

Socio-Technical Gaps Exposed by the United States Department of Education's Teacher Incentive Fund Grant Program

Christopher A. Thorn

Carnegie Foundation for the Advancement of Teaching, USA
thorn@carnegiefoundation.org

Abstract. The United States (U.S.) Department of Education's Teacher Incentive Fund provides financial support for local school districts, state agencies, and non-governmental agencies to design and implement performance-based compensation systems that reward teachers and leaders who improve student learning. These policies put pressure on data systems and the social systems of supporting organizations to adapt tools and procedures designed for administration and scheduling to deliver high-stakes outcome results. This misfit has led to system failure in some early implementations. It has also led to the development of new technical solutions and unlikely partnerships between social actors in these spaces.

Keywords. Data quality; performance pay; growth measures; educator effectiveness.

1 The U.S. Teacher Incentive Fund Program

Performance based compensation systems are not new to education. There have been prior attempts to explore and implement such programs – usually referred to as *merit pay* – in the 1970s and 1980s [1]. The notion is a simple one – tying pay (sometimes base pay but often one's bonus) to one or more performance measures. In the case of teachers, this could be measured using a number of different metrics – from improvements in student performance on high stakes tests to assuming additional duties or leadership roles in addition to classroom teaching. This paper will describe the reintroduction of performance pay in the education sector in the United States (U.S.) and the substantial change in requirements for technical systems, for analytical skills of teachers and administrators, and for collaboration across traditional organizational boundaries.

While there has been on-going experimentation in evaluating and compensating teachers since the 1980s, there was no large scale implementation of such policies until the creation of the Teacher Incentive Fund. The Teacher Incentive Fund (TIF) program was created in 2006 [2] as a competitive grant program administered by the U.S. Department of Education that was intended to allow schools, school districts, state agencies, and non-governmental partners to design and implement performance-based compensation programs. There were two rounds of grants made under the orig-

inal funding. The appropriation was allocated in two rounds of grants. The first, referred to by the Department of Education as Cohort 1, received a total of \$42 million awarded across 15 grantees in November 2006. The second group of 18 recipients – referred to as Cohort 2 – received \$38 million in June 2007 [3]. There was a third round funding announced in May 2010 [4] and awards made in September 2010. The third round of TIF grants - Cohort 3 – were made from an appropriation of \$437 million. The grants were awarded to 62 recipients – 50 grantees in the general program and 12 grantees who also agreed to participate in a randomized, controlled trial of their designed program against a simple whole-school bonus treatment [5]. The focus of this program is stated in the press release announcing the most recent call for proposals as follows, “TIF grants support local projects that reward teachers, principals and other school personnel who improve student achievement [6].” This basic notion has been operationalized differently across the three rounds of funding. Changes to the program requirements for Cohort 3 reflect a substantial shift from a reliance on compensation alone to a set of measures and interventions designed to achieve the desired core policy goal of improving outcomes for students.

The TIF focus on improving student outcomes – particularly the focus on growth in student learning – has led to a change in data needed to meet the requirements, the quality of that new and pre-existing data, timeliness of data reporting and quality assurance/verification around the use of those data. Data that had been collected and managed for student scheduling, payroll, etc. suddenly needed to reflect new relationships and radically higher-stakes use cases. In addition to their reliance on test-based measures of student progress, all TIF grantees are also implementing observational frameworks based on detailed rubrics to evaluate, score, and share evaluations of teacher and principal practice.

In this paper, I will focus on the grant priorities articulated in the 2010 request for TIF proposals (mentioned above). I will explore the socio-technical implications for implementing these reforms we have discovered while providing technical assistance to the U.S. Department of Education (U.S. ED) and the nearly 100 TIF grantees funded to date. In particular, I will explore the very detailed workshop materials from a presentation made by U.S. ED staff on June 3, 2010. Senior program and policy leaders from the Department held a Technical Assistance Workshop for any organizations considering applying for a Teacher Incentive Fund [7]. The workshop covered eligibility for the program and incentives for a subgroup to participate in a formal evaluation. The primary focus of the presentation was on the absolute and competitive priorities of the competition and how proposals would be evaluated based on the quality with which the grantees addressed these priorities.

TIF program managers at U.S. ED had engaged with the technical assistance providers and major national actors (reform groups, labor unions, etc.) to create more explicit grant priorities for Cohort 3. They also translated the priorities into what they called the Core Elements of the TIF program and stated that they would require all grantees to adequately implement these program features in their TIF projects [4]. The Core Elements provide the primary measurement framework for the required Annual Performance Report (APR) each grantee must submit. The APR also requires all grantees to set goals on all performance based on actual first year performance. Seen

as compliance measures, the Core Elements are indicators of the presence of required program features. At the same time, they also provide a framework for understanding the level or success of program implementation.

2 Core Elements

As stated above, the Core Elements provide the bulk of the program measures for each TIF grantee (in addition to financial review for compliance with spending rules). Each of the core elements has implications for information and communications systems. I explore how some of these requirements have translated into very real challenges for many TIF grantees.

2.1 Core Element A

A plan for effectively communicating to teachers, administrators, other school personnel, and the community-at-large the components of its [performance-based compensation systems] PBCS [4]

In the initial two TIF Cohorts, many grantees had no or very limited communication plans. The requirements for Cohort 3 include the development of an initial plan that identifies all stakeholders and then builds tailored communication channels and strategies to carry the project through its full lifecycle. Grantees are required to engage in on-going communication activities and have strategies in place for communicating changes to the program. Some of the larger grantees developed social media sites and many used role-specific outreach materials to target school staff, parents, and the community at large (see, for example [8]). Grantees now routinely include allocation of funds for hiring communication consultants to help guide the development and implementation of their plans.

The most challenging implications for technical systems in this Core Element are the growing importance of two-way communication in any *social* media channel. One of the most common approaches to two-way communication is the collection of questions (including online submissions, systematic notes taken from face-to-face presentations, etc.) that are then answered in a *Frequently Asked Questions* resource (see [9]). As school districts and schools integrate interactive communication channels into their communication strategies, the expectations of participants in the system for timeliness and transparency will very likely also go up.

2.2 Core Element B

The involvement and support of teachers, principals, and other personnel (including input from teachers, principals, and other personnel in the schools and LEAs to be served by the grant) and the involvement and support of unions in participating LEAs (where they are the designated exclusive representatives for the purpose of collective bargaining) that is needed to carry out the grant [4]

This is the only Core Element that does not explicitly have a technology component. However, beyond the initial agreement to participate in the project, most projects consider the active components of the communication plan as a portion of the evidence of compliance with Core Element B. Involvement usually includes email newsletters, websites, webinars, etc. that communicate the on-going development of the various aspects of the project and convey what role the participation of represented staff members have made by showcasing teacher involvement on specific projects or products.

2.3 Core Element C

Rigorous, transparent, and fair evaluation systems for teachers and principals that differentiate effectiveness using multiple rating categories that take into account student growth ... as a significant factor, as well as classroom observations conducted at least twice during the school year. The evaluation process must: (1) Use an objective, evidence-based rubric aligned with professional teaching or leadership standards and the LEA's coherent and integrated approach to strengthening the educator workforce; (2) provide for observations of each teacher or principal at least twice during the school year by individuals (who may include peer reviewers) who are provided specialized training; (3) incorporate the collection and evaluation of additional forms of evidence; and (4) ensure a high degree of inter-rater reliability... [4]

These four factors of a “[r]igorous, transparent, and fair” evaluation process require substantial infrastructure that is often not present in existing IT systems or actively supported by the local educational agency. Some of the concepts or practices (such as inter-rater reliability on observational measures) are quite foreign to district or school staff members.

The first requirement – a standards-aligned, objective, evidence-based rubric – is something that often does exist as a functional, embedded practice in many schools. There is a substantial research base around approaches to measuring teacher practice. There are several excellent pieces that provide a discussion of comprehensive approaches to teacher evaluation [10-12]. The Bill and Melinda Gates Foundation funded project *Measures of Effective Teaching* is in the midst of conducting a large-scale study of several popular evaluation frameworks (see [13] for a description of the launch of the project and [14] for methods of combining value-added outcomes with observational measures and student surveys).

What many larger districts find difficult about the implementation of this required element is the challenge of collecting, managing, and using this information at scale – across schools, taking adult mobility into account, shaping observation to reflect a teacher at different stages of his or her career, how frameworks should differ by subject and grade level, etc. The National Institute for Excellence in Teaching (NIET) has one of the longest track records of managing these complex data across multiple sites [11]. Their system – called the CODE (Comprehensive Online Data Entry) system - supports the collection and use of observational evaluation data, tracks inter-rater reliability of evaluators, and links the observational scores to intervention strategies aligned with each strand of the evaluation framework. Many TIF grantees have struggled to reinvent each aspect of such a system – from creating or adapting an evaluation framework to building and deploying the technical apparatus to collect and share the evaluation results. There is a clear market gap in this area. Many vendors are moving in to respond to this need, but the lack of expertise and clear requirements from the customer (schools, districts, and state agencies) suggests that vendors will likely have to go through a number of iterations to deliver tools at the level of sophistication provided by NIET to its partner schools.

2.4 Core Element D

A data-management system...that can link student achievement ... data to teacher and principal payroll and human resources systems [4]

Linking teachers to students by tested subjects has emerged as the most difficult technical challenge encountered across all TIF grants. The core of the problem is the misfit between the original requirements for student information systems and the new requirements measuring teacher performance using high stakes assessment results [15].

Student information systems (SIS) serve a number of functions. Scheduling has been a primary feature for many years. In the early grades, there is usually a room code and students are assigned to that room based on a number of considerations – classroom balance, number of students possible in that space, contract-based student teacher ratios, etc. In later grades, scheduling is a more complex practice that includes managing the mix of offered, required and elective courses in addition to the above requirements. Many state and district data systems in the U.S. revolve around a term that is used to describe the teacher primarily responsible for a particular classroom or course – “Teacher of Record” [16]. In the period before individual level teacher accountability policies, schools were not required to know exactly who the teacher of record for each room or course was. Scheduling typically happens before instruction begins in a future term (fall term is planned at the end of the spring term, etc.) and is refined once a new school year or term starts. It is not uncommon, therefore, for student scheduling systems to have teachers named “teacher one” or “department chair” in the scheduling system when the new school year starts. Many SISs are not linked to human resource systems directly and may never be updated with the identity of a new

teacher after scheduling is finished (see [15] for details on some of the challenges and potential solutions to this very common linkage problem).

Contrast this approach to a high stakes use of assessment data – such as calculating the value-added contribution of a given teacher to a classroom of students - that assumes a perfect match between teacher and assigned students – including tracking the mobility of students in and out of a particular classroom, formal or informal team teaching, the support of second language or special education aides, etc. There have been several high profile failures of the student teacher linkage systems that have led to errors in bonus pay-outs for TIF grantees. The most explosive to date was the incorrect assignment of student teacher linkages in Houston that led to 99 teachers receiving bonuses based on faulty data. Compounding that data quality error, the district leadership then asked all of the teachers who received the bonus to repay it [17]. This led to the admonition that was shared again and again at TIF meetings and in technical assistance sessions – “Don’t be Houston”. The case of Houston’s data quality failures made clear to most observers that 100% accuracy with 0% error is the only acceptable rate for such high stakes use of teacher effectiveness data.

This new requirement remains a challenge for most, if not all, shipping SIS products. Even with the most modern products that can support attribution of multiple adults providing support to each student in a particular subject, the social organization of work and the burden of keeping the technical system updated remains a challenge. In recent discussions with grantees, the conversation often revolves around how much accuracy is enough. When a student is supported with services from multiple adults in a particular curricular area, what is the smallest measurable unit that one needs to track? This remains an open question. System vendors will need to work with practitioners and evaluators to evaluate the payoff of the choices made about accuracy of complex service provision.

Another important scholarly debate emerged in 2009 that may have implications for linkage systems. Jessie Rothstein [18] argued that students were not assigned to classrooms at random. He showed that non-random assignment called into question the validity of estimates of teacher value-added scores, since students on different growth trajectories might be purposefully assigned to particular teachers. This would obscure the actual productivity of any given teacher. There was a rapid scholarly exchange [19] that showed that this problem could be addressed by using multiple prior test scores for students. One of the important issues highlighted by this debate is that students are not assigned to teachers at random and that these non-random assignments might actually produce optimal outcomes for students and teachers with complementary characteristics. Prospective (forward-looking) student-teacher assignment shares important characteristics with retrospective student-teacher linkage. One may be able to leverage the same system to support both the pro- and retrospective work.

2.5 Core Element E

A plan for ensuring that teachers and principals understand the specific measures of teacher and principal effectiveness included in the PBCS, and receive professional development that enables them to use data generated by these measures to improve their practice [4]

Finally, the TIF program requires that grantees use the results of teacher and principal effectiveness analysis to target the delivery of professional development resources based on those data. We have seen two systemic challenges associated with this requirement. First, while many schools and districts have previously implemented evaluation systems of some sort, the ability to break down evaluation strands into their sub-components and link those to specific professional development interventions is beyond the analytical capabilities or, at least, far beyond typical practice for most organizations. Second, TIF grantees have not previously tracked individual teacher professional development choices. This requirement to collect information on linked professional development could also mean that schools and districts must collect information of professional development chosen by teachers using their own funds. This substantially extends the requirements beyond simply tracking district- or school-level professional development offerings towards a more holistic understanding of how life-long-learning interacts with teacher and principal effectiveness.

3 Implications

The requirements of the Teacher Incentive Fund Core Elements have proven to be a challenge for many of the current round of grantees. Any actors trying to implement evaluation systems at a scale above individual schools have found every technical hurdle more or less difficult to overcome. The majority of the current Cohort is only now entering the implementation phase of their projects. While many of these educational organizations have piloted most of the infrastructure required to implement their evaluation frameworks, it is unlikely that all will succeed at scale.

It is also clear that most grantees have far to go to create meaningful links between teacher effectiveness measures and professional development resources that would be a good fit to address any identified challenges. Indeed, this approach seems to be a deficit-focused strategy. Another strategy, not present in any of the TIF requirements, would be to identify those practices associated with the most highly effective teachers and proactively promulgate those practices across all staff.

The successes and failures of the prior two TIF Cohorts were reflected in the Core Elements for Cohort 3. It is very likely that some of the challenges identified here will be taken into consideration when the U.S. Department of Education releases the new Request for Application for TIF Cohort 4 – likely to come in spring 2012. The story of TIF is an emerging one. It is likely to have a profound influence on the requirements for most core technical systems in the U.S. When combined with School Improvement and Race to the Top grants, the impact on the information infrastructure

for primary and secondary education and teacher preparation is likely to be far reaching and may lead to substantial shifts in who are important actors in this market.

References

1. English, F.W.: Merit pay: Reflections on education's lemon tree. *Educational Leadership*, 41 (4), 72, (1983)
2. U.S. Department of Education: Office of elementary and secondary education; overview information; teacher incentive fund; notice inviting applications for new awards for fiscal year (FY) 2006. Available from: <http://www.gpo.gov/fdsys/pkg/FR-2006-11-14/pdf/E6-19193.pdf>, (2006)
3. Teacher incentive fund: CFDA 84.374A project abstract. Available from: <http://www2.ed.gov/programs/teacherincentive/2007-awards.html>, (2010)
4. U.S. Department of Education: Office of elementary and secondary education: Overview information; teacher incentive fund: Notice inviting applications for new awards for fiscal year (FY) 2010. Available from: <http://www2.ed.gov/legislation/FedRegister/announcements/2010-2/052110c.html>, (2010)
5. Teacher incentive fund: FY 2010 funded applications. Available from: <http://www2.ed.gov/programs/teacherincentive/apps/index.html>, (2011)
6. U.S. Department of Education: Department begins competition for \$437 million in teacher incentive fund grants. Available from: <http://www2.ed.gov/news/pressreleases/2010/05/05202010.html>, (2010)
7. Teacher incentive fund: Main & evaluation competitions: Technical assistance workshop. Available from: <http://www2.ed.gov/programs/teacherincentive/dcta632010.pdf>, (2010)
8. Maricopa County Education Service Agency: Rewarding excellence in instruction and leadership. Available from: <http://www.maricopa.gov/Schools/service-home.aspx?sid=1>, (2012)
9. Hillsborough County Public Schools: Empowering effective teachers: Answers to your previous questions. Available from: http://communication.sdhc.k12.fl.us/empoweringteachers/?page_id=384, (2012)
10. Battelle for Kids: Why are some teachers more effective than others? The challenges and opportunities of defining "great" teaching. Battelle for Kids, Columbus, OH (2010)
11. Daley, G., Kim, L.: A teacher evaluation system that works. National Institute for Excellence in Teaching, Santa Monica, CA. Available from: http://www.tapsystem.org/publications/wp_eval.pdf, (2010)
12. Jerald, C.D., Van Hook, K.: More than measurement: The TAP system's lessons learned for designing better teacher evaluation systems. National Institute for Excellence in Teaching, Santa Monica, CA. Available from: http://www.tapsystem.org/publications/eval_lessons.pdf, (2011)
13. Bill & Melinda Gates Foundation: Foundation commits \$335 million to promote effective teaching and raise student achievement. Available from: <http://www.gatesfoundation.org/press-releases/Pages/intensive-partnership-for-effective-teaching-091119.aspx>, (2009)
14. Kane, T., Staiger, D.: Gathering feedback for teaching: Combining high-quality observations with student surveys and achievement gains. Bill and Melinda Gates Foundation, Seattle, WA (2012)

15. Thorn, C.A., Glover, R., Watson, J.: Information technology considerations guide to implementation: Resources for applied practice. U.S. Department of Education, Office of Elementary and Secondary Education, Washington, DC (2009)
16. Data Equality Campaign: Teacher of record definition framework. Available from: <http://dataqualitycampaign.org/resources/details/1047>, (2010)
17. Mellon, E.: 99 teachers told to repay part of bonus: HISD blames a computer glitch for overpayments. Houston Chronicle (TX) (March 9, 2007)
18. Rothstein, J.: Teacher quality in educational production: Tracking, decay, and student achievement. *The Quarterly Journal of Economics*, 125 (1), 175-214, (2010)
19. Koedel, C., Betts, J.R.: Does student sorting invalidate value-added models of teacher effectiveness? An extended analysis of the Rothstein critique. Department of Economics, University of Missouri-Columbia Report No.: WP 09-02, Columbia, MO (2009)