Mediterranean Journal of Social & Behavioral Research
2022, 6(3), 85-91
ISSN 2547-8559 (Print)
ISSN 2547-8567 (Online)
https://www.mjosbr.com/

Research Article

OPEN ACCESS

Towards Understanding Test Score Pollution in Ghana: Test Preparation Practices, Test Administration Situations and External Factors as Predictors

Simon Ntumi [1]*, Sheilla Agbenyo [2], Tapela Bulala [3]

[1]Department of Educational Foundations, Faculty of Educational Studies, University of Education, Winneba, West Africa, GHANA
[2] Bia Lamplighter College of Education, West Africa, GHANA
[3] Botswana University of Agriculture and Natural Resources, Southern Africa, BOTSWANA
***Corresponding Author:** sntumi@uew.edu.gh

Citation: Ntumi, S., Agbenyo, S., & Bulala, T. (2022). Towards Understanding Test Score Pollution in Ghana: Test Preparation Practices, Test Administration Situations and External Factors as Predictors. *Mediterranean Journal of Social & Behavioral Research*, 6(3), 85-91. https://doi.org/10.30935/mjosbr/12323

ABSTRACT

Background: Test score pollution explains how multifaceted factors affect the truthfulness of a test score interpretation. The pressure to raise test scores has resulted in practices which pollute the inferences we make from these scores. Issues of accurate testing remains relevant in the space of any testing environment in Ghana and beyond. This study explored the different sources of test score pollution considered test preparation practices (teacher factor), test administration situations (testing environment), and external factors (parents and community pressure).

Methods: The study was nested into the quantitative approach using descriptive survey. Basic school teachers (n=353) and parents (n=123) were selected from three districts (South, North and Central Tongu) in the Volta region using G*Power software. A validated and standardized instrument (with alpha coefficient of .783 and correlation coefficient of .823) was used to obtain the data. The obtained data was analyzed using SPSS v.25 and interpreted with linear multiple regression after the data had met all the required assumptions.

Findings: The results revealed that all the predictive factors that is test preparation practices (t=4.73, Sig.=.007, $CI_{95\%}$), test administration situations (t=4.20, Sig.=.006, $CI_{95\%}$) and parents and community pressure (t=2.69, Sig.=.000, $CI_{95\%}$) predicted test score pollution in the selected districts. However, among all the predictor variables, test administration situations (testing environment or conditions) were identified as having much influence on test score pollutions in the districts (R^2 =.652, 65.2%, Sig.=.000, β=.616, $CI_{95\%}$).

Conclusion: The study concluded that due to test score pollution, most test practices in Ghana are not at its optimal best. Clearly, the demand and the pressure to raise test scores results pollute and contaminate the interpretations, inferences and decisions that are made from these test scores.

Keywords: scores, test, practices, pollution, test administration

Received: 19 Mar. 2022 ◆ Accepted: 4 Aug. 2022

INTRODUCTION

Undoubtedly, test is an essential tool that helps to quantify constructs which helps one to make a value judgment about the degree to which a construct might probably exist in an individual (Akaranga & Ongong, 2013; Akyeampong, 2007; Amedahe, 2000; Anane, 2008). In Ghana and across the globe, evidence suggest that test results are increasingly becoming the benchmark for measuring an institution's success. We share in the statement that humans are living in an era of increasing interest in accountability which is driven by tests and its outcomes (Azizeh & Mansoor, 2010; Bachman et al., 2013). Standardized tests continue to be touted as the most important measure and predictive of student performance (Akyeampong, 2007; Buzick & Jones, 2015; Chalak & Tavakoli, 2010).

Practically, when scores go up, the administrators and politicians are happy and act as though achievement has gone up, but the smile is not that achievement had improved (Chapman & Snyder, 2019). In general, there is more rhetoric than evidence about the consequences of assessment and too often policy debates emphasize only on one side or the other of testing effects coin (Bachman et al., 2013; Cohen et al., 2017). In the case of Ghana, people are of the view that student's pass their examinations without much evidence to show in their potentials in the job place and as society would want them to function. These perceptions seem to be normal, especially, when a test is seen as an attempt to determine how an individual would function in a set of

actual situations (Azizeh & Mansoor, 2010; Bachman et al., 2013; Dolly & Williams, 1983).

It is worthwhile to note that the high pressure to produce high test score in many educational institutions have urged many teachers and testers to train and prepare test takers to complete their tests just for higher scores (Bachman et al., 2013; Dreisbacha, & Keogh, 2012). Literature suggests that demand to raise test scores has resulted in conditions and practices which pollutes and contaminates the interpretations and inferences that are made from test scores (Akaranga & Ongong, 2013; Bachman et al., 2013; Filson & Brown, 2018; Flippo & Caverly, 2008). This implies that the increased use of standardized achievement tests has come the pressure to raise scores, which in turn leads to increase in test score pollution. The pollution seriously affects the truthfulness of test score interpretations and puts into doubt the reasonableness of the many of the uses of the test scores (Flippo et al., 2019; Genshaft & Kirwin, 2018; George & Mallery, 2012).

Studies in the field of general education have established the influence of testing on the process of teaching and learning (Gerald, 2018; Grant, 2015; Gyimah, 2012). Test results across the world are used as an indicator of the performance of teachers, schools, and the accountability of the education system (Haladyna, 2016; Haladyna et al., 1991).

In the current climate of dissatisfaction with public education, the standardized achievement tests scores have been the operational definition for educational achievement and raising test scores has been equated with educational improvement (Bachman et al., 2013; Chalak & Tavakoli, 2010). The misuse and overuse of test results, high pressure to produce high test score, high stakes nature of many uses, equating test scores with educational improvements have urged many teachers and testers to train and prepare test takers to complete their tests. The demand to raise test scores has resulted in conditions and practices which pollutes and contaminates the interpretations and inferences that are made from test scores (Filson & Brown, 2018; Haladyna et al., 1991, 2018).

In preparing students for standardized achievement testing, test preparation practices (teacher role), test administration situations (testing environment or conditions) and external factors (parents and community activities) could be indicative of students' performance for that matter their true score (Haladyna et al., 2018; Hargett, 2016; NCME, 2014). For example, in test preparation, many teachers utilize a variety of strategies which aimed at maximizing student performance. Within the measurement and testing literature, several researchers have written about the potential consequences, both good and bad, of using particular test preparation practices (Akaranga & Ongong, 2013).

Largely, some authors have asserted that factors such as test anxiety, motivation, self-esteem, inhibition, stress, fatigue, concentration, attention, interest, setting, policies of the school, administration, location, and the examiner effect are among the factors that have impact on test scores (Chalak & Tavakoli, 2010; Haladyna et al., 2018; Jones et al., 2019). These factors are specific to the administration of the test and pollute or contaminate test results or interpretations that are made from test scores (Buzick & Jones, 2015; Grant, 2015; Haladyna et al., 2018). The society accessibility to test results also pushes schools to provide any support necessary for raising test scores which results in practices that leads to test score pollution (Haladyna et al., 2018; Mereku, 2000).

MATERIALS AND METHODS

Research Design and Sampling Procedure

Quantitative approach using descriptive survey was adopted for this study. Quantitative approach was deemed appropriate for this study based on the rationale that we wanted to quantify social phenomena and collect and analyze numerical data that will reflect the phenomenon under investigation (Hargett, 2016; Mitchell, 2007).

The total population of the study was made up of all the basic teachers and parents in the districts. Sample size for the study was 352 teachers and 123 parents. G*Power software was used to use to obtain the sample. The rational for using the software is based on the assumption that it enables researchers to do analyses for many different t-tests, regression test, F tests, Chi-square ($\chi 2$) tests, z-tests, and some exact tests. G*Power also enable researchers to compute effect sizes and to display graphically the results of power analyses

Instrumentation and Data Analysis Procedure

The instrument used for data collection contained indicators that measured test preparation practices (teacher factor), test administration situation (testing conditions or environment) and external factors (parents and the community pressure). The instrument was validated and proven reliable and standardized for data collection. To estimate the validity, content and construct validity were employed.

To evaluate the reliability evidence, internal consistency using alpha coefficient and correlation coefficient were computed. For alpha coefficient, .783 was obtained and .823 was obtained for correlation coefficient. The items on the questionnaire were close ended and were used to measure the predictors. The items on the questionnaire were multiply scored on a four-point Likert type scale. The items on the Likert scale scored ranging from four (4) for strongly agreed to one (1) for strongly disagree for positive statements. Negative statements that were captured were scaled in the reverse form in the coding process.

The obtained data was collated and edited without altering the responses. After coding, the data was entered into the computer and processed using the statistical package for social sciences (SPSS v.25) and interpreted with the linear multiple regression (LMR) using the stepwise selection. Using the stepwise selection, we combined the predictors variables in a forward and backward selection matter.

In the approach, we began with a null model, then we added the single independent variable that makes the greatest contribution toward explaining the dependent variable, and then iterates the process. Additionally, a check was performed after each such step to see whether one of the variables has now become irrelevant because of its relationship to the other variables. In our case, all the predictors were relevant as such were not removed.

The justification for selecting the multiple linear regression (LMR) using the stepwise selection approach was to show the direction and magnitude of the predictive variables (TAS[1], TPP[2], and PCP[3]) on the dependent variable (test score pollution). The use of the LMR allowed us to identify the unique contribution of each predictor to the outcome variable (TSP).

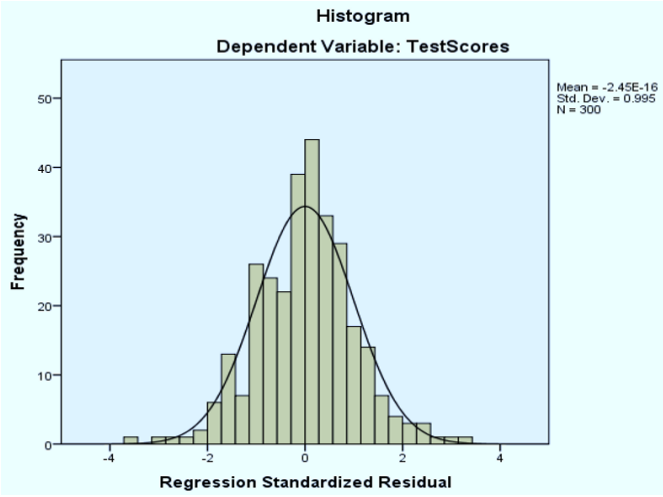

Figure 1. Normality plot

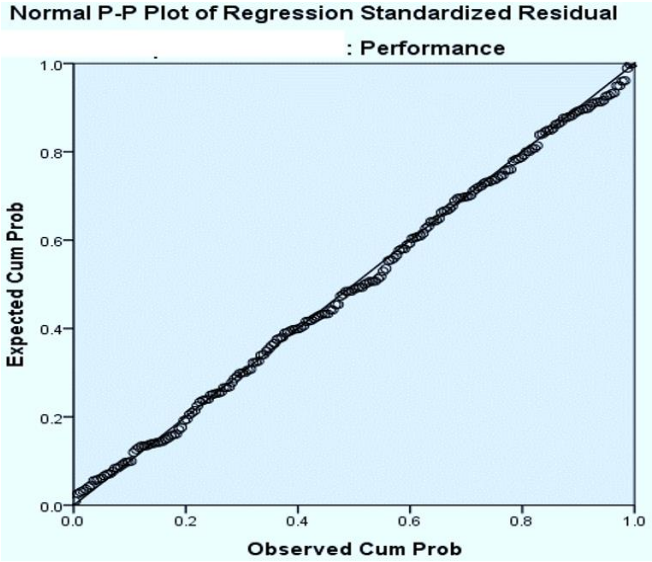

Figure 2. Diagnostic test of normality and linearity

FINDINGS

Statistically, we sought to find out how test preparation practices (teacher factor), test administration situations (testing conditions or environment) (test administration situation) and other external factors (parents and community pressure) are associated with test score pollution. To achieve this, LMR was deemed appropriate for the analysis. However, prior to performing multiple regression test, certain assumptions **must** be met. All these assumptions were **tested,** and the data was proven normal. One of these assumptions is normality of the study variables.

Figure 1 presents the normality of the multiple regression test. The clustering of the variables at the center of the normality curve shows that the data was normal and multiple regression could be performed. According to Nitko (2001), a straight normal probability plot is an indication of normality and linearity. Gerald (2018) noted that when multiple regression assumptions are met, it produces a reliable result.

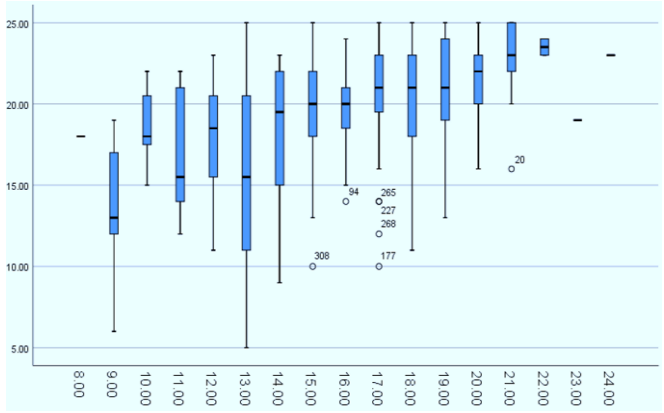

Figure 3. Means plot test of normality and Linearity

Table 1. Results of multicollinearity diagnostic test of study variables

Predictor variables (factors)	Factor1 (X1)	Factor2 (X2)	Factor3 (X3)
Test administration situation (TAS)	1.00	.039	.042
Test preparation practices (TPP)		1.00	.436*
Parents & community pressure (PCP)			1.00

Note. Source: Field survey (2021); Sample: Teachers (n=353) & Parents (n=123); Dependent variable: TSP; & Independent variables: TPP[1], TAS[2], PCP[3]

From **Figure 2**, a reasonable straight line could be seen from the plot demonstrating normality and linearity of the data. This, therefore, means that conducting multiple regression test was justified.

Similar interpretation and understanding are recounted and displaced in **Figure 3**. The histogram plot of standardized predicted values versus standardized residuals, showed that the data met the assumptions of normality of variance and linearity, and the residuals were approximately normally distributed.

Table 1 indicates the results of multicollinearity diagnostic test of the variables. The problem of multicollinearity is said to exist when independent variables used in the study (X1, X2, and X3) are highly correlated with each other. The study followed literature to test this assumption. It is assumed that correlation coefficient of 0.70 or more between independent variables is assumed to demonstrate evidence of multicollinearity problem (NCME, 2014; Nolen et al., 2016). From **Table 1**, the highest correlation coefficient is .436 (X2*X3) which is less than 0.70. and the least high correlation coefficient is .039 (X1*X2). This gives evidence that there was no problem of multicollinearity in the dataset. Having tested for the assumptions, running multiple regression was deemed appropriate.

Table 2 illustrates the descriptive statistics (means and standard deviations) of the associated factors that predicted test score in the three districts. From the analysis, the maximum score after the data was transformed was 60.00 (max.=60.00) and the minimum score was 15.00 (min.=15.00). This implies that mean values must fall with the minimum and the maximum range (15.00-60.00). On the Kurtosis values, the results show that the variables follow a normal distribution this is based on the reason that the kurtosis values were within the acceptable limit for normal distribution of ±2 (Nitko, 2001) indicating that the data was normal (not skewed).

The results in **Table 2** indicate that test administration situation (TAS) recorded the highest mean and standard deviation (M=45.62, SD= 1.097, K=.519, T=352, P=123, $CI_{95\%}$,). This was followed by test preparation practices (TPP) with the mean and standard deviation

Table 2. Descriptive statistics of the predictive variables

Associated factors	Mean	Standard deviation	Skewness		Kurtosis		n-sample
			Stat.	Standard error	Stat.	Standard error	
Test administration situation (TAS)	45.62	3.223	1.9416	-1.100	.519	1.39	T=353, P=123
Test preparation practices (TPP)	37.78	1.097	1.8423	-.081	.794	-.030	T=353, P=123
Parents & community pressure (PCP)	22.83	7.090	2.0534	.009	.954	-.400	T=353, P=123

Note. Source: Field survey data (2021); Max. score: 60.00; Min. score: 15.00; CI95%; & Sample: Teachers (n=353) & Parents (n=123)

Table 3. Multiple regression analysis of the predictors of test score pollution

Model	Unstandardized coefficients		Standardized coefficients	Cal. t-value	p-value p>|t|	Rks	n
	B	Standard error	Beta (β)				
(Constant)	72.508	9.069		2.61	.000($CI_{95\%}$)	-	T=353, P=123
TAS	1.425	2.134	.616	4.20	.006($CI_{95\%}$)	HP	T=353, P=123
TPP	3.381	3.104	.473	4.73	.007($CI_{95\%}$)	MP	T=353, P=123
PCP	3.249	.133	.396	2.69	.000($CI_{95\%}$)	LP	T=353, P=123

Note. Dependent variable: TSP; Independent variables: TPP1, TAS2, PCP3; HP-High predictor; MP-Moderate predictor; LP-Low predictor, n-Sample; T-Teachers; P-Parents; Rks-Remarks; & CI-Confidence interval

Table 4. Results of multiple regression analysis of contribution of each the predictors

Predictors	R	R^2	Adjusted R^2	Change statistics	Rank order	Rks
				R^2 change		
Test administration situation (TAS)	.604[a]	.377	.673	.652 (65.2%)	1st	HP
Test preparation practices (TPP)	.822[b]	.705	.763	.218 (21.8%)	2nd	MP
Parents & community pressure (PCP)	.895[c]	.874	.897	.130 (13.0%)	3rd	LP

Note. Source: Field survey (2021); Sample: Teachers (n=353) & Parents (n=123); CI95%; Dependent variable: TSP; Independent variables: TPPa, TASb, PCPc; HP-High predictor; MP-Moderate Predictor, & LP-Low predictor

(M=37.78, SD=7.090, K=.794, T=352, P=123, $CI_{95\%,}$). Parents and community pressure (PCP) recorded the least mean and standard deviation (M=22.83, SD= 3.223, K=.954, T=352, P=123, $CI_{95\%,}$). From this descriptive analysis, one could conclude that test administration situational (TAS) factors are more associated with test score pollution than all the other predictors. However, in order to determine whether these differences in the mean scores were statistically significant, we further conducted regression analysis to give more statistical evidence to confirm the findings.

Table 3 presents the results for the multiple regression analysis. From the results, it is evident that all the three independent variables (test administration situation, test preparation practices and parents and community pressure) were statistically significant at p-value of 0.05 indicting that all the predictive factors contribute to test scores pollution. For example, test administration situation as predictive factor was statistically significance (t=4.20, β=.616, Sig.=.000, $CI_{95\%,}$, T=352, P=123), test preparation practices also produced a significance result (t=4.73, β=.473, Sig.=.007, $CI_{95\%,}$ T=352, P=123), and finally, parents and community pressure yielded a significance result (t=2.69, β=.396, Sig.=.000, $CI_{95\%}$, T=352, P=123). Nonetheless, when measuring the standardized beta values, the highest factors upon the dependent variable are test administration situation (beta=.616), test preparation practices (beta=.473) and finally, parents and community pressure (beta=.396). when assessing the beta values, the results suggest that among all the predictive factors, test administration situation (testing conditions or environment) was identified as having much influence on test score pollution in the selected districts.

Table 4 shows how each of the predictors contributed to test scores pollution (in percentage terms) in districts. Using the R^2 square change statistics from **Table 4**, it was evident that X1 (test administration situation) contributed more than all the other factors. That is R^2 change statistic value of .652 ($CI_{95\%}$) representing 65.2%. Factor X2 (test preparation practices) contributed R^2 change statistic value of .218 ($CI_{95\%}$) representing 21.8% indicating the second contributor. Factor X3 (parents and community pressure) contributed R^2 change statistic value of .130 representing 13% showing the third contributor. The implication to this analysis is that predictor1 (X1-test administration situation) was identified as the best predictor of scores pollution in the selected districts in the Volta region of Ghana. This suggest and explains that to improve rigorous and valid testing practices in the districts, much attention should be given to situations or conditions that surround test administration.

DISCUSSION

The study established evidence to suggest that test scores are polluted with activities of test administration situation, teachers as well as external influence of parents and community. The results give abundant evidence to believe that scores of students ride on situations or conditions that surround test administration. That is to mean that how tests are administered reduce the fairness and the validity of the test scores. The paper is discussed using the three predictors (TAS, TPP, and PCP) in the study.

Test Administration Situations (Testing Conditions or Environment)

In our study, we were guided by the fact that test administration situation refers to all those factors that are specific to the organization and administration of tests and their situations. They are factors affecting test scores, their inferences, interpretations, and validity (Haladyna et al., 2018; NCME, 2014). In our study, it was revealing that how tests are administered contribute highly to test score pollution leading to error scores. The results give ample evidence to assert that in

Ghana, it does appear that little effort is put in place to control or hold some variables constant during test administration process.

Contextually, our findings place itself with other authors who accrued similar evidences in their studies. For example, the work of NCME (2014) and Thomas et al. (2001) found that factors such as test anxiety, motivation, self-esteem, inhibition, stress, fatigue, concentration, attention, interest, setting, policies of the school, administration, location, and the examiner effect are among the factors that have impact on test scores. One common striking evidence among all these studies is that these factors are specific to the administration of the test and may pollute or contaminate test results or interpretations that are made from test scores.

Sequel to other related work, test anxiety, time of the test, examiners attitudes were regarded as some of the highly possible situational factors that contribute to test score pollution. For instance, in a validation study by Mitchell (2007), it was contended that test anxiety was deemed as an important situational factor and has received great deal of attention. Test anxiety as test administration situations can affect the results of tests and scores (Amedahe, 2000; Tabachnick et al., 2007; Thomas et al., 2001; Wall, 2018).

Relatedly, another factor that was found to pollutes test score in test administration situation was time limit. Timed testing compared to extending time limits or increasing or providing extra time may affect the results obtained through tests (Anane, 2008; Buzick & Jones, 2015). On the other hand, teachers' attitude toward teaching, tests and even students may affect the results of the tests. It can have an impact on student achievements (Cohen et al., 2017; Genshaft & Kirwin, 2018). This factor as a situational factor may increase or decrease performance of the students. Standardized tests may create some kind of discouraging climate for teachers which in turn affect the profession and their attitude toward the test and consequentially may contaminate the results of test.

Test Preparation Practices (Teacher Factor)

The results from the study suggest that teachers in their quest to prepare their students for tests could pollute the scores. Most teachers are influenced by the activities of the school authorities to produce higher scores for their students. Most of these teachers engage in these practices in order to save their image as well as to protect their job. The evidences accrued in this study were reiterated in the work of Chalak and Tavakoli (2010) and Chapman and Snyder (2019), who similarly pointed out that test preparation activities may be regarded as ethical (such as training in test-wiseness skills, checking answer sheets to make sure that each has been properly completed, increasing motivation) or non-ethical (such as scoring high, presenting items similar, identical or parallel to those on the test) but the important point is that test preparation activities affect the validity of uses, inferences and interpretations that we make from test scores. Teachers in their quest to produce higher score for their students are likely to be engaged in unethical practices that could contaminate or pollute the true score of students. To this end, these practices of teachers could affect test score and are considered as test contaminants.

External Factors (Parents and Community Pressure)

On the external factors which carved and used in this study as parents and community pressure was recounted as one of the predictive factors that contribute to test scores pollution. The results gave abundant evidence to believe that pressures exacted by parents and the community, or the society compel many students to contaminate their scores. Most of these students engaged in these malpractices activities in order to satisfy the desires and curiosities of society and parents. Similar to this, Tamakloe et al. (2005) believed that the most important source of test score pollution is attributed to external factors. Also, Dolly and Williams (1983) claimed that the most dangerous pollution is the misinterpretation and over-interpretation of test scores which lead to many of the other sources of contamination.

Akin to the above related empirical propositions, the study of Chapman and Snyder (2019) also found that external factors could have a great effect on the test scores of students. This therefore suggest that to maintain an optimal level valid scores for students, a great deal of attention needs to be given to the external factors. The findings from this study are not far from that of NCME (2014) who examined different study skills and concluded that ineffective instructions by parents and community can lead to learners' test scores pollution.

CONCLUDING REMARKS

From the study, we could infer that the pollution of test score could be as a result of some measurement errors such as teacher influence, test administration situation and parent and community pressure. This is to suggest that most of the scores of students are not representative of their true abilities and skills. To promote or to provide accurate and reliable test results, it is incumbent on testing authorities to pay much attention to measurement errors (test administration situation, test preparation practices and parents and community pressure) that could dilute students' scores. Obviously, the reason for the test score pollutions in the three districts could be as a result of teachers viewing West African Examination Councils (WAEC) as a tool of accountability of the quality of teaching in their schools. In line with the test administration situations or the testing environment or conditions, it is highly possible that most of the testing environments are not rigorous enough to contain students and invigilators behaviors and attitudes. To worsen the phenomenon, pressure and comparisons from parents and community members could compel teachers and students to travel every mile to get their scores better. Until all these measurements errors (test administration situation, test preparation practices and parents and community pressure) that pollute test scores are control to some extent, most students' scores will be hypothetical and may not reflect their true abilities and skills.

Authors contributions: SN: conceived the study, drafted the methodology, performed all analysis as well as the graphs, and concluded the paper; **SA:** drafted the introduction; **TB & SN:** discussed the paper; & **SN, SA, & TB:** reviewed multiple drafts and suggested additions, amendments, and variations. All the authors approve final version of the article.

Funding: The authors received no financial support for the research and/or authorship of this article.

Acknowledgements: The authors would like to thank the reviewers for their intellectual stimulation and positive criticism throughout the development of the manuscript. The authors would also like to thank to the research assistants for their assistant during data collection.

Declaration of interest: Authors declare no competing interest.

Data availability: Data generated or analysed during this study are available from the authors on request.

REFERENCES

Akaranga, S. I., & Ongong, J. J. (2013). The phenomenon of examination malpractice: An example of Nairobi and Kenyatta Universities. *Journal of Education and Practice, 4*(18), 87-96.

Akyeampong, K. (2007). *50 years of educational progress and challenge in Ghana.* http://sro.sussex.ac.uk/id/eprint/2415/

Amedahe, F. K. (2000). *Continuous assessment* [Unpublished paper]. University of Cape Coast, Ghana.

Anane, E. (2008). *High stakes testing, curriculum implementation and instruction.* LAP LAMBERT Academic Publishing.

Azizeh, C., & Mansoor, T. (2010). *Sources of test score pollution: State of the art* [Master's thesis, Islamic Azad University].

Bachman, L., Cushing, S. T., & Purpura, J. E. (2013). *Designing a research questionnaire to explore test taker characteristics* [Paper presentation]. TESOL.

Buzick, H. M., & Jones, N. D. (2015). Using test scores from students with disabilities in teacher evaluation. *Educational Measurement: Issues and Practice, 34*(3), 28-38.https://doi.org/10.1111/emip.12076

Chalak, A., & Tavakoli, M. (2010). Sources of test score pollution: State of the art. *Curriculum Planning Knowledge & Research in Educational Sciences, 7*(26), 13-34.

Chapman, D. W., & Snyder, C. W. (2019). Can high stakes national testing improve instruction: Re-examining conventional wisdom. *International Journal of Educational Development, 20*(5), 457-474. https://doi.org/10.1016/S0738-0593(00)00020-1

Cohen, L., Manion, L., & Morrison, K. (2017). Validity and reliability. In L. Cohen, L. Manion, & K. Morrison (Eds.), *Research methods in education* (pp. 245-284). Routledge. https://doi.org/10.4324/9781315456539-14

Dolly, J., & Williams, K. (1983). *Teaching test-wiseness* [Paper presentation]. The Annual Meeting of the Northern Rocky Mountain Educational Research Association.

Dreisbacha, M., & Keogh, B. (2012). Test-wiseness as a factor in readiness test performance of young Mexican-American children. *Journal of Educational Psychology, 74*(2), 224-229. https://doi.org/10.1037/0022-0663.74.2.224

Filson, Y., & Brown, M. (2018). Effects of test anxiety and testing conditions on intelligence scores of elementary school boys and girls. *Psychological Reports, 32*(2), 643-649. https://doi.org/10.2466/pr0.1973.32.2.643

Flippo, H., & Caverly, D. C. (2008). *Handbook of college reading and study strategy research.* Rouledge. https://doi.org/10.4324/9780203894941

Flippo, R. F., Becker, M. J., & Wark, D. M. (2019). Test taking. In R. F. Flippo, & D. C. Caverly (Eds.), *Handbook of college reading and study strategy research* (pp. 121-144). Routledge.

Genshaft, J., & Kirwin, P. (2018). Improving the study skills of at-risk students. *Special Services in Schools, 5*(3), 109-130. https://doi.org/10.1300/J008v05n03_07

George, D., & Mallery, P. (2012). *IBM SPSS statistics 21 step by step*. Allyn.

Gerald, B. (2018). A brief review of independent, dependent and one sample t-test. *International Journal of Applied Mathematics and Theoretical Physics, 4*(2), 50-54. https://doi.org/10.11648/j.ijamtp.20180402.13

Grant, S. G. (2015). High-stakes testing: How are social studies teachers responding? In W. C. parker (Ed.), *Social studies today: Research and practice* (pp. 53-62). Routledge. https://doi.org/10.4324/9781315726885-11

Gyimah, K. A. (2012). *An evaluation of the practice of continuous assessment in the secondary schools in the Ashanti Region of Ghana* [Unpublished master's thesis]. University of Cape Coast, Ghana.

Haladyna, T. (2016). Test score pollution: Implications for limited English proficient students. In *Proceedings of the 2nd National Research Symposium on Limited English Proficient Student Issues: Focus on Evaluation and Measurement* (pp. 135-163).

Haladyna, T. M., Nolen, S. B., & Haas, N. S. (1991). Raising standardized achievement test scores and the origins of test score pollution. *Educational Researcher, 20*(5), 2-7. https://doi.org/10.3102/0013189X020005002

Haladyna, T., Haas, N., & Nolen, S. (2018). *Test score pollution* [Paper presentation]. The Annual Meeting of the American Educational Research Association.

Hargett, G. (2016). Test score pollution: Implications for limited English proficient students. In *Proceedings of the 2nd National Research Symposium on Limited English Proficient Student Issues: Focus on Evaluation and Measurement.*

Hoover, J. P. (2012). A dozen ways to raise students' test performance. *Principal, 81*(3), 17-18.

Jones, M. G., Jones, B. D., Hardin, B., Chapman, L., Yarbrough, T., & Davis, M. (2019). The impact of high-stakes testing on teachers and students in North Carolina. *Phi Delta Kappan, 81*(3), 199-203.

Kubiszyn, T., & Borich, G. D. (2016). *Educational testing and measurement.* John Wiley & Sons.

Maylone, N. (2014). Do tests show more than "test think"? *The Education Digest, 69*(8), 16-20.

Mereku, K. (2000). *The BECE grading system committee report: Implications for minimum educational qualifications for basic education certificate.* SAGE.

Mitchell, N. (2007). Error labelled four schools as failing. Spanish-language tests were thrown out by computer. *Rocky Mountain News.* http://www.rockymountainnews.com

Nathans, L. L., Oswald, F. L., & Nimon, K. (2012). Interpreting multiple linear regression: A guidebook of variable importance. *Practical Assessment, Research, and Evaluation, 17*(1), 9.

NCME. (2014). Standards for educational and psychological testing. Reduce standardized testing to improve accountability, school quality. *National Council on Measurement in Education.* http://idea.gseis.ucla.edu/newsroom/educationnewsroundup

Newman, L. A., Madaus, J. W., Lalor, A. R., & Javitz, H. S. (2020). Effect of accessing supports on higher education persistence of students with disabilities. *Journal of Diversity in Higher Education, 42*(1), 6-16.

Nitko, A. J. (2001). *Educational assessment of students.* Prentice-Hall.

Nolen, S. B., Haladyna, T. M., & Haas, N. S. (2016). Uses and abuses of achievement test scores. *Educational Measurement: Issues and Practice, 11*(2), 9-15. https://doi.org/10.1111/j.1745-3992.1992.tb00234.x

Oakland, T. (2017). The effects of test-wiseness materials on standardized test performance of preschool disadvantaged children. *Journal of School Psychology, 10*(4), 355-360. https://doi.org/10.1016/0022-4405(72)90004-0

Petty, N., & Harrell, E. (2017). Effect of programmed instruction related to motivation, anxiety, and test wiseness on group IQ test performance. *Journal of Educational Psychology, 69*(5), 630-635. https://doi.org/10.1037/0022-0663.69.5.630

Roger, D. (2012). *An empirical investigation of the effects of the adoption of a secured process on the uniform certified public accountant examination* [PhD dissertation, The University of Mississippi].

Smith, M. L. (2012). Put to the test: The effects of external testing on teachers. *Educational Researcher, 20*(5), 8-11. https://doi.org/10.3102/0013189X020005008

Stricker, L. J. (2013). Test-wiseness on personality scales. *Journal of Applied Psychology, 53*(3), 1-17. https://doi.org/10.1037/h0027380

Tabachnick, B. G., Fidell, L. S., & Ullman, J. B. (2007). *Using multivariate statistics*. Pearson.

Tamakloe, E. K., Atta, E. T., & Amedahe, F. K. (2005). *Principles and methods of teaching*. Black Mask Ltd.

Thomas, A., Hertzig, M., Dryman, I., & Fernandez, P. (2001). Examiner effect in IQ testing of Puerto Rican working-class children. *American Journal of Orthopsychiat, 4*(5), 809-821. https://doi.org/10.1111/j.1939-0025.1971.tb00747.x

Wagaman, J. (2019). *Raising high school reading test scores: Factors that affect reading scores*. http://www.suite101

Wall, D. (2018). The impact of high-stakes testing on teaching and learning: Can this be predicted or controlled? *System, 28*(4), 499-509. https://doi.org/10.1016/S0346-251X(00)00035-X

❖❖❖