

Teacher Perceptions of the Value of Teacher Evaluations: New Jersey's ACHIEVE NJ

This manuscript has been peer-reviewed, accepted, and endorsed by the National Council of Professors of Educational Administration (NCPEA) as a significant contribution to the scholarship and practice of school administration and K-12 education.



Kathe Callahan
Kean University

Leila Sadeghi
Kean University

The Teacher Effectiveness and Accountability for the Children of New Jersey (TEACHNJ) Act was adopted by the New Jersey legislature in August 2012 with the intent to raise student achievement by improving the overall quality of instruction. As a result of this act, new teacher evaluation systems, known as ACHIEVE NJ, have been introduced in school districts across the state in an effort to more accurately assess teacher performance and to customize professional development opportunities for teachers based on observed areas of need. The overarching question that informs our research is what impact will ACHIEVE NJ have on the overall value of teacher evaluations and the quality of professional development opportunities offered to teachers. Data collected through survey research presents the pre-implementation practices (2012-2013 school year) as well as one year post-implementation practices (2013-2014) provides a snapshot of what is taking place in school districts throughout New Jersey. The findings reflect teachers' perceptions of the value of teacher evaluation practices, the quality of the professional development opportunities and the value the school administration places on teacher evaluations.

Race to the Top

The funding to support TEACH NJ and ACHIEVE NJ comes from the federal reform initiative Race to the Top (RTT). Background on RTT provides insight as to why so many states, including New Jersey, introduced legislation to reform their teacher evaluation systems and tenure decision processes. In July 2009, the Obama administration launched its \$4.35 billion Race to the Top (RTT) Fund, one of the largest competitive grant programs in the history of public education in the United States. As such, it significantly altered the level of federal involvement in public education through the sheer size of its financial investment and through the articulation of specific federal priorities that were to be met through RTT funding.

In an effort to secure RTT funds, at a time when state budgets were eviscerated by the economic crisis, many states enacted legislation that would reform the standards for teacher evaluations and tenure decisions. Many states rushed through hastily crafted legislation to secure federal dollars that were needed to close the revenue gap and forestall drastic cuts in personnel.

TEACH NJ

The Teacher Effectiveness and Accountability for the Children of New Jersey Act (TEACH NJ) was signed into law on August 6, 2012.

In 2011, after two failed rounds of competition, New Jersey was awarded \$38 million to “reform” education. According to the NJ Department of Education, RTT funds will be used to pilot and develop a new educator evaluation system, which is the foundation of the TEACH NJ tenure reform act. The TEACHNJ Act calls for a four level evaluation system of teachers that links individual student data to teachers and creates a more difficult process for teachers to earn tenure.

Under the new law, teachers work for four years, with one of those years under the guidance of a mentor, before the tenure decision is made. During their first four years, new teachers must consistently earn good grades on annual performance evaluations in order to attain tenure. TEACH NJ also targets teachers who have already earned tenure. In a major change to educational policy, tenured teachers may lose their jobs after two consecutive years of ineffective evaluations. Prior to the legislation, school districts could dismiss tenured teachers for “inefficiency,” but the process for doing so took years and could often cost districts hundreds of thousands of dollars, leading many school districts to avoid the process all together. Now, teachers have 105 days after a school district files tenure revocation papers with the state to appeal the decision. Under the new law arbitration will take place outside of the courts and costs will be capped at \$7,500. In addition, the legal costs will be paid by the state. This reduction in administrative and financial burdens is thought to be an incentive for school districts to pursue the dismissal on ineffective teachers.

ACHIEVE NJ

Beginning in September 2013, all New Jersey’s teachers will be evaluated on an annual basis. The evaluations will be based on multiple observations of classroom performance as well as student growth objectives (SGO). Rather than relying on absolute standardized test scores, a statistical formula will determine student growth from year to year (called value-added) and compare that growth to that of their peers. Every teacher will receive a summative rating of

“highly effective,” “effective,” “partially effective,” or “ineffective” which replaces the binary system that rated teachers as “satisfactory” or “unsatisfactory.” The summative evaluation is based on 20 percent SGO and 80 percent teaching practice.

Teacher Evaluations

In theory, a teacher evaluation system should measure a teacher’s strengths and weaknesses through an accurate and consistent process that provides timely and useful feedback. The evaluation and feedback should inform instruction and professional development opportunities to improve classroom instruction and educational outcomes (Marzano, 2012). According to Kelley and Maslow (2005), “Teacher evaluation systems ideally should foster improvement in both professional development opportunities and teaching practices” (p.1). However, in the real world theory often fails to inform practice. Marshall (2005) demonstrated that “the theory of action behind supervision and evaluation is flawed and the conventional process rarely changes what teachers do in the classrooms” (p.274).

Inadequate assessments are all too common, which means poor performance is not addressed, teaching excellence goes unrecognized, new teachers do not receive the feedback they need, and professional development is not aligned with areas of need. The evaluation process can play an important role in developing teachers’ instructional capacity, which in turn contributes to the academic achievement of students (Sergiovanni & Starrat, 2002), however teacher evaluations, as currently conducted, fall short. Overall, teacher observations are brief and infrequent and they fail to differentiate among teachers.

Proponents of education reform rightfully argue that the current teacher evaluation systems are inadequate (Darling-Hammond, 2013; Marzano, 2012; Weisburg, et al 2009; Danielson, 2001). Often, these evaluations involve a short “walk through” visit by the principal or other administrator. The evaluators rely on a rubric that serves as a checklist of what they observe in the classroom. These rubrics tend to focus on trivial items that can be measured and have little to do with learning outcomes, school improvement efforts or professional development opportunities (Donaldson, 2008; Varlas, 2009).

Researchers found that teachers in Chicago had positive perceptions of the overall teacher evaluation process, especially when they valued the leadership of their principals and principal-teacher trust was rated as high (Jiang, Spote & Luppescu, 2015). In addition, they found the evaluation process contributed to teacher stress and decreased satisfaction in the teaching profession. They also learned that “teachers had negative perceptions about the inclusion of student growth metrics” (Jiang, et al 2015, p. 113).

Decades of research show there is a significant relationship between teacher effectiveness and student learning (Darling-Hammond, 2015; Darling-Hammond, 2013; Danielson, 2001; Tucker and Stronge, 2005). According to Darling-Hammond (2000), the “effects of well-prepared teachers on student achievement can be stronger than the influences of student background factors, such as poverty, language background, and minority status” (p. 39). And yet, existing teacher evaluation systems often illustrate no relationship between teacher effectiveness and student outcomes. On paper, almost every teacher is a good teacher, even at schools where student outcomes are dismal. In New York City, a school system with 89,000 teachers, only 1.8 percent of teachers were rated unsatisfactory (Brill, 2009) and in Chicago, where roughly 25 percent of high school students do not graduate on time, and 33 percent of fourth graders are not reading at grade level, 99.7 percent of teachers are evaluated as satisfactory to distinguished

(Rich, 2012). Weisburg and his colleagues (2009) conducted research on the rigor of teacher evaluations of 12 school districts in four different states and found “less than one percent of surveyed teachers received a negative rating on their most recent evaluations (p.10).”

According to Morgaen Donaldson (2009) “Multiple factors, often working in tandem, produce this effect. External constraints decrease evaluators’ inclination to evaluate rigorously – vague district standards, poor evaluation instruments, overly restrictive collective bargaining agreements, and a lack of time all contribute to this problem” (p.2). Internal constraints including a school culture that discourages negative ratings and a district culture that offers little oversight and few incentives contribute to the inflated teacher ratings.

The American Federation for Teachers (AFT, 2010) and the National Education Association (NEA, 2010) have acknowledged the need to reform teacher evaluation systems as the existing systems are inadequate. Both associations highlight the importance of using multiple measures to assess teacher effectiveness, such as classroom observations and district wide assessments, as well as additional opportunities for feedback. They also emphasize the importance of targeted professional development.

Measuring teacher performance is complicated and there is no formula for what makes a good teacher, which means there is no formula for what should be included in the evaluation. Evaluation systems have multiple purposes. Danielson (2012) believes that teacher evaluations should focus on accountability and improvement while Marzano (2012) identifies the dual purpose of teacher evaluations as measurement and development. Both experts agree that one system of evaluation cannot effectively serve both purposes. “Although efforts to move quickly in designing and implementing more effective teacher evaluations systems are laudable, we need to acknowledge a crucial issue – that measuring teachers and developing teachers are different purposes with different implications. An evaluation system designed primarily for measurement will look quite different from a system designed primarily for development” (Marzano, 2012 p. 15).

Professional Development

Research demonstrates that professional development opportunities, when properly designed and implemented, have the potential to enhance classroom practices and ultimately improve student learning outcomes (Fullan et al, 2006; Guskey, 2002). The key is providing professional development that is timely, relevant and effectively delivered. Professional development that is provided in an effective way can have a measurable impact on school improvement and student achievement (Schmoker, 2006; Mathers, Olivia & Lane, 2008). Historically, professional development programs were developed with little input from teachers. Research shows that when professional development programs are mandated, and there is a “pre-determined political agenda for instructional change and teachers’ perspectives are not valued during professional learning” little professional development takes place (Grierson & Woloshyn, 2013, p. 403). However, when teachers have the opportunity to inform the professional development training agenda, positive learning outcomes are realized and the transfer of knowledge is more effective (Alexander & Swafford, 2012; Edmond & Hayler, 2013; Alderman, 2004; Gregoire, 2003).

Moore (2002) conducted a study of 224 teachers and 23 administrators to assess their perception of the New Jersey Professional Development Initiative. The findings highlighted “considerable disjuncture between what teachers value and what they do in the area of professional development” (p. 156). According to Moore, professional development was a

“compliance vehicle” (p. 158) with teachers attending random workshops to accumulate the mandatory 100 hours of professional development required by the initiative. The focus was compliance, not professional or personal growth. Similarly, a recent report by McKinsey & Company (2012) found that most school districts tend to offer the same set of training courses each year without reflecting on what worked and what did not.

Chappuis, Chappuis and Stiggins (2009) find “it’s essential to emphasize the long-term, ongoing nature of professional development as opposed to a short-term, commercially promised quick fixes” (p. 57). A one-time professional development seminar for hundreds of teachers is not as effective as ongoing and personalized professional development that is found in professional learning communities and realized through peer coaching (Rhodes & Beneicke 2002). Research demonstrates that professional development is most effective when it is offered on-site, is job embedded, sustained over time, centers on active learning, and focuses on student outcomes (Chappuis et al, 2009; Sparks, 2003).

While there is a substantial body of research on professional development that identifies the essential characteristics of professional development, there is growing evidence that only a small percentage of what is known to work is actually being implemented (Hawley & Valli, 2000; Spicer, 2008).

Methodology

This research explores the current teacher evaluation and professional development practices in the state of New Jersey. The survey was designed to ascertain teacher perceptions of 1) the evaluation system in their school, 2) the level of communication between teachers and administrators, and 3) the availability, frequency and effectiveness of professional development opportunities. In addition, we wanted to ascertain if teachers are encouraged to participate in professional development activities as a result of their evaluations.

The original survey was pre-tested with a random sample (N=50) of New Jersey schoolteachers. Based on the feedback from the pre-test phase, the survey was revised and administered to a random sample (N=1235) of New Jersey schoolteachers during the spring of 2012 and yielded a 21% response rate (254 completed surveys).¹ Sixty-six percent of the survey respondents were female and 34 percent were male. In terms of tenure, 72 percent of the respondents were tenured teachers, while 28 percent were untenured.

The second survey was distributed to a random sample (N=1560) of schoolteachers in New Jersey during the spring of 2014 and yielded a 23% response rate (364 completed surveys). Seventy-five percent of the respondents were female and 25 percent were male. Most (89%) respondents worked in public school districts. There was less disparity with tenured versus non-tenured respondents in the 2014 survey, with 58 percent being tenured and 42% being non-tenured.

The original survey enabled us to gather baseline data for the 2012-13 school year; the year prior to the implementation of the new teacher evaluation system across all of the state’s districts. The 2014 survey provides data on teacher evaluations and professional development following the first full year of ACHIEVE NJ implementation. The survey will be replicated annually through the 2016-2017 school year.

¹ The researchers did not include data from partially completed surveys.

Findings

After analyzing the data we categorized the responses into four themes: formal evaluation process, impact of evaluation on teaching practice, perceived administrative value, and professional development needs.

Formal Evaluation Process

We asked our respondents to indicate how often they received a formal evaluation by their school principal or assistant principal, other teachers or members of the school management team, or from an external individual such as a supervisor from central office (See Figure 1). During the pre-implementation year, 21 percent of respondents indicated having never been evaluated by their principal or assistant principal during the school year; while 10 percent of the respondents during the post-implementation year indicated they never received an evaluation from their principal or assistant principal. In the pre-implementation year 15 percent indicated having been evaluated three or more times, while in the post implementation year over 30 percent of the respondents had been observed three or more times. In 2012, 23 percent of the respondents strongly agreed that the evaluation was a fair assessment of the quality of their work, while 14 percent strongly agreed the evaluation was helpful. In 2014, 22 percent of the respondents strongly agreed that the evaluation they received was a fair assessment of the quality of their work, and 12 percent strongly agreed that the evaluation was helpful.

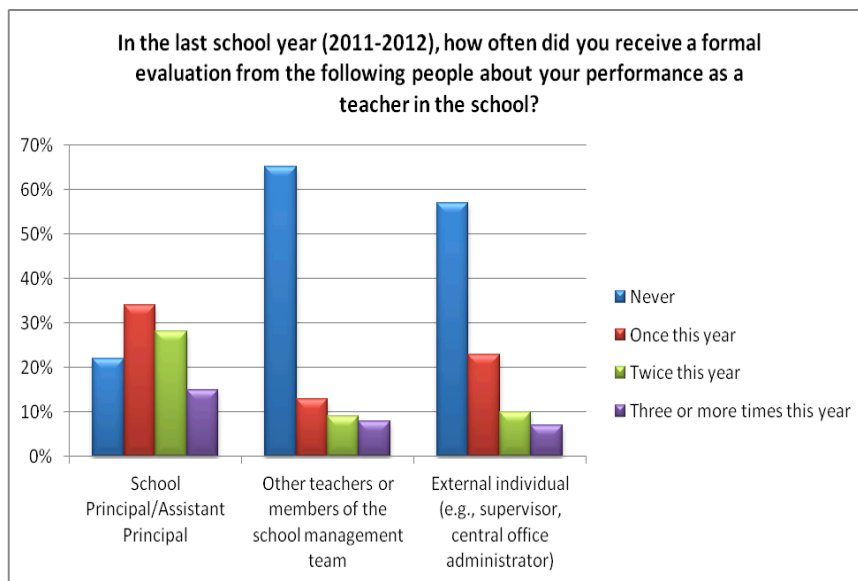


Figure 1A. Formal Evaluation Process Pre-implementation of ACHIEVENJ

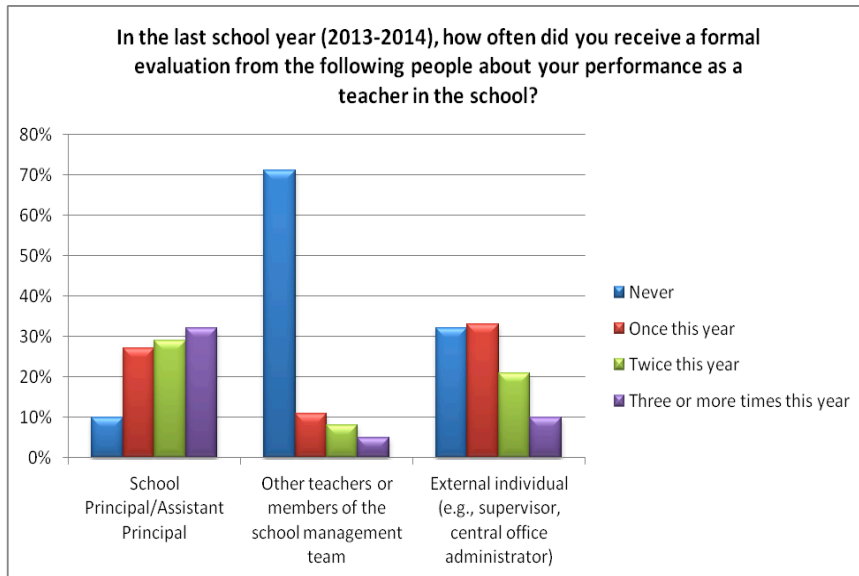


Figure 1B. Formal Evaluation Process One-Year Post-implementation of ACHIEVENJ

Impact of Evaluation on Teaching Practice

We asked the respondents to what extent the formal evaluation they received led to changes in teaching children with special needs, raising student test scores, handling student discipline, knowledge of subject pedagogy, and classroom management. Across all five categories, in 2012 and 2014, over half of the respondents felt the evaluation had no impact and resulted in no change (See Figure 2)

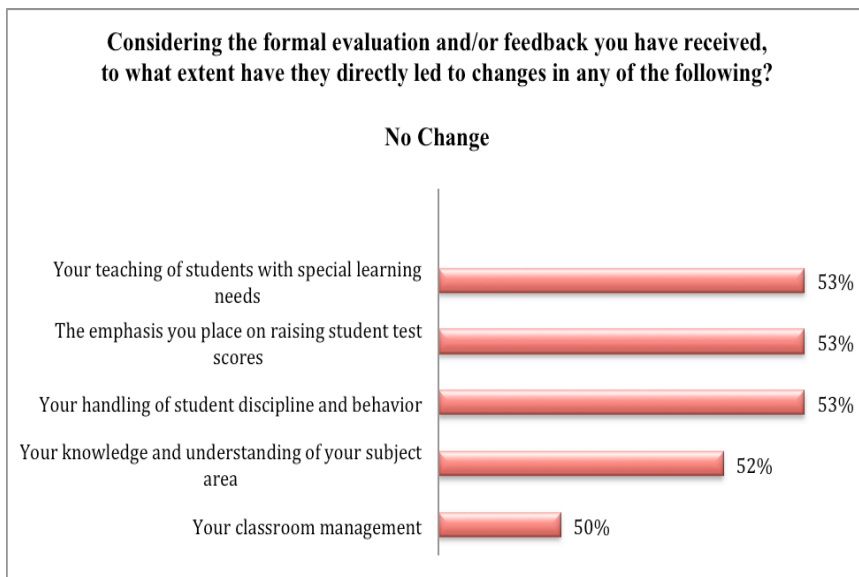


Figure 2A. Perceived Effects of Formal Evaluation Pre-implementation of ACHIEVENJ

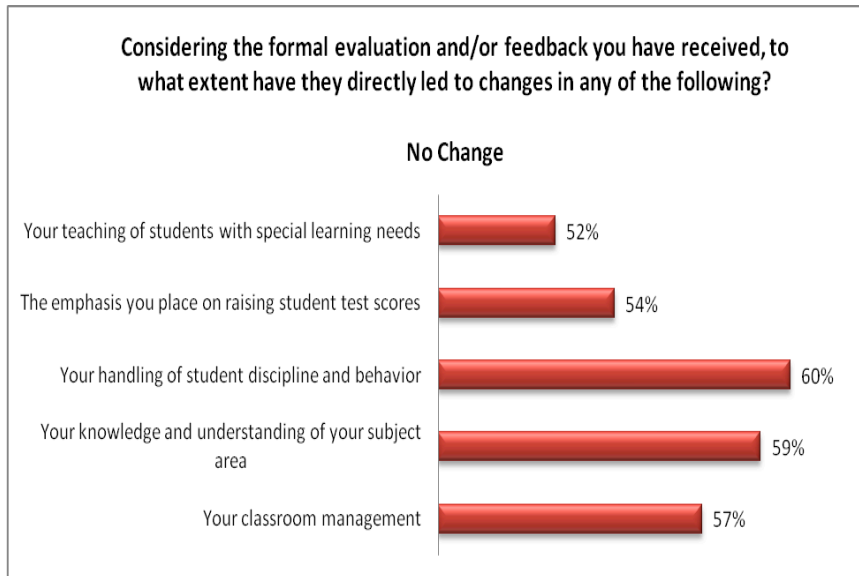


Figure 2B. Perceived Effects of Formal Evaluation One-Year Post-implementation of ACHIEVENJ

Perceived Administration Value

In an effort to develop a better understanding of the administrative value of the teacher evaluations we asked respondents to indicate how and if the outcomes of evaluations impact personnel decisions (See Figure 3). In 2012, 31 percent of the respondents agreed/strongly agreed that a teacher would be dismissed because of sustained poor performance, while in 2014 it increased to 43 percent. In 2012, slightly more than 39 percent agreed /strongly agreed that administrators work with teachers to develop individual professional development plans and this increased to 42 percent in 2014. In 2012, 41 percent of the respondents agreed/strongly agreed that the administration offers no incentives for improved teaching practices and in 2014, 42 percent of the respondents agreed/strongly agreed with this statement. In 2012, 44 percent agreed/strongly agreed that the formal evaluation had little effect on the way they teach and in 2014, 42 percent felt the same way.

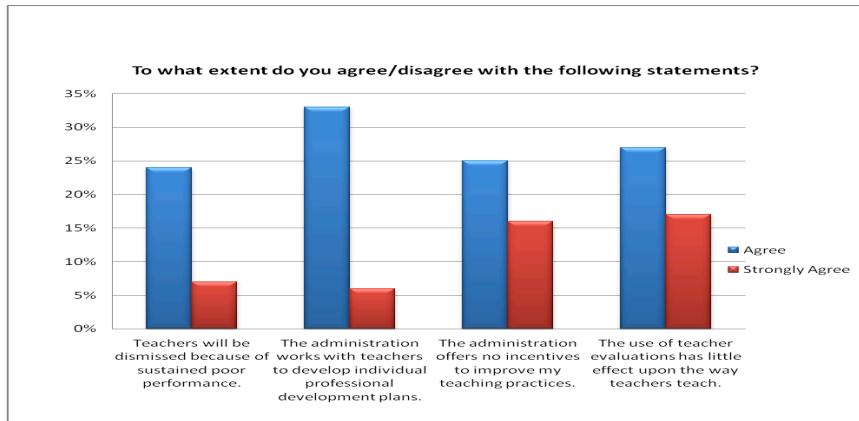


Figure 3A. Perceived Administrative Value Pre-implementation of ACHIEVENJ

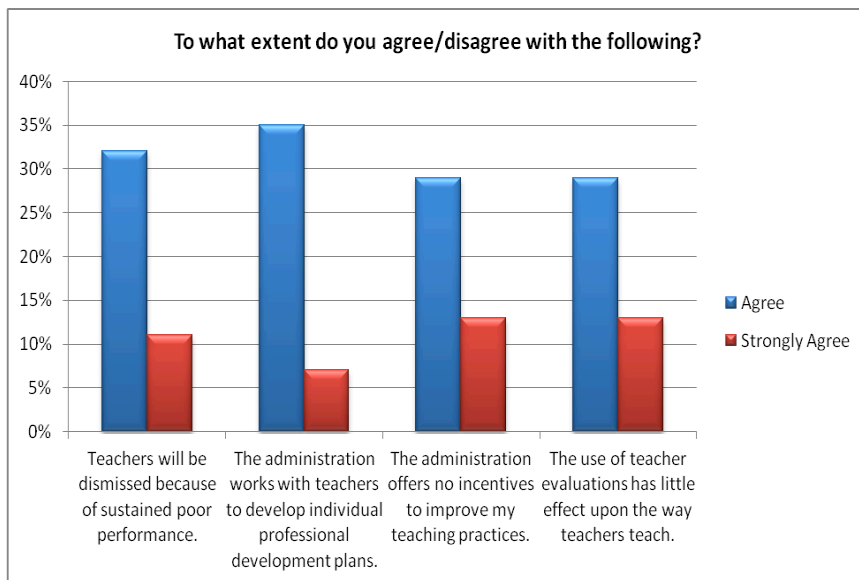


Figure 3B. Perceived Administrative Value One-Year Post-implementation of ACHIEVENJ

Professional Development Needs

The survey asked a series of questions related to professional development. Overall, a majority of respondents (59%) in 2012 indicated that they wanted more professional development but felt there were barriers that prevented them from doing so (See Figure 4). In 2014, 56 percent indicated they wanted more professional development opportunities. Forty percent of respondents indicated they could not participate in professional development because it conflicted with their work schedules. Additionally, 39 percent did not attend professional development because they could not afford it, and 36 percent indicated their district would not reimburse them. Twenty seven percent felt their administration did not support their participation, and only five percent agreed that their administration worked with teachers to develop appropriate professional development that matched their needs.

Additionally, we asked teachers if they participated in professional development activities such as having their colleagues conducting peer observations and the perceived value of this type of professional development on their teaching pedagogy.

When asked, in 2012, if teachers participate in mentoring/peer observations, over 50 percent of the teachers indicated that they did and 66 percent found it had a moderate to large impact on their professional development as a teacher. In 2014, the percentages changed to 53 percent of teachers participating in mentoring/peer observation, with 61percent of those teachers indicating that it had a moderate to large impact on their professional development as a teacher.

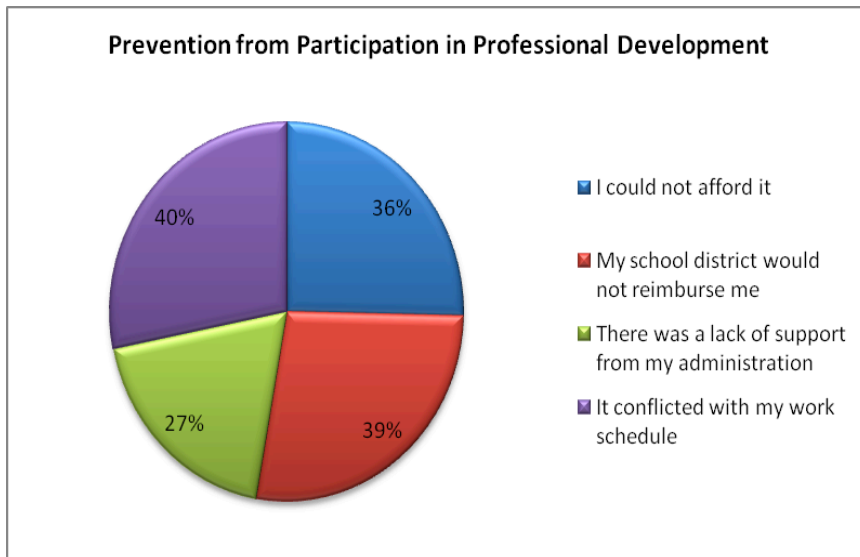


Figure 4A. Barriers to Professional Development Pre-implementation of ACHIEVENJ

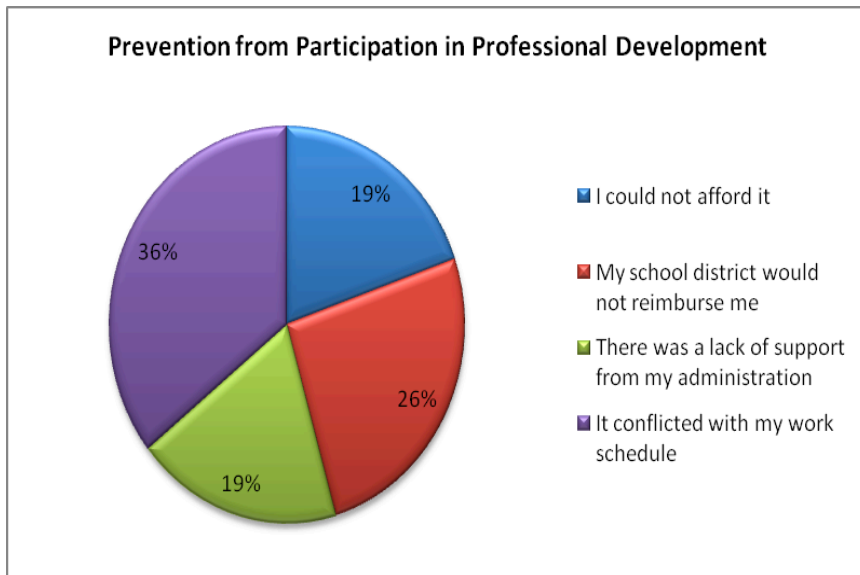


Figure 4B. Barriers to Professional Development One-Year Post-implementation of ACHIEVENJ

Discussion

In 2012 we found that formal evaluations were conducted infrequently with a varying degree of accuracy and impact. Nearly half of the teachers indicated the formal evaluations did not lead to improvements in their classroom as measured by five different indicators. In addition, a majority of the teachers thought the formal evaluations they received were not an accurate assessment of their teaching abilities. Some of the teachers were not observed at all and many indicated they were only observed once, and often not by a school administrator.

In 2014, there was evidence that the frequency of observations had increased. During the pre-implementation year, 15 percent of the respondents indicated they had been observed three or more times. In 2014, 30 percent of the respondents indicated they had been observed three or more times. While the number and frequency of observations increased, the perception of fairness decreased a percentage point from 23 to 22 percent as did the perception of helpfulness, which decreased from 14 to 12 percent. Another way to assess the frequency of observation is to look at the percentage of respondents who indicate they have never been observed. In 2012, 21 percent of the respondents had never been observed. In 2014 it decreased to 10 percent of the respondents never receiving a formal observation.

While teachers indicated that they were observed more often, they also noted the value of the observation was diminished. Open-ended comments reflected numerous concerns about the formulaic nature of classroom observations. Several teachers noted that their principals were more focused on entering observations in real time than on teacher-centered observations. They appeared more focused on entering information on tablets, then in actually observing. Teachers noted the technology and demands of observing numerous required elements made the observation scripted.

In 2012, teachers questioned the administrative value of formal teacher evaluations with many questioning the rewards and sanctions associated with the outcome of the evaluations. Thirty-one percent agreed/strongly agreed that the poor performers were not sanctioned. In 2014, it increased to 43 percent. While the perception that teachers will be sanctioned for poor performance increased by 12 percentage points, the percent that indicated that effective teaching would be rewarded increased one percentage point from 41 to 42 percent.

In 2012, 59% wanted more professional development. In 2014, 56% wanted more. A key component of ACHIEVE NJ was to align professional development opportunities with observed areas of needs, yet only 5% of post implementation respondents indicated that administrators designed professional development based on observed need. A majority of the teachers raised concerns about the lack of resources for professional development. Other teachers or administrators provide much of the professional development offered internally. While teaching colleagues can be a valuable source of professional development for inexperienced teachers, survey results indicated that this was best realized through informal mentoring, not formal professional development. In 2012 and 2014, over 60% of the respondents indicated that informal mentoring had a moderate to large impact on their teaching practice.

What does this tell us about the state of teacher evaluations in New Jersey? Teacher observations are conducted more frequently, but the value of the observations has not improved. The frequent observations are more rigid, following a script. Professional development opportunities have changed little. With that observation we recall what Marzano (2012) said about professional development—measuring teachers and developing teachers are quite different. An evaluation system designed to measure teachers is quite different than an evaluation

system designed to assess professional development needs. We also need to remember that effective professional development opportunities are contingent upon sufficient financial resources and there is a genuine concern that the funding available to develop high quality teachers is insufficient.

The regulations associated with ACHIEVE NJ have turned what was once an organic, albeit infrequent, process into a scripted one. Teachers in New Jersey are demoralized and one of the contributing factors is the emphasis on rating teachers. School boards and school administrators should not lose sight of the original intent of TEACH NJ and that is to improve the educational outcomes for all students. Teacher evaluation systems are not perfect and effective teachers are not the product of formulas. Research shows us that much of what effective teachers do cannot be measured by categorical ratings. However, that is not to say we should not attempt to define what effective teachers do and make every effort to replicate it. We need to move beyond checklists and rubrics that fail to acknowledge teaching excellence and we need to identify and offer professional development strategies that are most effective to improving teaching pedagogy and ultimately improving student achievement.

References

- Alexander, A. and Swaffield, S. (2012) Teacher leadership and professional development: perspectives, connections and prospects, *Professional Development in Education*, 38(2), 159-167.
- American Federation for Teachers. (2010). *A continuous improvement model for teacher development and evaluation*. Retrieved from <http://www.aft.org/pdfs/teachers/improvemodelwhitepaper011210.pdf>
- Beatty, B.R. (2000). Teachers leading their own professional growth: Self-directed reflection and collaboration. *Journal of In-Service Education*, 26, 73-91.
- Brill, S. (2009). The rubber room. *The New Yorker*, August 31, 2009.
- Chappuis, S., Chappuis, J. and Stiggins, R. (2009). Supporting teacher learning teams. *Educational Leadership*, 66(5), 56-60.
- Danielson, C. (2001). New trends in teacher evaluations. *Educational Leadership*, 58(5), 12-15.
- Danielson, C. (2012). Observing classroom practice. *Educational Leadership*, 70(3), 32-37.
- Darling-Hammond, L. (2013). *Getting teacher evaluations right*. New York: Teachers College Press.
- Darling-Hammond, L. (2000). Teacher quality and student achievement: A review of state policy evidence. *Education Policy Analysis*, 8(1), 15-17.
- Donaldson, M. L. (2009). *So long, Lake Wobegon?* Center for American Progress, Washington, D.C.
- Dufour, R. and Dufour, R. (2012). *The school leader's guide to professional learning communities at work*. Solution Tree Press.
- Edmond, N. and Hayler, M. (2013). On either side of the teacher: perspectives on professionalism in education, *Journal of Education for Teaching*, 9(1), 209-221.
- Fullan, M. (2008). *The six secrets of change*. San Francisco, CA: Jossey-Bass.
- Gabriel, R. and Allington, R. (2012). The wrong \$45 million question. *Educational Leadership*, 70(3), 44-49.
- Grierson, A. L. and Woloshyn, V. E. (2013). Walking the talk: Supporting teachers' growth with differentiated professional learning. *Professional Development in Education*, 39:3, 401-419.
- Guskey, T.R. (2002). Does it make a difference? Evaluating professional development. *Educational Leadership*, 59(6), 45-51.
- Guskey, T.R. (2000). *Evaluating professional development*. Thousand Oaks, CA: Corwin Press.
- Hawley, W. and Valli, L. (2000). Learner-centered professional development. *Research Bulletin*, 27(1), 7-10.
- Jayaram, K., Moffit, A., and Scott, D. (2012). *Breaking the habit of ineffective professional development for teachers*. McKinsey on Society, McKinsey & Company.
- Jian, J.Y., Spote, S.E., and Luppescu, S. (2015) Teacher perspectives on evaluation reform: Chicago's REACH students, *Educational Researcher*, 44(2), 105-116.
- Marshall, M. (2005). It's time to rethink teacher supervision and evaluation. *Phi Delta Kappan*, 86, 727-730.
- Marzano, R. J. (2012). The two purposes of teacher evaluation. *Educational Leadership*, 70(3), 14-19.

- National Education Association. (2010). *Teacher assessment and evaluation*. Retrieved from National Education Association's website: http://www.nea.org/assets/docs/HE/TeachrAssmntWhtPaperTransform10_2.pdf
- Rhodes, C. and Beneicke, S. (2002). Coaching, mentoring and peer-networking: Challenges for the management of teacher professional development in schools. *Journal of In-Service Education*, 28:2, 297-310.
- Rich, M. (2012). *National school debate is on in Chicago*. New York Times, September 11, 2012.
- Robbins, P. (1995) *Peer coaching: Quality through collaborative work*. in J. Block, S.F. Everson and T.R. Guskey (eds) *School improvement programs: A handbook for educational leaders*. New York: Scholastic.
- Sergiovanni, T.J. and Starrat, R.J. (2002). *Supervision: A redefinition* (7th ed.) Boston, MA: McGraw Hill.
- Sparks, D. (2003). Change agent: An interview with Michael Fullan. *Journal of Staff Development*, 24(1), 55-58.
- Toch, T. (2008). Fixing teacher evaluations. *Educational Leadership*, 66(2), 32-37.
- Tucker, P.D. and Stronge, J.H. (2005). *Linking teacher evaluation and student learning*. Alexandria, VA: ASCD
- Varlas, L. (2009). Making teacher evaluations meaningful. *Education Update*, 51:11, ASCD, Alexandria, VA.
- Weissberg, D., Sexton, S., Mulhern, J. and Keeling, D. (2009). *The widget effect: Our national failure to acknowledge and act on differences in teacher effectiveness*. The New Teacher Project.