



TEACHER
INCENTIVE
FUND

Ways to Evaluate the Success of Your Teacher Incentive Fund Project in Meeting TIF Goals

Anthony Milanowski and Matthew Finster

February 2016



Introduction

The Teacher Incentive Fund grants have provided financial and technical assistance support to develop and implement sustainable performance-based compensation and human capital management systems (HCMS) for educators. This brief outlines some simple methods that grantees could add to their local evaluations to find out how well they are promoting attainment of the four overall TIF goals:

1. Improving student achievement by increasing teacher and principal effectiveness
2. Reforming teacher and principal compensation systems so that teachers and principals are rewarded for increases in student achievement
3. Increasing the number of effective teachers teaching poor, minority, and disadvantaged students in hard-to-staff subjects
4. Creating sustainable performance-based compensation systems

Though some grantees will be addressing these goals through collecting information for the Annual Performance Report (APR), we think grantees might be interested in some simple ways to collect evidence about how well they are accomplishing these goals. These goals could easily be addressed as part of a grantee's local TIF evaluation. Showing stakeholders how the TIF grant is achieving important goals would not only add to the knowledge base about performance-based compensation, but also help grantees build a case for sustaining important components of the TIF program.

TIF grantees might want to consider using some of the approaches described in this brief to track progress toward TIF's goal of improving educator effectiveness, student achievement, and equity. The methods described in this brief will not provide definitive evidence that allows attribution of success to the TIF program alone. That would require more sophisticated evaluation designs, such as used in the TIF 3 national evaluation.¹ However, these methods can help grantees determine if they are moving toward improving effectiveness, student achievement, and equity.

1. *Improving student achievement by increasing teacher and principal effectiveness*

TIF educator evaluation systems are intended to provide valid and reliable measures of educator effectiveness. Thus, a basic way to track progress toward increasing teacher and principal effectiveness is to compare the proportions of educators rated effective and highly effective across years. Many grantees are doing this as part of their Government Performance and Results Act data reporting and APRs.

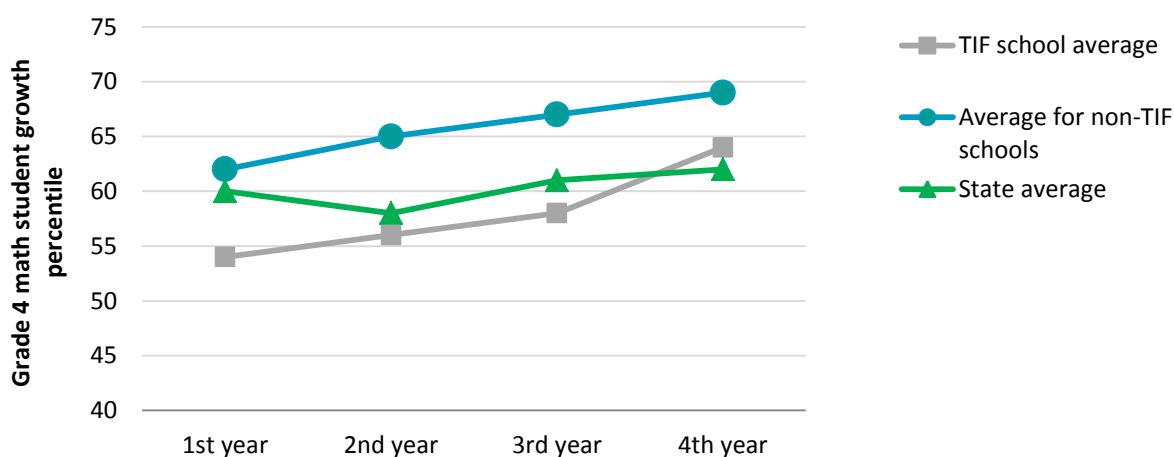
However, grantees may want to go beyond simply tracking the percentages of educators rated effective or highly effective. If educator effectiveness is improving, student achievement should also be improving. So along with looking at overall ratings or proportions of teachers in effective and highly effective categories, it

¹See *Evaluation of the Teacher Incentive Fund: Implementation and Impacts of Pay-for-Performance after Two Years. Executive Summary* at <http://ies.ed.gov/ncee/pubs/20154020/pdf/20154021.pdf>.

would be useful to also track school-level growth measures, such as school value-added or average school student growth percentiles. If educator effectiveness is improving, the effectiveness of schools participating in TIF should also be improving.

For example, Figure 1 below shows a hypothetical comparison of trends in school median student growth percentiles between the average for the schools participating in TIF, schools in the district not participating in TIF, and schools statewide. The graph shows that, on average in the TIF schools, student learning improved, but as importantly, the average for the TIF schools also improved faster than the average for non-TIF schools and faster than student learning was growing in the average school in the state. While this does not show that TIF caused student growth to improve, it does suggest that TIF schools, for whatever reasons, are moving in the right direction and doing so as fast or faster than other schools.²

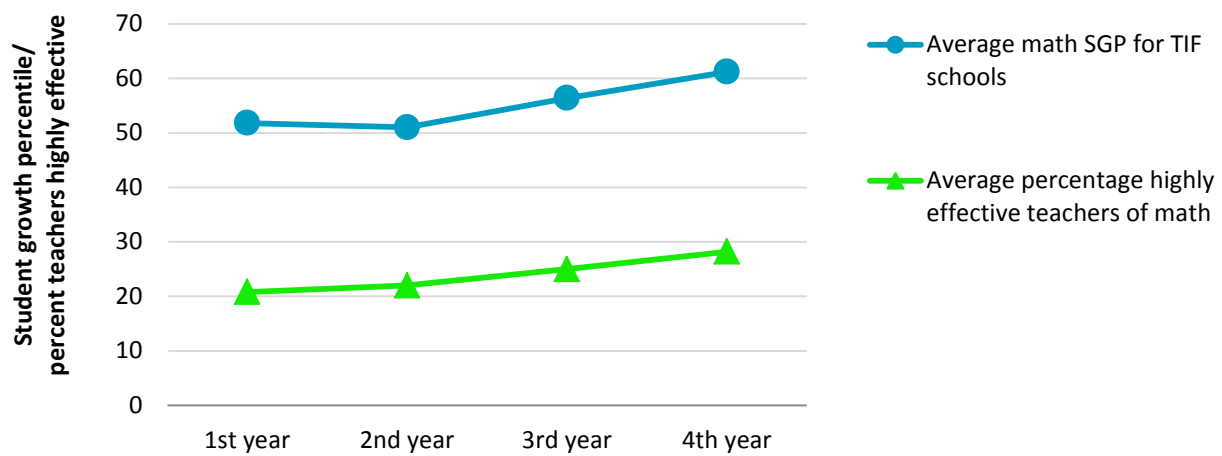
Figure 1. Trend in median mathematics student growth percentiles of TIF schools compared to non-TIF schools and statewide average



Another trend comparison that can provide evidence about TIF and school performance is to graph the average percentage of teachers rated highly effective with a measure of school effectiveness. Figure 2 below shows a hypothetical proportion of teachers who teach mathematics who were rated highly effective in each TIF year and the average math student growth percentile for the schools participating in the TIF grant. It shows that as more teachers were rated highly effective, the average math student growth percentile increased. While this does not prove that TIF caused student growth to increase, a pattern like this is what we would expect to see if TIF helped to increase educator effectiveness. Note that this graph also provides some evidence that evaluation ratings are capturing real increases in teacher effectiveness. If both evaluation ratings and average student growth are going up, we have more confidence that increases in evaluation ratings are showing that teacher effectiveness is improving.

² The choice of a comparison group that closely matches the characteristics of the TIF schools can improve the quality of the evidence provided by this kind of analysis. There are several methods of identifying comparison schools that go beyond looking at non-TIF schools in the district or statewide averages. The TIF TA team or your evaluator can provide suggestions about using more complex methods to determine comparisons.

Figure 2. Trend in median mathematics student growth percentile and percentage of teachers of mathematics rated highly effective



Some grantees might also have the opportunity to track trends in student achievement beginning before their TIF grant. Adding pre-grant data points to the trend comparison can strengthen the evidence for concluding that TIF had an impact on student achievement. This is because schools participating in TIF may already have been on a path of increasing student achievement. Data showing that the progress toward greater student achievement has been faster in TIF schools than in others after TIF began can help to rule out this alternative explanation of an upward trend in student learning. Figure 3 illustrates a hypothetical trend comparison that provides evidence of a positive TIF effect.

Figure 3. Trend in mathematics benchmark assessment scores of TIF schools compared to non-TIF schools

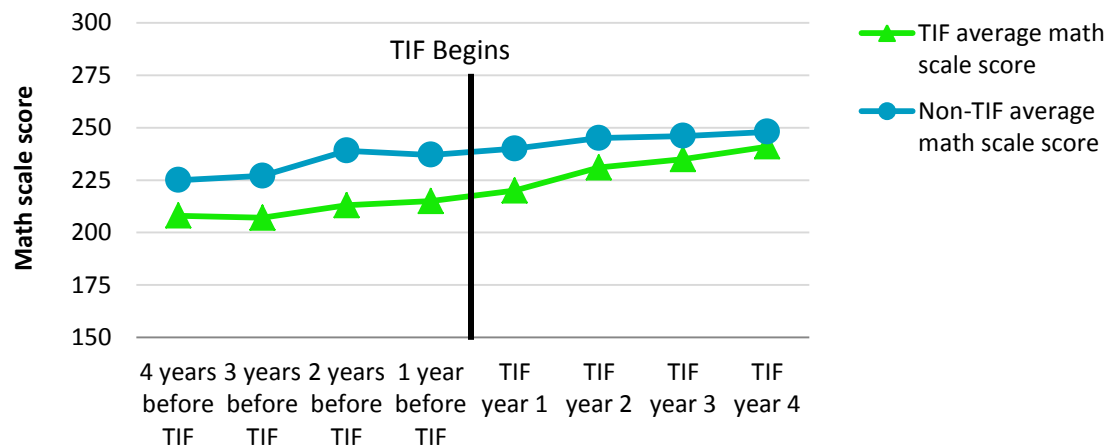


Figure 3 shows that the math scale score has increased in both TIF and non-TIF schools since TIF began, but that the increase was greater in TIF schools. Also, even though the trend was toward higher student achievement in TIF schools before the grant began, the trend accelerated after TIF's first year. This comparison provides better evidence for a positive effect of TIF because, in this hypothetical case, student achievement at TIF schools both improved faster after than before TIF and faster than at non-TIF schools.

Thus, it is less likely that the upward trend for TIF schools is due solely to characteristics of TIF schools in place before TIF began or to conditions affecting both TIF and non-TIF schools.

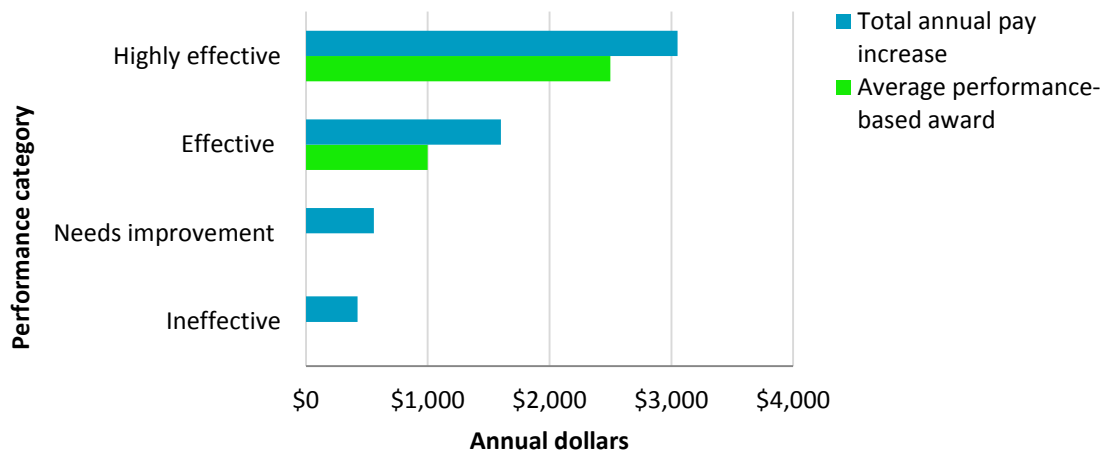
Unfortunately, changes in state tests often make it difficult for grantees to track student achievement trends over time. These changes make comparing the percentage of students scoring at the proficient level or above difficult, though not impossible. Grantees using tests other than state assessments (such as benchmark assessments purchased from test vendors) that have not changed scales over time can make comparisons like that shown in Figure 3. If certain conditions are met, grantees relying on state tests might be able to make comparisons over time based on the relative position of TIF schools in the distribution of school average test scores across the state. The idea here is that if TIF is supporting improvements in educator effectiveness, all else equal, the percentage of students scoring proficient or above in these schools should rise relative to the state average, or to non-TIF schools. As in Figure 3, comparing the trend for TIF schools before and after TIF to the trend in a comparison group helps separate the effects of TIF from other factors already operating in TIF schools and in the broader district or state. The TIF technical assistance team, or the grantee's program evaluator, can advise on whether this approach could work for each grant. Additional information on test transition issues can be found on the TIF Community of Practice (COP) at: <http://www.tifcommunity.org/resources/policy-issues-and-strategies-transition-common-core-assessments>.

2. Reforming teacher and principal compensation systems so that teachers and principals are rewarded for increases in student achievement

While it is fair to conclude that grantees providing performance-based compensation based on educator effectiveness measures that include student achievement growth are addressing this goal, there are several interesting ways to show that educators who increase achievement receive greater rewards than those who do not do as well.

Some grantees have shown this by comparing the average bonus or salary increase amounts for teachers with different effectiveness ratings, as shown in Figure 4.

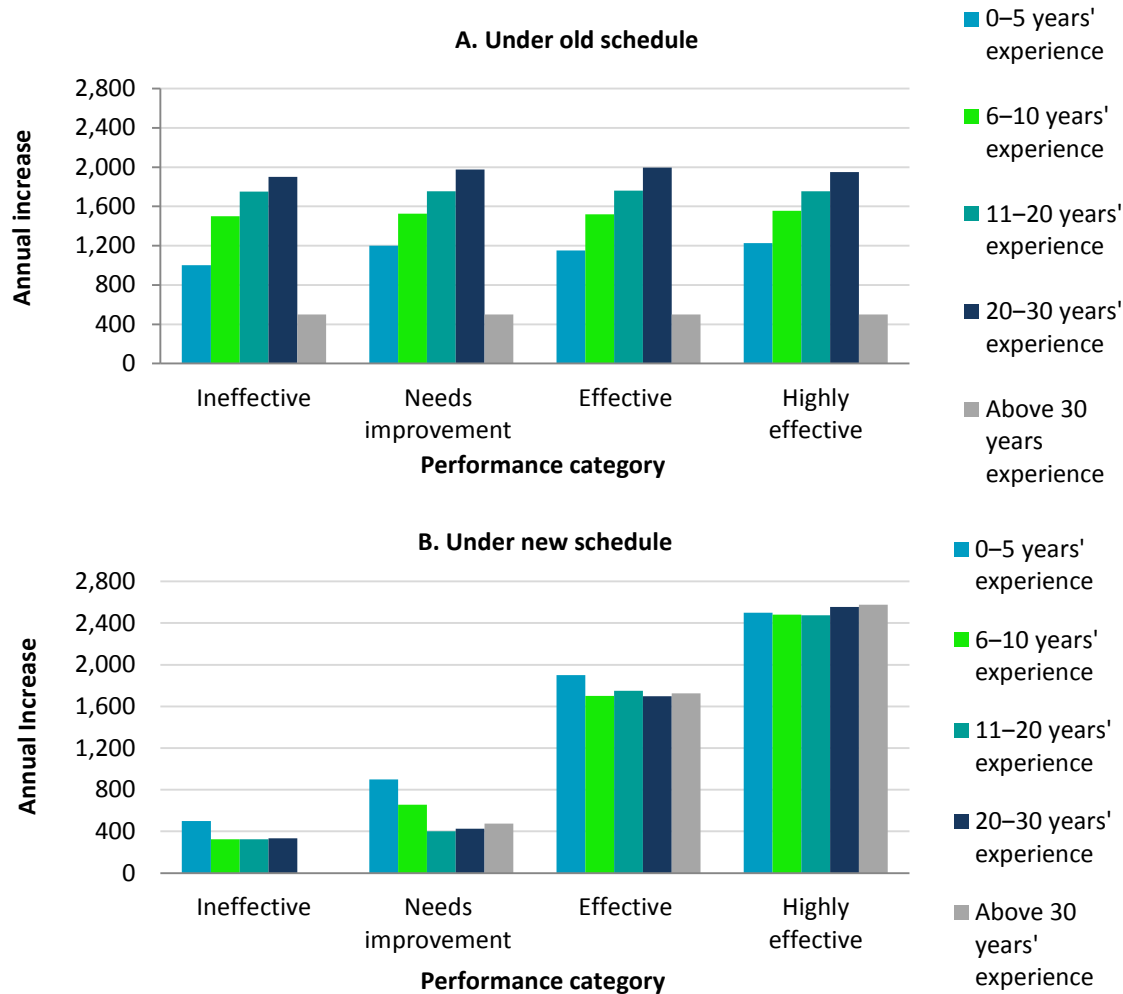
Figure 4. Comparison of compensation increases by teacher effectiveness



This graph shows the average performance-based compensation by performance category and compares it with the performance-based compensation to the total pay increase (performance-based compensation plus longevity bonuses or the annual value of progression on the salary schedule). In this hypothetical case, it is clear that effective and highly effective teachers are receiving substantially larger increases in compensation than teachers rated ineffective or in need of improvement.

For grantees that have implemented or are implementing a new salary schedule, the distributions of pay increases by experience and performance before and after the new schedule was implemented can be compared. Figure 5 is adapted from an analysis done for the Tennessee TIF program evaluation. Comparing the two graphs shows how the new schedule increased the differentiation of pay increases based on performance.

Figure 5. Comparison of average salary increases under old seniority-based schedule and new performance-based schedule



Comparing the two graphs clearly shows that seniority by itself is a lot less important in determining pay increases under the new schedule. In Panel A, there is a clear difference in salary increases by experience within each performance category, while the increases across categories are fairly similar. In Panel B, increases differ substantially across performance categories, but much less by experience within categories.

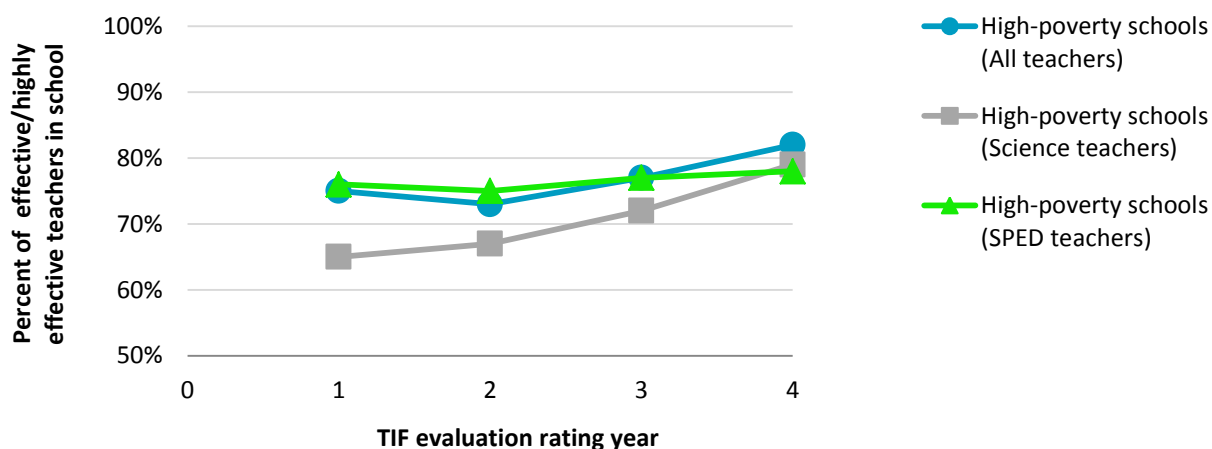
3. *Increasing the number of effective teachers teaching poor, minority, and disadvantaged students in hard-to-staff subjects*

As research has shown, the importance of teachers and principals to student achievement and the importance of access to effective educators to closing achievement gaps have become more apparent. TIF projects have worked to improve the access of poor, minority, or disadvantaged students to effective educators by providing incentives for recruiting and retaining effective educators in schools serving these students and by providing supports such as professional development, coaching, and feedback to improve

educator effectiveness. Grantees can easily examine whether percentage of effective and highly effective educators in these schools has increased, both in general and in hard-to-staff subjects.

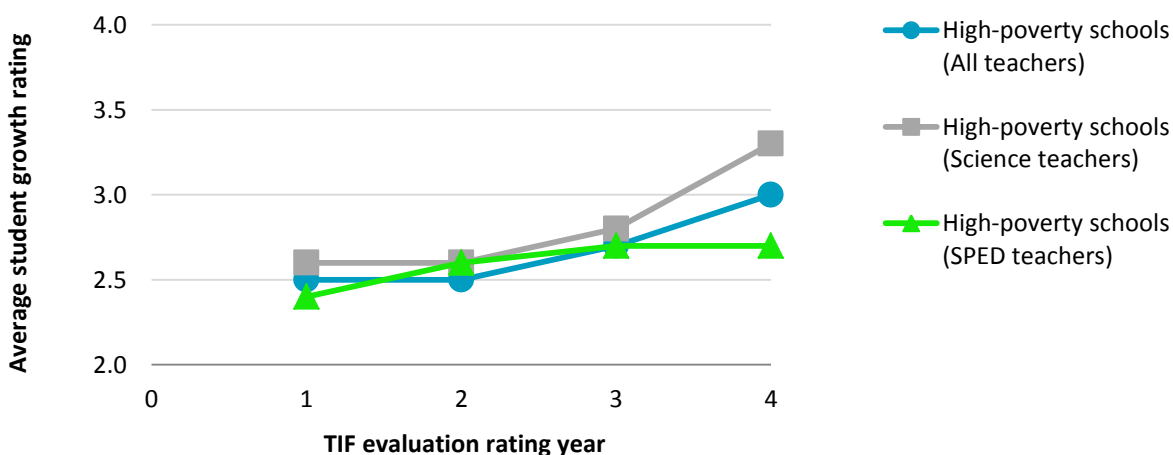
Figure 6 below illustrates a hypothetical case in which the percentage of teachers rated effective or highly effective in the grantee’s high-poverty schools has increased overall, and it has increased for one hard-to-staff subject, science, but not appreciably for another, special education. This suggests the need to look for additional strategies to recruit and/or retain effective and highly effective special education teachers.

Figure 6. Trend in evaluation ratings for teachers in hard-to-staff subjects and high-poverty schools



Another way to look at this question could be to plot the average student achievement growth score for teachers in high-poverty schools. Note that because the assessments used for different subjects and grades differ, the levels of growth might not be comparable. But the trends are still of interest. Hopefully, the average growth rating of each group of teachers will be going up over time, as illustrated by Figure 7 below.

Figure 7. Trend in student achievement growth ratings for teachers in hard-to-staff subjects and high-poverty schools



In this figure, the average student growth rating has increased for teachers in high-poverty schools and for science teachers in those schools, but not as much for special education teachers. Again, this suggests looking

more closely at retention of effective special education teachers and the supports provided to help them improve student achievement.

4. *Creating sustainable performance-based compensation systems*

Planning for sustaining performance-based compensation, as well as other aspects of HCMS TIF grants supported, has been challenging for many grantees. Several TIF COP resources are available to help grantees plan what is needed to sustain the parts of the TIF, such as *Sustaining Your TIF Efforts; A Reflection Guide*. Here, we focus on a few key areas that grantees could track as part of their ongoing TIF evaluations to assess the potential for sustainability and identify areas where additional work might be needed to sustain.

Stakeholder Support

Typically, TIF evaluations assess educator attitudes, an important indicator of support from key stakeholders. In this context, it would be helpful to track educators' attitudes over time about:

- Performance-based compensation in general
- Specific aspects of the current performance-based compensation system
- Fairness of performance measures
- Workload, time pressure, and stress
- Whether they believe performance-based compensation should continue

If these attitudes are reasonably positive, or at least improving, the chances of sustaining the grantee's performance-based compensation system are greater. Where attitudes are not favorable or declining, grantees should consider what might be done to address the concerns of key stakeholders.

Tracking these attitudes could be done by adding some relatively simple questions to annual teacher and administrator surveys or adding them to interviews or focus groups. Several grantees have developed usable survey questions, examples of which are shown in the table below. Including one or two items from each category, or similar items of a grantee's own design, would help determine if the program was attaining sufficient support to be sustained and addressing important reasons (e.g., fairness, increased workload) for educator concern.

Table 1. Example survey items about educators’ attitudes toward performance-based compensation

Subject	Item	Grantee/Study
Opportunity to earn performance-based compensation	I value the opportunity to earn a bonus as a way to be recognized for the impact I have on student performance.	Charleston
	I appreciate having the opportunity to receive additional compensation through <PBC program name>.	Maricopa
	I feel that the bonus/stipend system in my district does a good job at rewarding effective teachers.	Tennessee
	Performance-based compensation should reward teachers for achieving high performance ratings on their evaluations.	CTAC-Delhi
Pay increase/bonus size	Bonus/stipend payments are large enough in value to motivate changes in a teacher’s professional practice.	Tennessee
Fairness of performance measures	The performance measures used to determine my performance-based compensation under <PCBS program name> are fair.	
	The performance targets of <PCBS program name> are fair and equitable.	Maine
	The requirements for performance pay are reliable, consistently applied, and fair.	Washoe
Workload, Stress, Autonomy	The <PCBS program name> has substantially increased my workload.	Illinois evaluation implementation study TIF 3 national evaluation
	The <PCBS program name> has caused me a lot of stress.	
	I have less freedom to teach the way I would like to teach due to the <PCBS> program.	
Retention	The <PCBS program name> has encouraged me to remain with the district.	Harrison
	The opportunity to earn a <PBC program name> award makes it more likely that I will remain in my district.	Maricopa
	The potential for performance pay influenced my decision to stay at my school or transfer to another school.	Washoe

PBC = performance-based compensation.

Since school administrators are both potential recipients of performance-based compensation and play important roles in administering the program for teachers, it is advisable to check for their support, as well as the support of teachers.

Support Systems Capacity

By the third or fourth year of the grant, most grantees will have developed the basic systems for measuring performance, making performance-based compensation decisions, paying educators, and supporting performance improvement. However, the TIF evaluator could also help project staff assess the overall

capacity to sustain key parts of the grant by interviewing key staff in departments that support the performance-based compensation system to get their impressions of the challenges they see in supporting and sustaining TIF. While in many cases TIF program directors will be well aware of the challenges and capacities of support systems, in larger grants, or those encompassing multiple districts, it could be useful to get an overview from the evaluator of the views of those in charge of these systems. What changes have they had to make to support TIF? Have they had to relocate staff or struggle with extra workload? Have new systems put in place to support TIF made some processes more efficient? This information could be useful in deciding what additional resources might be needed to sustain the performance-based compensation system and other activities supported by TIF.

Human Capital Management System Implementation

Having a comprehensive HCMS based on a valid and reliable evaluation model in place will make sustaining a performance-based compensation system more likely. Program evaluations can assess whether this foundation is in place and can contribute to sustainability by examining the state of HCMS development. Several TIF 4 program evaluations looked specifically at the implementation of the HCMS that supports performance-based compensation. The evaluation done for district consortium-led Gilchrist County, Florida, used a survey questionnaire and document review to assess where each participating district was in developing its HCMS.³ Similarly, the evaluator of the Center for Education Innovation’s Partnership for Innovation in Compensation for Charter School grants worked with the grantee to develop an HCMS scorecard to track the progress of the independent charter schools in the grant toward a comprehensive HCMS.

Sustainability Priorities

The TIF evaluator could also be asked to collect information about the aspect of the grant stakeholders would most like to see continue. Adding a question such as “What parts of the TIF program do you think the district should continue?” to interviews, focus groups, or surveys could be useful in getting a wide-angle picture of what stakeholders want to sustain. The TIF COP includes a sustainability checklist that could be useful in developing questions. See <http://www.tifcommunity.org/resources/sustainable-reforms-presentation-and-handouts>.

³ The COP resource, *Designing and Implementing Human Capital Management Systems in Educator Evaluation*, describes how this grantee sought to help participating districts align their HCMS.

Conclusion

Even though TIF grantees may not be able to do rigorous experiments to assess the impacts of their grant activities on the goals of the TIF program, they can use their program evaluations to provide meaningful information on whether they are achieving their goals. Grantees can use simple comparisons over time and, in some cases, comparisons to external standards such as statewide or districtwide trends to see whether measures of the key TIF goals are at least moving in the right direction. This information is also likely to be important when making a case for sustaining important elements of TIF and for deciding which elements should be priorities for sustainability efforts. Grantees can also use their TIF evaluations to collect information about whether the number of effective teachers teaching poor, minority, and disadvantaged students in hard-to-staff subjects has increased. Last, surveys, interviews, and focus groups done as part of TIF program evaluations provide an opportunity to assess the support of educators for sustaining performance-based compensation and other human capital management activities the TIF grant has supported.

Additional Resources

The TIF technical assistance team is available to help your grant and your program evaluator refine how you measure and accomplish the TIF objectives. There are also resources on program evaluation available on the TIF COP website. They include:

Program Evaluation for the Design and Implementation of Performance-Based Compensation Systems

Year of Publication: 2012

Authors: Peter Witham, Curtis Jones, Anthony Milanowski, Christopher Thorn, Steven Kimball

Program Evaluation of Human Capital Management Systems (HCMS)

Year of Publication: 2014

Authors: Matthew Finster, Grant Sim, Ashley Varrato

The Role of Program Evaluation in Communication Strategies

Year of Publication: 2012

Authors: Peter Witham, Clarissa McKithen, Jenna Aurand Scott

Other resources that could be of interest include:

Evaluating Performance Appraisal Programs: An Overview

Year of Publication: 1999

Author: U.S. Office of Personnel Management

<https://www.opm.gov/policy-data-oversight/performance-management/reference-materials/more-topics/eval.pdf>

Evaluating Human Resource Management Programs

Year of Publication: 2007

Authors: Jack Edwards, John C. Scott, Nambury S. Raju