

# FairTest

---

## National Center for Fair & Open Testing

March 29, 2012

These comments go to the overall thrust of the Teacher Incentive Fund proposed “priorities, requirements, definitions, and selection criteria” (hereafter, “guidance”) and to Proposed Priorities 1 and 2. (TIF draft regs are at <http://www.regulations.gov/#!documentDetail;D=ED-2012-OESE-0001-0001>.)

The evidence shows that performance-based compensations systems (PBCS) do not lead to educational improvements, while they often create negative consequences (Adams, S.J., Heywood, J.S., and Rothstein, R., 2009; *Teachers, Performance Pay and Accountability: What Education Should Learn from Other Sectors*. Economic Policy Institute; Springer, M.G., et al., 2010, *Teacher Pay for Performance*, National Center on Performance Incentives). We recognize this is guidance for an “incentive” fund, but incentives need not require a PBCS. Incentives for teachers to gain more knowledge and skills through such things as becoming a mentor or a peer evaluator make sense, while the basic thrust of PBCS, particularly ones that include student test score gains, is controverted by the preponderance of evidence.

The National Center for Fair & Open Testing (FairTest) therefore recommends that the Department withdraw its current proposed guidance and fundamentally recast the basic thrust of the Teacher Incentive Fund.

FairTest is specifically concerned that the proposed TIF regulations will have perverse effects on teacher quality, instruction and student learning due to the mandatory incorporation of “student growth” measures, the most complex recent versions of which are the so-called “value added measures” (VAM), into teacher evaluation systems.

FairTest recommends that no use of “student growth” or VAM be required or privileged in the TIF or similar programs. Such use should be, at most, an option chosen by the grantee. Our brief rationale, with key references included, follows.

Value-added and student growth measures are based, in practice and in the draft regulations, entirely on traditional standardized tests. Under NCLB, “teaching to the test” has become a serious problem recognized by both President Obama and Secretary Duncan. Unfortunately, the use of VAM is likely to exacerbate teaching to the test. (These points are made in detail and with citations in “Flawed Massachusetts’ Teacher Evaluation Proposal Risks Further Damage to Teaching and Learning, by the Massachusetts Working Group on Teacher Evaluation of the National Center for Fair & Open Testing; [http://www.fairtest.org/sites/default/files/fairtest%20report%206611.indd\\_.pdf](http://www.fairtest.org/sites/default/files/fairtest%20report%206611.indd_.pdf).) In addition,

**P.O. Box 300204, Jamaica Plain, MA 02108**  
**fairtest@fairtest.org 617-477-9792 <http://fairtest.org>**

problems such as cheating are likely to intensify, just as the state and federal pressures to boost test scores have been cited explicitly as a cause of cheating in Atlanta by state investigators (<http://www.ajc.com/news/volume-1-of-special-1000798.html>).

Further, schools already push out children likely to score low (Advancement Project, et.al., 2010. *Federal Policy, ESEA Reauthorization, and the School-to-Prison Pipeline*. <http://www.fairtest.org/position-paper-nclb-and-school-prison-pipeline>). Pressure to do so will only increase if test-based VAM results become a “significant factor” in educator evaluations.

Additional problems, such as teacher reluctance to teach “difficult” students, are likely to increase. That category includes children with disabilities, English language learners, racial minorities and students from low-income communities. Indeed, teachers are likely to seek to leave schools with large numbers of any such students. This is because, despite claims to the contrary, VAM cannot fully control for issues such as the systemic influence of family and community poverty and resulting “peer influence” that affects classrooms. Absent random assignment of students—which is impossible as it would have to be done across districts to address poverty and other factors known to heavily influence educational performance --VAM will remain biased, and teachers know it. All these factors will increase already near-record rates of teacher dissatisfaction (see Scholastic and the Gates Foundation, *Primary Sources 2012: America’s Teachers on the Teaching Profession*). They will further reduce morale and lead to many good teachers leaving for fear of the consequences of “evaluations” via VAM or student growth that are close to random in results (see “Flawed Massachusetts’ Teacher Evaluation Proposal Risks Further Damage to Teaching and Learning,” cited above).

The underlying presumption is that the use of VAM or “student growth” measures will lead to improved student learning. A recent study by Chetty, Friedman and Rockoff ([http://obs.rc.fas.harvard.edu/chetty/value\\_added.pdf](http://obs.rc.fas.harvard.edu/chetty/value_added.pdf)) is frequently relied upon to support this claim. However, this study uses data from years before high-stakes testing corrupted test results (Koretz, D. *Measuring Up*. 2008, Harvard University Press). The authors acknowledge that teaching to the test and cheating are likely to inflate test scores and reduce the results they found:

“An important limitation of our analysis is that teachers were not incentivized based on test scores in the school district and time period we study. The signal content of value-added might be lower when it is used to evaluate teachers because of behavioral responses such as cheating or teaching to the test.”

They further found that “the top 2% of teachers ranked by VA have patterns of test score gains that are consistent with test manipulation... Correspondingly, these high VA outlier teachers also have much smaller long-term impacts than one would predict based on their VA.” There is every reason to believe that the Chetty, et al., findings will largely evaporate in the face of teaching to the test, removal of low-scorers, and cheating.

In light of the strong preponderance of the evidence, many organizations and experts have expressed great caution regarding the use of VAM as mandated and “significant” parts of teacher or principal evaluation. The Forum on Educational Accountability (FEA), for example, has expressed its concerns, on behalf of many groups, in a public letter to Congress (at [http://www.edaccountability.org/What\\_Should\\_Congress\\_Do\\_About\\_Teacher\\_Evaluation-FEA\\_letter\\_9-20-11.pdf](http://www.edaccountability.org/What_Should_Congress_Do_About_Teacher_Evaluation-FEA_letter_9-20-11.pdf)). In addition to expressing concerns about VAM, FEA also points out,

“There is not enough evidence that evaluation systems are ready for large-scale state use or that they are a more important use of school improvement funds relative to other efforts.” There is an absence of evidence that spending funds to develop extensive and expensive teacher evaluation systems is a more effective use of limited funds than would be other uses. The Department can establish incentives for helpful teacher development without requiring the creation of evaluation systems that are likely to be unsustainable in the absence of continuing federal resources. FairTest would support the development of teacher evaluation systems as a possible use of TIF funds (absent mandates to use student growth measures).

Monty Neill  
Executive Director  
monty@fairtest.org