The influence of K-12 language teachers' demographic profiles on authentic assessment practice

Hailay Tesfay Gebremariam[*] **& Abate Demissie Gedamu**

[1]*Department of Ethiopian Languages and Literature (Amharic), College of Social Science and Humanities, Arba Minch University, Arba Minch, Ethiopia*
[2]*Department of English Language and Literature, College of Social Science and Humanities, Arba Minch University, Arba Minch, Ethiopia*

ABSTRACT
While much literature has been used to investigate authentic assessment practice in the field of education, particularly in the context of language classrooms, a study investigating the demographic profiles of teachers and their relation to authentic assessment practice in language classrooms is crucial and needs further exploration. To this end, the current study investigated the K-12 language teachers' demographic profiles in relation to their authentic assessment practice, in terms of gender, teaching experience, and educational status. The study included 282 K-12 teachers enrolled in a summer program in Ethiopia. The study design was a descriptive survey, and data were obtained via a questionnaire. Descriptive and inferential statistics were used to analyze the gathered data. Results showed that language teachers had significant differences ($p > .05$) in classroom questioning and discussion (hereafter referred to as CQD), as well as substantial differences ($p > .05$) in peer and self-assessment (hereafter referred to as PSA) and feedback process (hereafter referred to as FB). Additionally, significant differences ($p < .05$) were observed in the effect of gender, teacher experience, and educational status on all three subsections. The findings show that although language teachers had authentic assessment practice, there were no consistent results regarding the influence of their background information. These findings contribute to our understanding of how demographic profiles of language teachers affect authentic assessment practice in K-12 education. Based on the results of the study, recommendations have been made for further consideration.

Keywords: Authentic assessment; demographic variables; K-12 language teachers; language education; teachers' practice

INTRODUCTION

In education, authentic assessment is used to encourage lifelong learning (Kangaslampi et al., 2022). This type of assessment can help students in acquire knowledge and ideas that will last a lifetime, and it shapes people's attitudes towards teaching (Fanrong & Bin, 2022; Gebremariam & Gedamu, 2022; Swatevacharkul & Boonma, 2021). If students approach their learning through authentic assessment, they can enhance their capacity for self-regulation, self-evaluation, and metacognitive skills (Farrell, 2017; Popham, 2017). To measure learning

objectives, teachers must consider their own demographic variables in relation to their authentic assessment practice (Farrell, 2017). Meaningful assessment cannot be separated from learning and teaching; teachers use standards to evaluate their own educational approaches and instruction implementation processes (Boud et al., 2018; Capan et al., 2020; Winarso, 2018). Students assess their own and their peers' learning abilities and provide feedback (Hichour, 2022; Ngui et al., 2022). They are also expected to take a problem-solving approach to their learning (Hussain et al., 2021;

Tesfay, 2017; Vogt et al., 2020). To facilitate this, there is a need for training that allows teachers and students to compare their views of the learning context to the learning lesson's aims and objectives (Lysaght & O'Leary, 2013).

Authentic assessment researchers (e.g., Mirian & Zulnaidi, 2020; Moradian et al., 2021; Nieminen et al., 2022; Onalan & Gursoy, 2020; Wyatt, 2018; Zhang et al., 2022) suggest that gender, educational status, and teaching experience should be taken into account when investigating teachers' assessment practices. For instance, Mirian and Zulnaidi (2020) and Zhang et al. (2022) reported gender and teaching experience differences in teachers' assessment practices, while Hichour (2022) found no gender differences. Additionally, Alotaibi (2019), Karas (2019), Nasr et al. (2018), Onalan and Gursoy (2020), Wyatt (2018), and Zhang et al. (2022) reported gender differences in teachers' classroom assessment practices and environment, while other researchers (e.g., Nourdad & Banagozar, 2022; Qadi, 2021; Veyis, 2020) did not detect any gender differences. Asare (2021) found no statistically significant interaction effects of educational status, gender, and teaching experience on teachers' formative assessment practices; however, a statistically significant main effect for gender was found. In contrast, educational status and teachers' experience did not influence teachers' formative assessment practices in the classroom. These findings suggest that the research is inconclusive regarding gender effects, which may be due to inconsistencies in the measurement of either the teachers' assessment practices or the gender variables. Nevertheless, past researchers have suggested that gender, teaching experience, and educational status may play a role in teachers' authentic assessment practices.

Recent related studies (e.g., Asare, 2021; Gebremariam, 2023; Nasr et al., 2018; Veyis, 2020) have focused on the influence of language teachers' demographic variables on their assessment practices. It is believed that language teachers' authentic assessment practice is important for language instruction, and studies (e.g., Popham, 2017; Yan & Pastore, 2022) have suggested that teachers' background factors are critical for students' learning development. Therefore, exploring the authentic assessment practice of language teachers in the study area, particularly in the Game Zone, is important to gain deeper insights. In light of the importance of authentic assessment practice, this can be seen as an input that would improve educational quality and resilience construction (Gebremariam, 2023; Gedamu & Shewangezaw, 2020). The majority of K-12 language teachers who participated in the study are expected to have diplomas and first degrees from public and private colleges. Thus, the current study aims to explore whether K-12 language teachers' demographic variables are associated with their authentic assessment practices. To this end, a research question was set: Is K-12 language teachers' demographic profile associated with authentic assessment practice? The authors believe that the study findings will help improve authentic assessment practices in real language classrooms by identifying teachers' gender, teaching experience, and educational status, and by highlighting potential challenges at the grassroots level.

In addition, some studies (e.g., Alotaibi, 2019; Karas, 2019; Nasr et al., 2018; Yan et al., 2022) have confirmed that it is important to understand how key demographic variables, including gender, educational status, and teaching experience, influence teachers' assessment practices. These teacher demographic variables are essential for us to appreciate the outcomes of all aspects of our educational structure. Concerning the influence of gender on assessment practices, previous studies have generated mixed findings. For instance, Yan et al.'s (2022) study on teachers' formative assessment practices revealed no statistically significant differences, while Ngui et al.'s (2022) work on formative assessment practices also discovered no statistically significant differences. However, Asare's (2021) and Brown's (2019) studies on the assessment practice of language teachers reported significant differences. Similarly, Benkirane et al.'s (2019) and Xue's (2022) reported significant differences based on the assessment practices. In terms of educational status, Van der Steen et al.'s (2022) work on teachers' classroom assessment skills revealed significant differences. Likewise, Yan et al.'s (2022) and Nasr et al.'s (2018) studies on assessment practices of teachers reported a significant influence of teachers' authentic assessment practice. Conversely, Asare's (2021) and Alotaibi's (2019) studies reported no significant differences. Regarding teaching experience, Asare's (2021) study reported no significant differences; however, Molloy et al.'s (2020) reported significant differences. These aforementioned studies tend to confirm that the influence of gender, educational status, and teaching experience may be unique depending on the study sample and geographical location.

The literature indicated limited studies and inconsistent results in terms of teachers' authentic assessment practices. For instance, in some studies (e.g., Karas, 2019; Wyatt, 2018; Xue, 2022), teachers' demographic variables were found to be a significant determinant of teachers' authentic assessment practice, while in others (e.g., Qadi, 2021; Zhang et al., 2022), they had no influence on their assessment practice. Similarly, research findings concerning authentic assessment practice (Brown et al., 2019; Veyis, 2020) on teachers' authentic assessment practices are inconclusive. A gap exists in the literature examining how language

teachers' demographic variables influence authentic assessment practice in the Ethiopian context. The purpose of the current study is to add to the existing literature by exploring the influence of K-12 language teachers' demographic profiles on authentic assessment practices in the Ethiopian context.

In the Ethiopian context, several studies (e.g., Brown, 2019; Gebremariam & Gedamu, 2022; Muianga, 2023; Subheesh & Sethy, 2020; Tesfay, 2017) have been conducted on assessment practices at K-12. However, these studies have failed to investigate the influence of key demographic variables such as gender, educational status, and teaching experience on teachers' authentic assessment practices. This gap in the literature needs to be filled. It is against this backdrop that this study seeks to investigate the influence of gender, educational status, teaching experience (Benkirane et al., 2019), and their various interactions on teachers' authentic assessment practices (Mussawy et al., 2021), particularly at the K-12 educational level. This is because K-12 seeks to lay a strong foundation for inquiry, creativity, and innovation, and lifelong learning in general, and provides building blocks for higher levels of education. Findings from this study would provide teachers and researchers with a better understanding of variables related to teachers' authentic assessment practices. The study's research hypothesis is that teachers' gender, educational status, teaching experience, and their various interactions will have no statistically significant influence on teachers' authentic assessment practices.

Some scholars contend that a complete theoretical foundation is not in place to support process-oriented authentic assessment practices. Thus, they attempt to propose different conceptual frameworks that can support the effective use of such assessments (Monteiro et al., 2021; Takele & Melese, 2022; Swatevacharkul & Boonma, 2021). Additional research into language teachers' authentic assessment can enhance the theory by creating a new link in the model of authentic assessment development in relation to teachers' demographic backgrounds. This might lead to the creation of specific assessment practices that can deliberately result among language teachers, and the findings of the current study might contribute to educational

practice theory in the Ethiopian context by exploring the teacher demographic profile and practice in real-world classrooms (Gebremariam, 2023). If the theory can be expanded this way, perhaps the introduction of authentic assessment practices could evolve further.

As a result, this study seeks to explore the influence of language teachers' demographic profile on authentic assessment practice at selected K-12 schools in Ethiopia. Primarily, this study aims to address the following research questions:

1. What level of authentic assessment practice do language teachers have?
2. Does teachers' demographic profile influence their authentic assessment practice?

METHOD

In line with the previously reviewed related literature, this study utilized a descriptive survey research design to address the research questions. This design offered sufficient opportunity for the sample to obtain spontaneous reactions for the interpretation and analysis of the phenomenon relevant to the focus of the study. The study identified teachers' authentic assessment practices in language classrooms in terms of gender, teacher experience, and educational status. Data was collected from 282 language teachers through an authentic assessment practice questionnaire to confirm the level of authentic assessment practice in terms of gender, teacher experience, and educational status.

Participants

Teachers enrolled in the summer programs at one Ethiophian University's first and second-degree programs participated in the study by providing information. Of the 1032 summer trainees enrolled in the university's first and second-degree programs, 282 teachers engaged in the study by offering comments. The participating teachers were chosen using the complete sample technique; 207 student teachers were randomly selected from the 977 trainees for the BA/ed degree, and 75 trainees for the second degree, depending on their year of study and class identity. Once permission was obtained and the purpose was explained, the constructed questionnaire was distributed to the chosen classes.

Table 1
Demographic profile of participants.

Attribute	Category	N	%
Sex	M	112	39.7
	F	170	60.3
Educational status	Diploma	169	59.9
	BA/Ed Degree	93	33.0
	MA/MEd degree	20	7.1
Teaching Experience	1-5 Years	48	17.0
	6-10 years	38	13.5
	Above 10 Years	192	68.1

Data Instrument

An authentic assessment practice questionnaire was employed to measure language teachers' authentic assessment practices in the K-12 language education context. The questionnaire was adopted from Lysaght & O'Leary (2017) and consisted of 58 items on a five-point Likert scale ranging from 5 (always applied) to 1 (very limited applied), in order to assess the learning process. Each option represented the application of the item from always applied (above 75%) to very limited applied (0-10%).

The measure was designed to test the awareness and educational implementation skills of teachers in their respective schools, both individually and collaboratively, taking background differences into account. The questionnaire from Lysaght & O'Leary (2017) was classified into four subsections; however, for this study, only three subdimensions were included, with 42 questions. The questionnaire was then translated into Amharic by two language teachers, and the equivalence of the two teachers' interpretations was evaluated by two other teaching experts. Based on the opinions of both evaluators, the questionnaire was compressed into one version and adapted to the Ethiopian education policy and the culture of teaching and learning applications. This adaptation was tested by 37 language teachers who were not included in the main data, and was then used as the data collection tool for the study.

The questionnaire was split into two sections. The first section provided demographic information on the participants, including gender, teacher experience, and educational status, which were independent factors of this study and thus included in the data and outcomes analysis. The second section provided a test of the teachers' authentic assessment practice. Three themes were used to group the questions: CQD (16 questions), FB (14 questions), and PSA (12 questions). The reliability level of the original questionnaire had a Cronbach Alpha value of 0.884, and the Amharic version had an internal consistency value of 0.874. For each subscale, reliability ratings were calculated using Cronbach's alpha; CQD (0.638), FB (0.707), and PSA (0.908). The results also showed that the propositions of the questionnaire were at an adequate level, indicating that the tool was reliable and measured the constructs it was intended to determine.

Data Analysis Techniques

The data collection procedure had multiple stages. Before obtaining the data, school teachers were first informed briefly about the relevance and context of the study and then asked for their consent to participate. Next, the questionnaire was administered to help the participants fill it out correctly.

After collecting the data from the language teachers, the quality of the data was checked and filtered; it was sorted by category and entered into the statistical database (SPSS-25). Descriptive statistics were then used to verify the normality of the data, and a one-sample t-test was calculated to confirm the level of authentic assessment practice of language teachers. Before using descriptive and inferential statistics, the data collected was checked against some basic assumptions of the statistical instruments used to analyze the data. The distribution of the scores of the quantitative data at items and scale levels showed a normal distribution since the skewness and kurtosis values were between +1.5 and -1.5, and there were no significant extreme outliers that could influence the mean scores for data analysis. Moreover, the Levene statistic test of homogeneity of variance for the subscales of the assessment practice showed no significant differences $(df(2,280) = 0.043, p > 0.05)$. Furthermore, the normality probability plots (Normal Q-Q Plots) showed straight lines that indicated normal distributions for the assessment practice.

Mean values, standard deviation, and ANOVA tests were applied to analyze the data obtained through the questionnaires. Specifically, standard deviation and mean scores at an item level and item aggregate mean values were employed to address assessment practice and subscales. Since mean values alone could not distinguish whether there were statistically significant differences among the mean values of the dimensions, the ANOVA test was employed; the Tukey HSD test analysis was then run to compare the mean scores. Finally, a five percent ($\alpha = 0.05$) significance level was employed throughout the study. In addition, the interpretation of the participants on their assessment practice used the common following scale indicated by Magulod (2019): 4.20-5.00 (Very High/Strongly Agree); 3.40-4.19 (High/Agree); 2.60-3.39 (Moderate/Undecided); 1.80-2.59 (Low/Disagree); and 1.00-1.79 (Very Low/Strongly Disagree).

Although the distribution of data was good, the sampling of participants did not use a uniform procedure. The sample sizes for gender, teacher experience, and educational status were also not comparable. Therefore, data was analyzed using non-parametric analysis due to non-compliance.

FINDINGS

The data collected from the language trainee teachers was analyzed through a questionnaire to answer the research questions. Thus, to answer the research questions, the data was analyzed as follows.

Language Teachers' Authentic Assessment Practice

To understand the level of language teachers' authentic assessment practice, data were analyzed to answer the first research question: "What level of authentic assessment practice do language teachers have?" Mean values and standard deviations were applied to analyze and determine the participants' assessment practice. Based on the data obtained from the questionnaire, the language teachers' practice of authentic assessment is provided in Table 2 as indicated below.

Table 2
Authentic assessment practice dimensions

Dimensions	N	Mean	SD	Level
CQD	282	3.41	0.50	High
FB	282	3.32	0.62	Moderate
PSA	282	3.28	0.73	Moderate
Total	282	3.34	0.49	Moderate

Table 2 shows the descriptive statistics of the language teachers' assessment practices regarding the three dimensions and scales. According to Magulod's (2019) cut-point of the level of assessment practice, the mean score for CQD (M = 3.41, SD = 0.50) was at a high level, followed by FB (M = 3.32, SD = 0.62) at a moderate level, and PSA (M = 3.28, SD = 0.73) at a moderate level. The overall scale mean value was M = 3.34 with a standard deviation of SD = 0.49. However, the mean values alone could not indicate whether there were statistically significant differences among the mean values of the four dimensions. To this end, an ANOVA test was conducted to determine if there were substantial variations in the teachers' ratings of the three dimensions of authentic assessment practice, as shown in Table 3.

Table 3
ANOVA summary of authentic assessment practice dimensions.

		Sum of Squares	Df	Mean Square	F	Sig.
QCD	Between Groups	1.602	1	1.602	4.956	0.027
	Within Groups	90.514	280	.323		
FB	Between Groups	.135	1	.135	0.359	0.549
	Within Groups	103.513	275	.376		
PSA	Between Groups	.510	1	.510	0.883	0.348
	Within Groups	160.541	278	.577		
Total	Between Groups	.780	1	.780	2.586	0.109
	Within Groups	84.439	280	.302		

As shown in Table 3, a one-way between-groups analysis of variance (ANOVA) was conducted to examine whether there were statistically significant differences among the mean scores of teachers' authentic assessment practice dimensions. The results showed that the assessment practice dimension mean scores of QCD statistically significantly differed (F (1, 280) = 1.602, p = 0.027), whereas the FB and PSA dimensions mean scores were not significantly different (F (1, 275) = 0.135, p = 0.549) and (F (1, 278) = 0.51, p = 0.348), respectively. The total value of the ANOVA analysis was not significantly different (F (1, 280) = 0.78, p = 0.109). Furthermore, the effect size (η2) of QCD was very low. This result indicates that only QCD had a difference in the teachers' ratings of specific dimensions compared to those of the two other dimensions. However, this result does not show which dimensions contributed significantly to the differences. To identify the dimensions that contributed significantly to the difference, post hoc comparisons of the dimensions with the Tukey HSD test were computed, as presented in Table 4.

Table 4
Tukey HSD dimensions multiple comparisons.

	Mean	1	2	3	4	Ranks
QCD	3.41	---	0.687**	0.541**	0.866**	3
FB	3.32		---	0.524**	0.841**	1
PSA	3.28			---	0.839**	2
Total	3.34				---	4
	**. Correlation is significant at the 0.01 level (2-tailed).					

The Tukey HSD test comparison of mean scores showed a statistically significant difference between the learning intention and success mean score (M = 3.31) and classroom environment (M = 3.41), QCD (M = 3.32) and FB (M = 3.37) scores, respectively. Similarly, there was a statistically significant mean score difference between the QCD (M = 3.41) and PSA (M = 3.28) dimensions, p < .01. Thus, the QCD (M = 3.41) dimension of authentic assessment practice had the highest impact, while the FB (M = 3.32) and PSA (M = 3.28) dimensions had equal weight placed second and third, respectively. According to Magulod (2019), the QCD (M = 3.41) scale showed high confidence while the FB and PSA scales fell in a moderate range of authentic assessment practice. Therefore,

the results showed that language teachers had moderate confidence in their assessment practice across the three dimensions.

Teachers' Demographic Profile and Their Authentic Assessment Practice

The second question of the research was "Does teachers' demographic profiles influence their authentic assessment practice?" This was to determine whether language teachers' demographic profile had an influence on authentic assessment practice in the Ethiopian context. The demographic variables considered included gender, teacher experience, and educational status. The data obtained from the close-ended questionnaire was calculated and presented as follows.

Table 5
Authentic assessment practice by demographic variables

Dependent	Experience	Edustatus	Sex	Mean	SD	N
Authentic assessment practice	1-5 years	Diploma	Male	3.85	0.52	24
			Female	3.86	0.28	12
		BA/Bed	Male	3.07	0.54	5
			Female	3.37	0.51	3
		MA/MEd	Male	3.17	0.07	2
	6-10 years	Diploma	Male	3.14	0.66	5
			Female	3.68	0.45	17
		BA/Bed	Male	3.17	0.96	5
			Female	3.24	0.32	9
		MA/MEd	Male	4.33	.	1
			Female	3.45	.	1
	>10 years	Diploma	Male	3.37	0.42	29
			Female	3.54	0.52	72
		BA/Bed	Male	3.10	0.56	28
			Female	3.33	0.45	42
		MA/MEd	Male	3.34	0.32	7
			Female	2.98	0.61	9
	Total	Diploma	Male	3.55	0.54	58
			Female	3.60	0.50	101
		BA/Bed	Male	3.11	0.60	38
			Female	3.32	0.43	54
		MA/MEd	Male	3.41	0.43	10
			Female	3.02	0.59	10

Table 5 shows descriptive statistics of the influence of language teachers' demographic profile on authentic assessment practice. The mean score of each category of demographic variables on authentic assessment practice based on the calculation (teaching experience*educational status*gender respectively) is as follows:

- 1-5 years*diploma*male, M = 3.85
- 1-5 years*diploma*female, x = 3.86
- 1-5 years*BA/BEd*male, M = 3.07
- 1-5 years*BA/BEd*female, M = 3.37
- 1-5 years*MA/MEd*male, M = 3.17
- 6-10 years*diploma*male, M = 3.14
- 6-10 years*diploma*female, M = 3.68
- 6-10 years*BA/BEd*male, M = 3.17
- 6-10 years*BA/BEd*female, M = 3.24
- 6-10 years*MA/MEd*male, M = 4.33
- 6-10 years*MA/MEd*female, M = 3.45
- >10 years*diploma*male, M = 3.37
- >10 years*diploma*female, M = 3.54
- >10 years*BA/BEd*male, M = 3.10
- >10 years* BA/BEd*female, M = 3.33
- >10 years* MA/MEd*male, M = 3.34
- >10 years* MA/MEd*male, M = 2.98

Besides, the total mean values were:
- experience*diploma*male = 3.55;
- experience*diploma*female = 3.66;
- experience*BA/BEd*male = 3.11;
- experience*BA/Ed*female = 3.32;
- experience*MA/MEd*male = 3.41;
- experience*MA/MEd*female 3.02.

The above mean scores, ranging from 3.40 to 4.19, are considered high, while scores ranging from 2.60 to 3.39 are considered moderate. However, the mean values alone cannot distinguish whether there are statistically significant differences among the demographic values. To this end, an ANOVA test was run to determine if there were substantial variations in the teachers' demographic profiles in terms of authentic assessment practice scores, as shown in Table 6.

Table 6

Tests of Between-Subjects Effects teachers' demographic profiles and their assessment practice

Dependent Variable: Assessment practice

Source	Type III Sum of Squares	Df	Mean Square	F	Sig.
Corrected Model	18.236[a]	16	1.140	4.538	0.000
Intercept	679.714	1	679.714	2706.355	0.000
Exprience	0.654	2	0.327	1.303	0.273
Educational status	3.868	2	1.934	7.699	0.001
Sex	0.010	1	0.010	0.041	0.839
Exprience * Eduucational status	2.698	4	0.674	2.685	0.032
Exprience * Sex	0.063	2	0.032	0.126	0.882
Educationa status * Sex	1.315	2	0.657	2.617	0.075
Exprience * Educational status *Sex	0.652	3	0.217	0.865	0.460
Error	65.551	261	0.251		
Total	3390.799	278			
Corrected Total	83.787	277			

a. R Squared = 0.218 (Adjusted R Squared = 0.170)

Table 6 shows that the tests of between-subjects effects regarding language teachers' demographic profiles and their assessment practice in language classrooms were examined using a univariate analysis to determine if there were statistically significant differences among the mean scores of teachers' authentic assessment practice dimensions. The corrected model of the statistical results was $F_{(2, 278)} = 4.538$, p = .000. In particular, the assessment practice dimensions total mean scores significantly differed in terms of teachers' experience ($F_{(2, 278)} = 1.303$, p = 0.273), educational status ($F_{(1, 278)} = 7.699$, p = 0.001), and sex ($F_{(1, 278)} = 0.041$, p = 0.839). Except in educational status, the statistical results of experience and sex were not significant.

Moreover, based on the three demographic profiles of language teachers, the inferential statistics of the dimensions were calculated, and showed as; experience with educational status ($F_{(4, 277)} = 2.685$, p = 0.032). In contrast, the mean scores related to experience with sex ($F_{(2, 278)} = .126$, p = .882), educational status with sex ($F_{(2, 280)} = 0.084$, p = 0.772), and between the three dimensions experience with educational status, experience with sex and educational status with sex ($F_{(3, 279)} = 0.865$, p = 0.460) were not significant. Furthermore, this result does not show which demographic variables contributed significantly to the differences. To identify the dimensions that contributed significantly to the difference, post hoc comparisons of the dimensions with the Tukey HSD test were computed, as presented in Table 7.

Table 7

Correlations between the demographic profiles of language teachers and assessment practice

		QCD	FB	PSA	Exprience	Edustatus	Sex
QCD	Pearson Correlation	1					
	Sig. (2-tailed)						
	N	282					
FB	Pearson Correlation	0.687**	1				
	Sig. (2-tailed)	0.000					
	N	277	277				
PSA	Pearson Correlation	0.541**	0.524**	1			
	Sig. (2-tailed)	0.000	0.000				
	N	280	275	280			
Exprience	Pearson Correlation	-0.173**	-0.134*	-0.257**	1		
	Sig. (2-tailed)	0.004	0.027	0.000			
	N	278	273	276	278		
Educational status	Pearson Correlation	-0.240**	-0.221**	-0.295**	0.161**	1	
	Sig. (2-tailed)	0.000	0.000	0.000	0.007		
	N	282	277	280	278	282	
Sex	Pearson Correlation	0.132*	0.036	0.056	0.215**	-0.060	1
	Sig. (2-tailed)	0.027	0.549	0.348	0.000	0.315	
	N	282	277	280	278	282	282

**. Correlation is significant at the 0.01 level (2-tailed).
*. Correlation is significant at the 0.05 level (2-tailed).

Regarding Table 7, the correlations between the assessment practice dimensions and the teachers' demographic variables within and between subjects were calculated. The mean scores of FB and QCD (n = 282; r = 0.687**), PSA and FB (n = 277; r = 0.524**), PSA and QCD (n = 277; r = 0.541**), QCD and experience (n = 280; r = -0.173**), experience and FB (n = 275; r = -0.134*), experience and PSA (n = 280; r = -0.257**), educational status and QCD (n = 278; r = -0.240**), educational status and FB (n = 273; r = -.221**), educational status and PSA (n = 276; r = -0.295**), educational status and experience (n = 278; r = 0.161**), sex and QCD (n = 282; r = 0.132*), sex and FB (n = 277; r = 0.036), sex and PSA (n = 280; r = 0.056), sex and experience (n = 282; r = 0.215**) and sex and educational status (n = 282; r = -0.060) were all significantly correlated, except for the correlations between sex and FB (r = 0.036), sex and PSA (r = .056), and sex and educational status (r = -0.060) of language teachers. Additionally, seven correlational results showed a negatively statistically significant difference between the authentic assessment practices (QCD, FB and PSA) and the demographic variables (experience, educational status, and sex). In contrast, a positive correlation, although not significantly different, was observed between the authentic assessment practice and gender.

DISCUSSION

In education, particularly in language education, authentic assessment has received growing interest in recent decades due to its potential to enhance students' learning (Benkirane et al., 2019; Jiang, 2020). However, it heavily relies on teachers' knowledge and practice, and thus, K-12 language teachers' assessment for learning practice is a fundamental factor that contributes to the effectiveness of assessment for learning practice (Gebremariam & Gedamu, 2022, 2023; Takele & Melese, 2022). While there is sufficient support for teachers to design authentic assessment activities, no guidelines exist to help them tie these different activities in language classrooms (Van der Steen et al., 2022). Furthermore, teachers spend 30% to 50% of their time assessing students' learning improvement (Mussawy et al., 2021; Sumardi & Guci, 2023; Vogt et al., 2020). Despite its paramount significance to language education, this empirical study explored the influence of K-12 language teachers on authentic assessment practice in the Ethiopian context.

The results of the analysis of the first question revealed that the level of difference in the cut-point (Magulod, 2019) of the authentic assessment practice in all three implementation subsections was only significant in one (QCD) (P = 0.027), and in the rest two dimensions there was no significant

difference (p > 0.05). This indicates that when the results of language teachers' authentic assessment practice level are compared with previous studies (Gebremariam & Gedamu, 2023; Hussain et al., 2021; Kangaslampi et al., 2022; Mohamed et al., 2021), it seems to have a complementary effect. Hussain et al. (2021), in particular, assert that authentic assessment evaluates students' skills in the application of course material rather than their theoretical comprehension.

However, Monteiro et al. (2021) concluded that most teachers conceive assessment as being for learning improvement, while their assessment practice is concerned with students' accountability. Their result was associated with the teachers' misconception of authentic assessment because their practice is not matched with the concept of assessment for learning. Moreover, Asare (2021), Nasr et al. (2018), Yan et al. (2022), and Sultana (2019) explain that assessments are practical and measure the effectiveness of educational implementation. They state that knowledge of assessment practices is a priority so that teachers can understand the purpose and properly implement the practice. On the other hand, Capan et al. (2020) seem to have found the opposite in terms of the principles of assessment implementation. This, according to Tesfay (2017), has the implication that teachers use assessment to determine learning outcomes. According to Onalan & Gursoy (2020) and Fanrong and Bin (2022), teachers know the purpose of assessment; however, they do not implement it based on practice principles.

From the literature, authentic assessment is considered one of the most promising variables to support teacher decision-making and improve education and student learning (Gebremariam & Gedamu, 2023; Van der Steen et al., 2022). However, due to demographic factors and teachers' understanding issues, authentic assessment does not always meet these expectations. According to Muianga (2023), conceptions of authentic assessment and assessment activities should be consciously and coherently planned, aligned with other aspects of the curriculum and the students' and teachers' demographic profiles.

With regard to the second question, concerning the authentic assessment practice of language teachers based on their gender, teaching experience, and educational status, the data obtained from the questionnaire was analyzed. Thus, it was understood that there is a significant (p < 0.05) difference in educational status. However, in terms of gender and experience observed, there was no significant difference (P ≥ 0.05).

In the second question, we show that, as described in the studies of Gebremariam and Gedamu (2023), and Kangaslampi et al. (2022), the authentic assessment practice is seen from the perspective of gender, teaching experience, and

educational status differences. According to Lysaght & O'Leary (2017), teachers' ability to teach is related to teaching experience and educational status; those with better experience and higher levels of education are likely to have greater views and skills. Thus, it is partially in line with the conclusion reached by Brown et al. (2019). The previous studies (Brown, 2019; Brown et al., 2019; Fitriyah et al., 2022) did not show significant differences in teachers' background information. Most teachers use authentic assessment practice only to measure results. This suggests that compared to the results of this study, teachers do not strive to have an authentic assessment practice; their main focus is on collecting student results.

Of course, the results of this study differ from those of Boud et al. (2018), Höltge et al. (2019), and Hussain et al. (2021), as they show a significant difference in the total scores of all three backgrounds. Although this was not consistent across subcomponent scores, the total score showed a difference across all authentic assessment applications. Furthermore, language teachers are apprehensive about using assessment practice in classrooms for learning improvement (Brown, 2019). This study's findings thus support previous studies (Nieminen et al., 2023; Nourdad & Banagozar, 2022), which found that many teachers lack the skill and knowledge of assessment practices in language classrooms and prefer to begin practicing the same traditional approaches to learning (Fitriyah et al., 2022; Hichour, 2022; Höltge et al., 2019). Gebremariam and Gedamu (2023) revealed that teachers are inadequately trained in classroom implementation assessment. Teachers' literacy in assessing learning practices is limited (Sultana, 2019).

Several literatures have been used to investigate assessment of learning in the context of language classrooms (Mohamed et al., 2021; Molloy et al., 2020; Sumardi & Guci, 2023). This study presents important conceptual frameworks systematically that serve as a foundation for exploring authentic assessment practice to achieve the desired goal. Hence, this study examined language teachers' authentic assessment practices, teachers' demographic profiles, and the justification of the current study. To sum up, the main objective of this study was to understand the influence of language teachers' demographic profiles on authentic assessment practice; it was examined in terms of gender, teaching experience, and educational status. To test the main objective of the study, two questions were formulated and data were collected from 282 language teachers. The results obtained from the research data were analyzed using descriptive and inferential analytical methods. The results obtained from the data analysis showed that the teachers had high and moderate practices of all subsections of authentic assessment practice. In

terms of gender, teaching experience, and educational status, although no consistent level of significance was observed in educational status only; the other two dimensions did not show any significant (P ≥ 0.05) differences. In terms of total scores, no significant (p < 0.05) differences were observed.

CONCLUSION

The current study investigated the influence of teachers' demographic profiles on authentic assessment practices among Ethiopian K-12 language teachers, and the findings provide unique and compelling evidence in the Ethiopian context regarding the study variables: language teachers' demographic profiles and authentic assessment practices. It is evident that language teachers have adopted authentic assessment practices, likely due to their focus on the implementation process of learning objectives for assessment, rather than the theory. To verify this supposition, further investigation is required. Additionally, background data shows that educators' perceptions of authentic assessment practices are partially influenced by gender, teacher experience, and educational status, further highlighting the need for ongoing in-service training for teachers.

In general, despite the findings regarding the demographic profiles and authentic assessment practices of language teachers, there were some limitations. Suggestions for future research are thus employed. The views of authentic assessment practice cannot be verified immediately; it is necessary to conduct continuous studies with different designs and approaches. Additionally, prospective teachers graduating from teacher education institutions at different times should be assigned to different schools after completing their training. In addition to the theoretical lessons given by these teachers in educational institutions, it is beneficial to conduct a series of studies and trainings on authentic assessment practice. Specifically, trainees should be equipped with knowledge and practice by providing them with appropriate authentic assessment and practice education, thus enabling them to independently recognize and apply it.

REFERENCES

Alotaibi, K. A. (2019). Teachers' perceptions on factors influence adoption of formative assessment. *Journal of Education and Learning, 8*(1), 74-86. https://doi.org/10.5539/jel.v8n1p74

Asare, E. (2021). Basic school teachers' formative assessment practices: Influence of demographic variables. *International Journal of Current Approaches in Language,*

Education and Social Sciences, *3*(1), 57-68. https://doi.org/10.35452/caless.2021.3

Benkirane, L., Hamza, M., Sbihi, W. & Arabi, E. (2019). Perception of learning assessment methods by students at the end of their initial training at the Faculty of Dentistry of Casablanca. *Education Research International*, *12*, 1-6. https://doi.org/10.1155/2019/8463169

Boud, D., Dawson, P., Bearman, M., Bennett, S., Joughin, G., & Molloy, E. (2018). Reframing assessment research: Through a practice perspective. *Studies in Higher Education*, *43*(7), 1107–1118. https://doi.org/10.1080/03075079.2016.1202913

Brown, G. T. L. (2019). Is assessment for learning really assessment? *Frontiers in Education*, *4*, 64. https://doi.org/10.3389/feduc.2019.00064

Brown, G. T. L., Gebril, A., & Michaelides, M. P. (2019). Teachers' conceptions of assessment: a global phenomenon or a global localism. *Frontiers in Education*, *4*, 16. https://doi.org/10.3389/feduc.2019.00016

Capan, M. M., Lettner, S., Bäwert, A., Puttinger, C., & Holzinger, A. (2020). Pursue today and assess tomorrow-How students' subjective perceptions influence their preference for self- and peer assessments. *BMC Medical Education*, *20*(1), 479. https://doi.org/10.1186/s12909-020-02383-z

Fanrong, W., & Bin, S. (2022). Language assessment literacy of teachers. *Frontiers in Psycholpgy, 13,* 864582. https://doi.org/10.3389/fpsyg.2022.864582

Farrell, A. (2017). *Assessment for life-long learning*. Academic Practice, University of Dublin.

Fitriyah, I., Massitoh, F., & Widiati, U. (2022). Classroom-based language assessment literacy and professional development need between novice and experienced EFL teachers. *Indonesian Journal of Applied Linguistics*, *12*(1), 124-134. https://doi.org/10.17509/ijal.v12i1.46539

Gebremariam, H. T. (2023). Language teacher's experience and preference towards online learning platforms during Covid-19 pandemic. *Education Research International*, *2023*. https://doi.org/10.1155/2023/9932873

Gebremariam, H. T., & Gedamu, A. D. (2022). Assessment for learning strategies: Amharic language teachers' practice and challenges in Ethiopia. *International Journal of Language Education, 6*(2), 128-140. https://doi.org/10.26858/ijole.v6i2.20505

Gebremariam, H. T., & Gedamu, A. D. (2023). Primary school teachers' assessment for

learning practice for students' learning improvement. *Frontiers in Education*, 8, 1145195. https://doi.org/10.3389/feduc.2023.1145195

Gedamu, A. D. & Shewangezaw, L. G. (2020). Secondary school teachers' and students' perspectives on cooperative group teaching assessment challenges in Ethiopia. *African Journal of Teacher Education*, 9, 104-119. https://doi.org/10.21083/ajote.v9i0.6083

Hichour, H. (2022). Teachers' Experience in E-assessment: Case Study of EFL Teachers in Algerian Universities. *Arab World English Journal*, *13*(1), 450-461. https://doi.org/10.24093/awej/vol13no1.29

Höltge, L., Ehm, J., Hartmann, U., & Hasselhorn, M. *(2019).* Teachers' self-efficacy beliefs regarding assessment and promotion of school-relevant skills of preschool children. *Early child development and care, 189*(2019). 339-351. https://doi.org/10.25656/01:17607

Hussain, S., Idris M., & Akhtar, Z. (2021). Perceptions of teacher educators and prospective teachers on the assessment literacy and practices. *Gomal University Journal of Research*, *37*(1), 71-83. https://doi.org/10.51380/gujr-37-01-07

Jiang, Y. (2020). Teacher classroom questioning practice and assessment literacy: case studies of four English language teachers in Chinese universities. *Frontiers in Education*, *5*, 23. https://doi.org/10.3389/feduc.2020.00023

Kangaslampi, R., Asikainen, H., & Virtanen, V. (2022). Students' perceptions of self-assessment and their approaches to learning in university mathematics. *LUMAT: International Journal on Math, Science and Technology Education*, *10*(1), 1-22.

Karas, M. J. (2019). English language teacher self-efficacy beliefs. *Electronic Thesis and Dissertation Repository*. 6091. https://ir.lib.uwo.ca/etd/6091

Lysaght, Z., & O'Leary, M. (2013). An instrument to audit teachers' use of assessment for learning. *Irish Educational Studies*, *32*(2), 217-232. https://doi.org/10.1080/03323315.2013.784636

Lysaght, Z., & O'Leary, M. (2017). Scaling up, writ small: Using an assessment for learning audit instrument to stimulate site-based professional development, one school at a time. *Assessment in Education: Principles, Policy and Practice*, *24*(2), 271-289, https://doi.org/10.1080/0969594X.2017.1296813

Magulod, G. C. (2019). Learning styles, study habits, and academic performance of Filipino university students in applied science courses: Implications for instruction. *Journal of Technology and Science Education*, 9(2), 184–198.

Mirian, B. M., & Zulnaidi, H. (2020). Mathematics teachers' conceptions of assessment: Gender and academic qualification comparisons. *International Journal of Instruction, 13*(2), 239-252.

Mohamed, M., Abd Aziz, M. S., & Ismail, K. (2021). "Assessment for learning" practices amongst the primary school English language teachers: A mixed methods approach. *Journal of Social Science and Humanities*, 29, 1875–1900. https://doi.org/10.47836/pjssh.29.3.21

Molloy, E., Boud, D., & Henderson, M. (2020). Developing a learning-centered framework for feedback literacy. *Assessessment and Evaluation in Higher Education*, 45, 527–540. https://doi.org/10.1080/02602938.2019.1667955

Monteiro, V., Mata, L., & Santos, N. N. (2021). Assessment conceptions and practices: perspectives of primary school teachers and students. *Frontiers in Education*, 6, 631185. https://dx.doi.org/10.3389/feduc.2021.631185

Moradian, M. R., Miri, M., & Alamdar, P. (2021). The role of dialogic interaction in EFL writing assessment: A sociocultural perspective. *Indonesian Journal of Applied Linguistics*, 10(3), 677-686. https://doi.org/10.17509/ijal.v10i3.31754

Muianga, F. (2023). English language teachers' conceptions of assessment. *Frontiers in Education*, 7, 972005. https://doi.org/10.3389/feduc.2022.972005

Mussawy, S. A. J., Rossman, G., & Haqiqat, S. A. Q. (2021). Students' and teachers' perceptions and experiences of classroom assessment: A case study of a public university in Afghanistan. *Higher Learning Research Communications, 11*(2) 22–39. https://doi.org/10.18870/hlrc.v11i2.1244

Nasr, M., Bagheri, M. S., Sadighi, F., & Rassaei, E. (2018). Iranian EFL teachers' perceptions of assessment for learning regarding monitoring and scaffolding practices as a function of their demographics, *Cogent Education*, 5, 1, 1558916. https://doi.org/10.1080/2331186X.2018.1558916

Ngui, W., Pang, V., & Hiew, W. (2022). E-portfolio as an academic writing assessment tool in higher education: strengths and challenges.

Indonesian Journal of Applied Linguistics, 12(2), 556-568. https://doi.org/10.17509/ijal.v12i2.40122

Nieminen, J. H., Bearman, M., & Tai, J. (2023). How is theory used in assessment and feedback research? A critical review. *Assessessment and Evaluation in Higher Education*, 48, 77–94. https://doi.org/10.1080/02602938.2022.2047154

Nourdad, N., & Banagozar, M. A. (2022). The effect of e-portfolio assessment on EFL vocabulary learning and retention. *Indonesian Journal of Applied Linguistics*, 12(2), 466- 475. https://doi.org/10.17509/ijal.v12i2.44232

Onalan, O., & Gursoy E. (2020). Private school EFL teachers' perceptions on assessment: A case study in Turkey. *Porta Linguarem*, 33, 249-264. https://www.researchgate.net/publication/342515341

Popham, W. J. (2017). *Classroom assessment: What teachers need to know* (2nd ed.). Pearson Education, Inc.

Qadi, A. (2021). An exploratory study of EFL teachers' assessment conceptions at a Saudi university. *Arab World English Journal*, 12(3), 464-485. https://doi.org/10.24093/awej/vol12no3.32

Subheesh, N. P., & Sethy, S. S. (2020). Learning through assessment and feedback practices: A critical review of engineering education settings. *EURASIA,* 16, em1829. https://doi.org/10.29333/ejmste/114157

Sultana, N. (2019). Language assessment literacy; an uncharted area for the English language teachers in Bangiladish. *Language Testing in Asia*, 9(1), 1-14. https://doi.org/10.1186/s40468-019-0077-8

Sumardi, S., & Guci, R. I. (2023). HOTS-based language assessment literacy: Challenges and prospects in English language teaching. *Indonesian Journal of Applied Linguistics*, 12(3), 818-827. https://doi.org/10.17509/ijal.v12i3.44261

Swatevacharkul, R., & Boonma, N. (2021). Learner autonomy assessment of English language teaching students in an international program in Thailand. *Indonesian Journal of Applied Linguistics*, 10(3), 749-759. https://doi.org/10.17509/ijal.v10i3.31764

Takele, M., & Melese, W. (2022). Primary school teachers' conceptions and practices of assessment and their Relationships. *Cogent Education*, 9(1), 2090185.

https://doi.org/10.1080/2331186X.2022.20901
85

Tesfay, H. (2017). Investigating the practices of assessment methods in Amharic language writing skill context: The case of selected higher education in Ethiopia. *Educational Research and Reviews*, *12*(8), 488-493. https://doi.org/10.5897/ERR2017.3169

Van der Steen, J., van Schilt-Mol, T., van der Vleuten, C., & Joosten-ten, B. D. (2022). Supporting teachers in improving formative decision-making: Design principles for formative assessment plans. *Frontiers in Education*, *7*, 925352. https://doi.org/10.3389/feduc.2022.925352

Veyis, F. (2020). Investigation of the self-efficacy beliefs of Turkish language and literature teachers in practicing constructivist approach in terms of arious variables. *International Education Studies*, *13*(7), 47-56. https://doi.org/10.5539/ies.v13n7p47

Vogt, K., Tsagari, D., Csépes, I., Green, A., & Sifakis, N. (2020). Linking learners' perspectives on language assessment practices to teachers' assessment literacy enhancement (tale): Insights from four European Countries. *Language Assessment Quarterly*, *17*(4), 410–433. https://doi.org/10.1080/15434303.2020.177671 4

Winarso, W. (2018). Authentic assessment for academic performance: Study on the attitudes, skills, and knowledge of grade 8 Mathematics students. *Malikussaleh Journal of Mathematics Learning, 1*(1), 1-8. https://doi.org/10.29103/mjml.v1i1.5799

Wyatt, M. (2018). Language teachers' self-efficacy beliefs: A review of the literature (2005-2016). *Australian Journal of Teacher Education, 43*(4). 92-120. http://ro.ecu.edu.au/ajte/vol43/iss4/6

Xue, Y. (2022). The role of EFL teachers' self-efficacy and emotional resilience in appraisal of learners' success. *Frontiers in Psychol*ogy, *12*, 817388. https://doi.org/10.3389/fpsyg.2021.817388

Yan, Z., & Pastore, S. (2022). Are teachers literate in formative assessment? The development and validation of the Teacher Formative Assessment Literacy Scale. *Studies in Educational Evaluation, 74*(1), 101183. https://doi.org/10.1016/j.stueduc.2022.101183

Yan, Z., Chiu, M. M., & Cheng, E. C. K. (2022). Predicting teachers' formative assessment practices: Teacher personal and contextual factors. *Teaching and Teacher Education*, 114, 103718. https://doi.org/10.1016/j.tate.2022.103718

Zhang, X. S., Zhang, L. J., Parr, J. M., & Biebricher, C. (2022). Exploring teachers' attitudes and self-efficacy beliefs for implementing student self-assessment of English as a foreign language writing. *Frontiers in Psychology, 13*, 956160. https://doi.org/10.3389/fpsyg.2022.956160