Teachers Competence in the Educational Assessment of Students: The Case of Secondary School Teachers in the Amhara National Regional State

Asnakew Tagele[1] and Lake Bedilu[2]

Received: 18 April 2016; Accepted: 09 February 2017

Abstract: The purpose of this study was to assess secondary school teachers' competence in the educational assessment of students in Amhara National Regional State. Data was collected using a modified form of the "Teacher Assessment Literacy Questionnaire," a four-option multiple choice test from 814 teachers. The questionnaire consisted of seven competence areas and a total of 31 items. Means and standard deviations, and proportion (percentage) were used to describe the collected data. Moreover, t-test and one-way ANOVA and repeated measures ANOVA were used to analyze the data. The result showed that secondary school teachers in Amhara National Regional State failed to demonstrate competence in the educational assessment of students. They failed to demonstrate competence in the overall and in each of the seven standards of educational assessment of students. The t-test showed that those who took a separate course in measurement and evaluation during their training scored statistically significantly higher than those who did not take a course where the mean for those who took a course being greater by 1.28 points. Although it was not statistically significant, language teachers scored higher than those who were teaching other subjects, where the mean difference was one point. The results implied that the status of training prospective teachers in student assessment should be analyzed seriously. The Regional Education Bureau may consider developing in-service training programs in educational assessment for teachers.

[1]Lecturer, Department of Psychology, College of Education and Behavioral Sciences, Bahir Dar University; P.O. Box 79; e-mail asnakewtagele@gmail.com
[2]Assistant Professor, Department of Psychology, College of Education and Behavioral Sciences, Bahir Dar University; P.O.Box 79; e-mail lakeymca@gmail.com

Introduction

A comprehensive research by Darling-Hammond (2000) showed that schools can make a difference, and a substantial portion of that difference is attributable to teachers. This review indicated that differential teacher effectiveness is a strong determinant of differences in student learning, far outweighing the effects of differences in class size and heterogeneity. Sanders and Rivers (as cited in Darling-Hammond, 2000) revealed that students who are assigned to ineffective teachers have significantly lower achievement and gains in achievement than those who are assigned to highly effective teachers. Hence, to improve student learning, teachers will need to increase their knowledge and skills.

Assessment is a critical and an integral aspect of teaching and learning process. Khan (2012) indicated that the quality of learning is determined by the quality of assessment practices in the classroom. Assessment is used to make educational decisions about students; to give feedback to students about their progress (strengths and weaknesses); to judge instructional effectiveness and curricular adequacy; and to inform policy (Sanders and Vogel, 1993). Ojerinde (2009) pointed out that assessment is at the heart of education as assessment results are used to gauge students' academic strengths and weaknesses.

Assessment is one of the most powerful educational tools for promoting effective learning (Assessment Reform Group, 1999). It is early in 1962 when Robert Ebel indicated that the assessment of educational achievement is essential to effective education. In a review of research on assessment and classroom learning, Black and William (1998) synthesized evidence from over 250 studies linking assessment and learning. The outcome was a clear and indisputable message: that initiatives designed to enhance the effectiveness of the way assessment are used in the classroom to promote learning can raise

student achievement. Thus, student success is largely dependent on teacher practice.

In support of this relationship between teacher practice and student success, Rice as cited in Greenstein (2010) asserts that "Teacher quality matters: It is the most important school-related factor influencing student achievement" (p. 1). Black and William (1998) also concluded that teacher's skill in classroom assessment enhances student achievement. To better understand this relationship, they indicated that a teacher who begins at the 50th percentile in his/her skill at using classroom assessments and a student in his/her class who begins at the 50th percentile in his/her achievement. If the teacher increases his/her effectiveness at using classroom assessment to the 84th percentile over time, one would predict that the student's achievement would increase to the 63rd percentile. Similarly, if the teacher increases from the 50th to the 99th percentile in his/her skill at using classroom assessments, one would predict the student's achievement to increase to the 78th percentile (Marzano, 2006).

Assessing student performance is one of the most critical aspects of the job of a classroom teacher. It impacts nearly everything that teachers do. In their role in the classroom, Stiggins as cited in Plake and Impara (1997) estimated that teachers spend up to 50% of their instructional time in assessment-related activities such as quizzes, tests, questions, and projects. The percentage of time spent on assessment or assessment preparation is growing (Quilter, 1999). In the Ethiopian Educational System, a good portion of the budget also goes into formal testing. With so much time and money devoted to assessment, it's worth critically understanding the knowledge and skills teachers possess in the educational assessment of students.

Further, student assessment is an essential part of teaching, and that good teaching cannot exist without good student assessment (Eckhout, et al., cited in Kiomrs, Abdolmehdi, and Naser, 2011). However, studies show that teachers consistently use a variety of factors in their assessment practices and consequently make erroneous decisions. Most teachers lack effective assessment knowledge and skills; that is, when evaluating student academic achievement, teachers exhibit misconceptions about assessment practices (Cizek, Fitzgerald, and Rachor, 1996; McMillan, 2001; cited in Chen, 2005). As a result, the continuing need to develop the potential of classroom assessment to support learning has recently been stressed by some researchers in the field (Assessment Reform Group, 1999). In particular, Black and William (1998) have called for research which supports teachers in trying to establish new practices in formative assessment.

Hence, a shortfall in the competence of teachers in this area means that assessment benefits may not be realized. Teachers are responsible for highlighting students' strengths and need improvement areas, as these support or hinder students' learning. Thus, assessment has a central and paramount role in extending support to students' learning outcomes. Teachers' relevant assistance and guidance based the assessment data become very significant in enhancing students' learning. So teachers should have a clear understanding of relevant and effective assessment strategies and have to implement these strategies in the classroom. In other words, teachers should have competence in educational assessment of students if the goals of the educational system such as quality education for all are to be achieved.

Statement of the Problem

The need for developing standards to guide teachers' professional preparation and in-service training in the assessment was recognized in America as early as 1912 (Starch and Elliot, 1912). Popham (2009) indicated that in America, increasing numbers of professional development programs have dealt with assessment literacy for

teachers and advocated that assessment literacy must be a pivotal content area for staff development endeavors.

Professional organizations have also acknowledged the need for assessment literacy within the teaching profession. Teachers' organizations, as well as the educational measurement community, have come together to promote assessment literacy through seven assessment standards (American Federation of Teachers, National Council on Measurement in Education, and National Education Association [AFT, NCME, and NEA], 1990). The skills advocated in the Standards for Teacher Competence in Educational Assessment of Students (AFT, NCME, and NEA, 1990) are: (a) choosing appropriate assessment methods; (b) developing appropriate assessment methods; (c) administering, scoring, and interpreting assessment results; (d) using assessment results to make instructional or curricular decisions; (e) developing appropriate grading practices; (f) communicating assessment results; and (g) recognizing unethical, illegal, and otherwise inappropriate uses of assessment information. According to AFT, NCME and NEA (1990), some of these standards focus on classroom-based competencies while the other standards address assessment competencies underlying teacher participation in decisions related to assessment at the school, district, state, and national levels because of teachers' growing roles in education and policy decisions beyond the classroom. These competencies are the knowledge and skills critical to a teacher's role as an educator.

Assessment is a focus in today's educational agenda in Ethiopia, though it seems merely a fashionable focus for today's professional developers, not as a significant area of professional development interest. Moreover, the Ministry of Education also developed five standards on assessment under the professional standard for Ethiopian school teachers (MoE, 2012). However, these standards are comprehensive, and the five standards proposed overlap with each other.

Research studies addressing one or more of the seven standards have been conducted over years. However, very few studies (Impara, Plake, and Fager, 1993; Plake and Impara, 1997; Campbell, Murphy, and Holt, 2002; Mertler, 2005) have specifically examined in-service teachers' knowledge of assessment to meet the seven standards.

Although the research showed an important connection between the quality of teachers' classroom assessments and students' achievements (Stiggins, 1999), as indicated in the preceding section, no comprehensive research is done about our secondary school teachers' competence in assessment. Although the emphasis is given to the relevance of assessment in the Education and Training Policy (FDRE, 1994) and the school development program of the government, no research is conducted in the local context. A preliminary study at Bahir Dar town secondary school teachers indicated that teachers' competence in the educational assessment of student learning was very disappointing in that it was very poor (Lake, 2014). However, a large-scale assessment is necessary before proposing an intervention to develop assessment competence of teachers at the regional level. In other words, there is a pressing need to conduct a study that assesses the competence of secondary school teachers in Amhara National Regional State in the educational assessment of student learning.

Thus, based on the above background information and statement of the problem, the following research issues will be addressed:

- Teachers' knowledge about educational assessment;
- Whether there are differences in the competence of teachers among the seven competency areas in the educational assessment of students;
- Whether there is a difference in assessment literacy between male and female teachers;
- Whether there is a difference in assessment literacy between different subject teachers' competence in the educational assessment of students, and

- Whether there is a gap between teachers' competence in the educational assessment of students as assessed by the TLI against their actual competence as demonstrated in the exams they prepared.

As indicated above, the study has the following objectives:

- Assess the extent to which secondary school teachers in Amhara Region are competent in the seven standards for teacher competence in the educational assessment of students;
- Examine whether there exists a difference in the levels of competence of those teachers who took and didn't take a separate course in measurement and evaluation during pre-service training;
- Investigate whether there exists a difference in the levels of competence between male and female teachers;
- Investigate whether there exists a difference between different subject teachers' competence in the educational assessment of students, and
- Investigate whether there exists a gap between teachers' competence in the educational assessment of students as assessed by the TLI against their actual competence as demonstrated in the exams they prepared.

This study is timely and significant. Nowadays, focus is given to the quality of education which largely depends on the competence of teachers in all educational activities of which student assessment is a key. The findings of this study may provide information to teachers, schools, education offices, and teacher training institutions about the level of assessment literacy of secondary school teachers to improve teachers' competence in the educational assessment of students. The results of this research may also inform teachers to improve their student assessment competence through lifelong learning. Schools and education offices may use the results as resources to plan a continuous professional development program to improve teachers' competence in educational assessment. Teacher training institutions

may also consider reviewing their teacher education curricula to put emphasis on trainees' competence in educational assessment. Policy makers may also use this as resource for giving attention to student assessment skills in pre-service training. Moreover, the study may serve as a foundation for further consecutive studies in the area.

Apparently, the study is not without constraints. Teachers were allowed to take the questionnaire home or office, fill/complete it and bring it the as they finished/completed. As a result, they might have an opportunity to consult resources and/or discuss with colleagues to answer the questions. Thus, the lack of a controlled testing situation might result in a response that may not reflect the respondent's actual knowledge and skill in the area and might exaggerate the results.

Operational Definition

Assessment Literacy refers to the level of competence of teachers as evidenced by their responses to the teacher assessment literacy questionnaire. As a result, high score on the questionnaire reflects good literacy on the educational assessment of students.

Secondary School refers to the first cycle (Grades 9 – 10) of the secondary education.

Methodology

Population, Sample and Sampling Design of the Study

Based on data obtained from the Amhara National Regional State Education Bureau, there were 288 secondary schools in the region in the 2010/11 academic year. The Amhara Region consists of eleven zones. The numbers of schools in these zones differ ranging from eight in Waghimera to forty-seven in North Gonder. Hence, the region was categorized into zones and to adequately represent zones, the number of schools to be selected in each zone was determined using proportional allocation (proportional to the number of schools in each zone). Then schools in each zone were selected using systematic random sampling based on their names secured from the Education Bureau and arranged in alphabetical order to construct the sampling frame. Finally, using equal allocation, from each of the selected schools 20 teachers were selected again using systematic random sampling based on the list of their names obtained from the respective schools. Selected teachers but who were absent on that specific day of data collection were substituted by available teachers who were next in the series of names arranged in alphabetical order (sampling frame). A total of 1000 teachers were selected from 50 schools through this procedure.

Table 1: Number of Schools, Selected Schools, and Selected Teachers by Zone

No.	Zone	Number of Schools		Number of Teachers Selected
		N	*n*	
1	East Gojjam	29	5	100
2	Awi	13	2	40
3	Bahir Dar	12	2	40
4	South Gonder	32	6	120
5	South Wollo	42	7	140
6	West Gojjam	31	6	120
7	Oromiya	7	1	20
8	North Gonder	47	8	160
9	North Showa	36	6	120
10	North Wollo	31	6	120
11	Waghimera	8	1	20

Data Gathering Tools

A questionnaire is used to collect data from the respondents. The questionnaire has two parts: background information and the Teacher Assessment Literacy questionnaire.

The Background Information is composed of questions (items) asking about the participants' sex, qualification, subject taught, years of experience, educational level, and course status on educational measurement and evaluation (whether a teacher has or has not taken a course in measurement and evaluation during his/her pre-service training).

Teacher Assessment Literacy Questionnaire is composed of 31 multiple-choice items with four response alternatives. This questionnaire is adapted from Plake, Impara and Fager's (1993) Teacher Assessment Literacy Questionnaire. Plake, Impara and Fager

(1993) in cooperation with the National Council on Measurement in Education and the W.K. Kellogg Foundation developed an instrument (Teacher Assessment Literacy Questionnaire) to measure teachers' knowledge in the seven competency areas. The instrument was developed based on the "Standards for Teacher Competence in the Educational Assessment of Students" developed by the collaborative effort between the American Federation of Teachers (AFT), the National Education Association (NEA), and the National Council on Measurement in Education (NCME) in 1990. These standards are incorporated in the Standard five, i.e., assess, provide feedback and report on student learning part of the professional standard for Ethiopian school teachers developed by the Ministry of Education in 2012 (MoE, 2012). The original questionnaire is composed of 35 multiple-choice items (five multiple-choice test questions for each of the seven competency areas) designed to assess teacher assessment literacy. Plake, Impara, and Fager (1993) found a reliability of .54 for the entire test using the KR–20 method. However, four of the 35 items (item numbers 14, 20, 28, and 31) were found to be irrelevant to the Ethiopian context because either they refer to the interpretation of standardized tests, rules, or community-based curriculum issues that are not relevant to the Ethiopian education system. Thus, the questionnaire used in this study has only 31 of the 35 multiple-choice questions, and its reliability for the entire 31-items test was computed using the Spearman-Brown formula and found to be .59. This coefficient is higher than the one reported by Plake, Impara, and Fager (1993) and Lake (2014).

Data Collection Procedures

Data was collected from the respondents using the questionnaire mentioned above. The questionnaire was presented in English and was distributed to each respondent in person in their school and was collected in the same fashion. During the delivery of the questionnaire, the objectives of the questionnaire and the study, and instructions on how to fill the questionnaire were clearly communicated. Also, respondents were informed that their responses would be confidential and would be used only for research purpose.

Data Analysis

First, the responses of the 31 multiple-choice questions were coded as 0 if it is incorrectly responded and 1 if it is correctly responded. The data were entered into a computer and analyzed using SPSS version 20 for windows following this coding. In this process, cases with missing values were discarded and were not considered in the analysis. Thus, the analysis was carried out only on valid data, i.e., 805, excluding missing cases.

Means and standard deviations were used to describe the results of each of the seven standards or level of competency areas and of the overall competence for the whole respondents. One sample t-test with the expected mean value was used to measure teachers' competence in student assessment. In addition to this, independent samples t- test was used to investigate whether there existed a statistically significant difference between male and female teachers in their level of educational assessment competence; and a comparison was made between those teachers who took and didn't take an independent course in measurement and evaluation in their pre-service training. One-way ANOVA was used to see whether there existed a statistically significant difference among teachers who taught different subjects in their levels of competence in the educational assessment of students.

Results

The purpose of this study was to assess the level of secondary school teachers' competence in the educational assessment of students in Amhara National Regional State. A modified form of the Teacher Assessment Literacy Questionnaire having 31 items with four-option multiple choice items was used. The analysis was performed only on valid data, ignoring the missing ones. The analysis yielded the following:

First, a summary of the socio-demographic characteristics of teachers who participated in the study is displayed. This is followed by the descriptive and inferential statistical values of the analysis for the level of competence of teachers in student assessment. A percentage of teachers who correctly answered each of the items is displayed on a table, and finally, the values for the comparison tests are presented.

The Socio-Demographic Characteristics of Respondents

The sample teachers are categorized by their sex, educational level, subjects they teach, year of experience, and on whether they have taken independent courses in measurement and evaluation during their training. Of the 805 teachers who completed the questionnaire, 89 percent are males, and the rest 11 percent are females. A dominant number of respondents (95.4 % of them) have a Bachelor's degree, a few (3.1%) have a Master's degree and the rest very few teachers (1.5%) have a diploma. Around 70% of the respondents have served at least 10 years, and the rest have taught for more than 10 years. Most of the respondents (71.8%) have taken a separate course on measurement and evaluation during their pre-service educational training. Table 2 below indicates the socio-demographic characteristics of the respondents.

Table 2: Socio-Demographic Characteristics of the Respondents

Respondent Characteristics	*n*	*%*
Sex		
Male, Female	716, 89	89, 11
Educational Level		
Diploma, Bachelor's Degree, Master's Degree	12, 764, 25	1.5, 95.4, 3.1
Subjects Taught		
Natural Science	398	50.4
Social Science	171	21.7
Language	162	20.5
Others (IT, Economics, PHE,)	68	8.35
Experience in Years of Service		
1-5 Years	213	26.2
6-10 Years	355	43.6
11-15 Years	91	11.2
16-20 Years	50	6.1
21-25 Years	48	5.9
26-30 Years	30	3.7
>30 Years	27	3.3
Taken Course on Educational Measurement and Evaluation		
Yes (No)	572 (225)	71.8 (28.2)

Descriptive Statistics for Different Categories of Teachers

One purpose of this research is to make comparisons between groups of teachers. Thus, the means and standard deviations of categories of respondents is in Table 3 below. As displayed in the Table, the mean score of males is a greater than the mean score of females, and the mean score of those teachers who have taken a course on measurement and evaluation is greater than those who have not taken. The difference in the mean score of teachers who have and have not taken a course on measurement and evaluation is greater than the difference in the mean score of male and female teachers.

Table 3: Means and Standard Deviations of Different Groups of Respondents

Group	n	M	SD
Male	628	11.01	3.22
Female	68	10.43	3.88
Taken Course	498	11.32	3.29
Didn't Take Course	194	10.04	3.20

It is generally believed that teaching experience matters. Hence, teachers are classified based on their years of experience in teaching and compared to see if the year of service creates a difference in their competence of student assessment. As presented in Table 4 below, competence in assessment increases as years of service increases and then it decreases.

Table 4: Means and Standard Deviations Respondents by Years of Service

Year of Service	*n*	*M*	*SD*
1 – 5 years	188	10.38	3.10
6 – 10 years	303	11.19	3.45
11 – 15 years	78	11.72	3.30
16 – 20 years	45	11.24	3.47
> 20 years	91	10.57	2.83
Total	705	10.95	3.29

Teachers' Competence in the Educational Assessment of Students

As displayed in Table 5 below, the mean performance of teachers in the overall or specific competency areas is poor, and the standard deviations are small indicating similarities among teachers in answering the items. Overall, the mean performance on the 31-item instrument is 10.95 (with a standard deviation [SD] – 3.29) or 35 % correct. Across the seven competency areas, teachers demonstrate that the relative highest level of competency in the area of Choosing

Assessment Methods (M = 2.05) and the lowest level of competency in the area of Communicating Assessment Results (M = .87, SD = .78). To assess the level of competence one sample t-test is computed for the overall performance and for each of the standards. All mean values are statistically significantly below the respective expected means. Teacher performance across the seven competency areas is summarized in Table 5 below.

Table 5: Means and Standard Deviations by Standard and Total Possible Scores on Teacher Assessment Literacy Questionnaire

Competency Area	M	SD	t	Total Possible
Choosing Assessment Methods	2.03	1.00	-13.28	5
Developing Assessment Methods	1.97	1.05	-14.12	5
Administering, Scoring, and Interpreting Assessments	1.73	1.21	-6.30	4
Using Assessments for Decision Making	1.49	0.95	-14.99	4
Using Assessments for Grading	1.73	1.06	-20.35	5
Communicating Assessment Results	0.87	0.78	-44.04	4
Recognizing Unethical Practices	1.14	0.91	-26.45	4
Mean of Total Item Score	10.95	3.29	-36.69	31

All t values are significant at P < 0.001

To better understand the competence of teachers, Table 6 presents the proportion of teachers who correctly answered each of the items within each competency area. It shows the difficulty level of items (the percentage of teachers who correctly answered each of the questions).

The sample teachers answered correctly an average of slightly less than eleven items out of 31 items. A review of Table 4 below reveals that six of the 31items are very difficult for these teachers (items with the percentage correct less than 25) and other seven items are also difficult (items with the percentage correct less than 30). Fifty-five percent or more of the teachers answered five of the 31 items correctly.

One item each came from Standard 2-Developing Appropriate Assessment Methods, Standard 3-Administering, Scoring, and Interpreting Assessments and Standard 4-Using Assessments for Decision Making; two items came from Standard 1- Choosing Assessment Methods.

On six of the 31 items, fewer than 25% teachers answered the item correctly. One item each came from Standard 1-Choosing Assessment Methods, Standard 2-Developing Assessment Methods, and Standard 7-Recognizing Unethical Practices; three items came from Standard 6-Communicating Assessment Results. Sixty- nine percent of the respondents (between 60% and 80%) correctly answered only a single item and even those items of which 50% or above got correct were only five.

Table 6: Percentage of the Sample Teachers Who Answered Each of the Items Correctly

Competency Area	Item	Percentage
Choosing Assessment Methods	1	69
	2	21
	3	27
	4	55
	5	31
Developing Assessment Methods	6	57
	7	41
	8	41
	9	43
	10	14
Administering, Scoring, and Interpreting Assessments	11	43
	12	26
	13	59
	14	43
Using Assessments for Decision Making	15	59
	16	32
	17	31
	18	28
Using Assessments for Grading	19	28
	20	36
	21	42
	22	29
	23	37
Communicating Assessment Results	24	20
	25	24
	26	33
	27	9
Recognizing Unethical Practices	28	27
	29	24
	30	28
	31	34

Comparison of Teachers Assessment Literacy by Sex, Course Status, and Subject Area

The mean score for those who took an educational measurement and evaluation course (n = 498) was 11.32 and for those who didn't take (194) was 10.04. Moreover, as displayed in Table 7 below, the independent samples t-test revealed that there is a statistically significant difference between those who have taken a separate course in measurement and evaluation and those who haven't taken the course, $t = 4.68$, $P < 0.001$.

Table 7: Independent Samples T-Test between Males and Females, and Between Teachers Who Took and Didn't Take Course in Assessment

Source	Mean Difference	t	df	P
Sex	0.59	1.40	694	0.16
Course	1.28	4.63	690	0.000

An independent t-test was computed to see whether there existed a statistically significant difference between male and female teacher. However, the result was not significant. In addition to this, repeated measures ANOVA was used to see differences in the performance of teachers among the seven standards. First, the raw scores were transformed into z-scores and then one-way repeated ANOVA was computed. However, no statistically significant difference was found among the standards. Finally, the one-way ANOVA result revealed that there was no statistically significant difference among teachers who taught different subjects. However, the mean score of language teachers was relatively greater than the mean score of teachers who taught other subjects.

To see whether the years of service affect teachers' competence in student assessment, one- way ANOVA (Table 8) was used followed by multiple comparisons as it was significant.

Table 8: ANOVA taking Years of Service as an Independent Variable

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	141.692	4	35.423	3.316	.011
Within Groups	7476.855	700	10.681		
Total	7618.548	704			

As shown on Table 9, all the differences between each interval were found to be significantly different from the following or the preceding one. From this it could be concluded that teachers' assessment literacy increases as their years of service in teaching increase.

Table 9: Post Hoc ANOVA: Multiple Comparisons by Years of Experience

(I) Experience	(J) Experience	Mean Difference (I-J)	Std. Error	Sig.
1-5	6-10	-.81046*	.30342	.008
	11-15	-1.34029*	.44017	.002
11-15	>20	-1.14652*	.50430	.023

Discussion

The purpose of this study was to assess secondary school teachers' competence in the educational assessment of students in Amhara National Regional State. Many of the results of this study were parallel to those findings of earlier studies that used the original version of the instrument (Plake, 1993; Plake, Impara and Fager, 1993; Martler, 2003). However, teachers in the current study scored smaller than those in the earlier studies. The results of the analyzed data are indicated below:

With respect to the overall performance on the 31 items, the average score of the sample secondary school teachers' competence in the educational assessment of students was 10.95 (35 % correct). In other words, participant teachers answered on average nearly eleven out of 31items correctly. This score is quite smaller than the earlier findings (Plake, 1993; Martler, 2003). Plake (1993) and Martler (2003) reported that the average scores of the overall performance on the 35 items were equal to 23 and 22 respectively. Plake, Impara, and Fager (1993) conducted a national assessment of teachers competence in the educational assessment of students all over the United States of America on 555 teachers from elementary, middle, and high school levels and they found nearly 66 % correct which is closer to twice the competence level found in this study. The findings of the current study confirmed that teachers also lacked the competence in each of the seven competency areas.

Moreover, given that the Education and Training Policy of the Federal Democratic Republic of Ethiopia (1994) set 50% achievement as a minimum score in order to be promoted from one level to the following, most teachers participating in this study would receive a failing grade based on their demonstrated knowledge of educational assessment of students. This average score of nearly eleven of thirty-one items indicated that teachers lack the necessary knowledge and skill in student assessment which might be attributed to the inadequacy of their pre-service teacher training in student assessment.

When examining the extent to which teachers met the seven standards, the results of this study are similar to earlier findings (Mertler, 2005; Plake and Impara, 1997). In the present study, teachers had the most difficulty with Standard 6 (M = .87), communicating assessment results. Plake and Impara (1997) also found that in-service teachers scored the lowest (M = 2.70) on this standard. Likewise, in Mertler's (2005) study, teachers did not score high on this standard (M = 2.48). Although the findings are similar in that the mean score for the standard "communicating assessment results" is the lowest compared

to the mean scores for the remaining six standards, the mean score in the present study is far below than the means in the two earlier studies. The first standard was about choosing assessment methods, and failure in this standard demonstrates that teachers had difficulties in connecting assessment to clear purposes since choosing assessment very much depended on the purpose of assessment. As they indicated in the interview, they frequently used group work, quizzes, and tests; no other assessment procedures. They were however expected to choose from a dozen of assessment procedures. Their failure in the second standard showed that teachers lacked the skill to develop assessment methods and it depended on the failure in the first standard.

The proportion or percentile of teachers who correctly answered an item shows the difficulty or easiness of the item usually called the difficulty level of the item. Measurement and evaluation experts recommend that the average level of difficulty for a four-option multiple choice item should be between 60% and 80%. Moreover, in four-alternative multiple choice items, a test taker has the probability of .25 or 25% of getting an item correct by chance. Hence a proportion of less than .25 indicates that either the item is very difficult or it has ambiguity and that the test takers failed to understand it.

In the current study, the instrument used is a standardized one and hence ambiguity is a less probable explanation for the poor performance of the teacher respondents. Consequently, the scores on six items with the level of difficulty less than .25 in the current study are attributable to the poor competence of secondary school teachers in the Amhara National Regional State.

With respect to teachers' preparation in assessment, around 72% of teachers responding to the study reported that they had taken a separate measurement and evaluation course. Moreover, those who had taken a measurement and evaluation coursework scored statistically significantly higher on teacher assessment literacy

questionnaire than those who hadn't, with the mean achievement difference more than one point (1.28). This finding is similar to Plake's (1993) who reported a similar difference between those who have a previous training and who didn't have but he indicated that the difference was less than one point.

Although teachers who have taken a course in measurement and evaluation demonstrated a better competence than those who didn't, their mean achievement was still very low (only 11.32 or 36.5 % correct). It is common sense to assert that a single three-credit hour course is not enough to equip teachers with the necessary knowledge and skill required to implement educational assessment. Further, as many of these teachers took the course during the summer program as part of an in-service upgrading, the course offered was limited to a theoretical explanation of limited chapters such as the principles of test construction, administration, and scoring. The short instructional time made it difficult to cover chapters about test statistics, validity and reliability, item analysis, and communicating and reporting results, let alone to include other forms or procedures of assessment relevant to the practice.

Further, the pre-service teacher training program also lacks emphasis in assessment. It is common to offer a course merely on test construction without addressing topics on authentic and continuous assessment, and other important and current topics. Moreover, the delivery system is not practice oriented; students do not practice in developing and validating varieties of assessment tools. That is why teachers are observed to frequently use the traditional assessment tools such as quizzes and tests to assess students. Though these tools are useful, they assess whether students remember content knowledge, and not whether they have become effective learners. The current situation requires students to be problem solvers, they should not only master subject matter to perform well in school tests, they should also analyze, evaluate and even create knowledge. This calls for an assessment literate teacher, one who does not only understand

why assessments are important but who is also skilled in assessing students and in evaluating the curriculum. Such a teacher believes in passing on assessment literacy to students, developing in them the habit of critically thinking about what they learn and the way they learn.

The current result in any way shows that the training in pre-service teacher training program about student assessment was not adequate. In this respect, Gullickson (as cited in Sanders and Vogel, 1993) noted that colleges often provide some instruction in measurement and evaluation, but the time devoted to such instruction is limited. It is also true that no long-term continuous professional development is prepared on student assessment.

Teachers indicated that they frequently use tests and have difficulties in developing other assessment procedures. Moreover, they confirmed that they develop test items only from textbooks. Teachers were hesitant to complete and request additional time to return the questionnaire. When asked their reasons, they indicated that they failed to comprehend the questions and feared that the researchers may ridicule them for being assessment illiterate. Although assessing student performance is one of the most critical responsibilities of classroom teachers, yet, those who participated in this study do not feel adequately prepared for this task. They believe that they need assistance to apply assessment concepts and techniques.

Generally, the findings pinpoint that secondary school teachers in Amhara National Regional State do not have adequate knowledge in classroom assessment which could negatively constrain the quality of education in the region. Unless assessment literacy becomes a pivotal content area of professional development programs, the problem may worsen the already poor quality of education in the Region.

Conclusion

Secondary school teachers in Amhara National Regional State are not competent in the seven standards for teacher competence in assessment of students. Hence, the findings suggest that teachers have limitations in choosing, developing, administering, using and communicating appropriate classroom assessment. Consequently, teachers are in need of professional assistance in educational assessment methods.

Recommendations

Based on the findings of this study, the following are recommended to improve the overall competence of teachers in student assessment:

- The Amhara National Regional State needs to have a licensing and certification requirement of assessment literacy in line with the national professional standard of school teachers. Unless teachers attain specific competence standards in educational assessment, it will not be easy to improve student achievement.
- The Education Bureau of the Amhara Region should arrange an in-service training on educational assessment of students to teachers so that teachers might improve their competence.
- The continuous professional development activities undertaken in schools should give much emphasis to the educational assessment of students.
- Teacher training institutions need to revise their curriculum and provide more practical courses to teacher trainees on educational assessment in their teacher preparation program.
- Further research should be conducted to assess the depth and breadth of the problem at the national level including teachers at the elementary level to change/revise the curriculum of teacher training institutions.

References

American Federation of Teachers, National Council on Measurement in Education, and National Education Association (AFT, NCME, and NEA). (1990). *Standards for teacher competence in educational assessment of students*. Washington, DC: Author.

Assessment Reform Group (1999). *Assessment for learning: beyond the black box.* Cambridge: University of Cambridge School of Education.

Black, P., and William, D. (1998). Assessment and classroom learning. *Assessment in Education, 5*(1), 7–75.

Campbell, C., Murphy, A., and Holt, K. (2002). *Psychometric analysis of an assessment literacy instrument: Applicability to preservice teachers*. Paper presented at the annual meeting of the Mid-Western Educational Research Association, Columbus, OH.

Chen, P. P. (2005). Teacher candidates' literacy in assessment. *Academic Exchange Quarterly*.

Darling-Hammond, L., Wise, A. E. and Pease, S. R. (1983). Teacher evaluation in the organizational context: a review of the literature. *Review of educational research*. 53, 285-237.

Ebel, R. L. (1962). Measurement and the teacher. *Educational Leadership*, 20, 20-24.

FDRE. (1994). *Education and Training Policy of Federal Democratic Republic of Ethiopia.* Addis Ababa: Ministry of Education.

Greenstein, L. (2010). *What teachers really need to know about formative assessment*. Alexandria: Association for Supervision and Curriculum Development.

Khan, B. (2012). Relationship between assessment and students' learning. *International Journal of Social Sciences and Education*. 2(1), 576-588

Kiomrs, R., Abdolmehdi, R., and Naser, R. (2011). On the Interaction of Test Washback and Teacher Assessment Literacy: The Case of Iranian EFL Secondary School Teachers. *English Language Teaching*, 4 (1).

Lake, B. (2014). Secondary School Teachers' Competence in Educational Assessment of Students in Bahir Dar Town. *Bahir Dar Journal of Education*, 14(2), 54-63.

Marzano, R. J. (2006). *Classroom assessment and grading that work*. Alexandria: Association for Supervision and Curriculum Development.

Mertler, C. A. (2003) *Preservice versus inservice teachers' assessment literacy: Does classroom experience make a difference?* Paper presented at the annual meeting of the Mid-Western Educational Research Association, Columbus, OH (October).

Mertler, C. A. (2005). Secondary teachers' assessment literacy: Does classroom experience make a difference? *American Secondary Education*, 33, 76-92.

Mertler, C. A. (2009). Teachers' assessment knowledge and their perceptions of the impact of classroom assessment professional development. *Improving Schools*, 12(2), 101–113

MoE. (2012). *Professional standard for Ethiopian school teachers*. Addis Ababa.

Ojerinde, D. (2009). *Using assessment for the improvement of tertiary education in Nigeria: The Joint Admissions and Matriculation Board (JAMB) role*. Paper presented at the 35th International Association for Educational Assessment Conference, Brisbane, Australia.

Phye, G. D. (1997). Classroom assessment: A multidimensional perspective. In G. D. Phye (ed.), *Handbook of classroom assessment: Learning, adjustment, and achievement* (pp. 33-51). San Diego: Academic Press.

Plake, B. S., and Impara, J. C. (1997). Teacher assessment literacy: What do teachers know about assessment? In G. D. Phye (Ed.), *Handbook of classroom assessment: Learning, adjustment, and achievement* (pp. 53-68). San Diego: Academic Press.

Plake, B. S., Impara, J. C., and Fager, J. J. (1993). Assessment competencies of teachers: A national survey. *Educational Measurement: Issues and Practice*, *12*(4), 10-12, 39.

Quitter, S. M. (1999). Assessment literacy for teachers: Making a case for the study of test validity. *The Teacher Educator*, 34: 4, 235-243

Sanders, J. R., and Vogel, S. R. (1993). The Development of Standards for Teacher Competence in Educational Assessment of Students. Teacher Training in Measurement and Assessment Skills. Retrieved from http://digitalcommons.unl.edu/burosteachertraining/5

Popham, W. J. (2009). Assessment literacy for teachers: Faddish or fundamental? *Theory into Practice*, 48:4–11

Starch, W. D., and Elliot, E. (1912). Reliability of grading high school work in English. *School Review*, 20, 442-457.

Stiggins, R. J. (1995). Assessment literacy for the 21st century. *Phi Delta Kappan*, *77*(3), 238-245.

Stiggins, R. J. (1999). Are you assessment literate? *The High School Journal*, 6(5), 20-23.