadequate sample size.

Acknowledgement

The author of the paper is on Stipendium Hungaricum Scholarship programme of the Hungarian government in collaboration with the Namibian government through the Namibian Students Financial Assistance Fund (NSFAF).

References:

1. Abd-El-Khalick, F., Myers, J. Y., Summers, R., Brunner, J., Waight, N., Wahbeh, N., . . . Belarmino, J. (2017). A logitudinal analysis of the extent and Manner of representations of nature of science in U.S. high school Biology and Physics textbooks. *Journal of Research in Science Teaching*, *54*(1), 82-120.
2. Abd-El-Khalick, F., Waters, M., & Le, A.-P. (2008). Reprersentation of nature of science in high school chemistry textbooks over the past four decades. *Journal of Research in Science Teaching, 45*(7), 835-855.
3. Carver, R. H., & Nash, J. G. (2012). *Doing data analysis with SPSS version 18.* Boston: Brooks/Cole Cengage Learning.
4. Chen, J. A. (2012). Implicit theories, epistemic beliefs, and science motivation: A person-centered approach. *Learning and Individual Differences*, *22* , 724–735.
5. Chen, J. A., Metcalf, S. J., & Tutwiler, M. S. (2014). Motivation and beliefs about the nature of scientific knowledge within an immersive virtual ecosystems environment. *Contemporary Educational Psychology*, *39*(2), 112–123.
6. Chen, S. (2006). Development of an instrument to assess views on nature of science and attitudes towards teaching science. *Science Education, 90*, 803-819.
7. Cohen, L., Manion, L., & Morrison, K. (2007). *Research methods in education (Sixth Ed).* New York: Routledge.
8. Conley, A. M. M., Pintrich, P. R., Vekiri, I., & Harrison, D. (2004). Changes in epistemological beliefs in elementary science students. *Contemporary Educational Psychology*, *29*(2), 186–204.
9. Dagher, Z. R., & Erduran, S. (2016). Reconceptualizing the nature of science for science education. *Science & Education*, *25*, 147–164.
10. DeVellis, R. F. (2003). *Scale development: Theory and applications* (2nd ed.). Thousand Oaks, CA: Sage.
11. Dogan, N., & Abd-El-Khalick, F. (2008). Turkish grade 10 students' and science teachers'conceptions of nature of science: A national study. *Journal of Research in Science Teaching*, *45*(10), 1083-1112.
12. Elliot, A., & McGregor, H. A. (2001). A 2 x 2 achievement goal framework. *Journal of Personality and Social Psychology, 80*(3), 501-519.
13. Everitt, B. S. (2002). *The Cambridge Dictionary of statistics.* Cambridge, UK: Cambridge University Press.
14. Field, A. (2005). *Discovering Statistics Using SPSS 2nd Ed.* London: Sage.

15. Field, A. (2009). *Discovering Statistics Using SPSS. 3rd Edition*. London Sage Publications Ltd.
16. George, D., & Mallery, P. (2003). *SPSS for Windows step by step: A simple guide and reference. 11.0 update (4th ed.).* Boston: Allyn & Bacon.
17. Gliem, J. A., & Gliem, R. R. (2003). Calculating, Interpreting, and Reporting Cronbach's Alpha Reliability Coefficient for Likert-Type Scales. *Science and Education*. Columbus, OH, Ohio, USA. Retrieved September 19, 2017, from https://scholarworks.iupui.edu/bitstream/handle/1805/344/ Gliem+&+Gliem.pdf?sequence=1
18. 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*(10), 1159–1176.
19. Grandy, R., & Duschl, R. (2008). Consensus: Expanding the scientific method and school science. In R. Duschl, & R. Grandy, *Teaching scientific inquiry: Recommendations for research and implementation* (pp. 304-325). Rotterdam, The Netherlands: Sense.
20. Irzik, G., & Nola, R. (2011). A family resemblance approach to the nature of science for science education. *Science & Education, 20*, 591-607.
21. Khishfe, R., & Abd-El-Khalick, F. S. (2002). Influence of explicit and reflective versus implicit inquiry-oriented instruction on sixth graders' views of nature of science. *Journal of Research in Science Teaching, 39*(7), 551-578.
22. Lederman. (2007). Nature of science: Past, present and future. In S. K. Abell, & N. G. Lederman, *Handbook of research on science education* (pp. 831-880). Mahwah, NJ: Lawrence Erlbaum.
23. Lederman, Abd-El-Khalick, F., Bell, R. L., & Schwartz, R. S. (2002). Views of the nature of science questionnaire: Toward valid and meaningful assessment of learner's conceptions of the nature of science. *Journal of Research in Science Teaching, 39*(6), 497-521.
24. Lederman, J. S., Lederman, N. G., Bartos, S. A., Bartels, S. L., Antink-Meyer, A., & Schwartz, R. S. (2014). Meaningful assessment of students' understandings about scientific inquiry—The views about scientific inquiry (VASI) questionnaire. *Journal of Research in Science teaching, 51*(1), 65-83.
25. Lederman, N. G., & Abd-El-Khalick, F. (1998). Avoiding De-natured science: Activities that promote understanding of nature of science. In W. F. McComas, *The nature of science in science education:*

Rationales and strategies (pp. 83-126). Dordrecht, The Netherlands: Kluwer Academic Publishers.

26. Lederman, N. G., Antink, A., & Bartos, S. (2014). Nature of science, scientific inquiry and socio-scientific issues arising from genetics: A pathway to developing a scientifically literate citizenry. *Science & Education, 23*(2), 285-302.
27. Lederman, J., Lederman, N., Bartels, S., Pavez, J. J., Lavonen, J., Blanquet, E., … Yalaki, Y. (2017, August 21-25). Understandings of scientific inquiry : An international collaborative investigation of seventh grade students. Paper presented at European Science Education Research Association Conference.https://www.researchgate.net/publication/322603295_understandings_of_scientific_inquiry_an_international_collaborative_investigation_of_seventh_grade_students
28. Lederman, N. G., Wade, P. D., & Bell, R. L. (1998). Assessing understanding of nature of science: A historical perspective. In W. F. McComas, *The nature of science in science education: Rationales and strategies* (pp. 331-350). The Netherlands: Kluwer Academic .
29. Mathews, M. R. (2012). Changing the focus: From nature of science to features of science (FOS). In M. S. Khine, *Advances in nature of science research: Concepts and methodology* (pp. 3-26). Dordrecht, The Netherlands: Springer.
30. McComas, W. F. (1996). Ten myths of science: Reexamining what we think we know about the nature of science. *School Science and Mathematics, 96*(1), 10-16.
31. McComas, W. F. (2008). Seeking historical examples to illustrate key aspects of the nature of science. *Science & Education, 17,* 249-263.
32. McComas, W. F., Almazroa, H., & Clough, M. P. (1998). The nature of science in science education: An introduction. *Science & Education, 7*, 511-532.
33. Ministry of Education. (2010). *National Curriculum for Basic Education.* Okahandja: NIED.
34. Niaz, M. (2009). *Critical appraisal of physical science as a human enterprise: Dynamics of scientific progress.* Dordrecht, The Netherlands: Springer.
35. Nowak, K. H., Tiemann, R., & Upmeier 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.
36. Osborne, J., Collins, S., Ratcliffe, M., & Duschl, R. (2003). What "ideas-about science" should be taught in school science? A Delphi

study of the expert community. *Journal of Research in Science Teaching, 40*(7), 692-720.
37. Peters-Burton, E. (2016). Scientists taking a nature of science course: Beliefs and learning outcomes of career switchers. *School Science and Mathematics, 116*(3), 148-163.
38. Piedmont, R. L. (2014). Inter-item correlations. In A. C. Michalos, *Encyclopedia of Quality of Life and Well-Being Research* (pp. 3303–3304). Netherlands: Springer Netherlands.
39. Schwartz, R. S., Lederman, N. G., & Crawford, B. A. (2004). Developing views of nature of science in an authentic context: An explicit approach to bridging the gap between nature of science and scientific inquiry. *Science Education, 88*(4), 610-645.
40. Streiner, D. L. (2003). Starting from the beginning: an introduction to coefficient alpha and internal consistency. *Journal of Personality Assessment, 80*(1), 99-103.
41. Summers, R., & Abd-El-Khalick, F. (2017). Development and Validation of an instrument to assess students attitudes towards science across 5 through 10. *Journal of Research in Science Teaching*, 1-33. doi:10.1002/tea.21416
42. Tsai, C. C., Ho, H. N. J, Liang, J. C., & Lin, H. M. (2011). Scientific epistemic beliefs, conceptions of learning science and self-efficacy of learning science among high school students. *Learning and Instruction, 21*(6), 757–769.
43. Vhurumuku, E. (2010). The impact of explicit instruction on undergraduate students' understanding of the nature of science. *African Journal of Research in Mathematics, Science and Technology Education, 14*(1), 99-111.
44. Wong, S. L., & Hodson, D. (2009). From the horse's mouth: What scientists say about scientific investigation and scientific knowledge. *Science Education, 93*, 109-130

Project Management in Education: A Case Study for Student Learning Pathway Evaluation in a Lifelong Learning Program Using Fuzzy CPM and Fuzzy PERT

Omiros Iatrellis, MSc
Panos Fitsilis, PhD
Dimitrios Lampakis, PhD
University of Thessaly, Larissa, Greece
Achilles Kameas, PhD
Hellenic Open University Patras, Greece

Doi: 10.19044/ejes.v6no2a3 URL:http://dx.doi.org/10.19044/ejes.v6no2a3

Abstract

Time to degree and graduation are paramount concerns in higher education today and have caught the attention of educators, policy makers, and researchers in recent years. Delays in student flow through learning pathways may result in undesired consequences and this fact has made educators and the public start to regard timely degree completion as a critical measure of accountability for colleges and universities. Project management is defined as application of techniques to project activities to meet the project requirements. It has techniques that can help Higher Educational Institutions (HEIs) standardize education and quality, reduce costs, improve effectiveness and decrease prolonged education periods. The current study is carried out in a University's Division of Lifelong Learning to understand the student flow of the offered program with respect to the Fuzzy Critical Path Method (FCPM) and Fuzzy Program Evaluation Review Technique (FPERT). The reason for incorporating fuzzy sets is stemmed from that fuzzy numbers are more effective for high uncertainty processes as in learning environments. Activities affecting student flow through the learning pathways are identified and project network is drawn. 72 student' data about the times of each activities are obtained from the observed university. Findings through application of FCPM and FPERT, the expected time for completion of the project, slack times and the critical path are determined.

Keywords: Education, lifelong learning, time to degree, fuzzy critical path method, fuzzy program evaluation review technique.

1. Introduction:

A project can be considered to be the achievement of a specific objective, which involves a series of activities and tasks and it has to be completed within a set specification, having definite start and end dates (Munns & Bjeirmi, 1996). A standard project typically has the following four major phases each with its own agenda of tasks and issues: initiation, planning, implementing and completion processes (Serrano & Avilés, 2018). Planning the project is a critical phase of the lifecycle of a project, as defined processes and tasks can predetermine efficiency. Project management is the practice of initiating, planning, executing, controlling, and closing the work of a team to achieve specific goals and meet specific success criteria at the specified time. In its Glossary, A Guide to the Project Management Body of Knowledge (Project Management Institute Inc, 2000) defines the Triple Constraint which in the basic setup is the time, cost and scope with quality occasionally included as an adjunct to or substitute for scope or as a fourth constraint. In education the notion of an iron triangle has been posited, linking firstly access, quality and cost and latterly accessibility, quality and efficiency in order to help explain the interplay and interactions between specified components of higher education systems at different levels and to take account of emerging trends towards new approaches (Lane, 2014). Quality, one of the most important of all, is directly related to the efficiency and performance in all the facets of education sector (Iatrellis, Kameas. Achilles, & Panos, 2019). An important factor that affects the performance of educational processes in a HEI is the "Time to completion" or "Time to degree", which refers to the time it takes for a student to attain the desired degree (Wächter et al., 2015). The abovementioned metric has drawn increased public policy attention since its elongation involves substantial costs for stakeholders in terms of foregone earnings and additional tuition expenditures (Bound, Lovenheim, & Turner, 2010). Any delay in degree completion represents a waste of resources both at individual and at collective level, thus affecting the returns to investment in higher education (Casalone & Orientale, 2011). Among researchers, however, the length of time to degree attainment has received little attention. It is identified that student flow processes include:

1. Learning time: This component is the amount of time it takes the student to complete successfully a learning activity.
2. Scheduling time: This component is the amount of time from when a procedure is scheduled until it actually takes place e.g. internship approval
3. Teaching and processing time: this component is the time it takes to perform a service e.g. a computer networking lab, thesis or assignment supervision.
4. Assessment and examination reporting time: The component is the time it takes from exam application to grade announcement.

5. Registration time: This component is the amount of time it takes the undecided or undeclared student to select and complete the registration for a specific pathway (major, specialization field etc.).

The complexity and uncertainty of academic processes bring about different issues such as prolonged education periods, inefficient use of HEI resources, and unbalanced course or staff scheduling (Shulock & Koester, 2014). Prolonged education periods, are constituted by delays in pathway selections, excess courses, failures in “roadblock” courses, facility constraints especially for courses that have laboratory requirements, advising and scheduling bottlenecks (Jenkins & Cho, 2013). Therefore, an alternative for improving the performance of a specific department can make changes in metrics for student success such as time to completion and in the student flows. These changes may include eliminating unnecessary activities and identifying alternative process flows. By doing this, some operations research methods like CPM and PERT can help department director.

This study is conducted in a University's Division of Lifelong Learning to understand the student flow of a specific program with respect to the FCPM and FPERT. The reason for incorporating fuzzy sets is stemmed from that fuzzy numbers are more effective for high uncertainty processes as in education. Activities affecting student flow through the learning pathways are identified and network flow diagram is drawn. 72 student's data about the times of each activities are obtained from the observed program. Findings through application of these methods, the expected time for completion of the program, the critical path and slack times are determined.

The rest of the paper is organized as follows: Section 2 presents related literature while section 3 presents an overview of CPM and PERT methods alongside with their fuzzy versions. Section 4 deals with our case study in a lifelong learning department of a university, which offers a training program for the Cisco associate-level certification. Finally, our conclusions and limitation of the study are presented in section 5 as well as some suggestions for future works.

2. Literature review:

A HEI has to perform an amount of processes and develop a variety of procedures both to ensure the fulfillment of its duties and to improve the abovementioned rates. These processes and procedures are multifaceted (Tam, 2010) and raise constant challenges to management and administration at different levels (Ibrahim, 2015) (Iatrellis, Kameas, & Fitsilis, 2019). From student’s perspective, educational processes, especially those belonging to the higher educational sector, are acknowledged as complex and characterized by the plethora of academic choices, the local constraints, the student’s needs and the dynamic nature of the defining personalized educational goals (Vizcarro

& Vos, 2002) (Iatrellis, Kameas, & Fitsilis, 2017). Of course, process complexity is not unique in education. Other industries have processes that are every bit as complex and sometimes they are spread across large geographic areas and shorter or longer periods of time.

Nevertheless, PERT/CPM as well as their fuzzy versions are not used in education. A review of the literature uncovered only few instances of PERT/CPM in education and that used PERT/CPM for curriculum design and management; there was no attempt to change the processes of education.

3. Methods:

3.1 CPM and PERT

CPM and PERT are the two of contemporary planning and scheduling techniques that are widely used in the programming of projects. CPM assumes that activity durations are known with certainty. However, PERT assumes that activity durations are random variables (i.e. probabilistic). The first step in CPM/PERT is to construct a project network. In an activity-on-the-node network format, project activities are represented by nodes and precedence relations by arcs between the nodes. Figure 1 displays a precedence relation between two activities in an activity-on-the-node format. It is said that activity 2 is a successor of activity 1 and activity 1 is a predecessor of activity 2.

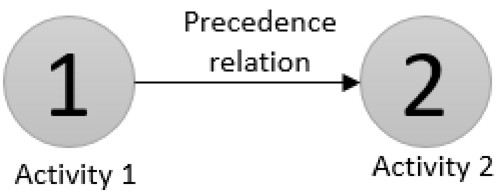

Figure 1: An example activity link in activity-on-the-node format

In an activity-on-the-arc network format, project activities are represented by arcs, as shown in figure 2. The nodes are events (or milestones) denoting the start and/or finish of a set of activities of the project and implicitly model the precedence relations between the nodes.

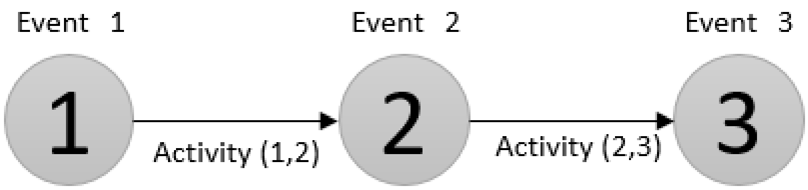

Figure 2: An example activity link in activity-on-the-arc format

The duration of the project is determined by the length of the critical path, which is the longest path of the network chart and consists entirely of critical nodes. The critical path can be identified by determining the following parameters for each task:

1. Earliest Start Time (ES): the earliest time at which the task can start, given that any predecessor tasks must be completed first.
2. Earliest Finish Time (EF): the earliest start time for the task plus the time required to complete the task.
3. Latest Finish Time (LF): the latest time at which the task can be completed without delaying the project.
4. Latest Start Time (LS): the latest finish time minus the time required to complete the task.

The slack time or float for a task is the time between its earliest and latest start time, or between its earliest and latest finish time, or, to put it another way, slack is the amount of time that a task can be delayed past its earliest start or earliest finish without delaying the project. If the earliest and latest end times are the same, the task is critical. The critical path is the path through the network in which none of the tasks have slack, that is, the path for which ES=LS and EF=LF for all tasks in the path. Any delay in the critical path delays the whole project. In order to reduce the duration of the project, it is necessary to reduce the total time required for the activities in the critical path.

A distinguishing feature of PERT is its ability to deal with uncertainty in activity completion times. For each activity, the model usually includes three time estimates:

a. Optimistic time t_0
b. Most likely time t_m
c. Pessimistic time t_p

The expected activity time (t_e) of a specific activity is calculated as in (1)

$$t_e = \frac{t_0 + 4t_m + t_p}{6} \qquad (1)$$

The variance for each is given by (2):

$$\boldsymbol{\sigma^2 = \left(\frac{t_p - t_0}{6}\right)^2} \qquad (2)$$

The completion time is determined by adding the times for the activities in each sequence. The project standard deviation can be calculated by determining the square root of the sum of the PERT variances. Eventually, a HEI decision maker can obtain the probability of completion before a specified date, the critical path activities that directly impact the completion time, the activities that have slack time and that can lend resources to critical path activities and the activities start and end dates (Trietsch & Baker, 2011).

3.2 Fuzzy versions of CPM and PERT:

Before explaining the steps of FCPM and FPERT, it is useful to review the fuzzy set theory, as developed by Zadeh (Zadeh, 1965), and the concept of fuzzy numbers persented by Dubois and Prade (Dubois & Prade, 2007). In a universe of discourse X, a fuzzy subset A of X is characterized by a membership function fA(x) which associates with each element x in X a real number in the interval [0, 1]. The function value fA(x) represents the grade of membership of x in A. The larger fA(x) the stronger the degree of belongingness for x in A.

A fuzzy number A (Dubois & Prade, 2007) in R (real line) is a trapezoidal fuzzy number if its membership function fA(x) is as in (3)

$$f_A(x) = \begin{cases} \frac{x-c}{a-c}, & c \le x \le a \\ 1 & a \le x \le b \\ \frac{x-d}{b-d}, & b \le x \le d \\ 0, & otherwise \end{cases} \qquad (3)$$

with $-\infty < c \le a \le b \le d < \infty$, the trapezoidal fuzzy number A can be represented by (c, a, b, d). Using this function, it is possible to assign a membership degree to each of the element in the universe of discourse X. Elements of the set could but are not required to be numbers as long as a degree of membership can be deduced from them. It is important to note the fact that membership grades are not probabilities since there is no requirement for their summation to be equal to 1.

A trapezoidal fuzzy number A of the universe of discourse X can be characterized by a trapezoidal membership function parameterized by a quadruple (c ,a,b,d) as shown in Fig 3, where a , b, c and d are real values.

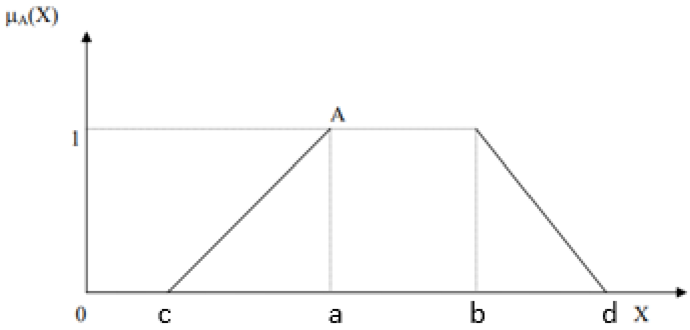

Fig.3 Membership function of trapezoidal fuzzy number A

From Fig.3, we can see that if c = a and b = d , then A is called a crisp interval; if c = a = b = d , then A is a crisp value. In Fig.4, if a = b , then A becomes a triangular fuzzy number as shown in Fig.3, and it can be parameterized by a triplet (c ,a, d).

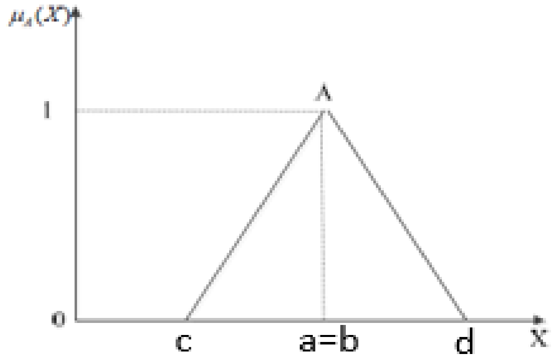

Fig.4 Membership function of triangular fuzzy number A

By the extension principle [20], the extended algebraic operations of any two trapezoidal fuzzy numbers A1 = (c1, a1, b1, d1) and A2 = (c2, a2, b2, d2) can be expressed as:

Addition ⨁:

A1 ⨁ A2 = (c1, a1, b1, d1) ⨁ (c2, a2, b2, d2) = (c1 + c2, a1 + a2, b1 + b2, d1 + d2)

Subtraction Θ:

A1 Θ A2 = (c1, a1, b1, d1) Θ (c2, a2, b2, d2) = (c1 − d2, a1 − b2, b1 − a2, d1 − c2)

The notation and process steps of FCPM are given in the following [13-16,18]:

Notation:

N: The set of all nodes in a project network
Aij : The activity between nodes *i* and *j*
FETij : The fuzzy activity time of *Aij*
FESj : The earliest fuzzy time of node *j*
ELFj : The latest fuzzy time of node *j*
FTSij : The total slack fuzzy time of *Aij*
S(*j*): The set of all successor activities of node *j*
NS(*j*): The set of all nodes connected to all successor activities of node *j*, i.e.,
NS(*j*) = *{k | Ajk ∈ S*(*j*), *k ∈ N}*
F(*j*): The set of all predecessor activities of node *j*
NP(*j*): The set of all nodes connected to all predecessor activities of node *j*, i.e.,
NP(*j*) = *{i | Aij ∈ F*(*j*), *i ∈ N}*
FCPM(*Pk*): The fuzzy completion time of path *Pk* in a project network.

Steps of FCPM

Fuzzy critical path analysis algorithm:

1. Identify activities in a project.
2. Establish precedence relationships of all activities.

3. Estimate the fuzzy activity time with respect to each activity.
4. Construct the project network.
5. Calculate β that means risk factor for each A_{ij} with (6).

$$\beta = \sum_i \sum_j \left(\frac{(a_{ij} - c_{ij})}{(a_{ij} - c_{ij}) + (d_{ij} - b_{ij})} \right) / t \quad (6)$$

where *t* denotes the set of all actives and the number of actives in a project network, respectively.

6. Let $FES1 = (0, 0, 0, 0)$ and calculate $FESj$, $j = 2, 3, . . . , n$, by using property:
   $FESj = \max\{FESi \oplus FETij \mid i \in NP(j), j \neq 1, j \in N\}$ (7)
7. Let $FLFn = FESn$ and calculate $FLFj$, $j = n-1, n-2, . . . , 2, 1$, by using property
   $FLFj = \min\{FLFk \ \Theta \ FETjk \mid k \in NS(j), j \neq n, j \in N\}$ (8)
8. Calculate $FTSij$ with respect to each activity in a project network by using property
   $FTSij = FLFj \_ (FESi \oplus FETij), 1 \leq i < j \leq n; i, j \in N$ (9)
9. Find all the possible paths and calculate $FCPM(Pk)$ by using property

$$FCPM(P_k) = \sum_{i,j \in P_k} 1 \leq i < j \leq n \ FTS_{ij} \in P \quad (10)$$

10. Find the fuzzy critical path by using definition (11) and theorem (12).
  Assume that there exists a path *PC* in a project network such that
  $FCPM(PC) = \min\{FCPM(Pi) \mid Pi \in P\}$ (11)
  then the path *PC* is a fuzzy critical path.
  Assume that the fuzzy activity times of all activities in a project network are trapezoidal fuzzy numbers, then there exists fuzzy critical path in the project network. (12)

11. Higher education decision makers can learn the probability that the project will be completed by a specified time using (13)

$Z = \frac{X-\mu}{\sigma}$ (13)

4. Application case:

The application is performed in a university's lifelong learning division. The university of Applied Sciences of Thessaly offers the Cisco Networking Academy program (CP) and prepares students for the widely accepted Cisco Associate-level certification. Candidates begin with Cisco Certified Entry Level Technician (CCENT) certification as an interim step to the Associate-level (see Fig. 5). A sample of 72 learners' data is used. The data is gathered through the university student information management system, interviews and direct observation method. The learner flow of the CP is as follows:

Candidates pursuing the CCENT can study on campus or at a distance. The first option is designed for both in-class and self-paced study and is divided into an introductory and a main module, while the second option provides a completely self-study pathway, which includes virtual learning instructor-led demonstrations and visual presentations. After obtaining the CCENT credential, which is a prerequisite for associate-level credentials, learners need to choose from several certification pathways, including cloud, collaboration, cyber ops, data center, industrial, routing and switching, security, service provider and wireless. An academic advisor is available to assist learners with selecting the best track for both the CP program and meeting their future career goals. Learners who successfully complete the training are eligible to earn Cisco Associate-level certification by passing one or two certification exams, depending on the track they choose. Finally, learners can complete the CP program and exit with three different ways: 1) leaving from university 2) register to a different lifelong program offered by the university, and 3) continue on the Professional level of CISCO certifications track.

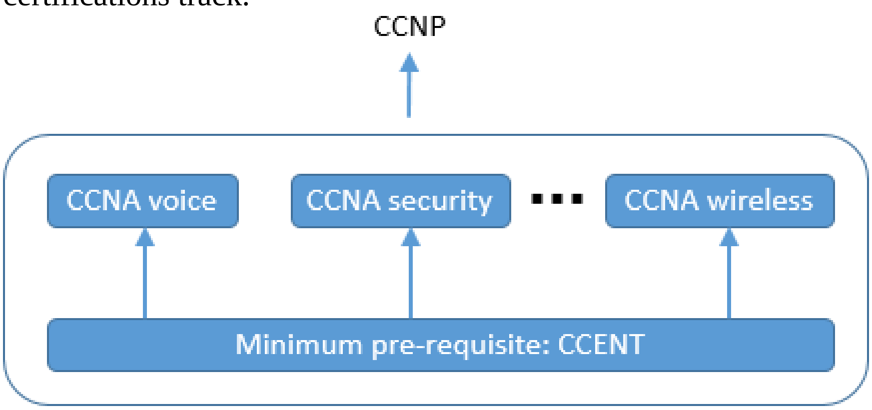

Fig.5 Cisco certification pathways

The activities of the CP learner flow are identified as shown in Table 1. The fuzzy activity times of the project and the project network chart are shown in Table 2 and Figure 6, respectively. In this paper, the fuzzy activity time is presented by trapezoidal fuzzy numbers (Han, Chung, & Liang, 2006), where a_{ij} and d_{ij} are minimum and maximum values of accessing activity time for A_{ij}, whereas b_{ij} and c_{ij} are the first quartile and third quartile of activity time for A_{ij}. If there is only one set of four historical data, the a_{ij}, b_{ij}, c_{ij} and d_{ij} can be sorted from minimum to maximum. Conversely, if one has no further information with respect to activity A_{ij}, the fuzzy activity $FAT_{ij}=(a_{ij}, b_{ij}, c_{ij}, d_{ij})$

can be evaluated subjectively by the CP decision maker based on his/her subjective judgement and experience.

Table 1. Activities of the CP

Symbol	Activity
A	Completing distance learning package (CCENT certification)
B	Completing on campus introductory training "computer basics" (in-class or self-paced)
C	Completing on campus CCENT training (in-class or self-paced)
D	EXAM (CCENT certification)
E	Academic Advising Services assisting and supporting learners in exploring available tracks and through the registration process
F	Selecting a Cisco Associate-level track
G	Completing training courses (Cisco Associate-level certification)
H	EXAM (Cisco Associate-level certification)
I	Transferring to another lifelong program offered by the institution
J	Registration
K	Terminating enrollment
L	Continue on the CCNP certification track

Table 2. FETs for each activity of the CP

Activity	Predecessor	FATs (in days)			
A	-	30	40	55	110
B	-	5	7	9	20
C	B	24	29	26	65
D	A, C	3	5	9	13
E	D	4	5	8	14
F	D	8	15	28	134
G	E, F	90	150	280	411
H	G	14	20	26	36
I	H	5	6	8	9
J	H	5	8	8	14
K	I	1	2	3	5

Fig.6 Project network chart of the CP

By using the equations in the related section above, EFT, LFT values of each note are determined as shown in Table 3.

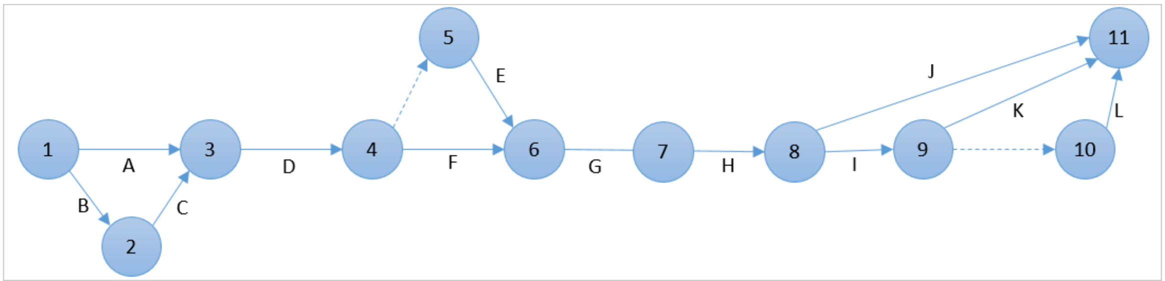

Table 3. FES and FLF of each node

Node	FES values				FLF Values			
1	0	0	0	0	-653	-223	99	457
2	5	7	9	20	-543	-168	139	487
3	5	7	9	20	-478	-142	168	511
4	8	12	18	33	-465	-133	173	514
5	8	12	18	33	-345	-113	183	518
6	16	27	46	167	-331	-105	188	522
7	106	177	326	578	80	175	338	612
8	120	197	352	614	116	201	358	626
9	125	203	360	623	125	206	364	630
10	125	203	360	623	122	203	360	626
11	130	209	366	631	130	209	366	631

Then total FTSs are calculated for each activity with (9) and shown in table 4:

Table 4. FTS of each node

Node	FTS values			
1	-563	-168	139	487
2	-588	-197	128	481
3	-543	-799	-227	278
4	-498	-151	161	506
5	-498	-151	161	506
6	-378	-131	171	510
7	106	177	326	578
8	-498	-151	161	506
9	-498	-154	161	505
10	-498	-151	161	506
11	130	209	366	631

All possible paths are found and FCPMs for each one are calculated with (10). They compared and FCPM (P_i) is obtained. FCON and R(FCPM(P_i)) values of each possible path are shown in table 5.

Table 5. FCPM and R(FCPM (Pi)) values of each possible path

#	Possible paths	FCPM (Pi)				R
1	1-3-4-5-6-7-8-9-11	-2354	-607	1108	3086	0,123492581
2	1-3-4-5-6-7-8-11	-2855	-764	1265	3587	0,022335678
3	1-3-4-5-6-7-8-9-10-11	-2357	-610	1104	3082	0,085119622
4	1-3-4-6-7-8-9-11	-2972	-781	1259	3587	0,014604829
5	1-3-4-6-7-8-11	-2474	-624	1098	3083	0,015223527
6	1-3-4-6-7-8-9-10-11	-2975	-784	1255	3583	0,088913013
7	1-2-3-4-5-6-7-8-9-11	-3370	-1531	1053	3875	0,364872082
8	1-2-3-4-5-6-7-8-11	-2872	-1374	892	3371	0,054279849
9	1-2-3-4-5-6-7-8-9-10-11	-3373	-1534	1049	3871	0,013983146
10	1-2-3-4-6-7-8-9-11	-2992	-1394	882	3367	0,020546752
11	1-2-3-4-6-7-8-11	-2992	-1394	882	3367	0,004323685
12	1-2-3-4-6-7-8-9-10-11	-2995	-1403	878	3361	0,013367419

As a result, R(FCPM(P_{11})) has minimum value among all possible paths' FCPMs. So, the critical path is 1-2-3-4-5-6-7-8-11 in this case. Time to completion period is approximately between 209 and 366 days i.e. (130, 209, 366, 631).

The activities on the critical path are: B (Completing on campus introductory training "computer basics"), C (Completing on campus CCENT training) , D (EXAM), E (Academic Advising Services), G (Completing training courses), I (Transferring to another lifelong program offered by the institution), K (Terminating enrollment). As important activities do not permit any flexibility in scheduling, any delay in any of the important activities B, C, D, E, G, I or K will delay the whole CP flow process.

Mean values and standard deviation which is calculated based on the total μ and σ values of the activities on the critical path are determined as μ=480.452, σ =180.34

5. Conclusion

The need to embrace economic policies has become inevitable in higher education sector. In order to reduce the unnecessary delays, eliminate unnecessary activities and optimize process flows in education, operations methods like CPM and PERT can play an important mission. Since deterministic approaches cannot afford to complex projects, fuzzy versions of CPM and PERT are applied in this study. A case study is carried out in a university to understand the student flow of the department for a specific program of study. Activities affecting student flow through the learning pathways are identified and project network is drawn. 72 students' data about the times of each activities are used in the analysis. Findings through application case, the expected time for completion of the project, slack times and the critical path are determined.

References:

1. Bound, J., Lovenheim, M., & Turner, S. (2010). Increasing Time to Baccalaureate Degree in the United States. *Population Studies Center, University of Michigan, Institute for Social Research*.
2. Casalone, G., & Orientale, P. (2011). Does time-to-degree matter ? The effect of delayed graduation on employment. *AlmaLaurea Inter-University Consortium.*
3. Dubois, D., & Prade, H. (2007). International Journal of Systems Science Operations on fuzzy numbers. *International Journal of System Science*, *9*((6)), 613–626. https://doi.org/10.1080/00207721.2010.492663
4. Han, T., Chung, C., & Liang, G. (2006). Application of Fuzzy Critical Path Method To Airport ' S Cargo Ground Operation Systems.

Science, *14*(3), 139–146.
5. Iatrellis, O., Kameas. Achilles, & Panos, F. (2019). A novel integrated approach to the execution of personalized and self-evolving learning pathways. *Educ Inf Technol (2019) 24:781-803*, *24*(ISSN 1360-2357). https://doi.org/10.1007/s10639-018-9802-7
6. Iatrellis, O., Kameas, A., & Fitsilis, P. (2017). Academic Advising Systems: A Systematic Literature Review of Empirical Evidence. *Education Sciences*, *7*(4), 90. https://doi.org/10.3390/educsci7040090
7. Iatrellis, O., Kameas, A., & Fitsilis, P. (2019). EDUC8 ontology: semantic modeling of multi-facet learning pathways. *Educ Inf Technol (2019)*. https://doi.org/10.1007/s10639-019-09877-4
8. Ibrahim, N. G. (2015). Process Improvement in Higher Education Institutions, (May).
9. Jenkins, D., & Cho, S.-W. (2013). Get With the Program ... and Finish It: Building Guided Pathways to Accelerate Student Completion. *New Directions for Community Colleges*, *2013*(164), 27–35. https://doi.org/10.1002/cc
10. Lane, A. (2014). Placing Students at the Heart of the Iron Triangle and the Interaction Equivalence Theorem Models. *Journal of Interactive Media in Education*, *2014*(2), Art. 5. https://doi.org/10.5334/jime.ac
11. Munns, A., & Bjeirmi, B. F. (1996). The role of project management in achieving project success. *International Journal of Project Management*, *14*(2), 81–87. https://doi.org/http://dx.doi.org/10.1016/0263-7863(95)00057-7
12. Project Management Institute Inc. (2000). *A guide to the project management body of knowledge (PMBOK® guide)*. *Book*. https://doi.org/10.5860/CHOICE.34-1636
13. Serrano, S. C., & Avilés, R. A. (2018). Project management techniques at the Complutense University: academic librarians' perceptions. *New Review of Academic Librarianship*, *4533*(January). https://doi.org/10.1080/13614533.2017.1406378
14. Shulock, N., & Koester, J. (2014). Maximizing resources for student success by reducing time- and credits-to-degree, (July). Retrieved from http://hcmstrategists.com/maximizingresources/images/Maximizing_Resources_Paper.pdf
15. Tam, M. (2010). Quality in Higher Education Measuring Quality and Performance in Higher Education Measuring Quality and Performance in Higher Education, (May 2013), 37–41. https://doi.org/10.1080/1353832012004507
16. Trietsch, D., & Baker, K. (2011). PERT 21: Fitting PERT/CPM for Use in the 21st Century. *International Journal of Project Management*, (September).

17. Vizcarro, C., & Vos, H. (2002). New developments in the assessment of learning : where are we ? *International Journal of Continuing Engineering Education and Life Long Learning*, *12*, 1–30.
18. Wächter, B., Kelo, M., Lam, Q. K. H., Effertz, P., Jost, C., & Kottowski, S. (2015). *University quality indicators: a critical assessment*. Retrieved from http://www.europarl.europa.eu
19. Zadeh, L. a. (1965). Fuzzy sets. *Information and Control*, *8*(3), 338–353. https://doi.org/10.1016/S0019-9958(65)90241-X.

Curriculum Supervision and Implementation in Kenya: The Role of Secondary School Heads

Dorothy Owuor Jonyo, PhD
Bonn Odera Jonyo, DBA
United States International University-Africa, Kenya

Doi: 10.19044/ejes.v6no2a4 URL:http://dx.doi.org/10.19044/ejes.v6no2a4

Abstract

This is a conceptual paper which reviewed the role of secondary school heads in curriculum supervision and implementation in Kenya. The main objective is on the process of execution of the curriculum with a particular emphasis on the role of secondary heads in the supervision. The paper highlights the theory which curriculum is grounded and the various perspectives by scholars over the years. The paper delves in the background, the rationale, implementation and the process including the relevant legislations. The method used was relevant scholarly publications and TSC Act. The theory used is a more instrumental way to legitimize what is considered the 'right way' to execute. From the discourse, it is apparent that there is a strong agreement among curriculum implementers that the major purpose of curriculum supervision include monitoring performance, sharing information and solving problems. We recommend that the procedure used by the supervisors should be discussed, and agreed upon by the supervisees and that curriculum cannot be discussed in isolation without elucidating the roles of the head teachers in secondary schools.

Keywords: Curriculum, supervision, implementation, process, practice, functions.

Introduction

According to Morrison (2010), complexity theory explains the curriculum phenomenon. Several perspectives and ground rules for curriculum theory building have been advanced and can be summarized by five generalizations: Curriculum theory ought to begin by defining its set of events. Curriculum theory ought to make clear its accepted values and sources for making decisions. Curriculum theory ought to specify the characteristics of curriculum design. Curriculum theory ought to describe the essential processes for making curriculum decisions and the interrelationships among those processes and lastly, Curriculum theory ought to provide for continuous regeneration of curriculum.

Curriculum supervision is a program that prepares individuals to superintend instructional and support personnel at the school building, facility or staff level. It includes instruction in the principles of staffing and organization, the management of learning activities, personnel relations, administrative duties related to departmental or unit management, and specific applications to various educational settings and curricula (National Center for Education Statistics, 2018). Curriculum supervision therefore involves observation of teaching and learning, assisting teachers in their professional development, both in individual and group context, evaluation of teachers, research and revision of the curriculum (Education Encyclopedia, 2017).

Most post-modernists have criticized the models of supervision as being rigid, classified and authoritarian. Supervision represses the teachers' independence and hypothesize further that rational-technical formations of supervision lessen effectiveness to routines which turn supervisors into domineering with the authority to diagnose teachers' pedagogical lapses and impose solutions (Glanz & Heinmann, 2018). Ovando (2000), on the other hand compliments effective supervision, and maintains that instructors, including teachers, curriculum specialists, and supervisors would collaborate in order to improve instruction. Cobbold, Kofie, Bordoh, and Eshun, (2015). advances that curriculum supervision, is a method of teaching the staff to act in a more cognizant way. Its goal is to provide curriculum implementers and supervisors with more information and unfathomable insights into what is happening around them. This increases options for instructors to work with students and superiors to facilitate effective partnership between curriculum leaders as supervisors, as teachers learn to identify and resolve their challenges. Subsequently supervisors get a new perspective about what happens in different classroom environments. Supportive supervision is a learning situation for both teachers and their supervisor.

However, to others (mostly staff and students), supervision could lead to some curriculum leaders overstepping their role expectations just to teach one a hard lesson or show where power lies. Very critical is the issue of trust between the supervisor and supervisee otherwise there is bound to be a very poor relationship, unfriendliness and suspicion.

Print (1993) defines curriculum as all planned activities which are consciously organized and systematically implemented under careful watch of instructors in the school. In other words, all conditions which yield learning by reason of the structure and organization, as well as peculiar practices of the school within and without the classroom, official timetable or the syllabus are deemed to have greater effect. Practices and activities which may be described as extracurricular or co-curricular are considered as important as those officially presented in documents.

Additional perspective of the curriculum is the totality of how learning experiences harmonizes with broad view or generalist perspective of curriculum development. It depicts curriculum as presenting opportunities and avenues by educational delivery systems through which learners attain knowledge forms, skills, values and attitudes which contribute to effective living, but not necessarily prescribed for certification. This means therefore that, learners gain experiences and develop, not only through prescribed content, but also through procedures of instructions, modes of enquiry, personal and organized interactions within the social and physical environments of learner. Curriculum experience may, hence, be gained through the planning of field trips, supervised projects, industrial attachments and competitive field games. In classroom scenarios, apart from the prescribed content itself, a wide range of techniques such as experimental procedures, role play, simulations, and group work are vital for total learner development.

Holloway (1995) proposes that, the supervisor should function in an ideal professional behavior and practice, both covertly in the managerial connection and overtly by role-playing for the supervisee.

This reduces interpersonal distance and makes exercise of power a collaborative process. This, however, requires the trust and respect of the supervisee(s) in order to engage in a more collaborative rather than antagonistic relationship.

RELEVANT LITERATURE

Rationale for Supervision

According to Hawkins and Shohet (2012), the critical focus of curriculum supervision envisages;

- Orderly environment for supervisees to replicate the content and process of their work.
- Development of skills, receiving information, other perspectives concerning the teacher's work, evaluation and feedback. This ensures that the teacher is validated and supported both as an individual and as a teacher.
- Exploration and expression of personal distress, re-stimulation, change as a consequence.
- Enhancing planning and utilization of the soft skills and professional capital of teachers.
- Ensure quality of work-specific references to functions or purposes of curriculum supervision. This includes improvement in classroom teaching and learning, assisting teachers in professional and group development, evaluation of teachers' work output, research and revision of the curriculum.

- Maintaining of standards or benchmarks, meeting delivery targets within timeframes, as well as checking resistance in teachers and learners. This calls for the need to be abreast with what constitute the historical overview of curriculum supervision.

The phenomenon of curriculum supervision therefore exists within school management, monitoring or even inspection. It must be emphasized that curriculum supervision as a field of educational endeavor with clearly delineated roles and responsibilities has also evolved slowly as a distinct practice, always in relation to the institutional, academic, cultural and professional dynamics that have historically generated the complex agenda of schooling (Education Encyclopedia, 2017). Once proper supervision has been done implementing the curriculum comes automatically as the teacher is armed with all the necessary tools.

METHODOLOGY

References was made on scholarly publications, TSC Act and several Ministry of Education Science and Technology reports.

Curriculum Implementation

Curriculum implementation involves the practice of formally approved courses of study, syllabuses and subjects. The learner acquires knowledge or experience through an organized environment with instructors with a set of prescribed rules and regulations. The centrality of the learner is emphasized in the curriculum implementation process. The process of implementation takes place as the learner acquires the planned experiences, knowledge, skills, ideas and attitudes (University of Zimbabwe, 1995). This perspective emphasizes also that curriculum implementation is an educational programme when put into effect.

THE PROCESS

Supervision/Implementation

The Teachers Service Commission (TSC) is committed to effectively manage the teaching service for quality education in line with provision of the Constitution of Kenya which recognizes that Secondary schools play a vital role in improving the quality of lives of Kenyans. Pursuant to section 11(f) and 35 (i) of TSC Act (2012), the Commission is mandated to monitor the conduct and performance of teachers in the teaching service. When signing the performance contract, the head teacher signs a 'Statement of Responsibility' where he /she states that as a head teacher, it is his/her undertaking to provide the required leadership in designing suitable plans and strategies to deliver educational services and enable this school realize its goals. In this regard, the Commission has introduced an open Performance

Appraisal System for teachers to strengthen supervision and to continuously monitor their performance in curriculum implementation at the institutional level. The appraisal system shall be used to provide feedback, improve communication, and clarify roles and responsibilities detailed below:

The school management is directly involved in curriculum implementation and supervision. This is the basic operation area of all school administration. Owen (1992) submits that all school managers must ensure the improvement in instruction by developing effective instructional leadership on the part of the head teacher. The head teacher must ensure that the educational objectives of their school and the means of achieving them are clearly spelt out by the professional staff. Ensure effective teaching is observed in the school through regular supervision of classroom instruction. As a school manager the head teacher should make teaching possible by stimulating desired changes in the professional behavior of the teachers. Effective instructional leadership demand that he or she must be a competent teacher and should keep abreast to recent developments in curriculum in general and instruction supervision in particular.

The Teacher Appraisal Performance and Development (TPAD) Tool for Heads (TSC/QAS/TPAD-HPSS/03), contains seven teaching standards with the first being professional knowledge and application where the head teacher should have the ability to prepare and supervise;

- Preparation, maintenance and use of professional documents.
- Learner's internal evaluation through development of testing policy for the school and management of national exams.
- Maintenance and tracking of learner's value added progress (VAP) records, Individualized Education Programs (IEP).

In particular, the head teacher must check the teaching standards by reference to scheme of work, lesson notes, and records of works, student's exercise books, and also participate in classroom observation. The head teacher has a role of selecting, producing and supervising locally available instructional materials/resources which will be beneficial to the teacher in curriculum delivery and improve instruction through effective instruction leadership This means that effective and efficient running of the school depends on the head teacher's instructional management role. The head teacher must schedule, assign work, coordinate and oversee performance and make sure that work is done in time.

As a head teacher one should have the ability to ensure effective time management through monitoring;

- Teacher presence and attendance to duty.
- Teacher lesson attendance.
- Adherence to school timetable · Attendance to all other school activities eg staff meetings, AGMs, academic clinics, co-curriculum activities.

Knowledge of child protection and safety legal provisions is imperative to curriculum implementation. It is necessary because when the learning environment have no safety measures and not child friendly, proper learning will not take place. It's the duty of the head teacher of the institution to ensure that there is compliance with the Children's act, Sexual offence Act. TSC Act, CORT, COCE. Equally important is learner discipline and teacher conduct in the school. All teachers and stakeholders should be sensitized on the above mentioned documents.

Co-Curricular activities in the learning institutions should be supported, coordinated together with life skills education for nurturing of talents among learners. Games, clubs and societies should be taken seriously and students encouraged to join and participate in the various activities in the school.

The success of any learning institution depends on the ability of its leadership to collaborate with patents/guardians and stake holders. This can be done by involvement in community development based activities, exchange programmes and national initiatives like tree planting, advocacy against FGM, drug and substance abuse, HIV and AIDS sensitization and prevention. All these lead to the improvement in teaching and learning, the basis upon which the head teacher performance is evaluated.

Secondary education mainly caters for primary school leavers ranging between 14-18 years. It is both terminal and preparatory for learners who proceed to the university or middle level colleges for further education and training. It therefore plays an important role in preparing learners either for the society or for further education and training. Ministry of Education Science & Technology (MoEST) launched the Transition Infrastructure Grants (TIG) in 2016 to have all primary school leavers enrolled into secondary schools. In July 2018, the Teachers Service Commission hired 8,672 teachers to support this Government policy on 100 per cent transition drive from primary to secondary schools launched in January 2019. The new teachers are supposed to plug a severe staffing shortage in public secondary and primary schools — estimated to be 155,000, said TSC chief executive Dr. Nancy Macharia. The commission had asked the government for Sh8.3 billion to recruit 12,626 secondary teachers annually for four years, translating to 50,504 teachers to support the transition but the Government only gave Sh4.7 billion, according the year's budget policy statement. This shows that transition from primary to secondary level has tremendously increased. The head teachers therefore need to double their supervisory role in order for students to realize high academic achievement and actualize for this goal. As a manager therefore, the head teacher should explain tasks to the staff and make follow up accordingly.

In terms of teacher appraisal, the head teacher should have the ability to manage the appraisal process, identify gaps and based on these gaps develop and implement a School Based Teacher Professional Development a (TPD) plan. The current introduction of performance contract by TSC requires head teachers to acquire more knowledge, skills and experience. According to Wideen, (1987), "schools now require that even experienced principals improve their leadership skills through seminars, workshops and refresher courses on classroom organization, teaching methods outcome based instruction and teacher evaluation". In analyzing the curriculum that the school wishes to deliver and the associated management structure needed to support it, Head teaches should strive to attend professional courses in matters related to curriculum instruction management. The country looks upon them to give guidance and direction otherwise the national examinations in secondary schools will always register improvement or decline depending on their commitment to the curriculum instruction management roles they undertake in their schools.

Peer classroom observation should be encouraged in the school but the head teacher can also do classroom observation in his school using the classroom observation tool designed by the TSC as part of its appraisal system with five observation areas namely;

- Introduction and lesson organization.
- Content delivery.
- Teaching methods/techniques.
- Learner involvement and communication.
- Classroom management.

There should always be a preparatory meeting before the lesson observation and a feedback session after the lesson observation. Establish good relationship with the teacher, informing the observing the teacher's performance in class, meeting the teacher in a quiet place after observation to advice the teacher accordingly. This should never be a fault finding session but a time to reflect on the strengths, weaknesses of the teacher and how to improve on them. The head teacher needs to explain to the teacher the purpose of the classroom visits to enable him/her understand the educational programme. Chiemela (2010) emphasizes that when the head teacher observes the teaching learning process he/she should have specific items (checklist) to observe like methods presentation, motivation for learning aids student interest and attention, classroom atmosphere, this ensures that the head teacher has a criterion for judging each area of the teacher's performance.

UNESCO (2018) observes that staff members will perform effectively when they are motivated in terms of; good salary, job security, regular consultation with the head, their work being appreciated, fact full discipline and when they receive sympathetic help when dealing with problems. Some of the ways in which a head teacher can meet his staff motivational needs according to the MoEST (2018) include:

- Giving teachers a chance to attend in- service training and other activities.
- Delegating important responsibilities such as chairing subject panels · Inviting role models to talk to teachers, students and supportive staff · Recognition of teachers during AGMs and other for a Good management and best practice requires head teacher to act and disseminate information promptly to avoid breakdown of communication. The head teacher should ensure good communication so that the teachers are fully informed of what is going on in the school. The manager should develop and maintain a system of communication that provide for an upward flow to benefit decision making, a downward flow to benefit the implementation of policy, and a horizontal flow to facilitate coordination of all departments of the organization. Emerson (1993) observes that successful communication requires that every member of the organization has all the information required at the appropriate time in order to undertake their duties. However successful communication demands that all members shall have the opportunity to question, comment, inform and engage. Staff relations are influenced by communication, failure to which there can be misunderstanding among the staff members. UNESCO (1993) observes that the head teachers require both oral and written communication skills in order to communicate effectively to all stake holders at different times.
- Effective communication helps to control behaviour.
- It provides feedback to the personnel on how to improve.
- It facilitates decision making and fulfil the social need of expression of feelings.

Oral communication can be done in assemblies, while written communication can be done through proposals, report, minutes, internal memo and correspondence, newsletter, suggestion box, fax, telex. To ensure effective communication, the head should ensure clarity of information and civil language. He/she should choose an adequate channel to communicate, communicate in time for the appropriate response and provide a proper atmosphere for feedback.

Wideen (1992) observes that, the teacher is the central figure in curriculum and school development. There can be no development without the development of the teacher. A teacher who is not growing professionally is unlikely to make significant development improvement in the classroom programmes. For education to meet both the rapid requirement of the demand for new curriculum and methods of teaching the teachers have to be empowered through training. In the Performance Contract signed by the head teachers and the Teachers Service Commission at the beginning of every year, Performance Criteria C4 states that the head teacher is expected to promote Teacher Professional Development (TPD) at the school level by;

Identifying teachers' professional gaps (challenging areas in pedagogy, subject content and learners' management that hinders them from effective delivery of curriculum) · Developing TPD programmes at school level to mitigate teachers' professional gaps.

- Monitoring teachers' progress in school planned TPD.
- Maintaining records of TPD.

The head teacher must accept accountability over all training, this may involve targeting training priorities, making time and resources available for training.

Conclusion

From the above discourse, it is evident clearly that we cannot talk of curriculum supervision and implementation without elucidating the head teacher roles. He /She is the engine behind the curriculum implementation and is responsible for all matters pertaining to the smooth running of the school. The priority of any government is to improve the quality of education offered and students achievements. The Kenya government therefore rely strongly on the school supervision system to monitor and evaluate both quality of schools and key measures of its success such as student achievement.

The Head teacher plays a major role in effective implementation because he is the person most likely to shape the Institution's conditions necessary for the success, such as development of shared goals, work structure, climate and procedures for monitoring results. He /She is an administrator, problem solver and a facilitator for change.

References:

1. Awiti, J. A., Onderi, H., & Raburu, P. (2016). Head' Teachers'' Supervision' of' Curriculum' Implementation:' Implications o' Provision' of' Quality Secondary Education in Kenya.
2. Chiemela, N. I. (2010). Instructional supervisory practices of Anambara state secondary
3. school principals. A comparison with desired instructional supervisory practices.
4. Unpublished Project. University of Alberta: Canada. Chan, 2004
5. Cobbold, C., Kofie, S., Bordoh, A., & Eshun, I. (2015). Perception of curriculum leaders and teachers on feedback and conditions essential for effective supervision in senior high
6. schools. *International Journal of Educational Research and Information Science*, 2(2), 21-26.
7. Complexity Theory and Macao's School Curriculum Management System. Available from:

https://www.researchgate.net/publication/261637035 [accessed May 14 2019].
8. Education Encyclopaedia (2017), https://www.newworldencyclopedia.org/entry/Education
9. Emerson, C (1993). Managing Staff in School. Paris: Athenaum press Ltd.
10. Glanz, J., & Heinmann, R. (2018). Encouraging reflective practice in educational supervision
11. through action research and appreciative inquiry. In S. Zepeda & J. Ponticell (Eds.), The
12. Wiley handbook of educational supervision. New York: Wiley-Blackwell.
13. Hawkins, P. & Shohet, R. (2012). 4 Supervision in the helping professions: An individual, group and organizational approach, Milton Keynes: Open University Press.
14. Holloway, E. (1995). Clinical supervision: A systems approach. Thousand Oaks: 2 Sage Publications Inc
15. Maranga, J. S. (1993). Guidelines for Training Educational Supervisors in Kenya, Columbia University.
16. Mbiti, D. M. (1999): Foundation of School Administration. Nairobi: 5 Oxford University Press.
17. MoEST (2017) Education Report
18. Morrison, K. (2010), Complexity Theory, School Leadership and Management: Questions for Theory and Practice; https://doi.org/10.1177/1741143209359711
19. National Center for Education Statistics, (2018).
20. Ovando, M., (2000), In J. Glanz and S. L. Behar-Horenstein. Debates in curriculum and supervision. Westport: C.T Publications
21. Owen, J (1992). Managing Education. The Purpose and Practice of Good Management in Schools. Longman, London.
22. Print, M. (1993). 5 Curriculum development and design (2nd ed.). St. Leonards, NSW: Allen & Unwin.
23. Teachers Service Commission. (2013). Annual Report 2012/2013. Nairobi: TSC.
24. Teachers Service Commission. (2014). TSC Act 2012 Revised Edition. Nairobi: National Council of Law Reporting.
25. Teachers Service Commission. (2015). Annual Report 2014/2015. Nairobi: TSC.
26. Teachers Service Commission. (2016). Annual Report
27. TSC/QAS/TPAD-HPSS/03 (2018) ... Teacher Performance Appraisal and Development.

28. 2019 Global Education Monitoring Report,m Regional Launch Nairobi, Kenya – 20 November 2018.
29. Wideen, F. M. (1992). The Staff Development for School Improvement. A Focus on the Teacher, New York: London Simon Fraser Press.