Nepalese Journal of EDUCATIONAL ASSESSMENT

NEJEA
ERO

Volume 2 • Issue 1
2017 June

Government of Nepal
Ministry of Education
Education Review Office
Sanothimi, Bhaktapur
Nepal

NEPALESE JOURNAL OF EDUCATIONAL ASSESSMENT
Volume 2 • Issue 1 • 2017 June

ISSN: 2467-9488

Nepalese Journal of Educational Assessment

A Publication of the Education Review Office, Sanothimi, Bhaktapur

Volume 2 • Issue 1
2017 June

Government of Nepal
Ministry of Education
Education Review Office
Sanothimi, Bhaktapur
Nepal

Year of Publication:

June 2017 (Ashad 2074 BS)

Nepalese Journal of Educational Assessment, 2017, Volume 2, Issue 1

Printed at: Dharma Printing Press, Kathmandu

ISSN: 2467-9488

Nepalese Journal of Educational Assessment

Education Review Office

Sanothimi, Bhaktapur, Nepal

Phone: +977-1-6633556, +977-1-6633557, Fax: +977-1-6633556

Email: nejea.ero@gmail.com; Website: www.ero.gov.np

NEJEA
ERO

Nepalese Journal of Educational Assessment
Volume 2, Issue 1, June 2017

EDITORIAL

Education Review Office (ERO) has been conducting various forms of educational assessments including National Assessment of Student Achievement at various Grades of school education and Performance Audit of educational institutions. Besides, ERO has been carrying out Early Grade Reading Assessment and conducting research studies on educational issues. The main aim of these assessments and studies is to provide feedback to the government as well as concerned agencies and persons for enhancing quality of and equity in education with improved accountability.

Assessment results can be interpreted using either norm referenced or criterion referenced method. The first one is an interpretation based on the achievement of other students or sometimes comparing the result with the average of the same group (the norm). The latter one is an interpretation of assessment results by comparing with the pre-determined criteria and standards. The previous assessments conducted by ERO at Grade 3, 5 and 8 two times each were norm referenced. But ERO has begun criterion referenced assessment at Grade 8 this year with the aim of continuing this kind of assessment at grade 5, 8 and 10 in the next round. Along with developing assessment framework for national assessment of student achievement, development of performance audit framework and early childhood learning and development assessment framework have also been completed which are expected to help enhancing the quality and creditability of these assessments.

This issue of journal shares some innovations and good practices from the assessments conducted by Education Review Offices together with the results of some other studies and works related to educational assessment. This journal is disseminating and extending the knowledge and technology developed in the field of educational assessment, especially student assessment and performance audit of educational institutions, therefore this can be considered as a forum for learning and knowledge building. It will also support creating an opportunity for sharing innovative practices, techniques and technologies, and reviewing the knowledge and practices in the field of educational assessment. At this point, we take the opportunity to acknowledge all the contributors including authors and reviewers of articles.

We believe that readers in the field of education in general and educational assessment in particular will be benefited from this issue of journal. Constructive suggestions and feedbacks from readers are always welcomed for further improving its forthcoming issues. We also invite scholarly articles from all authors to support our endeavour to build and share knowledge, techniques and innovative ideas in the field of educational assessment.

June 2017

Nepalese Journal of Educational Assessment

• Volume 2 • Issue 1 • 2017 June

Table of Contents

Second Language Vocabulary Assessment: A Quest into Instrumentation for Addressing Receptive and Productive Dimensions

Dr Binod Luitel

Associate Professor, CERID, Tribhuvan University

binodluitel@yahoo.co.in

Abstract

An attempt has been made in this paper to explore into the literature of second language vocabulary assessment and analyze the instruments used for testing word-meaning. Five different formats of tests and their variants have been intensively analyzed, and their strengths as well as limitations are examined – considering the potentiality of measuring receptive and productive dimensions of vocabulary knowledge. Different types of testing tool were found useful for assessing different aspects of vocabulary competence; and it was concluded that one particular tool cannot address all aspects. An integrated framework for word-meaning assessment is proposed in the end – which ensures the coverage of the two major dimensions. It is expected that the framework can measure the aspects of 'awareness', 'comprehension', 'abstraction', 'controlled production' and 'free production'.

Key terms: Vocabulary competence, receptive dimension, productive dimension, awareness, comprehension, abstraction, controlled production, free production, ACACF framework

Introduction

Needless to say, words carry meanings. While reading or listening, one can hardly grasp knowledge in any subject matter without knowing the vocabulary items used therein. In this way, we can see vocabulary as an essential component in language. Highlighting the importance of this component while learning a second language (L2), Read (2004, p. 146) states: "…Lexical items carry the basic information load of the meanings they (L2 learners) wish to comprehend and express."

Thus, the relevance of vocabulary instruction and testing is quite obvious in teaching-learning – as we have experienced that a learner is unable to express ideas unless s/he can use the vocabulary items that carry the ideas. So, vocabulary should be considered an essential part of teaching as well as assessment in general and language pedagogy in particular – whether we talk about it in the context of classroom instruction, or from research perspective. Read (2000, p. 40) has emphasized the dual purpose of vocabulary assessment – for teaching and research. While its assessment for measuring what learners have achieved in their learning career is important in

teaching, the task of exploring vocabulary acquisition on the basis of the data derived from testing has great relevance in research.

In the contemporary literature on vocabulary competence, we can see theoretical discussions regarding the quality of vocabulary knowledge – whereby 'receptive-productive continuum' has been an important phenomenon. Accordingly, different stages of development in vocabulary competence can be conceptualized. In language learning, it is experienced time and again that knowledge of a word-meaning is not always 'all or nothing' phenomenon. As Waring and Takaki (2003, p. 133) have admitted, there are several "…stages, levels or degrees of word knowledge. These could range from knowing only that you have seen or heard the word-form without being able to recall the meaning, to a full understanding of the word and its various nuances…."

In this way, the dimension of 'reception' involves learner's ability to recognize word-meaning association, while 'production' entails one's ability to use the item when needed. Nation (2001, p. 27) has stressed the point that, whatever the tiny aspect of lexical knowledge (word form, meaning, word associations, uses, collocations, constraints on use, etc.) being considered, vocabulary competence involves knowing these aspects receptively as well as productively. It is generally assumed that one becomes able to use an item productively only after acquiring it receptively but not the other way round. This continuum has also been referred in terms of the 'strength' of lexical knowledge (Laufer et al., 2004) – whereby the items at receptive level are considered having poorer strength compared to those at productive level.

Once we take this point in consideration, we will be in a position to realize the need for testing vocabulary items by using receptive as well as productive measuring instruments. And, the two types of assessment inform us with different pictures regarding the strength of learner's word meaning knowledge – which have very significant implications in the context of second language pedagogy.

Statement of the Problem

The need for considering 'receptive-productive continuum' discussed above demands the concern towards exploration into the various types of instruments that were employed for L2 vocabulary testing in the past. So, examining the tools employed in the course of vocabulary testing and word-meaning learning research, synthesizing the findings of such exploration, and then consolidating them towards modeling new framework/s of vocabulary assessment are some of the direct concerns in the context of L2 pedagogy. The study reported in this article is an attempt towards this direction.

Objectives of Study

The study has been undertaken to accomplish the following objectives:

- To explore into the literature on language testing, and analyze the specific tools employed for assessing the receptive as well as productive dimensions of learner's vocabulary competence in second language; and

- To model an integrated framework of vocabulary assessment for covering the two dimensions, based on the findings of exploration

Methodology

Literature was the principal source of secondary data for the collection of information required in this study. Books on English language teaching and testing, scholarly journals in the same field, and research reports/dissertations etc. were consulted for data collection. Different types of vocabulary questions described in these sources were then analyzed as required to address the research objectives. In this analysis, basically the characteristics of the testing tools were thoroughly examined and the write up was prepared with relevant sub-headings. Details of the qualitative description-cum analysis are presented in the section that follows. Major findings of analysis are synthesized, and then a new framework of assessment is presented based on the synthesis.

Exploring Literature: Instrument Analysis

The available literature was explored from 1960s to the running decade of the 21st century, and the relevant tools and instruments employed for vocabulary assessment were studied. Some important ones are described with critical scrutiny in this section, as presented in the sub-headings below.

Matching format

The practice of listing words in a column and meanings in the next one, and then assigning the task of showing connection between the two has prevailed for a long time in language teaching; and the same has been used as a vocabulary testing tool. This format has been demonstrated in the works of various scholars (e.g. Heaton, 1975; Ur, 1996) – whereby the testees are instructed to match the two sides on the basis of meaning relation – synonym, antonym, hyponym, or short definition, etc. This format is usable for testing multi-word lexical items as well. In the context of English as a foreign language (EFL) among Chinese learners, Li (2004) has used idiom-definition matching for the assessment of idiomatic words and phrases.

Matching test has been employed for research purpose as well – typically for assessing the learners' receptive/passive knowledge of word-meanings. Fan's (2001) use of this test, or Li's (2004) work mentioned above are examples. 'The Vocabulary Levels Test', a widely established

test (proposed by Nation, 1990), has also employed this tool for research purpose.

This kind of test seems highly credible so far as the measurement of learner's ability to recognize word-meaning is concerned. Thus, it works as a valid tool for receptive testing of vocabulary – whereby it can inform us the status of meaning abstraction in learner's mind. However, as it involves basically "…a recognition rather than a recall task…" (Read, 2000, p. 171), it cannot demonstrate the learner's ability to use the word for communication or overt expression. With a clear indication of being unable to assess the productive dimension of vocabulary competence, as such, we cannot expect that one's successful performance in this tool can predict the occurrence of the item in their speaking or writing. Moreover, it tests learner's word-meaning knowledge in a de-contextualized manner; so we cannot predict reading comprehension in context simply by seeing the learner's achievement tested through this tool. In order to address these lacunas, some other instruments are required.

'Yes/No' (or checklist) format

A long list of words is provided to the learners in this kind of test; and they are instructed to respond simply by indicating whether each of the words is known or unknown, by marking 'yes' or '√' (if the item is known) and 'no' or '×' (if unknown). Instead of asking if the items are known or unknown, testees can even be instructed to indicate whether they have heard or read the words. This format tries to correct the problem of reliability in marking 'yes' or 'no' by including several 'non-words' (sequence of letters that do not form any real word in the language) within the list of the words being tested. For example, if we are preparing a list of 50 English words for testing, some letter sequences like 'wascotil' and 'chetamor' (both are non-words) etc. can be included in the list. When the testee marks 'yes' in the real words, the cases are called 'hits' and if non-words are marked 'yes', they are called 'false alarms' (Meara et al., 1994). We can say whether one's response in the test is reliable or not after seeing the number of false alarms, if present at all in the response.

In a slightly modified fashion, Shu and others (1995) had implemented the 'checklist test' model of vocabulary testing by using two kinds of non-words within the list of real words: (i) letter sequences which do not contain any meaningful item (neither a word, nor any free or bound morpheme) – 'ushom' for example; and (ii) letter sequences having the real 'base' followed by meaningless derivation (thus becoming non-words), which they have called 'pseudo words', e.g. 'birdable'.

Seemingly a "simple minded idea" (Meara and Buxton 1987) rather than being a 'test', in the beginning, one can say this format simply collects the learner's reported responses but cannot give evidence of word-meaning knowledge. However, its use has much been appreciated in vocabulary learning research; and the practice of using non-words is found in reading passages as well. For example, Pulido and Hambrick (2008) have constructed such words according to

orthographic and morphological rules of Spanish, and used them in reading text.

One of the great problems in testing vocabulary size is related to covering a very large number of words within a limited time. A checklist test of the kind just described can solve this; that is why this format has also been called a "promising innovation" of L2 vocabulary testing (Meara et al., 1994, p. 298). A lengthy tool like multiple choice cannot help when we have to test a large number of items. Checklist format can be employed for seeing the learner's awareness on given items – even if a few items are included in the test. From a learner's positive response to a word in this tool, we can say that at least the learner is aware of the fact that the word exists, and has some vague idea regarding its meaning.

As Meara and Buxton (1987) pointed out, there was a good correlation between multiple-choice test and checklist test. This tool has been used very widely among native speakers since 1990 (Read, 2000, p. 88). Among the non-native speakers, the 'Eurocentres Vocabulary Size Test' developed in the 1980s used it for assessing the learners' vocabulary size. (ibid., p. 126-27)

Despite the benefits, there are some points of caution in using this format. The most important thing is that we can assess only the superficial knowledge of words with this tool. In-depth understanding cannot be measured, nor can it be used to test word-meanings in context. Therefore, it cannot be a valid tool for testing the quality of word-meaning knowledge. We must use other testing tools for assessing the depth of word-meaning knowledge.

Multiple choice formats

Tests in this format have been used for a long time in vocabulary testing. As Lado (1961, p. 188) had admitted, the multiple choice format has "…probably achieved its most spectacular success in vocabulary tests." Following Meara and others (1994, p. 297), this test has been "the most widely used type of vocabulary test".

This format involves two parts: the stem (or problem), and alternatives. The testees are instructed to respond after choosing the best option from the given alternatives. Considering the nature of contents presented in the stem and alternatives, some varieties of this format are found in vocabulary testing practices – including the ones introduced below.

i. Picture in the stem, followed by 4 alternative words as options – whereby the testee chooses the word that suits the given picture (Heaton, 1975)

ii. Definition in the stem, followed by 4 alternative words as options – whereby the testee chooses the word that suits the given definition (ibid.)

iii. One word in the stem, followed by 4 options that can include definitions or words (synonym,

antonym, or other type) – whereby the testee chooses the definition or word from the list of options that suits the word given in the stem (Heaton, ibid.; Lado, 1961)

iv. A sentence in the stem with an italicized word, followed by 4 alternative definitions or words as options – whereby the testee chooses the best option that suits the italicized word (Heaton, ibid.; Ur, 1996)

v. A sentence with blank space in the stem, followed by 4 alternative words as options – whereby the testee fills in the blank by choosing the appropriate word (Buchanan, 1992; Akbari, 2008)

vi. A short paragraph (with 2-3 sentences) in the stem with a blank space for a word, followed by 5 alternative words as options – whereby the testee makes choice that is appropriate in the blank space as per the context (Luitel, 2015b, p. 155)

Considering Read's (2000) two-way distinction under 'receptive' dimension of vocabulary competence (recognition versus comprehension), the first three instruments clearly fall under 'recognition' category while the last three can be categorized under 'comprehension' – as comprehension involves knowing the item in context and 'recognition' requires understanding word-meaning association isolated from context.

Among these, the first model looks more concrete than others, and accordingly it is applicable particularly for testing concrete words. The others can be applied for testing abstract words as well. The second and third models are helpful to test word-meaning association in absence of context. They can serve the purpose of testing the ability to recognize abstracted meaning of word. So, considering the focal content of the test, these two do not seem to be much different from matching test. The only difference is that the testee has to face a single stem for 4 options in these two models while matching test has, in essence, multiple stems for multiple options. Thus, it seems the difference is in format rather than in the content being tested.

The fourth, fifth and sixth models have operationalized the idea of testing word-meanings in context. While the fourth one is concerned with the testee's ability to recognize the word's meaning stated in terms of definition or synonym, the last two models aim to test the ability of recognizing the word that best fits in given context. Giving a short paragraph for context rather than being limited within a sentence, the last model has been developed to test more abstract, formal and academic word-meanings – whereby somehow extended linguistic expression is required to make the context more explicit for the occurrence of targeted item. Besides, this model includes 5 options – with a view to minimize the effect of 'luck' or 'random guessing' while responding.

Cloze formats

A text of appropriate length is selected for test preparation in this format. Then words in the selected text are deleted at fixed interval. In the traditional cloze format, there are practices of

deleting every 5th, 7th, or 9th word and leaving a blank space for each of the deleted items in the text. The testee is instructed to supply the missing word in the blank spaces. An example of this format looks like the following.

> Koko, a young gorilla, was three _____old when Patterson, a graduate student _____Psychology at Stanford University, first saw ____. Patterson wanted to try to communicate _____ with Koko. She asked the zoo _____ for permission to work with Koko. _____ first, the zoo authorities refused. Soon _____this, an illness spread through the _____ population at the zoo. Koko almost_____, and her mother could not take____of her. She had to be_____from the other gorillas. (Adapted from: Awasthi et al. 2009)

For proper responding, this test demands the comprehension of the surrounding information (Read, 2000) as well; so cloze has been used not only as a vocabulary testing tool but also for the testing of reading comprehension. The assumption is that it can work as an instrument for holistic testing of grammar, vocabulary and overall comprehension (Madsen, 1983, p. 7).

The applicability of this format has been well appreciated in testing; but an important problem is that most of the times the targeted item does not occur in the 'fixed interval'. Therefore, it is difficult to ensure the testing of the vocabulary items that are specifically targeted by the test designer. To resolve this problem, some modified versions of cloze test have been developed and used. The following are some examples.

(i) *Selective deletion cloze,* whereby, instead of the way followed in systematic deletion (which is a bit 'mechanical'), words in the text are deleted as per the designer's decision. This model was developed for testing the words actually targeted by the designer (Read, 2000). The target words can be decided as per the need for testing – considering what must be tested.

(ii) *'C-test',* whereby the test designer systematically deletes the second half of every second word from the text, and testees are instructed to complete the blanks. Having unnecessarily more deletions in the text, this format is highly demanding and causes confusion most of the times; so there are more demerits in it than in the original unmodified cloze format described above. In addition to being less practicable in this way, it cannot ensure that the items intended by the test designer are tested and the unnecessary ones (those not intended) are left without being tested. Thus, this model is said to be "least promising as a specific measure of vocabulary" (ibid., p. 111), unless modified.

(iii) *Modified C-test,* which can be called the 'C-test version' of the selective deletion cloze (Read, 2000) introduced above; so words are deleted as per the test designer's decision but the beginning part of the deleted word is maintained. This model can be used for testing comprehension of word's meaning as well as learner's accumulation of productive knowledge

of the word. If the deleted items are not so many, the possibility of learner's confrontation with many word-meanings at a time can be minimized in this model; and the confusion caused will be decreased.

(iv) *'Simple completion',* in which the beginning letter/s are supplied leaving the rest part of the word blank (Madsen, 1983). The focus here is the testing of word derivation rather than meaning recognition or word production – whereby the testees are instructed to supply the missing suffix of the base word. For example, blank space is left in the place of 'ive' for 'supportive' (i.e. the problem word is presented as 'support___' in sentence).

(v) *Modified simple completion* (Luitel, 2015a), whereby some features of simple completion are maintained; and only one word is deleted in the short paragraph (with 2-3 sentences) given to provide the context. Here, the beginning letter/s (1-3 letter/s depending on the length of target word) of the word that fits in the blank is/are provided. The focus of this model is on testing word-meaning rather than morphological derivation.

(vi) *Definition-resource cloze* (Malagarriga, 1998), whereby blank spaces are left for selected content words in the text; and definitions are given after the cloze text in the order that corresponds to the correct words to be supplied as answers in the blank. This model essentially tests the ability of producing words with the help of the given context and definitions.

Except for 'simple completion' model, we can notice strong lexical focus in all these cloze formats. While we may not become ready to prefer the use of traditional unmodified ('original') model exemplified above and the 'c-test' model on practical ground (as argued in the respective paragraphs), the others (selective deletion cloze, modified c-test, modified simple completion, and definition-resource cloze) can be used for preparing the testing tools as per need and as per the competence level of testees.

These cloze test models are valid as well as useful for assessing learners' vocabulary competence – whereby word-meanings are tested in "contextualized way" (Read, 2000, p.14). Among the 3 basic stages of the development of word-meaning knowledge pointed out by Laufer (1998) (which include: receptive, controlled productive and free productive), we can clearly notice the potentiality of assessing 'controlled productive' knowledge in these cloze modalities – whereby the tool can test how far the testee can produce a word when the occurrence of the word is prompted by a task.

As such, a controlled productive testing tool such as cloze format essentially "prompts subjects to produce predetermined target words" (Meara and Fitzpatrick, 2000, p. 20). The problem, however, is that using an instrument of this kind cannot help us in determining how far the learners can use the words voluntarily in free production. Therefore, some other sort of test is essential for eliciting free productive vocabulary from the respondent.

Translation-based formats

Relevance of translation in second/foreign language testing has been recognized for a long time. Some modalities of translation test that demand the testee's ability of linking the first language (L1) and target language, as described in the literature, include the following.

(i) *L2-L1 translation:* This involves the task of translating from target language to the testee's native language; and it can be used for testing receptive vocabulary. Lado (1961) pointed out that it can be a useful tool if other ways of testing could not be applicable. By providing the target word in the stem and giving alternative words in testee's L1, Lado's (ibid.) demonstrated the use of L2-L1 translation in multiple choice format. In the model implemented by Webb (2008), translation has been used for vocabulary testing in two ways: (a) Testees are given the L2 word forms and instructed to translate into their L1. (b) Similar to Lado's model just described, the testees are instructed to recognize the L1-translated version of L2 word form – that is part of the list of 4 options.

(ii) *L1-L2 translation:* Here the testee's native language words are given for translation into the target language. If other tools could not serve the purpose, its use for the measurement of productive vocabulary has been recommended (e.g. by Lado, ibid.) – whereby the L1 equivalent of the targeted L2 word is given and testees are instructed to supply the target language word.

With some adaptation in this model, the author has employed the idea of translation in vocabulary learning research among Nepalese learners of English (Luitel, 2005; 2012, 2015b) – whereby a short sentence, rather than isolated word, is given in L1 (where the L1 equivalent of the targeted L2 word appears). The testees are instructed to write in L2 by translating the sentence from L1. The research interest here is – whether the testee uses the targeted word in the L2 sentence while translating.

The author developed the model just mentioned taking into account two important theoretical conceptualizations: (i) Read's (2000) idea of two way distinction within 'productive' vocabulary knowledge that includes 'recall' and 'use' of vocabulary items, and (ii) Laufer's (1998) idea of 3 basic stages of word-meaning knowledge mentioned above: 'receptive', 'controlled productive' and 'free productive'. First, testing isolated words through L1-L2 translation could be a recall test – which can simply demonstrate successful rote memorization. Employing translation at sentence level, the author has been inclined towards testing how far the testee can use the word in context. Secondly, with reference to Laufer's framework, learner's ability to use a word in sentence (in whatsoever way demonstrated) can be considered 'free productive', since its use occurs here as a consequence of the learner's own choice rather than as a result of prompted stimulation or any sort of rote learning. We can be more confident about the learner's productive competence in word-meaning when such a

free choice of words is demonstrated by the learner in a more prolonged linguistic expression such as a sentence – whether written directly in L2 or translated from L1. This may not be the case for translating words in isolation.

(iii) *Dictation translation:* This has been used for testing productive vocabulary, whereby the teacher dictates the words orally in the testees' L1 and they should write the L2 equivalents by translating the dictated words (Ur, 1996). This can also be used as a receptive testing tool by instructing the testees to write the L1 equivalents of the L2 words that are dictated.

Overall, bilingual testing of vocabulary is justified on practical grounds (Read, 2000) when learners have a common native language and the teacher is well proficient in the learners' language. Translation test has been proved a useful research tool. In the work of Ellis (1995), for instance, L2-L1 translation was employed as a testing tool in vocabulary acquisition research among Japanese learners of English.

Findings of Analysis: A Synthesis

Each of the tools reviewed so far has their own typical characteristics, uses and limitations. As we understand, coverage of a wide range of items is possible only through checklist test. But this does not go beyond the estimation of word-meaning awareness – failing to concretely demonstrate the status of deep level knowledge. Though L1-L2 translation at sentence level is usable to see the learner's choice of targeted items in production, we must remember that translation is a special skill of linking two languages; so this tool cannot test word production unless the testee can establish link between L1 and L2 words. Use of cloze is desirable as a complementary tool to test productive knowledge of words, but this tool can test the production of targeted word only when it is 'prompted' by the clue. Multiple choice can assess the comprehension of word-meaning in context, but not production – so, using a test of this kind, we cannot say how successfully the testee can use the words in writing or speaking. In the same way, despite its success in informing the status of meaning abstraction, matching test cannot demonstrate the learner's ability to use the word for communication or overt expression.

Thus, different testing tools have attempted to assess learner's word-meaning knowledge in different ways; and a tool can test only some particular attributes of knowledge. Change in the tool (even if they are within the same category – whether 'receptive' or 'productive') causes difference in the attribute of word-meaning knowledge to be measured. Taken an instrument at a time, we can test either the receptive or productive dimension of vocabulary competence but definitely not the both. This reality implies the need for employing multiple tools and looking for some framework that can ensure somehow integrated assessment of several aspects of word-meaning knowledge.

Towards integrated assessment framework

In the light of insights from the analysis of tools, an integrated framework for vocabulary assessment is presented below.

Modeling

In addition to ignoring the need for critically examining the characteristics of different tools, people may not realize the importance of covering many aspects of vocabulary knowledge in testing. In such cases, instead of employing multiple testing tools, one will be contented by a single tool; and the choice of tool depends solely on the user's personal interest or preference.

But, as stressed above, one cannot claim that a typical instrument will suffice in assessing all the attributes of vocabulary knowledge. Therefore, in vocabulary learning research, it is often stressed that using a single test means the possibility of looking at only one type of word knowledge gain in learners – which is "a shame because it limits us to a one-dimensional picture" (Waring and Takaki, 2003, p. 133) of vocabulary competence. To quote from Read (2000, p. 149), "There is no one perfect test and it is necessary to develop a whole range of instruments to address the various purposes for vocabulary assessment." In this connection, it would be relevant to remember Laufer's (1998, p. 257) proposal for 'multiple test approach', whereby we are suggested to use "…a battery of tests where each test measures a different aspect of vocabulary knowledge." Thus, the important point is: If we want to examine different aspects of word-meaning knowledge for research purpose, a judicious selection from the list of several tests will be the sole strategy to follow.

With this consideration, an attempt has been made hereby to model an integrated framework of vocabulary assessment by judiciously incorporating the important formats examined above. This model incorporates 5 instruments, each of which aims to test one specific aspect – where three instruments are concerned with 'receptive' and the rest two with 'productive' dimension of word-meaning competence of L2 learners. Here, a content word is tested using all the 5 instruments; so, a vocabulary testing package developed in this framework will be useful to measure the status of learners' word-meaning knowledge more comprehensively. And, it shows more precisely whether the item tested has a poor or rich profile in 'receptive' as well as 'productive' dimension. In addition, it also demonstrates the profile in the tiny aspect of word-meaning knowledge specific to the tool employed in testing. As the figure below illustrates, the 5 tools in this framework are anchored to the same vocabulary item – whereby one question (or problem) is used for testing a word through each of the tools.

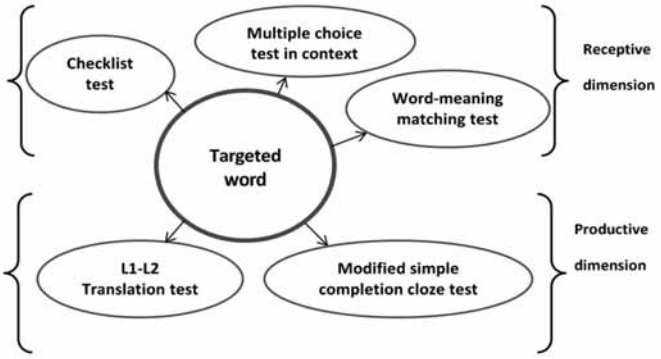


Fig. Depiction of integrated assessment modality

Exemplification

As the 5 instruments depicted in this framework have already been introduced previously, no description about them is required here. However, they are mentioned once again as given below – indicating the specific functions they perform in testing the 'receptive' or 'productive' dimension of word-meaning competence. Examples are presented for each of them.

(i) *Checklist test* (for measuring *word awareness* under 'receptive' dimension; and for covering a wide range of items in a test)

Example:

Tick (√) the item you know and signal with cross (×) if you do not know it.

abduct.	adjacent.	assembly. . . . .
abundant.	aptacol.	cendran. . . . .
academic.	ambiguous.	coincide. . . . .
actrosin	approximate.	compound. . . . .
acclaimed.	arotis.	chatemor. . . . .
accumulate.	arbitrary.	console. . . . .
antoric.	aspiration.	waskotil. . . . .

(ii) *Multiple choice test in context* (for measuring *comprehension of meaning* in context, under 'receptive' dimension)

Example:

Tick (√) the best answer.

1) Michael is expert in playing chess, and he has won several competitions and got medals. Last month he said to me, "I have ……. 21 gold medals altogether."
   a) dominated b) contemplated c) promoted d) investigated e) accumulated

2) I saw an orphan child in the hostel. He cried saying, "My friends' mothers come to meet them here, but my mother doesn't come." A lady teacher…… him saying, "I'm your mother." Then he became quiet, as he did not know his real mother.
   a) predicted b) contrasted c) precluded d) consoled e) exposed

(iii) *Word-meaning matching test* (for measuring *abstraction of meaning* under 'receptive' dimension)

Example:

Read the words in column A and the meanings in column B. Then fill suitable letters (a, b, c, etc.) from column B in the blanks after words in column A.

A	B
1. accumulate ….	(a) intended for giving punishment to people
2. gorgeous ……	(b) the group of people (or activities) that one lives in
3. console ……..	(c) strong feeling of friendship and support between people
4. milieu ……..	(d) to collect or gather something over a period of time
5. fraternity ……	(e) a proverb teaching some cultural value
6. punitive …….	(f) trying to make someone feel more comfortable
	(g) attractive and beautiful (giving pleasure)

(iv) *Modified simple completion cloze test* (for measuring *controlled production* of words, under 'productive' dimension)

Example:

Fill in the space to complete the word that starts with the letter/s as given in the beginning.

1) Usually, smugglers have linkage with some other persons who play indirect role in the business. Their income is distributed in the network. So, an individual culprit does not acc…….. the entire 'black money' collected from such business.

2) That old man struggled hard for educating his children, but none of them looked after him. When he was seriously ill, none of his relatives showed any concern on his problem. He started weeping; but no one co…….. him.

(v) *L1-L2 translation test* (for measuring *free production* of words, under 'productive' dimension)

Example:

Translate the Nepali sentences into English.

१. मेरो भाइले फुटबलमा जम्मा १५ वटा पदक आर्जन गरेको छ।

२. दु:ख पाएको मान्छेलाई हामीले सान्त्वना दिनु पर्दछ।

Discussion

As demonstrated above, this framework requires the use of multiple tools for testing different aspects of the same word-meaning. To explain how integration of the instruments is possible in this framework, it would be relevant to draw attention to the 2 words presented for testee's response in the checklist test exemplified above, and then see the same items in the rest 4 tools exemplified thereafter. These words are 'accumulate' and 'console', which have been targeted for testing in 'multiple choice', 'word-meaning matching', 'modified simple completion cloze', and 'L1-L2 translation'. As explained previously, occurrence of these words in translation test here is a bit indirect. Here, if the testee has gained productive competence in 'accumulate' and 'console', s/he can use these words while translating the Nepali sentences into English – for the Nepali equivalent 'cfh{g u/]sf]' in the first sentence and ';fGTjgf lbg'' in the second sentence respectively.

Due to the limitation of space, many items could not be demonstrated above for exemplifying their testing through 4 important tools (multiple choice, matching, cloze and translation). But it must be remembered that this can be done for all the items contained in the checklist test above, except for the non-words' (e.g. 'actrosin' or 'aptacol') that are included simply for ensuring the reliability of testee's response in the real words – as described earlier. On the whole, it can be claimed that most of the English content words are testable in the way demonstrated above.

Regarding the words 'accumulate' and 'console' as occurred in the testing tools above, we can say that the following aspects of vocabulary knowledge are measured with the application of this integrated framework.

(i) Whether the learner has heard or found the word, and whether s/he is aware of the vague idea about its (potential) meaning – i.e. awareness, measured through checklist test;

(ii) Whether the learner is able to comprehend the message contained in the word as it occurs in written paragraph – i.e. comprehension, measured through multiple choice test;

(iii) Whether the learner is able to recognize the meaning stated in terms of abstract definition, or whether s/he has conceptualized the association between word and meaning – i.e. abstraction of meaning, measured through matching test;

(iv) Whether the learner is able to produce the word in meaningful context if prompted by a task – i.e. controlled production, measured through cloze (modified simple completion cloze) test; and

(v) Whether the learner can use the word meaningfully on his/her own choice – i.e. free production, measured through translation test.

In this way, the entire range of receptive-productive continuum can be covered by this integrated framework of vocabulary testing, which can be abbreviated as ACACF framework (with the initial letters used in the terms 'awareness', 'comprehension', 'abstraction', 'controlled production' and 'free production'). With the application of this framework, the strength of word-meaning knowledge (which is considered as an important concern regarding the quality of vocabulary knowledge) can be assessed successfully.

Concluding Remark

Considering the need for innovative ways of vocabulary teaching for learner's enhanced language competence, one of the most important things is the accurate, precise and in-depth diagnosis of problem in learner's word-meaning knowledge – which is possible only by judiciously employing a range of instruments, as discussed in this paper. The integrated framework presented here can, in fact, revolutionize vocabulary assessment – in the context of classroom teaching as well as research in second/foreign language learning. It can be suggested that this framework is implementable for vocabulary testing not only in English but also in other languages being learned as second/foreign languages.

References[1]

Akbari, O. (2008). Teaching vocabulary items through contextualization and picture to elementary Iranian EFL students. (University Putra, Malaysia) *The Asian EFL Journal* 10.3. Pp. 53-77

Awasthi, J. R.; Bhattarai, G. R.; Khaniya, T. R.; Mishra, C.; Rai, V. S.; Giri, A.; Maharjan, L. B.; Bhattarai, A.; Bhandari, B. M.; and Gautam, G. R. (2009). *New generation English.* Kathmandu: Vidyarthi Prakashan Pvt. Ltd.

Buchanan, J. (1992). Guidelines for writing multiple choice vocabulary items. *Guidelines* 14.2. Pp. 54-60.

Ellis, R. (1995). Modified oral input and the acquisition of word meanings. *Applied Linguistics* 16 (4). Pp. 409-41.

Fan, M. Y. (2001). An investigation into the vocabulary needs of university students in Hong Kong. *Asian Journal of English Language Teaching* 11. Pp. 69-85.

Heaton, J. B. (1975). *Writing English language tests: A practical guide for teachers of English as a second or foreign language.* London: Longman.

Lado, R. (1961). *Language testing.* London: Longman.

Laufer, B. (1998). The development of passive and active vocabulary in a second language: Same or different? *Applied Linguistics* 19 (2). Pp. 255-71.

Laufer, B.; Elder, C.; Hill, K.; and Congdon, P. (2004). "Size and strength: Do we need both to measure vocabulary knowledge?" *Language Testing* 21 (2). Pp. 202-26.

Li, X. (2004). *An analysis of Chinese EFL learners' beliefs about the role of rote learning in vocabulary learning strategies.* A thesis submitted for the degree of Doctor of Philosophy. University of Sunderland.

Luitel, B. (2005). *Innovating a diagnostic curriculum for English vocabulary learning: Reading plus cooperative task-based approach (Empirical study with special reference to secondary level schooling in Nepal).* An unpublished Ph D Thesis. Hyderabad (India): University of Hyderabad.

Luitel, B. (2012). *Vocabulary in the English language pedagogy: Input, process and product.* A faculty research report submitted to University Grants Commission/Nepal. Sanothimi, Bhaktapur.

Luitel, B. (2015a). Towards translingual definition-resource cloze for second language vocabulary exercise designing. J*ournal of TESON* 2.1. Pp. 5-14.

Luitel, B. (2015b). Efficacy of TDRC versus puzzle task for word-meaning learning in English: A quasi-experiment. *Education and Development* 26. Pp. 134-56.

Madsen, H. S. (1983). *Techniques in Testing.* New York: Oxford University Press.

Malagarriga, E. (1998). Definition-resource cloze. In *New ways of classroom assessment,* ed. by Brown James Dean. New Ways in TESOL Series II: Innovative Classroom Techniques. Alexandria (USA): TESOL.

Meara, P. and Buxton, B. (1987). An alternative to multiple choice vocabulary tests. *Language Testing* 4 (2). Pp. 142-54.

Meara, P. and Fitzpatrick, T. (2000). Lex 30: An improved method of assessing productive vocabulary in an L2. *System* 28 (1). Pp. 19-30.

Meara, P.; Lightbown, P. M.; and Halter, R. H. (1994). The effect of cognates on the applicability of YES/NO vocabulary tests. *The Canadian Modern Language Review* 50 (2). Pp. 296-311.

Nation, I. S. P. (1990). *Teaching and learning vocabulary.* New York: Newbury House Publishers.

Nation, I. S. P. (2001). *Learning vocabulary in another language.* Cambridge: Cambridge University Press.

Pulido, D. and Hambrick, D. Z. (2008). The virtuous circle: Modeling individual differences in L2 reading and vocabulary development. Reading in a *Foreign Language* 20 (2). Pp. 164-190.

Read, J. (2000). *Assessing vocabulary.* Cambridge: Cambridge University Press.

Read, J. (2004). Research in teaching vocabulary. *Annual Review of Applied Linguistics* 24. Pp. 146-61.

Shu, H.; Anderson, R. C.; and Zhang, H. (1995). Incidental learning of word meanings while reading: A Chinese and American cross-cultural study. *Reading Research Quarterly* 30 (1). Pp. 76-95.

Ur, P. (1996). *A course in language teaching: Practice and theory.* Cambridge: Cambridge University Press.

Waring, R. and Takaki, M. (2003). At what rate do learners learn and retain new vocabulary from reading a graded reader? *Reading in a Foreign Language* 15 (2). Pp. 130-63.

Webb, S. (2008). The effects of context on incidental vocabulary learning. *Reading in a Foreign Language* 20 (2). Pp. 232-245.

[1]The authors' initial (and middle) names are abbreviated here as per the style followed by this journal. Their (full) names as appeared in the respective sources will be given in: http://www.binodluitelblog.wordpress.com. Also read *Salute to author identity in academic writing and research* (by Binod Luitel, published in Review Nepal on 15 October 2015 – see: http://reviewnepal.com/articles/salute-to-author-identity-in-academic-writing-and-research.html) to understand the rationale for respecting the identity of authors while referring to their works in academic writings .

A Review of the Results of National Assessments of Student Achievement in Nepal

Dr Lekha Nath Poudel

Joint Secretary

Education Review Office, Government of Nepal, Ministry of Education

lnp_001@hotmail.com

Abstract

This article presents a review of the process and results of National Assessments of Student Achievement conducted by Education Review Office in 2011 at Grade 8, 2012 at Grade 3 and 5, 2013 at Grade 8, and 2015 at Grade 3 and 5. The review is based on the reports of these four assessments published by Education Review Office. Although National Assessment of Student Achievement has been considered as one of the valid and reliable sources of student achievement, there is a need for proper institutional arrangement as well as capacity development for further improvement in assessment quality. The article also indicates that there is a need for responding appropriately the concerns about the systematic reforms towards improvement in the quality of and equity in education based on the assessment results. In the meantime, recently initiated criteria and standard referenced assessment could provide more accurate interpretation of student achievement if managed appropriately.

Key terms: National assessment, Large-scale assessment, Education Review Office, Item Response Theory (IRT), Classical Test Theory (CTT)

Introduction

Assessment is an integral part of education system. One of the purposes of educational assessment is to provide feedback to the government as well as other relevant persons and agencies for improvement in education system or any part of the system. Similarly, another important purpose of assessment is to judge the quality of educational products and outcomes. The first purpose has developmental interest, and we usually call it as formative assessment whereas the latter one has evaluative purpose, which is generally called summative assessment. Students' assessment is one of the most important assessments in education, which has both formative and summative purposes. Formative assessment of students works towards the improvement of pedagogical practices whereas summative assessment aims at providing students with grades and certifying their qualifications. Three type of student assessments − classroom based assessment, public examinations and national assessment of student achievement − have been practiced in

most of the education systems across the globe with three different purposes − assessment 'as', assessment 'for' and assessment 'of ' learning (Clarke, 2012). The main objective of earlier two assessments is to improve students learning and the objective of latter one is to assess the level of learning or the achievement of students.

Globally large-scale assessments of student achievement have been practiced at national as well as regional and international level. Several countries have been participating in one or more than one international assessments such as PISA (Programme for International Student Assessment), PIRLS (Progress in International Reading Literacy Study), TIMSS (Trends in International Mathematics and Science Study). Besides, there are a number of regional assessment programs such as PILNA (Pacific Islands Literacy and Numeracy Assessment), SACMEQ (Southern and Eastern African Consortium for Monitoring Educational Quality), LLECE (The Latin American Laboratory for Assessment of the Quality of Education**),** PASEC (The Analysis Programme of the CONFEMEN Education Systems) available to the countries of the region. Nepal started large-scale national assessment of student achievement of school students in 1990s, mostly commissioning the work to a consulting company. Those assessments were carried out using Classical Test Theory (CTT) with norm referenced analysis. Although assessments were conducted and results were published, institutional arrangement and capacity development regarding large-scale assessment were absent there. Realization of the need for an assessment agency, Government of Nepal, in 2010, established an agency – Education Review Office (ERO) – responsible for undertaking educational assessments including student assessment, performance audit of schools as well as other educational institutions, and educational research, studies and evaluation studies.

ERO has been conducting National Assessment of Student Achievement (NASA) at various Grades since its establishment. Until 2015, ERO conducted two rounds of assessment for the students of Grades 3, 5 and 8 in various subjects. Assessments were conducted using modern test theory - Item Response Theory (IRT); and reports were prepared mostly using norm referenced analysis.

This article is an attempt of a brief review and comparative analysis of these two rounds of assessment results. The reviews and comparisons presented in this article are based on the four published reports of the National Assessment of Student Achievement conducted by ERO in 2011, 2012, 2013 and 2015. ERO conducted NASA at Grade 8 in 2011 for Mathematics, Nepali and Social Studies subjects, and in 2013 for Mathematics, Nepali and Science subjects. In the years 2012 and 2015, assessments were conducted at Grades 3 and 5 in which Nepali and Mathematics subjects were repeated in both the Grades and in both the years whereas at Grade 5 one additional subject (English language) was also added in both the years.

This article, after describing the objectives of National Assessment of Student Achievement in general and the objectives of NASA conducted by ERO, reviews the methods and process used

to conduct NASA. The next section presents a summary of key results from the four assessments and then identifies major implications of the results. Before mapping for future practice of national assessment of student achievement in Nepal, it explores some concerns and challenges regarding National Assessment of Student Achievement in Nepal. Finally, it presents some concluding remarks.

Objectives of the National Assessment of Student Achievement

Among the various types of assessment discussed in introductory section, large-scale national assessment of student achievement aims at providing feedback to the government and other stakeholders for improvement in education system by improving education policy and programs (Poudel, 2016). National assessment provides necessary information on students' achievements – which helps in planning and execution of decisions. Greaney & Kellaghan (2008) rightly mention that national assessment of student achievement mainly answers the following questions related to student achievement:

- How well are students learning?
- Is there evidence of particular strengths and weaknesses in students' knowledge and skills?
- Do certain sub-groups of students perform poorly?
- What factors are associated with student achievement?
- Do the achievements of students change over time?

The responses provided to the above questions through assessment are generally based on the standard defined by the national curriculum. Therefore, we develop and carry out national assessment to assess the overall performance level against the curricular objectives and standards. Policy makers and practitioners may utilize the information about the overall achievement level and factors contributing learning as the feedback towards reform in policy and practices (Poudel, 2016). Murphy et al. (1996) also justified the need for national assessment of student achievements by mentioning, "well-designed national assessments can help countries make informed decisions about interventions to improve educational quality" (p. 2). In this way, national assessment of student achievement facilitates to improve the accountability of education system towards students' learning so that it helps improving the quality of and equity in education.

The overall objective of National Assessment of Student Achievement is to identify students' level of learning in respective subjects and Grades. We may elaborate the objective of National Assessment of Student Achievement conducted by ERO as follows:

- Identify the students' achievement levels in respective Grades and subjects;
- Analyse the differences and inequalities of the achievement of students among various geographical regions, districts, development region, language groups and gender;
- Identify the factors influencing students' learning;
- Analyse the changes in students' achievements in certain duration.

(See, ERO, 2013, 2015, 2015a, 2016)

Assessment Methods and Process

All assessments conducted by ERO from 2011 to 2015 were based on representative sample of schools and students. Item Response Theory (IRT) was used to calibrate the items and to equate the score. Analysis of results was basically norm referenced rather than criterion and standard referenced. This section briefly describes sampling method, tools development process, process and method of administering assessment, and the process used in analysing the results of those assessments.

Selection of schools and students (Sampling)

The population for each of these assessments were Nepali schools running respective Grades and students studying at respective Grades. For example, for assessment in 2011 at Grade 8, all the Nepali schools running Grade 8 and all the students studying at Grade 8 were the population. Sample of schools were taken using Stratified Random Sampling method. The strata for sampling were ecological zones (Mountain, Hill, Tarai and Kathmandu valley), development regions (eastern, mid, western, mid-western and far-western), school location (rural and urban) and school type (community/public and institutional/private).

Sample selection began with the selection of districts in each of the strata from ecological zones and development regions. First of all, 75 districts were categorised based on ecological zones and development regions and classified into three groups covering all ecological zones and development regions. While grouping the districts into three groups, three districts from the Kathmandu valley were repeated in each year. Assessment was conducted in 2011, 2012, 2013 in the districts of a group starting from group1 in 2011. Districts from group 1, group 2 and group 3 were selected for assessment in 2011, 2012 and 2013 respectively; then the districts from group 1 were repeated in 2015. Three groups of districts were as follows:

Table 1: Selection of sample districts

Group	Districts
Group 1	Ilam, Jhapa, Morang, Sankhuwasabha, Bojpur, Ramechhap, Sindhupalchok, Lalitpur, Bhaktapur, Kathmandu, Dhading, Rautahat, Bara, Gorakha, Tanahu, Syanja, Mustang, Nawalparasi, Pyuthan, Banke, Jajarkot, Kalikot, Bhajhang, Doti, Kanchanpur
Group 2	Dhankuta, Khotang, Saptari, Solukhumbu, Udayapur, Bhaktapur, Chitwan, Dolakha, Kathmandu, Lalitpur, Mahottari, Parsa, Makawanpur, Sindhuli, Baglung, Kapilbastu, Kaski, Manag, Myagdi, Bardiya, Jumla, Humla, Rolpa, Salyan, Achham, Kailali, Baitadi, Darchula
Group 3	Taplejung, Panchthar, Terthum, Sunsari, Siraha, Dhanusha, Okhaldhunga, Sarlahi, Kavre, Nuwakot, Rasuwa, Kathmandu, Lalitpur, Bhaktapur, Rupandehi, Palpa, Arghakhanchi, Gulmi, Parbat, Lamjung, Dang, Rukum, Dolpa, Surkhet, Dailekha, Mugu, Bajura, Dadeldhura

Source: ERO (2013, 2015, 2015a, 2016)

After selecting districts, schools in each district were selected covering both rural and urban as well as community (public) and institutional (private) schools. However, while selecting schools, very small schools having less than 10 students at related Grades were replaced with other schools having 10 or more than 10 students in the respective Grades. In each category of public/private and rural/urban schools, schools were selected randomly. Regarding the selection of students, except for the assessments of 2011 all the students from the respective Grades of each of the sample schools were included in the sample. In 2011, a fixed number of students were sampled in each selected schools using random sampling method. However, some of the students with disability were excluded from the sample due to the nature of assessment tools. Similarly, schools running below Grade 5 were not included in the sample even for the Grade 3 assessment – as the samples of both the grades 3 and 5 were taken from the same school. Due to the devastating earthquake, the assessment of 2015 at Grades 3 and 5 was administered 5 weeks later than that of planned date – when the students of Grades 3 and 5 were already enrolled in Grades 4 and 6 respectively. Therefore, as an exception, the sample students of Grades 3 and 5 were taken from the students enrolled in Grades 4 and 6 respectively, and the sample schools were selected from the schools running at least Grade 6. The following table presents the samples for each assessment.

Table 2: Distribution of Sample for the assessment

Year	Grades	Sample size			Remarks
		No. of districts	No. of schools	No. of students	
2011	8	25 (Group 1 as above)	1201	48,682	
2012	3 and 5	28 (Group 2 as above)	1690	80,232	Both the grades 3 and 5 from the same school
2013	8	28 (Group 3 as above)	1199	44,067	
2015	3 and 5	23 (Group 1 as above)	1543	73,878	Both the grades 3 and 5 from the same school

Source: ERO (2013, 2015, 2015a, 2016)

Tool development

As the assessment tools, each of the assessments conducted by ERO used test items in each assessed subject as well as background questionnaires to the students, teachers and head teachers.

Test items in each subject

Item preparation started with the analysis of the curriculum of each subject. Before writing the items, ERO organized training on item writing for subject teachers. A team of item writers was selected from the participants of the training together with some subject experts and university teachers. The team drafted sufficient number of items based on curricular objectives and competencies. Items written by the team of item writers were screened, edited and revised through workshops of subject experts. Finally, subject committee selected at least six sets of items for pre-testing. Pre-testing of items was carried out among the students of the same Grade where at least 1800 students participated in each subject. Items were analysed using pre-test results by calculating item difficulty and item total correlation, and items with appropriate difficulties were selected for final test. Items were of objective as well as subjective types.

In order to compare results with some international assessment results, some items were calibrated using Item Response Theory (IRT) modelling from the sets of released items of PIRLS (For Grades 3 and 5 in language) and TIMSS (For Grade 8 in Mathematics and Science). Besides, some items from previous assessments were also included in each set of items. Finally, three sets of items were prepared choosing some items as the linking items so that three sets of items could be made comparable. The major principles adopted while selecting the items were as follows:

- To ensure content validity, items were selected from the curriculum and tried to cover contents

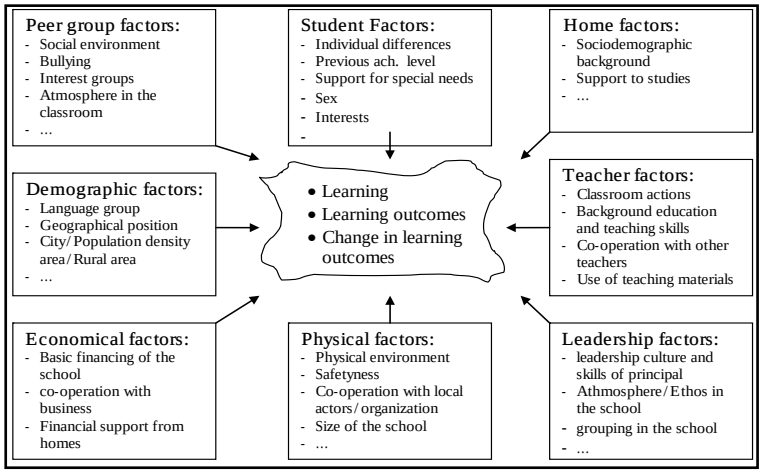
as much as possible.

- Using adopted version of Bloom's taxonomy, items were selected and classified into four categories: knowledge, comprehension, application and higher ability to represent various cognitive levels.
- Reliability of test was ensured by selecting items with proper difficulty level and discrimination power.
- To compare results, items were calibrated and adopted from relevant international assessments as well as previous assessments.

Background questionnaires

Background questionnaires for students, teachers and head teachers were prepared to identify the factors influencing students' achievement. Background questionnaires included contextual information related to students, family, school and classroom practices. It includes peer group factors, student factors, home factors, demographic factors, teacher factors, economic factors, physical factors, and leadership factors (see, ERO, 2013, 2015, 2015a, 2016; Metsamuuronen, 2009). The following figure presents the overall conceptual framework for background questionnaires used in National Assessment of Student Assessment conducted by ERO from 2011 to 2015.

Figure 1 Conceptual framework for the background information



(Source: ERO, 2011; adapted and contextualized from Metsämuuronen, 2009)

ERO organized several workshops of experts, teachers and head teachers to draft the background questionnaires. While developing background questionnaires, contextual questionnaires from some of the international assessments such as TIMSS and PISA were also studied, and relevant questionnaires were adopted with some revision and contextualisation. A team of experts worked for reviewing and editing the drafted questionnaires. The questionnaires were piloted and revised accordingly before using them.

Questionnaires related to students' attitude towards subjects were also prepared using adopted version of Fennema-Sherman attitude scale (Fennema and Sherman, 1976; Kadijevich, 2006), and included in the set of background information questionnaires. Similarly, indicators related to socio-economic status including father's and mother's education, father's and mother's occupation, home possession, home accessories, attending private school were also included in the background information questionnaires.

Test administration, marking and data entry

Tests were administered in all sample schools on the same day and time in which background questionnaires were also filled by the students who participated in the test, subject teachers and headteacher of those schools where the test was conducted. Answer sheets and filled questionnaires were collected in the marking centre. Marking schemes were prepared and scorers were trained before scoring. Marks of individual item of each student and information from background information questionnaires were recorded in an Optical Mark Recognition (OMR) sheet. Similarly, the information from the background information questionnaires of teachers and head teachers were also recorded in separate OMR sheets. Subject experts and ERO technical personnel regularly monitored the whole process of marking and OMR writing. After reviewing each OMR sheet, data were tabulated by scanning each OMR sheet using OMR. Data were reviewed and cleaned before preparing final data for the analysis.

Analysis of results

A framework for data analysis was developed in the year 2011 before starting the analysis in which the same framework with some revisions was used to analyse the data in subsequent years. However, along with adopting basic framework for analysis from 2011, some modifications were made and additional analysis was carried out in subsequent years. Data were analysed using various statistical tools and concepts. The software used to analyse data were SPSS and OPML. Results were analysed using some descriptive statistics such as frequency, percentage, mean, standard deviation, Chi-square, correlation as well as some inferential statistics such as General Linear Model (GLM) including t-test, Analysis of Variance (ANOVA), Analysis of Covariance (ANCOVA), regression analysis, Multivariate methods including factor analysis. Statistical significance was identified by calculating p-value and Alpha (α) level, and effect size was calculated using Cohen's d and Cohen's f.

The analysis framework includes mainly three categories of analysis in each subject: basic results, results based on diversity factors and selected explanatory factors for assessment results. As the basic assessment results, analysis was carried out with overall distribution of scores, achievement in various content areas, achievement in various cognitive domains, item types and achievement, and comparison of achievements with previous database as well as with international standards such as TIMSS or PIRLS as appropriate. Similarly, assessment results were analysed with various categories of diversity factors including district, ecological zone, development region, school type, school location, home language, ethnicity/caste, gender. Assessment results were analysed also to explain the influence of various factors related to student, family and school. The factors included in the analysis were socio-economic status (SES) including education and occupation of parents, home accessories and possessions, type of school attended, working beyond school hours, attitude towards subject, student's age, support received for study, availability of textbooks, homework, and positive (motivating teachers' behaviour and school environment) and negative (e.g. bullying) activities in the school.

Reporting and dissemination of results

ERO prepared two types of reports from each of the assessments: the detailed report and summary report. The detailed report describes technical details including the methods and process followed, analysis of results, major findings and implications. The summary report is a public report written in Nepali language used for wider dissemination. The summary report presents a summary of methods and process as well as major results, findings and implications without adding detailed technical descriptions.

The results of each of the assessments were shared to the department heads of central level agencies and joint-secretaries of the Ministry of Education in the presence of Secretary, Ministry of Education, and then prepared reports were submitted formally to Education Policy Committee (EPC) chaired by the Minister for Education. Besides, report of each assessment was shared to the Annual Review Meeting for school Sector Reform Program (SSRP) organized by the Ministry of Education. In the next stage of dissemination, ERO invited media persons and shared the assessment results in which medias reported the results through their newspapers, magazines, radio/FM, televisions as well as in electronic form. Both the reports from each assessment were printed and distributed to concerned agencies, whereas the summary reports were sent to all the Resource Centres (RCs). During the year ERO organized several dissemination and discussion sessions in various locations in the country. Besides, ERO also included sessions related to the dissemination of NASA results in various orientation programs related to national assessment of students' achievement and performance audit of educational institutions organized in various districts.

Summary of Key Results

This section presents a summary of key results of the assessments conducted by ERO in 2011, 2012, 2013 and 2015 at each of the Grades 3, 5 and 8 in two times. It presents general trends as well as noticeable changes or differences between the results of two assessments of the same grades. Although assessments of different years were made comparable using IRT modelling by calibrating the items and equating the scores, percentage of score may not be comparable perfectly. However, it shows general trends of the result. In IRT based analysis, there is a practice of calculating ability score (latent ability-theta value) to compare the score of different assessments (see, Baker, 1992, 2001). The NASA reports in some cases calculated theta value but in most of the cases the analysis was done using percentage of scores. The analysis of NASA results in this article also uses the same data of NASA reports, mostly the percentage of scores. The following are the summary of key results of NASA conducted by ERO in two times at Grades 3, 5 and 8 from 2011 to 2015.

- ***Student population are concentrated more on low performing area.*** Except for the assessment conducted in 2015 at grade 3 in Mathematics subject, students' population in all subjects and years are not normally distributed. Instead, the distributions are concentrated towards low performing areas. It also shows a wide difference in achievements between the high and low performing students.

- ***Students performed poorly in higher level of cognitive domain***. Students are better in lower level of cognitive domain in comparison to higher levels of cognitive domain, that is, students' achievement score decreases as the level of cognitive domain increases from lower to higher level. For example, students are good in knowledge level in comparison to comprehension, application and higher level. It indicates that teaching learning is concentrated more on memorising the information and facts rather than developing understanding and applying to solve the daily problem.

- ***Achievement varies within the various content areas of the same subject and Grade.*** There are wide variations in achievement scores of various content areas of the same subject. For example, in Nepali language subject, variations in achievement are seen among reading, writing, grammar and vocabulary. Similarly, in Mathematics subject there are variations in students' achievement score among the various contents domains such as algebra, arithmetic, geometry, sets and statistics.

- ***There are differences in student achievement among the districts, ecological zones and development regions.*** In each assessment, wide differences in students' achievement scores are found among districts, ecological zones and development regions. For example, district variation in students' achievement score in assessment conducted in 2013 at Grade 8 in Mathematics shows up to 47 percent points difference between high (59 percent) and low

(12 percent) performing districts; similarly, there is 26 percent points difference between the high (60 percent) and low (34 percent) performing districts in 2016 at Grade 5 in Nepali. While comparing ecological zones, Kathmandu valley has always got highest achievement scores and the difference between the scores of students from Kathmandu Valley and other ecological zones are very wide. When we exclude the Kathmandu valley, achievements of Mid-western and Far-western regions are low in comparison to Western, Central and Eastern regions. Except for the assessment conducted in 2011 at Grade 8 in Mathematics and in 2015 at Grade 3 in Mathematics, western region got the highest achievement score among the five development regions.

- ***There are wide differences in achievement score between community (public) and institutional (private) school students***. Wide gaps in achievement between public and private school students are found in each assessment and each subject. Public schools are concentrated more towards low achieving areas whereas private schools are more concentrated towards high achieving areas. However, there are some cases of high achieving public schools and low achieving private schools.

- ***Remarkable gap in students' achievement is found between urban and rural schools.*** Every year, there is a wide difference in students' achievement between urban and rural schools as the achievement of the students from urban schools is higher than that of rural schools. This difference is found in each subject and Grade.

- ***There are differences in achievement associated with home language and caste/ethnicity.*** When we compare the achievement scores of the students having their home language Nepali and non-Nepali, students having Nepali as their home language have achieved higher. However, if we see the individual language, there are several cases of higher achievement among the students having non-Nepali as their home language. Comparing the achievement of various caste/ethnic groups of students, achievements of Brahman/Kshetri and Janajati students are in the first and second highest position respectively. Except in Nepali language subject, Dalits' achievement is the lowest, whereas in most of the cases Madhesis are poor in Nepali language subject and relatively better in Mathematics.

- ***Gender gap in students' achievement has narrowed down noticeably.*** Expect for a small difference in the achievement of girls and boys in Mathematics, Science and Social Studies at Grade 8, and in Mathematics at Grade 5, gender parity has almost been achieved. Girls' achievement in Nepali subject is higher than the achievement of boys.

- ***The socio-economic status (SES) of the family and students' achievement are directly correlated.*** In each assessment, socio-economic condition of family was assessed through seven indicators: father's and mother's education, father's and mother's occupation, home accessories, home possessions and the type of school attended. It is found that there is a

positive relationship between SES and student achievement in each assessment, that is, higher SES gives higher level of student achievement. As SES is the aggregate of the effects of all the seven indicators, we may identify the effect of each indicator. For example, parents' education and students' learning achievement have positive correlation. Moreover, mothers' education status and students' learning achievement have very strong positive correlation. Similarly, availability of home possessions and accessories have positive impact on students' achievement.

- ***Over age of the student has some negative effects on achievement.*** In terms of their age**,** the students in 8-10 years at Grade 3, 10-12 years at Grade 5, and 12-14 years at Grade 8 have achieved better compared to other age groups of students. Result shows that when the age of students exceeds the right age (as mentioned above) their achievement decreases. Similarly, students younger than right age have also achieved lower than those who are in right age. Here, the right age of students is assumed 5- 6 years at Grade 1.

- ***There are remarkable effects of household work and paid work on students' achievement.*** Participation in simple household work with parents and siblings up to 2 hours a day does not give any negative effects on students' achievement; rather, it has shown some positive effects on achievement. However, engagement in household work more than 2 hours daily and participation in paid work of any duration has given negative effects on students' achievement.

- ***Availability of textbooks has positive effects on students' achievement***. Those students who did not get textbooks have got poor achievement score than those who received textbooks. Data shows that 3 to 5 percent students in the sample did not get textbook even in the end of the session.

- ***Regular home works with proper feedback helps for better achievement.*** Along with providing regular home works, regular feedbacks after checking home works is necessary for better achievement of students. However, the amount of time to be spent in doing home works should be appropriate according to the age of the students as excessive amount of homework may have negative effects on students learning.

- ***Availability of support for study at students' home help improve student learning.*** Support from siblings and parents are important to improve students' learning. Students who got some support from their siblings or parents have higher achievement score than those who did not get any support.

- ***Students' attitudes towards the school, teacher and subject have some effects on their achievement.*** Positive attitude of students towards school, teacher and subject helps getting better achievement. Students, in most cases, develop their positive or negative attitude towards schools and teachers from the school's environment and teachers' behaviour. Pleasant

and motivating environment and activities of school, and motivating behaviour of teachers help students to develop positive attitude towards school, teacher as well as the subject. Similarly, participatory and child friendly classroom practice also provides an opportunity of developing positive attitude towards subject as well as teacher and school.

- ***School bullying hinders students learning.*** A large number of students (40-50%) have been experienced bullying at schools. Some students have been experiencing severe bullying (4 or 5 out of 5 types of bullying). There are wide differences in students' achievements between the students who experienced bullying and those who did not experience bullying. Particularly, the difference in achievement is very high between severely bullied and non-bullied students, as severely bullied students achieved very poor scores.

Implications of the Results

The above summary of key results shows that the national assessments conducted by ERO have provided a wide range of information on students' learning as well as factors influencing the achievement of students. In this context, this section summaries the key and common implications of the four assessments conducted by ERO in 2011, 2012, 2013 and 2015 – which were conducted two times for each of the Grades 3, 5 and 8. Key results indicate that, in order to improve student achievement, there is a need for designing and implementing appropriate strategies by the government, schools and other relevant stakeholders in the following areas:

- Increasing learning achievement and reducing the persisting inequalities across districts, ecological zones, development regions, rural and urban locations,
- Promoting higher cognitive ability and improving reading ability,
- Balancing instructional activities to all content domains and areas,
- Managing various diversities including linguistic, cultural, socio-economic, gender through developing inclusive and child friendly school and classroom environment,
- Reorganizing and revising teacher preparation and teacher development strategies,
- Making school accountable for student learning,
- Learning from high performing schools and reducing the difference in achievement between public and private schools,
- Enrolling students of proper age and designing some alternative strategies for over aged students,

- Ensuring parents' and community participation in school, particularly in students' learning, and designing and carrying out parental education,
- Ensuring that children of school (Basic) age should not be engaged more than 2 hours a day in the household works,
- Eliminating child labour and ensuring that school (Basic) age children are not involved in any paid job,
- Designing some extra supports to those students whose parents and other family members are unable to support their study at home,
- Ensuring the access of a set of textbooks for each student on the first day of new session, and making sure that they have access to some additional reading materials,
- Eliminating the incidence of bullying at school,
- Improving assessment system focusing on particularly classroom based assessment for formative purpose.

Some Concerns and Challenges

ERO completed two rounds of assessments at grades 3, 5 and 8 and prepared reports of each of the four assessments. The summary of methods and process used to conduct those assessments, key results and major implications have been presented in this article. As "a 'national assessment' is a survey of schools and students that is designed to provide evidence about the levels of student achievement in identified curriculum areas for a whole education system or for a clearly defined part of an education system" (Postlethwaite and Kellaghan, 2008, p1), these four assessments carried out in each of the Grades 3, 5 and 8 two times have been successfully conducted considering the basic features of a national assessment. These assessments were conducted focusing mainly on describing and evaluating the quality of student learning outcomes that have been produced by schools (ibid). Schools, parents and other educational stakeholders have also realized the need for national assessment for the purpose of assessing educational outcomes by means of standardized test. Moreover, concerned agencies and persons have been taking assessment results as one of the valid and reliable information for planning, programming, and reviewing educational programmes and outcomes. However, the following concerns and challenges are still visible in the practice of national assessment in Nepal:

- There is a concern about the systemic reforms based on assessment results and finding as there is a realization that sincere and systematic efforts of reforms based on the assessment finding is lacking, although Department of Education has prepared an action plan to initiate reforms based on NASA results and findings. Therefore, the challenge is to establish a mechanism for

continued improvement in policy and practice towards enhancing the quality of and equity in education based on the findings of NASA.

- There is a challenge on institutional arrangement and national and institutional capacity development for assessment. Concern has also been related to restructuring ERO along with the structure of educational administration as per the constitution.
- The need for developing criteria and standard referenced assessment is realized, instead of the existing norm referenced assessment for proper identification of students' various level.
- Developing a clearly defined national assessment framework together with criteria and standards is another strong need.
- A challenge is realized regarding interpretation and comparison of the results of assessment using latent ability score rather than raw score.
- Preparation for participation in international assessments such as PISA and TIMSS is another challenge.

NASA: An Indication for Future Practices

The quality of a national assessment system mostly depends on three quality drivers, which are: Enabling context for assessment, system alignment, and assessment quality (Clarke, 2012). Although NASAs conducted by ERO have maintained reasonable level of assessment quality, weak enabling context and lack of proper alignment in system have been impeding the development of quality system of national assessment (see, Poudel, 2016). Therefore, there is a need for proper institutional arrangement as well as establishment of a mechanism for continued improvement in policy and practice based on the findings of national assessment of student achievement.

In order to define and identify proper level of student achievement, Education Review Office has developed assessment frameworks for Grade 8 in Mathematics, Science and Nepali (ERO, 2016a) and for Grade 5 in Nepali and Mathematics (ERO, 2017) in which the first round of national assessment based on the newly developed framework with criteria and standards has just been administered at Grade 8. The report of this assessment will be completed by October 2017. The assessment following the framework based on criteria and standards is being prepared for Grade 5 in Nepali and Mathematics and the same will be developed for Grade 10 in different subjects. For Grade 3, ERO is planning for an assessment covering Early Grade Reading and Early Grade Mathematics. The School Sector Development Plan, 2016-2022, the seven year plan of the government of Nepal for school education, has also suggested to conduct national assessment at Grades 5, 8 and 10 in a certain interval and Early Grade Reading and Mathematics assessment in grade 3 (MOE, 2016).

Regarding the participation in international assessment, ERO conducted a study entitled 'Feasibility study of Nepal's participation in international assessment' (Khanal et al, 2016). This study suggests strengthening national assessment system at this moment rather than participating in the international assessments like PISA and TIMSS. Moreover, this study also suggests that after preparing the technical and human resource required for participation in international assessment, we may first participate in PISA for Development and prePIRLS.

Concluding Remarks

Since the establishment, ERO has been conducting national assessment of student achievement in various Grades and subjects. This article presented a review of the process and key results of the national assessments of student achievement conducted in 2011 at Grade 8, in 2012 at Grades 3 and 5, in 2013 at Grade 8 and in 2015 at Grades 3 and 5. This article has been developed by reviewing the reports of these assessments conducted in four different years. All the assessments were conducted to assess the students' achievements against the approved curriculum of the government of Nepal. Assessments were conducted using Item Response Theory, but analyses were mostly norm referenced.

Assessment results show that there are various disparities in students' achievement across districts, development regions, socio-economic status of parents, language, caste/ethnicity, rural-urban location, public-private categories of schools. A number of influencing factors on students' achievement and their effects have been identified. Major identified influential factors on students' achievement are age of the students, household and paid work, availability of textbooks, working beyond school hours, home works with proper feedback, availability of support for study at home, students' attitude towards the school, teacher and subjects, and school bullying. Results indicate that most of the differences in achievement as well as the factors influencing students' achievements have been repeated though levels of variations and their effects may vary.

Results of the National Assessment of Student Achievement have been considered as one of the valid and reliable information for planning, programming and reviewing educational programmes and outcomes. In addition to policy implications of national assessment, it works as a monitoring and validating mechanism for outcomes, and helps raising public awareness on the students' achievement and factors influencing the achievement. However, there is a concern regarding systemic reforms based on the assessment results and findings. Another concern and challenge regarding national assessment is on institutional arrangement and capacity development. In the meantime, ERO has begun developing criteria and standard referenced assessment instead of norm referenced assessment which can help to interpret student achievement more accurately if the criteria and standards are applied properly.

Reference

Baker, F. B. (1992). *Item response theory: Parameter estimation techniques.* New York: Marcel Dekker.

Baker, F. B. (2001). *The basics of item response theory*. ERIC Clearinghouse on Assessment and Evaluation.

Clarke, M. (2012). *What matters most for student assessment systems? A framework paper.* Washington DC: The World Bank.

ERO. (2013). *Report of national assessment of student achievement 2011, grade 8*. Sanothimi: Education Review Office.

ERO. (2015). *Report of national assessment of student achievement 2013, grade 8.* Sanothimi: Education Review Office.

ERO. (2015a). *Reports of national assessment of student achievement 2012, grades 3 and 5.* Sanothimi: Education Review Office.

ERO. (2016). *Report of national assessment of student achievement 2015, grade 3 and 5.* Sanothimi: Education Review Office.

ERO. (2016a) *Assessment framework: National assessment of student achievement for Grade 8.* Sanothimi: Education Review Office (ERO).

ERO. (2017). *Assessment framework: National assessment of student achievement for Grade 5.* Sanothimi: Education Review Office (ERO).

Fennama, E., & Sherman, J. A. (1976). Fennema-Sherman mathematics attitude scales: Instruments designed to measure attitudes toward the learning of mathematics by females and males. *Journal for Research in Mathematics Education, 7(5), 324-326.*

Greaney, V. and Kellaghan, T. (eds.) (2008). *Assessing national achievement levels in education.* Washington DC: The World Bank.

Kadijevich, Dj. (2006). Developing trustworthy TIMSS background measures: A case study on mathematics attitude. *The Teaching of Mathematics, 9, 2, 41-51.*

Khanal, P. et al. (2016). F*easibility study of Nepal's participation in international assessment.* Sanothimi: ERO.

MOE. (2016). *School Sector Development Plan, 2016-2022.* Kathmandu: MOE.

Murphy, P., Greaney, V., Marlaine, L., & Rojas, C. (1996). 'Introduction', in P. Murphy, V. Greaney, L. Marlaine & C. Rojas, (eds.) *National assessments: Testing the system, 1-8.* Washington DC: The World Bank.

Postlethwaite, T. N.; Kellaghan, T. (2008). *National assessments of educational achievement.* Paris: The International Institute for Educational Planning (IIEP), UNESCO.

Poudel, L. N. (2016). Reviewing the Practice of National Assessment of Student Achievement in Nepal. *Nepalese Journal of Educational Assessment, 1(1), pp. 1-16.*

Beyond the National Assessment of Students' Learning: Feasibility of Nepal's Participation in International Testing

Dr. Peshal Khanal

Associate Professor

Central Department of Education, Tribhuvan University

peshal.khanal@tucded.edu.np

Abstract

This article reports on a study that assesses the readiness and feasibility of Nepal's participation in international assessments. The study draws upon the analysis of existing resources, technical preparedness and institutional capacity to conduct large-scale testing of student learning. Using a multi-method approach, this study provides a comparative overview of the contents and domains of international tests such as the Program for International Student Assessment (PISA), Trends in International Mathematics and Science Study (TIMSS), Progress in International Reading and Language Study (PIRLS), and of Nepali school curricula. As this study reveals, there is a significant overlap of the content and domain of learning between international assessments and school curricula of Nepal. However, adaptation and alignment of the content of national curricula in line with the international testing requirements and redesigning the existing school testing are needed to prepare the Nepali students for international testing. In addition, institutional, technical and human resource preparedness and pedagogical reform are the key prerequisites for Nepal's involvement in international assessments.

Key terms: National Assessment of Student Achievement (NASA), PISA, TIMSS, PIRLS, National Curriculum Framework (NCF), Education Review Office (ERO)

Introduction

Several countries in the globe have been participating in the international or regional testing along with their national assessments. In Nepal, ERO has been conducting National Assessment of Student Achievement in a certain time interval at various Grades of school education. In the meantime, there are some discussions in Nepal on the possibilities of participation in international testings. This article is an attempt of assessing the feasibility of Nepal's participation in international testings like TIMSS, PISA, PIRLS.

This article provides an overview of international assessments, particularly of Program for

International Student Assessment (PISA), Trends in International Mathematics and Science Study (TIMSS) and Progress in International Reading and Language Study (PIRLS), and assesses the feasibility of methods and processes used in these assessments in view of Nepal's participation in these assessments. For this purpose, it analyses the comparability of contents between the national curricula of Nepal and major international assessments – PISA, TIMSS and PIRLS. In addition, it compares the cost of National Assessment of Students' Achievement (NASA) with the cost for participation in international assessments and assesses the appropriateness of Nepal's participation in international assessments considering the possible implication in budget. Assessing both benefits and challenges of administering international standardized tests in developing country contexts, it draws some implications for Nepal's preparation and participation in international assessments.

Methodology

This study employs a multi-method approach, using a wide range of techniques including document and data review, financial analysis, interviews, web-based research, and policy study. Specifically, this study draws on the information and data obtained from the following four methodological components.

Review study

A review study was based on the available policies, secondary data, published report and other relevant literature. This study reviewed international assessment practices, looking into the frameworks, programs and practices of the international tests such as PISA, TIMSS and PIRLS. In addition, the review study also looked into the national assessment policies and programs, focusing particularly on various assessments carried out by ERO in the past.

Subject-wise comparative study

A deskwork was done by subject experts to compare the contents and domains to be tested in PIRLS, TIMSS with the Nepali curriculum of grades 5 and 8 in Mathematics, Language (Nepali and English) and Science. Similarly, comparison was made on the contents and domains to be tested in PISA with the Nepali curriculum of grade 10 in Mathematics, Language (Nepali and English) and Science. The subject experts thoroughly looked into the curriculum and contents, and specified the contents that are overlapped and compatible as well as the contents that are unique and different, not compatible for international testing.

Feasibility study

A detailed study was undertaken to assess the feasibility of the methods and process used

in international assessment in the context of Nepal, particularly based on existing capacity on conducting students' assessment. This was done first by reviewing the test requirements for different international studies, and then assessing institutional and human resource capabilities of ERO and other relevant national institutions to carry out international tests in Nepali context. For this purpose, a survey-cum-interview study was carried out to assess institutional capacity of ERO and other relevant institutions in Nepal.

Cost-analysis

A financial study was undertaken comparing the cost of NASA with the cost needed to participate in international assessments and assessing the appropriateness of participation in an international assessment considering the possible implications in budget. For this purpose, financial estimation was made to participate separately in different international tests (e.g. PISA, TIMSS) by analyzing the financial policy and requirements of the individual assessment. For the budgetary requirements and funding options for administering international testing in Nepal, financial records and audit reports of the previous NASA study were reviewed along with the interviews with the concerned authority and staff at ERO.

Personal communication and interview

Personal communication was made with the international assessment agencies to collect assessment materials and information about procedures that are not publicly available (for example, budget and other internal policy documents). Open ended interviews were carried out with three professional experts of testing and measurement – two from Tribhuvan University and one from Kathmandu University. These interviews were conducted to collect experts' opinion on the relevance and feasibility of Nepal's participation in international assessments.

Overview of National and International Assessments

Assessment is a tool for determining the quality of educational outcomes. The objectives of student assessment are three-fold (Poudel, 2016, p. 1): to provide feedback for classroom teaching and improve student learning, to certify the grade and qualification of students, and to monitor and evaluate the quality of education system. Accordingly, assessment is broadly classified into three types – classroom-based assessment, public examination and national assessment. Classroom assessment has relatively a narrow scope that focuses more on formative aspect of students' learning in a particular school or classroom. The fundamental goal of this assessment is to assess learning and provide feedback to teachers and students for improving their learning. Public examination, on the other hand, is administered usually by external authority (such as National Examination Board) to certify the grade and qualification of students. Both classroom assessment and national examination are considered mandatory for students, requiring all to take part in these tests.

National assessment is administered over a fixed period of time (e.g. every two or three years) in order to determine the status of educational achievement of students in identified curriculum areas (e.g. reading or literacy, mathematics, science). Assessed mostly among a sample of students, this kind of assessment plays a critical role in determining the performance of the education in the country and value for money in the education sector.

In a national assessment, a) achievement is assessed using standardized instruments, administration and scoring procedures; b) assessment instruments are administered to an agreed population of students or, more commonly, to a probability sample of students who are selected to be representative of the population; c) individual student achievements are aggregated to the system level. Reliable data may be also obtained for sub-populations if samples are sufficiently large (e.g., students categorized by the state/province in which they attend school; students attending private schools and students attending public schools); d) background information, provided by participating students, teachers, and sometimes parents, is usually collected in questionnaires to provide insights into relationships between achievement and a variety of factors (e.g., school and classroom resources and practices, student characteristics, family characteristics) [Department for International Development (DFID), 2010].

With the establishment of ERO in 2010, Nepal laid foundation for conducting national assessment and applying the assessment result for the diagnosis of systematic problems of schooling and thereby taking necessary remedial actions. Along with the system development for national assessment, ERO successfully complemented NASA 2011, NASA 2012, NASA 2013 and NASA 2015 (see ERO, 2013, 2015a, 2015b, 2016). The assessment and analysis framework used in NASA assessments are designed and verified by national and international experts. Example of computer-based IRT modeling used in NASA test analysis suggests that national assessment has been technically sound, authentic and reliable.

National assessment differs from the kind of assessment found in regular program of school and colleges. These assessments are on the level of knowledge, skill or understanding of individual students in the classroom as a form of formative or summative decision about the students' learning in a particular curricular domain. Classroom assessment does not meet the following characteristics of national assessment (DFID, 2010): a) standardized instruments and procedures; b) administration to an entire student population or representative sample thereof; c) system or sub-system level aggregation of results; and d) systematic relation of performance data to background characteristics of students. A national assessment is also different from public examinations in the sense that the latter play a crucial role in certifying student achievement, selecting students for further study and standardizing what is taught and learned in schools (Greaney and Kellaghan, 2008). This is a kind of 'high stake' testing which has a high consequence for students' future career.

Along with the regular cycle of national assessment, there has also been a growing

opportunity for Nepal to participate in standardized international assessments like Program for International Student Assessment (PISA) (see http://www.oecd.org/pisa/), Trends in International Mathematics and Science Study (TIMSS) and Progress in International Reading and Language Study (PIRLS) (see http://www.iea.nl/, and https://timssandpirls.bc.edu/). PISA is run by Organization for Economic Cooperation and Development (OECD) which was first conducted in 2000 and repeated in every three years. This survey assesses learners aged 15 who are nearing the end of secondary education. It assesses performance in reading, mathematics, science and problem solving. The last round of PISA test was undertaken in 2015; and its result will be published in December 2016. Managed by the International Association for the Evaluation of Educational Achievement (IEA), TIMSS was first conducted by IEA in 1995 and now repeated in every 4 years and tests learners of 4th and 8th graders (10 and 14 years old). In 2001, the IEA offered another international survey, which was called PIRLS. This is repeated every five years, and it focuses on 10 year old learners' abilities in reading and literacy. The last round of this survey was conducted in 2011.

PISA for Development and pre-PIRLS are new testing opportunity designed particularly for the developing countries. *PISA for Development is a new initiative of OCEC* to 'increase participation of developing countries' in PISA assessments. This is done by developing contextual questionnaires and data-collection instruments that better capture diverse situations in emerging and developing countries, adjusting the PISA test instruments so that they are sensitive to a wider range of performance levels, and establishing methods and approaches to include out-of-school students in the PISA assessment. Pre-PIRLS is another preparatory option for the developing countries designed by IEA to participate in PIRLS which reflects the same conception of reading as PIRLS, except that it is less difficult. Depending upon a country's educational development, pre-PIRLS can be given at the fourth, fifth, or sixth grade. The following table provides a summary of the three international assessments – PISA, TIMSS and PIRLS.

Table 1: A comparative overview of major international assessments – PISA, TIMSS and PIRLS

	PISA	TIMSS	PIRLS
Full Name	Program for International Student Assessment	Trends in International Mathematics & Science Study	Progress in International Reading Literacy Study
Assesses	Reading, mathematics, science, problem solving	Mathematics and science	Reading
Age	15	10 and 14	10

	PISA	TIMSS	PIRLS
Grade	Grade 9 (UK Year 10)	Grade 4 and Grade 8 (UK Year 5 and 9)	Grade 4 (UK Year 5)
Last assessment	2015	2015	2011
Next assessment	2018	2018	2016
When	Autumn	March-June	March-June
Purpose	Evaluates education systems by assessing to what extent students at the end of compulsory education can apply their knowledge to real-life situations and be equipped for society	Measures trends in maths and science achievement Describes educational context, including home support, students' attitudes, curriculum, teachers' training, classroom activities	Measures trends in reading comprehension Investigates the experiences young children have at home and school in learning to read
Focus	Skills-based	Curriculum-based	Curriculum-based
Type of test	Criterion-referenced	Criterion-referenced	Criterion-referenced
Achievement levels reported	Reading 1-5 levels, Mathematics 1-6 levels, Science 1-6 levels	Low, intermediate, high, advanced	Low, intermediate, high, advanced
Supplementary information	Background information obtained from learners in a questionnaire. Focuses on characteristics of learners, attitudes to subjects, motivation and learning strategies	Background information obtained from learners in a questionnaire. Information also collected about teachers, activities of schools and teachers' classroom behaviour	Background information obtained from learners in a questionnaire. Information also collected about teachers, activities of schools and teachers' classroom behaviour
Organization	Organisation for Economic Cooperation and Development (OECD)	International Association for the Evaluation of Educational Achievement (IEA)	International Association for the Evaluation of Educational Achievement (IEA)

	PISA	TIMSS	PIRLS
Countries	79 countries and economies in 2015	59 countries in 2015	55 countries and 7 sub-national entities in 2011
Test length	120 minutes, plus 15 minute background questionnaire	72 minutes at Grade 4, 90 minutes at grade 8 plus 15 minute background questionnaire	80 minutes, plus 15 minute background questionnaire
No. Learners assessed	More than 5,000 learners in each country/Jurisdiction	At least 4,000 learners in each country/ jurisdiction	About 3,500-4000 learners in each country/jurisdiction
Development process	Developed by international experts and PISA Consortium test developers. Test items reviewed by country representatives for cultural bias and relevance to PISA's goals	TIMSS Science & Math Item Review Committee and National Research Coordinators from participating countries develop frameworks through iterative process	PIRLS Reading Development Group and National Research Coordinators from participating countries update frameworks for each PIRLS administration and reviews test items for cultural bias
Variants	A new PISA-based test for schools was developed for 2014. It provided results for schools but not aggregated at national level - OECD ran a new survey of adult skills (age 16-65); results released in 2013. PISA for Development is another initiative of the OECD that aims to identify how PISA can be support evidence-based policy making in emerging and developing economies.	TIMSS Advanced in maths and physics for age 18 – Grade 12 (UK Year 13) or 1st year of university	Depending on country's educational development, can be taken later than age 10 Pre-PIRLS version: less difficult e-PIRLS based on computer-based literacy

Source: Cambridge (2015)

Contents of International Assessments and National Curricula: A Comparative Overview

There is a significant overlap of content and domain of learning between the framework of international assessment and national curricula. However, there is much to be done in order to make the contents and domain of learning congruent. Two of such tasks are - adaptation and alignment of the content of national curricula in international testing context and redesigning existing examination and testing system of schools to prepare students for internationally designed testing. The summary of content-specific comparative analysis is presented in the following sub-headings.

Mathematics Grade V and VIII - TIMSS and National Curriculum Framework (NCF*)*

Both TIMSS and NCF (Grade V and VIII) (CDC, 2007, 2062BS, 2071BS) have similar contents in the areas of number system, geometry, and measures and data display. There is around 90% overlap of contents in all areas of mathematics. Therefore, there is little problem regarding the gap or content mismatch while adopting TIMSS test items for Nepalese students' assessment at the end of grade V. However, the Nepalese curriculum practices, including instruction and assessment, need to be reviewed and revised to meet the requirement of international testings. The areas of adaptation are: context/problem generation, mathematization of the context/problem, calculation, and verification with the context to adopt TIMSS level assessment.

Mathematics Grade X and PISA

Both PISA and NCF have aimed to develop knowledge and skills in the content areas such as change and relationships, space and shape, quantity, uncertainty and data. There is approximately 90% similar content in PISA and NCF. Therefore, the content mismatch between PISA and NCF is not a great issue. However, since the PISA assessment is based on context, process, and content, the Nepalese curriculum practices, including instruction and assessment, need to be adapted and revised in relation to context/problem generation, mathematization of the context/problem, calculation, and verification with the context while adopting PISA assessment items.

Science Grade V - TIMSS and NCF

There is a high overlap of the contents between TIMSS and NCF Grade V Science. Regarding the content weightage, living beings, matter, energy; environment, and Earth and universe have 35%, 32% and 17% weightage in NCF. However, the equivalent is 45% (life science), 35% (physical science) and 20% (Earth Science) in TIMSS. The additional content in Grade V in NCF

is information technology (8%) and simple technologies (8%). The content breakdown overlaps from 70% to 80%. Regarding the area and domain of learning, knowledge, understanding and higher abilities carry 20%, 30% and 50% weightage respectively in NCF whereas the content breakdown in TIMMS is 40% (knowing), 40% (applying) and 20% (reasoning).

Science Grade VIII - TIMSS and NCF

The overlap of contents of grade 8 Science in both TIMSS and NCF is high. Physics, chemistry, biology and earth science have 25%, 20%, 35% and 20% contents respectively in TIMSS though the same contents have 25%, 15%, 15%, 20% weightage respectively in Nepali school curriculum. The school curriculum has additional 25% weightage allocated for practical activities. Approximately 70% to 80% contents overlap between NCF and TIMSS. In terms of learning areas or domain, knowledge, understanding and higher abilities have 20%, 30% and 50% weightage respectively in NCF where the corresponding domain covers 35% for knowing, 35% for applying and 30% for reasoning in TIMSS. The NCF has additional emphasis on 'higher abilities' domain.

Science PISA and Grade X

While PISA is not a curriculum-based test, a significant overlap of the content and domain of learning has been observed between PISA and NCF. Secondary school science curriculum of Grade 10 consists of 30% physics, 15% chemistry, 22.5% biology and 7.5% astronomy and geology portion, whereas PISA has 36%, 36% and 28% weightage allocated for physical, living and earth and space sectors. PISA has less emphasis on chemistry. While the NCF has subject-based focus on knowledge, understanding, application, and higher abilities of science learning, the PISA has a focus on three competency areas: 1) explain phenomenon scientifically, 2) evaluate and design scientific enquiry, and 3) interpret data and evidence scientifically. This suggests that NCF requires preparing students for higher order competency such as explaining, evaluating, designing and interpreting the scientific phenomena and design.

PISA Reading and NCF

PISA reading is focused on understanding, using, reflecting and engaging with written texts in order to develop students' knowledge and potential and participate in society. While reading is core part of the PISA, the Grade 10 language curricula (both English and Nepali) focus on all the four language skills – listening, speaking, reading and writing. The PISA literary assessments select items from different situations (e.g. personal 30%, educational 25%, occupational 15% and public 30%), the Grade 10 language curricula (both English and Nepali) focus more on the selected educational contents, and little on personal and public situation. While most of the reading exercises in Grade 10 language curriculum are based on continuous texts formed by sentence

organization into paragraphs, PISA reading assessment has 60% items from continuous text, 30% from non-continuous (list, tables, graphs, diagrams, advertisements, schedules, catalogues, indexes and forms), mixed (5%) and multiple (5%) text format. Text in PISA includes description, narration, exposition, argumentation and instruction while Grade 10 language curriculum largely includes description and narration. Rather, Grade 10 curriculum focuses reading exercises on different genres such as poetry, prose, drama, letter writing, fiction, autobiography and story. While the main mode of delivery for the previous PISA assessment was paper-and-pencil, PISA 2015 has also an option for computer-based delivery. This option does not exist in NCF Grade 10 language curriculum. This indicates that NCF requires alignment of curricula in line with PISA reading in order to prepare Nepali students for PISA testing.

PIRLS Reading and NCF

Reading literacy, according to PIRLS 2016, is the ability to understand and use those written language forms required by society and/or valued by the individual. While reading to learn is a key learning domain of PIRLS, Grade V curriculum (both English and Nepali) aims to develop all four skills of language – listening, speaking, reading and writing. In PIRLS reading has two particular purposes: Literary experience (50%) and acquire and use information (50%). Similarly, reading comprehension has focused on retrieving explicitly stated information (20%), make straightforward inferences (30%), interpret and integrate ideas and information (30%) and evaluate and critique. On the other hand, NCF (Nepali and English) focuses more on functional use of reading, read different kinds of texts (stories and facts) both for enjoyment and to extract specific information, interpret charts, tables, diagrams, develop reading skills (skimming and more detailed reading). The curriculum for the development of reading skills is associated with grammar skills and includes exercises such as completing, matching, ordering, choosing, and composing sentences. The reading skills will be developed in tandem with the writing and other learning activities. In this sense, Nepali curricula are in greater alignment with PIRLS reading test. Nevertheless, the competency of English as a second or foreign language is a serious concern for making decision about Nepal's participation in PIRLS English reading test.

Feasibility of Participation in International Testing

A country's entry into the international assessments such as TIMSSS, PISA and PIRLS requires a number of prerequisites met at country level. Primarily, there should be adequate national capacity, including infrastructure, techonology and human resource, for implementing the large-scale standarized tests. In addition, there should be clear policy and legal provisions to facilitate the administratation of the international testing at country level. Such provisions should clearly spell out the reform strategies that are adopted after analysing the test results of such large-scale testing. Another fundamental requirement is the financial investment to take part in international testing, which is generally high as compared to the cost of national assessment. The following sections further analyze each of these considerations.

Requirement for undertaking TIMSS, PISA and PIRLS and national capacity

All major international assessments – PISA, TIMSS and PIRLS – require the participating country to assign a national institution (national center for TIMSS and PIRLS, for example) to undertake this assessment under the broader framework and guidelines provided by the test organizations (e.g. OECD for PISA and IEA for TIMSS and PIRLS). The head of this institution – National Research Coordinator (NRC for TIMSS and PIRLS) or National Project Manager (NPM for PISA) plays a major role in international assessment projects and works as the main contact person in the country. NRC has following responsibility in order to conduct TIMSS and PIRLS:

- Organization and management of the national center to ensure that all tasks are fulfilled
- Employment and supervision of the staff
- Assurance of the availability of required hardware, software and other necessary equipment/ materials
- Participation in the international NRC meetings.

In Nepal, ERO could be the national institution to undertake this assessment and the head of ERO could act as NRC or NPM. However, as the ERO head and staff are the government's employees, their independent role as suggested by the international testing organizations could be questionable.

Similar to TIMSS and PERLS, the National Project Managers work with the OECD contractor on all issues related to the implementation of PISA in their country. They play an important role not just in the successful implementation of PISA in accordance with OECD quality standards, but also in the development and review of PISA reports and publications.

There should be other staff in each country to successfully conduct the study, such as Sampling Coordinator, Data Manager, Translators/Translation Reviewers, School Coordinators, Test Administrators, National Quality Control Monitor, Scorers, Staff for data entry and Office staff.

The overall tasks of the national team are to prepare school samples, contact schools and sampling classes for the assessment using particular software (e.g. WinW3S for TIMSS and PIRLS), contributing to preparing the achievement items, translating, reviewing and producing the assessment materials, submitting the survey instruments for international translation verification, administering the assessment with the supervision of international quality control monitor, scoring the assessment, creating the data files entering test information into the software,

submitting the national database and related materials to the international data processing center, preparing national analysis of the data, writing the national report and disseminating the national report in synchronization with the release of the international report. In addition, the national center will need to complete a national context questionnaire.

In order to accomplish these responsibilities, the existing human resources and technical set-up at ERO is inadequate – particularly in accomplishing the technical task of data entry, processing, and reporting at par with the standard of international assessments. Because of the inadequacy of human resources and other technical stuffs, ERO has been undertaking NASAs by hiring national consultant firms; and both parties share the NASA activities and tasks. The following table summarizes the NASA tasks shared by ERO and consultant firms.

Table 2: NASA tasks shared by ERO and consultant firms

S.N.	NASA components	Role of ERO	Role of the outsourced consultants	Remarks
1.	Pre-work: item writing, pre-testing, item selection	√		Worked with experts, school and university teachers and
2.	Questions printing, packing		√	
3.	Test administration in schools	√	√	DEO also supported test administration at school
4.	Scoring answer book		√	
5.	Data entry		√	
6.	Data analysis and delivery of key results	√		
7.	Report writing	√		Some expert service also used
8.	Editing, publishing and disseminating	√		Some expert service also used
9.	Item-bank updating	√		

The above table shows that much of the technical work of NASA has been undertaken by the outsourced consulting firms. The consulting firms, however, have to be equipped with theoretical and technical skills required for using cutting-edge technology of assessment. As suggested by an ERO official, the consulting firms should have "adequate knowledge and skill on three parametric

Item Response Theory, use of IRT models in R and STATA, curriculum and assessment, criteria based test development, data analysis and reporting".

The existing human resource strength at ERO is largely non-technical (see Table 3), experienced more in administrative tasks. None of the ERO staff members is professionally expert in testing and assessment, although some are trained to undertake technical work of large-scale assessment. Notably, only one officer at NASA was found trained and skilled in analyzing test score using IRT modelling during the time of this study (April 2016).

Table 3: ERO's human resource for carrying out large-scale assessment

S.N.	Position	No	Qualification	Assessment experiences
1.	ERO Head	1	PhD	Item development to report writing, curriculum and evaluation, Overall planning and implementation including item development, writing reports, M&E
2.	Under-secretary	1	Masters/MPhil	Overall planning and monitoring, orientation and dissemination, partial support to NASA including writing reports,
3.	Section head, NASA	1	Masters	Planning and coordination of NASA activities, including items and tools development to test administration and analysis and reporting (Currently vacant).
4.	Section officer	4	Masters/MPhil	Test planning to report writing, test development and analyzing, reporting, training & orientation, (only one officer is trained in item analysis including IRT modelling), coordination, field-visit, training and orientation.
5.	Under-secretary	2	Masters	Assigned for another job other than NASA
5.	Section officer	7	Masters	Assigned for another job other than NASA

Comparing the human resources required for international assessments with the existing capacity of ERO, one can conclude that ERO needs institutional autonomy and additional technical staff for participation in international assessments.

Requirements and Preparedness for Participation in International Assessment

For international testing, there is a requirement of strong institutional and human resource foundation for undertaking the assessment. For example, both PISA and TIMSS assessments require a national institution with adequate technological infrastructure, including computers and software, as well as human resource qualified and trained in the areas of large scale national and international testing. There should be assurance of competent testing staff and availability of required hardware and other necessary equipment and materials. While the staff such as sampling coordinator, data manager, translators, test administer, quality control monitor, scorers, staff for data entry and analysis are inadequate at ERO, human resource management seems a key prerequisite for participation in the international testing. The additional requirement is to train all staff in computer-assisted testing, including IRT modeling for test analysis.

This study has also solicited the ideas of ERO officials and assessment experts regarding the policy, legal and financial requirements in order to strengthen the existing capacity of ERO for the sustainable implementation of NASA as well as for Nepal's groundwork to participate in international testing. The following information has been obtained:

Policy requirements

- ERO should be run as an autonomous institution in order to take decision independently, keeping itself away from political and administrative control and direct regulation of the Ministry of Education.

- The policy should make some structural adjustment in existing system of allocating the quota of administrative staff and their recruitment through a transparent procedure. The ERO should also have a policy and adequate budget for hiring and training of technical staff required for international assessments. Rather than the project-based, fixed-term appointment, they should be deputed as permanent staff of ERO so that assessment can be conducted as a regular cycle, informed by regular research and evaluation.

- There should be a clear policy for the capacity development of ERO staff with a view to keeping them abreast with the new and state-of-the art methods and tools of testing and assessment.

Legal requirements

- Under the Ministry of Education, ERO has been established as a permanent center for assessing student achievement at national level. And recently, ERO has obtained a legal status through the eighth amendment to the Education Act which could pave the way for fulfilling the legal

requirements for the effective functioning of ERO. Therefore, ERO has been established as a permanent legal institution. In this context, Education Regulations should clearly specify mandates, roles and responsibilities of ERO.

Technical and human resource requirements

- ERO requires the state-of-the-art IT and software facilities required for IRT modeling or other form of modern testing approaches. There are a few trained data analysts, mostly inadequately trained in IRT modeling. For previous NASAs, technical works were undertaken by the outsourced consulting firms and staffs. In this context, ERO, which could be the national center for undertaking international assessments, is essential to be equipped with adequate technological infrastructure and IT system as a par with the international testing standards. There is a need of adequately trained staff at ERO for developing, administrating and analyzing the test.

Financial requirements

Conducting international assessments also require significant financial investment on the part of the participating country. For example, the international overhead cost for new participants in PISA 2018 is EUR 182 000 payable over four years at EUR 45,500 per year from 2016 to 2018 (OECD, 2016). The total amount is equivalent to NRs 2,18,40,000.00 (1 EUR = NRs 120). In addition, costs for the national implementation of the program should be borne entirely by the participating countries. This includes both the staff cost and operating cost of the assessment. There will be additional cost for attending the regular meetings to be held periodically.

Similarly, the basic fee per grade for participation in paper TIMSS or IEA eTIMSS in 2019 is 225,000 ICU (IEA's International currency unit) = US$ 112,500 plus EURO 112,500 [International Association for the Evaluation of Educational Achievement (IEA), 2016]. This is equivalent to NRs 2,56,50,000 for one grade and NRs 5,13,00,000 for two grades (1 US$ = NRs 108 and 1 EUR = NRs 120).

Projecting the total participation cost using different international indicators and study data, Nepal has to allocate the following amount of budget:

Table 4: Financial requirements for Nepal's participation in international assessments

International Assessment	International fees (in NRs)	Total estimated cost for Nepal (in NRs)
PISA 2018	2,18,40,000.00	8,73,60,000.00
TIMSS 2019 (Grade 4 & 8)	5,13,00,000.00	20,52,00,000.00
PIRLS 2016	2,68,00,000.00	10,72,00,000.00

The comparison of the above cost with the Nepal's NASA 2013 expenditures shows that international testing comes with high expenditure of public money. The total cost of NASA 2013 which was administered in Mathematics, Nepali language and Science subjects of Grade 8 is NRs 2,15,42,000.00. Taking the NASA 2013 cost as a benchmark (i.e. NRs 2,15,42,000.00), the above table shows that four-times more budget is needed for Nepal to participate in the PISA 2018, nearly 10 times more for the TIMSS 2019 and nearly five times more for the PIRLS 2016. This suggests that the cost of Nepal's participation in international assessments is too high to justify in terms of the value for public money.

Benefits and Challenges of Participation in International Testing

Participation in international assessments invites both benefits and challenges. The possible **benefits** include:

- For Nepal, participation in international studies enables policy makers and educators to compare the performance of Nepalese students with international students.
- Not only studies of student outcomes are important but, equally, studies on teacher training, teaching techniques and background information of students and schools are important in identifying the factors which help students learn and achieve their potential.
- International studies provide schools and students with an opportunity to experience cutting-edge assessments that are designed by international experts.
- Participation in international assessment also helps to develop technical and human resource capacity to undertake state-of-the-art methods and techniques of assessing student achievement, which in turn contribute to building better and reliable system of national assessment.

The following **challenges** are envisaged if Nepal intends to take part in international assessments:

- Nepal's current need is that educational investments and policies should focus on access, quality and management, school-age children in a significant number are still outside the school and learning opportunity within the schools is questionable. So, Nepal has a challenge to provide the access of schooling to all and meet the basic requirements for quality schooling before taking part in internationally comparable testing.

- All international assessments are designed by developed countries in order to compare the achievement of students among themselves. The decision on test requirements and test objects is made on the basis of the educational context of these countries. So, such testing may not be relevant for developing countries.

- Taking consideration of contextual factors of schooling, it is argued that international assessment has laid its emphasis merely on learning outcomes of students no matter how the context, input and process variables are different. Considering the unequal and diverse societies of Nepal in terms of their geography, location, socio-cultural and demographic characteristics, Nepal may require equity-based assessment more than the standardized and uniform tests.

- International assessments are largely based on the state-of-the-art technology for administering the test and analyzing the scores. The latest version of these assessments requires students to take computer-based tests. Taking decision for joining international assessments without such preparation would be problematic.

Conclusion and Implications

Based on the analysis of existing capacity for undertaking technology-based large-scale international assessments, curricular compatibility, existing data on access and achievement of students and the value for money, this article concludes that Nepal needs to strengthen the national assessment system and developing a strong technological and human resource foundation before taking part in international assessments. For this purpose, the following recommendations are made:

- ERO needs to be developed as an autonomous and independent organization, keeping it away from the direct regulation of the Ministry of Education.

- ERO is required to be equipped with more sophisticated infrastructure and information technology.

- ERO should have adequate permanent positions to recruit qualified and trained staff in different areas of testing. It is also recommended that the relevant representatives of ERO should be provided opportunities to attend various capacity building trainings provided by international testing organizations and contractors.

- The result of the periodic national assessment should be disseminated to the school level, and schools should be encouraged to develop learning improvement plan and take action for increasing achievement level of students.

- While strengthening national assessment system, some comparison with international assessment could be done by calibrating some relevant items from international items and equating the scores using IRT modeling. Although such practices have already begun in previous NASA projects, this should be made more rigorous and comprehensive in future assessments.

Although a greater overlap has been observed between the framework of national curriculum and international testing, a further alignment of curricular contents and learning domains is necessary to participate in international assessments. While the national curricula of Nepal should have a focus on cultural, linguistic and ethnic diversity of its population, standardizing curricular contents with a greater global-local linkage is an additional requirement.

In view of the fact that international assessments require more analytical, critical and judgmental capacity of the students to respond to the different high-ability questions, the domination of teacher-centered mode of delivery in Nepal should be substituted by more participatory, interactive and technology-assisted instruction whereby students take part in construction of knowledge and critical analysis of it.

Both TIMSS and PISA are now developed as computer-based tests, though there is a paper-based option available for underdeveloped countries. This suggests that if Nepal aims to participate in PISA or TIMSS, the paper-based option may not be available in the next assessment cycle. Nepal therefore requires its schools to equip with technological infrastructure including electricity, computers, projectors, printers and required learning software and prepare teachers for ICT-supported instruction and assessment. Considering all these needs and requirements, institutional, technical, and human resource preparedness, curricular alignment and pedagogical reform are the key prerequisites for Nepal's participation in international assessments. With the fulfillment of basic pre-requisites of international testing, Nepal could participate in PISA for Development (PISA-D) and pre-PIRLS which are preparatory tests developed particularly for the developing countries, and depending on the lessons learnt, further decisions can be made to participate in major international testing such as PISA, TIMSS and PIRLS.

Acknowledgements

As the article is based on the study for ERO about Nepal's fesibility of the participation in some international assessments, the author offers sincere thanks to ERO for providing the opportunity to carrying out this study. Sincere thanks also goes to Bed Prasad Dhakal, Kamal Acharya and Ashok Sapkota for the comparative analysis of the Nepali curricula with those of PISA, TIMMS and PARLS.

References

Cambridge. (2015). *International surveys: PISA, TIMSS, PIRLS.* Retrieved from http://www.cie.org.uk/images/271193-international-surveys-pisa-timss-pirls.pdf

CDC. (2007). *National curriculum framework for school education.* Sanothimi: CDC.

CDC. (2062BS). *Primary education curriculum.* Sanothimi: CDC.

CDC. (2071BS). *Secondary education curriculum.* Sanothimi: CDC.

DFID. (2010). *National and international assessment of student achievement* (Guidance note, A DFID paper). Retrieved from https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/67619/nat-int-assess-stdnt-ach.pdf

ERO. (2013). *Report of national assessment of student achievement 2011, grade 8.* Sanothimi: ERO.

ERO. (2015a). *Report of national assessment of student achievement 2012, grade 3 and 5*. Sanothimi: ERO.

ERO. (2015b). *Report of national assessment of student achievement 2013, grade 8.* Sanothimi: ERO.

ERO. (2016). *Report of national assessment of student achievement 2013, grade 3 and 5*. Sanothimi: ERO.

Greaney. V. and Kellaghan, T. (eds.) (2008). *Assessing national achievement levels in education*. Washington, DC: The World Bank.

IEA. (2016). *Meeting the needs of a range of countries*. Retrieved from http://timssandpirls.bc.edu/pirls2011/downloads/prePIRLS.pdf

OECD. (2016). *PISA 2015 assessment and analytical framework: Science, reading, mathematic and financial Literacy.* Paris: OECD Publishing.

Poudel, L.N. (2016). Reviewing the practice of national assessment of student achievement in Nepal, *Nepalese Journal of Educational Assessment,* 1(1): 1 – 16.

Devising a Framework for Assessing Early Childhood Learning and Development in Nepal

Dr. Kishor Shrestha (Professor & ECD Export), Devina Pradhananga (ECD Export)
and Gopal Prasad Bhattarai(Assessment Expert)

Abstract

This paper is an attempt to devise a framework for assessing learning and development of young children in Nepal. After reviewing early learning and development standards (ELDs) available so far at national and international level, expected learning standards covering main domains of learning and development that are considered relevant to assess in the context of Nepal have been presented. Along with them, learning and development tasks expected to be performed by children and tasks of assessment have also been nuanced.

Key terms: assessment framework, developmental task, domain of development, ELDS, ECD, IDELA, learning standard

Introduction

It is well evident that proper learning and development at early years of an individual lays solid foundation for all-round development and well being of a person throughout life in terms of physical, mental, intellectual, social and emotional development. Early childhood development has also shown to have lasting implications for learning, health, and well-being (Walker et al., 2007). Being cognizant of these far reaching implications in human development, individual parents and nation states both, through their public policy, place due emphasis on children's development at the start of school. While the start of school is the entry point for education system, children's development at the preschool age is the manifestation of years of influence, resulted from health, nutrition status, and exposure to stimulation and emotional support from the first days of life. Investment in early child development not only provides greater returns through improved cognitive, social-emotional development, school readiness, health and nutritional status, but also is instrumental in achieving enrollment, retention, high achievement and completion of primary school as well as developing attitude towards inclusive and non–discriminatory values. Such wider benefits of early childhood development program will have significant contribution on achieving the Sustainable Development Goals (SDGs), which has been identified as critical element of reaching proposed education goals, with the proposed target placing emphasis on children's development at the start of school (Raikes, 2016).

An increasing concern is found on measuring early learning and development skills in order to understand to what extent the desired goals have been achieved, whether school

readiness skills have been developed in children and what further intervention would be required. Accurate, reliable assessment of child using Early Learning Development Standard (ELDS) provides the evidence based information to support policy formulation; identify the additional interventions needed for curricula, instruction and teacher training. At the same time, it will be helpful in identifying children's achievement as well as the gaps in achievement against the set standard. Besides, assessment on nationally representative sample population or in census form is important for tracking equity, because without knowing where children begin at the start of formal schooling, it is impossible to know whether education systems are successful in closing the gaps between children as the school years progress (Raikes, 2016). As ELDS is developed considering developmenally appropriate milestones in Nepali context, this assessment will provide a strong evidence on the entry level of individual child for schooling and learning including the child's holistic development status.

To guide the assessment of early learning and development and foreground it with certain conceptual base, developing an assessment framework by spelling out main domains to assess, methods and tools for assessment has increasingly been important. Assessment framework is essential also for devising standards, indicators, measurement tools, and assessment guidance. Assessment frameworks have been able to guide coordinated action in countries, regions, and at the global level, because they provide a platform for stated agreements on what is most critical to measure; clarify the purposes for measurement; spur the creation, adaptation, and adoption of measurement tools that can be used to improve services for children; and finally, encourage regular collection of indicators to track progress. Measurement benefits from efficiencies of scale: By aligning efforts, better and more complete data are obtained, and innovations are shared more rapidly (Raikes et al, 2014). Assessment framework will also serve as a basis to make the data comparable that is generated from one to another assessment. Given the reasons, developing a framework and assessing children's learning and development based on the framework has been a global trend.

As Nepal is planning to initiate assessment of early learning and development of the children at Early Childhood Development Center (ECD), it is imperative to have a national framework for this purpose to standardize the process, generate uniform and comparable data for monitoring the progress over the period of time, harmonize the assessment actions and use the assessment results to improve the quality of ECD program. In this context, this article is an attempt to devise a framework for assessing learning and development of young children in Nepal. Before suggesting a framework, ECD context of Nepal has been presented, early learning and development standards (ELDS) available so far at national and international level have been reviewed and international practices on assessment of early learning and development have also been discussed to seek insight from the different perspectives.

Methodology

This paper is an output of document review and consultations with key personnel involved in ECD planning and program implementation. In order to be familiar with what the children at preschool age are expected to learn and the expected requirements to be available in ECD centers in the context of Nepal, Early Childhood Development Guidelines (Curriculum) 2062, ELDS for Children From 48 to 60 Months Old 2069 and Minimum Standards for ECD center (DOE, 2062, 2069 B.S.) have been reviewed since they spell out the necessary knowledge, skills and competencies of ECD children, and standards for infrastructures, facilities and materials needed for ensuring appropriate learning and development activities to take place in ECD centers. Besides, expected learning standards for preschool age children available mainly in the publications of UNICEF, Save the Children and The World Bank have also been reviewed to identify globally accepted learning standards for this age group. While preparing the draft of this framework, the consultative meeting with the program planner, implementer was conducted to know what preschool age children should learn, what aspects are to be assessed and what type of assessment tools to apply.

ECD Context in Nepal

The academic focused preschool program started since 1980s was shifted towards holistic development of child from 2000 under the EFA framework. Realizing the need and importance of ECD programs, MOE started the ECD program following the new shift. Observing the benifit and various impacts on children's development and learning, ECD centers in later years expanded rapidly throughout the country to improve its access for the children aged 3-4 years old. As a result, their number excluding preprimary classes run by institutional schools has reached to 30448 with 92 percent gross enrolment rate of target age population in the academic year 2073 BS. Of the total enrolment in Grade 1, about 64 percent children enter in this grade with the experience of ECD and PPC (DOE, 2016).

With a view to promote a comprehensive approach to Early Childhood Education and Development (ECED) programs for safeguarding the rights of children, National Policy on Early Childhood Development in 2004 emphasized the full fledge development of children's physical, socio-emotional, cognitive, spiritual and moral potentials. The National Strategic Plan for Early Childhood Development in Nepal (2004) was developed to speed up implementation of the National Policy on ECED, which aims to ensure its alignment with Education for All (EFA) program. Aiming at maintaining standard in service and facilities to be provided in ECD centers, there is also the Operation and Management Guideline developed by the MOE/DOE. To ensure minimum learning and development standards in all ECD centers, ECED curriculum, Early Learning Development Standards (ELDS) and ECED Minimum Standards are in place.

Despite these initiatives regarding the development of ECD program, effort for the

assessment of early learning and development of children in ECD against the standards expected in the aforesaid documents are yet to make. To facilitate the process, a national framework for assessment and tools for carrying out it is seen an urgent need in Nepal. Developing an assessment framework is also essential to measure, track and monitor the progress against indicator 1 under SDG target 4.2[1] (Rebello Britto & Hancioglu, 2016) as each member country has to report where one stands against them. Given the context, an endeavor for devising a framework for assessing early learning and development of children is made after reviewing the early learning and development standards available to date at national and international level. Review of the ELDS will lead to identify the aspects and domains to assess which will provide a basis for devising an assessment framework.

Reviewing Early Learning and Development Standards (ELDS)

Parents, program planners and policy makers have expectations of certain developmental and learning tasks from the children at each stage of their development. ECD program also expects certain level of learning tasks from the children as a result of interventions made for some period of time which are broadly termed as 'early learning and development standards' (ELDS). They can be mentioned as well–defined expectations for child outcomes (National Research Council, 2008) after the intervention of ECD. To state them specifically, ELDS are the statements that describe expectations for the learning and development of young children across the domains of health and physical well–being, social and emotional well-being, approaches to learning, language development and symbol systems, and general knowledge about the world around them (Council of Chief State School Officers and Early Childhood Education Assessment Consortium, 2007). In similar vein UNICEF Romania (2010) mentions that ELDS represent a set of statements that reflect the expectations concerning what children should know and be able to do–which also reflect certain perspective about the child and his/her development. These are defined for supporting child's growth and development from birth to when they start school. They incorporate new theories concerning children's development generated by psychology, brain research, child development and education and the national values to be promoted. These ELDS will be helpful to seek insights on the aspects and domains that are important to assess ECD.

While preparing children to be ready for school by keeping them developmentally on track is the main development purpose of ECD program, global consensus is also found among the ECD experts that readiness for school should be understood more broadly than cognitive skills. Accordingly, ELDS are found to be formulated as involving several developmental areas, including motor, language and early literacy, math and problem solving, socio-emotional development, and

1 Indicator 1: Percentage of children under 5 years of age who are developmentally on track in health, learning and psycho-social well-being. SDG Target 4.2: By 2030 ensure that all girls and boys have access to quality early childhood development, care and pre-primary education so that they are ready for primary education.

approaches to learning. Competence in all these areas will ensure that children are ready to benefit from educational activities offered in the school environment.

Some skills included in ELDS are relevant across domains and encourage learning and development in many areas (National Research Council, 2000). Self-regulation, or the ability to focus attention and behavior, is hypothesized to be relevant across all domains because it is so central to what children learn and experience. Therefore, self-regulation is understood to play an especially central role in school readiness (Eisenberg, Valiente, & Eggum, 2010). Language development also has a strong influence on many domains of development, including mathematics, literacy and social interactions. Social interactions are especially sensitive to expressive language because children with better language skills can communicate better with peers. Expressive language plays a role across multiple domains of development, including social/emotional, literacy and cognitive. Finally, some areas of development, such as executive function, seem to facilitate acquisition of new skills and knowledge, and thus show strong relationships with many domains of learning from early childhood onward (Blair & Razza, 2007). This means that measurement of early childhood development should include all domains of development, rather than focusing on early academic skills alone, and for some areas, effects of early childhood development may become apparent for several years to come.

Considering the importance of the holistic development in early years for early learning and development, ELDS is developed as the nationl standard covering all necessary domains considering the national expectation on what children should know and be able to do. For the assessment of children under ELDS, different organizations have developed the assessment tools. Among them, International Development and Early Learning Assessment (IDELA) developed by Save the Children (Pisani et al, 2015), East Asia-Pacific Early Child Development Scales (EAP-ECDS) developed by the UNICEF (Rebello Britto & Hancioglu, 2016) and by the World Bank (Laat, 2015) are considered important as they cover all the domains, nation specific ELDS have been built around them, and are considered to meet the global standards. Main domains included in the ELDS developed by them have been presented briefly in table below.

Table 1: International ELDS

Domains	ELDS by		
	UNICEF: (Rebello Britto & Hancioglu, 2016)	Save the Children (Pisani et al, 2015	World Bank: (Laat, 2015)
Physical development	• Gross motor development • Fine motor development • Sensory motor development	• Gross motor development • Fine motor development	• Physical growth, nutrition • Fine and motor skills
Health, personal care and hygiene	• Health and nutrition • Personal care and hygiene • Personal security skills	• Personal safety	
Social development	• Interaction with peers • Interaction with others • Accepting and respecting	• Peer relations • Conflict resolution	• Sharing • Peer interactions
Emotional development	• Self concept development • Emotional self control development • Emotional expression development	• Emotional awareness • Empathy • Self awareness	• Perspective taking • Understanding feeling • Self awareness
Approaches to learning	• Curiosity and interest • Initiative • Persistency in activity • Creativity	• Persistence, initiative and engagement	Executive function and self regulation • Emotional control • Focus attention and behavior,working memory

Language and communication	• Receptive communication (listening and understanding) • Expressive communication • (Talking)	Emergent literacy and language • Print awareness • Expressive vocabulary • Letter identification • Emergent writing • Initial sound discrimination • Listening and comprehension	Early literacy skills • Receptive and expressive language • Listening and comprehension • Alphabet knowledge, • Name writing
Cognitive development	• Logical thinking and problem solving • Elementary mathematical knowledge and skills (numbers, numeracy, operations, space concepts, geometrical shapes) • World knowledge and understanding (living world, earth, space, scientific method)	Emergent numeracy • Measurement comparison • Classification/ sorting • Number identification • Shape identification • One to one correspondence • Simple operations • Simple problem solving	Early mathematic skills • Naming numbers, • Compare quantities • Perform basic addition • Shape and spatial awareness

Building upon these ELDS, national standards are found to have been developed relevant to the specific context. For the Nepalese context, Department of Education (DOE) (2062 BS) has also initiated to set standards of learning and development for ECD.

Review of ELDS in Nepalese Context

For the purpose of maintaining standards, improving management and harmonizing developmental support in ECD, DOE has developed standards on early learning and development through Early Childhood Development Guidelines (Curriculum) 2062 and ELDS for Children from 48 to 60 Months old 2069 (DOE, 2062, 2069 BS). Both documents spell out ELDS among ECD children and standards regarding infrastructures, facilities and materials needed for ensuring

appropriate learning and development activities to take place in ECD centers. The ELDS set by domains in the documents for the given purpose are presented hereafter.

Table 2: ELDS in ECD Curriculum of Nepal

Domains	ELDS	Domains	ELDS
Physical development	• Gross motor skills • Fine motor skills • Sensory motor skills • Health hygienic skills • Safety behaviors • Healthy environment awareness • Good food habit	Emergent mathematics and numeracy	• Naming numbers and counting • Measuring and compare length, height, width, weight and volume
Social development	• Interaction with peers and elders • Respect and affection • Participation and Initiation	Creativity	• Forming shape and size • Imitating and role playing • Follow and make rhythm
Emotional development	• Self awareness • Feels pride on own self • Expression of feeling and emotions • Control of emotions • Respond to other feeling • Empathy • Appreciation • Motivation and curiosity to learn	Cultural development	• Knowing and following family/class rules norms and ECD center's routine • Knowing and appreciating local culture, cultural and natural heritage • Understanding and respecting differences • Showing respect to national emblems and symbols.
Cognitive development	• Inquisitive to do and learn • Problem solving • Logical thinking and imagination • Reasoning • Classification/sorting, ordering and sequencing	Language development	• Listening and responding • Identifying/recognizing sounds • Following/giving short instructions • Comprehending

	• Comparison and discrimination • Scientific thinking and reasoning • Recognition, identification and description • Awareness of direction, distance, location and spatial relations • Awareness of time		• Communicating owns thoughts, feelings interests • Asking questions • Demonstrating pre reading • Demonstrating pre writing • Copying letters/words/ patterns

ELDS for Nepalese Context

Based on the domain-wise ELDS as stated in the national documents mentioned above, some ELDS by each of the following representative domains have been identified appropriate for assessment purpose, which are considered comparable with IDELA developed by Save the Children and UNICEF for global use.

- **Physical development:** This domain encompasses children's physical health and ability to engage in daily activities.
- **Social development:** This domain refers to the child's ability to form positive relationships that give meaning to children's experiences at home, school, and larger community.
- **Emotional development:** This domain refers to a child's disposition and addresses the emotional competence, rather than skill, for becoming involved in learning and acquiring knowledge.
- **Cognitive development:** This domain includes children's ability to understand and think about the physical and social world. In particular, this domain focuses on children's knowledge of objects in the world around them, their logic and mathematical knowledge, their knowledge of agreed-upon social conventions such as numbers and colors, and their understanding and appreciation of the arts in their lives.
- **Language development:** This domain encompasses children's understanding and use of language, emerging reading and writing skills, and ability to communicate effectively.
- **Cultural development:** This domain was felt important to be included so as to address the value related to cultural diversity and respect for cultural heritage for the future of Nepal.

Major ELDS under each domain have been presented in the table below:

Table 3: ELDS for Nepal

Domains of Development	Sub Domains/ Aspects	Development/ Learning Standards	Development/Learning Actions/Tasks
Physical development, health and well being	Physical development a) Gross motor skills	• Demonstrate co-ordination of gross motor body parts	• Walk along straight line balancing the whole body • Stretch, bend and touch their feet. • Move some steps backward • Jump some steps using both feet • Hop, turn by turn on one foot • Crawl on the ground • Climb up/down ladder, slopes, steps • Carry/lift a small chair of her/his size • Throw/catch toys, ball, purse, small bags to and from short distance • Swing • Climb up/ down from branch/rope/bar • Child can run for a purpose and stop suddenly with control
Physical development, health and well being	b) Fine motor skills	• Coordinate and use fine motor body-parts • Demonstrate hand /eye co-ordination • Use tool/ instruments objects with control	• Tear paper, leaves in to pieces as instructed • Thread various objects such as beads, corncobs • Do up buttons • Click fingers • Produce a sound by whistling • Color within a boundary using crayons • Do simple weaving • Draw lines, circles, patterns using pencils • Fold paper, cut and paste small objects • Mould and manipulate wet sand, dough or clay
		• Show reactions/ respond to something using senses	• Imitate/identify/differentiate sounds • Follow given rhythm • Show a reaction on object/things to feeling something (smooth/rough/hard/soft/hot/cold/ sharp/blunt)

Domains of Development	Sub Domains/ Aspects	Development/ Learning Standards	Development/Learning Actions/Tasks
			• See and distinguish objects far and near using each eye in turn • Respond to tasting something (sweet/sour, bitter/ hot)
	Health and hygiene a) Personal hygiene	• Demonstrate healthy and hygienic behavior/ practices	• Maintain neat and clean (nose/ hand/nail/feet/ teeth) • Comb up hair • Aware of dirt/clean objects/dress/foods • Follows proper toilet habits/practices
	b) Safe practice	• Have understanding of safe practices • Avoid harmful/ dangerous objects	• Shows awareness of danger of fire/electricity/ sharp objects/broken glasses/poisons • Familiar with signs of danger/poison and avoid them
	c) Food habit	• Have some understanding of healthy/hygienic food habits/ practices	• Follow healthy and hygienic food habits in practice • Avoid bad food habits
Social development	Social development a) Relationship with peers and adults	• Demonstrate proper relationship with familiar/ unfamiliar peers\ adults	• Interact easily with known family members • Show respect to family members • Interact positively with known community members e.g. shopkeeper, teacher • Greet and interacts with unfamiliar adults appropriately • Ask for help from unfamiliar adult when needed
	b) Life skills	• Introduce own selves • Take responsibilities in daily activities	• Tell owns name/address properly • Take out/put thins properly in right place • Shows curiosity to find out new things and keen to learn

Domains of Development	Sub Domains/ Aspects	Development/ Learning Standards	Development/Learning Actions/Tasks
Social development		• Participate in learning and development activities • Demonstrate self dependent in doing individual and in group work • Involve peers in performing given activities	• Listen to other before doing something and do following the instruction give • Take and wait for turn of own • Initiate and lead activities • Perform activities independently and confidently • Involve in group work with peers
	c) Cooperation	• Involve in class/ center's daily activities • Share things feelings/thoughts in groups/with peers	• Help at home bringing things such as books, mug, clothes • Help in class/center to bring and put things back in their places • Take responsibility for owns belongings • Offers own things and thoughts to friends as they needed
	d) Social behavior	• Demonstrate appropriate behavior with peers/adults • Adjust in familiar and unfamiliar environments	• Show sensitivity to the feelings and needs of others • Acknowledge own mistakes • Show respect according to relationship • Mix up own selves in any group • Express comfort in any setting
	e) Self concept	• Demonstrate sense of self • Show pride of own existence	• Introduce oneself mentioning name, gender home address etc. • Express owns likes and dislikes • Takes pride of owns abilities and skills • Takes pride of owns family, goods, home etc

Domains of Development	Sub Domains/ Aspects	Development/ Learning Standards	Development/Learning Actions/Tasks
Emotional development	f) Emotional expression	• Demonstrate different emotions • Respond to the feeling of others and empathies	• Express different emotions according to situations • Demonstrate satisfaction or pride of what s/he does, achieves, has got etc. • Demonstrate proper reaction when provoked, teased and appreciated etc • Appreciate or compliment others • Show empathy to others
	g) Emotional security and confidence	• Demonstrate security in different situations • Demonstrate confidence to new experience and learning	• Demonstrate confidence/security in familiar/ unfamiliar settings • Knows/realizes that s/he is valued/accepted by others when s/he is called by name • Demonstrate proper coping strategies for different situations • Show independence in individual/group work • Express oneself confidently with eye contact
	h) Self control/ balance	• Demonstrate balance response over the emotions	• Show modified expression/behavior of emotions for different situations • Help/support other in trouble • Show situational expressions
Cognitive development	Intellectual development a) Cognitive skills and learning process	• Shows inquisitiveness (exploration/ investigation) • Demonstrate persistence/ concentration • Solve problems and predict outcomes through experience	• Use materials/instruments to explore/discover • Concentrate in self selected tasks • Focus and concentrate on directed tasks • Use different strategies to solve problems • Show/tell logic/reason behind actions • Understand reason/cause of event/situations • Use previous knowledge/skills in day to day situations • Review/reflect of event/activities of day

Domains of Development	Sub Domains/ Aspects	Development/ Learning Standards	Development/Learning Actions/Tasks
		• Reflect on own learning and use for further learning	
	b) Classification and ordering	• Show understanding of shape/size/form • Recognize/ describe different shape • Match/compare objects • Arrange in sequential order and classify	• Arrange shapes in orders from small to large • Describe shapes/objects using terms smaller/ bigger/smallest/biggest/largest correctly • Compare objects by shape and size • Describes geometric shapes: circle/triangle/ square • Form shape together to make other shape • Differentiate and name colour • Copy pattern/shape
Cognitive development	Cognitive knowledge a) Scientific knowledge/ exploration	• Observe and describe weather/ sky • Describe parts of body using correct names • Differentiate living and non-living • Know the habitat of living beings • Identify and describe everyday materials • Different some tools and materials of daily use	• Recognize/describe the nature of the day-sunny, cloudy, rainy, windy, hot cold etc • Identify clothes and thing suitable for the weather • Tell about/draw picture of sun/moon/star • Describe functions of different organs/parts of body: eye, nose, ear, mouth, hands etc • Differentiate/describe basic feature of living/non living beings • Name the habitat of most familiar animal/insects • Identify objects by types of materials: plastic, glass, wood, metal, clay etc • Explain basic source and use of water • Show basic understanding of the concepts like float, melt, freeze, sink etc • Describe minimum functions of local tools (Janto, Dhiki, Kodalo etc), machines (generator, refrigerator, thresher etc) and devices(phone, TV, radio etc)

Domains of Development	Sub Domains/ Aspects	Development/ Learning Standards	Development/Learning Actions/Tasks
Cognitive development		• Describe types of transportation available in the context	• Describe most common means of transportation like rickshaw, bicycle, horse, cart, boat etc
	b) Distance and direction	• Show an understanding of distance/direction • Follow and show distance and direction • Show an understanding on spatial relations	• Show basic directions like up and down, right and left, in front of and behind • Understand the difference between far/near, in between etc • Show the placement of objects like under, above, inside etc
	c) Concept of time	• Describe activities by time • Differentiate different times of day	• Show an understanding on the time of day: morning/day/evening/night/afternoon • Show an understanding on the concept like before/after while describing an event • Describe difference of time like yesterday, today, tomorrow • Know days of week and months of year
	Mathematics a) Numeracy	• Demonstrate knowledge of numbers and counting	• Recognize and trace number 1-9. • Show an understanding on the use of less than/ more than in describing quantity of objects • Count with 1 to 1 corresponding to objects up to 1-10 • Use basic number and counting operations like few, less, more, equal etc • Count from 1 -10 • Recognize and trace the numbers 1-9

Domains of Development	Sub Domains/ Aspects	Development/ Learning Standards	Development/Learning Actions/Tasks
Cognitive development	b) Measurement	• Demonstrate an understanding of volume (amount, quantity), height, weight, length.	• Describe length and size of object • Use the terms like longer, short, taller, equal while describing objects • Measure steps up to 10, walking across a room • Measure sand, soil, water using containers/pots
	Creativity a) Creative arts	• Use of materials to create and draw	• Create pictures using different materials • Make paper crafts. • Make model of different objects and patterns from clay, dough, blocks and using colours
	b) Imagination, drama and role play	• Use owns imagination in drama and role play • Differentiate imaginary situation and imitation	• Play the role of the characters they heard in story • Dramatize his/her own story creatively • Explain whether the story is real or fantasy
	C) Music and movement	• Create music using different instruments, objects and voice	• Use simple instrument eg stick, shaker, Chhin Baja to make rhythm. • Use voice and body parts to make music and rhythmic movement
Cultural development	Values a) Family and community	• Know and follow the values and norms of their family and community • Follow and value their own family, community and ECD center routine	• Introduce himself or herself using name, gender, locality and nationality • Involve in their family events such as birthday, marriage etc. • Follow the daily routine of ECD center showing respect for local and national events and festivals.

Domains of Development	Sub Domains/ Aspects	Development/ Learning Standards	Development/Learning Actions/Tasks
Cultural development	b) Natural and cultural heritage	• Show knowledge and appreciation of local rules and culture. • Demonstrate the respect for natural and cultural heritage	• Know about specific aspect of festival celebrated by her/his family community, e.g. special food • Talk about different ethnic, cultural dress worn in her/his family and community • Name three heritages of the locality. i.e. River, Pond, Mountain, Cave • Say basic do's and don'ts in the heritage
	Nation and nationality a) Unity and diversity	• Understand and respect people's different needs and culture • Understand children born from Nepali parents are Nepali • Show respect to Nepali emblems	• Mixes up and play with children from different culture, ethnicity, community with respect • Identify and appreciate different ways of life in different region Mountain, Hill and Tarai of Nepal • Identify national flag, bird, animal etc • Sing national anthem and song
Language development	Communication a) Listening	• Listen to and respond properly	• Listen to and distinguish common sounds(human, animals, birds) • Listen to and follow instructions in familiar language • Listen to and respond in familiar language • Listen to story for short time
	b) Communicating with others	• Speak simple short sentences • Communicate with each others (familiar peers and adults)	• Speak fluently in familiar language using appropriate word with respect to geography, caste/ethnicity, culture, community • Participate in conversation in small groups • Ask and respond the questions • Express views and opinions • Address and speak with elders and younger children according to local norms and values

Domains of Development	Sub Domains/ Aspects	Development/ Learning Standards	Development/Learning Actions/Tasks
			• Wait for their turn, listen to others and respond properly
	Literacy a) Pre reading	• Demonstrate pre reading skills	• Interpret the picture and pictorial story • Match letters and words with pictures. • Recognise alphabets in Nepali and mother tongue
	b) Pre writing	• Demonstrate pre writing skills	• Draw pictures • Draw and copy patterns • Draw, make a line, half circle, full circle, pattern • Begin to form alphabets

Review of IDELA

Considering the importance of International Development and Early Learning Assessment (IDELA), the framework and tools developed by Save the Children (SC) have been widely reviewed and consulted while formulating ELDS, assessment framework and tools for assessing the ELDS.This framework is considered useful for some reasons. First, it is a holistic, rigorous, open source instrument that is feasible and easily adapted to different national and cultural contexts. Second, it was developed and used in the under-developed and developing countries of Asia and Africa for the program evaluation and monitoring receiving development support from SC. Third, IDELA has widely been used by international and national partners in a number of urban settings as well in middle and high income countries, including Eastern and Central Europe, Australia, Egypt, and the Philippines, with positive and promising results (Pisani, Borisova & Jo Dowd, 2015). On top of all, it covers key areas of development representing key early learning and development competencies that most often appear in national ECCD curricula and standards. IDELA is a skill-oriented assessment, tapping into the degree of mastery of specific skills and as such items were developed in way that allows us to see progress over time in all skill areas. Scoring is continuous in most instances (rather than a yes/no response) and a number of items contain integrated stop rules that allow for questions to be answered meaningfully by children of varying abilities and ages. The following are core domains and skills formulated for IDELA for assessment purpose.

Table 4: IDELA Framework

Gross and Fine Motor Development	Emergent Literacy and Language	Emergent Numeracy	Socio emotional Development	Approaches to Learning
• Hopping on one foot • Copying a shape • Drawing figure • Folding paper	• Print awareness • Expressive vocabulary • Letter identification • Emergent writing • Initial sound discrimination • Listening • Comprehension	• Measurement and comparison • Classification/ sorting • Number and shape identification • One to one correspondence • Simple operations • Simple problem solving	• Peer relations • Emotional awareness • Empathy • Conflict resolution	• Persistence • Motivation • Engagement

Part of the IDELA includes items for the assessment of core ELDS of varying skills and supplementary items covering the domains like executive function and health and hygiene knowledge. Tools have been designed for direct assessment of children mostly through interview

with children and observation of them. Embarking on it, an assessment framework for measuring ELDS as formulate above has been suggested hereafter.

A Framework for Assessing the ELDS in Nepal

Considering the importance of the domains and ELDS under them as presented above that contribute to children's long-term development, learning and well being; they have been included as key domains to measure. These areas of development represent key early learning and development competencies that most often appear in national ECCD curricula and standards. As elaborated in the table above, domains are the broad areas to be measured. Within each domain, there are sub-domains, aspects and particular skills to be measured that are presented as learning standards and tasks.

For the purpose of assessing the ELDS presented above, the following framework has been presented along with tasks and activities. Based on this framework, assessment tools and materials have to be developed.

Table 5: Assessment Framework for ELDS

Domains	Subdomains	Aspects	Standards	Tasks to perform	Tasks of Assessment for children
Physical development, health and well being	Physical development	Gross motor skills	• Demonstrate coordination of body parts	• Walk along straight line balancing the whole body • Stretch, bend and touch their feet. • Move some steps backward • Jump some steps using both feet • Hop, turn by turn on one foot • Crawl on the ground • Climb up/down ladder, slopes, steps • Carry/lift a small chair of her/his size • Throw/catch toys, ball, purse, small bags to and from short distance • Swing • Climb up/ down from branch/rope/bar	• Walking along 10 steps straight line balancing body • Touching feet with hands by standing • Moving some steps back • Hopping on right and left foot. • Lifting a plastic chair • Throwing the ball towards wall or instructed direction. • Running, and stopping when instructed
		Fine motor skills	• Coordinate and use fine motor body-parts • Demonstrate hand /eye co-ordination • Use tool/ instruments objects with control •	• Tear paper, leaves in to pieces as instructed • Thread various objects such as beads, corncobs • Do up buttons • Click fingers • Produce a sound by whistling • Color within a boundary using crayons • Do simple weaving • Draw lines, circles, patterns using pencils • Fold paper, cut and paste small objects • Mould and manipulate wet sand, dough or clay	• Tearing the given piece of paper into pieces of equal sizes. • Threading the pieces of flower to make a garland. • Clicking finger as instructed • Coloring the given rectangular/ /circular shape in paper • Drawing a straight line and circle on the paper with pencil • Making the paper some folds of equal sizes • Pasting the given pictures in the piece of paper

Domains	Subdomains	Aspects	Standards	Tasks to perform	Tasks of Assessment for children
Physical development, health and well being		Sensory motor skills	• Show reactions/ respond to something using senses	• Imitate/identify/ differentiate sounds • Follow given rhythm • Show a reaction on object/things to feeling something (smooth/rough/ hard/soft/hot/cold/sharp/ blunt) • See and distinguish objects far and near using each eye in turn • Respond to tasting something (sweet/sour, bitter/ hot)	• Identifying what the given sound is. • Producing the same sound as heard. • Distinguishing blunt and sharp by touching ends/ points of two pencils. • Identifying the distance of given objects which is far and which one is near.
	Health and hygiene	Personal hygiene	• Demonstrate healthy and hygienic behavior/ practices	• Maintain neat and clean (nose/ hand/nail/feet/teeth) • Comb up hair • Aware of dirt/clean objects/dress/foods • Follows proper toilet habits/practices	• Identifying clean/ dirty things/clothes/objects • Knowing the place to put dirty objects/wastage/ garbage.
		Safe practice	• Have understanding of safe practices • Avoid harmful/ dangerous objects	• Shows awareness of danger of fire/electricity/ sharp objects/broken glasses/poisons • Familiar with signs of danger/poison and avoid them	• Recognizing the sign of danger/poison when shown
		Food habit	• Have some understanding of healthy/hygienic food habits/ practices	• Follow healthy and hygienic food habits in practice • Avoid bad food habits	• Knowing/ identifying names of foods good for health
Language development	Communication	Listening	• Listen to and respond properly	• Listen to and distinguish common sounds (human, animals, birds) • Listen to and follow instructions in familiar language	• Carrying out activities following short simple instructions. • Repeating/copying heard words/short sentence/line of songs

Domains	Subdomains	Aspects	Standards	Tasks to perform	Tasks of Assessment for children
Language development				• Listen to and respond in familiar language • Listen to story for short time	• Identifying names of characters from the story/ events heard
		Communicating with others	• Speak simple short sentences • Communicate with each other (familiar peers and adults)	• Speak fluently in familiar language using appropriate word with respect to geography, caste/ethnicity, culture, community • Participate in conversation in small groups • Ask and respond the questions • Express views and opinions • Address and speak with elders and younger children according to local norms and values • Wait for their turn, listen to others and respond properly	• Answering the questions asked about child's name/ age/home/parents etc
	Literacy	Pre reading	• Demonstrate pre reading skills	• Interpret the picture and pictorial story • Match letters and words with pictures. • Recognise alphabets in Nepali and mother tongue	• Recognizing and naming the pictures/objects shown • Matching the pictures with their names given. • Making the words from the first letters/ alphabets given. • Recognize and reading out the alphabets/letter shown.
		Pre writing	• Demonstrate pre writing skills	• Draw pictures • Draw and copy patterns • Draw, make a line, half circle, full circle, pattern • Begin to form alphabets	• Copy the alphabets/letter shown • Draw horizontal/vertical line and make points/dots, half/full circle with pencil in paper. • Write owns first name in paper

Domains	Subdomains	Aspects	Standards	Tasks to perform	Tasks of Assessment for children
Cognitive development	Intellectual development	Cognitive skills and learning processes	• Explore and investigate. • Reflect on their learning use it to support	• Use puzzles for exploration and investigation • Use previous knowledge and skills for different day to day situations	• Fix puzzle pieces at appropriate places • Tell names of the days
		Classification and ordering	• Recognise and describe geometrical and other shapes •	• Describe familiar geometric shapes squares, triangle, circle	• Matchthe geometric shapes and tell their names
	Cognitive knowledge	Scientific exploration Sky /weather	• Describe sun/ moon/sky	• Name and draw picture of sun, moon, star and sky	• Name of given pictures • Sun, moon, star, sky
		Living and non-living beings	• Know the parts of body and their functions	• Name the parts of body • Differentiate the living and non-living beings	• Name the shown/drawn parts of body • Classify the pictures by living and nonliving being
		Scientific knowledge-materials	• Identify and describe everyday materials	• Identify utensils of everyday use • Identify the materials by their quality/property	• Name given utensils of everyday use • Spoon, plate, glass, scissors, frying pan, cookers by looking at the pictures • Classify the materials that float and sink
		Scientific exploration-technology	• Differentiate some materials and tools	• Name technological devices available in the context	• Tell the name of 3 means of communication (telephone, radio, TV, computer)
		Scientific exploration-transportation	• Describe types of transportation	• Name types of transportation available in the context	• Name the picture and tell what they are used for • Tell the name of three means of transportation

Domains	Subdomains	Aspects	Standards	Tasks to perform	Tasks of Assessment for children
		Distance and direction	• Follow and use direction	• Follow and give directions	• Give instructions to follow the directions such as: Raise your right/left hand, Take one step backward/ forward
		Concept of time	• Talk about different activities at time	• Name the days of week/ months of year	• Tell the name of days in week • Tell the name of first months of year
	Mathematics	Numeracy	• Recognize and write numbers 1-9	• Recognise and write numbers 1-9	• Recognize the numbers shown/given in sheet • Write the numbers 1-9
		Measurement	• Differentiate shorter/taller/ longer	• Perform comparing and measuring tasks	• Identify the tall/short men/ longer/shorter stick
	Creativity	Creative arts	• Create/draw shapes/sketches of objects/things	• Draw and name the pictures	• Draw picture of things of own's interest • Name the pictures • Colour the shape/picture
		Imagination, drama/role play	• Use imagination in drama/role play • Create music using instrument/objects	• Make/ create voice/ rhythm	• Make given voice/rhythm
Social & emotional development	Social development	Relationship with peers and unfamiliar adults	• Demonstrate relationships with peers and unfamiliar adults	• Greeting and interacting with peers and adults	• Greeting/addressing to peers and adults
		Life skills	• Expressing likes/ dislikes • Seek and provide support from/ to others in performing tasks	• Express own's likes/ dislikes with reason in choosing things • Perform task working collaboratively	• Give reason in choosing/ selecting thing • Build/form shape from building blocks
		Cooperation and coordination	• Take responsibility for own belongings	• Involve in ECD canter's chores and activities • Put things/belongings properly	• Follow the routine activities of ECD center • Put things where they were or are to be

Domains	Subdomains	Aspects	Standards	Tasks to perform	Tasks of Assessment for children
Social & emotional development		Social behaviour	• Demonstrate appropriate behaviour with peers and adults	• Mix up well while performing tasks	• Perform given tasks in group
		Self concept	• Demonstrate sense of self • Show pride of own existence	• Express own's likes and dislikes	• Tell own's likes and dislikes on books, food, dress etc.
	Emotional development	Emotional expressions	• Express emotions as per the situations • Show appreciation to others	• Clap to show appreciation to others	
		Emotional security and confidence	• Demonstrate confidence in new experience	• Demonstrate independence in individual and group work	• Perform the task individually and show/tell others how you performed it
		Self control balance	• Demonstrate balance over the emotions	• Modify the behaviour and expressions of emotions for different situations	
Cultural development	Values	Family and community	• Follow and value routines/norms of own class/family	• Demonstrate an understanding on basic dos and dons of own family/class	• Tell the dos and donts of: a) class • family
		Natural and cultural heritage	• Show an appreciation to local rules and cultures • Demonstrate respect for natural/ cultural heritage	• Know particular festival celebrated in his/her family/local community • Know dress/costumes of Nepali people of their locality • Know some don'ts in temple/mosques/ monastery of the locality	• Tell main festival of their family/community • Tell the dress of Nepali women of their community • Tell the name temple/ mosques/monasteries of their locality
		Nation and Nationality	• Show respect to Nepali emblems- Flag and Anthem	• Identify national flag • Knows Nepal's national song/anthem	• Sing one two lines of Nepali national song • Recognize Nepal's national flag

Conclusion

Though this framework incorporates main learning standards to assess by the domains accepted globally, all the learning tasks specific to the context may not have been covered. At the same time, the learning tasks mentioned in the framework may not be feasible to assess through single tool. So, it needs to consider assessing what is important and practicable to the specific context while designing assessment tools and administering them. While some of the learning tasks can be assessed through group tests, others are possible only through individual test.

Acknowledgement

This article is based on the framework for assessing early childhood learning and development in Nepal developed for ERO by Dr. Kishor Shrestha, Devina Pradhananga and Gopal Prasad Bhattarai. The authors offer sincere thanks to ERO for this opportunity of developing the framework.

References

Blair, C., & Razza, R.P. (2007). Relating Effortful Control, Executive Function, and False Belief Understanding to Emerging Math and Literacy Ability in Kindergarten. *Child Development, 78(2),647-663.*

Council of Chief State School Officers and Early Childhood Education Assessment Consortium. (2007). *The Words We Use: A Glossary of Terms for Early Childhood Education Standards and Assessment.* Available at: http://www.ccsso.org/Projects/scass/projects/early_childhood_education_assessment_consortium/publications_and_products/2892.cfm [accessed March 26 2017.

Department of Education, (DOE). (2062 BS). *Early Childhood Development Guidelines (Curriculum).* Sanothimi: Author

------- (2067 BS).*National Mimimum Standard for ECD Centers.* Sanothimi.Author.

-------.(2016). *Flash I Report 2073(2016-17).* Sanothimi: Author.

-------(2069 BS). *Early Learning and Development Standards for the Children from 48 to 60 Months Old*. Sanothimi: Author

Laat, J. D. (2015). *Measuring Early Childhood Outcomes.* Washington D.C: The World Bank.

National Research Council. (2008). *Early Childhood Assessment: Why, What, and How.* Committee on Developmental Outcomes and Assessments for Young Children, C.E. Snow and S.B. Van Hemel (Eds). Board on Children, Youth, and Families, Board on Testing and Assessment, Division of Behavioral and Social Sciences and Education. Washington, DC:

The National Academies Press.

Pisani, L., Borisova, I. & Jo Dowd, A. (2015). *International Development and Early Learning Assessment. Technical Working Paper.* NA:Save the Children.

Raikes, A., Rebello Britto, P. & Dua, T. (2014). *A Measurement Framework for Early Childhood: Birth to 8 Years of Age. A discussion Paper.* Washington D.c.: Institute of Medicine.

Raikes, H. A. (2016). *Measuring of child development and learning. Paper commissioned for the Global Education Monitoring Report 2016, Education for people and planet: Creating sustainable Futures for All.* Paris: UNESCO.

Rebello Britto, P. & Hancioglu, A. (2016). *Measuring Early Childhood Development in the SDG.* Webinar Presentations. New York: UNICEF.

UNICEF. (2010). *Early Learning and Development Standards for Children from Birth to 7 Years Old.* Romania: Vanemode.

Walker, S. P., Wachs, T. D., Gardner, J. M., Lozoff, B., Wasserman, G. A., Pollitt, E., Carter, J. A. & the International Child Development Steering Group. (2007). Child Development: Risk Factors for Adverse Outcomes in Developing Countries. *Lancet,* Vol. 369, pp. 145–157.

Factors Affecting Students' High and Low Learning Achievement

Dr. Rishi Ram Rijal

Associate Professor

Tribhuvan University, Mahendra Ratna Campus, Kathmandu

rishi.rijal@yahoo.com

Abstract

Student learning achievement encompasses the various knowledge and skills including getting mastery over the substantial knowledge as determined by the curriculum, capacity of academic skill and educational experiences. Some schools are successful in achieving the high student learning achievement whereas some others are unable to accomplish the success though they have the similar intakes, infrastructure and policies. There are various such factors which affect them severely. The high achieving schools usually have strong educational leadership, high expectations on pupils' achievement, safe environment for teaching and learning, frequent monitoring of students' progress and the provision of remedial teaching. The school administration involves teachers, parents and school management committee in making decisions and develops a sense of ownership in them in both decision and their implementations. The schools with high student learning achievement engages the teachers, students as well as parents to support their learning but the schools with low student learning achievement lack partially or completely these qualities.

Key terms: factors affecting learning, learning achievement, school culture, school leadership, time-on-task.

Introduction

It is very difficult to find any unanimous definition of student learning achievement because it encompasses not only getting mastery over the substantial knowledge as determined by the curriculum, but it also embraces the capacity of academic skills and educational experiences through reading and writing at a particular class or level. In Bhattarai's (2016,p. 39) view, students' learning achievement refers to the success in learning the subject matter taught, acquisition of knowledge and cognitive skills as expected by the curriculum or the mastery over the given content in the curricula. Citing Yuba Community College District Academic Senate (2005), he further says that it can also be defined as in terms of knowledge, cognitive skills and abilities that students have attained as a result of their involvement in a particular set of educational experiences. These definitions indicate that student learning achievement incorporates a variety of knowledge and skills like cognitive skills, proficiency in knowledge and understanding in a

particular subject through reading, writing, mathematics and science developed through formal schooling, however, it has been equated with the test score in numerical value.

Formal schooling has the goals of preparing the young people for equipping with the knowledge, skills and attitude essential for individual prosperity. Through the general and specific subject disciplines, they develop the literacy, numeracy and social norms and values that have to be essential to maintain their future life. For many people, the numeracy, literacy and socialization are the ultimate goals whereas for some others, the basic formal education will be the foundation of their further education. The skills and abilities which they develop in their learning and be able to perform as per the objectives set in the prescribed curriculum and syllabus are their learning achievements. These learning achievements are determined through some standardized tests but their factors affecting such high and low learning achievements cannot be shown with the result of test-scores of National Assessment of Student Achievement (NASA) only.

In this vein, through the in-depth study of multiple case studies, factors affecting learning achievement of students have been explored and presented as the summary of the report in this article. This study has sought to answer the questions like what are the factors which make some schools more successful and others less successful in student learning achievement. What suggestion could be made through the analysis and interpretation of the results that the successful schools have achieved? So, it focuses to find out the differences seen in the schools where there are similar intake and investment and to explore the measures of factors that facilitate the schools to achieve the high performance and other factors which affect the student low learning achievement. To find out the differences seen in the schools where there are similar intake and investment

Theoretical and Conceptual Review

Theory and concept

The students' learning achievement indicates the ability that the students have got on completion of some particular level or grade. They are expected to perform tasks that the course objectives have been set for the class, level or subject. In Danielson's (2009, p. 94) view, to gain the better learning achievement, students must be involved in learning being mentally active, making connection, formulating hypotheses, linking new understanding to what is known, participating in in-depth structural reflection and being engaged in collaboration.

Cullingford (1995, p. 179) has mentioned seven different attributes which affect the students' learning achievement. They include orderly and secure environment of school, trust between students and other teaching and non-teaching staff, awareness of the agenda of events affecting their lives (curriculum and beyond), personal involvement in learning (interactive teaching), understanding of the purposes of activities availability of opportunities and challenges and sharing of the agenda by all children. In this regard, those schools which are incoherent suffer

from the lack of the dialogue among and between all stakeholders of school system, lack of any focus for the collaborative energy of staff.

Cullingford (1995, p. 181), in the same vein, believes that the schools are the organizations in which order is always formed and reformed in different ways. There is the structure which describes the behavior of people involved directly and indirectly in the school as the stakeholders. Every organization has to serve many purposes of its own. The individual people are helped to develop their own belief systems which are firmly established in the school system through the behaviors of other people like principals, their favorite teachers and intimate friends.

Giving pupils rights and responsibilities and enabling them to play an active role in the life of the school is important for raising self-esteem and encouraging children to take responsibility of their own learning. Partnership that encourage and foster parental support for learning have positive effects on achievement, successful schools make demands on parents as well as encouragement and involvement. Effective schools are 'learning organizations' where teachers and managers as well as pupils continue to be learners, improve their practice and keep up with change.

Teachers are the members of school organization. They are expected to perform important roles at the school level including working co-operatively with colleagues, serving on committees and working with administrators and parents. Their work, working style and their initiation and collaboration in the school make significant difference in the students' learning achievements and schools' better performances. The environment of schools and teaching and learning culture are not only connected to students' learning but they also affect the adults' learning and development. In this context, Cullingford (1995, p. 10) has talked about the characteristics of an effective teacher as: a concern for other people; a willingness to work hard; a willingness to learn and to reflect teaching.

The teacher's main role is teaching academic content to students and evaluating their students' progress. Though there are some contradictions in their roles like individual help of the students, the distance between the teachers and the students to maintain their authority and to establish the closeness between them. School organizations are called cellular because every teacher is independently responsible for organizing leadership and teaching within the 'cells' of their own classrooms. There are head teachers, teachers and in some schools department chairs. Every classroom is regarded as cells within which the teacher is responsible for organizing the students' learning, managing the resources, disciplines, time and activities including their evaluation and progress. Teaching is purposeful, well organized and clear about objectives, well prepared, appropriately paced and structured, questioning focuses of pupils attention.

The result of reflection is demonstrated not only so much in the way the teacher presents himself in the high learning achievement of the students but it can also be seen in a shared working

atmosphere; in an awareness of the needs of each students; in a purposeful-well organized classroom and a celebration of success. Such signs can not be understood with what a teacher says but how he involves his students in the classroom teaching and learning activities. A teacher cannot be so effective with a lack of self-awareness, not knowing what he is doing whether it is right or wrong; defensiveness because they cannot bear any criticism however constructive and they react as the blame while giving feedback after observation.

The school environment and the teachers' planning, preparation and use of instructional materials make some schools more effective than others. Schools are human system that is influenced not only by the people who learn and work in them but by the larger community and society. According to Arends (2001, p. 412) schools are places where individuals do not have independence and freedom to work in disconnected ways. All the stakeholders including students, teachers, school administration, parents as well as School Management Committee (SMC) are more or less interdependent.

Though the individual students come to school to promote their purposeful learning, it is not possible only with his/her efforts. Where there is co-operation and collaboration between and among all stakeholders, the ultimate goal of schools that is improved students' learning achievement can easily be achieved and the schools' performance, head teacher's influence, teachers' acknowledgement as well as the students learning achievement becomes higher. In other words, a teacher or a student being intelligent and working independently can not improve the whole performance of the school securing higher average score of students' learning achievement. The all stakeholders' efforts and accountability should become the synergy developed by them acting in concert which can have important consequences for students' learning.

The primary goal of schools is to provide a purposeful learning environment that leads to the development of self-regulated learners (Arends, 2001, p. 413). Like in other organizations in schools too, there must be co-ordination of effort among all the stakeholders including teachers, school staffs, curriculum coordinators, administrators, counselors as well as students, parents and school management committee members. In Pollard's (2006, p.405) view, there is a clear priority for teaching and learning as the school's primary purpose. Four factors; time spent on learning, amount of homework, effective learning time, and learning time for different subjects are measures indicating the practical implementation of this focus.

The stakeholders of the schools like students, teachers and other administrative staffs must be present. The students must have compulsory attendance. Their parents must be supportive to help and involve in the school activities. It is believed that those parents who have high socio-economic status and who have high education can help and support their children in their learning and have high learning achievement (eg, ERO, 2014, 2016). That school which has high attendance of teachers and devotion of the school administration also has high learning achievement. In this regard, Pollard (2006, p.405) says that those schools where teachers share a common vision and

have shared values and have created a collaborative professional community characterized by dialogue and a sense of belonging produce higher students' learning achievement

A school's culture consists of the ways of its members thinking about their actions and reflection of their beliefs, values and history (Arends, 2001, p. 412). He further says that school culture greatly influences what goes on in schools and determines expectations and roles for beginning teachers (p. 412). Every school has its own norms, roles and the culture of teaching which are the guidelines of the school for the purpose of getting work accomplished. Norms are the expectations that the stakeholders of school system have for them in specific social setting of the school. The professional teachers are free to teach as they please within the confines of their classrooms which is known as the autonomy norm. Similarly, the teachers in the school do not interfere in the work of other teachers. This norm is called the hands-off norm (Arends, 2001). Though the teachers of a school have their autonomy norm and hands-off norm, the effective school encourages them to work together for the improvement of the school's performance as the whole. They have strong influence on the teacher's working life and the students' learning achievement.

Teachers have to do much more than the work in the classrooms which include preparation, planning, materials collection and designing, evaluation tools designing, homework checking, record keeping as well as the caring of the students. That's why teachers' time-on-task requires much more time than the personnel who work in the other organizations except school organizations. So, most teachers spend less than 40 percent of their work week on instructional activity and that a sizable portion of their work consists of meetings and exchanges with other adults such as parents, the principal and professional colleagues (Arends, 2001, p. 418). The activities they have to perform in schools include: Instruction, organizing, reviewing, testing, and monitoring. In addition to them, they do other work with students like study hall supervision, assemblies and clubs, control and supervision. They have to interact with colleagues and others in planned meetings and unscheduled meetings including desk and routine work along with travel time and private time.

Students' higher learning achievement is through a synergy which is at work in the school that produces result that exceed what a teacher working alone could achieve. In other words, the quality of teaching can be achieved through the co-operation and collaboration of the school team in addition to the money used for the physical facilities, books and other resources. In the same way, higher achievement of student learning are strongly associated with the aspects of school's social organization, such as the degree to which a common ethos exist, the extent to which teachers hold common attitudes, and the degree to which they behave in consistent ways toward their students (Arends, 2001, p. 420). Such kind of result was also consistent with the result of the study carried out by Ruther (1979), Brookover et al. (1979), Firestone and Rosenblum (1988), Rosenholz (1989) as some schools develop cultures and communities that support student learning where other school do not. Pollard (2006, p. 405) holds the view that in the school there

must be 'shared vision and goals' necessary for lifting aspirations and creating consistency of practice through whole-school policies and contracts.

Effective school's features can be divided into two categories: those that deal with the social organization and those that deal with instructional and curriculum patterns. Social Organization includes: clear academic and social behavior goals, order and discipline, high expectations teacher efficacy, pervasive caring, public rewards and incentives, administrative leadership and community support whereas the instructional and curriculum pattern includes: high academic learning time, frequent and monitored homework, frequent monitoring of student progress, coherently organized curriculum, varieties of teaching strategies, Opportunities for student responsibility. There is a general culture which has high expectations of everyone; teachers, pupils and parents. Better pupil outcomes follow from positive reinforcement, clear feedback, rewards and clear rules for behaviors. These are more successful than punishment or criticism.

Joice, Herst, McKibbin (1993) and Arends (2001) say that the achievement gained by students are significantly higher than in schools where teachers take collective responsibility for their students' academic successes and failures rather than blaming students (p. 421). The amount of staff co-operative makes a difference in student achievement. The findings emphasize the importance of teachers' work lives, particularly the beliefs they hold about their responsibility for student learning. Effective schools are more like caring communities than efficient bureaucracies and that caring communities are more likely to occur in small settings rather than larger ones. The teacher is the most important ingredient in the mix of factors that influence a child's education. It is teachers working together, not alone, which seems to be the crucial reason (Pollard, 2001, p. 422). Effective schools are those places where all stakeholders have common goals, teachers have organized their curriculum work coherently, common rules and norms guide teachers' expectations for students, homework policies and disciplines.

Review of empirical studies

With the beginning of National Assessment of Student Achievement (NASA) in 2013, under the Education Review Office (ERO) in the Ministry of Education (MOE) in Nepal, it assessed the learning outcomes of grade eight students to know where they were in Mathematics, Nepali and Social Studies. This report found that there were great difference in achievement between and among the students, schools, district and development regions. Institutional schools were found having higher learning achievements of students in comparison to community schools. Students were also found good in remembering and recalling the learnt materials and weak in comprehending, applying and analyzing the learnt materials. The students from the low socio-economic condition had low learning achievement. Similarly, the students' home languages also were found to have weak effect in some subjects.

After the discussion of the results and their affecting factors, it was found that the high and low achieving schools differed in terms of the number of study days, number of students, teacher-student ratio, time-on-task of teachers and engagement of students in their class work and homework, homework checking and giving instant feedback, parents' involvement in school activities and student learning support at their own homes, parents' literacy, economic and social condition, etc. the high achieving schools used more time on class work and homework checking, higher study days, smaller student number and smaller teacher-student ratio, teachers' and students' higher attendance, positive attitude towards their teaching and learning, higher literacy rate of their parents and high socio-economic status.

The result of the assessment conducted in 2012 in class 3 and 5 by ERO also showed that there is unbalance in students' learning achievement across the regions, districts, institutional and community schools but the variable of gender and ethnicity and rural-urban areas did not have any effect on student learning achievement. In Mathematics, the achievement level in Algebra and numeracy was remarkably lower than arithmetic and geometry. The result also showed that there was low capability to solve tasks that required higher ability. Like in earlier NASA, there was wider disparity in the achievement between and among different districts, regions and school types. In the same vein, there was wider variation in the student learning achievement from the different linguistic communities, socio-economic status and parents' involvement in schools' activities like helping their children at home and taking part in school meetings, etc. It could also be seen as the negative effect of the lack of books, bullying and unfair treatment from the teachers.

As in the previous assessments, the result of NASA 2015 conducted by ERO showed that the average achievement has remained at 45 to 52 percent varying by subjects and grades (ERO, 2016, p. 3). In Nepali the third graders have achieved 52 percent which is 46 for 5th graders. Similarly, in Mathematics, the higher graders have achieved 48 percent but lower graders 45 percent.

The result of such learning achievement have also been seen as in the earlier results disproportionately distributed achievement level, wider variation in terms of location, lower level of cognitive ability, variation due to home language, strong link of socio-economic and educational background of parents, their involvement in students' learning and school activities, availability of books in time, regularity of students and teachers and stagnant learning achievement over the years. These reviews of the results and factors affecting student learning achievement show that the affecting factors and the results of various standardized tests taken by ERO and the other authorized units seem to be similar. Though it seems like some corrective measures have been taken on the part of government, the situation and learning achievement have not yet been reformed as expected.

Conceptual framework

The above review indicates that the in-depth studies are to be carried out for that very reason. In this context, multiple case studies were carried out of four schools of Sarlahi, Dhanusha and Sunsari districts last year (2016). I have developed a conceptual framework for the study by synthesizing the reviews of theories, concepts and studies on student achievement. The conceptual framework indicates various factors affecting students' learning achievement.

Figure 1 : Factors affecting students learning achievement

Methodology

The case studies were the research design in which the in-depth information has been generated through these techniques like observation, interview, focus group discussion, document analysis (Cohen, Manion and Morrison, 2011). The cases were studied from all dimensions for the authenticity, genuineness and trustworthiness. Therefore, in this research study, the information is generated not only from the students, teachers, head teachers and other parents, guardians, SMC chairs. The districts and schools were determined by the client. The tools for data collection were separately designed for the students, parents, SMC members, teachers, headteachers. The interview guidelines were designed for discussing with the head teachers individually. These guidelines were used to talk to teachers, parents, SMC members and students in groups. Some guidelines were also be prepared for the classroom observation and observation of the overall schools. These tools were prepared in both English and Nepali languages so that the respondents could understand what they were supposed to share. After designing the tools, they were implemented in a school to test whether the tools are too difficult to understand for the respondents and whether they are able to generate expected information. The result of the tools have been analyzed to finalize them and to be used in the assigned field.

Presentation and Analysis of Data

After going through the available literature of effective schools, the NASA results and other related literature along with the information received through various sources like focus group discussion with students, classroom observation, general observation of schools and its records, talking to the subject teachers, head teachers, PTA representatives and the chairs of school management committee, the information have been thematized into different categories. The educational leadership, high expectations on students' learning achievement, safe and orderly environment of school, emphasis on acquisition of basic skills, frequent monitoring of student progress, number of school days, number of schools, student teacher ratio have been categorized under the school related factors. Similarly, teachers' time-on –task, time on homework, homework checking and giving instant feedback, teachers' regularity, dynamic personality of the teachers have been categorized as the teacher related factors. In the same vein, students' regularity, attitude towards learning and their efforts have been categorized as the student related factors. Family's socio-economic and educational status and their involvement in school activities and support in their home have been categorized as the family related factors affecting student learning achievement. These themes and the result on these themes have been discussed in the following section of this data presentation and findings.

Strong educational leadership

The working culture and high achievement of an organization is dependent on its leadership. If the leadership is purposive, visionary and transparent, the working atmosphere and the success

rate of that organization become high. Applying this principle in the two sample schools, it can be said that both school 'A' and 'B' have the same infrastructure and human resources like subject wise teachers who are both B.Sc. and both of them have the experience of over twenty years. Similarly, both of them manage substitute teachers in the absence of subject teachers but one of them who has been teaching in the high achieving school (the School 'A') encourages to give home work everyday, check it minutely and for giving instant feedback but the head teacher of school 'B' does not encourage to give home work and feedback everyday. The head teacher of school 'A' also teaches specialized subject: science in grade eight but the teacher in school 'B' though engages classes of others in their absence, he does not have regular class in routine.

The leadership of school 'A' seems to be strong and purposeful. According to the other stakeholders like teachers, students, SMC members and the head teacher himself, when he came to this school, there were nearly one hundred students in the whole school but when we went, there were 350 students. According to the PTA representative, he goes to the guardians' houses and convinces them to send their children to school. But such behavior was not found in the head teacher of school 'B'. The head teacher of school 'A' has engaged the other teachers in other various committees. His authority is accepted by all teachers and staff without any hesitation because all of the teachers are younger than him whereas, the authority of the head teacher of school 'B' was not accepted by all other teachers because some teachers go to the boarding schools in which they have invested. Similarly, the head teacher of school 'A' not only calls the SMC members to talk about the improvement of student achievement but also frequently visits the students' guardians for their involvement in the school activities and students' learning support. The head teacher of school 'B, though he claims he had been trying to increase the co-operation, collaboration and involvement of the teachers, parents and SMC chairs and members, he has not been able to do so.

Those schools have been less successful where the teachers feel isolated in the school; they have no influence on others as well. The influence of the head teachers that they bring to the post causes a great deal of difference in the students' learning achievement and teachers' professional development. However, the school 'B' being led by the head teacher seeks merely to please the outside world and seeks merely to placate the staff.

High expectation on pupil's achievement

This factor has high significance because if the teachers, students and the school administration do not expect high learning achievement, their work, working styles and the teachers' planning and preparation and implementation of these activities along with the use of instructional materials will be badly affected. In this regard, though the teachers of school 'A' did not have any written lesson plans, the head teacher had given them ready made diary to make their plans and daily lesson plans in point form, but the teachers of school 'B' who think they do not need any lesson plan because of their experience of over 20 years teaching the same subject did not have any plans and they prepared the lessons just before going to the classroom in their leisure

time. The students also do not expect so high achievement because their morale was found low because of their study in community school. This shows that the head teachers, teachers as well as the students do not have the expectation of high learning achievement.

Safe and orderly environment

Environment where the teachers, students and other non-teaching staff work should be safe and orderly. The environment becomes safe and ordered because of the community where the school lies, where there is compound in the school premise and where the school leadership involves every stakeholder to prepare a code of conduct and its implementation in which they themselves are bound together. Regarding this factor, both of the schools 'A' and 'B' seem similar. However, the school 'B' has made a code of conduct but it has not been strictly implemented in the practicality because some teachers have invested in some boarding schools and they entice some good students of the school in the name of scholarships and they leave the school before 4 o'clock as well.

Emphasis on the acquisition of basic skills

Basic skills are the foundation of further learning. If the schools give emphasis on acquiring such basic skills, the students' foundations will be strong. Since the students in these both schools came to study from the low background in terms of the social, economic and educational background, their foundation is always poor. The school 'A' has tried to develop the students' basic skills through running Early Childhood Development (ECD) classes and by teaching some basic subject in English medium but such efforts were not found in school 'B' even though the infrastructure of the school was rather better than that of school 'A'.

Frequent monitoring of student progress

It is necessary to any body or any institution to know where I am or where it is in terms of progress. This can be possible only through frequent monitoring and follow up. For such monitoring, the teachers and head teachers must be self-reflective. Regarding this factor, the school 'A' seems to be far ahead than the school 'B' because the school 'A' frequently takes examination and test in between the first quarterly, second quarterly as well as the final examination of every class. However, the school 'B' only takes examining as specified in the annual calendar as the first, second quarterly and final examination. Both of the school didn't have any provision of keeping students' portfolios to main the records of students' progress and achievement.

Number of study days, students' number and teacher student-ratio

The students' learning achievement is affected because of various reasons. As the school

'A' lies in the distance near the border of India and a difficult corner to reach for the strikers, the number of school days seem higher than the case of school 'B' which lies only 13 kilometers from the head quarter of the district and there is frequent disturbances from the political parties. On the other hand, the school 'B' has +2 as well as campus level in the same building, the study days are more disturbed than the school 'A' where there is only basic education level. The number of students and teacher-student ratio is also smaller in school 'A' in comparison to school 'B'.

Teachers' time-on-task

The theory shows that less than 40% of teacher's time is spent in classroom instruction and other more than 60% time is spent in other tasks like planning, preparation, meeting with parents, homework checking and others. As the school 'A' is smaller, has less number of students and younger teachers, they spend more time on these tasks. But in the case of school 'B', the teachers spend more time on classroom traditional work than other tasks. That's why, the effect of this was directly seen on students' learning achievement. This is also because of the head teachers strictness and devotion to the school. Many of the teachers, according to the students of the school taken to the focus group discussion, do not enter the classrooms in right time and leave the classroom five minutes earlier than the bell.

Time on homework and checking it for giving instant feedbacks

One of the most affecting factors of students' learning achievement is giving homework, checking it rigorously and giving instant feedback. Regarding this factor, the teachers of school 'A' seem to be more active than the teachers of school 'B' according to the students of the focus group. Some teachers of school 'B' were claimed not to give homework even once a week. This shows that homework giving, checking as well as giving appropriate feedback have important role in learning achievement of the students.

Teachers-student regularity

Many studies (e.g. ERO 2012, 2014) have shown that the teachers and students regularity affect very much in the learning achievement of the student. While observing the classroom teaching and checking their attendance, the research shows that in school 'A' only two students were absent from the classroom whereas in the case of school 'B' nearly 40% of students were absent from the class. Similarly, in school 'A' there is the strict provision of giving prior notice of taking leave by the teachers but in the school 'B', some teachers were absent from the school when we went there and when asked whether they gave any notice or leave letter before staying in the short term leave, they could not show any. In the school'A' the teachers stay on leave with pre-notice to their head teacher but according to the head teacher of school 'B' it was a bit difficult to bind them in such a code because some of the teachers of school 'B' have invested in the boarding

schools near Janakpur town or in the vicinity. This shows that the presence and absence of the both teachers and students also affect student learning achievement.

Attitude towards learning

Attitude here means motivation and eagerness of learning more. If the students are motivated by the teachers or schools, they do not like to bunk off the classes and pay high attention to their learning. In this regard, though students of school 'A' had parents with much more low socio-economic status, were involved solely on agriculture and did not have education higher than the eighth grade, they had positive attitude towards their learning. But in the case of students of school 'B', some of them complained that their teachers' behaviors discouraged them to come to the school. Their morale also seemed very low for being the students of community school. Instead of encouragement in their learning, some teachers always encouraged them to take tuition classes with them in their residence.

Collaboration and involvement

Another most important attribute of gaining higher learning achievement is involving the teachers in decision making, policy designing and curriculum development and ensuring them that they can have their own programs designed for their classes and for the whole school performance. These activities will lead securing consistency and fairness in all teachers. In this context, Pollard (2006, p. 405) maintains that monitoring progress keeps track of whether the school is meeting its targets and goals; maintains awareness of targets and goals among staff, pupils and parents, informs planning and teaching; sends clear messages to pupils and that teachers are interested in their progress.

The most essential component of a high student learning achievement of a school is communication. Those schools which involve all teachers in decision-making and the development of the policy bind themselves without isolating any teacher from intensive discussion to make the school better. According to Cullingford (1995, p. 183), "the curriculum only becomes lively and useful when it is the matter of debate and when the resources that support its content are shared. He further adds that to be high student learning achiever school, teachers of the schools need to be willing to talk to each other about educational matters; they need to be open-minded and have a sense of their own responsibilities."

Dynamic personality of teachers

Teachers are the greatest asset of the schools. Without their active and enthusiastic participation the prescribed curriculum can not be a live one. Teaching is such a profession which never ends with a certain certified degree but begins from there. So these teachers who are young and eager to learn; they can transfer their lifelong learning habit in their students as well; they can

be up-to-date with the recent knowledge and skills and become more self-aware and set their own future direction. In the case of these schools, the head teacher of school 'A' said that the teachers of his school had to learn at least one book in a week and summarize it to help in his own subject. The teachers were also young and energetic because though they were teaching in the grades of basic level most of them held master degree, but the head teacher of school 'B' was aged one and most of the other teachers were also older than the teachers of school 'A'. Except taking part in some short term trainings, which could provide allowance, they did not mention any further development processes.

School neatness and premises

Though we reached the schools 'A' and 'B' both of the schools were not so neat and clean for the health of the students. The school 'B' was bigger, had enough non-teaching staff but they could not be properly handle the school work like keeping sanitation in the school premises. The school 'A' had a small ground with three buildings with tined roof but in the school 'B' there were four concrete two storey buildings with enough rooms for students. Both of the schools had fearless learning situation but in school 'B' some of the students could also be involved in bulling and destruction in higher grades. In both schools the head teachers claimed that they analyzed the needs of students but except taking some quarterly and final tests, the school 'B' does not involve in other progress tests.

Physical facilities

In school 'B' there was the provision of clean drinking water, science lab, various readymade instructional materials, good school infrastructure and the students who passed sent-up test also got the chance of taking coaching classes. It has enough level wise subject teachers' quota along with enough space for play ground and one donated school bus for the small children, primary health care and time and again the distribution of medicine for killing worms. However, because of the weak school administration, less time for the study, teachers' jagire mentality, enticement of good student to take their own invested boarding school, not so amicable relationship with SMC and PTA members, less homework given and checking with instant feedback the students' learning achievement in science of 8th grade was very low – the lowest in two districts Sarlahi and Dhanusha.

Remedial teaching, monitoring and supervision

In school 'A' even the school did not have enough rooms for the administration and classroom teaching, not enough provision of clean drinking water, no library and science lab, a very few ready made instructional materials, no provision of remedial classes, not enough subject wise teachers' quota, less use of progress tests, no monitoring from the line agency because of the difficult place from the head quarter of the district narrow place of school complex, a very few

instrument of games, sports and entertainment students' low attendance in the classroom; it had the highest student learning achievement, the reason behind it might be the purposeful leadership, less absence of teachers with giving prior notices, shared vision of the teaching staff or let's say the acceptance of the head teacher's authority and expertise in teaching as well as the trail of the head teacher to keep in tough with the students' guardians.

Those schools that keep careful and precise record of the students' work, including their personal, educational and social development, have higher student learning achievement. Better pupil outcomes follow from positive reinforcement, clear feedback, rewards and clear rules for behaviors. These are more successful than punishment or criticism.

Parents' support in student learning

Students spend one fourth part of their time in the schools every day. The remaining time is spent at their own homes with their parents and guardians. They are familiar with their habits, attitudes and dispositions. If the parents and guardians are involved in school activities, whether they are formal or informal, whether they enter the classrooms for the teachers' help or they come to know the progress of their children can be made sure and they can also be involved in decision making process, these schools can gain higher learning achievement. However, both of these schools lack these qualities.

Feedback and encouragement

Teaching is the job of carefully structuring the different teaching and learning activities, providing the students plenty of model, time and work to do with appropriate feedback and encouragement in the students' efforts. Though teaching seems to be intellectually challenging they can cope with the immediate problems with their own intuition and imaginations. Similarly, teaching is not only keeping students silent and disciplined in the classroom working students together. It is also very necessary to pay their attention to their individual needs.

Following the constructivism in teaching-learning process the interaction between the teachers and students is the main process of learning. The students should have the highest time to interact with other colleagues in the target subject but students must also have some particular area in one session. Effective learning involves higher order thinking, the students have to be concentrated on the particular thing they are about to learn. If they are dominated, criticized or punished in some minor weaknesses, they feel scared and cannot concentrate in the content they are learning. In such cases, those schools which have developed the culture of positive attitude and encouragement and praise for their efforts the students of these schools can have higher learning achievement on the one hand, as well as there is amicable relationship between teachers, parents, students and administrators.

Summary of Findings

On the basis of the above data and analysis, it can be said that though both schools have the same type of intake, similar infrastructure and same type of socio-economic condition as well as the educational status of parents, the schools differ in terms of student learning achievement in science subject of grade eight. Some of the findings have been summarized in the following table:

School 'A'	School 'B'
• Strong educational leadership • Higher expectations of student learning achievement • Safe and secured environment for both teachers and students • Emphasized foundational strengths • Regular monitoring of students' progress • Higher study days, smaller student teacher ratio, and greater teachers' time on task • Regular homework, rigorous homework checking and instant feedback • Teachers' and students' regularity in school • Because of the co-operation and collaboration of the teachers positive attitude of student in learning • Young and energetic and higher qualified teachers	• Weak educational leadership • Lack of high expectation on student learning achievement • Safe and secured environment for both teachers and students • Lack of foundational strengths • Only quarterly tests and final examination • Disturbances of various factors like strike, higher graders and other programs of the community members, larger teacher-student ratio, less teachers' time-on-task • Irregularity in homework, homework checking and feedback • More irregular teachers and students • Because of the less co-operative collaboration with teachers and enticement in taking tuition classes as well as to the boarding schools, good students had negative attitude towards their schools • Teachers taking in trainings only for allowance not for learning and teaching.

Conclusion

The most significant factor that affects student learning achievement is the purposeful and visionary leadership. The head teacher has to develop a sense of ownership with the teachers and school staff towards the school and the decisions they make. Similarly, he or she also involves not only teaching staff in the teaching learning work but also the parents and guardians in support of their children's learning. In the same way, many subject teachers are very weak though they are so called qualified and trained. If the head teacher regularly observes their classes and gives feedback on the one hand and chooses the best one among them (if they are to be chosen), the student learning achievement will be higher. We must accept that the students' home background

has the effect of their learning, but if the teachers spend the whole time in the classrooms, involve their students in the activities, if they get enough preparation, use essential materials and provide instant feedback, certainly the student learning achievement will be higher. The problems lie not in students but they lie in administration, subject expertise, shared vision, collaboration, teachers' reflection and parents' formal and informal involvement in their children's learning.

Almost all teachers whether they are new or experienced have never entered into other teachers' classroom, neither for the purpose of learning from the experienced and nor for giving critical feedback to improve classroom teaching behaviors. Both teachers as well as the students have rarely felt any reflective activities in the schools. The teachers in both schools say that they are overloaded. They have to teach 35/36 periods in a week that's why they cannot check their students' homework so minutely. However, those teachers who have to teach only 22/23 periods in a week also have the same complain.

The reality of complexity of the school organization has not been realized by many people including the teachers themselves, scholars, media personals as well as the SMC, PTA and school leadership. Because of this, teaching can be thought to be done even without doing any methodology courses, educational psychology, learning principles as well as foundational courses and hands-on experiences in the students' actual classrooms. Even the teachers who are so called trained have been found to have used their own experience through their student life of their initial classes of schooling. Therefore, there is a misunderstanding (Arends, 2001, p. 411) on the part of teachers, parents and even policy makers regarding school improvement efforts.

Many teachers in Nepalese context learn to teach with the indirect experience gained while they themselves were students by observing the teachers' classroom behaviors. Similarly, though many teachers have received a number of short and long term training, their teaching behaviors have been heavily dominated with their own initial concept of teaching. If they are asked to explain why it has been happening, they try to rationalize that they do not have such environment to use their knowledge and insight through their pre-service as well as the in-service trainings, they are not co-operated by their school leadership, their students do have low standard; they have to finish their assigned courses in time; they do not have sufficient facilities and instructional materials; but they do not have the knowledge of showing the discrimination between the real object, models and pictures and maps. They are accustomed to receive any ready made materials like annual work plans, teaching grids as well as question papers from the resource centers.

Acknowledgement

As this article is based on the study conducted for ERO in 2016, the author offers sincere thanks to ERO for providing opportunity of carrying out the case study on two high and two low performing schools. A sincere thanks also goes to Dr Netra Prasad Poudel for working as a co-researcher of this study.

References

Arends, R. (2001). *Learning to teach.* Singapore: McGrow Hill.

Bhattarai, G. P. (2016). Raising Students' Learning Achievement: Concerns and Context of Nepal. *In Nepalese Journal of Educational Assessment* vol. 1(1), 37-56.

Cohen, L., Manion, L. Morrison, K. (2011). *Research Methods in Education. London:* Routledge Falmer.

Cullingford, C. (1995). *The effective teacher.* London: Cassell

Danielson, C. (2009). *Talk about teaching.* New Delhi: Corwin.

ERO. (2011). *Where are we now? Student achievement in Mathematics, Nepali and Social Studies.* Sanothimi: ERO.

ERO. (2016). *National assessment of student achievement (NASA) 2015 Grade 3 and 5.* Sanothimi: ERO.

ERO. (2014). *National assessment of student achievement (NASA) Grade 3 and 5.* Sanothimi: ERO.

Pollard, A. (2006). *Reflective Teaching.* London: Continuum.

Request to the Writers for Academic Articles

The Education Review Office (ERO) invites academic articles from writers for the forthcoming issue of Nepalese Journal of Educational Assessment (NEJEA) on the area of educational assessment, which includes themes, principles, theories, practices, issues and innovations on student assessment and performance audit of educational institutions. Along with electronic copy of article writer (s) should provide with a brief CV.

Articles should:

- **be written in English language.**
- **consist minimum of 4000 to maximum of 6000 words typed using Times New Roman Font.**
- **be well organized in standard form along with proper citation and referencing in established format based on APA referencing guidelines.**
- **have a short and precise abstract of 100 to 150 words.**
- **be an original not published elsewhere earlier.**
- **be either conceptual or empirical, however, we encourage for the articles having a strong linkage between theory and/or empirical research with the implications for policy and/or practice on educational assessment.**

Editorial Board reserves the right to accept, reject or publish with editing the articles received.

Government of Nepal
Ministry of Education
Education Review Office
Sanothimi, Bhaktapur
Nepal
Phone: +977-1-6633556, +977-1-6633557
Fax: +977-1-6633556
Email: nejea.ero@gmail.com
Website: www.ero.gov.np

NEJEA
ERO

Volume 2 • Issue 1
2017 June