Factors Influencing Assessment Conceptions among Basic School Teachers: A Multiple Analysis of Variance

Shani Osman

Department of Social Sciences, Tumu College of Education, Ghana

Abstract: **This study examined teacher variables that influence assessment conceptions of basic school teachers in the Sissala East Municipality of Ghana. Gender, academic levels, class assigned, teaching experience, age and assessment-based training were investigated to determine the effects of these individual variables on the teachers' conception of assessment. The data for this study were collected and analyzed using a strictly quantitative approach. Teachers Conception of Assessment III was used to collect data from 204 teachers. Using multiple analysis of variance (MANOVA), out of the six variables included in the study, the results revealed that the main effect for gender was significant, Wilks' Lambda = .95, *F* (4, 198) = 2.53, *p* < .05; multivariate η^2 = 0.05. A follow-up Bonferroni adjustment analysis, the dependent variables were verified individually, and the only variable to attain a difference in statistical significance was improvement assessment conception, F (1, 201) = 6.56, p = .011, partial eta squared = .03. An analysis of the mean scores revealed that females recorded marginally higher levels of improvement conception (*M* = 4.84, *SD* = .55) than males (*M* = 4.63, *SD* = .59). It is therefore recommended that stake holders should ensure the development of ways to improve the assessment literacy of especially male teachers.**

Keywords: **assessment, conceptions of assessment, Ghana, Multiple Analysis of Variance.**

I. INTRODUCTION

Teaching and assessment are inextricably linked. Assessment is the "gathering, interpreting, and using evidence of student learning to support teacher decision making in a variety of ways" (McMillian, 2018, p.14). On their part Okyere, Kuranchie, Larbi and Twene (2018) view assessment as an act of obtaining data about students' learning, analyzing and synthesizing information to advance the quality of their learning. Assessment can be explained as a system of interpreting student achievement information utilizing a number of approaches or practices (Brown, 2011). National Council for Curriculum and Assessment (NaCCA) (2019) explained assessment as a process of of gathering and analyzing learners' information and using it in the decision-making process to enhance learning. This suggest that there can be no perfect decision making about students' learning and effective teaching bereft of classroom assessment. This is due to the fact that assessment is an essential component of any educational process. It determines where learners are now and what level they have reached; it provides feedback on their learning; it diagnoses learners' developmental needs; and it allows for the planning of curricular, resources, and activities (Alderson, 2005).

Xu and Brown (2016) note that assessment conceptions signify "the belief systems that teachers have about the nature and purposes of assessment, and that encompasses their cognitive and affective responses" (p. 56 Brown (2006; 2008) opined that the assessment conceptions teachers hold can be subdivided into: (1) school accountability- making schools and teachers accountable; (2) student accountability- ensuring learners are accountable for their learning; (3) Improvement-enhancing teaching and learning and; (4) irrelevance - assessment is irrelevant and of no impact on the work learners and teachers.

The assessment conception of improvement is based on the concept of "assessment for learning" or "formative assessment" (Brown, Lake, & Matters, 2011). Thus, teachers view assessments as supporting the role of increasing their teaching and student learning (Brown, 2006). The primary goal of the assessment is to improve teaching and learning via the use of formative assessment techniques and strategies in order to offer the teacher with the relevant feedback. The student accountability assessment conception requires students to take charge of their learning independently in attaining the credentials needed to go on to various stages of education. The principal recipients of this type of assessment information are employers and parents. Hence, the student accountability conception implies that assessment is used for the check students' achievements according to pre-established norms (Moiinvaziri, 2015). The assessment conception of school accountability indicates that schools and instructors are responsible for the learning of their students. In this way, assessment is used to measure the success of schools and teachers by taking the two to account for the shortfalls of the achievements of their students. Moiinvaziri (2015) observed that the notion of school accountability refers "the use of assessment to see how well teachers or schools are doing in relation to the established standards" (p. 76). The irrelevance conception, also known as "anti-purpose of assessment" opposes the purpose of assessment and subscribes to the notion that assessment is pivotal in the work of students and teachers. This is founded on the observation that assessment is untrustworthy and so does not benefit teachers or students,

hence, it has to be discounted (Brown, 2004; Brown et al., 2011; Harris, 2008).

Research on classroom assessment indicate that teachers' beliefs regarding the purposes of assessment are influenced by several independent variables, such as the teacher' gender (Brown & Gao, 2015; Ndalichako, 2015), the teaching experience of the teacher (Brown & Gao, 2015; Sahikarakas, 2012; Vardar, 2010), and the teacher's exposure to professional assessment training (DeLuca, Chavez & Cao, 2013; Smith, Hill, Cowie & Gilmore, 2014) among others. There appear to be inadequate studies on how these factors shape teachers' conception of assessment in Ghana

Research Questions

The study was guided by the following research question;

1. To what extent do basic school teachers' assessment conceptions differ based on teacher variables (e.g.,

    i.   level of teaching,
    ii.   teaching experience,
    iii.   training in assessment,
    iv.   gender, and
    v.   age)?

II. LITERATURE REVIEW

Teacher Variables Influencing Teachers' Conceptions of Assessment

Studies on classroom assessment reveal that, assessment conceptions of teachers are affected by certain independent factors such as teacher' teaching experience, age, gender and exposure to professional training in assessment among others. Brown (2004) found that teacher variables like "teacher gender, years of training, years of experience, and the role in school were irrelevant to mean scale scores on the teachers' conceptions of assessment inventory" (p.311). Also, school variables like school locality (urban or rural) and schools' socio-economic status were immaterial to the teachers' conceptions of assessment.

An investigation by Vardar (2010) involving 414 teachers in Turkey using the TCoA-IIIA, found no statistically significant difference in teachers' conceptions of assessment based on in-service training and teaching subject. However, significant differences existed in teachers' conceptions of assessment with respect to undergraduate institution teachers attended and years of teaching experience. Again, Yetkin (2018) found a no statistically significant difference in teachers' conceptions among 204 prospective Turkish English teachers with regards to age, teaching experience and gender. In contrast, Sahikarakas (2012) found a significant difference in the assessment conception of Language teachers with respect to teaching experience. Teachers with more experience, see assessment in a negative light than their peers with less experience. According to Sahikarakas (2012), the differences are due to the experienced teachers highly valuing themselves

to a level that there is no need for them to obtain evidence of their teaching efficacy through assessment.

A study by Brown and Gao (2015) found differences in teachers' assessment conception with regards to teaching experience and gender. Male teachers and those with twenty and above years' experience endorsed the notion of assessment as a control mechanism to inspect and monitor students, teachers, and the school to ensure effective teaching and learning. Ndalichako (2015) discovered that more female teachers possessed a positive view of classroom assessment compared to their male counterparts. A statistically significant difference was found between male and female teachers about the use of assessment to promote and sustain teaching.

On his part, Benson (2014) investigated 6th to 8th-grade teachers' beliefs about assessment and discovered that gender was unrelated to teachers' assessment conceptions. He found that female and male teachers held similar beliefs on assessment. In connection with teaching experience, Benson (2014) found that older teachers (above 43 years) and younger teachers (25 to 30 years) held related beliefs regarding the irrelevance conception of assessment. Similarly, Daniels, Poth, Papile and Hutchison (2014) discovered that gender or level teacher was trained for has no impact on Canadian pre-service teachers' conception of assessment.

Mehrgan, Hayati and Alavi (2017) examined the influences of EFL teachers' age, teaching experience, gender and educational background on their beliefs of formative assessment. Findings from the study indicated that age had no statistically important impact on teachers' views about formative assessment. Also, gender did not impact on the teachers' views about formative assessment. However, teachers' teaching experience significantly influences their beliefs about formative assessment. Yidana and Anti Partey (2018) investigating the effects of Economics teachers' age, experience, and gender on their conception of assessment, revealed that age and gender did not affect teachers' assessment conceptions. Fulmer, Tan and Lee (2017) study found no significant relationship with teachers' teaching experience and their conception of assessment. Also, no significant statistical differences existed between teachers according to the subject area or by the school.

Izci and Caliskan (2017) employing an action research method explored the influence of participating in an assessment course "Assessment and Evaluation in Education" on 118 soon-to-be teachers' assessment conceptions. The results revealed that except for the irrelevance conception, teachers' attendance in an assessment course and secured in-depth knowledge of assessment did not significantly change their assessment conceptions of improvement, student accountability, and school accountability. Similarly, studies by Brown and Hirschfeld (2008), Levy-Vered and Alhija (2015) and Vadar (2010) discovered that having more training in assessment or attending an assessment course did not enhance teachers' assessment conceptions. Nonetheless, some studies like DeLuca, Chavez and Cao (2013), and Smith, Hill, Cowie and

Gilmore (2014) revealed that teachers' conception of assessment improved after getting regular professional assessment training.

Some research studies have shown that the perceptions of teachers regarding assessment differ with the level at which they teach (Brown et al., 2011; Remesal, 2007), however, it remains uncertain whether this variation is as a result of the arrangement and policies related to various levels of education (e.g., primary and secondary), or whether it is linked to previous notions of teaching, learning, and assessment (Bonner, 2016). A study by Remesal (2007) of primary and secondary teachers in Spain revealed primary school teachers were much more inclined to consider assessment is for instructional purposes. In contrast, secondary school teachers often held an 'accounting' conception of assessment in certifying student performance. These perceived differences have been attributed to policy discrepancies at both school levels (Remesal, 2011). Brown et al. (2011) noted similar differences among Queensland, Australia, primary and secondary teachers in that primary teachers' assessment conceptions leaned in the direction of improvement conception while secondary teachers' conceptions inclined in favour of student accountability. These disparities can be due to policy differences at the primary and secondary levels where, at the time, a comprehensive, externally controlled school-based assessment framework only existed in some subject areas at the upper secondary level. However, the difference could be related to other factors (Brown et al., 2011).

From the above review, it appears that the formation of teacher conception about assessment, in general, is complicated and differs by individual teacher factors, policy context, and student development level. There is a paucity of studies on how teacher factors shape teachers' conception of assessment in Ghana.

III. METHODOLOGY

The data for this study were collected and analyzed using a strictly quantitative approach. The study's population included 796 professional basic school teachers from Sissala East Municipality's nine circuits, of whom 260 were chosen for analysis using Krejcie and Morgan's (1970) sample size table. However, the questionnaire was completed and returned by 224 teachers, resulting in a return rate of 86 percent.

Brown (2006) Conceptions of Assessment III (TCoA-IIIA) Abridged Survey instrument was the main instrument used to collect data. This questionnaire focused on the conceptions of assessment of the teachers under (1) School accountability, (2) Students' accountability, (3) Improvement, and (4) Irrelevance. The questionnaire presented these items in a 6-point Likert scale (1 for strongly disagree, 2 for slightly disagree, 3 for slightly agree, 4 for moderately agree, 5 for mostly agree, and 6 for strongly agree). There was no reverse scoring in the instrument.

The data was analyzed using the Statistical Product for Service Solutions (SPSS) software program. The data were analyzed for missing values before descriptive statistics analysis could be performed. No missing values were detected. Earlier, before entering the data, all questionnaires with any omission especially in demographic variables were not entered. This resulted in 10 questionnaires being rejected from the 224 returned questionnaires.

Moreover, the data was explored to determine whether data was normal or non-normal distribution as the distribution of data calls for completely different analytical methods. Both, the Kolmogorov-Smirnov and Shapiro-Wilk normality tests revealed a non-normal distribution of data in three of the four scales with the exception of irrelevance conception which significant value was above the p – value of 0.05. According to Pallant (2016), the breach of the assumption of normality should not lead to significant problems with large enough samples (> 30 or 40); this means that we may use parametric analytical tools even when the data are not distributable normally (Elliott & Woodward, 2007). However, as these tests were not the only means of testing normality, the researcher concentrated on finding further evidence to fulfill this assumption. Statistics of skewness and kurtosis were translated to z-scores by dividing each statistic by their standard errors respectively. Values of the skewness and kurtosis z-scores greater than 1.96 are significant and would indicate a potential problem (Pallant, 2016; Tabachnick & Fidell, 2013). The student accountability was the only conception that violated this recommendation with a kurtosis of .26(SE = .34) and a skewness of .84 (SE = .17). The review of histograms with normal curves plot were indicated that the data was normally distributed with a slightly skewed direction to proceed with the analysis.

Furthermore, seven univariate outliers and four multivariate outliers were discarded in order to optimize test results for normality which resulted in remaining 203 cases for further analysis. Then, the scale was analyzed for reliability. The inventory Cronbach's alpha coefficient was estimated as 0.74. This result showed a reasonable degree of reliability for the inventory and its objects. After the normality and reliability analyses were verified, descriptive statistics were employed to analyze the data. Mean values for the sub-scales (student accountability, school accountability, improvement and irrelevance) by teacher variable were calculated and interpreted. For inferential Multivariate Analysis of Variance (MANOVA) Tests were conducted since there were more than one dependent variable. MANOVA was desirable to the Analysis of Variance (ANOVA) and the Independent Sample-t Test because, according to Pallant (2016) and Tabachnick and Fidell (2013), a MANOVA is conducted on variable means to guard against increasing the Type 1 error rate when a series of t-test or ANOVAs are performed. In such conditions, the likelihood of experiencing Type 1 error may be large; finding significant differences after multiple analysis even though in reality there were no statistically meaningful difference.

Prior to MANOVA data analysis, the data were explored for the purpose of assessing whether or not the data met any of MANOVA 's assumptions. In order to guarantee normality,

11outliers were omitted. Secondly, The Mahalanobis distance revealed no violation of multivariate normality. Thirdly, linearity analysis shows no serious violation of the assumption of linearity. The assumption of multicollinearity was fulfilled as the dependent variables are correlated at low to moderate range with each other (i.e., up around .8; Pallant, 2016). Additionally, Box's Test of Covariance Equality was performed to examine if the data contravenes the supposition of homogeneity of covariance matrices. Moreover, the Levene's Test of Error Variance Equality was performed to check the requirement of equality of variance. It is widely recognized that when the value Sig is bigger than .001, it does not flout the assumption that the variance matrices are homogeneous. The data were subjected to the Manova test once all assumptions were fulfilled. The results of the Multivariate Test and Wilks' Lambdas were then calculated, verified and interpreted if the dependent variable met all of the assumptions.

IV. RESULTS

The research question *"To what extent do basic school teachers' assessment conceptions differ based on teacher variables (e.g., level of teaching, teaching experience, training in assessment, gender, and age)?"* sought to determine the influence that teacher variables have on their assessment conceptions. Descriptive statistics for assessment conceptions were calculated according to the independent variables: gender, age, level of education, teaching level, years of experience, and assessment training. For each level of the independent variable, Multivariate Variance Analysis were run to examine whether the teacher's assessment conceptions ratings varied greatly.

Conceptions of Assessment and Gender

Mean aggregate values were compared for the two levels of this variable: male and female. Table 1 provides a summary of the mean values for each gender level by the conception of assessment subcategory. The data showed a general trend whereby females recorded the highest average values in favour of school accountability, student accountability and improvement conceptions. However, males had the largest mean for conception of irrelevance. Standard deviations for each subcategory revealed that the biggest variation of responses was related to student accountability, whereas the least variation in ratings was associated to the assessment conception of improvement.

Table 1: Mean Scores of Conceptions of Assessment by Gender

Gender		School Accountability		Student Accountability		Improvement		Irrelevance	
	N	M	SD	M	SD	M	SD	M	SD
Male	99	4.76	.86	4.95	.81	4.63	.59	2.95	.71
Female	104	4.92	.81	5.04	.92	4.84	.55	2.73	.71

A MANOVA was performed to see if the mean differences were significant statistically. A Box's Test of Equality of Covariance was conducted to examine whether the data flouts the hypothesis of homogeneity of covariance matrices. A p-value of .450, based on an alpha value of 0.05, was correlated with the Box M value of 10.114. This suggested that each sex group of covariance matrices was identical to each other, and the assumption has been fulfilled. Again, the Levene's Test of Equality of Error Variance was conducted to check the requirements of equality of variance. The test results of Levene, as shown in Table 2 revealed that all the *p* values were greater than .05, indicating that the requirements of the equality of variance was satisfied. Furthermore, tests for linearity, normality and multivariate and univariate outliers were performed with no grave contraventions.

Table 2: Levene's Test of Equality of Error Variances (Gender)

	F	df1	df2	Sig.
School Accountability	.725	1	201	.396
Student Accountability	1.242	1	201	.266
Improvement	.000	1	201	.997
Irrelevance	.019	1	201	.890

A MANOVA was performed to examine mean differences between gender levels (male and female) and conception of assessment (School Accountability, Improvement, Student Accountability and Irrelevance) scores. The results revealed that the main effect for Gender was significant, Wilks' Lambda = .95, $F (4, 198) = 2.53$, $p < .05$; multivariate $\eta^2 = 0.05$. The effect size valued at.05 means that gender accounted for 5.0 percent of the variance in the dependent variable. The details of the MANOVA results are presented in Table 3.

Table 3: Manova Results for School Accountability, Student Accountability, Improvement, and Irrelevance by Gender

Variable	Wilks' Lambda	*F*	Hypothesis df	Error df	*p*	Partial Eta Squared
Gender	.951	2.53	4	198	.042	0.05

The significant value of the Wilks' Lambda statistics required a follow-up analysis test of a post hoc multiple comparisons. However, to guard against the Type 1 error (Pallant, 2016), a Bonferroni procedure was used to test each ANOVA at a p-value of 0.0125 (0.05 divided by the number of dependent variables). Based on this adjusted alpha value of 0.013), the univariate ANOVA results for the dependent variables were verified individually, and the only variable to attain a difference in statistical significance was Improvement, F (1, 201) = 6.56, p = .011, partial eta squared = .03. The details of each variable are shown in Table 4. An analysis of the mean scores revealed that females recorded marginally higher levels of Improvement conception (*M* = 4.84, *SD* = .55) than males (*M* = 4.63, *SD* = .59).

Table 4: Tests of Between-Subjects for Gender

Source	Dependent Variable	df	F	p.	η^2
Gender	School Accountability	1	1.86	.174	.01
	Student Accountability	1	.53	.466	.00
	Improvement	1	6.56	.011	.03
	Irrelevance	1	4.27	.040	.02

Conceptions of Assessment and Age Group

For the analysis of significant differences in mean scores according to age, age was categorized into three groups. These are low (21 – 30 years), mid (31 – 40 years), and high (41 and above years) age groups. Mean aggregate values were calculated for the three different levels of age for each assessment conception subgroup. Table 5 provides a summary of the mean values for each age group by the conception of assessment subcategories. The data showed a general trend whereby those in the mid (31 – 40 years) age group had the highest average values for all the dependent variables except school accountability, where those with a low age group scored the highest. Standard deviations for all subcategory revealed that the greatest variation of responses was related to student accountability for low (21 – 30 years) age group and high (40 – 60 years) group, whereas the least variation of responses was associated to the assessment conception of improvement.

Table 5: Comparison of mean scores of conceptions of assessment by age

Age (years)	N	School Accountability		Student Accountability		Improvement		Irrelevance	
		M	SD	M	SD	M	SD	M	SD
21 - 30	70	4.90	.89	4.90	.87	4.73	.50	2.88	.68
31 - 40	84	4.82	.78	5.08	.84	4.75	.62	2.91	.76
41 - 60	49	4.80	.85	4.97	.91	4.73	.52	2.77	.70

MANOVA was performed to see if the mean variations were significant statistically. A Box's Test of Equality of Covariance was conducted to examine if the data flouts the hypothesis of homogeneity of covariance matrices. A p-value of 0.320, based on an alpha value of 0.05, was correlated with the Box M value of 28.084. This suggested that each age group of covariance matrices was identical to each other, and the assumption has been fulfilled. Again, the Levene's Test of Equality of Error Variance was conducted to verify the assumption of and equality of variance. The test results of Levene, as shown in Table 6, revealed that all the *p* values were greater than .05, indicating that the requirements of the equality of variance was satisfied. Furthermore, tests for linearity, normality and multivariate and univariate outliers were performed with no grave violations.

Table 6: Levene's Test of Equality of Error Variances (Age Group)

	F	df1	df2	Sig.
School Accountability	1.043	2	200	.354
Student Accountability	.057	2	200	.945
Improvement	1.214	2	200	.299
Irrelevance	.742	2	200	.477

The findings from the tests of MANOVA show that the participant age did not differ statistically significantly from the combined dependent variables of F (8, 396) = 0.68 p =.7705.; Wilks Lambda =.97; partial eta squared = .01. This suggests that the linear combination of School Accountability, Student Accountability, Improvement, and Irrelevance was similar for each level of age group. The MANOVA results are presented in Table 7. Since there were no significant predictors, additional testing was not performed.

Table 7: MANOVA Results for School Accountability, Student Accountability, Improvement, and Irrelevance by Age

Variable	Wilks' Lambda	*F*	Hypothesis df	Error df	*p*	Partial Eta Squared
Age	.973	.68	8	394	.705	0.01

Conceptions of Assessment and Educational Level

Mean aggregate values have been examined for each conception subgroup for the two distinct levels of the independent variable, educational level. The two levels of this variable were: diploma and Bachelor and above. There were only three respondents who attained a master's degree, so this number was added to the bachelor group because of their small number. Table 8 provides a summary of the mean values for each educational level by the conception of assessment subcategories.

Table 8: Comparison of Mean Scores of Conceptions of Assessment by Educational Level

Highest Qualification	N	School Accountability		Student Accountability		Improvement		Irrelevance	
		M	SD	M	SD	M	SD	M	SD
Diploma	92	4.91	.80	4.99	.84	4.79	.58	2.85	.70
Bachelor & above	111	4.78	.85	5.00	.89	4.69	.58	2.84	.73

The data showed a general trend whereby teachers with only diploma qualification recorded the highest mean scores in terms of school accountability, improvement and irrelevance conceptions, whereas, those have bachelor's degree and above had the highest mean for student accountability conception. The greatest variation in ratings was related to student accountability, whereas the least variation in ratings was associated with the improvement conception.

A MANOVA was conducted to see if the mean differences were significant statistically. A Box's Test of Equality of Covariance was conducted to examine if the data flouts the

hypothesis of homogeneity of covariance matrices. A p-value of 0.656, based on an alpha value of 0.05, was correlated with the Box M value of 7.891. This suggested that each educational level group of covariance matrices was identical to each other, and the assumption has been fulfilled. Again, the Levene's Test of Equality of Error Variance was conducted to verify the assumption of and equality of variance. The test results of Levene, as shown in Table 9 revealed that all the *p* values were greater than .05, indicating that the assumption of the equality of variance was satisfied. Furthermore, tests for linearity, normality and multivariate and univariate outliers and multicollinearity were performed with no grave contraventions.

Table 9: Levene's Test of Equality of Error Variances (Educational Level)

	F	df1	df2	Sig.
School Accountability	.347	1	201	.556
Student Accountability	.230	1	201	.632
Improvement	.199	1	201	.656
Irrelevance	.551	1	201	.459

The MANOVA test results revealed that the main effect for educational level was not significant, $F(4, 198) = .72$, $p = .58$; Wilks' Lambda = .99; partial eta squared = .01, suggesting the linear combination of School Accountability, Student Accountability, Improvement and Irrelevance was not significantly different between the educational levels. The results are presented in Table 10.

Table 10: MANOVA Results for School Accountability, Student Accountability, Improvement, and Irrelevance by Educational Level

Variable	Wilks' Lambda	F	Hypothesis df	Error df	p	η^2
Gender	.986	.717	4	198	.58	0.01

Conceptions of Assessment and Class Level of Teaching

Mean total values were compared for each assessment conception subgroup at the three distinct levels of the independent variable. The three levels of this variable were: lower primary, upper primary and JHS. Table 11 provides a summary of the mean values for each level of teaching by the conception of assessment subcategories.

Table 11: Comparison of Mean Scores of Conceptions of Assessment by Age

Age (years)	N	School Accountability		Student Accountability		Improvement		Irrelevance	
		M	SD	M	SD	M	SD	M	SD
Lower Primary	60	5.01	.77	5.12	.82	4.88	.55	2.76	.71
Upper Primary	46	4.89	.83	4.93	.89	4.72	.51	2.82	.67
JHS	97	4.71	.86	4.94	.98	4.66	.62	2.91	.72

The data in Table 11 showed a general trend whereby those teaching at lower primary level had the highest mean values for the school accountability, student accountability and improvement conceptions sub-dimensions with the mean scores declining to their lowest level at the JHS level except for student accountability conception. However, the reverse is true for the irrelevance conception, where those at the JHS level had the highest and the lower primary level, the lowest

A MANOVA was performed to verify whether the mean differences were significant statistically. A Box's Test of Equality of Covariance was conducted to examine whether the data flouts the hypothesis of homogeneity of covariance matrices. A p-value of 0.748, based on an alpha value of 0.05, was correlated with the Box M value of 15.997. This suggested that each level of teaching group of covariance matrices was identical to each other, and the assumption has been fulfilled. Again, the Levene's Test of Equality of Error Variance was conducted to verify the assumption of and equality of variance. The test results of Levene, as shown in Table 12, revealed that all the *p* values were greater than .05, indicating that the assumption of the equality of variance was satisfied. Furthermore, tests for linearity, normality and multivariate and univariate outliers and multicollinearity were performed with no grave contraventions.

Table 12: Levene's Test of Equality of Error Variances (Class Level of Teaching)

	F	df1	df2	Sig.
School Accountability	1.210	2	200	.300
Student Accountability	.107	2	200	.899
Improvement	.924	2	200	.399
Irrelevance	.284	2	200	.753

The MANOVA results show that participant's level of teaching on the combined dependent variables, $F (8, 394) = .1.134$, $p = .339$; Wilks Lambda = .96; partial eta squared = .02 were not statistically significant. The results are shown in Table 13. Since there were no significant predictors, additional testing was not performed.

Table 13: MANOVA Results for School Accountability, Student Accountability, Improvement, and Irrelevance by Class Level of Teaching

Variable	Wilks' Lambda	F	Hypothesis df	Error df	p	Partial Eta Squared
Class Level of Teaching	.955	1.134	8	394	.399	0.02

Conceptions of Assessment and Years of Teaching Experience

For each assessment conception subgroup, mean aggregate values were calculated for the three separate levels of teaching experience. These were: low (less than five years), mid (5 – 10 years) and long (over ten years) teaching experience. Table 14 provides a summary of the mean values for each level of years of teaching experience by the conception of assessment subcategories. The data showed a general trend whereby those

with low (less than five years) teaching experience recorded the highest mean values in school accountability. Those with mid (5 – 10 years) teaching experience recorded the highest mean value in student accountability and irrelevance conceptions while those with high (over ten years) teaching experiences had the highest mean score in improvement conception.

Table 14: Comparison of Mean Scores of Conceptions of Assessment by Years of Teaching Experience

Teaching Experience (years)	N	School Accountability		Student Accountability		Improvement		Irrelevance	
		M	SD	M	SD	M	SD	M	SD
< 5 years (low)	68	4.89	.79	4.91	.84	4.74	.64	2.86	.66
5 – 10 years(mid)	62	4.87	.87	5.14	.86	4.72	.56	2.89	.79
>10 years (high)	73	4.76	.85	4.94	.90	4.75	.54	2.79	.70

Also, teachers with low teaching experience scored the lowest mean value for student accountability conception, and those with long teaching experience had the lowest mean value in school accountability and irrelevance conceptions. The subcategories standard deviations revealed that the greatest variation in responses were related to student accountability among the high teaching experience group. In contrast, the least variation in responses was in the conception of improvement among those with high teaching experience.

A MANOVA was performed for statistical significance of mean differences. A Box's Test of Covariance Equality was conducted to examine whether the data flouts the postulation of homogeneity of covariance matrices. A p-value of 0.086, based on an alpha value of 0.05, was correlated with the Box M value of 29.938. This suggested that each level of teaching group of covariance matrices was identical to each other, and the assumption has been fulfilled. Again, the Levene's Test of Error Variance Equality was performed to check the supposition of equality of variance. Results of the Levene test as shown in Table 15, revealed that all the p values were greater than .05, indicating that the assumption of the equality of variance was satisfied. Furthermore, tests for linearity, normality and multivariate and univariate outliers and multicollinearity were performed with no grave contraventions.

Table 15: Levene's Test of Equality of Error Variances (Years of Experience)

	F	df1	df2	Sig.
School Accountability	.299	2	200	.742
Student Accountability	.592	2	200	.554
Improvement	.438	2	200	.646
Irrelevance	2.521	2	200	.083

The MANOVA results show that there were no statistically meaningful differences among respondent's level of years of teaching experience on the combined dependent variables, F (8, 394) = .812, p = .593; Wilks Lambda = .97; partial eta squared = .02. The results are shown in Table 16. Since there

were no significant predictors, additional testing was not performed.

Table 16: MANOVA Results for School Accountability, Student Accountability, Improvement, and Irrelevance by Years of Teaching Experience

Variable	Wilks' Lambda	F	Hypothesis df	Error df	p	Partial Eta Squared
Teaching Experience	.968	.812	8	394	.593	0.02

Conceptions of Assessment and Training in Assessment.

Training in assessment was categorized into two levels as: training during pre-service only and training during and after pre-service. Table 15 shows the mean values for each category of training in assessment by the conception sub-dimensions. The data revealed that teachers with training in assessment during and after pre-service recorded the highest mean scores for all the levels of assessment conceptions.

Table 15: Comparison of Mean Scores of Conceptions of Assessment by Training in Assessment

Assessment Training	N	School Accountability		Student Accountability		Improvement		Irrelevance	
		M	SD	M	SD	M	SD	M	SD
During pre-service	102	4.81	.83	4.94	.91	4.66	.63	2.82	.71
During & after pre-service	101	4.86	.84	5.04	.83	4.82	.52	2.87	.73

A MANOVA was performed for statistical significance of mean differences. A Box's Test of Covariance Equality was conducted to examine whether the data contravenes the hypothesis of homogeneity of covariance matrices. A p-value of 0.686, based on an alpha value of 0.05, was correlated with the Box M value of 7.572. This suggested that each educational level group of covariance matrices was identical to each other, and the assumption has been fulfilled. Again, the Test of Error Variance Equality was conducted to verify the assumption of equality of variance. The test results of Levene, as shown in Table 16, revealed that all the p values were greater than .05, indicating that the assumption of the equality of variance was satisfied. Furthermore, tests for linearity, normality and multivariate and univariate outliers and multicollinearity were performed with no grave contraventions

Table 16 : Levene's Test of Equality of Error Variances (Assessment Training)

	F	df1	df2	Sig.
School Accountability	.986	1	201	.322
Student Accountability	3.441	1	201	.065
Improvement	.038	1	201	.845
Irrelevance	.001	1	201	.973

The MANOVA test results revealed that the main effect for training in assessment was not significant, $F(4, 198) = .1.22$, $p = .30$; Wilks' Lambda = .98; partial eta squared = .02, suggesting the linear combination of School Accountability, Student Accountability, Improvement and Irrelevance was not significantly different between the levels of assessment training. The results are presented in Table 17.

Table 17: Results of MANOVA for School Accountability, Student Accountability, Improvement, and Irrelevance by Educational Level

Variable	Wilks' Lambda	F	Hypothesis df	Error df	p	η^2
Gender	.976	1.219	4	198	.30	0.02

V. DISCUSSIONS

Teacher variables such as gender, educational level, years of teaching experience, teaching level, age and training in assessment were examined to determine if they had an influence on the teachers' conceptions of classroom assessment. As a result, a MANOVA test was applied to examine the differences.

In terms of gender and its influence on teachers' assessment conceptions, the MANOVA test revealed a statistically significant disparity in their conception of assessment between males and females Only the improvement conception of dependent variables was statistically significant after a further analysis using multivariate testing and Bonferroni adjustment. The results were not different from the qualitative findings where almost all the women viewed the purpose of assessment for improving teaching and learning compared with only half of their male counterparts. The results of this study are close to those of Ndalichako (2015), who found that more female teachers possessed a favourable view of classroom assessment compared to their male counterparts. The differences between male and female teachers were statistically significant about the use of assessments to promote and sustain teaching. Also, Brown and Gao (2015) found differences in teachers' assessment conceptions with regards to the gender of teachers. Male teachers embraced the notion that assessment should be used to control and monitor pupils, teachers and the school. By contrast, the findings of this study are contrary to those of Benson (2014), Mehrgan, Hayati and Alavi (2017), Yetkin (2017) and Yidana and Anti Partey (2018), who were unable to determine any gender impact on teachers' conceptions of classroom assessment. In comparison with this study, disparities in results can be attributed both to differences in methodologies used and in the contexts of educational systems. Furthermore, the composition of the gender of earlier studies compared to the current study could possibly be responsible for discrepancies in results.

To ascertain whether age impacts the respondents' assessment conceptions, they were grouped into three age groups of low (21 to 30 years), mid (31 to 40 years), and high (41 to 60 years) age groups. A MANOVA test to study the difference has not shown any significant difference between the various age categories and assessment conceptions. Such findings are consistent with those of the earlier assessment literature studies. Mehrgan, Hayati and Alavi's (2017) study of the influences of EFL teachers' age on their formative assessment beliefs revealed that age had no statistically significant impact on teachers' perception about formative assessment. Also, Yetkin (2018) found that prospective Turkish English teachers' conceptions significantly did not differ based on age, gender, and teaching experience. Similarly, Yidana and Anti Partey (2018) discovered that Economics teachers' age did not influence their assessment conception. In this study, descriptive findings indicated that the assessment conceptions of the three age groups are identical, although some minor mean differences have been established. All teacher groups expressed their agreement with conceptions of accountability and improvement and disagreed with the conception of irrelevance, as shown in the study's independent variable values. This means that the teachers viewed assessment as a means for improvement and accountability, whatever their age differences.

Teachers' educational level or attainment was investigated to find the extent it was related to teachers' conception of assessment. Teachers' educational attainment was categorized into two levels: those with a Diploma and those with a Bachelor's degree and above. The results suggest teachers with Diploma level education believed slightly that assessment measures school accountability and improvement conceptions than those earned a Bachelor's degree and higher. However, these differences were not significant. This study is in contrast to Calveric (2010), who found that those without a Bachelor's degree education assume that assessment measures serve student accountability purposes.

Teachers' grade level of teaching was examined to see whether the grade level of teaching makes a substantial impact on their assessment conceptions. The levels of teaching were grouped into three: Lower Primary, Upper Primary and JHS. No statistically significant difference between the grade level of teaching and conception of assessment was reported in the MANOVA test results. Some research studies have shown that teachers' conceptions about assessment differ according to the levels at which they teach (Brown et al., 2011; Remesal, 2007), however, it remains uncertain whether this variation is as a result of the arrangement and policies related to different levels of education (e.g., primary and secondary), or whether it is related to previous convictions about teaching, learning, and assessment (Bonner, 2016). A study by Remesal (2007) of primary and secondary teachers in Spain revealed primary school teachers were more inclined to consider assessment is for instructional purposes. In contrast, secondary school teachers often held an 'accounting' conception of assessment in certifying student performance. These perceived differences have been attributed to policy discrepancies at the primary and secondary school levels (Remesal, 2011). Similar differences were noted among Queensland, Australia, primary and secondary teachers by Brown et al. (2011), in that primary teachers' conceptions leaned towards improvement conception while secondary teachers' conceptions inclined towards

student accountability. The differences may be due to policy disparities between primary and secondary education in which a comprehensive, publicly controlled school-based assessment program existed at that time only at the upper secondary level. However, in this current study, the data showed a general trend whereby those teaching at lower primary level had the highest mean values for the school accountability, student accountability and improvement conceptions sub-dimensions with the mean scores declining to their lowest level at the JHS level except for student accountability conception. However, the reverse is true for the irrelevance conception, where those at the JHS level had the highest and the lower primary level, the lowest. The similar views held by teachers teaching at these three levels could be due to the fact that these teachers were trained generally for the basic school level and so could be assigned to teach at any of these levels in any academic or school year. In effect, it is possible that some of the teachers teaching at the JHS level, could have taught at the lower or upper primary levels and vice versa in the previous academic or schooling year.

To assess whether years of teaching experience factor affect assessment conceptions of participants, they were categorized into three teaching experience groups of low (> 5 years), mid (5-10 years) and high (over ten years) teaching experience groups. In this current study, descriptive statistics revealed that teachers with high (over ten years) teaching experience obtained the highest mean value in improvement conception, those with low (less than five years) teaching experience had the highest score for the school accountability. Those with mid (5 – 10 years) teaching experience recorded highest score for student accountability and irrelevance conceptions. A MANOVA conducted to analyze the differences found no substantial differences between the groups of teaching experience and assessment conceptions. This study is similar to Fulmer, Tan and Lee's (2017) study of Singaporean secondary school teachers in which no significant relationship was found with teachers' teaching experience and their conception of assessment. Similarly, Yetkin (2018) found a non-statistically significant difference in teachers' conceptions among 204 prospective Turkish English teachers based on teaching experience. Furthermore, Benson (2014) discovered that in terms of teaching experience, younger and older teachers held similar assessment conceptions, especially the irrelevance conception. In contrast, Sahikarakas (2012) found a significant difference in the assessment conception of Language teachers concerning teaching experience. Teachers with more experience have a negative view of assessment than their less experienced counterparts According to Sahikarakas (2012), the differences are due to the experienced teachers highly valuing themselves to a level that there is no need for them to obtain evidence of their teaching efficacy through assessment. Also, Yidana and Anti Partey (2018) found that Economics teachers' teaching experience does affect their conception of assessment. Teachers with over seven years of teaching experience in Economics have shown a positive conception, relative to those with fewer than three years of

experience. Besides, Brown and Gao (2015) found differences in teachers' assessment conception regarding teaching experience. The conception that assessment should be used to evaluate and track students, staff, and school to ensure successful teaching and learning was supported by teachers with twenty and over years of experience. The study by Vardar (2010) found that teachers have significant differences in assessment conception in relation to years of experience of teaching with the more experienced teacher category having the highest level of student accountability conception compared with other less experienced groups.

The degree to which teachers' training in assessment impacted their assessment conceptions was investigated. The findings revealed that teachers' assessment conceptions did not change with regards to the training in assessment. In this study, training in assessment was categorised into two groups; training during teachers' studies at undergraduate studies only and training in assessment during and after undergraduate studies. However, it appears that teachers with training in assessment during and after undergraduate studies recorded the highest mean scores for all the levels of assessment conceptions. The study findings are similar to studies by Brown and Hirschfeld (2008), Levy-Vered and Alhija (2015) and Vadar (2010), who discovered that having more training in assessment or attending an assessment course did not enhance teachers' assessment conceptions. However, the finding of the study is contrary to that of DeLuca, Chavez and Cao (2013), and Smith et al. (2014). They revealed that teachers' assessment conception improved after getting routine professional assessment training. Similarly, Yidana and Anti Partey (2018) found a positive relationship between Economics assessment conceptions and their professional training in assessment. They attributed this to the Ghana Education Service routine in-service training provided to teachers, as well as the quality and length of such professional training.

VI. CONCLUSION AND RECOMMENDATION

The study revealed that except gender, other teacher demographic characteristics such as educational level, age, teaching experience, class teaching level, and assessment training did not impact on the respondents' assessment conceptions. It is therefore recommended that stakeholders such as head teachers, the Ghana Education Service should consider the development of ways to improve the assessment literacy of male teachers.

REFERENCES

[1] Alderson, J. C. (2005). Principles and practice in language testing. Keynote address at RATE-QUEST Conference, Cluj, Romania, 5th August, 2005.

[2] Benson, T. L. (2014). Sixth through eighth grade teachers' conceptions (beliefs) about assessment practices. Unpublished D. Ed dissertation presented to the Wingate University School of Graduate and Adult Education, Union Country.

[3] Bonner, S. (2016). Teachers' perceptions about assessment: Competing narratives. In G. T. L. Brown & L. Harris (Eds.), Handbook of social conditions in assessment (pp. 21–40). New York: Routledge.

[4] Brown, G. T. L. (2004). Teachers' conceptions of assessment: Implications for policy and professional development. Assessment in Education: Policy, Principles and Practice, 11(3), 305-322.

[5] Brown, G. T. L. (2006). Teachers' conceptions of assessment: Validation of an abridged instrument. Psychological Reports, 99, 166-170.

[6] Brown, G. T. L. (2008). Conceptions of assessment: Understanding what assessment means to teachers and students. Nova Science Publishers.

[7] Brown, G. T. L., & Gao, L. (2015). Chinese teachers' conceptions of assessment for and of learning: Six competing and complementary purposes. Cogent Education, 2(1), 1-19.

[8] Brown, G. T. L., & Hirschfeld, G. H. F. (2008). Student' conceptions of assessment: Links to outcomes. Assessment in Education: Principles, Policy and Practice, 15(1), 3 – 17.

[9] Brown, G. T. L., Lake, R. & Matters, G. (2011). Queensland teachers' conceptions of assessment: The impact of policy priorities on teacher attitudes. Teaching and Teacher Education, 27, 210-220. doi: 10.1016/j.tate.2010.08.003

[10] Calveric, S. B. (2010). Elementary teachers' assessment beliefs and practices. (Doctoral dissertation, Virginia Commonwealth University, Virginia, USA).

[11] Daniels, L. M., Poth, C., Papile, C., & Hutchison, M. (2014). Validating the conceptions of assessment-III scale in Canadian pre-service teachers. Educational Assessment, 19(2), 139-158. doi:10.1080/10627197.2014.903654

[12] DeLuca, C., Chavez, T., & Cao, C. (2013). Establishing a foundation for valid teacher judgement on student learning: The role of pre-service assessment education. Assessment in Education: Principles, Policy & Practice, 20(1), 107 – 126.

[13] Elliott, A. C., & Woodward, W. A. (2007). Statistical analysis quick reference guidebook with SPSS examples (1st ed.). London: Sage Publications

[14] Fulmer, G. W., Tan, K. H. K., & Lee, C. H. I. (2017): Relationships among Singaporean secondary teachers' conceptions of assessment and school and policy contextual factors. Assessment in Education: Principles, Policy & Practice, 1 – 17. http://dx.doi.org/10.1080/0969594X.2017.1336427

[15] Harris, L. (2008). Secondary teachers' conceptions of the purpose of assessment and feedback. Paper presented to the Australian Association for Research in Education (AARE) Annual Conference, December 2008, Brisbane, Australia.

[16] Izci, K., & Caliskan, G. (2017). Development of prospective teachers' conceptions of assessment and choices of assessment task. International Journal of Research in Education and Science (IJRES), 3(2), 464 – 474.

[17] Krejcie, R. V. & Morgan, D.W. (1970). Determining sample size for research activities. Educational and Psychological Measurement, 30, 607-610.

[18] Levy-Vered, A., & Alhija, F. N. (2015). Modelling beginning teachers' assessment literacy: The contribution of training, self-efficacy, and conceptions of assessment. Educational Research and Evaluation, 21(5), 378 – 406.

[19] McMillan, J. H. (2018). Classroom assessment: Principles and practice that enhance student learning and motivation (7th ed.). New York: Pearson.

[20] Mehrgan, K., Hayati, A., & Alavi, S. M. (2017). Investigating the impacts of EFL teachers' age, educational background, instructional experience and gender on their beliefs about formative assessment. International Journal of Foreign Language Teaching & Research, 5(18), 143 – 160.

[21] Moiinvaziri, M. (2015). University teachers' conception of assessment: A structural equation modelling approach. Journal of Language, Linguistics and Literature, 1(3), 75-85.

[22] National Council for Curriculum and Assessment (NaCCA) (2019). Resource guide for the orientation of primary school teachers towards the implementation of the revised curriculum for primary schools. Accra: Ministry of Education.

[23] Ndalichako, J. L. (2015). Secondary school teachers' perceptions of assessment. International Journal of Information and Education Technology, 5(5), 326 – 332.

[24] Okyere, M., Kuranchie, A., Larbi, E., & Twene, C. (2018). Essentials of assessment in schools. Sunyani: Aduana Printing Press

[25] Pallant, J. (2016). SPSS survival manual, a step by step guide to data analysis using IBM SPSS (6th ed.). Berkshire, England: McGraw Hill

[26] Remesal, A. (2007). Educational reform and primary and secondary teachers' conceptions of assessment: The Spanish instance, building upon Black and Wiliam (2005). The Curriculum Journal, 18(1), 27-38.doi: 10.1080/09585170701292133

[27] Remesal, A. (2011). Primary and secondary teachers' conceptions of assessment: A qualitative study. Teaching and Teacher Education, 27, 472-482.

[28] Sahikarakas, S. (2012). The role of teaching experience on teachers' perceptions of Language assessment. Procedia – Social and Behavioural Science, 47(2), 1786 – 1792.

[29] Smith, L. F., Hill, M. F., Cowie, B., & Gilmore, A. (2014). Preparing teachers to use the enabling power of assessment. In C. Wyatt-Smith, V. Klenowski, & P. Colbert (Eds.), Designing assessment for quality learning (pp. 303–323). Dordrecht, The Netherlands: Springer.

[30] Tabachnick, B. G. & Fidell, S. (2013). Using multivariate statistics (6 ed). Boston: Pearson Education, Inc.

[31] Vardar, E. (2010). Sixth, seventh, and eighth grade teachers' conception of assessment. (Master's Thesis, Middle East Technical University, Ankara, Turkey).

[32] Xu, Y., & Brown, G. T. L. (2016). Teacher assessment literacy in practice: A reconceptualization. Teaching and Teacher Education, 58(2), 149 – 162.

[33] Yetkin, R. (2017). Pre-service English teachers' conception of assessment and their future assessment practices in a Turkish context. Master's thesis. Hacettepe University.

[34] Yetkin, R. (2018). Exploring prospective teachers' conceptions of assessment in Turkish context. European Journal of Education Studies, 4(5), 133-146. DOI: 10.5281/zenodo.1230554

[35] Yidana, M. B. & Anti Partey, P. (2018). Economics teachers' conceptions of classroom assessment: A study of senior high schools in the Central and Ashanti Regions of Ghana. International Journal for Innovation Education and Research, 6(10), 153 – 174. DOI: https://doi.org/10.31686/ijier.Vol6.Iss10.1176