P-Issn: 1693-1246
E-Issn: 2355-3812
December 2020

Jurnal Pendidikan Fisika Indonesia 16 (2) (2020) 122-131
DOI: 10/15294/jpfi.v16i2.25238

http://journal.unnes.ac.id/nju/index.php/jpfi

Formative Assessment Practices of Physics Teachers in Pakistan

M. Khan[1]*, T. U. Zaman[1], A. Saeed[2]

[1]Department of Science Education, Allama Iqbal Open University, Islamabad, Pakistan
[2]HIESS, Hamdard University, Karachi, Pakistan

Received: 27 August 2020. Accepted: 11 November 2020. Published : December 2020

Abstract

Formative assessment is an ongoing activity that helps to understand the gaps between a student's current understanding and the objectives to achieve. If combined with appropriate feedback, it has powerful positive impact on students' learning. The objective of the study was to explore the quality of classroom assessment practices of physics teachers in Pakistan. One hundred fifteen (115) principals, one hundred thirty-nine (139) physics teachers and eight hundred (800) 10t^h class science students were selected randomly from 162 target schools. Three (03) validated and reliable 5-point Liker-type scales were developed to collect the perceptions of the respondents. Eighteen hundred checked class tests of 20 physics teachers were also analysed to explore the quality of assessment. It was concluded that the quality of formative assessment is poor, and teachers must be trained to improve feedback process given to the students to enhance their learning.

Key words: formative assessment, quality indicators, written feedback

INTRODUCTION

According to Heritage, Kim, Vendlinski, and Herman (2009), formative assessment is a systematic process of gathering evidences and providing continuous feedback about learning during teaching-learning process. Formative assessment is an ongoing activity that engages both learners and teachers to gauge students' quality of learning and helps designing the necessary steps for improvement. It is an integral part of teaching - learning process. Shepard (2005) asserts that formative assessment carried out to improve students' learning and instructional effectiveness. Its main use is monitoring student's learning and providing the feedback required to adjust the lesson plan, so that each learner can master the concept (s) being taught (LaBay & Comm, 2004). It is effective if it helps in improving students' learning. Bennett (2011) asserts that formative assessment improves students' knowledge and skills. It is usually evidenced several times as teachers and students ask questions, report on their assignments, and make decisions about what to do next (National Research Council, 2001). It helps to understand the differences between a student's current understanding and the objectives to achieve. Furtak, Thompson, and van Es (2016) have divided formative assessment into four major components. These components include (i) designing tasks for formative assessment (ii) asking questions to elicit students' thinking (iii) interpreting students' ideas and (iv) providing feedback in order to move students' thinking forward.

Formative assessment is the heart of quality learning. In Pakistan, most of the classroom assessment check factual knowledge (Browne, 2016; Rind & Malik, 2019). Unfortunately, teachers don't have careful attitude in formative assessment practices (Hussain, Shaheen, Ahmad, & Islam, 2019) and don't realize its importance. There is no formal formative assessment guideline for schools in Pakistan (Asian Development Bank, 2019; Browne, 2016). Formative assessment is a key component of teaching-learning process as it helps teachers in classroom decision-making (Rahim, Venville, & Chapman, 2009). Quality of students' assessment and quality of instructions are linked together. Hughes (2012) indicated that monitoring students' progress affects quality of teaching and hence the quality education.

*Correspondence Address:
Ashfaq Ahmed Rd, H-8/2 H 8/2 H-8, Islamabad 44000, Pakistan
E-mail: hashirdanial@yahoo.com

Pouyioutas (2014) pointed out that receiving student feedback, student participation in academic feedback mechanisms, regular meetings for feedback, follow up of feedback, assessment on clear and agreed outcomes are the key indicators related to teaching-learning process.

Although National Education Policy (2009) has called for an appropriate balance between the use of formative and summative assessment approaches at school level but formative assessment is not made mandatory in school curriculum. Unfortunately, in Pakistan, most of the teachers do not use formative assessment. There is misconception about the formative assessment as teachers continue to use classroom assessments primarily for grading students (Rauf, Shamim, Aly, Chundrigar, & Alam, 2014) and not for learning (Rehmani, 2012). According to Weldmeskel (2015), elements of quality formative assessment include (i) formative feedback (ii) self-assessment and peer assessment, which contribute to learning improvement and (iii) the self-regulation of learning.

Formative feedback is the information passed to the learners for improving their learning. Hattie and Timperley (2007), Shute, Dennen, Kim, Donmez, and Wang (2008) and Lipnevich, McCallen, Miles, and Smith (2014) have recognized the significance of formative feedback on learning improvement. However, Gamlem and Smith (2013) pointed out that it was the least understood aspect in teaching- learning process. In feedback to students most of the teachers give marks only(Ghazal, Gul, Hanzala, Jessop, & Tharani, 2014; Khattak, 2012). It has been well documented that feedback in terms of comments tends to be more useful than grades (Weeden, Winter, & Broadfoot, 2002).

The timing of the feedback is also crucial. Feedback is more effective to enhance students' future performance if it is provided timely (Brookes, 2010). According to Rauf et al. (2014), quality of feedback is the most essential feature in the success of formative assessment and has the maximum impact on student learning and their academic achievement. Mehmood, Hussain, Khalid, and Azam (2012) have mentioned that feedback helps students become aware of any gaps that exist between desired objectives and their current knowledge/ skill/understanding.

In effective feedback practices, Sadler (2009) thinks that feedback is necessary for students to achieve the desired objectives. So far as the nature of feedback is concerned, Arbaugh et al. (2008) have maintained that the most useful type of feedback tests provides unambiguous comments about the errors and mistakes students make and specific guideline for improvement. Such type of feedback motivates students to focus on the learning task rather than on simply getting the right answer. However, the feedback is only helpful to the students if it is clear, specific, timely, accurate, focuses upon the objectives, immediate, based on first-hand data, suggests plans for improvement, relevant, balanced, understandable, of multiple cycles and confidential. Written feedback, mostly, is reduced to marks only (Ghazal et al., 2014; Khattak, 2012). The situation is even worst at secondary level where students hardly get any feedback on their learning. The present study explores the quality of formative assessment practices in secondary schools. The study was delimited to Public Sector Secondary Schools, located in district Bahawalpur, Pakistan. Public Sector Secondary Schools were selected due to the fact that majority of the students (more than 70%) study in these government schools (BISE.Bwp, 2018).

METHOD

The study was descriptive in nature and the quantitative data were collected through two different approaches; survey and analysis of checked class tests. The target population was 162 public sector secondary schools of district Bahawalpur, affiliated with Board of Intermediate and Secondary Education, Bahawalpur. The researchers used Proportional Stratified Random Sampling technique. One hundred thirty nine (139) physics teachers, one hundred fifteen (115) principals and eight hundred (800) 10th class physics students (400 male, 400 female) were randomly selected from 115 randomly selected secondary schools. From 139 selected physics teachers, 20 teachers were randomly selected to analyse their checked class tests. Thirty (30) physics students of each teacher were selected on random basis. Three class tests of each student were analysed. In this way total 1800 class tests of 600 students were analysed. Multiple methods of data collection were adopted for the purpose of triangulation and to get the insight into the problem related to the quality of assessment.

Two kinds of tools were developed :

(i) Quality Indicators Questionnaires for physics teachers,physics students and principals
(ii) A checklist for written test analysis

Questionnaires were developed on 5-Point Likert-Type Scale pattern. Both positive and negative statements were included to reduce extreme biasness (Salazar, 2015), response speed and to promote actual cognitive reasoning in respondents (Podsakoff, MacKenzie, Lee, & Podsakoff, 2003). Quality indicators of assessment included monitoring of students' learning, use of formative assessment, use of various assessment methods, follow up of the feedback, sharing of check answer sheets and provision of timely feedback. The elements of feedback included 'Giving Marks only', 'Pointing out Mistakes', 'Correcting Mistakes', 'Pointing out what is Missing in the written answers', 'Correcting Mistakes plus Indicating Missing', 'Descriptive Feedback', and 'Oral Feedback'.

After the collection of statements, related to the quality of assessment, the tools were got validated from six (06) experts of the related field. Some statements were rephrased, few were deleted and the tools were finalized for pilot testing. On the basis of the feedback of experts and pilot testing, the final versions of the tools were prepared. Questionnaire for students was developed in English and its Urdu version was also developed for its easy understanding. A systematic translation procedure was followed in this regard. Evidences for construct validity are gathered through various aspects. The extent to which the internal components of test match the defined construct is most often estimated by factor analysis. Factor analysis is a mathematical technique used for data reduction and to identify super- variables which are not observable (Lingard & Rowlinson, 2006).

Exploratory factor analysis (EFA) and confirmatory factor analysis (CFA) are two main categories of factor analysis. The Confirmatory factor analysis offers evidence that all items or variables are aligned in a proper way with the right latent variables within the construct being measured. Various fit statistics are used in order to determine whether the proposed model provides adequate fit for the data. If the fit statistics are acceptable, the parameter estimates can then be examined. According to Holtzman and Vezzu (2011), "The ratio of each parameter estimate to its standard error is distributed as a t-statistic and is significant at the 0.05 level if the value exceeds 1.96 and at the 0.01 level if the value exceeds 2.56" (p.3) . In the present study, the confirmatory factor analysis (CFA) was run on the data collected from 139 physics teachers and 800 students. Model Estimates of Confirmatory Factor Analysis (CFA) confirmed constructs related to the questionnaire designed for teachers and principals. The factor 'classroom assessment' was confirmed with their specified variables (quality indicators). All of the fit indices fulfil the respective common acceptance levels. The t-statistics were all greater than 2.56, indicating that the measurement model exhibited a good fit with the data. Based on the CFA, all quality indicators (variables) were retained and used in further analysis.

For Likert Type Scales, Cronbach's alpha is required to be used in order to estimate internal consistency (Gliem & Gliem, 2003). Reliability Coefficient (α) was estimated at 0.79, 0.83 and 0.79 of questionnaires for principals, teachers and students respectively. Alpha coefficient greater than 0.70 is acceptable (Gliem & Gliem, 2003). A check-list was also developed and validated to record the quality of checked written tests.

After the development of the tools, the respondents were approached personally to collect the data. Response rate was hundred percent. The class tests, marked by the Physics teachers were collected for analysis. Statistics used for data analysis were mean, standard deviation, ANOVA, Scheffé test and Chi-Square. In order to compare three groups (principals- teachers- students) one-way Analysis of Variance (ANOVA) was used. Analysis of variance (ANOVA) is a parametric test used to determine whether the difference between two or more means is statistically significant at the selected level of Alpha or not. The researcher used one-way ANOVA. One- way ANOVA is used to test the significant difference for only one variable (Newby, 2014). ANOVA just

tells that there is difference if it exists among three or more groups but cannot tell between what group pairs or due to what treatment group this significant difference exists. There are different post-hoc tests available which can be used. The researcher used Scheffé test to compare pairs of group means in order to assess where the differences lie because it is one of the safest tests of all post-hoc tests (Gravetter & Wallnau, 2014). Chi-Square (χ^2) is a test of significance suitable if the data comes in "frequencies occurring in two or more mutually exclusive categories and compare observed frequencies with expected frequencies (Mills & Gay, 2019) . It is a statistical test used to estimate if two or more classifications of the samples are independent or not.

RESULT AND DISCUSSION

The data collected from the respondents and class tests were tabulated and analysed using various statistical techniques.

Table 1. Comparison between the Responses of Principals, Teachers and Students on the statements related to formative Assessment.

Quality Indicator	Respondents						F_{cal}	F_{tab}
	Principals (n=115)		Teachers (n=139)		Students (n=800)			
	Mean	St.Dev	Mean	St.Dev	Mean	St.Dev		
Monitoring Learning	3.36	1.50	3.98	1.25	2.00	0.92	261.20	3.00
Follow up of Feedback	3.01	1.52	4.06	1.08	3.08	1.27	36.48	3.00
Variety of Assessment Methods	3.70	1.50	4.29	1.03	3.76	1.29	10.04	3.00
Sharing of Answer Sheets	4.25	0.95	4.55	0.78	4.19	0.87	9.87	3.00
Timely Feedback	3.89	1.35	4.52	0.80	3.93	1.02	19.95	3.00
Marks only	3.47	0.98	1.99	0.85	4.16	0.90	349.43	3.00
Mention Mistakes	3.47	0.99	4.22	0.96	2.90	1.12	93.29	3.00
Correct Mistakes	3.17	1.12	4.08	1.12	2.87	1.24	58.87	3.00
Mention Missing	3.08	1.8	4.33	0.83	3.11	1.14	74.86	3.00
Correct Mistakes plus Mention Missing	2.87	0.86	4.27	0.83	2.46	1.09	180.51	3.00
Oral Feedback	2.79	1.24	4.54	0.74	3.00	1.24	105.86	3.00

α = Significance level, F_{cal} =Calculated value of F, F_{tab} =Tabulated value of F, St. Dev =Standard Deviation

Table 1 indicates the comparison of responses of principals, teachers and students on the statements related to the quality of formative assessment. Each statement was analyzed separately. The data were analyzed using one-way ANOVA. For each statement, it was found that there was at least one significant difference among the means of principals, Physics teachers and Physics students at significant level of 0.05. Consequently, the Scheffé test was used to compare pairs of group means in order to assess where differences lied.

Table 2. Results of Scheffé Test

Quality Indicator	F- Values for Scheffé Test		
	Principals – Teachers	Principals – Students	Students-Teachers
Monitoring Learning	7.75	58.17	146.32
Follow up of Feedback	21.98	*0.18	36.02
Variety of Assessment Methods	6.49	*0.10	9.96
Sharing of Answer Sheets	394.30	*2.49	655.48
Timely Feedback	11.54	*0.06	19.25
Marks only	84.60	29.05	341.03
Mention Mistakes	14.97	23.92	87.45
Correct Mistakes	17.80	49.60	58.41
Mention Missing	41.00	*0.05	72.88
Correct Mistakes plus Mention Missing	174.78	7.85	179.55
Oral Feedback	67.90	*1.56	98.93

F_{tab} =Tabulated value of F, α = 3.06 *non-significant values

Table 2 shows the results of Scheffé test. Scheffé test was used to check whether the difference between means of principals-teachers, principals- students and students- teachers was significant. It was found from the results of Scheffé test that means of all three groups (for monitoring learning) were significantly different from each other. Mean score of physics teachers (M=3.98) was significantly higher different from the mean score of principals (M=3.36) and students (M=2.00). Mean scores reflected that teachers and principals agreed that physics teachers monitored students' learning during classroom teaching whereas students disagreed with this statement and perceived that Physics teachers did not monitor learning.

For 'Follow up of Feedback', it was found from one-way ANOVA test that there was at least one significant difference among the means of principals, teachers and students at significant level of 0.05. The results of Scheffé test indicated that means of principals (M=3.01) and Means of students (M=3.08) were not significantly different from each other. Mean score of teachers (M=4.01) was significantly different from the mean score of principals (M=3.01) and students (M=3.08). Mean scores reflected that students and principals remained neutral regarding the statement related to 'follow-up of feedback' given to students. However, Physics teachers perceived that they followed up feedback given to students.

Similarly, comparison of responses of principals, teachers and students on the statements related to 'variety of assessment methods', 'Sharing of Answer Sheets', 'Timely Feedback', 'Marks only', 'Indicating Mistakes', 'Correcting Mistakes', 'Mentioning Missing', 'Correcting Mistakes plus Mentioning Missing' and 'Oral Feedback' was made using one-way ANOVA and it was found that there was at least one significant difference among the means of principals, teachers and students at significant level of 0.05. It was found from the results of Scheffé test that means of principals and Means of students related to 'Variety of Assessment Methods', 'Sharing of Answer Sheets', 'Timely Feedback', 'Mention Missing,' and 'oral feedback' were not significantly different from each other. Mean scores reflect that teachers, students and principals agreed that Science teachers share marked answer sheets with students. Mean scores reflected that teachers disagreed with the statement and pointed out that they did not give only marks to the students. However, students and principals agreed that most of the

time, Science teachers gave marks only to students on class tests. A mismatch was found between teachers and principals and students. Students perceived that Science teachers did not mention mistakes to students on class tests.

There was a mismatch found between Physics teachers, principals and students. Mean scores reflected that physics teachers agreed with the statement and perceived that they corrected students' mistakes. Similarly, principals perceived that sometimes physics teachers corrected mistakes, but students disagreed and perceived that physics teachers rarely/sometime corrected mistakes in class tests. There was a mismatch found between teachers and principals and students perceptions related to the correction of mistakes. Mean scores reflected that physics teachers agreed with the statement and perceived that they often mentioned missing and corrected mistakes. However, principals disagreed that physics teachers sometimes corrected mistakes and mentioned missing also. On the other hand, students claimed that physics teachers rarely corrected mistakes and mentioned missing also. There was a mismatch found among Physics teachers and principals and students.

Table 3. Analysis of Written Class Tests

Quality indicator	Tests Analyzed	Presence of indicator				df	A	Chi-Square (X^2)	
		Yes		No				C.V.	T.V.
		Freq	%	Freq	%				
'Marks Only'	1800	1133	63%	667	37%	1	0.05	121	3.84
'Point Out Mistakes'		231	13%	1569	87%			995	
'Correct Mistakes'		176	10%	1624	90%			1165	
'Mention Missing'		136	8%	1664	92%			1297	
'Correct Mistakes plus Mention Missing'		124	7%	1676	93%			1338	
'Descriptive Feedback'		87	5%	1713	95%			1468	

Note. C.V. =Calculated Value, T.V. = Table Value, C.V. > T.V. at significant level of 0.05

Table 3 shows the results related to different quality indicators of written feedback, given to the students on classroom written tests. As shown in Table 3, sixty-three percent (63%) (1133 out of 1800 tests) tests, marked by the physics teachers, indicated that teachers gave only marks in the class tests. As shown in Table 3, calculated value of Chi –Square (121) is greater than Table value (3.84) at the significance level of 0.05. It means the difference between the frequencies 'Yes' and 'No' is statistically significant. It shows that in most of the cases, physics teachers give marks only to the students and do not include other elements of feedback.

As shown in Table 3 above, calculated value of Chi –Square is greater than Table value (3.84) for all quality indicators at the significance level of 0.05. It means the difference between the frequencies 'Yes' and 'No' is statistically significant in all cases. As indicated above, in 13% class

tests, Physics teachers pointed out mistakes. Ninety percent (90%) tests, marked by the Physics teachers, indicated that teachers did not point out mistakes on the class tests. It means most of the physics teachers do not correct students' mistakes they (students) make in written tests. Only in 7% class tests, physics teachers provided both elements (error correction and mention missing also) of feedback. As shown in Table 3 above, 95% (1713 out of 1800 tests) texts, marked by the teachers, indicated that teachers did not provide descriptive Feedback to students on class tests. It means most of the physics teachers do not correct students' mistakes they make in written tests.

Sample of marked tests

Next three figures show the class tests marked by three different physics teachers for the purpose of formative assessment.

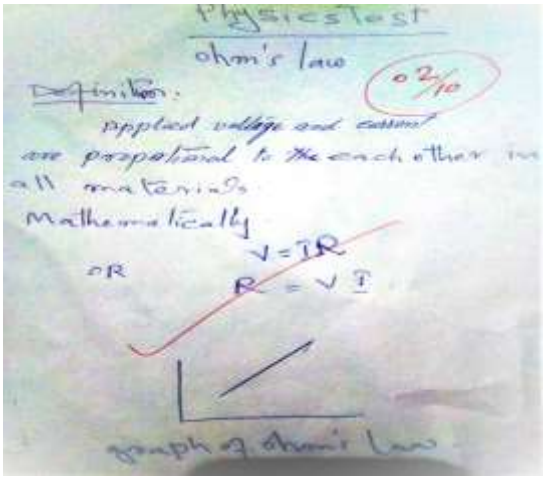

Figure 1. Class test marked by a Physics teacher (teacher-1)

Figure 1 shows a test, marked by a physics teacher (teacher-1). It was a routine class test conducted for the purpose of formative assessment. As indicated in figure 1 above, physics teacher (teacher-1) just awarded total marks (02/10) to the student. It did not indicate what were the student's mistakes, what was missing in it, how student could improve (formative feedback) etc.

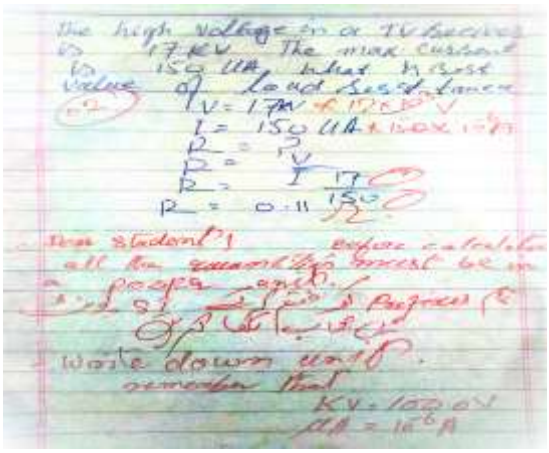

Figure 2. Class test marked by a physics teacher (teacher-2)

Figure 2 shows a test, marked by another physics teacher (teacher-2). As shown in figure 2, physics teacher, did not only awarded marks but also indicated what was missing, provided the required correct answer/response, and gave the comprehensive written feedback to the student. The feedback given in the test, shown in figure 2 was better than the feedback given in the test shown in figure 1 as more quality indicators of feedback were given in it.

Figure 3. Class test marked by a physics teacher (teacher-3)

Figure 3 shows another test marked by a different physics teacher (teacher-3). This physics teacher pointed out mistakes, corrected them and guided the student in a better way.

It means formative assessment is useful when it gives a comprehensive feedback to the students and includes its various components.

Figure 4. Presence of Quality Indicators of written Feedback

Formative assessment aims at monitoring teaching - learning process in order to provide on-going feedback to teachers and students. It can serve to provide constructive information only if it is comprehensive. The findings of the present study reveal that there is a poor classroom assessment system in Pakistan. Formative assessment practices usually focus on recalling information and provide only little feedback to students. Teachers select questions directly from exercises at the end of the chapters. According to Browne (2016), in Pakistan, classroom/formative assessment is not used to support learning. No mechanism is followed to monitor the quality of classroom assessment practices systematically. Browne (2016) observes that Generally, In Pakistan, classroom formative assessment is an information collection exercise as its much focus is on what is taught, and not on what is learnt. Feedback, mostly, is reduced to marks only. Khattak (2012) and Ghazal et al. (2014) reached on the same conclusion. The findings of Browne (2016) that teachers correct student' mistakes in feedback contradict with the findings of the present study as no such evidence has been collected. From the above analysis, it is concluded that quality of feedback is very poor in terms of the mentioned components. Only marks, the weakest form of feedback, are given to students. The results of the present study have been confirmed by the previous studies (Ghazal et al., 2014; Khattak, 2012; Norouzian & Farahani, 2012).

CONCLUSION

The present study discusses the quality of on-going formative assessment practices by physics teachers at secondary level in district Bahawalpur, Pakistan. Perceptions of principals, teachers and students have been compared and actual classroom assessment practices have been explored through the examination of checked written class tests. Mismatches have been found in the perceptions of principals, teachers and students. Similarly, mismatches have been found in the perceptions of physics teachers and the findings of the analysis of written class tests. It has been concluded that physics teachers do not monitor students' learning process in a proper way. However, it cannot be concluded whether teachers follow feedback, given to students or not. Variety of formative assessment methods are used during classroom teaching. Checked answer sheets are shared with students in time. Oral feedback is not given all the time. It can be concluded that physics teachers' feedback to students has been reduced to marks only. Such a feedback cannot be considered as an effective feedback and such type of formative assessment is considered as a poor assessment. Teachers should understand that productive formative assessment approaches with comprehensive feedback mechanism is imperative for meaningful learning.

REFERENCES

Arbaugh, J. B., Cleveland-Innes, M., Diaz, S. R., Garrison, D. R., Ice, P., Richardson, J. C., & Swan, K. P. (2008). Developing a community of inquiry instrument: Testing a measure of the community of inquiry framework using a multi-institutional sample. The internet and higher education, 11(3-4), 133-136.

Asian Development Bank. (2019). School Education in Pakistan: A Sector Assessment.

Bennett, R. E. (2011). Formative assessment: A critical review. Assessment in education: principles, policy & practice, 18(1), 5-25.

Board of Intermediate and Secondary Education (2018). Result Gazette Secondary School Certificate

(annual) Examinnation,2018. Retrieved from https:// www.bisebwp.edu.pk /Attachments/ 96479350-8daa-483a-97ea-dda52946c78dSSC _A18_FirstPages_2018.pdf

Brookes, M. (2010). An evaluation of the impact of formative feedback podcasts on the student learning experience. Journal of Hospitality, Leisure, Sports and Tourism Education (Pre-2012), 9(1), 53.

Browne, E. (2016). Evidence on formative classroom assessment for learning.

Furtak, E. M., Thompson, J., & van Es, E. (2016). Formative assessment and noticing: toward a synthesized framework for attending and responding during instruction. Paper presented at the Annual Meeting of the American Educational Research Association, Washington, DC.

Gamlem, S. M., & Smith, K. (2013). Student perceptions of classroom feedback. Assessment in Education: Principles, Policy & Practice, 20(2), 150-169.

Ghazal, L., Gul, R. B., Hanzala, M., Jessop, T., & Tharani, A. (2014). Graduate students' perceptions of written feedback at a private university in Pakistan. International Journal of Higher Education, 3(2), 13.

Gliem, J. A., & Gliem, R. R. (2003). Calculating, interpreting, and reporting Cronbach's alpha reliability coefficient for Likert-type scales.

Gravetter, F., & Wallnau, L. (2014). Introduction to the t statistic. Essentials of statistics for the behavioral sciences, 8, 252.

Hattie, J., & Timperley, H. (2007). The power of feedback. Review of educational research, 77(1), 81-112.

Heritage, M., Kim, J., Vendlinski, T., & Herman, J. (2009). From evidence to action: A seamless process in formative assessment? Educational Measurement: Issues and Practice, 28(3), 24-31.

Holtzman, S., & Vezzu, S. (2011). Confirmatory factor analysis and structural equation modeling of noncognitive assessments using PROC CALIS. NorthEast SAS Users Group (NESUG), 2011 proceedings, 11-14.

Hughes, J. R. (2012). How does the use of science notebooks as formative assessment increase students' conceptual understandings in science?

Hussain, S., Shaheen, N., Ahmad, N., & Islam, S. U. (2019). Teachers' classroom assessment practices: challenges and opportunities to classroom teachers in Pakistan. Dialogue, 14(1), 88.

Khattak, S. G. (2012). Assessment in schools in Pakistan. SA-eDUC, 9(2).

LaBay, D. G., & Comm, C. L. (2004). Student expectations regarding online learning: Implications for distance learning programs. Journal of College Teaching & Learning (TLC), 1(10).

Lingard, H. C., & Rowlinson, S. (2006). Sample size in factor analysis: why size matters. Hong Kong: University of Hong Kong.

Lipnevich, A. A., McCallen, L. N., Miles, K. P., & Smith, J. K. (2014). Mind the gap! Students' use of exemplars and detailed rubrics as formative assessment. Instructional Science, 42(4), 539-559.

Mehmood, T., Hussain, T., Khalid, M., & Azam, R. (2012). Impact of formative assessment on academic achievement of secondary school students. International Journal of Business and Social Science, 3(17).

Mills, G. E., & Gay, L. R. (2019). Educational research: Competencies for analysis and applications: ERIC.

National Education Policy. (2009). National Education Policy 2009. Ministry of Education Islamabad, Pakistan.

National Research Council. (2001). Classroom assessment and the national science education standards: National Academies Press.

Norouzian, R., & Farahani, A. A. K. (2012). Written error feedback from perception to practice: A feedback on feedback. Journal of Language Teaching and Research, 3(1), 11.

Podsakoff, P. M., MacKenzie, S. B., Lee, J.-Y., & Podsakoff, N. P. (2003). Common method biases in behavioral research: a critical review of the literature and recommended remedies. Journal of applied psychology, 88(5), 879.

Pouyioutas, P. (2014). Key Performance Indicators for the Teaching in Higher Education Quality Model (THEQM). Paper presented at the International Conference Edulearn Barcelona.

Rahim, S. S. A., Venville, G., & Chapman, A. (2009). Classroom assessment: Juxtaposing teachers' beliefs with classroom practices. Paper presented at the Australian Association For Research In Education: International Education Research Conference.

Rauf, A., Shamim, M. S., Aly, S. M., Chundrigar, T., & Alam, S. N. (2014). Formative assessment in undergraduate medical education: concept, implementation and hurdles. J Pak Med Assoc, 64(64), 72-75.

Rehmani, A. (2012). Changing assessment practices in Pakistani schools: A case of AKU-EB middle school assessment framework.

Rind, I. A., & Malik, A. (2019). The examination trends at the secondary and higher secondary level in Pakistan. Social Sciences & Humanities Open, 1(1), 100002.

Sadler, D. R. (2009). Transforming holistic assessment and grading into a vehicle for complex learning Assessment, learning and judgement in higher education (pp. 1-19): Springer.

Salazar, M. S. (2015). The dilemma of combining positive and negative items in scales. Psicothema, 27(2), 192-199.

Scheerens, J., Luyten, H., & Van Ravens, J. (2011). Perspectives on educational quality: Illustrative outcomes on primary and secondary schooling in the Netherlands: Springer Science & Business Media.

Shepard, L. A. (2005). Linking formative assessment to scaffolding. Educational leadership, 63(3), 66-70.

Shute, V. J., Dennen, V. P., Kim, Y.-J., Donmez, O., & Wang, C.-Y. (2008). 21st century assessment to promote 21st century learning: The benefits of blinking. Unpublished manuscript, Florida State University, Tallahassee.

Weeden, P., Winter, J., & Broadfoot, P. (2002). Assessment: Psychology Press.

Weldmeskel, F. M. (2015). The use of quality formative assessment to improve student learning in West Ethiopian universities.