Tabaran Institute of Higher Education
International Journal of Language Testing
Vol. 12, No. 1, March 2022
ISSN 2476-5880

Examining Teacher Assessment Literacy and Instructional Improvement of Iranian High School Teachers on Various Fields of Study

Saeed Ahmadi[1], Sadegh Ghaffary[2], *Mohsen Shafaghi[3]

Received: October 2021 Accepted: December 2021

Abstract

Assessment literacy refers to the knowledge teachers have about definitions of testing and the employment of this knowledge to classroom practices. This study inspected the extent to which high school teachers of English, mathematics, science, and social studies utilize assessment literacy and instructional improvement practices in their classrooms and the probable relationship between these two factors. To this end, ninety-seven teachers from sixty high schools across Zanjan province in Iran were selected based on convenient sampling. The instrument used was a 32-item questionnaire designed by researchers and validated by experts. A confirmatory factor analysis (CFA) was also conducted in order to validate the factor structure of the questionnaire. The results showed a strong positive relationship between assessment literacy and instructional improvement practices in all four subject areas through correlations. This suggests that showing an understanding of assessment literacy results in a greater understanding of instructional improvement in the classroom. Moreover, a one-way ANOVA was performed in order to determine if there is a significant difference in the use of assessment literacy practices and instructional improvement among different subject areas. The data revealed no statistically significant differences between neither assessment literacy nor instructional improvement scores by subject area. The study can be beneficial to school districts and teachers since the review of existing teaching practices and an increase in ongoing professional development in assessment literacy theories and practices may improve instructional improvement in all subject areas.

Keywords: Assessment literacy; Instructional improvement; Correlation; One-way ANOVA; Confirmatory Factor Analysis (CFA)

[1] Assistant professor in TEFL, English Language Department, Kangan Branch, Islamic Azad University, Kangan, Iran. Email: ahmadi_efl@yahoo.com
[2] Ph.D. candidate in TEFL, English Language Department, Urmia University, Urmia, Iran. Email: sadeghghaffari@gmail.com
*[3] Ph.D. in TEFL, English Language Department, Allameh Tabataba'i University, Tehran, Iran. Email: mohsenshafaghi@gmail.com

1. Introduction

According to Popham (2006), the only reason for educational assessments to exist is to "make more defensible educational decisions regarding… students" (p. 1). Stiggins (2008) states "We assess for two reasons: (1) to gather evidence to inform instructional decisions; (2) to encourage students to try to learn. Both purposes must be well served for schools to be effective" (p. 3). Both researchers agree that assessments only guide decisions *by* educators-they do not make decisions *for* educators. However, as Popham (2006) points out, assessments only provide data. Educators need to make inferences about student learning from the data itself. These inferences and the decisions that stem from them are vital parts of the assessment process. Unfortunately, assessment concepts seem to be poorly understood by many practicing educators, affecting their ability to effectively teach and support curricula.

Stiggins (2008), Popham (2006), and Brookhart (2011) all agree that fundamentally, assessment is more than just tests and grades. It is a process by which teachers design specific methods for collecting data about student learning, analyzing that data, then adjusting instruction that will be most effective for student learning based upon that data.

If assessment is a process by which teachers collect data about the success of student learning of a curriculum, then the process must be tightly aligned and integrated within the instructional framework itself. Stiggins (2008) asserts that assessments in the classroom, school, and institutional level only serve their purpose if the assessment is intentionally designed to illustrate how well students are mastering a standard. One study (Herman, Webb, & Zuniga, 2007) found that if instructors are to agree on the inferences from assessment data, then alignment of test items with curricular standards is paramount. Standards provide the common language that ties the data to learning, and without alignment there is no definitive way to interpret the information meaningfully.

Similarly, Stiggins (2008) stresses that teachers cannot accommodate the needs of students in their classroom through instruction without being literate in the language of assessment. It is the alignment of curriculum, standards, assessments, and instruction through practicing assessment literacy that provides the confidence and competence required for effective instruction. Therefore, this study will examine how teachers understand and apply assessment literacy in the classroom and how those practices influence their decisions about curriculum and instruction.

Even though the literature suggests assessment literacy is a vital skill needed to effectively teach students (Black & Wiliam, 1998, 2006; Sadler, 1989; Stiggins, 2010), the high school teacher's use of assessment literacy as an effective process of assessment used to decide about improving curriculum and instruction is not known. Additionally, the extent to which assessment literacy relates to improving student learning at the high school level is not known. According to Black and Wiliam (1998), to have a long-time positive effect on instruction, teachers must have a direct alignment between the design of their assessments with the design of their instructional strategies.

This study will add to the current literature examining if the understanding of assessment literacy is a viable component in decision making for improving teaching and learning. In this study, assessment literacy, as set out by Popham (2009) and Stiggins (2002), will be defined and measured through constructs of collaboration, design, method, data collection, and analysis. Similarly, the core constructs of instructional improvement–professional development opportunities for teachers to help them improve their performance in the classroom-are defined through data-based decision-making, standards-aligned curriculum, learner-centered instruction, and understanding both formative and summative assessments.

2. Review of Literature

2.1 Assessment Literacy

The literature in assessment literacy is ample with studies about EFL teachers (Buchanan, 2000; Estaji & Ghiasvand, 2021; Giraldo & Murcia, 2019; Harding & Kremmel, 2016; Velan, Rakesh, Mark, & Wakefield, 2002; Vogt & Tsagari, 2014; Xie & Tan, 2019). In a quite recent study, in order to bring modifications for teacher education reforms, Watmani, Asadollahfam, and Behin (2020) explored the EFL teachers' assessment literacy with a total of two hundred EFL teachers and students. The findings showed variations among EFL teachers with an English teaching background and those without such background in terms of their assessment literacy and perceptions of Assessment Literacy components.

In another study, Nurdiana (2020) conducted research to explore language teacher assessment literacy and the way it has been measured. The findings of the study supported the need to train language teachers on language assessment because they lacked sufficient knowledge of language assessment. Moreover, in spite of the fact that some instructors were literate in assessment, they did not implement the knowledge in their classroom; therefore, they needed training on selecting proper assessments for their students, designing tests, and alternative assessments.

Moreover, Chan and Luo (2020) conducted a research study to discover whether teachers know how to assess students' written reflections. The study examined the understanding of six novice university teachers about the assessment of reflective writings from three aspects. The results of the study showed teachers' lack of knowledge to perform a reasonable assessment of reflection in social and practical aspects, despite their strong comprehension of assessing reflection in conceptual aspect.

2.2 Instructional Improvement

Instructional improvement is a term in literature to describe professional development opportunities for teachers aimed at helping them to improve their performance in the classroom (Walter, 2006). Instructional improvement through assessment literacy is not unique to any one subject area. Several studies have looked at the impact of assessments within individual subject areas. Malone (2013) notes that researchers in the areas of English as a Foreign Language (EFL) tended to give more summative tests rather than formative ones owing to a lack of assessment

ISSN 2476-5880

literacy and its uses in the classroom. Pearson, Knight, Cannady, Henderson, and McNeill (2015) found that assessment literacy was crucial when designing effective science assessments incorporating reading and writing elements from the English/Language Arts disciplines, but even with a solid grasp of assessment literacy, developing effective formative assessments was difficult and time-consuming.

Salkind (2011) notes that math teachers need solid assessment literacy if they are to understand assessment both as a tool and as a process in effective teaching. Even across different subject areas, commonalities exist showing that understanding assessment literacy is crucial to instruction (Alemi, Miri, & Mozafarnezhad, 2019; Campbell & Collins, 2007; DeLuca, 2012; Fan, Wang, & Wang, 2011; Greenberg & Walsh, 2012; Hill, Ell, Grudnoff, & Limbrick, 2014; Lam, 2015; Leahy & Wiliam, 2012; Schneider & Randel, 2010; Smith, 2011).

The quantitative research studies in assessment literacy linked to instructional improvement (Buchanan, 2000; Velan et al., 2002) all suffered from methodological issues involving small self-selected populations and an inability to extrapolate the results to a more generalized learning population.

Studies about the importance of assessment literacy for teachers and its impact on instruction are replete with anecdotes and stories. However, data connecting teachers' use of assessment concepts and the instructional decisions they make have not advanced much since Black and Wiliam's (1998) call for more research to study the relationships between teachers' understanding of assessment and their roles in education.

In a similar vein, Gotch and French (2014) called for further research for developing an efficient and reliable instrument to measure teachers' assessment literacy of contemporary demands. As a result, this study aims to help fill that gap by investigating the relationship between teachers' assessment literacy practices and how those practices relate to instruction in the classroom. That is to say, the purpose of this correlational research study is to investigate high school English, math, social studies, and science teachers' practices of assessment literacy using constructs of collaboration, assessment design, item design, data collection methods, and data analysis.

The current study will compare those practices with the information teachers use to make instructional decisions, the extent to which they practice using a standards-aligned curriculum, their use of learner-centered instruction, and their use of formative and summative assessment in the classroom. As reviewed in the above studies, there is a scarcity of research in the teaching context of Iran considering high school teachers' use of assessment literacy as a process of effective assessment for making decisions about improving curriculum and instruction. Therefore, the present study aims at filling this gap by addressing the following research questions:

 ISSN 2476-5880

1. Is there a relationship between assessment literacy and instructional improvement among high school teachers of English, math, science, and social studies?
2. Is there any significant difference among teachers of English, math, science, and social studies in their use of assessment literacy practices and instructional improvement practices?

3. Method

3.1 Participants and Setting

The participants in the survey were a convenience sample of public high school teachers of English, math, science, and social studies across 14 educational districts encompassing 60 high schools, comprising 97 teachers in Zanjan, Iran. All of these teachers had a minimum of a bachelor's degree from an accredited university. Their teaching experiences ranged from 3 to 19 years. They were both male and female. Among the 97 respondents who participated in the survey, 29 primarily identified as English teachers, 25 as math, 23 as science, and 20 as social studies. The participants took part in the study voluntarily and were assured of the anonymity and confidentiality of the data and research findings. The teachers were asked to fill out a 32-item questionnaire which was shared online on WhatsApp. The number of participants in each of the four subject areas is presented in Table 1.

Table 1
Number of Respondents According to Instructional Area

Instructional Area	*n*
English	29
Math	25
Science	23
Social Studies	20
Total	97

The survey did not collect personally identifiable information on each participant other than demographic information collected in the first section of the study. Individual names, phone numbers, and email addresses were not collected. Specific instructions in the first section of the study asked participants to ensure they are only completing the survey once as to not adversely affect the data. The technical capabilities of the survey software also ensured that teachers can not submit data more than once.

3.2 Instrumentation

The instrument used in this research was developed in three phases. In the first phase, the research focused on creating the definitions of the instrument to ensure content validity in line with the research. During the second phase, the instrument items and scale scores were developed in line

 ISSN 2476-5880

with the conceptual framework of the study. In the third phase, a review of the instrument was sought from assessment experts to ensure the validity of the instrument.

The instrument (see Appendix) is a questionnaire comprising of three sections including 6 demographic questions and 32 Likert-type statements for a total of 38 items, as seen in Table 2.

Table 2

Survey Questions and Questionnaire Format

Sections	Factor	Survey items
Section 1	Informed Consent	
Section 2	Demographic Information	
	District of instruction	1
	Subject area of instruction	2
	Educational level	3
	Years of teaching experience	4
	Amount of professional development in assessment	5
	Amount of professional development in instructional improvement	6
Section 3	Likert-scale statements	
Section 3.1	Assessment literacy practices	
	Collaboration	7, 8, 9, 10
	Assessment design	11, 12, 13, 14
	Item design and data collection methods	15, 16, 17, 18
	Data analysis	19, 20, 21, 22
Section 3.2	Instructional improvement practices	
	Decision-making	23,24,25,26
	Standards-aligned curriculum	27, 28, 29, 30
	Learner-centered instruction	31, 32, 33, 34
	Formative and summative assessments	35, 36, 37, 38

Table 2 explains the three major sections of the survey. The first section contains the informed consent statement and approval. The second section is a set of six questions addressing the demographics of the respondents with questions about the district in which they teach, the subject area which they teach, their education level, and length of time teaching. The third section is a set of 32 five-point Likert-scale statements divided into two series. The first series of 16 statements focuses on ascertaining teachers' understanding of assessment literacy.

This set is organized into four subsections ascertaining respondents' practices with assessment literacy concepts including collaboration, assessment design, item design, and data collection methods, and data analysis. The second series of 16 statements focuses on ascertaining

 ISSN 2476-5880

teachers' understanding of instructional methods. This series is organized into four subsections ascertaining information about respondents' practices of instructional improvement including instructional decision-making, standards-aligned curriculum, learner-centered instruction, and formative and summative assessment practices. Each statement in the second section was answered using a Likert- scale ranging from 1-5.

3.2.1 Reliability Analysis

Reliability for the survey was measured using a Cronbach's alpha. Cronbach's alpha is appropriate to use when the analysis wants to determine whether the items are consistent with one another when representing a single construct (Salkind, 2011). The goal for reliability in this study is to ensure a value of $\alpha \geq .7$.

The independent variable of "Assessment Literacy" was divided into four constructs ascertaining respondents' practices of the components of assessment literacy: Data analysis, Item design and data collection methods, Assessment design, and Collaboration. Each construct was measured using four survey items. The reliability of the variable, as well as each construct, was measured using a Cronbach's alpha (see Table 3).

Table 3
Reliability of Assessment Literacy Constructs

Construct	Cronbac s' α	*n* of items	Questionnaire items
Collaboration	.87	4	7, 8, 9, 10
Assessment Design	.71	4	11, 12, 13, 14
Item Design and Data Collection Methods	-.75	4	15, 16, 17, 18
Data Analysis	.91	4	19, 20, 21, 22
Cumulative Assessment Literacy	.77	16	7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22
Adjusted Cumulative Assessment Literacy	.85	12	7, 8, 9, 10, 11, 12, 13, 14, 19, 20, 21, 22

"Cumulative Assessment Literacy" construct had a high level of internal consistency, as determined by a Cronbach's alpha of .77. Likewise, the items of Collaboration, Assessment Design, Item Design and Data Collection Methods, and Data Analysis all had alpha scores above the threshold of .7, indicating a high level of internal consistency within those constructs.

ISSN 2476-5880

To ensure internal reliability of the rest of the assessment literacy elements of the survey, Cronbach's alpha scores were rerun with all the items. Accordingly, a high level of internal consistency was determined by a Cronbach's alpha of .85 reported in Table 3 as "Adjusted Cumulative Assessment Literacy." This adjusted variable will be referred to as "Assessment Literacy" in the remainder of this study.

The dependent variable of "Instructional Improvement" was divided into four constructs ascertaining respondents' practices of instructional improvement: Decision-Making, Standards-Aligned Curriculum, Learner-Centered Instruction, and Formative and Summative Assessments. Each of those constructs was measured using four survey items. The reliability of the entire variable, as well as each construct, was measured using a Cronbach's alpha (see Table 4).

Table 4
Reliability of Instructional Improvement Constructs

Construct	Cronbach's α	*n* of items	Survey items
Decision-Making	.76	4	23, 24, 25, 26
Standards-Aligned Curriculum	.93	4	27, 28, 29, 30
Learner-Centered Instruction	. 71	4	31, 32, 33, 34
Formative and Summative assessments	.73	4	35, 36, 37, 38
Cumulative instructional improvement	.83	16	23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38
Adjusted cumulative instructional improvement	.85	8	27, 28, 29, 30, 35, 36, 37, 38

The "Cumulative Instructional Improvement" construct had a high level of internal consistency, as specified by a Cronbach's alpha of .83. Similarly, the items for Decision-Making, Standards-Aligned Curriculum, Learner-Centered Instruction, and Formative and Summative Assessments had alpha scores above the threshold of .70, indicating a high level of internal consistency within those constructs.

To ensure internal reliability of the rest of the assessment literacy elements of the survey, Cronbach's alpha scores were rerun with all the items. Therefore, a high level of internal consistency was specified by a Cronbach's alpha of .85 which is reported in Table 4 as "Adjusted Cumulative Instructional Improvement." This adjusted variable will be referred to as "Instructional Improvement" in the remainder of this study.

ISSN 2476-5880

3.2.2 Validity Analysis

In this section, the validity analysis of the instrument is explained in the form of content validity and construct validity. It is worth noting that the validity of the survey was confirmed by six university professors in the fields of applied linguistics (N=3) and educational assessment (N=3). The survey ensured content validity by confirming the structure and wording of items through a review by the university professors in the field of educational assessment in order to ensure the survey items measure the intended domains. Creswell (2012) indicates that such evidence based on test content is a reliable method for ensuring instrument validity. Based upon the feedback received from the university professors, the wordings of two questions in the assessment literacy measure and three questions in the instructional improvement measure were modified.

In addition, we employed Waltz and Bausell's (1981) Content Validity Index (CVI) to evaluate the content validity of the questionnaire. The CVI was derived from the ratings of content experts (university professors) of the questionnaire items in terms of three criteria, namely, simplicity, relevancy, and clarity on a 4-point ordinal scale ranging from 1(not compatible) to 4(highly compatible). The CVI was obtained by adding the number of content experts rating the items as 3 or 4, divided by the total number of content experts (N=6). Next, the average of the CVI scores on simplicity, relevancy, and clarity across the experts was calculated and used as the total CVI. The obtained CVI scores of all the items except items 33 and 34 turned to be greater than the cut-off point of 0.79 indicating that all the items yielded satisfactory CVI values (Waltz & Bausell, 1981). However, items 33 and 34 obtaining a CVI score of 0.66 were also acceptable with regard to Lynn (1986) whose minimum required CVI values range from 0.60 to 0.69.

To examine the construct validity of the questionnaire, we performed confirmatory factor analysis (CFA) employing the Maximum Likelihood Estimation Method via Amos graphics 24. In CFA, the hypothesized scale is utilized to examine a priori the number of latent factors and their related indicators (Hair, Black, Babin, & Anderson, 2010). The schematic illustration of the questionnaire together with the standardized path coefficients between the two constructs and their respective sub-scales is depicted in Figure 1 in which CFA substantiated the two-factor questionnaire entailing assessment literacy practices and instructional improvement practices with their corresponding constructs.

ISSN 2476-5880

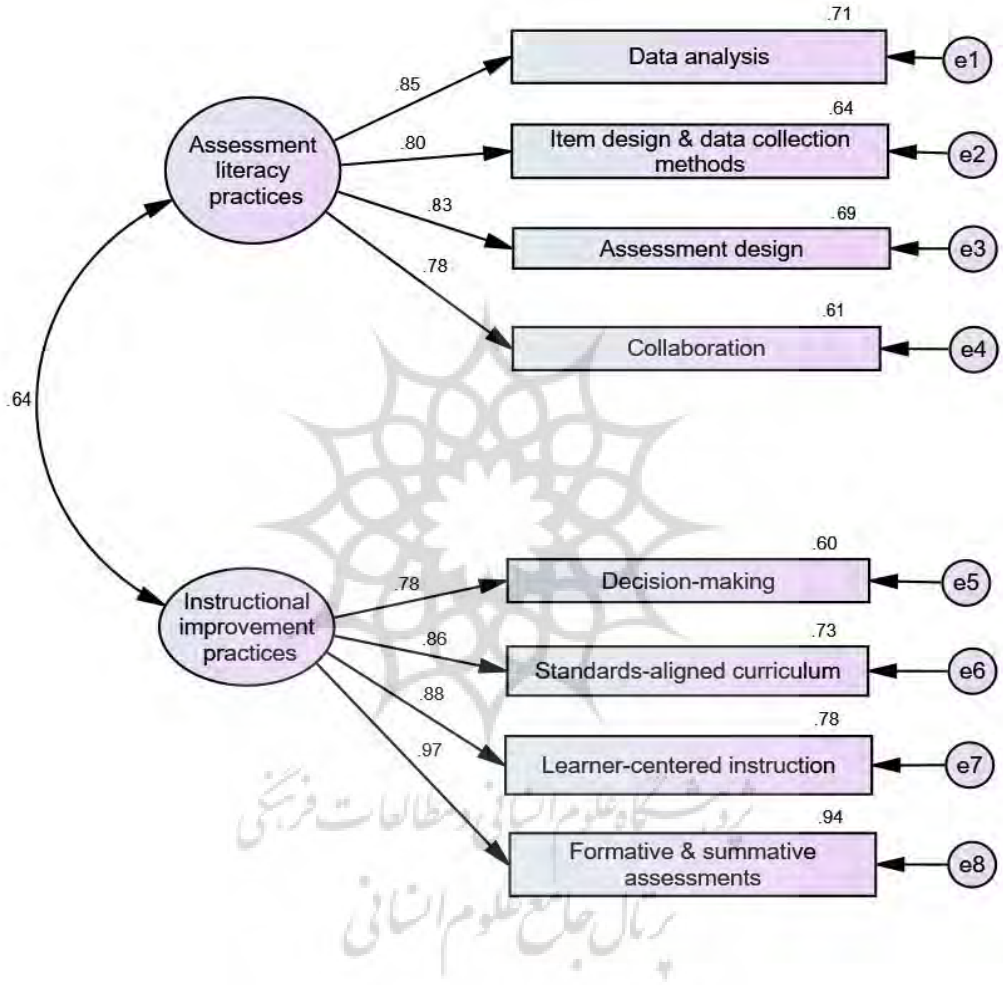

Figure 1: CFA of the Questionnaire in Standardized Estimates

As shown in Figure 1, the standardized estimates of the constructs were significant ($p<0.05$) and relatively strong. To validate whether the questionnaire well fit the data, the goodness-of-fit indices including CMIN /DF (chi-square fit statistics/degree of freedom), RMSEA (Root Mean Square Error of Approximation), IFI (Incremental Fit Index), TLI (Tucker-Lewis Index), CFI (Comparative Fit Index), and GFI (Goodness-of-Fit Index) were utilized. The minimum cut-off point for the questionnaire validation is less than 3 for CMIN /DF (Kline, 2011), and the threshold

ISSN 2476-5880

values for GFI and RMSEA are .9, and .08, respectively (Sharma, 1996). Also, IFI, TLI, and CFI all with minimum cut-off values of .9 (Byrene, 2010; Hair et al., 2010) were calculated to assess the questionnaire fit.

Table 5
Fit Indexes of the Questionnaire

Indexes	GFI	IFI	TLI	CFI	RMSEA	CMIN/ DF
Measurement model	0.940	0.960	0.942	0.960	0.049	**1.58**
Acceptable fit	>.90	>.90	>.90	>.90	< .08	**< 3**

As depicted in Table 5, all fit indices (CMIN /DF = 1.58; RMSEA=.049; GFI =0 .94; CFI =.0.96; IFI =0.96; TLI= 0.94) fell within the acceptable range. This indicates that the questionnaire can be considered as a valid scale for tapping assessment literacy practices and instructional improvement practices among Iranian EFL teachers. In other words, the fit indices confirmed the construct validity of the questionnaire of the study and the fit of the data to the CFA model. In light of the fit indices obtained, there was no need for respecification of the questionnaire using post-hoc modification indices suggested by the software.

3.3 Procedures

The process of developing and validating the questionnaire started with a standard and stepwise procedure. After going through the literature related to the four main constructs of assessment literacy (data analysis, item design and data collection methods, assessment design, and collaboration) as well as the four main constructs of instructional improvement (decision making, standards-aligned curriculum, learner-centered instruction, and formative and summative assessment), the outcomes derived from the literature were put together along with the discussions with the experts in the field regarding the constructs. The final result was a questionnaire consisting of 32 items related to all the eight fundamental constructs which were under the focus. Moreover, based upon the feedback received from the university professors, the wording of two questions in the assessment literacy measure and three questions in the instructional improvement measure were modified.

In order to effectively and efficiently collect data for this research, the questionnaire was administered to teachers online using Qualtrics survey software. The electronic link to the online questionnaire was distributed to teachers via WhatsApp. The authors pilot tested the pretest questionnaire between January 12, 2021, and January 30, 2021, using a sample of 30 teachers. They viewed pilot testing as an opportunity to establish the face validity of the research measures

 ISSN 2476-5880

and refine certain items. The solicitation electronic link, informing target population members about the study and encouraging them to participate, was distributed on February 22, 2021.

Data from the questionnaire was captured in a spreadsheet capable of being analyzed by the SPSS statistical software. SPSS was then used to test correlations and to define indicators for the relationships studied. The subject area data from Section 2 of the questionnaire was analyzed using SPSS as categorical data. Descriptive statistics were reported to compare the frequencies of each subject area. These constructs were used with the data from Section 3 in a one-way ANOVA test. The data from Section 3 of the questionnaire for each respondent was collected in SPSS as ordinal data on a scale from 1-5. The mean is reported for each construct as well as the variables to determine the rates in which teachers from each subject area practice the components of assessment literacy as well as the components of instructional improvement.

In order to explore the relationship between assessment literacy practices and instructional improvement practices, a composite score for each respondent's constructs was generated by calculating the mean from all items in each category. This composite score for each construct was reported on a continuous scale from 1-5 along with the measure of variability. Using this method, there are a total of two scores for Section 3 of the survey for each respondent; one for assessment literacy practices and one for instructional improvement practices.

Since the mean scores are reported on a continuous scale, a Pearson correlation is appropriate for this analysis (Salkind, 2011). Correlations were used to determine if a positive relationship exists between assessment literacy practices and instructional improvement practices for teachers in each subject area. A correlation coefficient was generated comparing the scores for each group of subject area teachers to determine if a positive relationship exists between assessment literacy and instructional improvement for each subject area. The data was analyzed in the context of subject area constructs from Section 2 in order to examine if teachers differ in their use of assessment literacy practices to improve instruction. A one-way ANOVA was performed in order to determine if the scores indicate a significant difference in the use of assessment literacy practices and instructional improvement between the different subject areas.

4. Results

The first research question asked, "Is there a relationship between assessment literacy and instructional improvement among high school teachers of English, math, science, and social studies?" Assessment Literacy is the independent variable and Instructional Improvement is the dependent variable in the study. The relationship between the two variables was measured using Pearson's Product correlations. Pearson's correlations were run between the Assessment Literacy and Instructional Improvement scores for all respondents and are reported for each subject area (see Table 6).

 ISSN 2476-5880

Table 6
Pearson Correlations between Instructional Improvement and Assessment Literacy by Subject Area

Instructional improvement	Assessment literacy			
	English	Math	Science	Social studies
English	.76***			
Math		.51*		
Science			.65***	
Social studies				.64*

Note. *=statistically significant at $p \leq .05$ level; ***= statistically significant at $p < .001$ level

For English teachers, there was a statistically significant, strong relationship between assessment literacy practices and instructional improvement practices among the teachers, r (46) = .76, p < .001, with practices in assessment literacy explaining 58.3% of the variation in instructional improvement practices.

For math teachers, there was a statistically significant, moderately strong relationship between assessment literacy practices and instructional improvement practices among the teachers, r (22) = .51, p = .011, with practices in assessment literacy explaining 26.2% of the variation in instructional improvement practices.

For science teachers, there was a statistically significant, moderately strong relationship between assessment literacy practices and instructional improvement practices among the teachers, r (42) = .65, p < .001, with practices in assessment literacy explaining 42.3% of the variation in instructional improvement practices.

For social studies teachers, there was a statistically significant, moderately strong relationship between assessment literacy practices and instructional improvement practices among the teachers, r (34) = .64, p < .001, with practices in assessment literacy explaining 40.7% of the variation in instructional improvement practices.

The second research question asked, "Is there a significant difference between high school teachers of English, math, science, and social studies in their use of assessment literacy practices and instructional improvement practices?" To answer this question, one-way ANOVA tests were conducted to determine if teachers' use of assessment literacy practices and instructional improvement practices were significantly different for respondents in each subject area.

4.1 Assessment Literacy Practices

A one-way ANOVA test was conducted to determine if teachers' use of assessment literacy practices was significantly different for respondents in each subject area. Levene's test for homogeneity of variances confirmed the homogeneity of variances (p = .255). Data are presented as ± standard deviation. Use of assessment literacy practices differed between English teachers (M = 3.81, SD = .59), math teachers (M = 3.49, SD = .54), science teachers (M = 3.84, SD =.48) and

ISSN 2476-5880

social studies teachers (*M* = 3.78, *SD* = .66). There were no statistically significant differences in assessment literacy scores between the different subject areas, *F* (3, 148) = 2.205, p = .090, y^2 = .043. The ANOVA results are presented in Table 7.

Table 7
Results of ANOVA for Assessment Literacy Across Subject Areas

Construct	Sum of Squares	df	Mean	F	Sig
Assessment Literacy	2.18	3	.72	2.20	.09

Because the differences in group means were not statistically significant (p =.09), the subject area makes no statistically significant difference in the use of assessment literacy practices.

4.2 Instructional Improvement

A one-way ANOVA test was carried out to determine if teachers' use of instructional improvement practices was significantly different for respondents in each subject area. Levene's test for homogeneity of variances confirmed homogeneity of variances (p = .234). Data is presented as ± standard deviation. Use of instructional improvement practices differed between English teachers (*M* = 4.02, *SD* = .52), math teachers (*M* = 3.89, *SD* = .55), science teachers (*M*= 4.04, *SD* = .54) and social studies teachers (*M* = 4.06, *SD* = .65). There were no statistically significant differences in assessment literacy scores between the different subject areas, F (3, 148) = .512, p = .675, y^2 = .010. The ANOVA results are presented in Table 8.

Table 8
Results of ANOVA for Instructional Improvement Across Subject Areas

Construct	Sum of Squares	df	Mean	F	Sig
Assessment Literacy	.49	3	.16	.51	.67

Because the group means were not statistically significantly different (p = .67), the subject area makes no statistically significant difference in the use of assessment literacy practices.

5. Discussion

EFL teachers spend nearly half of their time assessing students' performance. Thus, teachers need to be literate in language assessment as regards how to construct a test or recognize the appropriate method to assess their students' learning. This study was thus conducted to investigate teacher assessment literacy and instructional improvement. After a meticulous study of the relevant

 ISSN 2476-5880

literature, two research questions were raised. The first one inspected "Is there a relationship between assessment literacy and instructional improvement among high school teachers of English, math, science, and social studies?" To answer this research question, Pearson's Product Correlations were used. The data correlating the assessment literacy and instructional improvement variables for each subject showed that there was a statistically positive relationship between teachers who use assessment literacy practices and those who use instructional improvement practices in the classrooms across each of the four subject areas. This showed that regardless of the subject area, the more teachers use best practices of assessment literacy in the classroom, the more they will employ instructional improvement practices. The findings are in line with Pearson et al.'s (2015) study suggesting that assessment literacy is crucial for designing effective science assessments and that instructional improvement of teachers affect their assessment literacy.

One caveat for this type of correlation is that it is dependent upon realizing that an understanding of assessment literacy should come before the instructional improvements under the tenets of backward design. This is in congruence with Wiggins and McTighe (1998), and Stiggin's (2008) manifesto on balanced literacy assessment proposing that teachers' better understanding of assessment literacy will enable them to step toward enduring instructional improvements.

If teachers are using assessment literacy well, they must collaborate ahead of time on how the assessments will be designed, keeping the outcomes in mind. Understanding the objectives and outcomes of student learning must come first, followed by formative assessments carefully designed to measure those objectives and outcomes. Once the data from those assessments are analyzed, broader instructional improvement practices, such as selecting teaching strategies and content for students, can be selected. The assessment data drives instruction and student learning.

However, because the teachers who design the tests are also the ones giving the instruction, it becomes very easy to fall into a "teach to the test" mentality where instruction is emphasized in areas that will be tested. Teachers will design tests only around concepts they want to teach rather than what the standards-aligned curriculum requires. The alignment of curriculum and instruction driving assessment and student learning is a legitimate concern that has been plaguing education systems for many years. This agrees with Guskey's (2009) and Ainsworth's (2003) call for teachers to first adhere to a standards-aligned curriculum as part of improving instruction.

Collaborative discussions regarding curriculum planning and student-centered instructional strategies should not revolve around how to modify assessments to meet teacher preferences or merely allow students to only get a higher score. Rather, designing assessments that best measure learning based on the content standards should evolve with quality teacher understanding and training. Discussions about creating appropriate instructional practices can then follow when the common assessment terms of how learning will be measured are agreed upon by all teachers.

There is a subtle difference that requires a shift from teachers' preferences to repeat practices based on their familiarity to meeting the needs of student learning. In some districts, this will require very crucial conversations between educators of all levels to ensure intentions,

 ISSN 2476-5880

expectations, and understanding of assessment design are driving selection and use of effective instructional practices and not vice versa. Standards-aligned curriculum, assessment, and instructional practices based on learning standards are the heart of the relationship between assessment literacy and instructional improvement.

The second research question inquired "To what extent do high school teachers of English, math, science, and social studies differ in their use of assessment literacy practices and instructional improvement practices?" The subject area shows no statistically significant differences for scores of assessment literacy practices or instructional improvement practices. There are variations in each subject's use of the practices, as seen through the descriptive statistics but they are not enough to justify the statement that one subject seems to be better suited to particular assessment literacy and instructional practices than another. This is in keeping with theories proposing that the benefits of assessment and instructional practices apply to all subject areas (Popham, 2006; Stiggins, 2008). Nevertheless, the findings are incongruent with Salkind's (2011) research who argued that math teachers need solid assessment literacy compared with teachers from other fields of study. In addition, the findings are contrary to Watmani et al.'s (2020) study which revealed the superiority of EFL teachers with an English teaching background over those without such background in terms of their assessment literacy and perceptions of Assessment Literacy components. The findings of this study, however, revealed no statistically significant difference among teachers of the four fields of study mentioned in terms of their assessment literacy.

Knowing vocabulary and plot characteristics does not mean that one comprehends everything they read. Understanding the nature of numbers does not ensure math fluency. Appreciating the scientific method does not make one a scientist. In the same vein, teachers with an understanding of assessment are not necessarily assessment literate. The difference lies in practice. While the findings of this study indicate that teachers are putting elements of assessment literacy into practice in their classrooms, it also shows that these practices are not always being used. This reduced amount of practice illustrates two primary issues: teachers are not understanding what assessment literacy practices are or how to apply them into instruction, and teachers are not progressing or being monitored in the effective use of assessment practices.

The data clearly shows that because of the positive correlative relationship between the two variables, teachers who use assessment literacy practices in their classrooms are more likely to get involved into instructional improvement practices in their classrooms. This aligns with and reinforces Popham's theories about teachers' needs for assessment literacy and conforms with Stiggins' theories that assessment literacy is an integral part of the learning and teaching process. These theories indicate that practicing assessment literacy is crucial for teachers to be effective in the classroom, and the data from this study supports and reinforces those theories. The results of this study indicate that understanding and using assessment literacy is an area of focus for schools' professional development systems.

ISSN 2476-5880

At the earliest phases of a teacher's career, it is important that assessment literacy practices are specifically taught and reinforced so new teachers walk into their classrooms with effective tools for improving instruction at their disposal. Teacher preparation programs should investigate the rigor of their assessment theory coursework. They also should investigate the possibility of including projects that require future teachers not only to demonstrate their understanding of effective assessment practices but also how such practices lead to greater instructional improvement. However, teaching and reinforcing these practices are not solely the domain of teacher preparation programs.

Practicing teachers may know that collaboration, assessment design, and data analysis practices are good to use in their classrooms, but without ongoing coaching and development in their use, those practices tend to fall by the wayside. Administrators must ensure they are at least as familiar with the practices of assessment literacy as they expect their teachers to be and must then work to ensure that their teachers are understanding and implementing those same practices through continued professional development.

This study shows, however, that due to the strong connection between Popham's assessment-centric theories, Ainsworth's theories on standards-aligned curriculum, and Stiggins' theories about needing effective implementation of both formative and summative assessments in the classroom, teachers need to understand and to use assessment literacy, in order to practice instructional improvement in the classroom. Yet, those practices cannot be developed and used in isolation regardless of instructional improvement outcomes. The two become linked through the use of a standards-aligned curriculum. This finding is in line with Wiggins and McTighe (1998), Ainsworth (2003), and Guskey (2009) who proposed that by linking assessment literacy to instructional improvement practices, teachers will understand the standards-based outcome of a task before beginning the task. Whether the task is designing an assessment or using a specific instructional practice, ongoing professional development systems should include training and reference as to how to use standards to connect assessment literacy practices with instructional improvement practices to have the biggest impact.

Simply teaching within a subject area has no statistically significant impact on the use of assessment literacy practices or instructional improvement practices in the classroom. As schools and districts design ongoing professional development training programs, it follows that due to the positive relationship between assessment literacy and instructional improvement, a focus on developing and promoting assessment literacy practices benefits all teachers. Based on this research, schools and districts will benefit by devoting time, energy, and resources into developing the assessment literacy practices of all teachers in order to see an increase in the use of best instructional improvement practices in all classrooms, regardless of the subject area. Because the subject area has no significant bearing on the improvement, schools may find it useful to implement professional development for assessments in cross-curricular teams, so those who need to become more assessment literate will do so from the best practitioners of assessment rather than just their departmental peers.

 ISSN 2476-5880

Being assessment literate is more than just testing students and collecting data. It is about what we do with that data which matters. The simple solution for teachers and schools is to rely on assessment companies and publishers to create student assessments, the conditions for administering them, the formulas for analyzing the data, and the measures of judging student success or failure. That simple solution, however, is not the best practice. Teachers must realize the importance of being assessment literate, and the positive impacts it has on their students, their professions, and the future of education. The value of education demands assessment literacy from those who practice it most – teachers in the classroom. Wiggins and McTighe (1998) asked, "How will we know if students have achieved the desired results and met the standards?" (p. 12). With a better understanding of assessment literacy in their toolkit, teachers will have the knowledge, data, and capability to answer that question each and every day.

6. Conclusion

The findings of these studies reveal teachers' lack of knowledge in assessment areas and propose extensive training for teachers to improve their level of understanding of assessment areas. On the other hand, research studies have limitations. When people are asked whether more knowledge is required about a topic, they might answer positively, influenced by their fear of being considered unprofessional or simply by their interest. Apart from theoretical implications, this study also bears some pedagogical implications. Language policy-makers and assessment experts should try their best to familiarize teachers with the recent views of assessment literacy. Teachers should also have the opportunity to employ principles of assessment literacy in their pedagogy and development. It is also recommended that schools and universities run special courses that familiarize teachers with the theories and practices of assessment literacy.

This study specifically looked at the link between assessment literacy practices and instructional practices in the classroom in a correlational study and was not designed to determine causation behind such practices. Further research, therefore, should continue to investigate the specific reasons as to why teachers use or do not use specific practices in their classrooms to further explain the results of the survey data revealed in this study. If it follows that assessment literacy practices lead to more instructional improvement practices, knowing why teachers do not use specific assessment practices would add to this investigation. Qualitative or mixed-methods studies would be appropriate to continue this line of research.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

References

Ainsworth, L. (2003). *"Unwrapping" the standards: A simple process to make standards manageable*. Englewood, CO: Lead and Learn Press.

Alemi, M., Miri, M., & Mozafarnezhad, A. (2019). Investigating the effects of online concurrent group dynamic assessment on enhancing grammatical accuracy of EFL Learners. Tabaran Institute of Higher Education. *International Journal of Language Testing, 9*(2), 29-43.

Black, P., & Wiliam, D. (1998). Assessment and classroom learning. Assessment in Education. *Principles, Policy and Practice, 5*(1), 7-23.

Black, P., & Wiliam, D. (2006). Developing a theory of formative assessment. In J. Gardner (Ed.), *Assessment and learning* (pp. 81-100). London: Sage.

Brookhart, S. M. (2011). Educational assessment knowledge and skills for teachers. *Educational Measurement: Issues and Practice, 30*(1), 3-12.

Buchanan, T. (2000). The efficacy of a World-Wide-Web mediated formative assessment. *Journal of Computer Assisted Learning, 16*(9), 193-200.

Byrene, B. M. (2010). *Structural equation modeling with AMOS: Basic concepts, applications, and programming (multivariate applications series)*. New York: Taylor & Francis Group.

Campbell, C., & Collins, V. L. (2007). Identifying essential topics in general and special education introductory assessment textbooks. *Educational Measurement: Issues and Practice, 26*(1), 9–18.

Chan, C. & Luo, J. (2020). An exploratory study on teacher assessment literacy: Do novice university teachers know how to assess students' written reflection? *Teachers and Teaching, 26*(2), 214-228.

Creswell, J. W. (2012). *Qualitative inquiry and research design: Choosing among five approaches.* Thousand Oaks, California, USA: Sage.

DeLuca, C. (2012). Preparing teachers for the age of accountability: Toward a framework for assessment education. *Action in Teacher Education, 34*(5), 576–591.

Estaji, M. & Ghiasvand, F. (2021). Assessment perceptions and practices in academic domain: The design and validation of an assessment identity questionnaire (TAIQ) for EFL teachers. Tabaran Institute of Higher Education. *International Journal of Language Testing, 11*(1), 103-131.

Fan, Y. C., Wang, T. H., & Wang, K. H. (2011). A web-based model for developing assessment literacy of secondary in-service teachers. *Computers in Education, 57*(2), 1727–1740.

Giraldo, F., & Murcia, D. (2019). Language assessment literacy and the professional development of pre-service language teachers. *Colombian Applied Linguistics Journal, 21*(2), 162-185. https://doi.org/10.14483/22487085.14514

Gotch, C. M., & French, B. F. (2014). A systematic review of assessment literacy measures. *Educational Measurement: Issues and Practice, 33*(4), 14-18.

Greenberg, J., & Walsh, K. (2012). *What teacher preparation programs teach about K-12 assessment: A review.* Washington, DC: National Council on Teacher Quality.

Guskey, T. (2009). Closing the knowledge gap on effective professional development. *Educational Horizons, 87*(4), 224-233.

Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2010). *Multivariate data analysis.* New Jersey: Pearson Prentice Hall.

Harding, L., & Kremmel, B. (2016). Teacher assessment literacy and professional development. In D. Tsagari & J. Banerjee (Eds.), *Handbook of second language assessment* (pp. 413–428). De Gruyter Mouton. https:// doi.org/10.1515/9781614513827-027

Herman, J. L., Webb, N. M., & Zuniga, S. A. (2007). Measurement issues in the alignment of standards and assessments. *Applied Measurement in Education, 20*(6), 101–126. doi:10.1080/08957340709336732

Hill, M. F., Ell, F., Grudnoff, L., & Limbrick, L. (2014). Practice what you preach: Initial teacher education students learning about assessment. *Assessment Matters, 7*(2), 90–112.

Kline, R. B. (2011). *Principles and practice of structural equation modeling (3rd ed.)*. New York: Guilford Publication.

Lam, R. (2015). Language assessment training in Hong Kong: Implications for language assessment literacy. *Language Testing, 32*(2), 169–197.

Leahy, S., & Wiliam, D. (2012). From teachers to schools: Scaling up professional development for formative assessment. In J. Gardner (Ed.), *Assessment and learning* (pp. 49–72). London: SAGE.

Lynn, M. R. (1986). Determination and quantification of content validity. *Nursing Research, 36*(6), 382-385.

Malone, M. E. (2013). The essentials of assessment literacy: Contrasts between testers and users. *Language Testing, 30*(3), 329–344.

Nurdiana, L. (2020). Language teacher assessment literacy: A current review. *Journal of English Language and Culture, 11*(1), 66-74.

Pearson, P. D., Knight, A. M., Cannady, M. A., Henderson, J. B., & Mcneill, K. L. (2015). Assessment at the intersection of science and literacy. *Theory into Practice, 14*(5), 228-237. doi:10.1080/00405841.2015.1044372

Popham, W. J. (2006). Needed: A dose of assessment literacy. *Educational Leadership, 63*(6), 84-85.

Popham, W. J. (2009). Assessment literacy for teachers: Faddish or fundamental? *Theory into Practice, 48*(1), 4–11.

Sadler, D. R. (1989). Formative assessment and the design of instructional systems. *Instructional Science, 18*(3), 119-140.

Salkind, N. (2011). *Internal and external validity.* The SAGE dictionary of quantitative management research, pp. 148-149.

Schneider, M. C., & Randel, B. (2010). Research on characteristics of effective professional development programs for enhancing educators' skills in formative assessment. In H. L. Andrade & G. J. Cizek (Eds.), *Handbook of formative assessment* (pp. 251–276). New York: Routledge.

Sharma, S. (1996). *Applied multivariate techniques*. John Wiley & Sons Inc., New York.

Smith, K. (2011). Professional development of teachers as prerequisite for EFL to be successfully implemented in the classroom. *Studies in Educational Evaluation, 37*(1), 55–61.

Stiggins, R. J. (2002). Assessment Crisis: The absence of assessment FOR learning. *Phi Delta Kappan, 83*(10), 758-765.

Stiggins, R. J. (2008). *Assessment manifesto: A call for the development of balanced assessment systems.* Portland, OR: ETS Assessment Training Institute.

Stiggins, R. J. (2010). Essential formative assessment competencies for teachers and school leaders. In H. L. Andrade & G. J. Cizek (Eds.), *Handbook of formative assessment* (pp. 233–250). New York, NY: Taylor & Francis.

Velan, G. M., Rakesh, K. K., Mark, D., & Wakefield, D. (2002). Web-based self-assessments in Pathology with question perception. *Pathology, 34* (3), 282-284.

Vogt, K., & Tsagari, D. (2014). Assessment literacy of foreign language teachers: Findings of a European study. *Language Assessment Quarterly, 11*(4), 374–402.

Walter, S. (2006). Instructional improvement: Building capacity for the professional development of librarians as teachers. *References & User Services Quarterly, 45* (3), 1-20.

Waltz, C. F., & Bausell, R. B. (1981). *Nursing research: Design, statistics, and computer analysis*. Philadelphia: FA Davis Company.

Watmani, R., Asadollahfam, H., & Behin, B. (2020). Demystifying language assessment literacy among high school teachers of English as a foreign language in Iran: Implications for teacher education reforms. Tabaran Institute of Higher Education. *International Journal of Language Testing, 10*(2), 129-144.

Wiggins, G., & McTighe, J. (1998). *Understanding by design.* Alexandria, VA: Association for Supervision and Curriculum Development.

Xie, Q., & Tan, S. (2019). Preparing primary English teachers in Hong Kong: Focusing on language assessment literacy. *The Journal of Asia TEFL, 16*(2), 653–673. https://doi. org/10.18823/asiatefl.2019.16.2.14.653

Appendix: Teacher Assessment Literacy and Instructional Improvement Survey

This research focuses on an investigation of the impact of assessment literacy on instructional improvement in public high school classrooms. For this study, assessment literacy refers to the process of understanding and using assessments as needed for the performance of your responsibilities as a classroom teacher. Instructional improvement refers to the practices of teachers to help students meet educational outcomes.

This survey is designed to be sent and returned electronically using the link provided. Each individual should complete the survey only once.

The survey has three sections:

1. Informed Consent
2. Demographic Information
3. Survey Questions
   - 3.1 Assessment Literacy Practices
   - 3.2 Instructional Improvement Practices

Participation in this research is voluntary and all information will be confidential. No personally identifiable such as name or email address will be collected.

Section 1–Informed Consent

- I understand the purpose of this study is: to examine the relationship between high school teachers' assessment literacy practices and instructional methods as a vehicle to guide instructional improvement.
- I understand that I will participate in an online survey. I understand that my participation is voluntary and that I can withdraw at any time without penalty.
- I understand that my individual responses will be kept confidential by the researcher.
- I understand that the results of this study may be presented at a professional conference, or published in an academic journal, but that my identity will in no way be revealed in such a report.
- I understand this research study focuses on the relationship between assessment literacy practices and instructional improvement practices. In case of additional information or questions, I may contact the researcher.

I have read the consent information above. I understand the nature of the study and by selecting one of the choices below, I give or do not give my consent for participating in this study.

a. I consent to participate in this research study.
b. I do not consent to participate in this research study.

E-Signature of the participant:

Section 2-Demographic Information

Participants' Information	Number/Name	Frequency (Percent)
1. School District		
2. Field of Study		
3. Academic Degree		
4. Teaching Experience (years)		
5. Assessment Literacy Sessions		
6. Instructional Improvement Sessions		

ISSN 2476-5880

Section 3-Questionnaire Items

Instructions: For each of the following statements regarding assessment practices, respond by indicating the frequency of use on the scale below.

1 = *Never*
2 = *Rarely*
3 = *Sometimes*
4 = *Often*
5 = *Always*

3.1 Assessment Literacy Practices	1 Never	2 Rarely	3 Sometimes	4 Often	5 Always
3.1.1 Collaboration					
7. I collaborate with my colleagues when preparing a classroom assessment.					
8. My colleagues and I discuss what data should be collected from a given assessment.					
9. My colleagues and I discuss how to use data that has been collected from a given assessment.					
10. My colleagues and I discuss which decisions will be based on the data collected from a given assessment.					
3.l.2 Assessment design					
11. I have a specific instructional purpose in mind when designing a classroom assessment.					
12. I clearly define success on a classroom assessment before the assessment is given.					
13. I align individual assessment items on classroom assessments to specific learning targets.					
14. I use a specific, research-based approach when designing classroom assessments.					
3.l.3 Item design and data collection methods					
15. My assessments consist of objective items ONLY (i.e., multiple-choice, true- false, matching, fill-in-the-blank).					
16. My assessments consist of subjective items ONLY (i.e., short answer, extended response/essay, problem-solving, performance).					
17. My assessments consist of a combination of objective and subjective items on classroom assessments.					

ISSN 2476-5880

18. I use alternative assessments (i.e., observations, interviews, portfolios) to assess students.					
3.l.4 Data analysis					
19. I perform item analyses on classroom assessments.					
20. I use data from item analysis of classroom assessments to make inferences about student learning in my classroom.					
21. I use data from item analysis of classroom assessments to make inferences about student growth in my classroom.					
22. I use data from item analysis of classroom assessments to make inferences about my own teaching.					
3.2 Instructional Improvement Practices					
3.2.l Decision-making					
23. I use assessment data to rank students in order.					
24. I use assessment data to divide students into categories (i.e., performance groups, instructional groups, or class selection).					
25. I use assessment data to evaluate effectiveness of a course curriculum, pacing, or content.					
26. I use assessment data to evaluate effectiveness of my own instructional strategies or methods.					
3.2.2 Standards-aligned curriculum					
27. I review and revise alignment lesson objectives to ensure alignment with learning standards.					
28. I review and revise alignment assessment items to ensure alignment with learning standards.					
29. I review and revise instructional strategies to ensure alignment with learning standards.					
30. I review and revise subject content to ensure alignment with learning standards.					
3.2.3 Learner-centered instruction					
31. I use a variety of assessment types in my classroom instruction.					
32. My students feel motivated to learn in my classroom.					
33. My students get pressure from sources outside the classroom to perform well on classroom assessments.					

 ISSN 2476-5880

34. I get pressure from administrators for my students to perform well on classroom assessments.					
3.2.4 Formative and summative assessments					
35. I use formative assessments to measure student progress during an instructional unit or period of time.					
36. I evaluate and adjust my instructional methods based on data from formative assessments.					
37. I use summative assessments to measure student progress after an instructional unit or period of time.					
38. I evaluate and adjust my instructional methods based on data from summative assessments.					

Thank you for taking the time to complete this survey.