



RESEARCH PAPER

Rubric Effects on Assessment: An Analysis

Muhammad Mooneeb Ali ¹ Tayyaba Yasmin ² Jawaria Afzaal ³

1. Lecturer, Department of English Government College of Science Wahdat road Lahore, Punjab, Pakistan
2. Lecturer, Department of English, University of Education Lahore, Punjab, Pakistan
3. Lecturer, Department of English, The University of Lahore, Lahore, Punjab, Pakistan

PAPER INFO

Received:
April 27, 2020

Accepted:
June 15, 2020

Online:
June 30, 2020

Keywords:
Assessment
Rubric Effects
Testing

Corresponding Author:

mooneebai@hotmail.com

ABSTRACT

Testing is an integral part of language education. This article aims at analyzing the previously conducted researches related to the effects of rubric on assessment. Assessment has always been an important aspect in language testing. In the arena of research studies related to language testing Assessment fall into two categories i.e. holistic assessment and rubric based assessment. Data is collected through corpus where quantitative research method has been used. The data collected from ten research articles from discipline related journals of language testing and assessment. The findings suggest that rubric has rigorous impact on assessment. It helps to make valid assessment. The overall findings also help to highlight the value of rubric in academic context.

Introduction

The function of review articles is very pivotal for recognizing new research directions, to build up support for prevailing notions and theories and to classify the patterns among the current research studies. There is a need to write reviews after publication of thirty or forty research articles in any discipline (Price, 1965). One of the major purposes of review studies is to evaluate the existing work of literature. Review articles play an important role for creation of knowledge and dissemination. This justifies and demonstrates the importance of the current study because less attention has been paid to review studies in the field of language testing. Language testing has always been an important area in the field of education as tests and tasks not only help in finding out the problem faced by students by also help to improve methods of teaching and learning. It has long been used in education for various high stakes decision-making purposes, e.g., selection and placement, guiding and correcting learning, and grading achievement (Crooks, 1988). The current study endeavors to focus on the researches in the field of language testing and assessment. Palomba & Banta (1999) "Assessment is the systematic collection, review, and use of information

about educational programs undertaken for the purpose of improving learning and development” (p.4). Writing assessment has three major shifts in methods used in assessing writing. The first wave of writing assessment (1950-1970) sought objective tests with indirect measures of assessment. The second wave (1970-1986) focused on holistically scored tests where the students' actual writing began to be assessed. And the third wave (since 1986) shifted toward assessing a collection of student work (i.e. portfolio assessment) and programmatic assessment (Hout & Neal 2006).

According to ten articles studied assessment can be categorized into two types: holistic assessment and Rubric based assessment. Holistic assessment is used where learning or performance objectives are inter-related and complex and the extent of learning or performance is measured against established standards (Biggs, 2007) At the present time, the majority of language arts educators believe that learning is a holistic process that involves the social construction of meaning. In language classrooms, students are actively engaged in authentic listening, speaking, reading and writing experiences. Therefore Holistic assessment cannot serve the purpose as accuracy and organization is key feature of writing assessment (Cox, 2002). It is a fact that high quality assessment and evaluation of any performance depends on accurate and reliable measurement of key performance. So for explicit, organized and channelized feedback rubrics are used by the assessors.

Defining Rubric

Andrade (2000) defines a rubric as ‘a scoring tool that lists the criteria for a piece of work’ and one that ‘articulates gradations of quality for each criterion, from excellent to poor’ (p.1). Rubrics are descriptive scoring plans that are developed by teachers, or students and teachers or other evaluators to guide the examination of the products or processes of students’ efforts (Moskal, 2000). Rubrics are assessment tools, which facilitate the process of evaluation and reporting of student achievement by educators. They are ‘descriptive scoring schemes’, ‘a set of scoring guidelines’ and get their importance as a direct result of the emphasis on usage of constructed responses and performance based tasks in business programs (Hafner and Hafner, 2003; Quinlan, 2006; Glickman-Bond and Rose, 2006)

Significance of the Rubric

Traditionally, educators have kept assessment criteria of productive skills to themselves. They often expect students to know what makes a good presentation or piece of writing (Andrade, 2000). In today’s world, knowledge is generated, shaped and communicated so fast that it has become a fundamental priority and a basic attribute of prosperous societies (Sert, 2008).. In this educational setting, where the simple recitation of knowledge is not considered as meaningful, constructivist approach that permits the reformulation and reconstruction of knowledge is progressively gaining more popularity in many countries (Sert, 2008). This model of learning calls for authentic performance based assessments in classrooms including the language arts classrooms (Cox, 2002).

Performance assessment requires a student to perform a task or make up his or her own response (Linn & Gronlund, 2000; Wiggins & McTighe, 2005; Brualdi, 2002). Therefore, they are well-suited for evaluating complex learning outcomes such as critical thinking, communication and problem solving skills (Perlman, 2002). In contrast to most traditional forms of testing, performance-based assessments don't have clear-cut right or wrong answers. Rather, there are degrees to which a person is successful or not. For this reason, teachers should assess performance in a way that will permit them to consider those varying degrees. This can be accomplished by creating rubrics (Brualdi, 2002). In fact the use of rubrics as tools of assessment of student performance is more popular nowadays than ever before (Rezaei & Lovorn, 2010).

To inform students about the assessment process and to enable them to succeed, teachers mostly prefer using rubrics as a tool to communicate expectations for an assignment, provide focused feedback on work in progress, and to grade students' performance (Andrade, 2000, Goodrich, 1997). For this reason, teachers should assess performance in a way that will permit them to consider those varying degrees. This can be accomplished by creating rubrics (Brualdi, 2002). In fact the use of rubrics as tools of assessment of student performance is more popular nowadays than ever before (Rezaei & Lovorn, 2010). A rubric offers specific criteria for describing student performance at different levels of proficiency for a certain task (Burns, Roe, & Smith, 2002).

Aims of the previous studies in Language Testing

These research studies found to fall into two categories one effects of rubric from teachers perspective and effects from students perspective.

This section covers the aims of the ten articles studied in language testing arena. It will try to discuss about the set aims of different studies examined and also will attempts to give the brief analysis and description if some studies have similar aims or if some have different aims related to the effect of rubric on language testing.

All the studies examined the effects of rubric in different contexts and perspectives. Four studies discussed the impact of rubric on the assessment from student's perspective. Manal (2011) studies the impact of rubric in the writing class room, but from the perspective of teachers' and students'. Howell (2011) explored the impact of grading rubric use on student academic performance. Julas (2014) aimed to find out if rubrics have positive impact on the students or negative. Three studies focuses on the impact of rubrics but from the view point of writing. Jeong (2015) studies the impact of rubric use in assessing short EFL descriptive writing. Turgut and Kayaoğlu (2015) investigated the effect of using rubrics as an instructional tool on students' writing performance in English as a foreign language. De silva(2012) expressed the effects of provision of rubrics used for assessing to students before they do their assessment tasks. Another studies examined rubric from e learning aspect. Hack (2013) discussed the use of rubric on distance learning program. In Pakistani context Qasim (2015) studied the viewpoints of Pakistani university teachers on the effectiveness of rubrics to assess their students' writing. Ghalib and Al-Hattami (2015) investigated the performance of holistic and analytic scoring rubrics in the context of

EFL writing. Reddy and Andrade (2010) explored use of rubrics at the post-secondary level.

The aims of the previous studies can be better understood by the figure 1:

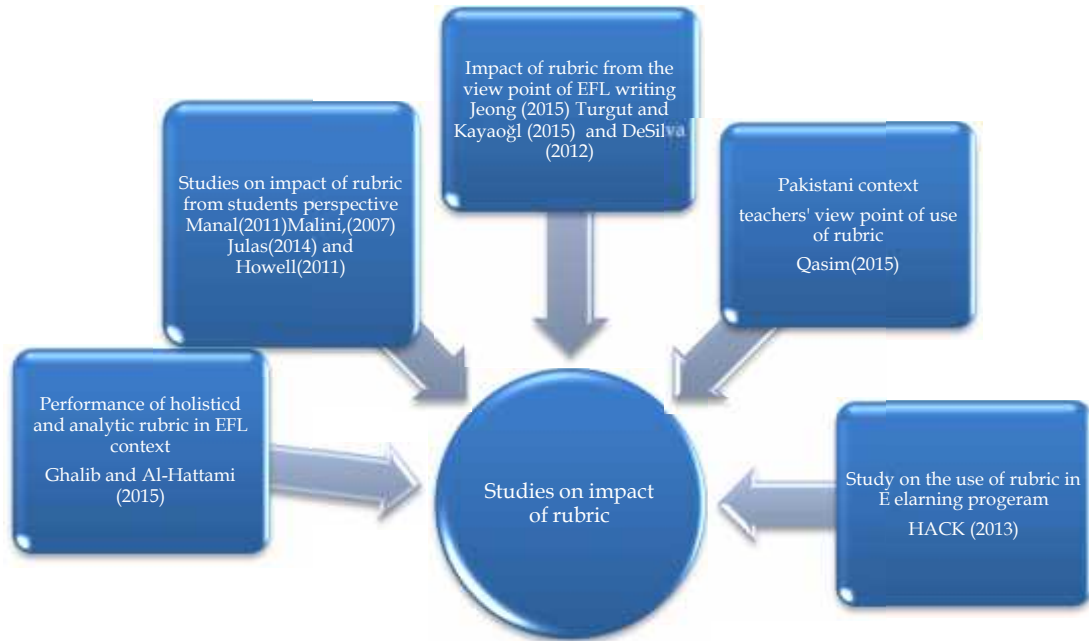


Figure 1: Classification of Research studies on Impact or Rubric

Setting and origin of the reviewed Studies

The next objective is to search out the origin and setting of reviewed studies. Apart from two all these reviewed studies have been conducted in different EFL setting like (Manal 2011, Lebanon; Julas, 2014 Bangladesh; Jeong, 2015 Korea; Qasim, 2015 Pakistan; Ghalib and Al-Hattami. Yemen; Reddy and Andrade, 2010 India; Turgut and Kayaoğlu, 2011 Turkey; De Silva, 2012 Sri Lanka) Only two studies i.e. Hack (2013) Howell (2011) were conducted in America

These studies have been carried out in outer as well as expanding circle.

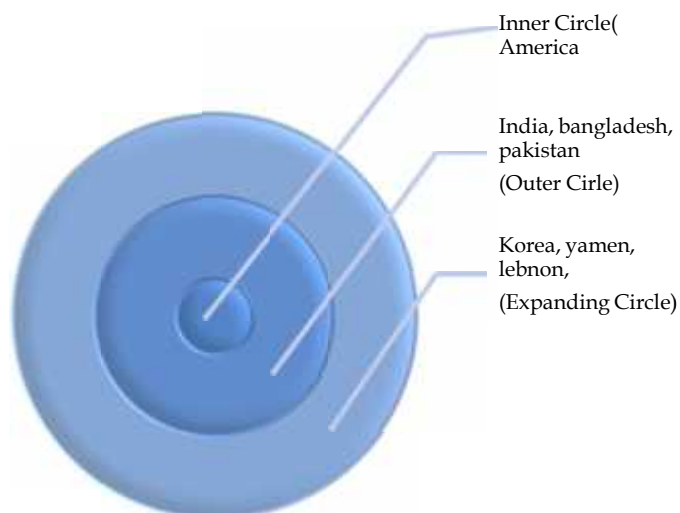


Figure 2: Origin of the Research Studies

Material and Methods

This section explains the methodology, Population, sampling, research design data collection, data analysis discussion and findings of the study. In the current study, previous 6 year's research studies in the field of language testing related to the rubrics have been randomly reviewed to find out the relevance of different articles with the current study. This review basically is done with the aspect of pertinent literature and purposes of different relevant researches to draw conclusion and findings.

Research Data Collection

The corpus used for this study consists of ten articles in the field of language testing. Each article is of 20-22 pages average published in the different international journals. The collection of the data was done from the different online journals like

Journal on Excellence in College Teaching, Language testing in Asia Journal of language and linguistic studies English language Teaching Published by Canadian Center of Science and Education. The sampling is confined only to the publications of language testing with particular focus on studies conducted on raters. The research data spans over a period of 6 years out of which the articles were randomly chosen for writing a review article. All the studies were found out using different websites like, bookzz.org, Google Scholar, HEC Digital Library, and Google Search Engine. From multiple important terms and key words only pertinent key words were selected and used like, Language testing, rubrics, assessment, effect of rubric on assessment. Later on different databases were also used like bookzz.org, booksc.org, linkspringer.com, tandfonline.com. The main articles related to this research were selected by reading the abstracts and the key words for multiple researches and that pertinent to the subject matter were given preference. This random selection depended upon subject and topic relevance, quality, and time span of publication.

The articles reviewed for this review had different techniques and dimensions for data collection. In the 6 studies data was collected from the students. Howell (2011) collected data from 80 undergraduates enrolled in an elective course at a research university during spring and fall 2009 to find out the effects of rubric. Ghalib and Al-Hattami (2015) collected the data from 30 participants attending an English undergraduate program in a Yemeni university. Desilva (2012) collected data from secondary school students in a government school in Sri Lanka. Julas(2014) collected data in two phase first from 18 students and then later on 46 students to ascertain his point of view in relation with the effects of rubrics. Turguta and Kayaoğlu(2015) collected data from the EFL students studying in a university in turkey. Hack (2013) collected data from the students on a distance learning module to find out the impact of the rubric.

Another 4 studies data was collected from the teacher i.e. Jeong (2011) collection of Data for his study came from teacher-raters' essay ratings, rating justifications and interviews. Manal (2011) gathered the data from a third grade writing classroom in a private school in Lebanon and also did interviews from the teachers in order to find out both teacher and students perspective. Reddya and Andrade (2010) gathered the data from student and teachers point of view about rubrics. In another study in Pakistani context Qasim (2015), the data was collected from 6 teachers 3 literature and 3 linguistics university teachers about the impact of rubric.

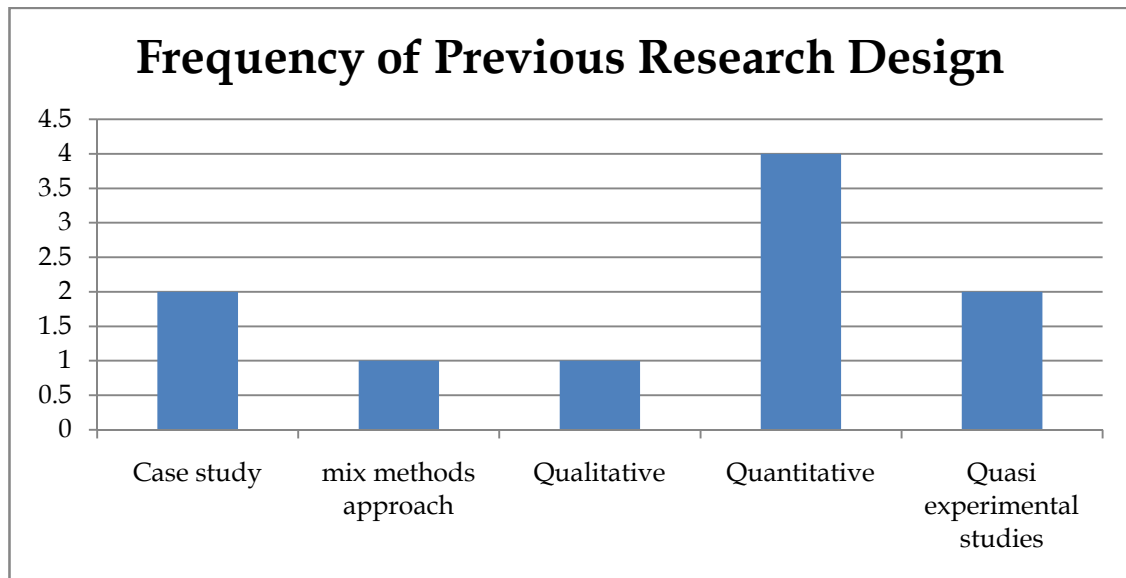
Research Design of the previous Studies

The research designs for the studies reviewed were different in nature from each other.

Two Studies i.e. Howell (2011) used a quasi-experimental evaluation research design in order to conduct his study in which he explored the impacts of rubrics. Turgut, F., & Kayaoğlu, M. N. (2015) used quasi-experimental research design for their specific study. Quasi experimental A quasi-experiment is an empirical study used to estimate the causal impact of an intervention on its target population without random assignment. Quasi-experimental research share Quasi-experimental research shares similarities with the traditional experimental design or randomized controlled trial, but they specifically lack the element of random assignment to treatment or control. Instead, quasi-experimental designs typically allow the researcher to control the assignment to the treatment condition, but using some criterion other than random assignment (e.g., an eligibility cutoff mark).

Four studies focused on quantitative research design. Julas (2012) used quantitative approach to carry out the impact of rubrics from students' perspective. Ghalib and Al-Hattami (2015) used quantitative research design for their study. De silva (2012) used quantitative methods as a research design for the study. Hack(2013) also used quantitative methods as a research design for his study. Ready and Andrade (2010) used qualitative methods for their specific study. Manal (2011) Qasim (2015) focused on case study research design.

The research design used in this study is a case study. A case study typically involves the observation of a single entity such as an individual, a class, a situation, a school or even a community. Jeong (2015) in his study used a mixed-methods approach, using both quantitative and qualitative methods to strengthen the methodology for this study.



Population of the previous studies

This section discusses the aspect about the population of the ten articles studied for the review. Howell(2011) took the population of the 200 student session 2009 in American university. The university in which the course was taught is located in a large Southeastern U.S. town. Jeong(2015) selected the teacher-raters for this study and those were native EFL instructors from a large private Korean university. All were experienced teacher-raters who had a wide range of teaching experience from 5 ~ 12 years in higher education.

Turgut, F., &Kayaoğlu, M. N. (2015).explored and selected 38 university students aged 18-20, attending the intensive English preparatory program at the School of Foreign Languages at Karadeniz Technical University in Trabzon, Turkey. Julas (2012)18 students of BRAC University. Another questionnaire survey was conducted among 46 Medhabikash students of BRAC University and also 8 teachers from the same University.Ghalib1 & Abdulghani A. Al-Hattami (2015) selected the participants of the study which consisted of 30 male and female Yemeni undergraduate students of English at the Faculty of Arts, Taiz University. MANAL(2011) chosen population which consisted of an elementary English teacher and 21 third grade students who learn English as a second language. Qasim (2015) chosen the population which comprised of three language teachers and three literature teachers to find out impact of rubric on teachers. All the teachers teach English language and literature at both undergraduate and graduate level as a foreign language. Three participant teachers

had taught English for 1-10 years while the other three participants had 10-15 years of English teaching experience. Desilva (2012) selected the population which consisted of 70 secondary school students in a national school in Sri Lanka. Two classes were selected for this purpose and students and their teacher of English were briefed about the study. Hack (2013) population for the current study are the 50 student assignments and also five tutors to find out effects of rubric

Instrument used for data collection in the reviewed studies

This section gives details regarding the research tool employed for data collection purposes in the reviewed studies. The instruments used for evaluation of writing skill were written tests i.e. essays. Howell (2011) provided and employed an English test in which only essay was used as an instrument for the students, this essay was following the general techniques of writing an essay. Jeong (2011) provided an anonymous topic to the students and then rubric which followed a standard was given to the raters. Similarly Qasim (2015) also provided an essay to the students and later on he conducted interview as tool of research to finalize his findings

In the remaining studies instruments used for data collection varied. In the research conducted by Turguta, and Kayaoğlu (2015) the instrument was essay but it was given to the university students and they also conducted interviews of students. In the research conducted by Julas (2014) both questionnaire which were provided to the students and teacher and essay on an anonymous topic were the tools of research. Ghalib & Al-Hattami (2015) used essay as a writing sample for the population and essay was their tool for data collection. Manal (2011) tool for data collection were interview of the teacher as well as written essay for the assessment. However, De silva (2012) collected writing samples from 70 students by giving a specific topic of an essay. Qasim (2015) used interview as their tool for data collection as their research was about teachers' perception regarding impact of rubric. Hack (2013) also employed essay as the tool for the collection of data

Instrument development (standardized or developed by the researcher)

Some researchers employed self-developed topics used in essay. In Howell's study (2011) only a self-employed topic was given to the students to write an essay. Similarly Jeong (2011) gave an anonymous topic and later on a rubric which was also self-employed to the students to find out impact of rubrics on students. Turguta, & Kayaoğlu (2015) firstly developed a questionnaire of the interview which was self-developed and before that gave topic of essay to students which was also not standardized. A structured questionnaire was developed by Julas (2014) which was self-developed and later on an interview questionnaire was also structured by him to find on impact of rubric on students. Ghalib & Al-Hattami (2015) gave an essay to students which were self-developed to evaluate the impact of holistic and analytical rubric on students. Manal (2011) gave an anonymous topic to the student for writing an essay which was not standardized and they also conducted an interview which was self-structured. Qasim (2015) conducted evaluators interview and the specific interview was self-structured and not standardized. De silva (2012) diagnosed a speaking and writing

test to students which was self developed idealising an I.E.L.T.s rubric. Hack(2013) gave a topic for writing assessment which was developed by himself.

Validation concerns of reviewed research studies

This section deals with the validation concerns of the reviewed studies. Most of the researchers have used non-standardized essay for the evaluation of the students and some researchers have taken interviews which were structured but was not following any standardized technique. These techniques can also be utilized by future researchers. However the validation issues of rubrics and consistency of rating have been resolved by measuring inter-rater reliability through Correlation Coefficients. As for as the validation concerns of Ghalib& Al-Hattami (2015) study, three different types of statistical analysis were carried out to investigate teachers' internal consistency, based on: including fit statistics, proportions of large standard residuals between observed and expected scores and a single rater-rest of the raters (SR/ROR) correlation. Most of the researchers used raters as well as reasonable number of students in their population to carry out the validity of the research. No research was conducted on the natives and Essays, rubrics; evaluation instruction was given keeping in view the student level and the rater's educational background and experience.

Findings of the previous studies

This part describes and discusses about the finding of the previous articles reviewed. There are different finding of the ten article studied but all are according to their pertinent aims

Howell (2011) suggests that that no real group differences existed for the impact of rubric with respect to *subject*, *Pre-Test*, and *Gender*. Jeong's (2015) findings show that the inexperienced raters in the study were strongly influenced by the mechanical and grammatical aspects of students' writing, and this focus didnot change by rating with a rubric. Turguta, & Kayaoğlu's (2015) results revealed that students who received the rubric outperformed the students in the controlgroup. Julas (2014) found out that students strongly believe that rubrics have impact on their performance. Ghalib& Al-Hattami (2015) Found out that findings showed that there were no significant difference, $F(2, 87) = 0.373, p = 0.690$, among the three raters when they used analytical rubric to grade students' performance. However, the raters scorings did significantly differ, $F(2, 87) = 4.833, p < .05$, when they used holistic rubric. Post Hoc analysis was run to find where the differences lie. The results showed that the difference was between rater 2 and rater 3 at $P < 0.05$. Manal (2011) findings showed that teachers and students are in favor of using rubrics due to the positive impact that they have on the teaching-learning process. Qasim(2015) Found out and indicated that the university teachers, who teach literature, give significant importance to detail and elaboration. Four respondents, out of six, agree that rubrics are effective for grading writing and should be used to address mechanical errors in students' work. They perceived rubric as an effective tool to assess students' progress and teachers' practices. De silva (2012) found that the group which received rubrics plus explanation of what is meant by each criterion and how the grading is done performed significantly better than the group

which received the rubric without any explanation. Hack (2013) found by having the feedback from students indicated that they found the rubrics very helpful in clarifying performance and promoting self-assessment, whilst the tutors felt that it was a time efficient and informative method of providing feedback.

Conclusion

In nutshell, research studies in the field of language testing were reviewed in this study. This study was delimited to the studies related to raters only. These review studies enabled researcher to realize the objectives of the current study. Conclusion is based on the findings of these studies.

These studies have some limitations. First, the number of raters in some of the studies curtaining to each group was small for the results to be generalizable to the entire rater population; even with the representativeness of the classroom teachers and sometimes to gauge the perspective of the students the population was not much to be generalized. Then in some of the researches the context and the techniques used for finding the importance and impact of rubric was not satisfactory. Interviewing teachers can give subjectivity to the research the better way is to conduct test with rubric and without rubric in order to find out difference in assessment. In Pakistani context there is a great scope of studies about rating the raters in language testing and it can be extended to broader level. In order to overcome the differences in evaluation with and without rubrics, there is a dire need for the development of rubrics for every class and also arranging workshops and training courses for raters.

References

- Andrade, H.G. (2000). Using Rubrics to Promote Thinking and Learning. *Educational Leadership*, 57(5), 13-18.
- Brualdi.A. (2002).Implementing Performance Assessment in the Classroom. In C. Boston (Ed.), *Understanding scoring rubrics: A guide for teachers* (pp.1-4). Maryland, U.S.A: Eric Clearinghouse on Assessment and Evaluation.
- Burns, P. C., Roe, B. D. & Smith, S. H. (2002).*Teaching Reading in Today's Elementary Schools, 8th Ed.* Boston: Houghton Mifflin Company.
- Catherine Hack, C. (2013). Using Rubric to Improve Marking Reliability and to Clarify Good Performance. Higher Education Academy.
- Cox, C. (2002).*Teaching language arts: A Students-and Response-Centered Classroom.* Boston: Allyn& Bacon.
- Crooks, T. J. (1988) The Impact of Classroom Evaluation Practices on Students, *Review of Educational Research*, 58, 438-81.
- Glickman-Bond, J. and Rose, K. (2006) *Creating and Using Rubrics in Today's Classrooms*, Norwood, MA: Christopher-Gordon Publishers, Inc.
- Goodrich, H. (1997). Understanding Rubrics. *Educational Leadership*, 54(4), 14-17.
- Huot, B. & Neal, M. (2006).Writing Assessment: A techno-history. In C. A. MacArthur, S. Graham, & J. Fitzgerald (Eds.), *Handbook of Writing Research* (pp. 417-432). New York, NY: Guilford Press.
- Ghalib, K. &Hattami, A. (2015).Holistic versus Analytical Evaluation of EFL Writing.*English language Teaching* 8(7) 122-130.
- Hafner, J. C. and Hafner, P. M. (2003) Quantitative Analysis of the Rubric as an Assessment Tool: An Empirical Study of Student Peer-Group Rating. *International Journal of Science Education* 25,(12),1509-28.
- Howell, R. J. (2011). Exploring the impact of Grading Rubrics on Academic Performance: Findings from a Quasi-Experimental, pre-post Evaluation. *Journal on Excellence in College Teaching*, 22 (2), 31-49.
- Julas, M.(2012). Impact of the use of Rubrics on the Performance of the Students. *Language Testing in Asia* 4(5) doi:10.1186/s40468-015-00121-7.
- Jeong, H.(2015). Rubrics in the Classroom: Do Teachers really follow them? *Language Testing in Asia*5(6)doi:10.1186/s40468-015-0013-5.

- Linn, R.L., & Gronlund, N.E. (2000). *Measurement and Assessment in Teaching*. Upper Saddle River, New Jersey: Prentice-Hall.
- Moskal, B. (2000). Scoring rubrics: What, when and how? *Practical Assessment, Research & Evaluation*, 7(3). Retrieved from <http://pareonline.net/getvn.asp?v=7&n=3>
- Manal, S. (2011) Teachers' and students's perspective on using rubrics in the writing class room. *Assessment and Evaluation in Higher Education*. 30(4)228-235
- Malini, R & Heidi Andrade, H. (2010). rubric use in higher education. *Assessment & Evaluation in Higher Education* 35 (4) 435-448.
- Palomba, C.A. & Banta, T.W., (1999). *Assessment Essentials: Planning, Implementing, and Improving Assessment in Higher Education*. San Francisco: Jossey-Bass, p. 4.
- Perlman, C. (2002). An Introduction to Performance Assessment Coring Rubrics. In C. Boston (Ed.), *Understanding scoring rubrics: A guide for teachers* (pp.5-12). Maryland, U.S.A: Eric Clearinghouse on Assessment and Evaluation.
- Price, D. (1965). Networks of Scientific Papers. *Science*, 149, 510-515.
- Qasim, A. Z. (2015). Using Rubrics to Assess Writing: Pros and Cons in Pakistani Teachers' Opinions. *Journal of Literature, Language and Linguistics* 16, 51-58.
- Quinlan, A. M. (2006) *A complete Guide to Rubrics', Assessments Made Easy for Teachers, K-College*. Lanham, MD: Rowman and Littlefield Education.
- Rezaei, A.R., & Lovorn, M. (2010). Reliability and Validity of Rubrics for Assessment through Writing. *Assessing Writing*, 15, 18-39.
- Sert, N. (2008). Constructivism in the Elementary School Curricula. *Journal of Theory and Practice in Education*, 4(2), 291-316.
- Turgut, F., & Kayaoğlu, M. N. (2015). Using Rubrics as an Instructional Tool in EFL Writing courses. *Journal of Language and Linguistic Studies*, 11(1), 47-58.
- Wiggins, G., & McTighe, J. (2005). *Understanding by Design*. Alexandria, VA: Association for Supervision and Curriculum.