Fall 2011

Designing an Effective Music Teacher Evaluation System (Part One)

Amy Clements-Cortés
Wilfrid Laurier University, notesbyamy2@yahoo.ca

Recommended Citation

Clements-Cortés, A. (2011). Designing an effective music teacher evaluation system: Part one. Canadian Music Educator, 53(1), 13-17.

Edited by Lee Willingham

Designing an Effective Music Teacher Evaluation System (Part One)

by Amy Clements-Cortès

Peer Review Corner features articles that have been submitted for review by a panel of music educators. The jury completes a "blind" review of manuscripts, offers suggestions for revision, and the revised article is either accepted or rejected based upon consultation with the journal editor and the others on the editorial board. If you wish to submit an article for review, please send it to Dr. Lee Willingham (lwillingham@wlu.ca).

Abstract: Danielson & McGreal (2000) state that an effective teacher evaluation system must contain three elements: a1) a coherent definition of the domain of teaching – "What," b2) techniques for assessing all aspects of teaching – "How," and c3) trained evaluators who can make consistent judgments about performance – "Who." In part one I have examined some of the issues associated with these three areas and in part two which will appear in the next journal I designed present a potential music teacher evaluation system that contains these elements for the school system in Ontario, Canada. On the whole, the proposed system is designed to help teachers feel that they have been fairly evaluated by professionals with relevant pedagogical knowledge, as well as from other parties who have an interest in their role as a teacher, and that they have had a voice in their evaluation. Careful consideration to levels of performance, weighting, and score combining were taken into account when designing the system alongside the evaluation being informed by multiple data sources.

Introduction

Music education is a publically funded piece of Canadian education, and like other tax supported activities, the accountability of teaching is a topic of interest to a variety of stakeholders including: parents, school administrators, teachers, politicians and students. With such a wide variety of people interested in the results of the evaluation for their own interests and concerns, designing an effective system that satisfies all parties is challenging. Diverse audiences have unique questions and concerns they want addressed. It is particularly complicated to develop an adequate and fair evaluation system for music educators.

There are a number of tensions inherent in the evaluation of music teachers that question the underlying assumption that teacher evaluation is a positive exercise, and it may not be seen as constructive or beneficial by some parties including educators. For example, teacher evaluation can become politicized. The establishment of the "Office for Standards in Education" in the U.K. reportedly has resulted in the undermining of staff morale in such a large scale that few have confidence in their own professional judgement anymore and few can be persuaded to enter the profession. Similar issues arise in Canadian education when teachers are held accountable for standardized test results.

As a discipline, music makes high claims to provide a unique, creative learning environment where a variety of academic, social, health and self-esteem benefits may be gained. For example, there are many claims on the Coalition for Music Education in Canada website (http://musicmakesus.ca/ educate/) including highlights from local newspaper articles reporting studies such as: "Music lessons get kids into college," and "Music is good for the health" both posted on April 15, 2011.

Music celebrates its ability to motivate students through different modes of learning. Music is different. Yet, a tension arises when advocates for music education argue that music should be core, or mainstream. When that goal is achieved, music teaching and learning find themselves under the scrutiny of accountability processes. Teachers are evaluated by criteria designed for other modes and disciplines.

Recognizing this assumption several questions surface: "Why would teachers, especially music teachers, impose evaluation upon themselves? What is to be gained from this experience?" Preparing for an evaluation requires considerable effort on the part of several parties, most predominately the teacher. That being said, well-designed evaluation systems may contribute to a teacher's overall continued development as an educator, role model and musician; therefore resulting in a reflexive practitioner striving to implement best teaching practices. Self assessment along with a well-designed systematic process of evaluating teaching practice in music education must, in fact, result in improved teaching and learning.

What Comprises a Teacher Evaluation System?

A teacher evaluation system is "a complete approach to the evaluation of teachers including its purpose, the rules and regulations that apply, the target group to be evaluated, the domains to be covered, the procedures and methods to be employed, the instruments to be used, the persons to be involved, and the types of reports and feedback to be provided" (Teacher evaluation kit: complete glossary, 2004). Haefele (1993) states that ultimately a teacher evaluation system should: provide constructive feedback to educators; recognize and reinforce outstanding service; provide direction for staff development practices; and unify teachers and administrators in their collective efforts to educate students.

Weisberg, Sexton, Mulhern & Keeling (2009) explain that in the United States a report by the *New Teacher Project* focusing on thousands of teachers and administrators spanning four states concluded that the current teacher evaluation systems used result in all teachers receiving the same top ratings. Gabriel (2010) quoted U.S. Secretary of Education Arne Duncan: "Today in our country, 99 percent of our teachers are above average" (p. 24). These types of evaluations are obviously not useful. If everyone receives the same top rating teachers who are truly performing at the top may feel negated by the loose rating system, and those who are not engaged in best teaching practices have no motivation to improve.

Effective teacher evaluation, according to Danielson & McGreal (2000), must contain three elements – the what, how, and who which includes: a coherent definition of the domain of teaching; techniques for assessing all aspects of teaching; and, trained evaluators who can make consistent judgements about performance. It is crucial that these definitions remain in the general domain and avoid specificity or uniformity. A coherent definition of "the domain of teaching" for music must honour the unique aspects of the learning that the teaching motivates. Assessment techniques must

the contributions of extra-curricular instruction such as band and choir are dismissed by most evaluation systems

be the right tools for this type of learning, and perhaps most importantly, the evaluators must be qualified to make assessments, evaluate, and report within the context of the music teaching domain. In music teacher evaluation consistent definitions may result in uniform teaching, mitigating against the very claims the music makes as being student centred and creative.

Danielson (2007) reiterates these fundamental principles to ensuring high-quality teaching in a teacher evaluation system. He expands on the three elements above by: discussing the importance of the definition of teaching to be grounded in research, giving teachers the opportunity to provide evidence on all the different criteria of how they are meeting their requirements (and that criteria is made known to teachers in advance), and ensuring the evaluators are sufficiently trained to guarantee consistent judgement. Issues that are associated with each of these areas will be presented, alongside a potential teacher evaluation system for music educators that contain these elements.

Issues

What

One of the main problems with many teacher evaluation systems today is that they were developed in the 1970's, and generally reflect what people believed about teaching at that time. As Danielson (2007) explains "In many schools, evaluation is something done to teachers, with the teachers themselves playing an essentially passive role" (p. 181). Further he states that learning is done by the learner, and if teachers are to develop and grow from their evaluations they need to play active roles. As attitudes evolve, and new approaches to teaching emerge, an evaluation system must also change alongside, to reflect current outlooks. Nolin, Rowand, Farris & Carpenter's (1994) survey, while somewhat outdated, found that 99 percent of elementary teachers acknowledge that subject matter should be considered when evaluating performance, while only 65 percent stated that it was considered to a large extent in their evaluations. Teachers want to have a voice in their evaluations, a voice that is not only heard by the evaluator but formally included in the written performance document and record. Even though the Ministry of Education in Ontario revised the evaluation system in 2010, it remains fairly similar to the earlier models (Ministry of Education, 2010).

Another issue involves reaching consensus on the definition of teaching. What one person believes to be "good teaching" may not be considered good teaching to another. Therefore, there could be a lack of agreement on what should be taught at various stages, and the strategies used to teach. It may be challenging to reach a point where the majority of people agree with what is to be taught at various stages, but the definition should be based on the research area and current literature. That being said, it is important to note that students are individuals who develop and acquire knowledge and skills at their own pace and not in a uniform manner. This is where a skilled teacher can adapt the curriculum to meet the needs of individual students. Within a school district, interested parties such as teachers, parents, and students (at a certain level) should be consulted about elements that must be part of the definition of teaching. This will address some of the sociological issues such as power and expectations. While it is essential that teachers' thoughts be taken into account as experts in the field, parents and students also have important contributions to make towards defining teaching. Some may disagree with this statement; however students and parents are at the core of teaching in the school system. We must understand from their perspectives what constitutes the art and science of "teaching.". Perhaps the balance lies in the amount of weight each party is given in terms of their influence over the definition of teaching.

In Ontario, the teacher performance appraisal documents assess 16 teaching competencies in five domains: commitment to pupils and pupil learning; professional knowledge; teaching practice; leadership and community, and ongoing professional learning (Ministry of Education, 2010). These domains are clearly identified, and are based on the literature. For example, Danielson & McGreal (2000) suggest four domains, which although they are named differently, encompass the same responsibilities and themes as those outlined in the Ontario standards. New teachers are appraised twice in the first twelve months of teaching and experienced teachers once every 5 years as the minimum guideline.

In designing an evaluation system three items should be kept in mind: levels of performance; weighting; and score combining. Levels of performance are needed to describe different levels of functioning. The issues inherent in these items are: providing descriptions to clearly define points on a rating scale, deciding if all of the evaluative criteria will be considered equal and given equal weight in the evaluation, and deciding if all of the scores will be combined. These decisions have implications. For example, a teacher could have high scores on several areas of the evaluation, but do very poorly in one area. If all the scores are combined into one total score, the area that needs continued development might be missed or overlooked.

How

Issues that develop in the "how" element of teacher evaluations include time, expense, choice of data sources to include, and the determination of whose role it is to co-ordinate the data sources. The amount of time and budget available for teacher evaluation will directly impact the type of evaluations that can be done, and will essentially provide the framework. In today's current financial climate, while desirable for music educators and others in specialized fields to have an evaluation completed by an individual external to the school and a specialist in their area, it is not realistic. In Ontario, the Ministry acknowledges that teacher appraisal is completed by the vice-principal, principal, or supervisory officer as these individuals have authority to do so under current union agreements and legislation. As evaluation continues to evolve and become more useful, it is hoped that funds will be provided to hire external professionals to conduct part of teacher evaluations especially in specialized fields like the arts.

When planning a system, time must be allotted for data gathering, decision making and training (Peterson, 1995). Typical teacher evaluations involve the principal or department head visiting the teacher in the classroom one to two times per year and writing a narrative about the episode. Sometimes this is expanded into what is known as the clinical supervision model, which was utilized in Ontario beginning in the 1970's (Magarrey, 2002). This method involves three steps: a pre-observation meeting, the observation, and a post-observation meeting. Currently the

steps outlined by the Ministry of Ontario (2010) involve the three steps listed above and conclude with a summative report that includes a rating of the teacher's overall performance. Although this is an improvement over the teacher simply receiving a written narrative about his/her teaching, it still primarily relies on direct observation as the means of evaluating performance; and direct observation alone cannot accurately assess all that encompasses teaching.

Specifically related to the arts, music researchers acknowledge that evaluation devices used for teacher evaluation nationwide do not work well with performing arts instruction (Grant & Drafall, 1991; Taebel, 1990). For example, the contributions of extra-curricular instruction such as band and choir are dismissed by most evaluation systems. Maranzano (2000) feels that traditional evaluation approaches do not supply evaluators with enough comprehensive information to make important educational decisions about music teacher performance. This is ultimately one of the largest challenges for music educators. Music educators argue that music should be mainstream but that it is not useful or beneficial to be evaluated by criteria designed for other subjects and disciplines.

Multiple data sources

The literature suggests that one of the most effective ways of obtaining the most complete picture of a teacher is to use multiple data sources (Peterson, 2000; Danielson & McGreal, 2000). Gathering data from numerous sources is obviously more time consuming and expensive than conducting an observation. Ostrander (1995) concluded that the most equitable and comprehensive performance appraisals potentially involve multiple judges, each offering a distinctive perspective on teacher effectiveness. Multiple judges could involve such persons as parents, other teachers, and students. Implementing this suggestion and using multiple judges will require additional planning and co-ordination of the material. For example, if the evaluation utilizes student surveys as a form of input and as one of the "judges," it must be determined in the planning stage if the students' comments will be of equal weight to the other data, such as the evaluator's observations. Collective agreements and current legislation prevent schools and principals from soliciting data from multiple data sources such as student surveys.

One of the obvious challenges in the utilization of multiple data sources is determining which sources to use. According to Peterson (2000) each method used should be subjected to tests of logic, practicality, reliability, validity, cost and user acceptance, and should have demonstrated success, for example through documentation in journals. There are several sources of data that should be avoided, including such items as testimonials, graduate follow-up, and peer consensus, as these do not meet the criteria established by Peterson. An acceptable combination of data sources could include any of the following elements: classroom observation, student surveys, peer review of materials, student achievement data, parent reports, teacher portfolios and teacher tests.

Classroom observation

Classroom observation has been one of the primary sources of obtaining data to complete teacher evaluations. Peterson (2000) recommends that classroom observations be 30-50 minutes in length, and that there be 4-5 observations over a 60-day period and that these visits should be unannounced. Unannounced visits while beneficial will not bode well with current collective agreements and therefore are not realistic in today's schools in Canada. Overall, Peterson's suggested approach appears appropriate, especially for new teachers (those in their first two years of teaching). The frequency of the observations he suggests is not likely to be widely implemented due to budget constraints, and may not be warranted for more experienced teachers.

There are some inherent problems with the classroom observation method. Initially one must choose to utilize either an open or a closed system. An open system includes a freely written script or recording of the events as they occur; whereas in a closed system, specific behaviours are identified in advance as codeable. Peterson (2000) outlines several styles of classroom observation, and acknowledges that each style raises questions. In "gazing about," the observer sits with students while the teacher is teaching a lesson. Additionally, he/she may walk around and ask students questions to see if they understand the lesson. Afterwards the observer writes a report about what was seen, and may or may not discuss these observations with the teacher. The main problem here is that there is no factual data written until after the observation, and that means that data may then be changed, manipulated, or simply not reflected accurately. As well, the evaluator's presence is more prominent in this model, than in ones where the evaluator observes silently, and this may influence the type of data being collected.

In "participant observation" the observer records at certain intervals and at other times is engaged with the students participating in the lesson. The central drawback is that it is difficult to participate and record at the same time, and so data are limited. Being a participant as well as an observer may end up changing some of the observations as the observer is also contributing to them.

The "diary description" observation method involves keeping a running log of all that is seen. The problem here is that it is difficult to record everything going on for the whole lesson, and in the end, observers must be selective about what they include.

"Time and event sampling" involves observing for a period of time such as thirty seconds, and then recording for another time period such as three minutes. This is not very useful for conducting teacher appraisals, but is much better suited to research studies aimed at observing some type of pattern.

"Problem point lists" involve the evaluator looking around the classroom and documenting areas for improvement. This is problematic because by definition, problems or generally negative comments are recorded, and there may or may not be a reflective statement written about the observation.

Ultimately, "systematic observation" appears to be the most appropriate way to conduct classroom observation, and it presents the least amount of concerns. Peterson & Kauchak (1982) outline five components of systematic observation: the observer is a neutral outsider to the school system; the observer is trained in observation techniques; observations are taken from a reliable number and timing of visits; the focus of the observation is limited to a few categories of events; and, the recording systems are systematic, have reliability in practice and analyze data within an established framework. As a neutral outsider may not be avail-

Principals, who are often the persons performing the evaluations, obviously cannot be experts in all subject areas, and this can undermine the evaluation process

able due to budgetary constraints, many teacher evaluations continue to be done by the principal, and therefore it may be difficult to achieve all five of the points outlined by Peterson & Kauchak.

One challenge with classroom observation with respect to music is that non-music specialists who are conducting the evaluation might not understand the lesson being taught and the steps required for skill acquisition in music. It might be beneficial for the teacher being observed to sit with the evaluator and comment on the various parts of the lesson the evaluator assessed in order to help him/her complete a more accurate evaluation through increased understanding of the lesson observed and the goals behind the activities.

Student surveys

Student surveys could take place in the form of a written survey or even a focus group. A focus group might be more time consuming and may make it more difficult to place information received into "neat and tidy" categories but would most likely produce the most useful information. When designing a student survey, it is useful to include "opportunity to learn" questions, for example, "I learn new things in this class." Questions that should be avoided are those which ask for judgements that the student is not really qualified to give, such as if the teacher knows his/her subject area, or if every student in the class is called upon equally. Also, questions that blur the line of responsibility for learning such as "the teacher makes me want to do my best" do not provide appropriate information (Tucker, Stronge & Gareis, 2002; Peterson, 2000). In Ontario, teachers might consider conducting a survey as a source of evidence when they are documenting in their log of teaching practice (Appendices F & G, Ministry of Education, 2010).

Peer review of materials

McCarthy & Peterson (1987) carried out a study regarding peer review of materials in a school district, and concluded that this data source holds considerable promise for teacher evaluation. Typically in this method, judgements are made by two or three peers who examine documents prepared by the teacher being reviewed. Documents could include such items as lesson plans, student recitals and performances, and examples of student work. One obvious problem here is that the teacher may select only work from his/her top students, and assemble documents that are highly favourable to his/her evaluation. To avoid any biased reviews, reviewers should not know the teacher they are reviewing, and to maximize the effectiveness of this method, the peers should currently be teaching students at the same level. Peer review of materials is particularly attractive to music educators as they can have their documents assessed by other specialists in the field, and perhaps feel that these assessments hold more significance given these peers' skills and expertise. The peer-evaluation role given to teachers is empowering; however, because teachers are unionized employees, union representatives may oppose these types of reviews if they have not been addressed in the collective agreement. Additionally, teachers may already feel overburdened with their workload and therefore not feel they have time to conduct such reviews fairly; or, they may have limited desire to take part. One way to implement a peer review process to limit the impact on workload would be to rotate this responsibility so that one year teachers are part of a team that reviews, and the next year they prepare their material for review by others.

Student achievement data

In using student achievement data in evaluation, it is often difficult to find valid measurements, and the logical connection between student learning and teacher performances are indirect and have mixed causality. Results obtained from standardized tests of students, which some persons feel reflect a teacher's ability to teach, have their own issues. For example, in preparing students to do well on an up-coming standardized test, some teachers may direct their teaching to emphasize material specific to the test. The central problem with using results of standardized tests for teacher evaluations is that they do not account for the knowledge that the students did or did not have before they were under the leadership of their current teacher. For music educators in Ontario and Canada, standardized tests do not evaluate music education and therefore have no basis in teachers' performance evaluation. Glazerman, Loeb, Goldhaber, Staiger, Raudenbush & Whitehurst (2010) maintain that reports by education researchers are calling for caution in the use of teacher evaluations being based on student test scores. Their argument is that at the individual level value-added scores often misclassify teachers in an unfair way.

The most accurate way to gather valid student achievement data is to assess the gains, increases or changes in students under the influence of the teacher. This means that gains must be adjusted to students' prior learning. Standardized tests are not able to accomplish this goal, and there is a danger of misinterpreting the results if evaluators feel that they do. Student achievement data should not focus only on post-instruction tests as an indicator, as there is a large influence of previous achievement on the part of the student. Obtaining accurate measures in this area is further complicated by a student's background, which can greatly influence levels of student performance. Also, class size and the rate at which a student learns new information affect this type of data. As children are unique they do not necessarily learn new material at the same rate as each other, and some students may pass through certain cognitive stages faster than others, thus enhancing their ability to learn new material faster. For music educators it is particularly tricky. Some students take private lessons outside of class and therefore are more advanced than other students in the class. The teacher not only has the role of teaching the class but often of mentoring these students with advanced skills to ensure they are learning new material and are equally challenged as those students with less background. This creates extra work for the music teacher, and this work is rarely understood or acknowledged by evaluators.

An example of a system that provides accurate measurements is the Tennessee Value-Added Assessment System (TVAAS) (as cited in Stronge & Tucker, 2000). Essentially this is a system that tests each student in each grade through a number of subjects. Value-added analysis takes the gains each student makes from year to year and compares the gains made by a normative sample for the same subject for those same grades. In essence, it is an achievement test and linear regression analysis that measures academic gains made by individual children, adjusted for each student's multi-year performance history. The system uses each child's previous scores as a blocking factor to control for variables such as socio-economic status, resulting in a standardized measure in relation to the expected gain. A limitation of this system is that because of its multi-year approach, it is expensive to implement.

Parent reports

Stronge & Ostrander (1997) recommended the use of parental surveys in teacher evaluation in order to bridge deficiencies in existing processes. Epstein (1985) concluded that parental contributions are valuable in teacher evaluation, as parents have an important perspective on performance. Again in Ontario and Canada it is not realistic to

have parental surveys included in teacher evaluations due to current legislation and collective agreements.

Teacher portfolios

Portfolios give teachers a voice in their evaluation. The unfortunate drawback to this method is that assembling a portfolio is time consuming and takes considerable effort on the part of the teacher. For music educators who often provide numerous extra-curricular activities for students on their own time, this is just another piece of work that gets added to their plates. Those negatives aside, there are many benefits of using teacher portfolios. Tucker, Stronge, & Gareis (2002) outline that portfolios can help teachers monitor their own performance, assist evaluators to acquire a more accurate picture of the teacher's performance, and benefit students with improved instructional techniques.

Deciding what items to include in a portfolio can be difficult. Painter (2001) explains that portfolios should include any evidence used by teachers to document or support how they meet teaching standards. Furthermore, judging the content of portfolios is equally difficult, and will require initial training for evaluators on how to assess them in a standardized manner. Furthermore, their use for summative evaluation can cause the items included to be distorted. Distortion could potentially occur when a teacher selects material that reflects positively on them, but does not necessarily reflect the daily work he/she does, such as including a well-designed lesson plan, when typically the teacher does not prepare ahead of time. Teachers should be informed and evaluators made aware of the fact that while portfolios help expand the evaluation of the teacher, they can by no means represent everything that a teacher does, and if the teacher tries to prepare a portfolio that does they will be highly frustrated with the process.

In Ontario teachers have the option of preparing a log of teaching practice in an attempt to document information and noteworthy examples of their performance in the domains that they are being evaluated upon. One way to make the assembly of such a portfolio less cumbersome is for teachers to add to it once every one to two months. For example, adding commentary from an external adjudicator at a music festival about a teacher's band or choral performance. By slowly adding to a portfolio when the teacher's evaluation is getting closer he/she has material to review and may select the best pieces for the portfolio.

Teacher tests

In Ontario the push from politicians and parents to make teachers accountable has fuelled the idea of using teacher tests to determine competency (Ontario Teachers Federation, 2005). Peterson (1995) recognizes that while they are advocated for, teacher tests are rarely included in evaluation practices. This could be due to the difficulty in designing such tests. For example, content measured on the test may not be matched to the knowledge required by the teacher in the classroom.

The idea behind using tests is valid. For example, if the test could determine whether or not teachers are up-to-date on their current pedagogical knowledge, then it would make sense to use this method. However, this is often not the case. Many questions could potentially be raised such as "How many tests would there have to be to determine pedagogical knowledge accurately?" For some of the same reasons that standardized tests of students should not be used to gauge teacher effectiveness, so too teacher tests have not been proven to determine who will be an effective teacher and who will not (National Centre for Fair and Open Testing, 2005).

Who

Who conducts the evaluation is an important factor in making the evaluation credible. Principals, who are often the persons performing the evaluations, obviously cannot be experts in all subject areas, and this can undermine the evaluation process. For example, a teacher's knowledge of content and content-related pedagogy is highly relevant to teaching, and teachers may be more knowledgeable about these matters than the administrators who are expected to evaluate their performance. Therefore, teachers may feel that the evaluator is really not qualified to assess them, and this may in fact be true. In the end, the evaluation will do nothing in terms of contributing to the overall growth of the teacher. With respect to music educators Maranzano (2000) acknowledges that the absence of music specialists performing the evaluation lowers the overall reliability of the evaluation process.

From a sociological perspective there appears to be a conflict in having the administrator/principal, who is seen as a leader, also hold the role of sole evaluator. Many principals may naturally inflate the evaluations, feeling that the results for teachers under their leadership are a reflection of their role as supervisors.

Another problem centres on the lack of consistent judgements by evaluators. Even when criteria are clearly defined, and systems with rating scales are employed which are typically more objective, there can still be issues. For example, when one evaluator might give a score of 3 out of 5, another might give a rating of 4. Therefore, evaluators should be trained in conducting evaluations, and it should be understood that there will always be some degree of subjectivity in teacher evaluations.

The tricky area for music educator evaluations appears to fall in the area of professional knowledge and professional practice. How can a principal or vice-principal accurately assess these skills? For example, how can he/she evaluate whether a teacher's professional knowledge informs appropriate music pedagogy, or evaluation of student learning while being evidence-based. It would seem that in order to accurately appraise these skills one must have some degree of specialization in the area of music and music education, or how can the evaluator accurately determine that a teacher knows his or her subject matter and uses effective strategies to teach that subject? With all this being said at present it will remain that principals will be the ones conducting the evaluations and therefore it is important for teachers to be proactive in a sense in order to facilitate accurate and helpful evaluations of their work.

Part two, which will appear in the next issue of the journal, will present the proposed teacher evaluation system.

References

Coalition for Music Education in Canada. (2011). (http://musicmakesus.ca/educate/)

Danielson, C. (2007). Enhancing professional practice: A framework for teaching (2nd ed.). ASCD, Alexandra, VA.

Danielson, C. & McGreal, T. (2000). Teacher evaluation to enhance professional practice. Alexandra, VA: Association for supervision and curriculum development.

Epstein, J. L. (1985). A question of merit: Principals' and parents' evaluations of teachers. Educational Researcher, 14(7), 3-8.

Gabriel, T. (2010, September 2). A celebratory road trip by education secretary. New York Times, pp. A24.

Glazerman, S. Loeb, S., Goldhaber, D., Staiger, D., Raudenbush, S. & Whitehurst, G. (2010). Evaluating teachers: The important role of value-added. Brown Centre on Education Policy at Brookings, Washington: DC: The Brookings Institution.

Grant, J. & Drafall, L. (1991). Teacher effectiveness research: A review and comparison. Bulletin of the Council for Research in Music Education, 108, 31-48.

Haefele, D. L. (1993). Evaluating teachers: A call for change. Journal of Personnel Evaluation in Education, 7, 21-31.

Magarrey, M. (2002). Research in Ontario secondary schools: Changing perceptions of the nature of teacher evaluation in Ontario secondary schools. Retrieved from: http://www.oise.utoronto.ca/field-centres/TVC/RossReports/vol7no3.htm.

Maranzano, C. (2000). Music teacher performance evaluation: A call for more inclusive models. Journal of Personnel Evaluation in Music Education, 14(3), 267-274.

McCarthey, S. & Peterson, K. D. (1987). Measurement of teacher merit for salary purposes. Journal of Personnel Evaluations in Education, 1, 259-267.

Ministry of Education: Ministry of training, colleges and universities (2002). Teacher performance appraisal manual. Retrieved from: http://www.edu.gov.on.ca/eng/teacher/appraisefs.html

Ministry of Education: Ministry of training, colleges and universities (2010). Teacher performance appraisal manual. Retrieved from: http://www.edu.gov.on.ca/eng/teacher/pdfs/TPA_Manual_English_september2010l.pdf

National Centre for Fair & Open Testing. Teacher Tests. (2005). From: http://fairtest.org/empl/tt.htm

Nolin, M. J., Rowand, C., Farris, E., & Carpenter, J. (1994). Public elementary teachers' views on teacher performance evaluations. U.S. Department of Education. Office of Educational Research and Improvement. NCES 94-097.

Ontario Teachers Federation (2005). The limitations of teacher testing. Retrieved From: http://www.ctf-fce.ca/en/action/2000/testing.htm

Ostrander, L. P. (1995). Multiple judges of teacher effectiveness: Comparing teacher self-assessments with the perspectives of principles, students and parents. (Doctoral Dissertation, University of Virginia, 1995, UMI No. AAT 9525110).

Painter, B. (2001). Using teacher portfolios. Educational Leadership, 58(5), 31-34.

Peterson, K. D. (2000). Teacher evaluation: A comprehensive guide to new directions and practices, (2nd ed.). Thousand Oaks, CA: Corwin.

Peterson, K. D. (1995). Teacher evaluation: A comprehensive guide to new directions and practices. Thousand Oaks, CA: Corwin.

Peterson, K. D. & Kauchak, D. (1982). Teacher evaluation: Perspectives, practices and promises. Salt Lake City: University of Utah.

Stronge, J. H. & Tucker, P. D. (2000). Teacher evaluation and student achievement. Washington, DC: National education association.

Stronge, J. H. & Ostrander, L. (1997). Client surveys in teacher evaluation. In J. H. Stronge (Ed.), Evaluating teaching: A guide to current thinking and best practice (pp. 129-161). Thousand Oaks, CA: Corwin.

Teacher evaluation kit: Complete glossary (2004). Retrieved from Western Michigan University, Evaluation Centre's Web site: http://www.wmich.edu/evalctr /ess/glossary/glossary.htm

Taebel, D. (1990). The evaluation of music teachers and teaching. In Colwell, R. J. (Ed.), Handbook of research on music teaching and learning (pp. 310-330). New York: Schirmer Books.

Tucker, P., Stronge, J., H., Gareis, C. R. (2002). Handbook on teacher portfolios for evaluation and professional development. Larchmont, NY: Eye on Education.

Weisberg, D., Sexton, S., Mulhern, J. & Keeling, D. (2009). The widget effect: Our national failure to acknowledge and act on differences in teacher effectiveness. New York, NY: The New Teacher Project.

Dr. Amy Clements-Cortés is assistant professor in music therapy at the University of Windsor. She has extensive clinical experience working with older adults in long-term and palliative care, adult mental health, complex continuing care, developmental delay and children with varied diagnoses. She holds a Masters and PhD from the University of Toronto, and an undergraduate degree in music therapy from the University of Windsor. Amy teaches voice through her company Notes By Amy and sits on the Board of Directors for the CAMT where she holds the position of internship chair. Her research has been focused in long term care, geriatrics, cognitive impairment, stress reduction, sleeping problems, pain, and palliative care. Her latest book, Episodes of Relationship Completion Through Song: Case Studies of Music Therapy Research in Palliative Care, is available on www.Amazon.com. The accompanying CD may be purchased on her own website www.notesbyamy.com.

a "toonful" perspective

Nadine Pickthall has taught instrumental, vocal and recorder music and homeroom subjects to intermediate, junior and primary classes in the York Region District School Board since 1993. She is a graduate of the University of Toronto and has a Specialist in music and a Specialist in Art. She plays the French horn and piano. Nadine has a strong interest in the arts. She has taken art courses through the University of Toronto, the "Portfolio for Animators" program at Sheridan College and cartooning at Seneca College. Nadine was a University of Toronto editorial cartoonist for two years and has also had work published in magazines and newspapers.