

11

Rewarding
Teachers*Caroline M. Hoxby**Eric A. Hanushek*

Almost everyone agrees that teachers are the key factor in whether a child learns at school. While a teacher cannot control who enters her classroom—children arrive with a wide variety of attainments and family resources—she can have a powerful effect on the gains a child makes while in her classroom. Scientific measurement of the gains that a teacher produces is the essence of value-added calculations, described in the earlier chapter on the subject. The beauty of value-added calculations is that they permit a state like Arkansas to reward teachers directly for raising pupil achievement. The teacher reward schemes discussed in this section are founded on value-added calculations. Their key virtue is that they can provide the vast majority of teachers in a state with consistent, powerful incentives to raise achievement. Yet, the schemes are also flexible: they can be designed to include school-wide incentives, instructional team incentives (an instructional team is, for instance, all fourth grade teachers in a school), incentives to teach in schools with particularly low-achieving children, and incentives to teach classes that are particularly hard to staff.

In the schemes described here, a substantial portion of a teacher's salary would, after a phase-in period, reflect her contribution to student achievement. By creating rewards for people who are talented and effective as teachers, the incentive systems described here would not only raise performance among the state's current pool of teachers but would also provide strong rewards for successful people to join and remain in the teaching profession.

This chapter describes how a teacher incentive scheme based on value-added calculations could be implemented in practice.

Many "merit pay" plans have been introduced previously around the nation. These have generally been very limited and have not relied on objective performance information. Experience with such plans has erroneously suggested that incentive plans will not work in schools. We begin with a series of frequently asked questions, many of which arise from an imperfect understanding of how incentives based on a value-added calculation would work. We describe how a system of incentives based on teachers' value-added can be refined to create a well-rounded system that achieves Arkansas' broad goals for student achievement.

Frequently Asked Questions

Do teachers really differ significantly in their value-added?

Yes. For example, the top 10 percent of teachers raise their students' achievement by about half a standard deviation more *each year* than do the lowest-performing teachers. That is roughly equivalent to half a grade's worth of learning. It is also roughly equivalent to half the difference between the achievement of the typical American child and that of a seriously disadvantaged child.

Aren't most of the differences in teachers' value-added really a

Rewarding Teachers**157**

reflection of getting a good or bad class? Isn't it true that, just by luck, some classes are dominated by fast learners and others by slow learners?

Value-added calculations are designed to eliminate the confusion between achievement that is attributable to a student's current teacher and achievement that reflects a student's home environment, his previous teachers and schools, and his peers. Still, value-added calculations are better and more reliable when they are based on a few years of achievement data rather than a single year. That's why a good reward system can give especially substantial awards to teachers with at least four years of experience.

If calculations of teachers' value-added are not really reliable until a teacher has been teaching for a few years, should beginning teachers be included in a reward system based on value-added?

It is true that value-added calculations are not terribly reliable for beginning teachers, such as those for whom we have only one or two years of data.

There is an additional reason why beginning teachers should be treated differently than experienced teachers. The typical teacher actually does improve her teaching (her value-added) substantially over her first few years of teaching. By year three or four, a teacher's value-added tends to settle down in a range if she is not given incentives to make further improvements.

On the whole, a teacher's beginning years should probably be focused on developing skills with the help of her principal, mentors, and educational institutions. If she succeeds in acquiring and practicing skills, she can look forward to rewards based on her high value-added once she has some experience.

Won't a reward system based on value-added cause teachers in the same school to undermine one another, each teacher seeking to increase her reward by decreasing those of others?

In a statewide reward system based on value-added calculations, it really does not make sense for a teacher to do anything that would undermine other teachers in her school. A teacher's value-added will be measured relative to a fixed standard that is set statewide. A reward system based on value-added can be based on absolute rewards (so much reward for every point of value-added); it need not be based on *relative* rewards (rewards for, say, only the top twenty percent of teachers even if all teachers attained higher gains).¹

In order to create incentives for whole schools to “pull together” to raise student achievement, the state can base part of a teacher's reward on the average value-added attained by teachers in her school. Similarly, to give incentives for all teachers in a grade or subject to “pull together,” part of a teacher's reward can be based on the average value-added attained by teachers in her instructional team, which might be fourth grade teachers or high school English teachers.

In a recent experiment in Little Rock's Meadowcliff School, teachers were rewarded based on their individual value-added. They appear to have been overwhelmingly supportive of rewards based on student gains, objective measures of performance, and individual teachers' contributions.

Why not just have all rewards based on school-wide performance or group-wide performance?

The data show that many teachers whose own value-added is high teach in schools where the other teachers' value-added is low. Thus, a teacher with consistently high value-added might never be rewarded if rewards are based entirely on school-wide

1. Hanley Chiang and Caroline M. Hoxby, “From Teacher Effects to Teacher Rewards: The Empirics of Computing and Rewarding Teachers' Contributions to Student Achievement” (working paper, NBER, 2005). This paper provides an explanation of how a fixed standard system can be set up.

Rewarding Teachers

159

performance. We should not discourage high value-added teachers from working in underperforming schools where they are needed as exemplars and mentors. Instead, they should be encouraged to seek jobs in schools where their talents will be most useful. A wise system of rewards balances incentives to have schools “pull together” with incentives for individual teachers to work in difficult environments.

How can a reward system based on teachers’ value-added offer absolute, not relative, rewards? In other words, can such a system be a “rising tide” that lifts all teachers’ pay?

If the state will commit that the pool of rewards will grow with the average value-added in the state, a reward system based on value-added can be a “rising tide.” Voters are wary of raising teachers’ pay with no strings attached, so it is hard for teachers to attain higher pay within the current structure, where salaries are based purely on seniority and credentials. Thus, having a pool of rewards that scales up with value-added may be one of the only ways for the teaching workforce in an entire state to improve as a whole, both in performance and compensation.

Don’t we already have reward systems for teachers in Arkansas, such as the teacher-of-the-year award, the federal program that recognizes great math and science teachers, and the Teacher Advancement Program?

The teacher-of-the-year and the great math and science teacher programs do a good job at focusing attention on teachers whose pedagogical and leadership skills are remarkable. These programs not only can, but should, run alongside a program of teacher incentives based on value-added. However, they are complements to a value-added program, not a substitute for one.

The teacher-of-the-year and great math and science teacher programs work through a combination of nomination and demonstration. A committee examines recommendations by princi-

pals, peers, or parents, as well as demonstrations of teaching (portfolios of teaching materials, videotapes, some student results). Thus, these programs are inherently not scalable. It would be impossible for committees to do a good job of evaluating all Arkansas teachers each year by these methods. In contrast, a system of rewards based on value-added information can include the vast majority of teachers.

The programs mentioned identify exemplary teachers whom others can imitate. But identifying an exemplar is inherently something done once, not again and again. For instance, no one would recommend that the same teacher be picked as the teacher-of-the-year, year after year, even if one teacher deserved that status. Thus, the bonus for being recognized by one of the programs mentioned above is typically a one-time event, not a way for a valuable teacher to experience higher pay year after year. By contrast, an incentive system based on teachers' value-added can create higher career pay for successful teachers.

In contrast, the Teacher Advancement Program *is* a model of regular rewards for teachers that includes significant compensation based on teachers' and schools' value-added. The Teacher Advancement Program is close to the model recommended here. However, it is currently only being implemented in five Arkansas districts; it is not a statewide system. A statewide system offers much larger promise for student achievement because it can change the entire teaching profession in Arkansas.

Why have rewards based on value-added when teachers can be rewarded for attaining higher credentials, like a master's degree. In Arkansas, teachers already qualify for annual bonuses if they are National Board certified. Why not just expand that program?

States can and do create programs that reward teachers based on *credentials*. An additional salary step for a master's degree is one obvious example. A salary step for National Board Certifi-

Rewarding Teachers

161

cation (which is based on a teacher's exhibiting her preparation in her subject area and her teaching materials) is another example. Rewards that are based on credentials are based on the supposition that, in the process of attaining a credential, a teacher acquires subject area knowledge or pedagogical skills that make her better at her job. However, research does not support the supposition that teachers with master's degrees or more hours of professional development actually contribute more to student learning. Systems of incentives based on value-added directly reward teachers for *outcomes* and thus avoid the question of how credentials translate into classroom results.

What if a school's students have such high levels of achievement that it is hard for teachers to produce substantial gains in them, at least as shown on statewide exams?

In the long run, statewide exams should be designed so that students at every level of incoming achievement are suitably challenged and can demonstrate substantial gains. However, it is difficult to design pencil-and-paper exams that fulfill such goals; to a large extent, such design requires computer-assisted technology whereby a student who does well on early questions can be presented with more and more difficult questions.

Given the exams that exist today in Arkansas and many other places, it makes sense for a state to base a teacher's rewards partly on the basis of her value-added and partly on the basis of her students attaining high absolute levels of achievement. With such a combined system, teachers won't be penalized because they teach students for whom the statewide exams are not terribly challenging.

Can a reward system place extra emphasis on achievement gains among students whose initial performance is low?

Some people argue that the state has a greater interest in raising achievement from low levels to proficiency, than in raising

proficient achievement to advanced levels, and this contention is reinforced by the federal No Child Left Behind Act. These arguments are based on the idea that a state's citizens cannot participate effectively in democracy, the workforce, or society if they do not achieve a certain level of competence in core skills.

In fact, most statewide exams, including Arkansas', already partly incorporate this concern, because they are designed to be especially informative about whether a student is performing at a level that is "Failing," "Basic," or "Proficient." Most statewide exams are less informative about a student's degree of excellence beyond proficiency. If a state wants to provide added recognition for teachers who raise achievement among students whose initial achievement is low, the incentive system can incorporate a "multiplier" for value-added that is higher for teachers who work in schools that serve such students. For instance, a point of value-added could earn 1.25 times more of a reward if the students were initially scoring at the "Below Basic" level than if the students were initially scoring at the "Proficient" level.

Is it best for a reward system to be based purely on quantitative measures of a teacher's contribution or should rewards be based on qualitative information, too?

A good reward system will incorporate qualitative information about a teacher's performance, such as administrator evaluations and parent evaluations. By incorporating such evaluations, a teacher's whole contribution, on many dimensions, can be considered. However, if a reward system puts too much weight on such evaluations, it is likely to be viewed as subjective, prey to favoritism, or manipulation. Therefore, the majority of a teacher's reward should be based on quantitative information about student performance (mainly value-added but probably also rewards for a high absolute level of performance, as mentioned above).

In the chapter in this volume on value-added measures, we

Rewarding Teachers**163**

noted that it was possible to make value-added calculations for a variety of annual outcomes, not just test scores. Thus, a reward system could incorporate quantitative information on attendance, disciplinary infractions, and so on.

How can teachers who teach classes in music, art, foreign language or other non-tested subjects be rewarded?

Such teachers' rewards must depend on group-wide or school-wide incentives, individual teacher performance as measured by student outcomes other than statewide exams, and administrator and parent evaluations. If Arkansas or some of its districts develop good evaluation metrics for such teachers, it will be more possible to give them strong individual incentives. Arkansas may also wish to employ national "off-the-shelf" tests in subjects that the state does not presently test.

Need Arkansas employ a single teacher reward system for the whole state?

Arkansas should make value-added calculations at the state level, and should, at minimum, impose certain constraints on the reward plans that districts adopt. For instance, a good constraint is that more than half of a teacher's reward should be based on her individual value-added, assuming that she is a teacher for whom reliable value-added calculations are available (for instance, a teacher with four or more years of experience who teaches an academic subject). Arkansas may also wish to impose rules on how the reward pool is allocated among schools where students are initially low-, medium-, or high-achieving. For instance, the "multiplier" noted above should be statewide.

The benefit of having a single statewide system is that teachers will learn it quickly and it will be "portable" if teachers move within the state. The benefit of allowing districts to have some latitude in the design of their reward systems is that districts with particularly good designs may benefit (by drawing good teachers)

and may therefore spur innovation, eventually moving the state as a whole toward the better designs.

Recommendations

1. *Any new state funding that can be allocated to teacher pay should be used to fund the reward scheme.*

Rewards for teachers are likely to improve achievement, but they would be difficult to implement in a revenue-neutral manner in the short run, assuming that districts are held harmless on total state aid. In the longer run, the reward scheme can be made more revenue-neutral by allowing it to grow relative to other forms of state aid and teacher pay.

2. *Individual rewards must be salient for good teachers.*

Rewards should be sufficiently large to attract and retain outstanding teachers—undoubtedly in the thousands of dollars as opposed to hundreds of dollars. The purpose of a reward system is to draw successful teachers into schools and keep them there. Small rewards for many teachers may seem “fair” and “fiscally sound.” Large rewards given to few may draw attention to exemplary teachers. But neither can transform teaching into a profession that competes successfully with other careers that reward better performance with higher career pay.

3. *The reward scheme should directly incorporate quantitative information about student performance wherever practical and should base the majority of a teacher’s reward on her individual value-added.*

Rewards should be based mainly on what an individual teacher can affect—the educational value that she adds to her students. Although rewards cannot depend primarily on individual value-added for beginning teachers and those who

Rewarding Teachers**165**

teach non-tested subjects, their rewards should, to the extent possible, be based on quantitative information on student outcomes, not paper credentials, preparation or subjective judgment.

4. *Quantitative assessments of value-added for individual teachers should incorporate several years of performance.*

Some variation in value-added calculations occurs simply because of fluctuations with small numbers of students. In order to have a reliable assessment of teacher performance, the quantitative portion of value-added calculations should rely on a moving window of student performance over the previous four years. This approach eliminates large annual swings in teacher incentive pay and emphasizes consistent performance. It also automatically avoids placing undue emphasis on the first years of performance for new teachers.

Teachers without four years of classroom experience or those in subject areas where students are not regularly tested should have their value-added assessed by school leaders. This assessment should emphasize quantitative information where feasible and should focus, as in tested subjects, on the performance gains attributable to each teacher.

5. *Arkansas should balance statewide and district-specific considerations.*

Many options exist for the design of the precise incentive system, and there are advantages to permitting local variation in the design. A sensible approach would have the state prescribe the essential elements of the plan and create a “default plan” for districts that want to adopt a conventional plan that will be understood quickly by its teachers. Individual districts would, however, have the discretion to modify the default plan (within the prescribed limits set by the state) in order to make it fit their needs. An example of a prescribed limit is a

minimum on the extent to which a teacher's reward must depend on state-computed value-added when it is available. Districts could use their discretion to develop their own weightings of performance information (consistent with state minimums) and mechanisms for using information of a qualitative nature.

Other aspects of the reward plan that should be statewide are the multipliers, if any, designed to give teachers an especially strong incentive to raise achievement in schools where students' incoming scores are low, in geographic areas where vacancies typically occur, or in subject areas that face critical shortages. Statewide multipliers may also exist at the other end of the spectrum to ensure that teachers want to generate true excellence in students whose scores put them near the "ceiling" of the Arkansas state tests.

6. *Evaluation.*

To ensure that the state and others learn from the state's reward scheme, data should be gathered that facilitate evaluation by external researchers. Those data should include information on the achievement levels and value-added, rewards earned, and districts' plans for individual rewards.