Teacher's performance evaluation system: applications, challenges and research areas

Aranuwa Felix Ola* and Sellapan Palaniappan

School of Science and Engineering, Malaysia University of Science and Technology, Selangor, Malaysia.

ABSTRACT

Worldwide national policies on higher education are giving increasing importance to improve the quality of education on offer. Consequently, the evaluation of teachers' performance is especially relevant for the academic institutions as it helps to define efficient plans to guarantee quality of teachers and the teaching learning process. However, a key challenge is to understand the complex range of features associated with teacher evaluation, the contentiuos issues on accuracy and reliability of classical evaluation methods coupled with both mixed empirical evidence about their effects on students' learning and conflicts of interest between key actors of education systems. This paper is an overview of evaluation of teachers' performance in higher institutions, challenges in the progress, the main research areas and its applications to develop an effective and improved teachers' performance evaluation system.

Introduction

Performance evaluation is a systematic process of evaluating an individual worker's job performance and effectiveness in relation to certain pre-established criteria and organizational objectives [1],[28]. Following this definition, teachers' performance evaluation can be defined as a systematic process of evaluating teachers' performance and competence in relation to certain pre-established criteria, standards and school objectives. According to Anderson and Van Dyke, (2000)[3], the strength of good education in any educational institution depends on the quality of the academic staff in that institution; and there is no satisfactory substitute for competent staff that possesses sound educational philosophy and dynamic leadership style. William (2012)[49] posited that teacher quality is among the most important within-school factors affecting student's achievement. As the most significant resource in schools, teachers are vital to improve student outcomes and raise education standards. Hence, improving the efficiency and equity of schooling depends, in large measure, on ensuring that teachers are highly skilled, well resourced, and motivated to perform at their best. From this perspective, teachers' performance evaluation is a vital step in the drive to improve the effectiveness of learning systems and raise educational standards. According to Denisi and Pritchard (2006) [13], a central reason for the employment of performance evaluation is performance improvement (initially at the level of the individual workforce, and ultimately at the level of the organization). Other fundamental reasons include basis for employment decisions (e.g. promotions, career advancement, performance reward, sanctions, e.t.c). Additionally, performance evaluation can aid in the formulation of criteria and selection of individuals who are best suited to perform a required organizational tasks [28]. It can be part of guiding and monitoring employee career development and improvement.

By definition evaluation is the process of examining a subject and rating it based on its important features [25]. While evaluation in education system can be defined as the systematic determination of merit, worth, and significance of a learning process by using criteria against a set of standards [8],[14]. In the same vein, it can be defined as a systematic acquisition and assessment of information to provide useful feedback about some objects [46],[50]. The three definitions above agreed that evaluation is a systematic endeavour and the term 'object' or 'subject' here could be a program, policy, technology, person, need, or activity. However, the later definition emphasizes acquiring and assessing information because all evaluation work involves collecting and sifting through data, making judgments about the validity of the information and of inferences we derive from it. According to Nakpodia (2011)[34], evaluation is an intervention strategy that has received significant attention in academic, business and political circles for information gathering process, ascertaining the decision to be made, selecting related information, collecting and analyzing information in order to report summary data useful to decision makers in selecting among alternatives.

Teachers' performance evaluation system provides teachers with meaningful appraisals that encourage professional learning and growth. It provide useful feedback to a variety of audiences including, school leaders, administrators, staff, and other relevant constituencies for decision making. The process is designed to foster teachers' development and identify opportunities for necessary adjustment or additional support where required. Performance evaluation process represents one element of vision of achieving high levels of student performance [38],[43]. Maigul (1998)[26] noted that performance evaluation results in quality improvement, beside quality improvement and maintenance of standards, performance evaluation is one of keys to validating policies, plans and procedures operating within an organization.

With the national policies mandating high stakes evaluation of teachers and the learning systems at the forefront of popular school reforms agenda, in which all academic institutions are

increasingly required to monitor the performance of their learning and teaching system [10],[18],[38],[49]. There is no doubt educational institution both in developing and developed countries have an obligation to deliver value for investment to the bodies that fund them. Moreso that university performance are often judged by the quality and reputation of the awards they provide and product they produce [23]. This then gives rise to a need to collate, analyze and interpret data, for apt evaluation of teacher's performance in the higher institutions in order to have evidence to inform academic policies that are aimed at.

Literature Review

Evaluation of teachers' practice and performance is not a new trend, but what is new is the deep interest to enhance ways of teachers' evaluation in higher institutions. Key challenges is to understand the complex range of features associated with teachers' evaluation, the controversial issues on accuracy and reliability of classical evaluation methods, practical feasibility coupled with both mixed empirical evidence about their effects on students' learning and conflicts of interest between key actors of education systems [20],[21]. Researchers and school administrators at different levels have proposed and used wide-ranging approaches to assessing teachers' performance. However, the efficiency and dependability of the classical methods has been controversial. [4], [15],[29], [34],[40]. As a result there was no uniform standard method or computerized solution for evaluating teachers' performance that covers indicating factors affecting the quality of university and other tertiary education [10].

Early in the history of teachers' performance evaluation, educators were evaluated based on traits or characteristics which may or may not have been related to performance, and yet no significant body of knowledge confirms to the fact that effective teaching performance is dependent on specific traits. As a result, this form of evaluation was discarded [22],[48]. Formative and summative evaluation have been discussed, proposed and used by different researchers at different levels for measuring teaching effectiveness, quality of teaching and teachers' performance in higher institutions respectively [2],[6], [21],[26], [49]. Formative evaluation refers to a qualitative evaluation on the teacher current practice (teaching assessment), aimed at identifying strengths and weaknesses and providing adequate professional development opportunities for the areas in need of improvement. Formative evaluation involves the use of classroom observations, student evaluation report e.t.c., as tools to measure the performance and effectiveness of a teacher. The overall intention is to provide informative feedback to assist faculty in improving the effectiveness of their teaching [47]. Summative evaluation is an indispensable source of documentation and recognizable way to evaluate a teacher, providing summary statements of a teacher's capabilities through inspection, examination or interviews, in order to measure aptitude and knowledge, to ensure that required standards are met, or to promote level of performance for immediate recognition [21]. It is used to determine the worth and career advancement of a teacher, assess that teachers are adopting the actions and best practices which improve student outcomes. Teacher summative evaluation gives crucial information about the teacher being evaluated relatively to what is considered as standards. Hence, summative evaluation is an indispensable source of documentation to hold teachers accountable for their professionalism. In many institutions, it involves the use of annual performance evaluation report

(APER) and interview as instruments to measure teachers' performance and effectiveness [34]. Various research studies are conducted to support the validity, reliability and usefulness of these tools at different levels; however, there are considerable debates about their reliability and how teacher evaluation should be used to improve schools standard and standards of education [15]; [29],[34]. For instance, critics claimed that student rating systems are too subjective (biased and one-sided) [6], [31].

Traditionally faculties have been skeptical about this, the negative feelings often spring from the fear that student ratings may not be reliable and may be used or misused for summative decision making purposes and such misuse can breed distrust between faculty and administrators, and resentment on the part of instructors. Additionally, they often believe that students do not take evaluations seriously and ratings may encourage grade leniency [6],[41]. Okoro (1991)[37] observes that students sometimes fill in what they think the teacher would like rather than how they feel about him or what they assessed. He also reasons that some teachers may treat students very leniently and may spend a lot of time joking with them in order to obtain favourable ratings. However, researchers recommended that when such data are going to be used, the instrument must be subjected to rigorous validity tests and analysis. Further, student rating data should be used in combination with other criteria in order to provide a better assessment, since any single measure of the evaluation will only emphasize one important element at the expense of others [17].

In the work of Nakpodia (2011)[34], he reported methodological weakness in the current ways of evaluating the performance and progress of teachers with respect to the Annual Performance Evaluation Report (APER). He claimed that this form of evaluation does not actually take consideration of substantial evaluation of individual staff progress performances, but the items of information on the report are used to determine whether the staff satisfies two out of the four conditions on which the promotion of academic staff is said to be normally based. The two conditions are: evidence of scholarly research publications; and evidence of effective service to the institution. He added some of the items in the rating scale are ambiguous and cannot be easily evaluated while; others are extraneous and irrelevant. However, he concluded that students' rating alone cannot provide all the relevant information required to evaluate the lecturers. Hence, a supplementary instrument should be used to obtain information dealing with such aspects of professional development e.g academic advancement, research publications and participation in academic conferences, workshops and seminars. He suggests however, that if APER is going to be used, ambiguous sections should be eliminated from the instruments.

Another challenge in teachers' performance evaluation issue is that the work of a teacher in higher institutions involves considerably more than the pedagogical activities associated with student learning, but also extend to university community development. The question is how can this performance be measured. It is therefore appropriate that teacher evaluation models consider professional responsibilities though less directly related to the teaching itself. This recognizes the fact that the demands on schools and teachers are becoming more complex and teachers have their areas of responsibility broadened. Some examples are: working and planning in teams; projects between schools; management and shared leadership;

building community partnerships for learning; and participation in professional development [35].

From the facts above and according to Steele et al., (2010)[42], teacher evaluation systems should employ a diverse set of measures to capture the complex nature of the art and science of teaching and learning systems, which is inherently a multidimensional construct. Hence, the obvious need for improved teachers' evaluation systems that includes a spread of verifiable and comparable teachers' performance that distinguish teachers' quality and effectiveness.

Approaches to Teacher Evaluation

Teacher evaluation can be approached from three different but related angles [16]: (i) measurement of inputs, (ii) processes, and (iii) outputs. Inputs are what a teacher brings to his or her position, generally measured as teacher educational background, experience, pedagogical and content knowledge, certification and licensure. These measures are sometimes discussed in the literature as "teacher quality". Processes, on the other hand, refer to the interaction that occurs in a classroom between teachers and students. It also may include a teacher's professional activities within the larger school and community. Outputs can be referred to as "teacher effectiveness". It represents the results of classroom processes and career development, such as impact on student achievement, graduation rates, student behaviour, engagement, attitudes, quality and social-emotional well-being. Other outcomes may involve contributions to the school community in the form of taking on school leadership roles, educating other teachers, or strengthening relationships with other constituencies. Therfore, a meaningful teacher evaluation is expected to involve an accurate appraisal of the effectiveness of teaching, its skills and career development or advancement, followed by feedback, support and opportunities for improvement [36].

Teachers' Performance Evaluation Procedures, Metrics and Application

For effective and improved teachers' performance evaluation there is need for adequate procedures for a given objectives, such as the extent of the existence of national standards of education guidelines, or the culture of evaluation. Followed by development plan, internal regulations and activity plan at the institution level. According to OECD (2009)[36], proper evaluation of teachers' performance procedures requires the establishment of reference standards and criteria, choice of instruments and sources of information, which are discussed as follows:

Establishment of reference standards

A fair and reliable teachers' evaluation model needs reference standards to evaluate teachers relatively. The main reference standard for teachers' evaluation typically is the teachers' professional qualifications or profiles (level of education, experience, pedagogical and content knowledge, certification and licensure [16]. The key element and fundamental precondition of these must be clearly and concisely stated to know what are expected from teachers at different levels. Teachers' profiles often express levels of performance appropriate to beginning teachers, experienced teachers, and those with higher responsibilities. It is important to note that professional profiles provide the common basis to organize the key elements of the teaching profession such as initial teacher education, teacher certification, teachers' ongoing professional development and career advancement.

Establishment of criteria

Another essential basis for good practice in evaluation is the existence of clear and measurable criteria which must be consistently applied by competent (trained and experienced) evaluators [12],[45]. This requires the development of explicit guidelines about what is expected from professional practice. Teacher evaluation procedures require setting up evaluation criteria to determine the level of performance of individual teachers for each of the aspects assessed. This typically implies the development of indicators and/or standardized forms to record teachers' performance. Two major types of teachers' performance evaluation methods identified and widely used in literatures are: formative and summative evaluations [6],[21],[26], [30],[39],[44],[49]. An additional criterion is the weighting of the different aspects assessed in order to compute an overall quantitative rating. This permits evaluators to associate the quantitative rating to a qualitative scale, e.g. "unsatisfactory", "satisfactory", "average", "proficient", "basic", "above average", "distinguished" e.t.c.). According to OECD (2009)[36], some teachers' evaluation models establish quotas for the proportion of teachers who can be rated at the top of the scale (e.g. "distinguished"). Such approach runs counter to a criterion-referenced evaluation. However it might be justified in cases of an incipient culture and tradition of evaluation models to preclude a situation whereby most teachers end up being rated at the top of the scale, in which case the model loses its purpose. A number of teachers' evaluation systems in the United States have set a list of criteria based on Danielson's framework for Teaching [9],[19],[32]. The four US districts of Cincinnati, Washoe, Coventry and Vaughn adopted customized versions of the framework's competency model. So did the province of Quebec in Canada, Chile's four domains and twenty criteria of assessment were also largely inspired by the framework [5]. UNESCO's analysis of the European and Latin American teacher evaluation systems emphasizes the content knowledge, the pedagogical skills, the abilities to assess teachers and the professional responsibilities vis-à-vis the school, the students and their families as key domains to evaluate teachers. However, one should note that the analysis does not mention the engagement in professional development as a common teaching standard in European systems, with a subsequent risk to undervalue the teacher's engagement and willingness to enhance his own practice. Nevertheless, England has recently implemented a framework for professional standards, close to Danielson's one, which includes professional development criteria for the five levels of teaching performance (the award of Qualified Teacher Status, teachers on the main scale, Post Threshold Teachers, Excellent Teachers, and Advanced Skills Teachers) [45].

Instruments and Information Sources

Since the way of gathering evidence about a particular teacher may influence the evaluation results, the choice of instruments is of chief importance in designing and implementing systems to evaluate teacher performance. Gathering multiple sources of evidence about teacher meets the need for accuracy and fairness of the evaluation process, taking into account the complexity of what a competent teacher should know and be able to do. A range of instruments and information sources are typically used to evaluate teachers. Some of which are: classroom observation, student evaluation form, inspection and interview, student outcomes, questionnaires and survey, annual performance evaluation form to mention a few.

Validity and Considerations in Measuring Teachers' Performance

Determining what type of teacher evaluation method is best for a given purpose includes taking account of the validity and reliability of the instrument or process being used. According to Millett, Stickler, Payne, and Dwyer, (2007)[33], validity is the most fundamental consideration in assuring the quality of any evaluation. Validity refers to the degree to which an interpretation of a test score, or in this case, a score from a measure of teachers' performance evaluation, is supported by evidence. For a measure of teacher performance and effectiveness to be valid, evidence must support that the measure actually assesses the dimension of teachers' performance and effectiveness it claims to measure and not something else. In addition, evidence that the measure is valid for the purpose for which it will be used is essential. Instruments cannot be valid in and of themselves; an instrument or assessment must be validated for particular purposes [24]. For example, an observation-based score might be validated for professional development purposes but might not be validated for compensation purposes. Determining the validity of an instrument requires taking account of the evidence regarding what the instrument measures and how the scores are being used. This requires the user of the instrument to be well-informed about these issues and to only make judgments about the degree to which there is sufficient evidence to use a particular instrument for the purpose under consideration. In addition to concerns about validity, there are other measurement concerns. Blanton et al. (2003)[7] identified six criteria that are particularly useful in informing this discussion which are elaborated in [11]. They are:

Comprehensiveness: This refers to the degree to which a measure captures all of the various aspects of teachers' performance evaluation. Less comprehensive measures might only capture important elements in one aspect at the expense of others.

Generality: refers to how well an instrument captures the full range of contexts in the teaching profession. An instrument is said to have a high level of generality if it measure important contents across contexts.

Utility: refers to how useful scores from an instrument are for a specific purpose. For example, scores from an instrument that ignores teaching context may not be useful in identifying contexts that appear to support more effective teaching. The experience of other researchers or practitioners with an instrument makes it possible to better anticipate its potential uses and limitations.

Practicality: refers to the logistical issues associated with a measure. These include the developmental work required to adapt an existing model or measure for one's own purpose [7].

Reliability: refers to the degree to which an instrument measures something consistently. For example, it might be important to know whether scores on an instrument measuring teacher effectiveness vary by time of year, time of day, grade level, or subject matter.

Credibility: is a specific type of validity, face validity that is particularly important in measures of teacher performance and effectiveness. A measure is said to be credible if it is viewed as reasonable and appropriate by stakeholders from different groups (e.g., teachers, experts, and administrators). These aspects of measurement: validity, comprehensiveness, generality, utility, practicality, reliability, and credibility must be duly considered in any teacher evaluation system.

Conclussion

This paper presented an overview of evaluation of teachers' performance in higher institutions, challenges in the progress, the main research areas and its applications to develop an effective and improved teachers' performance evaluation system. Classification approaches as well various techniques, instruments and scenarios for teacher evaluation system are discussed. An improvement in the teachers' performance evaluation system can be achieved by combining the multiple sources of information and considering other fundamental principles.

References

[1] Abu-Doleh, J. and Weir, D. (2007). Dimensions of Performance Appraisal Systems in Jordanian Private and Public Organizations. International Journal of Human Resource Management, 18(1), 75-84.

[2] Apampa, K.M., Wills, G., and Argles, D (2010). User Security Issues in Summative E-Assessment Security. International Journal of Digital Society (IJDS), Vol., 1, Issue 2, June 2010 *School of Electronics and Computer Science, University of Southampton, UK.*

[3] Anderson, L.W and. Van Dyke L.A (2000). *Secondary School Administration*: Houghton Mifflin Company, Boston, U.S.A.

[4] Archibong, I. A and Nja, M.E. (2011). Towards Improved Teaching Effectiveness in Nigerian Public Universities: Instrument Design and Validation. Journal of the Higher Education Studies, Canada. Vol. 1, No. 2; December 2011 URL: http://dx.doi.org/10.5539/hes.v1n2p78.

[5] Avalos, B. and Assael, J. (2006) "Moving from resistance to agreement: The case of the Chilean teacher performance evaluation", *International Journal of Educational Research*, Vol. 45, No. 4-5, pp 254-266.

[6] Bain, K. (2004). What the Best College Teachers Do. Cambridge, MA: Harvard Press.

[7] Blanton,M.R. (2003). The Galaxy Luminosity Functionand Luminosity Density. The Astrophysical Journal, 592:819-838, Aug. 2003.The American Astronomical Society. All rights reserved. Printed in U.S.A.

[8] Bramley, P. and Newby, A. C. (1984). The Evaluation Of Training Part I: Clarifying The Concept. *Journal of European & Industrial Training*, 8,6, 10-16. Retrieved from: http://www.nwlink.com/~donclark/hrd/sat6.html. 17/6/13.

[9] Borman, G. and Kimball, S. (2005) "Teacher Quality and Educational Equality: Do Teachers with Higher Standards-Based Evaluation Ratings Close Student Achievement Gaps?", *The Elementary School Journal*, Vol. 106, No. 1, pp 3-20.

[10] Chaudhari, O.K., Khot, P.G., and Deshmukh, K.C (2012). Soft Computing Model for Academic Performance of Teachers Using Fuzzy Logic.British Journal of Applied Science & Technology 2(2): 213-226, 2012.

[11] Coggshall, J. G. (2007). Determining Teacher Effectiveness - Center on Innovation and Communication Framework for Measuring Teacher Quality and Effectiveness: Bringing Coherence to the Conversation. Washington, DC.

[12] Drake, J. M (1984). Improving Teacher Performance through Evaluation and Supervision. Paper presented at the annual meeting of the National Association of Secondary School Principals, February 1984. ED 250 782

[13] Denisi, A. and Pritchard, R. (2006, July). Performance Appraisal, Performance Management, and improving individual performance: A motivational framework. *Management and Organization Review, 2(2)*, 253-277.

[14] Foxon, M. (1989). Evaluation of training and development programs: A review of the literature. *Australian Journal of Educational Technology*, 5(2), 89-104. Retrieved from: http://www.nwlink.com/~donclark/hrd/sat6.html. 17/6/13.

[15] Glazerman, S., Goldhaber, D., Loeb, S., Raudenbush, S., Staiger D. and Whitehurst, G.J R. (2010). *Evaluating Teachers: The Important Role of Value-Added.* The Brookings Brown Center Task Group on Teacher Quality Pg 2. November 17, 2010.

[16] Goe, L, Bell, C, and Little O (2008). Approaches to Evaluating Teacher Effectiveness: A Research Synthesis June 2008. National Comprehensive Center for Teacher Quality, sponsored under government cooperative agreement number S283B050051.

[17] Guarini, C. and Stacy, B. (2012). Review of Gathering Feedback for Teachers. Boulder, CO: National Education Policy Center. Retrieved from http://nepc.colorado.edu/files/pb-options-1-teval_1.pdf

[18] Hénard, F., and Roseveare, D.(2012).Fostering Quality Teaching in Higher Education:Policies and Practices.An IMHE Guide for Higher Education Institutions.

[19] Heneman, H.; Milanowski, A.; Kimball, S. and Odden, A. (2006) "Standards-Based Teacher Evaluation as a Foundation for Knowledge- and Skill-Based Pay", Consortium for Policy Research in Education (CPRE) Policy Briefs RB-45.

[20] Holland, P. (2005) "The Case for Expanding Standards for Teacher Evaluation to Include an Instructional Supervision Perspective", *Journal of Personnel Evaluation in Education*, Vol. 18, No. 1, pp 67-77.

[21] Isoré, M. (2009), "Teacher Evaluation: Current Practices in OECD Countries and a Literature Review", *OECD Education Working Papers*, No. 23, OECD Publishing. doi: 10.1787/223283631428.

[22] Jernbaun, K (2012). Evaluation of College Instructors' Teaching Competencies Towards Proposed Human Resource Development Program. Retrieved from http://www.studymode.com/essays/Evaluation-Of-College-Instructors-Teaching-Competencies-1293750.html

[23] Jusoff K., S. A. A. Samah, and P. M. Isa, "Promoting university community's creative citizenry," in Proceeding of world academy of science, 2008, pp. 1-6.

[24] Kane, M.T. (2006). Current Concerns in Validity Theory. Journal of Educational Measurement Volume 38, Issue 4, pages 319–342.

[25] Keifer, K (2013): A Definition of Evaluation. Colorado State University. Copyright © 1993-2013. Available online at: http://writing.colostate.edu/about/copyright.cfm

[26] Kulik, J.A. (2001). Student ratings: validity, utility, and controversy. In: Theall M., Abrami, P.C. Mets, L.A. (Eds.), *New Directions for Institutional Research*. Jossey-Bass, 109, 9–24.

[27] Malgui, H.W. (1998). *The Role of Research and Evaluation in Survival of Education in Nigeria*. Journal of Quality in Education 3, 177-184.

[28] Manasa, K. & Reddy, N. (2009). Role of Training in Improving Performance. The IUP Journal of Soft Skills, 3, 72-80.

[29] Mardikyan, S., Badur, B., (2011). Analyzing Teaching Performance of Instructors Using Data Mining Techniques. Journal of Informatics in Education, 2011, Vol. 10, No. 2, 245–257 245.

[30] Marsh, H. W., & Roche, L. A. (1997). The use of students' evaluations and an individually structured intervention to enhance university teaching effectiveness. *American educational research journal*, 30(1), 217-251.

[31] McKeachie, W., (1997). Student Ratings; The validity of use. American Psychologist v52, No 11, 1218-1225.

[32] Milanowski, A. (2004) "The Relationship Between Teacher Performance Evaluation Scores and Student Achievement: Evidence From Cincinatti", *Peabody Journal of Education*, Vol. 79, No. 4, pp 33-53.

[33] Millett, C. M,. Stickler, L. M, Payne D. G and Dwyer, C. A. (2007.) A Culture of Evidence: Postsecondary Assessments and Learning Outcomes

[34] Nakpodia, E.D (2011). A Critique of the Methods of Evaluating the Competency of Lecturers in Nigerian Tertiary Institutions- African Journal of Education and Technology, Volume 1 Number 1, April 2011; pp. 53-59.

[35] OECD (2007). On the Edge: Securing a Sustainable Future for Higher Education (2007), OECD/IMHEHEFCE.

[36] OECD (2009). Teacher Evaluation: A Conceptual Framework and Examples of Country Practices. This paper was prepared for presentation at the OECD-Mexico Workshop towards a Teacher Evaluation Framework in Mexico: International Practices, Criteria and Mechanisms, held in Mexico City on 1-2 December 2009.

[37] Okoro, O.M. (1991) *Program Evaluation in Education.* Obosi Nigeria: Pacific Publishers, Uruowulu-Obosi, Anambra State, Nigeria.

[38] Ontario Ministry of Education (2010).Teacher Performance Appraisal Technical Requirements Manual . This publication is available on the Ministry of Education's website at http://www.edu.gov.on.ca/eng/teacher/pdfs/TPA_Manual_English_september2010l.pdf

[39] Ory, J. (2000). Teaching Evaluation: past, Present, and Future. New Directions for teaching and learning, No. 83, Fall 2000.

[40] Rothstein, R., Ladd, H. F., Ravitch, D., Baker, E. L., Barton, P. E. Darling-Hammond, L., Haertel, E., Linn, R. L., Shavelson, R. J., & Shepard, L. A. (2010). *Problems with the Use of Student Test Scores to Evaluate Teachers*. Economic Policy Institute. Retrieved from http://www.epi.org/publication/bp278/

[41] Sojka, J., Gupta, A.K., and Deeter-Schmetz, D.R. (2002). Student and Faculty perception of Student Evaluation of Teaching: A Study of Similarities and Differences. College Teaching 50(2): 44-49.

[42] Steele, J. L., Hamilton, L. S., and Stecher, B. M. (2010). Incorporating Student Performance Measures into Teacher Evaluation Systems. Palo Alto, CA: RAND. ISBN: 978-0-8330-5250-6. Retrieved from: http://www.rand.org/content/dam/rand/pubs/technical_reports/2010/RAND_TR917.

[43] Stronge, J. and Tucker, P. (2003) Handbook on Teacher Evaluation: Assessing and Improving Performance, Eye On Education Publications, 2003.

[44] Theall, M. and Franklin, J. (20010. Looking for bias in all the wrong places: A search of truth or a witch hunt in student ratings of instruction? New Directions for Instructional research 109: 45-56.

[45] Training and Development Agency (TDA) for Schools (2007b) "Professional Standards for Teachers: Why Sit Still in Your Career?", TDA, United Kingdom, 2007.

[46] Trochim,W.M.K (2006). Introduction to Evaluation. Web Centre for Social Research Method Retrieved 30/6/2013 from http://www.socialresearchmethods.net/kb/intreval.htm.

[47] UNESCO (2007). *Evaluación del Desempeño y Carrera Profesional Docente: Una panorámica de América y Europa*, Oficina Regional de Educación para américa Latina y el Caribe, UNESCO Santiago, 2007.

[48] Vollmer J. (1987). Our Schools Need to Change: http://www.studymode.com/essays/Evaluation-Of-College-Instructors-Teaching-Competencies-1293750.html, http://www.jamievollmer.com/about.html.

[49] William M. (2012). Research-Based Options for Education Policymaking-"Teacher Evaluation". National Education Policy Centre (NEPC). Available online at: http://nepc.colorado.edu/publication/options and http://nepc.colorado.edu

[50] Wikipedia, 2012: Retrieved on 14th June, 2013.