

Teacher Evaluation: Rationale, Issues and General Principles

Marshall Hughes¹

¹Alexander College, Larnaca, Cyprus

December 05, 2018

Abstract

This paper presents a review of recent literature in the field of teacher evaluation or appraisal. It considers the problems commonly associated with more traditional methods of teacher evaluation and synthesises these with the author's own professional experience as an international school Principal, to offer a six-point summary of the main criticisms. Based on the inverse of these criticisms, the paper then proposes six general principles for an improved, evidence-based approach to teacher evaluation.

Teacher Evaluation: Rationale, Issues and General Principles.

Dr.M.Hughes

Alexander College, Larnaca, Cyprus.

Abstract

This paper presents a review of recent literature in the field of teacher evaluation or appraisal. It considers the problems commonly associated with more traditional methods of teacher evaluation and synthesises these with the author's own professional experience as an international school Principal, to offer a six-point summary of the main criticisms. Based on the inverse of these criticisms, the paper then proposes six general principles for an improved, evidence-based approach to teacher evaluation.

Key words: teacher evaluation, appraisal, lesson observation, observation instrument.

"The monitoring of effective instruction is at the heart of effective instruction"

Lortie (1975)

The impact of good teaching on student achievement

For well over a decade the research has consistently indicated that an individual classroom teacher can have a positive, powerful effect on the learning of his or her students (Marazno and Toth 2013). Murphy (2013) concurs, stating that the large impact a good teacher can make on a pupil's academic outcomes is now well established. He cites a number of studies in support of this assertion. (Aaronson, Barrow, and Sander, 2007, Rivkin, Hanushek and Kain, 2005 and Rockoff 2004). The Organisation for Economic Cooperation and Development (OECD), (2005) report that raising teaching performance is the policy direction most likely to lead to substantial gains in student learning. Years of research on teacher quality, support the fact that effective teachers not only make students feel good about school and learning, but also that their work actually results in increased student achievement (Tucker and Stronge 2005). In a large-scale study, Wright et al (1997, p.63) concluded that "the most important factor affecting student learning is the teacher". The inverse is also true. The findings of the Dallas Public School Accountability System noted that the negative effect of poor teaching was discernible statistically for approximately three subsequent years. (cited in Tucker and Stronge 2005 p.5). Marshall (2013) cites another eleven studies which document the positive relationship between the quality of instruction and student achievement; the supporting evidence is overwhelming.

The need for monitoring and evaluation

It follows then that if high quality instruction has such a powerful effect, establishing and maintaining good teaching must be a priority for all schools. In short, they should implement some form of quality assurance. As Murphy (2013, p.3) points out “There is now widespread acceptance among researchers within the UK and internationally that good teaching is at the heart of good schools, and must therefore be at the heart of any school improvement programme”. The effective monitoring and evaluation of teaching is central to the continuous improvement of the effectiveness of teaching in a school (OECD 2009) . In this regard there can be few more apposite quotes than the one from Lortie (1975) at the top of this paper; it is over 40 years old and still resonates today.

So far, so good. We know that good teaching (whatever that is) leads to an improvement in student achievement (however that is measured) and that because of this, schools should monitor and evaluate the quality of their teaching (however that is done). But...there is an abundance of research which supports the notion that current systems of teacher evaluation are *not* working well. Weisber et al (2009) state that teacher evaluation systems have traditionally failed to provide accurate and credible information about the effectiveness of individual teacher’s instructional performance. A report for the Bill and Melinda Gates Foundation entitled ‘Gathering Feedback for Teaching’ reported on the failings by saying:

The nation’s collective failure to invest in high-quality professional feedback to teachers is inconsistent with decades of research reporting large disparities in student learning gains in different teachers’ classrooms (even within the same school). The quality of instruction matters. And our schools pay too little attention to it.

Kane and Staiger (2012, p3)

Existing systems rarely help teachers improve or clearly distinguish those who are succeeding from those who are struggling (Darling-Hammond, 2013). Toth and Rochman (2008, p 1) said that traditional evaluation practices are “superficial, capricious and often don’t even address the quality of instruction much less measure students’ learning”. This does not sound encouraging, so what is going wrong?

Problems with traditional methods of teacher evaluation

As long ago as the early 1980’s, teachers in America were critical of their evaluation systems and were calling for a more specific and rigorous approach. (Wise et al 1984, cited in Marzano 2013). Darling-Hammond (2013, p.6) believes that “most teachers want more from an evaluation system. They crave useful feedback and the challenge and counsel that would enable them to improve.” A group of expert teachers, the Accomplished California Teachers (2010) network, identified six main problems with teacher evaluation which they had encountered across a number of districts. Marshall (2013) devotes an entire chapter to a consideration of what he considers to be the seven main problems underlying the need for a new approach. Both lists are very similar and will be synthesised, together with other research and the author’s experiences as a teacher, vice-principal and principal, to produce the following 6-point summary of problems arising from traditional methods of teacher evaluation. Since “teacher evaluation traditionally has been based on the act of teaching and documented almost exclusively through the use of classroom observation” (Tucker and Stronge 2005, p.7) much of what follows is written in relation to this particular technique.

There is no shared and agreed definition of what constitutes ‘good teaching’ “Teachers and administrators rarely have a good understanding of what good instruction looks like” (Marshall 2013, p23). If this is true (and Marshall presents supporting statistical evidence) then it follows that the contents of any ‘observation instruments’ which are used to evaluate instruction, might generate similar differences of opinion. Darling-Hammond (2013) notes that even when instruments list elements of effective teaching, they often fail to elaborate on what constitutes evidence of these, thus hampering a fair and objective assessment of teaching. If observers are looking for one thing, and teachers think they should be doing something different, clearly any evaluation will be seriously compromised.

Lack of a clear, shared understanding as to the purpose of the evaluation.

A common-sense assumption about observation is that its purpose is to provide feedback and guidance for improving professional practice. However, "teachers rarely get feedback and teachers and principals have few authentic conversations about teaching and learning" (Marshall 2013, p.46) The ACT (2010) found that in a survey of 1010 teachers, only 26% reported that their most recent formal evaluation was useful and effective. 41% felt it was just a formality whilst 32% said it was, at best, well-intentioned but not particularly helpful. These are not encouraging statistics. Tucker and Stronge (2005, p.5) note that many authors add the additional purpose of "documenting accountability". Where this is the case, outcomes can be linked to contract renewal or tenure (Winters and Cohen 2012). New Federal requirements in the USA encourage the use of evaluations "to inform decisions about tenure and continuation, compensation and promotion, advanced certification and dismissal" (Darling-Hammond 2013, p.2) making teachers understandably nervous about the whole process. In the UK, the National Union of Teachers stated that: *The various purposes for which observations may be required has meant that in some schools the total number of observations has increased, and the context in which they are carried out can be felt to be hostile and bullying, rather than supportive and developmental.* (NUT 2010, p1)

1. *Current systems of evaluation are time consuming and unwieldy*Marzano (2013, p.13) notes that "live classroom observations are very time-consuming and expensive". Darling-Hammond (2013, p.4) concurs stating that "in many schools. . . . principals have little time or training for evaluation or support". She goes on to say that evaluations can go awry because schools create "unmanageable systems". Marshall (2013) argues that because each cycle of observation takes at least four hours of the principal's time, it is difficult to schedule and coordinate. Consequently it is very difficult for the principal to see each teacher more than once a year. (see point 6 below) He also notes that observation write-ups of 4 or 5 pages are the norm in many districts.
2. *Little or no connection to student outcomes/learning*Tucker and Stronge (2005) assert that most teacher evaluation focusses on the *act* of teaching rather than student outcomes. Most evaluations pay little or no attention to the performance of a teacher's students and hence provide little advice about how to support student learning. The ACT teachers noted that as long as a class is well managed and *seems* to be on task, not much else matters. (Darling-Hammond 2013) In other words, little or no learning might be taking place even though the teaching *appears* to be acceptable. This relates to point 1 above, for if the purpose of the evaluation is seen as assessing teacher performance, the real focus of student learning gets overlooked. Marshall (2013) notes that because principals are too busy scripting the lesson to get up and check on the learning, they are left with little choice but to focus on teaching inputs rather than learning outputs.
3. *Lack of appropriate training for the observer*In point 3 above, we have already drawn attention to the lack of training for administrators in the art of evaluation or support. An area to which policy often does not devote enough attention is that of skills development for evaluation (OECD 2009). Darling-Hammond (2013 p.1) states that "many principals have not had access to the professional development and support the need to become expert instructional leaders and evaluators of teaching". Principals are usually very experienced educators, and it is therefore assumed that they will make good evaluators. Without appropriate training this may not be the case and "well-intentioned programmes can revert to the blind leading the blind" (Coe et al.2014 p.25)
4. *The observer sees only a 'snapshot' of teaching which may be unrepresentative*

Most teachers plan and teach about 900 lessons per year and principals typically evaluate just one or 0.1%. The remaining 99.9% of the time each teacher is essentially alone with students (Marshall 2013). He goes on to say that "when teachers are left alone, mediocrity happens". (ibid. p.21) There are many problems with this. The infrequency of observation means that when it is announced, most teachers prepare and produce an entirely atypical 'all singing, all dancing' show lesson which bears little resemblance to normal practice. Even if the lesson is more usual, what happens if the teacher gets nervous, has a bad day or the children play-up? What happens if the computer crashes at the start of an IT- based lesson? Every teacher's worst fear is being judged on a bad moment or being taken out of context, so the desire to have more control over evaluation visits is quite understandable. But the downside is obvious, the principal may not be seeing

the kind of teaching students are experiencing on a daily basis, making the appraisal inaccurate. (Marshall 2013)

General principles for an improved approach to teacher evaluation

If we take the above negatives and reverse them, then logically we should end up with some general principles for a model of how effective supervision and evaluation *should* look. These would be:

Teachers and administrators should have a clear, shared understanding of what constitutes ‘good teaching’.

Little et al. (2009) offer a 5 point definition of teacher effectiveness. Coe et al. (2014) offer a list of six common components suggested by research that should be considered when assessing teacher quality. The DfE in the UK have recently issued a new set of ‘Teachers’ Standards’ which “define the minimum level of practice expected of trainees and teachers” and which are also “used to assess the performance of all teachers” (DfE 2011 p.3). For the purpose of this paper, the overall summary statement used by Coe (p.2) will be adopted. He states that:

“Great teaching is defined as that which leads to improved student progress”. This simple, clear, unequivocal statement establishes a sound basis from which to proceed and guides observers towards what they should be looking for (see Point 4).

1. *Teachers and administrators should have a shared understanding about the purpose of the evaluation.* Before we can think about the validity of any measures of teaching effectiveness, we need to be clear what those measures are intended to be used for. (Coe et al. 2014). The purposes of evaluation should be part of a vision that is clearly articulated and is part of the culture of the school. Obviously, new or struggling teachers will require a more customised programme of support where the purpose is somewhat different, but for the ‘average’ teacher, the purpose should be twofold. Firstly it should be about student learning outcomes and secondly it should assist teachers to reflect on and improve their professional practice. The latter of course, feeds into the former. Timely, informative feedback is vital to any improvement effort (Tucker 2005)
2. *Any evaluation system should not require excessive amounts of time or paperwork to complete.* Principals and teachers are busy people. It should be possible to design and implement a system which takes into account all six of these general principles and meets the needs of the school and its stakeholders, but which is also lean, efficient and useful. Darling-Hammond (2013 p.xi) states that “systems should be designed to be manageable and feasible, not so complex that they overwhelm the participants with requirements and paperwork”.
3. *Teacher evaluation should be linked to student outcomes/learning* In an effort to address the multi-dimensionality of teaching goals and responsibilities, teaching effectiveness is typically defined in terms of student learning. (Hobson and Talbot 2001). Wherever possible, it makes sense to judge the impact of teaching from its impact on assessed learning (Coe et al. 2014). If the *raison d’être* of schools is to raise student achievement, then logically student learning should be central to the process of teacher evaluation. Evaluation should focus on the *outcomes* of teaching rather than the *process* of teaching. If quality learning is taking place, then the teacher is doing their job and vice-versa.
4. *Appropriate training should be provided for anyone involved in the evaluation process.* The success of any teacher evaluation system greatly depends on the in-depth training of the evaluators. (OECD 2009). Murphy (2013 p.16) states that observers “should have good training so that they know what to look for, provide effective feedback and keep subjective opinions to a minimum”. Darling-Hammond (2013) calls for support structures to be put in place to ensure that evaluators are properly trained. Marshall (2013) states that to be effective, principals should be knowledgeable and perceptive observers. It seems axiomatic that if we want a job done well, then the person responsible for that job should receive appropriate training to successfully execute the work. Unfortunately, in education we assume that such competencies are a natural by-product of experience.
5. *Any evaluation system should be frequent and authentic enough to give an accurate, representative view of teaching and learning.* We have already established that a ‘once or twice a year’ classroom observation

of a ‘show’ lesson is not an appropriate vehicle for making sound, helpful judgements about the quality of teaching and learning. Nor is it an adequate sample on which to base meaningful feedback for teachers, which might improve practice. Indeed, it may well have the opposite effect. Evaluations should be frequent and unannounced so that typical teaching and learning is seen and a culture of openness and transparency is encouraged and embedded in the school. Evaluations should *not* be based solely on classroom observation. Darling-Hammond calls for methods which draw on multiple sources of evidence about student learning, arguing that these are essential to get a fair gauge on what a teacher has accomplished with his or her students. Marzano (2013 p.13) supports this saying “student growth should be measured in multiple ways”. Little (2009) offers a number of additional methods including value-added models, self-reflection, interviews, student evaluation and portfolios. Mindful of No.3 above, it is worth reiterating that the system must also be manageable and meaningful.

Finding a system or an ‘instrument’ which will meet with the above principles is not easy. Teaching is a “complex and uncertain endeavour” (Sato and Lensmire 2009) and even the detailed forms or instruments used by many districts cannot capture its complexity. Conversely, some are over simplistic and miss out important aspects of instruction. However, Coe et al.(2014) states that If the assessments available to us are not good enough, we need to improve them. As Murphy (2013 p.18) observes, “although it doesn’t matter greatly which particular rubric a school chooses to evaluate its teachers, it is very important that it has one”. A principal owes it to his or her students and teachers to find a system which works for their school.

References

- Aaronson, D., L. Barrow, and W. Sander (2007) “Teachers and Student Achievement in the Chicago Public High Schools.” *Journal of Labour Economics* 25 (1): 95–135
- Accomplished California Teachers (ACT) (2010) *A quality teacher in every classroom: Creating a teacher evaluation system that works for California*. Stanford CA. National Board Resource Centre, Stanford University.
- Coe, R., Aloisi, C., Higgins, S. and Major, L.E. (2014) What Makes Great Teaching? Review of the underpinning research. The Sutton Trust.
- Darling-Hammond, L. (2013) *Getting Teacher Evaluation Right*. Teachers College Press, NY.
- Department for Education (2011) Teachers’ Standards: guidance for school leaders, school staff and governing bodies. Reference No: DFE-00066-2011
- Kane, T.J. and Staiger, D.O. (2012) Gathering Feedback for Teaching: Combining High-Quality Observations with Student Surveys and Achievement Gains. Research Paper. MET Project for the Bill & Melinda Gates Foundation
- Hobson, S.M. and Talbot, D.M. (2001) Understanding student evaluations: what all faculty should know. *College Teaching* 49 (1): 26-31
- Little, O., Goe, L. and Bell, C. (2009) *A Practical Guide to Evaluating Teacher Effectiveness*. National Comprehensive Centre for Teacher Quality. Washington DC.
- Lortie, D.C (1975) *School teacher: A sociological study*. Chicago, University of Chicago Press
- Marshall, K. (2013) *Rethinking teacher supervision and evaluation* . Jossey-Bass, San Francisco USA.
- Marzano, R.J. and Toth, M.D. (2013) Teacher evaluation that makes a difference: a new model for teacher growth and student achievement. ASCD Alexandria, VA. USA
- Murphy, R. (2013) *Testing Teachers: what works best for teacher evaluation and appraisal*. The Sutton Trust
- NUT (1010) Classroom observation guidance. Position paper. National Union of Teachers.
- OECD (2005), *Teachers Matter: Attracting, Developing and Retaining Effective Teachers* , OECD, Paris.

- OECD (2009) *Teacher Evaluation: a conceptual framework and examples of country practices*. OECD Paris
- Rivkin, S.G., Hanushek, E.A., and Kain, J.F. (2005) Teachers, schools and academic achievement, *Econometrica*, 73(2): 415–458
- Rockoff, J.E (2004) “The Impact of Individual Teachers on Student Achievement: Evidence from Panel Data” *The American Economic Review*, Vol. 94, No. 2
- Sato, M., & Lensmire, T. (2009). *Poverty and Payne: Supporting teachers to work with children of poverty*. Phi Delta Kappan, 90(5), 365–370.
- Toch, T. and Rothman, R. (2008) *Rush to judgement: teacher evaluation in public education*. Washington DC: Education Sector
- Tucker, P., and Stronge, J.H. (2005) *Linking Teacher Evaluation and Student Learning*. ASCD, Alexandria, VA. USA
- Weisberg, D., Sexton, S., Mulhern, J. and Keeling, D. (2009) *The widget effect: our national failure to acknowledge and act on differences in teacher effectiveness*. Brooklyn NY, New Teacher Project
- Winters, M.A., and Cohen, J.M. (2012) *Grading New York. Accountability and Student Proficiency in America’s Largest School District*. Available at:
- <http://journals.sagepub.com/doi/abs/10.3102/0162373712440039> (accessed 11.02.18)
- Wright, S.P., Horn, S.P. and Sanders, W.L. (1997) Teacher and classroom context effects on student achievement: Implications for teacher evaluation. *Journal of Personal Evaluation in Education* 11: 57-67. Kluwer Academic Publishers, Boston.