Ethiopian Journal of Behavioural Studies, October 2025, 8(2), 89-131

Influences of Selected personal and Contextual Factors on Primary School Teachers' Formative Assessment Practices in Sheger City

Gadisa Terefe[1], Tilahun Gidey[2] and Dame Abera[3]

DOI: https://doi.org/10.63990/ejobs.v8i2.12279

Received: 1 December 2024; Accepted: 30 June 2025

Abstract

This study aimed to identify teachers' level of formative assessment practice and examine the main personal and contextual factors that influence this practice. A mixed methods research approach, involving an explanatory sequential design, was employed. A stratified random sampling technique was used to select participants for the survey. Teachers' survey questionnaires, principals' (key informants) interviews, classroom observation, and students' focus group discussion guides were used to gather data. Frequency counts, percentages, and multiple regression were used to analyse the quantitative data. The qualitative data were analysed using the deductive organization of responses from participants into themes, and the two data strands were then logically integrated. Findings showed that teachers' formative assessment practices varied widely. Although most teachers self-reported as proficient or advanced uses of formative assessment, classroom observations indicated mainly limited practices of formative assessment. The discrepancies between self-reported data and observed practices suggest potential overestimation in self-assessments. The major factors hampering a successful formative assessment practice in primary schools are teachers' low level of understanding of the strategies of formative assessment, minimal support from school leaders, the school's tendency to promote the summative assessment type, large class size, teachers teaching subjects they were not trained for or not adequately trained for, and the absence of instructional materials (including textbooks). Implications for practice and policies in the area of formative assessment were discussed.

Keywords: Formative assessment, assessment practice, personal factors, contextual factors, primary school teachers

[1] PhD candidate at School of Psychology, Addis Ababa University, CELS, Ethiopia
[2] Associate Professor at School of Psychology, Addis Ababa University, CELS, Ethiopia
[3] Associate Professor at School of Psychology, Addis Ababa University, CELS, Ethiopia

Introduction

Formative assessment (FA) is a process that teachers and/or students engage in to gather data on student progress and use it to inform decisions for the next lesson plans (Irons, 2008). It can take place at any point during the process of teaching and learning (Cagasan et al., 2020). FA is widely recognized for enhancing student learning and teaching effectiveness (Black & Wiliam, 1998; Brookhart, 2011). It involves gathering evidence to adjust instruction in real time (Irons, 2008).

There are four key FA strategies that characterise it: clarifying learning objectives to students, generating evidence of learning, providing feedback, and learners' self-and-peer assessment (Black & Wiliam, 1998; Brookhart, 2007). In order to gather evidence of student learning and use it to tailor instruction to the needs of specific students or the class as a whole, teachers need to use a variety of assessment techniques, such as questioning, observation, self-assessment, and peer assessment (Black & Wiliam, 2009). Giving students timely, targeted, and useful feedback enables them to recognize their strengths and limitations and modify their learning strategies (Wiliam, 2011). Teachers are required to be competent enough to use FA embedded in the classroom instruction. Several studies have investigated teachers' classroom implementation of FA.

A study by Stiggins (2005) on teachers' use of FA across different grade levels and subject matters revealed varying levels of practice. Some of them greatly implemented it, whereby they frequently used FA strategies to collect information about their students' learning processes and modify teaching, while others had low

practice levels with more reliance on summative tests. Popham (2008) and Ruiz-Primo et al. (2006) also carried out studies that examined how teachers used FAs in their classroom settings. They both found that most teachers practiced FAs on an average basis. Some teachers showed a high level of commitment towards using this method for instruction, consistently applying questioning techniques, self-assessments, and feedback as their guiding principles, whereas for others it was less common, despite some cases involving only minimal application of these strategies (Ruiz-Primo et al., 2006).

In Ethiopia, the recently formulated General Education Curriculum Framework stipulated that classroom assessments must be continuously used to improve instruction (MoE, 2020). It particularly highlighted the need for continuous assessment at the primary school level, with an emphasis on improving students' proficiency in language, mathematics, morality, aesthetics, physical development, and the environment. The ultimate goal is to enhance both student learning and teacher effectiveness. Existing local studies, however, showed a varied extent of FA use. For instance, Yidenek (2018) explored seven grade 9 EFL teachers' integrated approach to classroom assessment practices and found that the teachers used only summative and no formative assessment methods. Dereje et al. (2022), in their study on secondary school EFL teachers' classroom assessment conceptions and practices, reported that teachers who participated in the study used summative assessment tools, although they believed in the positive roles FA could play. Hailay and Abate (2022), on their part, studied assessment for learning practices and challenges and observed that close-ended self-report data showed the presence of FA practices, whereas qualitatively gathered data revealed no formative but only summative assessment uses in the schools. With their

mixed-methods study on the practices of FA in selected secondary schools, Janbo et al. (2020) found that teachers sometimes communicate learning objectives and provide formative feedback. The authors further reported that the teachers varied considerably in their FA practices. Moreover, a study on secondary school science teachers' practice of assessment for learning (Askalemariam, 2015) revealed similar findings with a very low practice of FA. The author also reported that the teachers did not use a variety of FA strategies due to a lack of science resources, a large class size, a shortage of time, inadequate school support, a lack of appropriate professional development activities, a lack of instructional materials, students' and teachers' negative perceptions towards FA, a lack of knowledge and skills about FA, and the large content of courses.

However, all of the above-mentioned local studies, except one study by Sintayehu (2016) conducted at the primary school level in Chagni, addressed teachers' FA practices at the tertiary and secondary education levels. For example, Askalemariam (2015), Dereje et al. (2022), Janbo et al. (2020), and Yidenek (2018) were carried out on classroom assessment practices in the secondary schools of the country.

Successful implementation of FA depends on the existing associated factors that either facilitate or inhibit its practice. The primary factors influencing FA practices that have been extensively studied in the literature fall mostly into two groups: contextual and personal factors (Zi Yan et al., 2021). Research has shown the most common personal factors influencing FA were the teachers' instrumental attitude, their understanding of it, and their teaching experience. The way teachers perceive and embrace FA (i.e., their attitude) greatly affects their utilization of it. According to Ruiz-

Primo and Furtak (2006), teachers who are convinced that FA is vital to improving student learning tend to employ it more often.

In connection with teachers' understanding of FA, research shows that teachers who possess a solid understanding of assessment principles and procedures are more inclined to employ FA regularly (Hattie & Timperly, 2007; Heritage, 2010). To this end, education and training that focus on FA can assist teachers in acquiring the necessary knowledge and skills to effectively use it (Brookhart, 2007; Heritage, 2010).

In addition, teaching experience was also reported by some studies to be associated with FA practice. For instance, Hattie and Timperely (2007), after analysing extensive research, reported that experienced teachers are more able to implement FA effectively using strategies such as explaining learning goals explicitly, providing students with timely feedback based on the achievement of those goals, and promoting intrinsic student motivation and self-regulation.

Researchers have further reported several contextual factors that can affect teachers' FA practices, such as instructional time, workload, assessment policy, class size, and support from leaders. Teachers who perceive that they have less time available used FA less frequently (Hailay & Abate, 2022). There is also an impact of existing curriculum and assessment policies on the mode teachers adopt for FA. For instance, schools that place a higher priority on summative assessments could give FAs less attention (Black & Wiliam, 1998; Ruiz-Primo & Furtak, 2006).

Similarly, class size affects the frequency with which teachers use FA strategies (Ruiz-Primo & Furtak, 2006). That is, in smaller classes (i.e., classes with fewer than

25 students), teachers use techniques such as question and answer, feedback, and peer assessment more frequently, perhaps due to individual involvement and increased opportunity to respond (Garfield & Ben-Zvi, 2007; Ruiz-Primo & Furtak, 2006).

Regarding support for teachers, Black and Wiliam (1998) and Ruiz-Primo and Furtak (2006) reported that teachers who received support from school leaders were more likely to use FA consistently.

This study was guided by Kozma's (2003) three-level model, which explains the factors influencing classroom practice. Kozma proposed three levels of influences on teachers' understanding and adoption of novel classroom instructional practices. In the model, he identified three distinct but interacting levels of contextual factors: micro, meso, and macro, which are believed to explicitly affect teachers' classroom practices.

The immediate classroom context is included in the micro-level factors. A range of effects from the classroom may fall into this category. This could involve the teacher's assessment literacy, the number of pupils, and the students' past success in the subject matter, among other distinctive characteristics of both the teacher and the students. Furthermore, access to in-class technologies and resources, such as boards and other materials that could aid in assessment, could be considered at the classroom level (Kozma, 2003).

The meso-level is made up of elements that directly affect the classroom yet are not located within it. These elements are usually classified as school-level factors. Meso-level elements that are strongly related to the school include, for instance, the school's climate supporting assessment procedures, policies, and assistance from

leadership regarding assessments, and resources for assessments across the entire school. Conversely, the meso-level encompasses a broader conceptual scope than the school alone since it can take into account elements that impact the classroom but are not physically present in it. Parents' demands and community expectations are examples of meso-level influences that are not school-specific (Kozma, 2003).

The macro level is made up of distal factors that do not directly affect the classroom but can indirectly affect the meso-level and the classroom as a result. Education policy at the federal, state, and local levels may be included at the macro-level, depending upon the framework of educational governance. This may also include policies from the school's affiliated social institutions in the case of private and non-governmental schools (e.g., community or religious organizations funding the school) (Kozma, 2003).

Therefore, because the present study mainly focused on examining the nature and extent of influences of various factors on the teachers' FA practices, it was preferred to be underpinned by Kozma's theoretical framework of influencers of teachers' classroom practice. The study sought to develop a comprehensive model of how personal and contextual factors simultaneously influence teachers' FA practice. Knowing the factors that can affect the proper implementation of FA assists teachers and other decision- makers in figuring out things to be improved in order to maximize its effective utilizations.

In Ethiopia, studies that have investigated the factors associated with proper FA practice are scarce. To the researchers' knowledge, Askalemariam (2016) and Hailay

and Abate (2022) are the only two, who have studied factors associated with FA practices. These available scarce studies reported what a few teachers mentioned as factors influencing FA practices, which included lack of resources, large class size, shortage of instructional time, inadequate school support, lack of appropriate professional development activities, students' and teachers' negative perceptions of FA, teachers' lack of knowledge and skills about FA, and the large content of subjects, without considering the relative contribution of these multiple factors that can explain teachers' FA practices (Askalemariam, 2016; Hailay & Abate, 2022). Thus, the present study aimed at examining the practices and challenges associated with FA among teachers in Sheger City.

Overall, this study highlights notable justifications. The justifications emphasise:

Contextual Significance and Timing: This study aligns with the time context and the educational transformations currently taking place in Ethiopia. The Ethiopian Ministry of Education has shown commitment to implementing FAs, as stipulated in the general education curriculum framework. Therefore, it is crucial to examine its current classroom practices.

Existing Research Gaps: The role that FA plays in improving students' learning is being recognised globally. However, there seems to be a dearth of research on this subject within the Ethiopian context at the primary education level. The literature reviewed mainly focused on secondary and higher education institutions, leaving a gap in understanding the primary school-level teachers' practice of FA.

Mixed Methods Approach: The study used a mixed-methods approach to gaining an in-depth insight into the teachers' FA practices. This approach enables the triangulation of information on the subject obtained from surveys, key informant

interviews, classroom observation and students' focus group discussions. This methodology also establishes the credibility of the findings and provides a deeper understanding of the beliefs, challenges, and practices of teachers towards FAs.

Consequences for Educational Stakeholders: The results of the study could have enormous significance for several educational stakeholders, including teachers, educational administrators, and policymakers. By delineating what works in FA, the study may improve teacher training programmes and instructional methods.

Additionally, the findings from this study can be useful for long-term educational planning and policy formulation. This research may also compare the local practice with the best practice globally and offer practical recommendations that promote the development of a framework in support of teachers in the implementation of the FA.

Teacher Empowerment and Professional Growth: Ultimately, the study aims to empower the teacher. It will do so by creating an environment that encourages the professional growth of the teachers through the effective use of FA. The teachers' understanding of the FA strategies will be improved, and they will adopt more student-centred methodologies in their classes. Moreover, the results might promote teamwork among the teachers. They may share the best practices and improve their teaching skills together. This will create a culture of continuous professional growth.

Studying primary school classroom practices, including FA practices, is particularly pertinent as they are the foundation phases of schooling where children begin to further develop and enhance their cognitive and affective capacities for learning new knowledge and skills, thereby establishing a base for future development (DeLuca et al., 2013). The main research questions were:

1. What is the level of FA practice in primary schools?

2. What personal (teachers' instrumental attitudes towards FA, teaching experience, understanding of FA) and contextual (support from school leaders, class size, time constraints, workload, and assessment policy) factors significantly predict teachers' FA practices?

Operational definition of key terms

In this study,

- **FA practice:** refers to levels of teachers' use of FA strategies and principles.
- **Personal Factors:** include teachers' individual traits and characteristics, which vary from person to person that can shape their implementations of FA. E.g., attitude, understanding, experience, etc.
- **Contextual factors:** involve factors that are external to the individual teacher but that have immediate influence upon FA practices – especially factors related to the school itself or its surrounding community. E.g., leaders' support, school policy on assessment, working condition, etc.
- **Instrumental attitude towards FA:** refers to teachers' opinions regarding the importance of FA in the teaching and learning process.
- **Feedback:** It is a comment or piece of information that learners receive from a teacher on the process or product of a learning task.

Methods

The study's questions were addressed using a mixed-methods approach. The main intentions of using a mixed research approach in this study were to deeply explore

teachers' formative assessment practices and the influencing factors, while also enriching quantitative findings with qualitative insights..

The particular mixed research design employed in this study was the explanatory sequential design (QUAN → qual). Surveys of teachers, classroom observation, FGD with selected students, and interviews with key informants (i.e., school principals) were used to gather data. As a result, quantitative data were collected and examined first (via a teacher survey). Subsequently, further data was gathered using key informant interviews, classroom observation and focus group discussions to get more information on teachers' FA practices and associated factors. The two data strands were mixed based on theme breakdowns to present the findings in a clear and understandable manner.

Description of the Study Site

Sheger City is a city that Oromia Regional State has organized by combining the former Sebeta, Burayu, Lega-tafo Lege-dadi, Sululta, and Gelan towns. It covers 160,000 hectares of land, with a reported population size of around 3,000,000. The city has 12 sub-cities and 36 districts (woredas). The city surrounds and is attached in all directions to the country's capital city, Addis Ababa, and hence could be considered an educational and economic centre. The education offices are established at sub-city levels. Sheger City embraces more experienced and qualified teachers with diversified cultures, coming from all directions and outlets to the city in search of a better life and educational opportunities, which may make it an important context for examining the phenomenon under investigation.

Study Participants and Sampling Techniques

Sampling Techniques

First, three sub-cities out of a total of twelve were chosen using the lottery method. These three sub-cities were Gefersa Guje, Gelan Guda, and Sebeta. Fortunately, there are more teachers available in these sub-cities than can be included in the study, and the distribution of teachers within the sub-cities is generally balanced concerning qualifications, gender, experience, and other characteristics. Only government-owned primary schools were included in the study due to resource limitations.

Instead of visiting numerous schools with fewer teachers, the first three schools in each sub-city with the highest number of teachers were selected. The intention was to focus on the larger schools that could accommodate more pupils in order to make the greatest use of the few resources. Within these three sub-cities, there were 858 teachers teaching in 44 primary schools. Thus, 272 was the sample size determined by applying Yamane's (1967) straightforward formula for calculating sample size: n= (N)/ (1+N (e^2)). Stratified random sampling was used to select participants after the sample size was determined. A proportionate allocation of participants was made based on the population size of each sub-city.

Ultimately, a total of 223 teachers (M = 107, F = 113, and 3 lacking gender information) selected by the lottery method from the primary schools in the three sub-cities filled in the questionnaire. Gefersa Guje sub-city accounted for 95 (42.6%) of the teachers who filled out the questionnaire; Sebeta and Gelan Guda sub-cities accounted for 30.5% and 26.9% of the teachers who responded, respectively, as is shown in Table 1.

Overall, the study involved teachers from nine primary schools—three from each sub-city. The elementary schools were located in the following sub-cities: Roge-Game, Haro Jila, and Kaleab from Sebeta; Alemgena, Daleti, and Melka Sebeta from Gelan Guda; and Kolobo, Chorisa, and Ethio-Yugoslav from Gefersa Guje. The teachers' distribution across a range of socio-demographic factors, such as sex, the subjects they teach, and their access to pre-service and in-service training, suggests that the samples appear to be representative of the target population and enables inferences or generalizations about the characteristics of the sample to the target population.

Moreover, the researchers purposefully selected four instructional leaders from four primary schools as key informants to participate in semi-structured interviews, with the aim of obtaining deeper and richer information on the teachers' FA practices. They were chosen because they were close to one of the researchers' homes, which made it easier to make appointments and go to schools, and because the primary schools in those sub-cities work in very similar contexts. Furthermore, to get more data on the FA practices of the teachers, seven classes at three randomly chosen schools were observed in the classroom (three classes at each of the two schools and one class at one school). With the assistance of the principals of each school, the teachers were chosen based on their willingness to be observed. They were teaching different school subjects, including mathematics, environmental science, Afan Oromo, Geda, and Safu. Besides, to gather more data on the teachers' practices of FA and corresponding challenges, two groups of primary school students, each consisting of six members, were selected from one of the primary schools with the help of the school's principal to participate in the FGDs. Accordingly, twelve students participated in the FGD. Of those twelve students, 6 were

female students. More active and mature students who could better express their views on the teachers' FA practices were selected purposefully from among grades 6 to 8.

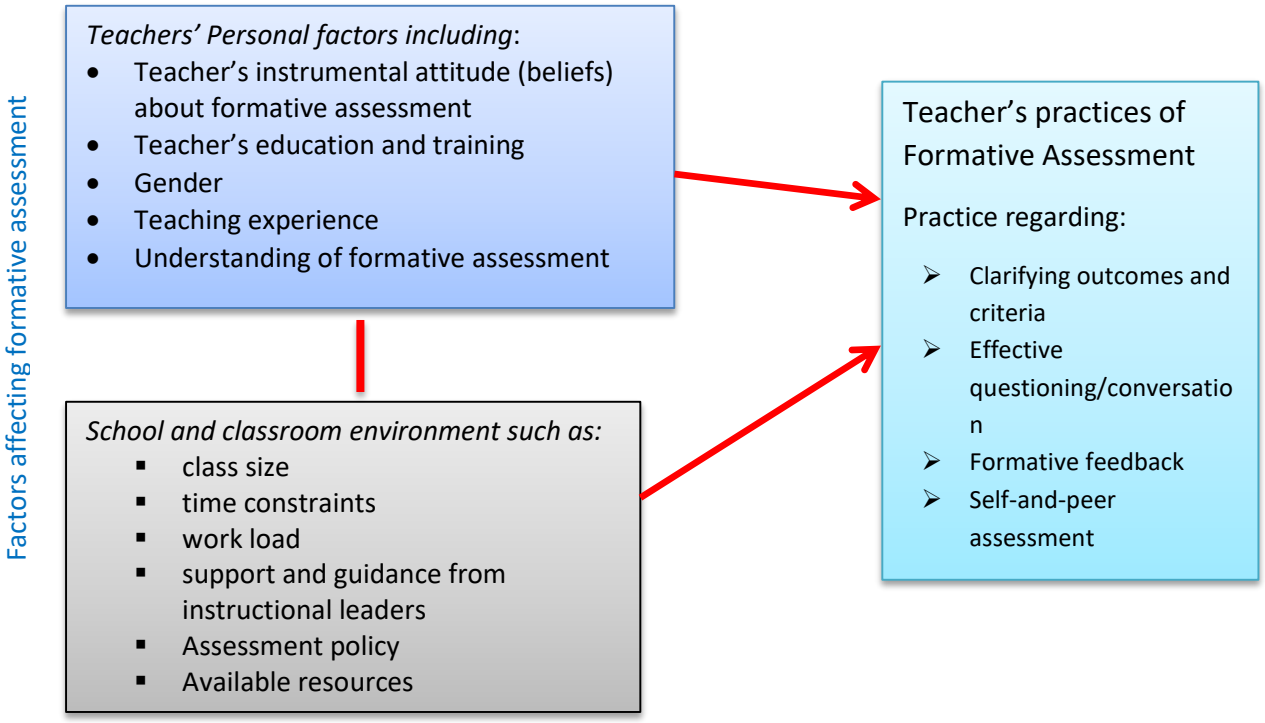
Variables of the Study

In this study, the following predictor and outcome variables were involved.

Predictors: teachers' understanding of FA, teaching experience, weekly workload, average number of students, assessment type required, usefulness of FA, internal school support, adequacy of instructional time.

Dependent Variable: Teachers' practice

Their interrelationships are indicated as follows:



Instruments

The instruments for gathering data for this study were a self-administered questionnaire, a guide for key informant semi-structured interviews, classroom observation checklist, and a focus group discussion guide.

Self-administered Questionnaire

The survey data was gathered via a self-administered questionnaire that contained items about teachers' socio-demographic characteristics, FA practice, and factors that may influence FA practice. The demographic and socio-economic part of the items within the questionnaire were developed based on a review of related literature, while the FA practice measuring items were adapted from scales, namely teachers' formative assessment literacy scale (TFALS) and teachers' FA practice scale (TFAPS).

Teachers' Formative Assessment Literacy Scale (TFALS)

Eight items from the Teachers' Formative Assessment Literacy Scale (TFALS) (Yan & Pastore, 2022a) were used to collect data on teachers' FA practices. This subscale was applied, combined with nine items from Yan and Pastore's teachers' formative assessment practice scale (TFAPS), which was developed in 2022(b).

A normative group of 585 elementary and secondary school teachers in Hong Kong participated in the development and validation of the overall TFALS scale (339 teachers from 9 primary schools and 246 teachers from 5 secondary schools) (Yan & Pastore, 2022a). Good internal consistency (a Cronbach alpha of .88) was reported for the practical dimension of the scale.

Teachers' Formative Assessment Practice Scale (TFAPS)

According to Yan and Pastore (2022b), there were two components of TFAPS: student-directed formative assessment (SdFA) and teacher-directed formative assessment (TdFA). All four of the SdFA items and five of the six TdFA items were employed in this study. Cronbach's alpha values of.75 and.70 for the TdFA and SdFA components, respectively, were reported, indicating satisfactory internal consistency measures (Yan & Pastore, 2022b). A five-point Likert scale was employed for all items, with 1 indicating never used, 2 indicating seldom used, 3 indicating occasionally used, 4 indicating regularly used, and 5 indicating very frequently used. The response levels or categories were also reduced from six to five to avoid overlap of some response options when translated into local languages.

Before collecting the survey data, these scales were translated into Afan Oromo (the language most participants speak). Two translators (an English language teacher from Sebeta College of Teacher Education and one of the researchers) translated the instruments from the source language (English) into Afan Oromo. Another college instructor who frequently taught teacher candidates measurement and evaluation courses conducted back translation. The content validity index (CVI), which was based on ratings by assessment specialists for items clarity and relevance to the issue under study, was used to verify the content validity of the scales.

Content Validity

First, the number of responses from 10 relevant experts who marked each item as "essential" was counted to determine the Content Validity Ratio (CVR) for each item using Lawshe's formula. Since none of the items had a CVR of less than 0.65, they were retained in the instrument in accordance with Lawshe's (1975) decision criteria. The self-report questionnaire had very good content validity in terms of relevance, as indicated by the range of observed CVR values, ranging from .80 to 1, with an overall content validity index (CVI) score of .92. However, in light of the experts' written comments, one of the two duplicate items was eliminated.

Reliability

Using Cronbach's alpha coefficient, the reliability of scales assessing the overall formative assessment practice, teacher-directed formative assessment (TdFA), and student-directed formative assessment (SdFA) was determined. Cronbach's alpha of.887 for the overall practice scale suggests that it has a high degree of internal consistency. The TdFA and SdFA sub-scales each had an alpha value of .826, showing high internal consistency among the items measuring the corresponding factors.

Confirmatory Factor Analysis (CFA)

A confirmatory factor analysis was conducted for the TdFA and SdFA sub-scales, in which the TdFA showed support for the proposed model while the SdFA showed discrepancies. In aggregate, the whole scale can be considered to have two factors applicable in the present study context.

For the TdFA sub-scale, all four fit measures supported the proposed model. The chi-square value was 4.094 (df = 5, N = 223), with p = .536, indicating an acceptable match between the model and observed data. The comparative fit index (CFI) and the normed fit index (NFI) yielded values of 1.00 and .991, respectively, suggesting an excellent fit. Additionally, the root mean square error of approximation (RMSEA) was.000, indicating an excellent fit (Loehlin, 2004).

In contrast, the results did not support the SdFA sub-scale model, highlighting the need for further exploration in future research). A chi square value of 21.301 (N = 223), with a p of =0.000 was obtained, that indicate a model-data non-fit. The CFI and NFI were 0.944 and 0.939, respectively, indicating a marginally poor model-data fit. Similarly, the RMSEA suggested a poor match with a value of 208.

The scoring and interpretation of the FA practice levels of teachers on the survey were conducted using the score categories and criteria proposed and used by Kanjee and Mthembu (2015), as presented in Table 1 below.

Table 1

Formative Assessment Practices: Performance levels, score categories and implications

Performance level	**Formative Assessment**	
	% Score	Implications for practice
Below basic	≤ 34	Teacher is unable to apply FA strategies in the class or use FA to identify learning gaps
Basic	35 – 64	Teacher is able to apply some FA strategies and identify gaps in learning, but is unable or unlikely to address these gaps
Proficient	65 – 79	Teacher is able to effectively apply all FA strategies and use FA to identify and address learning gaps
Advanced	80 -100	Teacher is able to effectively apply and use FA to improve learning and teaching practices

Source: Kanjee and Mthembu, 2015

All items measuring the DV were on a five-point Likert scale, where 1 meant never used, 2 meant seldom used, 3 meant occasionally used, 4 meant regularly used, and 5 meant very frequently used. Hence the scores on the DV ranged from 17 to 85.
The IVs were nominal, where most were dichotomous.

A Semi-structured Interview Guide

Four primary school principals participated in face-to-face, semi-structured interviews to gather data on the factors affecting teachers' FA practice. The purpose was to collect data that can supplement or help in explaining the quantitative data. Field

notes were taken, analysed and then translated to the English language. A minimum of one and a maximum of one hour and a half was spent on the interview. Interviews were held in the vice principals' or principals' offices. For this, an interview guide that addressed the research questions was developed and used. The interview guide specifically addressed if the teachers practised the FA practices in their classrooms and the challenges encountered in this case.

FGD Guide

A focus group discussion guide was developed based on the literature reviewed and was used to collect data from a group of students selected from more active and mature students in grades 6 to 8. A group of participants consisting of six members participated in the FGD. One FGD was conducted aiming at gathering students' views on their teachers' extent of use of FA. The FGD data was believed to supplement the data gathered through the quantitative tools. The student FGD took place for nearly one hour under the shade of a tree. Field notes were collected, analysed, and then translated into English. The guide with eight generic items was created with reference to the basic research questions.

The items in the FGD particularly focused on exploring students' views on the status of teachers' practices of FA.

Observation Checklist

A classroom observation checklist developed by Cagasan et al. (2020) was contextualized and used to collect data on participants' actual practices and experiences of FA. This tool had been developed and used to assess FA practices of teachers in the Philippines. It was thought to help in obtaining an in-depth understanding of teachers'

classroom FA practice. It focuses on exploring how teachers elicit information and use the elicited information to inform and improve student learning (Cagasan et.al, 2020). The Elicit (E) and Use (U) components of FA were the main focuses in developing the observation tool, the Classroom Observation of Formative Assessment (COFA). The elicit (E) component pertains to competencies in eliciting evidence of student learning to determine what students know and can do, while the use (U) component involves using evidence of student learning to move toward the learning goal (Cagasan et al., 2020).

Seven statements of indicative behaviour make up the COFA, each of which focuses on a different component of eliciting or using FA. Each indicative behaviour has a set of practices related to it. Observers indicate whether the indicative behaviours are present during class using COFA. They can also record supporting details for their evaluations or examples of practices in the tool's last column.

Method of Data Analysis

The study employed **descriptive statistics,** such as frequency, percentage, mean, and standard deviation, to describe the participants' background characteristics and the level of FA practices among the teachers.

Multiple regression analysis was applied to analyse relationships that could exist among a set of major factors, the independent variables (IVs) and the dependent variable (DV), as raised in research question 2.

In line with this, to address the effect of common method bias that could arise from the use of the same survey instrument to measure both the independent and the dependent variables, Harman's single factor test was used.

The qualitative data, on the other hand, were analysed using a thematic analysis approach to explore descriptive insights that complement findings from the quantitative data. To this end, the interview, classroom observation and FGD data were read and re-read to identify and label significant statements, structure and group the statements into certain categories or themes, cluster themes into broader categories, and finally assemble the detailed descriptions of the issues in the themes into clusters. Hence, the qualitative data analysis followed a deductive qualitative data analysis approach.

Ethical Considerations

Consent from institutions and individual participants was requested, and oral consents were acquired prior to data collection. Similarly, permission was obtained from the school principal and homeroom teachers regarding the students' FGD participation, in addition to a careful clarification of the purpose of the FGD for the students. The participants were told that they remain anonymous, that every piece of information is treated with the highest confidentiality, and that they have the option to decline participation at any time after data collection has begun. The researchers were also open about who they were and where they came from. In addition to this, participants were initially provided with a comprehensive explanation of the goals and procedures of the research.

Results

Data Screening and Tests of Statistical Assumptions

Before beginning the data analysis, all variables were screened for outliers, missing values, and other statistical assumption violations using SPSS frequencies, explore, and missing value analysis.

Firstly, six univariate outliers were identified and subsequently corrected. The univariate outlier detection was made through the frequency distribution analysis for each variable. Then, three variables—one continuous (FA practice) and two categorical (instrumental attitude toward FA and asking students to assess peers' work)—were found to be candidates for imputation during a screening process carried out to identify missing values. Since the sample mean likely to be the most accurate representation of the population mean, the mean substitution imputation technique was employed for the missing value analysis. Because the missing values in the current study's data were relatively small, this technique was chosen for missing value analysis. Values from the SPSS missing value analysis were used to replace 29 (13%) of the missing values on FA practice. The mean score for FA practice was found to be 69.499. Additionally, the SPSS analysis of missing values was used to replace the missing values for the following: the teacher's attitude toward FA that had 13 missing values (5.83% of the cases) and asking students to assess the work of their peers that had 11 missing values (4.93). The computed means for the attitude and asking students to assess their classmates' work variables were 1.799 and 3.563, respectively.

The FA practice, the scale variable, had a skewness value of .084, which was computed using the Kolmogorov-Smirnov and indicates that the distribution is approximately normal because it is within the allowed limits of +1 and -1. The Shapiro-

Wilk test of .959 for the FA practice also supported the same conclusion. Moreover, a non-significant homogeneity of variance test (P=.763) was observed, indicating the variances of the dependent variable are comparable across the levels of the independent variables.

Demographic, Personal and Contextual characteristics of respondents

Gender, teaching experience, instrumental attitude toward FA, class size, amount of time allocated for curriculum, workload, school policy about assessment, and internal school support were among the socio-demographic, personal and contextual factors considered, as presented in Table 2. Since the samples appear to be representative of the target population, the data in the table supported the possibility of making inferences or generalizations about the target population from the sample characteristics.

Table 2

Demographic, personal and contextual Characteristics of the Study Respondents

Variable	**N**	**%**
Sex		
Male	107	48
Female	113	51.4
Missing	3	1.3
Attitude towards FA		
Unfavourable	47	21.08
Favourable	176	78.92
Average Number of Students in a Class		
More than 50	180	80.7

30 to 50	35	15.7
Less than 30	7	3.1
Missing	1	0.4
Internal School Support		
Not sufficient	160	71.7
Sufficient	58	26.0
Missing	5	2.2
School internal policy		
Encourages Summative Assessment	66	29.6
Encourages Formative Asses	157	70.4
Weekly Workload		
Overloaded	94	42.2
Normal Load Range	126	56.5
Missing	3	1.3
Adequacy of Instructional Time		
Never Sufficient	84	37.7
Almost Sufficient	139	62.3

Teachers' Perceived Practice level of FA

This study used proficiency or competency levels of FA practice as determined by Kanjee and Mthembu (2015). They established four levels: below basic (0–34%), basic (35–64%), proficient (65-79%), and advanced (80–100%) based on FA practice scores. Table 3 presents an overview of FA practice as reported by teachers in the current study. Those in the below basic, basic, proficient, and advanced performance levels accounted for 1.84%, 20.63%, 30.49%, and 47.10%, respectively.

Table 3

Teachers' self-report result on their FA practice

Construct	% score	N of teachers	% of teachers	Implications
Level of practice of FA				
Below Basic	0-34	4	1.84	Failure to understand FA strategies & identify gaps
Basic	35-64	46	20.63	Partial understanding
Proficient	65-79	58	30.49	sufficient understanding
Advanced	80-100	105	47.10	comprehensive understanding

This indicates that the teachers greatly varied in their FA practices. The results also reveal that the majority of teachers have proficient to advanced levels of practice, although it needs caution in interpreting it, as the data used is self-response data, where the teachers may rate high values for their own practice, regardless of their actual FA practice status. Therefore, we used data from selected teachers' classroom observations to further understand their FA practice status.

Classroom observations conducted at selected schools revealed more of unsatisfactory practice of FA. At OPS1 (OPS-for observed primary school), for example, observations were made in Grade 1B for the subject Safu, Grade 2C for A/Oromo, and Grade 4B for Geda. In the Grade 1B Safu class, it was observed that the teacher had poor spelling skills in the medium of instruction, no lesson plan, and no clarification of the learning objectives. She gave no descriptive feedback to students. Eliciting evidence of learning from the whole class and peer- and self-assessment

strategies were also lacking. In grade 2C Afan Oromo and grade 4B Geda classes in the same school and grade 7 Mathematics and grade 8B Afan Oromo classes at OPS2, similar problems were observed: no clarification of learning objectives was made, no feedback was provided, and self- and peer assessment practices were lacking.

However, in the accelerated learning for Africa (ALfA) class observed at OPS2 and a grade 1B Afan Oromo class observed at OPS3, relatively better FA practices were observed, including clarifying the learning objectives, using various techniques (oral questioning, discussion, observation of students' work) to elicit evidence of students' learning, and giving some feedback to students compared to other classes, while there are still areas for improvement, for instance, eliciting evidence from the whole class and fully using the elicited evidence for improvement. The relatively better practice observed in the ALfA class may be due to the more practical teacher training given to facilitators of the ALfA classes by the Geneva Global, Ethiopia.

Influences of personal and contextual factors on FA practice

Standard multiple regression analysis was conducted to examine the predictive relationship between various independent variables and FA practice. The predictors were teaching experience, instrumental attitude towards FA, class size, internal school support, school policy on assessment, workload, adequacy of instructional time, and level of understanding of FA.

As the data for regression analysis were collected through a self-administered questionnaire, both procedural and statistical techniques were employed to avoid a common method bias. Firstly, the researcher assured the respondents that the procedural approach protects anonymity and maintains confidentiality to reducing social

desirability bias. Furthermore, it was emphasized in the instruction that there are no right or wrong answers to reduce acquiescence bias. Secondly, Harman's single factor test (a statistical technique) was conducted to check for the presence of common method bias (CMB). For this test, an exploratory factor analysis was conducted. The result, as indicated in Table 4, indicates that the first extracted factor accounted for only 18.14% of the variance, far less than the 50% or above cut-off contribution suggested for a CMB to exist, indicating the absence of a single factor accounting for a significant amount of variance and hence indicating the absence of a common method bias.

Table 4

Total Variance Explained						
Factor	Initial Eigenvalues			Extraction Sums of Squared Loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	1.814	18.141	18.141	1.225	12.249	12.249
2	1.571	15.715	33.856	1.025	10.250	22.499
3	1.246	12.458	46.314	.576	5.759	28.259
4	1.044	10.444	56.757	.378	3.784	32.042
5	.956	9.557	66.314			
6	.847	8.474	74.788			
7	.768	7.676	82.464			
8	.720	7.197	89.662			
9	.567	5.666	95.328			
10	.467	4.672	100.000			

Extraction Method: Principal Axis Factoring.

Results of Harman's single-factor test for common methods bias

As shown in Table 5, FA Understanding, internal school support, and assessment-related school policy (assessment type required) became significant predictors of FA Practice scores, while other variables (teaching experience, instrumental attitudes towards FA, class size, workload, and instructional time) showed non-significant

associations; however, as evidenced by the unstandardized coefficients (Table 5), FA understanding contributed the most to the prediction of FA practice (p<.05), while all the remaining predictors 'contributions were to a lesser extent (p>.05). The FA understanding had a partial correlation square value of .207, indicating that it uniquely contributed nearly 21% to the variance in the teachers' FA practice.

Table 5

Intercorreltions for FA practice and predictor variables (N=214)

Variable	1	2	3	4	5	6	7	8
Teaching experience (1)								
Attitude towards FA (2)	.153*							
Class size (3)	-.122	-.048						
Internal school support (4)	-.236*	.038	.143*					
Assessment type required (5)	.028	.017	-.043	.097				
Weekly workload (6)	-.054	.061	.001	.115	.016			
Adequacy of instructional time (7)	-.251**	-.022	.017	.130	.043	.238**		
T_Understanding (8)	-.031	.187**	.019	.036	.127	-.055	.063	
T_Practice (9)	-.082	.008	.089	.159*	.139*	-.045	.102	.456**

p<.05, * p<.01

Note: T_understanding= teachers' understanding, T_practice=teachers' FA practice
The collinearity diagnostics (Table 6) revealed acceptable levels of multicollinearity, with tolerance values ranging from .799 to .965 (values greater than .01) and conversely variance inflation factors (VIF) ranging from 1.036 to 1.252 (all less than 10).

Table 6

Coefficients and multicollinearity statistics for the regression model

Model		Unstandardized Coefficients		Standardized Coefficients	T	Sig.	Collinearity Statistics	
		B	Std. Error	Beta			Tolerance	VIF
1	(Constant)	36.808	5.728		6.426	.000		
	Teaching experience	-.083	.888	-.006	-.093	.926	.799	1.252
	Instrumental attitude towards FA	-2.439	2.132	-.073	-1.144	.254	.881	1.135
	Class size	1.671	2.037	.050	.820	.413	.964	1.037
	Internal school support	3.359	1.854	.114	1.812	.071	.906	1.104
	Assessment type required	1.502	1.734	.053	.866	.387	.965	1.036
	Weekly workload	-.676	1.650	-.026	-.410	.682	.912	1.097
	Adequacy of time	1.829	1.714	.068	1.067	.287	.881	1.136
	T_Understanding	1.113	.147	.473	7.582	.000	.925	1.081

a. Dependent Variable: T_Practice

Note: T_understanding= teachers' understanding; T_practice= teachers' practice

The R-squared value (shown in Table 7) was .264, signifying that 26.4% of the variance in FA practice was explained by the model. According to Cohen (1988), this constitutes a large effect. Regression results summarized in the ANOVA part of the same table (Table 7) revealed that the overall model was statistically significant, $F(9, 213) = 8.14$, $p < .001$.

Table 7

Regression Model Summary and ANOVA

R	R^2	Adjusted R^2	Std. E.E		Sum of Squares	df	Mean Square	F	Sig.
0.514	0.264	0.232	11.41	Regression	9533.428	9	1059.270	8.14	0.000b
				Residual	26557.773	204	130.185		
				Total	36091.201	213			

Predictors (Constant): teachers' understanding of FA, teaching experience, weekly workload, average number of students, assessment type required, usefulness of FA, internal school support, adequacy of instructional time.

Dependent Variable: Teachers' practice (T_Practice)

Findings from interviews conducted with key informants (with school principals and vice principals) to gather more data on teachers' FA practice and related challenges are summarized as follows.

Reflecting on the attempts to implement some FA techniques, such as oral questioning, work sample presentations, and supporting students in ability groupings, a vice principal from IPS1 (IPS=for interviewed primary school) mentioned factors affecting FA practices such as large class sizes, a lack of understanding of the principles and strategies of FA from the side of the teachers, a shortage of textbooks, and teaching school subjects not trained for (including teaching unrelated subjects). He further stressed that almost all teachers teaching in the Amharic as a medium of instruction program were not trained for that particular curriculum. The newly introduced school

subjects, such as Geda education, Safu education, technical, and career education, and visual and performing arts, did not have trained teachers and were taught by any randomly assigned teachers, he added.

A female vice principal from IPS2 outlined the challenges to FA practices as a large class size of up to 1:120 (which coincided with the observation data), an acute shortage of textbooks, and a shortage of in-service training for teachers. She also suggested ways out, including strengthening CPD, reducing class size, and strengthening internal school support and collaboration.

Similarly, IPS3's principal raised related challenges like a shortage of awareness, an acute shortage of instructional materials, especially textbooks (with an average book-to-student ratio of 1:5), and the presence of students who are unfamiliar with the medium of instruction.

Furthermore, a focus group discussion (FGD) conducted with selected 6^{th}, 7^{th}, and 8th-grade students from one primary school revealed that teachers used various assessment strategies and supported slow learners. Some teachers clarify learning objectives and use self- and peer-assessment methods. A recurring issue noted across data sources was that teachers were often assigned to subjects outside their area of expertise.Furthermore, FGD participants said they did not practically learn IT due to a lack of electric power service and computers.

Discussion

The study reveals significant diversity in teachers' FA techniques, with most teachers self-reporting proficient to advanced levels.

Classroom observation conducted in selected primary schools also revealed a considerable difference in teachers' FA practice, ranging from poor practice (no clarification of the learning intentions and success criteria to students, insufficient elicitation of evidence of learning, absence of feedback given to students, and no activation of students for self- and peer-assessment) to a relatively better practice of FA (implementation of some of the key strategies of FA, except the limitations in the use of elicited evidence to adjust the instruction).

In sum, classroom observations indicated predominantly unsatisfactory FA practices, including failure to clarify learning objectives, lack of formative feedback to students, and minimal use of peer/self-assessment. The present result is consistent with the majority of findings in earlier research, which showed a mixed picture of a) depending on traditional summative assessment practices (Chroinín & Cosgrave, 2013; Hailay & Abate, 2022; Poomoney & Govender, 2020; Sintayehu, 2016; Vlachou, 2018; Wong, 2014) and b) using some of the FA strategies infrequently (Kanjee, 2020; Askalemariam, 2015; Grob et al., 2019; Jambo et al., 2020; Johnson, 2019; Martin et al., 2022; Singh et al., 2022; Tuba & Yıldı, 2018; Vingsle, 2014;). Other research findings, such as those from Wang (2017) (from China) and Tolley, 2022 USA), reported praiseworthy FA practices, which ran counter to the current findings. Moreover, prediction of the teachers' FA practice from teachers' instrumental attitude towards FA, understanding of FA, pre-service training on assessment and evaluation, teaching experience, internal school support, adequacy of instructional time, class size,

assessment policy, and workload indicates a varying degree of contribution of the independent variables listed.

It was found that FA understanding is the most significant predictor of FA practice, while internal school support and adequacy of instructional time are also factors having a relatively higher association with FA practice. However, other variables such as teaching experience, attitude towards FA, and assessment type required showed non-significant associations. In general, 26.4% of the variance in FA practice was explained by the regression model, which indicates a large effect according to Cohen's criteria.

Interviews with school principals and vice principals revealed challenges to FA practices, consisting of large class sizes, in-service training shortages, teaching subjects not trained for, a lack of sufficient knowledge base of FA, and resources (including textbooks). Students' perspectives have also been added through focus group discussions (FGD), which reveals teachers' teaching of subjects they were not skilled in, and a lack of electricity and computers hindering practical IT classes. Therefore, in addition to the factors identified as significantly affecting FA practices (understanding, internal school support, and assessment related school policy) through the quantitative analysis, the qualitative exploration also disclosed further challenges to the implementation of FA.

Several previous study findings are consistent with the findings of the current study. For instance, in Zi Yan et al.'s (2021) systematic review, it was reported that teachers' implementation of FA may differ due to their different levels of understanding to apply FA techniques even with similar support and in the same environment.

Support and encouragement from principals, the school management team, and head teachers were found to be influential factors that motivate teachers to implement FA in previous studies (Brink & Bartz, 2017; Moss & Brookhart, 2019), consistent with the findings of the present study, which identified a considerable impact of internal school support for teachers. The non-significant impact of time constraints on FA implementation identified in the present study disagrees with the results of previous studies (for instance, Crichton & McDaid, 2016), which indicated a significant influence of time constraints on FA practice. Time is a less serious problem in the Sheger City schools, most probably because schools (i.e., teachers) in the city are working with the same learners both in the morning and the afternoon sessions (with no shift), unlike other schools in most parts of Ethiopia that serve students either in the morning or in the afternoon shifts due to a shortage of space, which in turn leads to time constraints.

The findings of the current study about the effect of assessment-related school policy on FA practice align with earlier research evidence which revealed the negative effects of traditional forms of summative (high-stakes) assessment on the teaching and learning process and especially on FA (Shepard, 2000; Wiliam et al., 2004).

Unlike the quantitative (self-response) data analysis findings, which indicated a relatively lower contribution of class size to the teachers' FA practice, the qualitative data analysis revealed that class size was one of the major challenges to FA practice. The finding from the qualitative data is consistent with Tebeje and Abiyu's (2015) and Brown and Gao's (2015), studies which revealed that large student-teacher ratios were practical constraints for teachers to the implementation of FA because of the difficulties

of class management and time. Also, findings from the quantitative self-response data analysis showed that school policy, teachers' instrumental attitude, and teaching experience, as mentioned above, had non-significant relationships with teachers' FA practice. Regarding the influence of teachers' teaching experience on their FA practice, previous studies showed varying findings. For instance, Bol et al. (1998, as cited in Alkharusi, 2011) reported that most experienced teachers indicated the use of some FA strategies more often than the least experienced teachers. On the other hand, Zi Yan et al. (2021) reported that teachers with fewer years of teaching experience had more FAP, explaining that such teachers were more likely to be trained in the latest FAP or more open to experimenting with novel assessment methods, such as FAP. Therefore, further studies need to be conducted to examine the effects that years of teaching experience have on FA practice.

The results of this study about the power of teachers' instrumental attitude towards FA on its implementation contrasted to most previous study reports, which revealed that the more positive instrumental attitude teachers held regarding the desirable consequences of practicing FA, the more willingly they were to implement FA (Brink & Bartz, 2017; Dixon & Haigh, 2009). Likewise, weekly workload was reported to have no impact on teachers' FA practices, consistent with a study by Askalemariam (2015).

Nonetheless, it was noted that teachers' FA practices were significantly influenced by resources (computers, books, and other facilities), which is supported by the findings of Askalemariam's (2015) study on science teachers in secondary schools.

Conclusion and Implications

The findings show that there is a variation among teachers regarding how they practice FA, with most self-reporting as proficient or advanced users. However, classroom observations indicated mainly unsatisfactory practices of FA, such as failure to explain learning objectives, lack of formative feedback to learners, and very minimal use of peer/self-assessment. Discrepancies between self-reported data and observed practices suggest potential overestimation in self-assessments. Some better FA practices (e.g., in ALfA class) were a result of specific teacher training provided by Geneva Global-Ethiopia, which demonstrated the need for teachers' professional development. Therefore, from self-reports, teachers seem to be performing well, but observational data underscore the need for improved training and implementation support to bridge the gap between perceived and actual FA competency.

The findings of this study further reveal that differences examined across socio-demographic factors on FA practices indicated significant differences in practice scores based on factors such as teachers' level of understanding of FA, assessment-related school policy, and support received from leaders. Further exploration of challenges to successful implementation of FA, through key informants' interviews, revealed teachers' low level of understanding of the strategies of FA resulted from shortage of in-service training, a very large class size, teachers were often assigned to subjects outside their training, including unrelated disciplines.

Therefore, the findings suggest that targeted professional development that focuses on improving FA understanding is of paramount importance. More specifically, a critical education and professional development program is of great need for teachers

with a low level of understanding of FA principles and strategies, teachers who have not participated in in-service training, female teachers, and those in schools who mainly value the traditional summative assessment over FA. The specific areas of FA requiring attention are clarification of learning intentions and success criteria; provision of descriptive feedback; peer- and self-assessment strategies; and use of elicited evidence.

Overall, teacher preparation programs ought to support the mental shift in teachers and provide them with the tools they need, enabling them to view FA as an integral part of their regular instruction rather than an extra task that takes up valuable teaching time. Creating significant links between formative and summative evaluation should be a primary priority of teacher education colleges. Furthermore, intensive internal support from the instructional leaders and collaboration among teachers are notably crucial.

Above all, all stakeholders (school principals, teachers, parents, respective educational offices, regional education bureau (REB), federal ministry of education (MoE), and development partners) must collaborate to create a conducive environment that is supportive of FA by addressing challenges to effective FA implementation identified in the study (issues of class size, teachers teaching subjects they were not trained for, the absence of instructional materials—including textbooks, and barriers in the instructional language for the learners) and thereby enhance students' learning.

Declarations of Conflict of Interest Statement

No competing interests were disclosed'

References

Alkharusi, H. (2011). Teachers' classroom assessment skills: Influence of gender, subject area, grade level, teaching experience and in-service assessment training. *Journal of Turkish Science Education*, *8*(2), 39-48. Retrieved from https://www.researchgate.net/publication/256353761

Askalemaiam, A. (2015). *Teachers' practices of assessment for learning in science education at East Gojjam preparatory schools, Amhara Regional State, Ethiopia* [Unpublished doctoral thesis, University of South Africa].

Black, P., & Wiliam, D. (1998). Assessment and classroom learning. *Assessment in Education: Principles, Policy and Practice*, *5*(1), 7-74.

Black, P., Harrison, C., Lee, C., Marshall, B., & Wiliam, D. (2004). Working inside the Black Box: Assessment for learning in the classroom. *Phi Delta Kappan*, *86*(1), 8-21. https://doi.org/10.1177/003172170408600105

Black, P., & Wiliam, D. (2009). Developing the theory of formative assessment. *Educational Assessment, Evaluation and Accountability*, *21*(1), 5-31. https://doi.org/10.1007/s11092-008-9068-5

Bol, L., Stephenson, P. L., O'Connell, A. A., & Nunnery, J. A. (1998). Influence of experience, grade level, and subject area on teachers' assessment practices. *The Journal of Educational Research, 91*(6), 323-330. https://www.jstor.org/stable/27542176

Brink, M., & Bartz, D. E. (2017). Effective use of formative assessment by high school teachers. *Practical Assessment, Research and Evaluation, 22,* 1-10. https://doi.org/10.7275/p86s-zc41

Brookhart, S. M. (2007). Expanding views about formative classroom assessment: A review of the literature. *Educational Measurement: Issues and Practice, 26*(4), 3-13.

Brookhart, S. M. (2011). Educational Assessment Knowledge and Skills for Teachers. *Educational Measurement*: *Issues and Practice, 30* (1), 3-12. https://doi.org/10.1111/j.1745-3992.2010.00195.x

Brown, T., & Gao, L. (2015). Chinese teachers' conceptions of assessment for and of learning: Six competing and complementary purposes. *Cogent Education, 2*(1), 993836, 10.1080/2331186X.2014.993836

Cagasan, L., Care, E., Robertson, P., & Luo, R. (2020). Developing a formative assessment protocol to examine formative assessment practices in the

Philippines. *Educational Assessment, 25*(1), 1-15. https://doi.org/10.1080/10627197.2020.1766960

Carla, C., & Johnson, C. (2019). A Study of the Implementation of Formative Assessment in Three Large Urban Districts. *American Educational Research Journal*, (56)6, 2408-2438. https://doi.org/10.3102/0002831219842347

Chróinín, D., & Cosgrave, C. (2013). Implementing formative assessment in primary physical education: Teacher perspectives and experiences. *Physical Education and Sport Pedagogy, 18*(2), 219–233. https://doi.org/10.1080/17408989.2012.666787

Crichton, H., & Mcdaid, A. (2016) Learning intentions and success criteria: learners' and teachers' views. *Curriculum Journal, 27*(2), 190- 203.

DeLuca, C., Tupper, J., & LaPointe-McEwan, D. (2013). Pedagogies for pre-service assessment in education: Supporting teacher candidates' assessment literacy development. *The Teacher Educator, 48*(2), 128–142. https:// 10.1080/08878730.2012.760024

Dereje, B., Dereje, T., & Jeylan, W. (2022). Ethiopian Secondary School EFL Teachers'Classroom Assessment Conceptions and Practices from Activity Theory Perspectives. *East African Journal of Education Studies, 5*(1), 105-116. https://doi.org/10.37284/eajes.5.1.574

Dixon, H., & Haigh, M. (2009). Changing mathematics teachers' conceptions of assessment and feedback. *Teacher Development, 13*(2), 173-186.

Garfield, J., & Ben-Zvi, D. (2007). How students learn statistics revisited: A current review of research on teaching and learning statistics. *International Statistical Review, 75*(3), 372-396.

Grob, R., Holmeier K. & Labudde P. (2019). Analysing formal formative assessment activities in the context of inquiry at primary and upper secondary schools in Switzerland. *International Journal of Science Education, 43*(3). https://doi.org/10.1080/09500693.2019.1663453

Hailay, T., & Abate, D. (2022). Assessment for learning strategies: Amharic Language Teachers' Practice and Challenges in Ethiopia. *International Journal of Language Education, 6*(1), 42-50. https://doi.org/10.17507/jltr.0901.06

Hattie, J. & Timperely, H. (2007). The Power of Feedback. *Review of Educational Research, 77*(1), 81-112, https://doi.org/10.3102/003465430298487

Heritage, M. (2010). *Formative assessment and next-generation assessment systems: Are we losing an opportunity*? Council of Chief State School Officers. (pp1-23)

Irons, A. (2008). Enhancing learning through formative assessment and feedback. Routledge.

Janbo, G., Wudu M., & Mekuria A. (2020). The practice of formative assessment in Ethiopian secondary school curriculum implementation: The case of West Arsi zone secondary schools. *Educational Assessment, 25* (4), 262-279. *https://doi.org/10.1080/10627197.2020.1766958*

Kanjee, A., & Mthembu, T. (2015). Assessment literacy of foundation phase teachers: An exploratory study. *South African Journal of Childhood Education, 5*(1), 142-168.

Lawshe, C. H. (1975). A quantitative approach to content validity. *Personnel Psychology, 28*(4), 563-575. https://doi.org/10.1111/j.1744-6570.1975.tb01393.x

Leigh, M. T. (2022). *Valuing Teachers' Evaluative Thinking*: *The Role of Teacher Knowledge and Practice in Formative Assessment.* Research Issues in Contemporary Education, 4(1), 21-42. University of Louisiana at Lafayette. https://files.eric.ed.gov/fulltext/EJ1244681.pdf

Kozma, R. (2003). Technology and learning: Supporting the development of the knowledge society. In R. K. Sawyer (Ed.), *Cambridge handbook of the learning sciences* (pp. 77–91). Cambridge University Press.

Martin, A. B., Smith, C. D., & Johnson, E. F. (2022). Examining Elementary School Teachers' Perceptions of and Use of Formative Assessment in Mathematics. *International Electronic Journal of Elementary Education, 14*(3), 417-425. https://doi.org/10.26822/iejee.2022.253

Ministry of Education. (2020). *General education curriculum Framework..*

Moss, C. M., & Brookhart, S. M. (2019). *Advancing formative assessment in every classroom:* A guide for instructional leaders (2nd ed.). ASCD.

Ottevanger, W., Akker, J. v., & Feiter, L. d. (2007). *Developing Science, Mathematics and ICT Education in Sub-Saharan Africa: Patterns and Promising Practices.* World Bank.

Poomoney, S., & Govender, K. (2020). Insights into Grade 2 teachers' enactment of formative assessment in mathematics in selected priority schools in Gauteng. *South African Journal of Education, 40* (2), https://doi.org/10.15700/saje.v40ns2a1780

Popham, W. J. (2008). *Transformative assessment*. ASCD.

Ruiz-Primo, M.A. (2006). Informal formative assessment and scientific inquiry: Exploring teachers' practices and student learning. *Educational Assessment, 11*(3-4), 205-235.

Shepard, L. A. (2000). *The Role of Classroom Assessment in Teaching and Learning.* The Regents of the University of California.

Singh, C. K. S., Ahmad, R., & Thompson, L. (2022). Exploring ESL teachers' alternativeassessment strategies and practices in the classroom. *Journal of Language and Linguistic Studies, 18*(1), 411-426. https://doi.org/10.52462/jlls.191

Sintayehu, T. (2016). The Practice of Continuous Assessment in Primary Schools: The Case of Chagni, Ethiopia. *Journal of Education and Practice, 7*(31), 24-30.

Stiggins, R. (2002). Assessment Crisis: The Absence of Assessment for Learning. *Phi Delta Kappan Internationa, 83*(10), 758-765. https://doi.org/10.1177/003172170208301010

Tebeje, M., & Abiyu, A. (2015). Improving the Implementation of Formative Continuous Assessment at College of Agriculture, Wolaita Sodo University, Ethiopia. *Journal of Education and Practice,6* (19), 24-28.

Tuba, A. E., & Hülya Y. (2018). Classroom Assessment Practices of Teachers in Turkey. *International Journal of Instruction, 11*(3). 587-602. https://doi.org/10.12973/iji.2018.11340a

Vingsle, C. (2014). *Formative assessment: Teacher knowledge and skills to make it happen*. Umea University.

Vlachou, M. (2018). Classroom assessment practices in middle school science lessons: A study among Greek science teachers. *Cogent Education, 5*(1), Article 1455633. https://doi.org/10.1080/2331186X.2018.1455633

Wang, X. (2017). A Chinese EFL teacher's classroom assessment practices. *Language Assessment Quarterly, 14*(4), 312-327. https://doi.org/10.1080/15434303.2017.1393819

Wiliam, D. (2011). *Embedded formative assessment*. Bloomington, in: Solution Tree Press.

Wiliam, D., Lee, C., Harrison, C., & Black, P. (2004). Teachers Developing Assessment for Learning: Impact on Students achievement. *Assessment in Education: principles, policy and practice, 11*(1), 49-65.10.1080/0969594042000208994

Wong, M. W. Y. (2014). Assessment for learning, a decade on: Self-reported assessment practices of secondary school music teachers in Hong Kong. *International Journal of Music Education, 32*(1), 70–83. https://doi.org/10.1177/0255761413504706

Yan, Z., & Pastore, S. (2022a). Are teachers literate in formative assessment? The development and validation of the Teacher Formative Assessment Literacy Scale. *Studies in Educational Evaluation*, Article 101183. https://doi.org/10.1016/j.stueduc.2022.101183

Yan, Z., & Pastore, S. (2022a). Are teachers literate in formative assessment? The development and validation of the Teacher Formative Assessment Literacy Scale. *Studies in Educational Evaluation.* https://doi.org/10.1016/j.stueduc.2022.101183101183

Yan, Z., & Pastore, S. (2022b). Assessing teachers' strategies in formative assessment: The Teacher Formative Assessment Practice Scale. *Journal of Psychoeducational Assessment*. Advance online publication. https://doi.org/10.1177/07342829221075121

Yidenek, T. (2018). Exploring EFL Teachers' an Integrated Approach to Classroom Assessment Practices: A secondary School in Focus (Unpublished master's thesis). Debre Berhan University.

Zi, Y., Cheng, Y., & Wong, K. (2021). Assessment for learning in the Hong Kong assessment reform: A case of policy borrowing. *Studies in Educational valuation*, 68, 100985. *https://doi.org/10.1016/j.stueduc.2021.100985*

Zi, Y., Chiu, M., & Cheng, K. C. (2022). A systematic review on factors influencing teachers' intentions and implementations regarding formative assessment. *Assessment in Education: Principles, Policy & Practice*. Advance online publication. https://doi.org/10.1080/0969594X.2021.1884042