



**TECHNIUM**  
**SOCIAL SCIENCES JOURNAL**

**Vol. 28, 2022**

**A new decade  
for social changes**

[www.techniumscience.com](http://www.techniumscience.com)

ISSN 2668-7798



9 772668 779000

## **The Impact of Using a Multidimensional Teacher Evaluation System on Increasing Teacher Professional Growth**

**Chakerian Taline, Dr. Samra Sami (Advisor)**

Doctorate student in science of education, Department of Human and Social Sciences, University of Saint Joseph, Beirut, Lebanon

[taline.chakerian@gmail.com](mailto:taline.chakerian@gmail.com)

**Abstract.** The aim of this article is to inspect the perceptions of administrators and teachers regarding the multiple data sources used to evaluate teachers, and to find out which source maximises teacher's professional growth as perceived by administrators and teachers provided by the Armenian High schools in Beirut in Lebanon. Both quantitative and qualitative data were gathered from seven Armenian High schools in Lebanon. The quantitative data was gathered through Teacher Evaluation Profile (TEP), where N=10 administrators and N=50 teachers participated in filling out the questionnaire while the qualitative data was gathered through interview questions, where N=7 administrators and N=14 teachers participated in the interviews. The results of the study convey the importance of using multidimensional evaluation to be able to gather a wholistic view about teacher's performance, as well as to maximise the teacher's professional growth. However, there are also other factors and procedures that are needed to be followed.

**Keywords.** Multidimensional evaluation system; Teacher evaluation; Professional Growth, Teacher Perception; Administrator Perception

### **I. Introduction**

As the history of the teacher evaluation system has shown, teacher evaluation processes started with a single observer, usually the principal, who entered classrooms with a rating checklist. O'Day (2002) explained that the problem with this type of evaluation was that it was a checklist-oriented process and did not include clear definitions or expectations to make the possible improvements that led to professional growth. Most of the items on this checklist were not related to instructional practices, but to other factors like the neatness of the classroom and bulletin boards (Donaldson, 2009). Noakes (2009) also explained that these types of evaluation, especially when short, did not have the capacity to reflect the teacher's true ability in educating.

Brandt, Mathers, Oliva, Brown-Sims, and Hess (2007) added that the limited time given to classroom evaluations naturally led to limited feedback for teachers. In addition to the concerns related to time and feedback, Brandt et al. (2007) explained that these traditional evaluation systems had limited professional standards. Thus, most of the teachers ended up being satisfactory according to these ratings. However, teachers perceived that they, as teachers, felt they were less effective than the ratings presented in the evaluations. As a result, these

evaluations did not even serve the purpose of keeping teachers accountable (Donaldson's, 2009).

A descriptive study was carried out by Brandt et al. (2007) about teacher evaluation policies at local, district, and state levels in the Midwest region: Arkansas, Illinois, Indiana, Iowa, Kansas, Michigan, Minnesota, Missouri, Nebraska, North Dakota, Ohio, South Dakota, and Wisconsin. This was done in an attempt to study both the policies in the schools found in that region and the processes of the teacher evaluation according to said policies. Hence, 140 participants from 216 school districts were included. The study found that the policies varied both in their practices and in their policy prescriptions. What was found was that the purpose of these evaluations was mainly to decide whether to dismiss or retain teachers. Therefore, new teachers were evaluated two or three times per year, while more experienced teachers were evaluated once every two or three years. The studies concluded that most evaluations were summative reports to aid in the decision-making process. The policies and procedures were not clearly described and they included vague terminology that could lead to different interpretations. These policies did not include clear consequences in case the teacher evaluation was unsatisfactory, and did not require any training for the evaluator. Overall, they found that these policies were more focused on the process of the evaluation, such as the people responsible for the evaluation, the time that should be given, and the number of observations. These were considered instead of focusing on the content and standards.

Hence, according to the research explained above, traditional evaluation systems depended on rating checklists that were not fully focused on instructional practices. Limited time was given to classroom observation, and therefore limited feedback was received by teachers. Without performance standards, evaluators could not provide feedback for improvement. Finally, most of the teachers ended up rated satisfactory. Principals could neither hold teachers accountable, nor could they help their professional growth.

Research has extensively proven that these methods of teacher evaluation do not lead to neither teacher growth nor student achievement, and consequently not even to school improvement (Noakes, 2009; Schochet & Chiang, 2010; Viadero, 2009; Danielson and McGreal, 2000). Since a well-designed teacher evaluation system leads to school improvement (Normore, 2005), thoughtful planning in this area becomes both a concern and a must.

An effective evaluation system, on the other hand, should aim for effective teaching through accountability and professional growth (Danielson, 1996). It should also have more than one person responsible for evaluation, and should use multiple data sources for accuracy and bias-free results (Elstad, Lejonberg, & Christophersen, 2015; Peterson, 2006). Using additional data sources would be helpful for both teachers and evaluators. Evaluators benefit because their responsibility and pressure of doing all the work alone decreases, and they are strengthened by additional information that helps their assessment. Additional data sources also help teachers in receiving various types of feedback, from different viewpoints. These provide less biased data, as explained by Peterson (2006). This is called a multi-dimensional evaluation system (Aramath, 2014), which different researchers have described in different ways. Darling-Hammond (2012) for example described it as a system that uses multiple measures, while Peterson (2006) described it as evaluating teachers using different kinds of data collecting methods. Therefore, a multi-dimensional teacher evaluation system is a well-designed evaluation system that assesses individual teachers using different methods to collect various types of data. These different methods include: classroom observation by the principal, teacher self-assessment, structured reflection, planned documents, teaching artifacts and other evidence such as parental communication logs, professional development activities, and samples of

student work, in addition to feedbacks received from parents, peers, and students (Danielson & McGreal, 2000). Each of these will be discussed separately including advantages, shortcomings, and research studies associated with each.

### **1.1. Different methods for evaluating teachers include**

*Observations:* This is a method widely used by school administrators. It usually involves principals or supervisors dropping into classrooms with a checklist to observe teachers and take notes, and then complete an observation form. For the observation to be beneficial, Webb and Norton (2003) explained that teachers should clearly know the purpose of the observation and the standards to be met.

Observations are the most commonly used form of teacher evaluation. They are the most important and accurate way of knowing what is happening in the classroom with regards to teacher-student interaction, student engagement, and the overall learning environment (Danielson & McGreal, 2000). However, Danielson and McGreal (2000) added, there are some aspects of instruction that observation alone cannot capture. For example, to know whether the teacher provides effective feedback to students, one can only find out by looking at what the teacher has commented on a student's work, or by listening to the teacher while they are meeting with a student.

Observations greatly differ in their purpose and form. Observations can be purchased, if they are created by researchers, or can be developed by the evaluators themselves according to their views of teaching. The evaluator can be the principal, an administrator, or an outside evaluator. The purpose of the evaluation can also differ, from measuring general teaching practices to subject-specific practices (Piburn & Sawada, 2000). They can follow a formal process, where the evaluator announces the name of the teacher, the schedules, and the date of the observation. This is called a Formal Observation. It can also be an informal observation, unannounced, where the observer can drop in to observe without prior notice (Marzano, 2013). It can occur once per year, or several times throughout the year. The length can also vary, such as attending the whole period to make note of all the teaching practices taking place in the classroom. It can also just be a short visit to develop a quick impression about what is happening in the classroom. It can serve as a formative assessment to provide specific feedback about teachers' strengths and the areas that need growth and improvement, or it can simply be a summative evaluation, used for accountability purposes. This can help in differentiating between effective and ineffective teachers, who need dismissal, or even those who need remediation (Geo, Belle, & Little, 2008).

In short, observations may vary in their forms and purposes. For the observation to provide useful and significant information about teachers' effectiveness, two important things should be considered: using valid and research – based instruments and having well trained observers, who can use the instruments in the best way possible.

Brandt, et al., (2007) found that classroom observations were mainly conducted by the principals. Most of them were conducted for accountability and the decisions of employment, rather than for helping the teacher grow and improve in their profession. Moreover, only eight percent of the evaluators were trained. They concluded that although the purposes of these evaluations were clear; however, there was little evidence about the validity and reliability of their use.

Harris and Sass (2007) raised an important question about the accuracy of the principals' judgments, in the absence of training. Other researchers also noted another problem with

observations, which was the subjectivity of the principal's ratings, even when using researcher-created instruments.

Yon, Burnap, and Kohut, (2002) also noted that through observations, principals or administrators could only rate the general behaviors of teachers. However, when assessing specific content knowledge, peer teacher observations were a better choice. That is because peers teaching the same subject have specialized knowledge and can make better judgments. Therefore, using both a principal's observation and peer-observations may increase the quality, providing more credible information.

While Marshall, Smart, & Alston (2016) stressed the importance of classroom observation in the evaluation process, Skedsmo and Huber (2018) argued that observation should not be the only tool used to measure teacher effectiveness, since teachers' effectiveness varies across teaching different subjects and different levels, in addition for the observation to be an effective tool, it should be conducted systematically and over time.

Thus, if the observation is the only method used in a school, then, to have a productive evaluation result, principals should consider clarifying the purpose of the observation by including both formative and summative purposes. The instrument of observation should also have clear goals, criteria, and standards. All of these should be explained to teachers beforehand, while feedback should be provided shortly after the observation. In addition, an opportunity should be provided to teachers to discuss the evaluation, after pinpointing the areas that need improvement, and offering help and support.

*Checklists:* Most of the teacher observations or other forms of teacher evaluations are done through checklists. In fact, teacher evaluation was first introduced in 1900 in a form of a checklist to survey teacher efficiency through rating scales (Callahan, 1962). At that time, checklists were the only means available for teacher evaluation. At this stage, the purpose of teacher evaluation was directed toward improving the school system than student achievement (Spooren & Mortelmans, 2006). However, the effectiveness of an instrument is not in the form but rather in the items included. Hinchey (2010) suggested that an evaluation rating scale should include teacher practices, aspects of instruction, and teacher-student interactions.

Checklists are an easy way of gathering evaluation data from a number of sources, including administrators, peer teachers, students, and parents (Peersman, 2014). Usually, checklists are rating scales that contain a list of items that explain teacher performance. However, these rating scales, explained Hinchey (2010), should include not only teacher practices, but also aspects of instruction and the interaction between teachers and students.

While checklists differ from one school to another according to the specific needs of the school and the purpose for which the checklist is to be used, Noakes (2009) presented the Patton's Utilization-Focused Evaluation (UFE) checklist, which includes criteria such as teacher-school readiness assessment, evaluators' readiness, identification of the focus of the evaluation and its design, as well as data collection and analysis. Noakes noted that this checklist supported both teacher growth and student learning. It also provided more opportunity for interaction between the evaluator and the teacher.

*Value-Added Model:* This method determines the teachers who are effective and successful at improving student learning (Little, Geo, & Bell, 2009). It mainly considers student achievement scores on a standardized test in a specific subject, to predict their achievement for the following year. However, at the end of the year, if most of the students in the specific classroom accomplish the standardized test with better results, the teacher becomes credited for being effective and successful at improving student learning. On the other hand, when most of

the students perform worse than predicted, the teacher is considered less effective and responsible for this gap.

Many researchers studied the use of the value-added model to measure teacher effectiveness, teacher qualifications, their personal characteristics, and its relationship to other measurement tools such as observations.

Hanushek and Rivkin (2010) studied the relationship between value-added scores and teacher characteristics such as education, gender, and experience. As such, they noticed that the observable teacher characteristics considered in their study could not explain teacher effectiveness. Harris (2012) on the other hand, found that tracking teachers' value-added measures had a positive impact on student learning.

Another study done by Schacter, Thum, and Zifkin (2006) compared value-added achievement scores with teacher ability to foster student creativity in the classrooms. They found that teachers, who encouraged student creativity through the use of specific strategies and activities, also improved the student's achievement. They concluded that the value-added model is capable of improving instruction; however, it should be combined with other measures such as high-quality observational data.

In consequence, relying on the value-added model alone, as a primary mean to measure teacher effectiveness makes teachers exclusively responsible for student achievement. This would oppose to considering other factors such as schools, educators, families, peers, and the students' themselves. Corcoran (2010) added that the perception that value-added system would provide clear data about the whole picture is not a research supported fact. However, when used with other measuring instruments, this model might have the potential to improve instruction. In addition, simply receiving regular feedback about student achievement is a positive step towards improving student achievement (Corcoran, 2010).

*360-Degree Evaluation System:* Danielson and McGreal (2000) defined it as the system which included the perceptions of others, in addition to teachers' direct evaluation. Its aim was to see the skills that teachers' incorporate from several points of view. They further explained that this was important because "students, parents, and colleagues can provide evidence of a teacher's performance that a principal might never witness" (p. 51). However, they concluded that careful attention must be paid in designing these evaluation systems so that the information that is being gathered is not "compromised by personality differences, favoritism, or other irrelevant matters" (p. 51).

This explains the importance of gathering feedback from students, parents and peers. Thus, what we may conclude is that using this method should be carefully implemented, considering the results of one source might be independent from the other, since each source measures a different viewpoint. For example, while an observation from the principal measures the quality of the instrument, parents may have little input in this area. This is because they are not present in the classroom and may offer little feedback about the teaching method used. Likewise, principals may understand little about the relationship between the teacher and the students by simply observing the classroom. In contrast, students might provide valuable information about the relationship built between them and the teacher. The research study done by Peterson in 1987 on using multiple sources, like principal survey reports, student feedback, parent surveys, peer, and teacher portfolios, showed that there was no connection between data sources. Thus, Peterson (1987) concluded that these results were due to measuring different viewpoints.

*Evaluations by Students:* This is a survey-type instrument that usually includes both open and close-ended questions. Close-ended questions usually use consistent scales (e.g.

caring or motivated) and specific teacher behavior (e.g. listening and providing feedback). It can also be completed in less than 20 minutes (Lowa State University, 2011). Danielson and McGreal (2000) explained that these surveys must be age-appropriate and carefully developed in order not to ask questions related to the teacher. Rather, they should be about what is happening in the classroom, including statements such as whether students know the rules of the classroom, are treated fairly and/or know what they are expected to do. Student evaluations have been mostly used in universities; however, it has recently become widespread in schools as well. Many researchers agree that student evaluation is somehow valuable. Ripley (2012) assures that students have the ability to indicate their most and least effective teachers, if they are asked the right questions. Webb and Norton (2003), also supports the idea that students feedback on teachers evaluation can be more valuable than supervisors evaluation. This might be simply because students spend more time with teachers and have more authentic experience with their practices in the classroom.

Spooren and Mortelmans (2006) conducted a study to determine the relationship between students' grades and teachers' evaluation scores. The study concluded that these evaluations provided valuable information about teachers. To find out if there is a relationship between students' high grades and teachers' high evaluation scores, Centra (2005) found little correlation, and concluded that giving higher grades to students will not improve teachers' evaluation scores knowing that there is another argument that stated the opposite: students lower their rates of teachers when they receive low grades.

Papanastasiou (1999), on the other hand, explained that student feedback about teachers do not lead to teacher improvement. Almutairi and Shraid (2021) also found that students may over-evaluate their teachers' performance in order to provide a better picture of their teacher.

Webb and Norton (2003) explained that students evaluation can provide valuable data if developed and conducted properly. It can help administrators eliminate the amount of time used in performing teacher evaluations.

*Feedback from Peers:* Peer evaluation is a form of evaluation where teachers set a specific date and time or the observation of one to the other. They go to class, take notes and meet together to evaluate instruction and teaching activities. This observation may also be followed with an improvement plan (Glickman, Gordon, and Ross-Gordan, 2005). The results of a research study done by Hamilton (2013) on peer-to-peer observations at Midwest, a suburban public high school located in United States, showed that peers observing each other led to increased collegial respect. Clark (2017) did a research study about informal peer observation, in a public school for 3 months, where teachers had to complete informal peer observations, fill out a documentation forms and answer questions. The research results were so rewarding that the researcher recommended all the school leaders and educators to be involved in this process as part of their professional growth plan. Yet, Almutairi and Shraid (2021) found that peer evaluations should be used only for formative purposes and should never be included in the summative evaluation, since they explained subjectivism and competition may negatively affect some teachers' evaluation scores.

*Feedback from Parents:* Parents are another important source of information. Their feedback is vital and should be included as another perspective in teacher evaluation. Nevertheless, the questions posed to parents should be carefully planned, explained Danielson and McGreal (2000). They must avoid asking details about teachers' educational methods, since they have no idea how the teacher is explaining the lesson. However, their input might be beneficial when asked about whether or not their kids are able to complete the homework needed independently, or whether the teacher is available when needed, in addition to whether

or not they can communicate with the teacher. This survey provides valuable information. Schools cannot fully know whether their teachers are effective communicators with parents except through parental feedback. They may not know whether students are going home knowing what is expected of them, except through these surveys.

*Portfolios:* Teacher portfolios are a collection of teacher lesson plans and artifacts that give evidence about what the teacher is doing in the classroom. They authentically reflect teacher performance, and provide the opportunity for interaction between the evaluator and teacher (Barton, 2010). However, they are time consuming for the teacher and do not reflect everyday classroom activities (Hinchey, 2010).

Marcoux, Brown, Irby, and Lara-Alecio (2003) had conducted research to see whether using portfolios as a method for teacher evaluation can lead to an increase in student achievement and teacher professional development. The results were positive and found that portfolio evaluations had some impact on student achievement and professional development of teachers.

Palazuelos and Conley (2008) studied the perceptions of teachers about portfolio evaluations and found that teachers felt motivated when working on preparing the activities and lessons they are presenting and implementing in their classrooms. They felt portfolios provided a more holistic presentation of their work than a brief visit to the classroom for an observation did. Oppositely, Geo and Croft (2009) also added that portfolios are not useful methods for evaluation, and do not lead to teacher improvement nor professional development.

Further research done by Heneman and Milanowski (2003) on implementing Danielson's standard based evaluation system, showed that teachers perceived the evaluation process to be effective. Yet, they mentioned preparing portfolios as one of the disadvantages, since it was time-consuming. All of the following might be included as part of the professional portfolio:

*a. Artifacts:* These are the assignments, worksheets and/or particular directions for projects or any other work prepared by the teacher and completed by the student (Danielson and McGreal, 2000). Teaching artifacts show the evaluator how the documented plans are to be implemented to reach the learners in the classroom. It shows students the challenge they are to be asked to engage in (Danielson and McGreal, 2000). However, they explain that these artifacts should be combined with teachers' observation, to find out whether these assignments fulfill the needs of that particular group of students. Many times, the assignments are great in the way they are prepared. However, they do not match the abilities of the group of students within the class, and at times, planned assignments need to be adapted to effectively reach students' needs (Peterson, 2006). These can be revealed through observations.

*b. Teacher Self-Assessment:* It requires teachers to reflect on their practices and assess themselves through a checklist that contains observable behaviors or practices. Through this checklist they might assess the extent to which they are using those practices or implementing the standards required (Geo, Belle & Little, 2008). Danielson and McGreal (2000) noted that teacher self-assessment might be an extremely accurate measure, since teachers know best about their educational skills, and such assessment can be a source of dialogue between the teacher and the evaluator concerning the quality of teaching.

*c. Structured Reflection:* This requires the teachers to think about and reflect on the activities they have done with students, or the lessons they have planned and implemented. How well things have gone and what changes are needed in the future for more effective practice might be some of the questions teachers need to think and reflect upon. Danielson and McGreal (2000) explain that if schools are planning to have evaluation systems and lead to

professional growth, systematic teacher reflection should by no means be part of the process. They further explain that the reflection itself is a highly productive tool for professional learning.

*d. Planning Documents:* Conducting successful instruction needs thoughtful planning. This is the reason why Danielson (2007) included an entire domain in her framework of teaching. It includes demonstrating knowledge of content and pedagogy, demonstrating knowledge of students, setting instructional outcomes, and demonstrating knowledge of resources.

Many other researchers have included items related to planning as part of their evaluation system (Noaks, 2009; Hinchey, 2010; Iowa state university, 2011; Marzano, 2013). This implies that while evaluating teachers, looking at how they have prepared their lessons should be a vital part of the evaluation.

Thus, using multidimensional evaluation processes while evaluating teachers helps both evaluators and teachers gather feedback from different perspectives. This may be more satisfactory and less subjective when trying to reflect on the effectiveness of the work being done.

### ***1.2 The Purpose of the Study***

The purpose of this study is to examine the views of administrators and teachers regarding the different sources used to evaluate teachers and to find out which source maximizes teacher's professional growth and improves their practice as perceived by administrators and teachers. A mixed method was used to examine perceptions in three areas: a) Present administrators and teachers perceptions regarding the data sources used to evaluate teachers and gather information about their performance, B) Evaluate which data source maximizes and helps the teachers to improve their practice, C) Analyze the extent of observation used in the teacher evaluation process, D) Make suggestions for improvement about the usage of the sources to maximize teachers' growth. This is important, especially for the administrators to recognize the strengths and weaknesses of their school in using different sources and ways to improve the evaluation processes in their school and recognize which approach to follow. In order to increase its impact on teacher professional growth.

### ***1.3 Research Questions***

Based on the theoretical framework and the reviewed literature, the research questions of this study were as follows:

- (1) How do administrators and teachers perceive the different sources used to gather data about teachers at school?
- (2) How do administrators and teachers perceive the extent of classroom observations used in their schools?
- (3) How much do these sources affect teachers' practice as perceived by administrators and teachers?
- (4) How can the usage of these data sources be improved to maximize teachers' growth?

## **II. Method**

### **2.1 Research Design**

This study used a mixed-method approach, including both quantitative and qualitative methods to gather data. Quantitative research was conducted for the purpose of collecting and analyzing numerical data in order to be able to describe, explain, predict, or control phenomena

to be studied (Gay, Airasian, & Mills, 2012). Qualitative research helps in controlling, analyzing, and interpreting different views to gain insight into a particular phenomenon to be studied (Gay, Airasian, & Mills, 2012).

Quantitative study was done to gather numerical data using the Teacher Evaluation Profile (TEP) questionnaire (Stiggins & Duke, 1988), qualitative data was used to confirm findings and support the quantitative data by interviewing both teachers and administrators, using a set of questions to explore their perceptions regarding the different sources used during the evaluation processes. The aim of using mixed-method research was to combine both quantitative and qualitative data in a single study. Its purpose was to combine and build on the strengths of both methods to understand a phenomenon more fully than is possible when using only one method (quantitative or qualitative) alone (Gay, Airasian, & Mills, 2012).

## **2.2 Sample Population**

This research included all the Armenian high schools in the Greater Beirut area that are comprised of Elementary (grades 1 to 6), and Secondary (grades 7-12) schools. Seven Armenian high schools participated in this research study.

Within these schools, all the administrators involved directly in evaluating teachers, N=10 and all the teachers who taught Math and Language Arts among the two levels: Elementary and Secondary, N=50 were invited to fill the survey and participate in this study.

Concerning the interview, one administrator from each school N=7, and two teachers N=14 were randomly selected to participate in the interviews. Each interview session lasted approximately 30 minutes.

## **2.3 Instrumentation**

Since this study involved mixed methods, two instruments were used to gather data. Quantitative data was gathered by using the modified version of the Teacher Evaluation Profile (TEP) questionnaire. Qualitative data was used through one-on-one online interviews through Zoom meetings with both teachers and administrators.

To provide context for the quantitative study, it is worthy to note that the TEP was first developed by Stiggins and Duke (1988) at the Northwest Regional Educational Laboratory. The modified version which was revised by other researchers (Colby, 2001; Doherty, 2009; Hughes, 2006; Rindler, 1994) used the Likert Response Scale from 1 to 5, with 1 being the least favorable and 5 being the most favorable.

Due to the Covid-19 pandemic, interviews were conducted through online Zoom meetings. Both administrators and teachers were asked to discuss the role of the administrators in the teacher evaluation processes, in using different sources to evaluate teachers, and in providing post evaluation feedback. They were also invited to make suggestions for improving the evaluation process to increase its impact on teacher professional growth.

The questions that will be used for the administrators include the following:

1. How much time is allocated for classroom observations?
2. What aspects do you look for in a formal observation?
3. What tools, rubrics format do you use when observing a formal observation?

The questions that were used for teachers were as follows:

1. How much time does the evaluator devote to observe your performance?
2. What sources does the evaluator use when evaluating you?
3. What are the most effective aspects of the evaluation processes?
4. What are the most ineffective aspects of the evaluation processes?

## 2.4 Data Collection

To collect data, permission to use the modified version of the Teacher Evaluation Profile (TEP) was gained, 50 teachers and 10 administrators were asked to fill in the Teacher Evaluation profile. In addition, interviews were conducted with 14 teachers and 7 administrators through Zoom meetings. However, findings of the quantitative data were not known while the interviews were taking place.

## III. Presentation and Discussion of the Results

Regarding the different sources of performance information considered as part of the evaluation, these sources included the observation of classroom performance, meetings with the evaluator, examination of artifacts: lesson plans, materials, home/school communications, examination of student performance, student evaluation, peer evaluation, and self-evaluation. Duke and Stiggins (1986) found that evidence for teacher observation was gathered through observations, examination of school records, and student achievement. They suggested that this data should be considered during evaluations and any one of these information sources should not be used alone. If used alone, the data sources might not provide accurate information about teachers' performance. Thus, Duke and Stiggins advised the evaluators to use multiple sources of information about teachers if the goal is their growth. Table 1 shows the mean score of the perceptions of the administrators and teachers regarding the sources of performance information considered as part of the teacher evaluation process.

**Table 1**

*Administrator and Teacher perceptions on the sources of performance information considered as part of the evaluation*

Attribute on TEP Question Number (Admin/Teacher)	Administrator Mean score	Teacher Mean Score
Observation of classroom performance	3.10	3.46
Meetings with evaluator	4.30	3.12
Examination of artifacts	4.20	3.64
Examination of student performance	4.60	3.78
Student evaluation	3.40	3.52
Peer evaluation	2.40	2.44
Self-evaluation	3	2.98

Table 2 shows the mean score of the perceptions of the administrators and teachers regarding the extent of the observations used in the evaluation process.

**Table 2**

*Administrator and Teacher Perceptions Regarding the Extent of Classroom Observation*

Attributes on TEP Question Number Admin/Teacher	Administrator Mean Score	Teacher Mean Score
Number of formal observations per year	2.10	2.56

---

Frequency of informal observations per year	3.10	2.22
---	------	------

---

To analyze this part quantitatively, questions from TEP 1-9 for both administrators and teachers were analyzed.

This study was also analyzed qualitatively, through interview questions 1, 2 and 3 for administrators: “How much time is allocated for classroom observations?”; “What aspects do you look for in a formal observation?” and “What tools, rubrics, and formats do you use during formal observations?” Interview questions 1,2,3 and 4 were asked for teachers: “How much time does the evaluator devote to observe your performance?”, “What sources does the evaluator use when evaluating you?”, “What are the most effective aspects of the evaluation process?”; “What are the most ineffective or weakest aspects of the evaluation process at your school?”

Concerning the observation of classroom performance, both administrators and teachers perceived it as an area that needs improvement. During the interviews, both administrators and teachers have mentioned that most of the evaluation is done through informal observations from outside the classroom for few seconds, while formal observations were done once per year on average. Thus, both administrations and teachers perceived that this area needed improvement.

This is consistent with the research results of Sheppard (2013), where upon gathering teacher and administrators’ perceptions, the majority of teachers report that the evaluation process is mainly concentrated on observations, which are neither long enough, nor frequent enough. The crucial role of conducting formal classroom observations is not doubtful (Marshall, et al., 2016), yet observations should not be the only tool used to measure teacher effectiveness (Skedsmo & Huber, 2018).

Meetings with the evaluator is perceived as an area of strength for administrators, and an area that needs improvement for teachers, meaning that teachers need more time and discussions with their evaluator. This is consistent with what Doherty (2009) found; that there was not enough communication between the evaluator and the teacher.

The importance of discussions between teachers and evaluator was also stressed by Lipowsky and Rzejek (2015) who highlighted this issue explaining that teachers need to receive feedback from their evaluators to be able to make improvements.

Examination of artifacts and examination of student performance were both considered as areas of strength by both administrators and teachers. One of the teachers explained during the interview: “We’re not being observed a lot, however, we are constantly being evaluated through our lesson plans, the assessments we prepare, and students’ performance results.”

Student evaluation is perceived as an area of weakness for administrators, since they explained during interviews that formal student evaluations were not conducted, except rarely for secondary classes. Teachers on the other hand, perceived it as an area of strength, explaining that although formal student evaluations were not done, administrators are very open to concern. This makes teachers believe that they are being evaluated greatly by the students. This is again consistent with what Sheppard (2013) found that peer observations and student evaluations are not considered as part of the process; although researchers have extensively found that students have the ability to make indications of their effective and ineffective teachers, if asked the right questions (Ripley, 2012).

Peer evaluation was considered as an area of weakness for both administrators and teachers, during interviews, teachers suggested increasing these types of activities in addition

to holding planned discussion sessions with their peers, highlighting its importance in their professional growth. As one of the teachers suggested, “If we want to improve professional growth, teachers teaching the same subjects should learn from each other, through observing each others’ classes. Then they should sit and discuss the problems that they encounter...”

This is consistent with what researchers have explained that when administrators carefully plan the sessions, where peers sit together to collaborate with each other (King, 2016); and when they create a culture where teachers feel free to express their ideas and disagreements (Anderson, 2012), and when administrators work on building trust between the team members (Hallam, Smith, Hite, Hite, & Wilcox, 2015), then peers use this time effectively and improve.

Finally, self-evaluation was considered by both administrators and teachers as an area of weakness and needed to be considered. During interviews, when administrators were asked to suggest improvements for the evaluation process, they clearly mentioned making peer-evaluation and self-evaluation mandatory. Other administrators explained that although they had done it for teachers, but it’s not part of their yearly routine.

Teachers also stressed the importance of self-evaluation and reflection during the interviews, some of them explaining that they use workshops to evaluate themselves and reflect on their own practice.

Thus, while its importance was recognized by both groups, teachers were not provided with planned opportunities to evaluate themselves. This is consistent with what many researchers explained that self-evaluations and reflections should not be left to choice but should be carefully planned helping teachers to set goals, then following up with them in the implication process, clarifying to teachers which content to reflect on, and accordingly plan for improvements (Glickman, Gordon, & Ross-Gordon, 2012; Owen, 2015; Philpott & Oates, 2017; Steeg, 2016).

The extent of classroom observations includes two areas, which are the number of formal observations per year, and frequency of informal observations per year. Both administrators and teachers perceive them as areas that need to be reconsidered.

#### **IV – Analysis of Results**

Regarding the different sources of performance information considered as part of the evaluation, both administrators and teachers perceived that most used sources were examination of artifacts and examination of student performance. Meeting with the evaluator was an area of strength for administrators, and an area that needed improvement for teachers. Clearly teachers needed more time spent with their evaluators. Likewise, student evaluation was perceived as an area that needed improvement for administrators and an area of strength for teachers, since administrators were open for student concerns which made teachers believe that they were being evaluated greatly from the students. On the other hand, student surveys were not conducted often, which made administrators believe that this area needed some improvement.

Respecting the areas of observation of classroom performance, peer-evaluation and self-evaluation are perceived by administrators and teachers as areas that need to be reconsidered. Hence, the only two areas that are regularly used as part of the evaluation process are the examination of artifacts and student performance. Thus, administrators need to consider including other sources like regular formal observations, peer-evaluations and self-evaluations as part of the evaluation process.

With regard to the areas related to classroom observations, the quantitative data suggested that both administrators and teachers considered the formal and informal observations as areas that needed to be reconsidered. The qualitative data reported that 57% of

the administrators explained that formal observations were done an average of twice per year, while the other 43% of the administrators explained that formal observations were done only when there was a problem or a particular point that needed clarification. Teachers also explained that formal observations were not conducted regularly and that most of the observations were done informally, from outside the classroom. In listing the most ineffective aspects of the evaluation process, teachers were concerned about not having a serious evaluation process that works on their needs and weaknesses. This meant that the evaluation process needed to be reconsidered. Finally, some teachers explained that since most of the evaluation information was being gathered from parents and students, too much emphasis was being placed on parents' and students' ideas.

When administrators were asked about the tools, rubrics, and formats they use while conducting a formal observation, some administrators did not answer due to the lack of formal observations taking place at their schools. While, others explained that the tools used were standard checklists, taken from the internet and adapted according to the school's needs and requirements. This was in addition to the evaluator's notes taken during observation. One administrator also stated having a personalized evaluation sheet used during classroom observations. However, this tool was not shared with teachers. Only administrators had access to it.

Thus, according to administrators and teachers, the areas of strength of the evaluation process in Armenian schools include having general but clear standards, in addition to including examination of artifacts of student performance in the evaluation process. On the other hand, areas of weaknesses that need improvement are: having a clear evaluation process which includes specific evaluation standards appropriate for teaching assignments and unique for individual needs; sharing these standards at the beginning of the school year with the teachers; revisiting these standards throughout the year to assess teacher needs; and developing an improvement plan. Including a variety of sources in the evaluation process would also benefit. Such changes would include: increasing the number of formal observations of classroom performance, including peer evaluation and self-evaluation, including student evaluation especially for secondary classes, and increasing the number of meetings and discussions held with the evaluator.

Different schools used different sources to evaluate their teachers and to maximize their professional growth. However, only one of the seven schools followed a different approach to evaluate and increase their teachers' growth. This approach was not using different evaluation sources to gather data, rather it was concentrated on observing the classrooms, yet the teachers in these schools were very positive about their evaluation system and perceived this system to help them maximize their growth. The approach was concentrated on having an outside evaluator who would train teachers on a specific skill, observe them in the classroom to see whether the goal has met and learning has taken place, provide immediate and specific feedback on teacher's performance, and in case the goal is not fully met, the plan for improvement is to observe the teacher again to provide another feedback. Thus, this approach was concentrated on providing a content specific workshop that works on one goal or skill at a time, followed by classroom observation by the same person who trained the teachers in the first place, followed by immediate and content specific feedback, followed by a plan for improvement and then another follow-up observation for that particular skill, if needed.

Hence, having a single observer was not the problem. The problem was rooted in the characteristics such as having observational skills, being knowledgeable in the content area. Knowing which standards or behaviors were observed or evaluated upon help reduce teacher's

stress and work on achieving those particular standards. More regular and purposeful classroom observations followed by immediate feedback, helped teachers learn where they are in the new skills learned and how they can improve. Getting observed on one skill at a time, instead of a list of behaviors, in addition to getting trainings or demonstrations on how to apply these skills in the classroom prior to observation. These observations were part of the formal evaluation where the concentration was on the improvement.

Used in the summative evaluation, these additional data sources would help teachers reflect on their own practice and plan for improvement, yet for the improvement to take place, teachers need more regular guidance and follow up.

### V. Conclusion

This study shows that having a multidimensional evaluation system is greatly appreciated by teachers and seen to be important by administrators since teachers gather different viewpoints about their performance, and administrators feel less responsible and pressured in making evaluation judgements. However, this approach was directly related to improve teaching. When asked for Improvements, teachers mentioned using more sources to collect data about their performance in order to gain a wholistic view about their own performance, yet the classroom observations were most common and perceived to be the most beneficial method for increasing teacher's growth, on the condition that these observations were followed by immediate and specific feedback from an observer who was knowledgeable in the subject area he/she is observing. Thus, using a multidimensional teacher evaluation was perceived to be important but not directly related to teacher's professional growth. To help teacher's professional growth two other important factors needed to be considered, which were the evaluator and the type of feedback provided to teachers.

### References

- [1] Almutairi, T. S., & Shraid, N. S. (2021). Teacher Evaluation by Different Internal Evaluators: Head of Departments, Teachers Themselves, Peers and Students. *International Journal of Evaluation and Research in Education*, 10(2), 588-596.
- [2] Anderson, J. (2012). States try to fix quirks in teacher evaluations. *New York Times*, 161(55687), 1. <http://www.nytimes.com/2012/02/20/education/states-addressproblems-with-teacher-evaluations>
- [3] Aramath, G. A. (2014). *Investigating practices of research-proven multidimensional teacher evaluation systems in Michigan schools*. Western Michigan University.
- [4] Barton, H., & Shana, N. (2010). *Principals' Perceptions of Teacher Evaluation Practices in an Urban School District*. ProQuest LLC. 789 East Eisenhower Parkway, PO Box 1346, Ann Arbor, MI 48106.
- [5] Brandt, C., Mathers, C., Oliva, M., Brown-Sims, M., & Hess, J. (2007). *Examining district guidance to schools on teacher evaluation policies in the Midwest region* (Publication No. 030). Washington, DC: Institute of Education Sciences U.S. Department of Education. Retrieved from <http://ies.ed.gov/ncee/edlabs/projects/project.asp?ProjectID=30>
- [6] Callahan, R. E. (1962). *Education and the cult of efficiency*. Chicago: The University of Chicago Press.

- [7] Centra, J. A. (2005). Will teachers receive higher student evaluations by giving higher grades and less coursework? *Educational Testing Service*. Retrieved from <http://www.ets.org/Media/Products/StudentEval.pdf>
- [8] Clark, C. A. (2017). Professional Growth: Informal Peer Observation.
- [9] Colby, S. A. (2001). *A comparison of the impact of state-mandated and locally developed teacher evaluation systems*. East Carolina University.
- [10] Corcoran, S. P. (2010). Can Teachers Be Evaluated by Their Students' Test Scores? Should They Be? The Use of Value-Added Measures of Teacher Effectiveness in Policy and Practice. Education Policy for Action Series. *Annenberg Institute for School Reform at Brown University (NJ)*.
- [11] Danielson, C. (1996). Enhancing professional practice: a framework for teaching (Virginia, Association for Supervision and Curriculum Development).
- [12] Danielson, C. (2007). *Enhancing professional practice: A framework for teaching*. ASCD.
- [13] Danielson, C., & McGreal, T. L. (2000). *Teacher evaluation to enhance professional practice*. Ascd.
- [14] Darling-Hammond, L. (2012). *Powerful teacher education: Lessons from exemplary programs*. John Wiley & Sons.
- [15] Doherty, J. F. (2009). *Perceptions of teacher and administrators in a Massachusetts suburban school district regarding the implementation of a standards-based teacher evaluation system*. Seton Hall University.
- [16] Donaldson, M. (2009). So long Lake Wobegon? Using teacher evaluation to improve teacher quality. Washington, DC: *Center for American Progress*. Retrieved from [http://www.americanprogress.org/issues/2009/06/pdf/teacher\\_evaluation.pdf](http://www.americanprogress.org/issues/2009/06/pdf/teacher_evaluation.pdf)
- [17] Duke, D. L., & Stiggins, R. J. (1986). *Teacher Evaluation: Five Keys to Growth*. NEA Publications Order Department, Academic Building, Sawmill Road, PO Box 509, West Haven, CT 06516 (Order No. 1536-3-00; \$6.95 prepaid; shipping and handling charges will be added on billed orders).
- [18] Elstad, E., Lejonberg, E., & Christophersen, K. A. (2015). Teaching evaluation as a contested practice: Teacher resistance to teaching evaluation schemes in Norway. *Education Inquiry*, 6(4), 27850.
- [19] Gay, L. R., Mills, G. E., & Airasian, P. W. (2012). *Educational research: Competencies for analysis and applications*.
- [20] Glickman, C. D., Gordon, S. P., & Ross-Gordon, J. M. (2005). *The basic guide to supervision and instructional leadership*. Boston: Pearson Education Company.
- [21] Glickman, C. D., Gordon, S. P., & Ross-Gordon, J. M. (2012). *The basic guide to supervision and instructional leadership*. Pearson Higher Ed.
- [22] Goe, L., Bell, C., & Little, O. (2008). Approaches to evaluating teacher effectiveness: A research synthesis. *National Comprehensive Center for Teacher Quality*.
- [23] Goe, L., & Croft, A. (2009). Methods of Evaluating Teacher Effectiveness. Research-to-Practice Brief. *National Comprehensive Center for Teacher Quality*.
- [24] Hallam, P. R., Smith, H. R., Hite, J. M., Hite, S. J., & Wilcox, B. R. (2015). Trust and collaboration in PLC teams: Teacher relationships, principal support, and collaborative benefits. *NASSP bulletin*, 99(3), 193-216.
- [25] Hamilton, E. R. (2013). His ideas are in my head: Peer-to-peer teacher observations as professional development. *Professional Development in Education*, 39(1), 42-64.

- [26] Hanushek, E. A., & Rivkin, S. G. (2010). Using Value-Added Measures of Teacher Quality. Brief 9. *National Center for Analysis of Longitudinal Data in Education Research*.
- [27] Harris, D. N. (2012). How Do Value-Added Indicators Compare to Other Measures of Teacher Effectiveness? What We Know Series: Value-Added Methods and Applications. Knowledge Brief 5. *Carnegie Foundation for the Advancement of Teaching*.
- [28] Harris, D. N., & Sass, T. R. (2007). Teacher Training, Teacher Quality, and Student Achievement. Working Paper 3. *National Center for Analysis of Longitudinal Data in Education Research*.
- [29] Heneman, H. G., & Milanowski, A. T. (2003). Continuing assessment of teacher reactions to a standards-based teacher evaluation system. *Journal of Personnel evaluation in Education*, 17(2), 173-195.
- [30] Hinchey, P. H. (2010). Getting teacher assessment right: What policymakers can learn from research. National Education Policy Center. Retrieved from <http://epicpolicy.org/publication/getting-teacher-assessment-right>
- [31] Hughes, V. M. (2006). *Teacher evaluation practices and teacher job satisfaction* (Doctoral dissertation, University of Missouri--Columbia).
- [32] King, A. H. (2016). Navigating Collaboration: A Multimodal Analysis of Turn-Taking in Co-Teaching. *Working Papers in TESOL & Applied Linguistics*, 16(2), 56-62.
- [33] Lipowsky, F., & Rzejak, D. (2015). Key features of effective professional development programmes for teachers. *Ricercazione*, 7(2), 27-51.
- [34] Little, O., Goe, L., & Bell, C. (2009). A Practical Guide to Evaluating Teacher Effectiveness. *National Comprehensive Center for Teacher Quality*.
- [35] Iowa State University. (2011, June 11). Student evaluation of teaching: Guidelines and recommendations for effective practice. Retrieved from <http://www.celt.iastate.edu/set/effective.html>
- [36] Marcoux, J., Brown, G., Irby, B. J., & Lara-Alecio, R. (2003). A case study on the use of portfolios in principal evaluation. Paper presented at the annual meeting of the American Educational Research Association, Chicago, IL.
- [37] Marshall, J. C., Smart, J., & Alston, D. M. (2016). Development and validation of Teacher Intentionality of Practice Scale (TIPS): A measure to evaluate and scaffold teacher effectiveness. *Teaching and Teacher Education*, 59, 159-168.
- [38] Marzano, R. J., & Toth, M. D. (2013). *Teacher evaluation that makes a difference: A new model for teacher growth and student achievement*. ASCD.
- [39] Noakes, L. A. (2009). Adapting the utilization-focused approach for teacher evaluation. *Journal of MultiDisciplinary Evaluation*, 6(11), 83-88.
- [40] Normore, A. H. (2005). Integrating personnel evaluation in the planning and evaluation of school improvement initiatives. *American Journal of Evaluation*, 26(3), 348-351.
- [41] O'Day, J. (2002). Complexity, accountability, and school improvement. *Harvard educational review*, 72(3), 293-329.
- [42] Owen, S. M. (2015). Teacher professional learning communities in innovative contexts: 'ah hah moments', 'passion' and 'making a difference' for student learning. *Professional Development in Education*, 41(1), 57-74.
- [43] Palazuelos, A. E., & Conley, S. (2008). Choice in teacher evaluation. Association of California School Administrators. Retrieved from

- <http://www.acsa.org/MainMenuCategories/ProfessionalLearning/PromisingPractices/HumanResources/Teacherevaluations.aspx>
- [44] Papanastasiou, E. C. (1999). *Teacher evaluation: Theories and practice*. Michigan State University.
- [45] Peersman, G. (2014). *Overview: Data collection and analysis methods in impact evaluation*. UNICEF Office of Research-Innocenti.
- [46] Philpott, C., & Oates, C. (2017). Teacher agency and professional learning communities; what can Learning Rounds in Scotland teach us?. *Professional development in education*, 43(3), 318-333.
- [47] Peterson, K. D. (2006). Using multiple data sources in teacher evaluation systems. In a. Stronge, J. H. *Evaluating teaching: A guide to current thinking and best practice*. Thousand Oaks, CA: Corwin Press.
- [48] Peterson, K. D. (1987). Teacher evaluation with multiple and variable lines of evidence. *American Educational Research Journal*, 24(2), 311-317. doi:10.3102/00028312024002311
- [49] Piburn, M., & Sawada, D. (2000). *Reformed Teaching Observation Protocol (RTOP) Reference Manual*. Technical Report.
- [50] Rindler, B. H. (1994). *The attributes of teacher evaluation systems that promote teacher growth as perceived by teachers of intensive English programs* (Doctoral dissertation, Boston University).
- [51] Ripley, A. (2012). Why kids should grade teachers. *The Atlantic*, 10, 1-8.
- [52] Schacter, J., Thum, Y. M., & Zifkin, D. (2006). How much does creative teaching enhance elementary school students' achievement?. *The Journal of Creative Behavior*, 40(1), 47-72.
- [53] Schochet, P. Z., & Chiang, H. S. (2010). Error Rates in Measuring Teacher and School Performance Based on Student Test Score Gains. NCEE 2010-4004. *National Center for Education Evaluation and Regional Assistance*.
- [54] Sheppard, J. D. (2013). Perceptions of teachers and administrators regarding the teacher evaluation process.
- [55] Skedsmo, G., & Huber, S. G. (2018). Teacher evaluation: the need for valid measures and increased teacher involvement. *Educational Assessment, Evaluation and Accountability*, 30(1), 1-5.
- [56] Spooren, P., & Mortelmans, D. (2006). Teacher professionalism and student evaluation of teaching: Will better teachers receive higher ratings and will better students give higher ratings?. *Educational Studies*, 32(2), 201-214. doi:10.1080/03055690600631101
- [57] Steeg, S. M. (2016). A Case Study of Teacher Reflection: Examining Teacher Participation in a Video-Based Professional Learning Community. *Journal of Language and Literacy Education*, 12(1), 122-141.
- [58] Stiggins, R. J., & Duke, D. L. (Eds.). (1988). *Case for Commitment to Teacher Growth, The: Research on Teacher Evaluation*. SUNY Press.
- [59] Viadero, D. (2009). Turnover in principalship focus of research. *Education Week*, 29(9), 1-14.
- [60] Webb, L. D., & Norton, M. S. (2003). *Human resources administration: Personnel issues and needs in education*. Prentice Hall.
- [61] Yon, M., Burnap, C., & Kohut, G. (2002). Evidence of effective teaching: Perceptions of peer reviewers. *College teaching*, 50(3), 104-110.