Paper Title What Impact Have Accountability Policies and Practices Had on the Retention of Teachers?

Author(s) Richard Ingersoll, University of Pennsylvania; Lisa Merrill, University of Pennsylvania; Henry May, University of Delaware

Session Title Hiring and Retaining Teachers

Session Type Paper

Presentation Date 5/1/2013

Presentation Location San Francisco, California

Descriptors Accountability, School Reform, Teacher Evaluation

Methodology Quantitative

Unit Division L - Educational Policy and Politics

-- Preliminary Draft --

What Impact Have Accountability Policies and Practices Had on the Retention of Mathematics and Science Teachers?[1]

By

Richard Ingersoll, Lisa Merrill and Henry May

of the

University of Pennsylvania

and the

Consortium for Policy Research in Education

[1] This research was supported by a grant (# 0814295) from the Research and Evaluation on Education in Science and Engineering (REESE) program of the National Science Foundation. Lisa Merrill's work was also supported, in part, by the Institute of Education Sciences, U.S. Department of Education, through Grant #R305B090015 to the University of Pennsylvania. Opinions in this paper reflect those of the authors and do not necessarily reflect those of the granting agencies. This paper will be presented at the 2013 Annual Meeting of the American Educational Research Association, in San Francisco.

Perhaps the most controversial and significant of contemporary education reforms has been the school accountability movement. This genre of reform has been initiated under a variety of labels, such as Standards-Based Reform and Performance-Based Accountability. But, regardless of the label, these reforms center on a series of common steps: the setting of performance standards for students and schools, the use of assessments, typically students' standardized test scores, to gauge student and school performance in regard to the standards, and the use of various combinations of incentives and sanctions to induce schools to improve their performance on various outcomes (see figure 1). Many states began to initiate such reforms throughout the 1990s, and in January 2002 accountability gained major impetus as a nation-wide reform with the advent of NCLB.

This study examines the impact of accountability on the ability of schools to retain qualified mathematics and science teachers. What impact have school accountability provisions, such as those mandated by NCLB, had on the ability of schools to staff mathematics and science courses with qualified teachers, as also mandated by NCLB? On the one hand, increases in assessment and scrutiny could apply new or undue pressure and tension, especially in lower performing schools, and hence, make it more difficult for schools to retain their mathematics and science teachers. On the other hand, an increased focus on performance, could lead to improved school leadership, management and school conditions, in turn fostering teachers' capacity to meet the standards, hence, make it easier for schools to keep qualified mathematics and science teachers? Or perhaps, the impacts of accountability and testing on teacher retention depend on the existing conditions and management in schools and, hence, how well the reforms are implemented? Answers to these questions have large and important implications for the success or failure of current testing reforms and the design of future federal and state accountability policy. They also have large implications for understanding how policy can assist schools in better ensuring that all classrooms are staffed with qualified mathematics and science teachers. The overall objective of this project is to enhance such knowledge.

Figure 1: The Theory of Educational Accountability

Set Performance Standards	→	Assess Performance on Standards	→	Pass Standards – Rewards Fail Standards – Sanctions	→	Improve Performance on Standards

Our data source is the National Center for Education Statistics' nationally representative Schools and Staffing Survey, along with its' supplement, the Teacher Follow-up Survey. SASS/TFS is the largest and most comprehensive data source available on elementary and secondary teachers and schools. We utilize data on public schools from three separate cycles of SASS/TFS representing the period from 1999 to 2009. We focus in particular on the 2003-04 SASS and 2004-05 TFS – data collected two-three years after the advent of NCLB. This cycle of SASS/TFS obtained an unusually rich set of items on a wide array of aspects, steps and components of accountability reform, from the perspective of both school leadership and school

faculty. 2003-04 is also a useful point at which to examine accountability because while these reforms had become widespread, they had only recently been nationally mandated, and hence, we would expect large variations in design and implementation across design and implementation.

In the middle of the 2003-04 school year, SASS asked a national sample of school-level administrators if, in the prior school year (2002-03), their school had been subject to school performance standards established by their district or state, whether their school had been subject to evaluations assessing their performance in regard to the standards, and how their school fared on the assessments. These administrators were then asked whether their school subsequently, in the current 2003-04 school year, received rewards, incentives, penalties or sanctions as a result of the school's performance. Subsequently, the TFS obtained data on which teachers, from the original 2003-04 SASS teacher sample, stayed in, or departed from, their schools, or from teaching altogether, by the following year – 2004-05. Hence, the 2003-05 SASS/TFS provides a clear timeline of the steps in Figure 1: schools' standards set and performance assessed in 2002-03; rewards or sanctions subsequently applied to schools in 2003-04; teacher retention or turnover between 2003-04 and 2004-05.

We especially focus on qualified math and science teachers, which we define as those with an undergraduate or graduate degree in math, in one of the sciences, or in related fields, such as engineering, math education, or science education. We do not count as qualified those who as assigned to teach math or science courses, or those with a certificate in math or science, absent having a degree in the field. We chose a major-based method of identification because it represents those teachers with a credential signifying human capital in the field – the subject of major policy concern. Hence, in our discussion to follow, the terms "math teacher" or "science teacher" refers to those with degrees in the field. Note that while we do not include measures of the performance or effectiveness of math and science teachers, in our analyses we do use a proxy measure of teachers' academic ability – the selectivity ranking of their undergraduate college or university.

Using these data, we seek to answer three sets of research questions:

1.) Levels of School Accountability

How many public schools have been subject to state or district mandated performance standards, and evaluations of their performance in regard to the standards? Has this changed over time before and after the advent of NCLB? How many schools passed, or did not pass, their evaluations in 2002-03, and moreover, how many schools subsequently received rewards or sanctions in 2003-04?

Moreover, what impact have these accountability reforms had teachers? For how many teachers are their students required to take state or district assessments in the teacher's subject? How many public school teachers have access to their students' scores on state or district achievement tests? How many teachers use state or district standards to guide their teaching, use their students' scores to adjust their classroom curricula in areas where their students had encountered problems, or use their students' scores to assess weaknesses in their own content

knowledge or teaching practices? Are teachers worried about their job security because of the performance of their students on state or local tests? What has been the impact of state or district standards on teachers' job satisfaction?

2.) The Effects of Accountability on Teacher Turnover

We then use multilevel logistic regression analysis to examine the effects of accountability reform on the likelihood that teachers subsequently stayed in or departed from their schools, or from teaching altogether. Did any of the steps of accountability illustrated in Figure 1 – setting of standards, the use of state or district assessments, how well their school fared, and the application of any subsequent incentives or sanctions – have an impact on the subsequent retention or turnover of teachers?

3.) The Role of School Organization in the Impact of Accountability on Teacher Turnover

Our expectation is that the effects of school accountability on teacher retention will be mixed and that these effects will depend on how these reforms and initiatives are implemented at the school level. Accountability policies do not necessarily mandate specific practices and, theoretically, can leave the leadership and teaching staff in schools with a considerable amount of discretion in determining how to meet performance targets. Our hypothesis is that such organizational effects will be both direct and indirect. Increases in accountability and performance pressures at the school level could directly shape teachers' decisions whether to continue in a particular school or to continue in teaching. But, accountability's effects on teacher retention could also depend on the school's organizational conditions, leadership, and resources – all of which could impact the school's capacity to meet the standards. Do school conditions long associated with the effectiveness of schools – such as the principal's management style, the amount of classroom resources and support provided to teachers, and the degree of autonomy teachers' have in their classrooms – shape the impact of accountability reforms on outcomes, such as teacher retention?

For instance, in prior research, we found that the degree of classroom teacher autonomy, and amount of the collective school wide faculty influence over decision making, have a strong impact on levels of teacher turnover, especially for math teachers (Ingersoll & Perda 2010; Ingersoll 2011; Ingersoll & May 2012). Hence, we expect that if accountability reforms lead to, or are accompanied by, constraints on individual teacher autonomy in classrooms, this in turn could lead to increases in turnover. On the other hand, while held accountable, if teachers are also provided with the autonomy in their classrooms they feel they need to meet the standards, this in turn could lead to decreases in turnover.

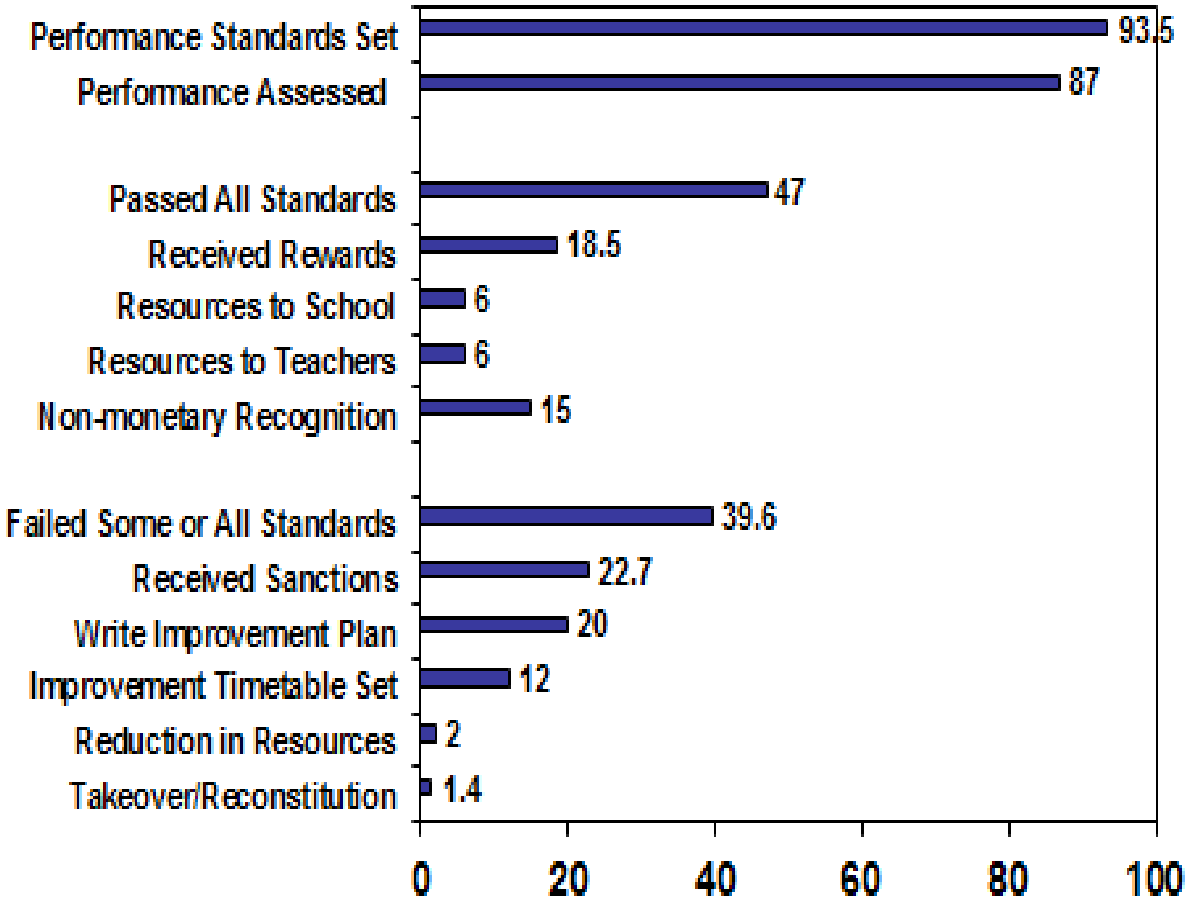
Results

Levels of School Accountability

Performance-based school accountability is a widespread reform. Even by the 1999-2000 school year (two years prior to the advent of NCLB in January 2002), 90 percent of public

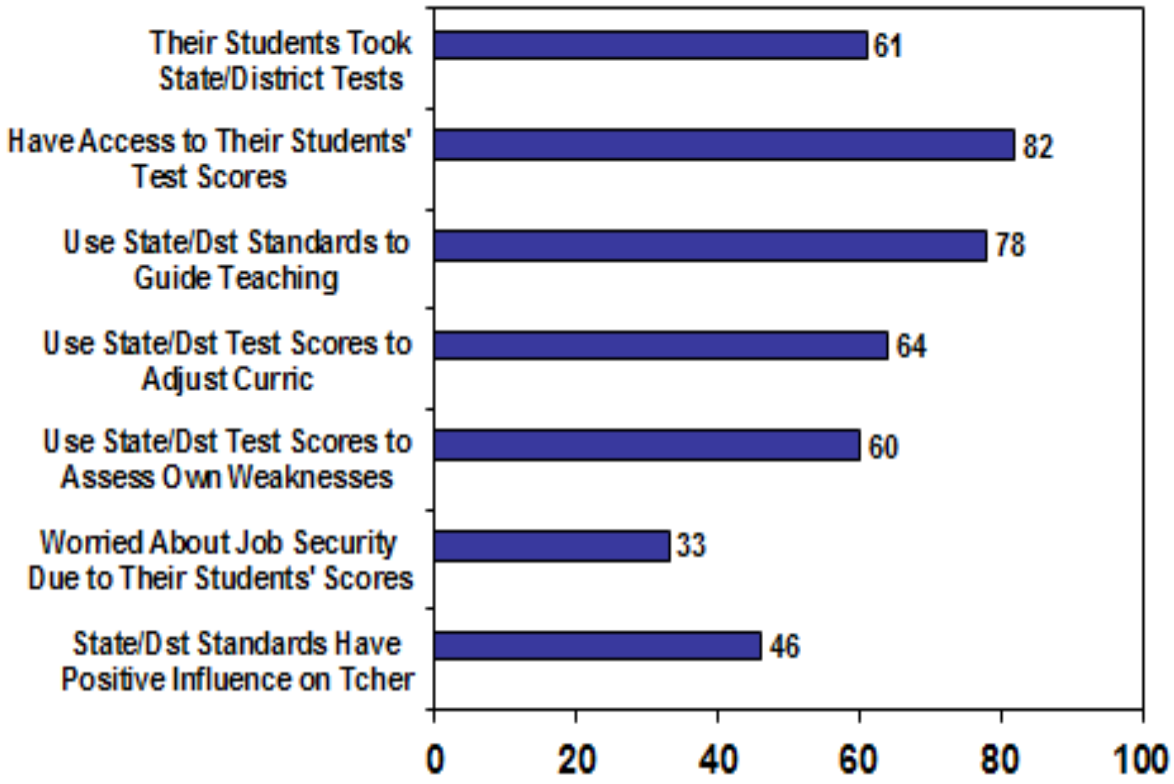
schools were subject to academic performance standards established by their district or state. By 2002-03, a year into NCLB, this had increased to 94 percent of all public schools and by 2007-08 virtually all public schools were subject to such standards. Setting standards is, of course only a first step in the theory of school accountability, as illustrated in figure 1. The data indicate that, of those schools with performance standards, almost all were also subject to assessments in regard to those standards (see Figure 2). Of those schools that were evaluated, in the 2002-03 school year over half (54%) passed all of the standards. Of those schools that passed all of the standards, a smaller portion – 39 percent – subsequently in 2003-04 received at least one of the following rewards or incentives, i.e. "carrots": non-monetary recognition; cash bonuses to the school/additional resources to support school-wide activities; or cash bonuses to teachers.

Figure 2: Percent Public Schools Subject to Different Accountability Steps, 2002-04



Of the 46 percent of evaluated schools that did not pass some, or all, of the performance standards in 2003-03, just over half were subsequently subject to at least one of four kinds of penalties or sanctions, i.e. "sticks." The sanctions measured varied in their degree of specificity, seriousness and frequency. They tended to be progressive and cumulative i.e. as the number and severity of sanctions received went up, the number of schools receiving them went down. The most common result of not meeting some, or all, of the standards was to be required to write a school improvement plan. Next most frequent was a tighter sanction – for the school to be put on an evaluation cycle with required improvements by specific dates. Far less common were more punitive consequences such as a reduction in resources to the school, or the school being subject to takeover or reconstitution of administration and teaching staff.

Figure 3: Percent Public School Teachers Impacted by Accountability Steps, 2003-04



Moreover, not only is accountability widespread, there is no question these reforms have had an impact on teachers (see Figure 3). Of course, not all fields and subjects, and at all grade levels, are required by the state or district to be tested. In 2003-04, 61 percent of teachers reported that the students they taught participated that year in required state or district assessments in the subject they taught. But regardless of whether a teacher's own students are tested in that teacher's subject, the fact and use of test scores have become ubiquitous parts of life in schools across all grades and all subjects. As early as the 2003-04 school year, 82 percent of all public school teachers reported that they had access to their students' scores on state or district achievement tests. Seventy-eight percent indicated that they used state or district standards to guide their teaching, 64 percent said that, additionally, they used their students' scores to adjust their classroom curricula in areas where their students had encountered problems, and 60 percent of all public school teachers reported they used their students' scores to assess weaknesses in their own content knowledge or teaching practices. Moreover, from an accountability theory perspective, this impact has appeared to have some "bite." By 2003-04, only one year into NCLB, fully one third of all public teachers, reported they were "somewhat" or "strongly" worried about their job security because of the performance of their students on state or local tests. On the other hand, from the perspective of teachers and their job satisfaction, this impact has not always been viewed as positive or benign. Indeed, in 2003-04 less than half (46 %) reported that state or district standards had a positive influence on their satisfaction with teaching.

Math and science teachers differed from other teachers in notable ways in regard to accountability and testing. Not surprisingly, math is one of the most frequently tested subjects. In addition, compared to others, math teachers were also more likely to have access to their student test scores, and slightly more likely to use state or district standards to guide their teaching. Science teachers also differed in these ways, but less so than math teachers. Moreover, both math and science teachers were less likely than others to report that state and district standards had a positive impact on their satisfaction with teaching. In 2003-04, almost two thirds reported that state or district standards had not had a positive influence on their job satisfaction.

Our second research question arises here: Did any of these steps and components of accountability – the setting of standards, the use of state or district assessments, how well their school fared, and the application of any subsequent incentives or sanctions – have an impact on the retention or turnover of teachers?

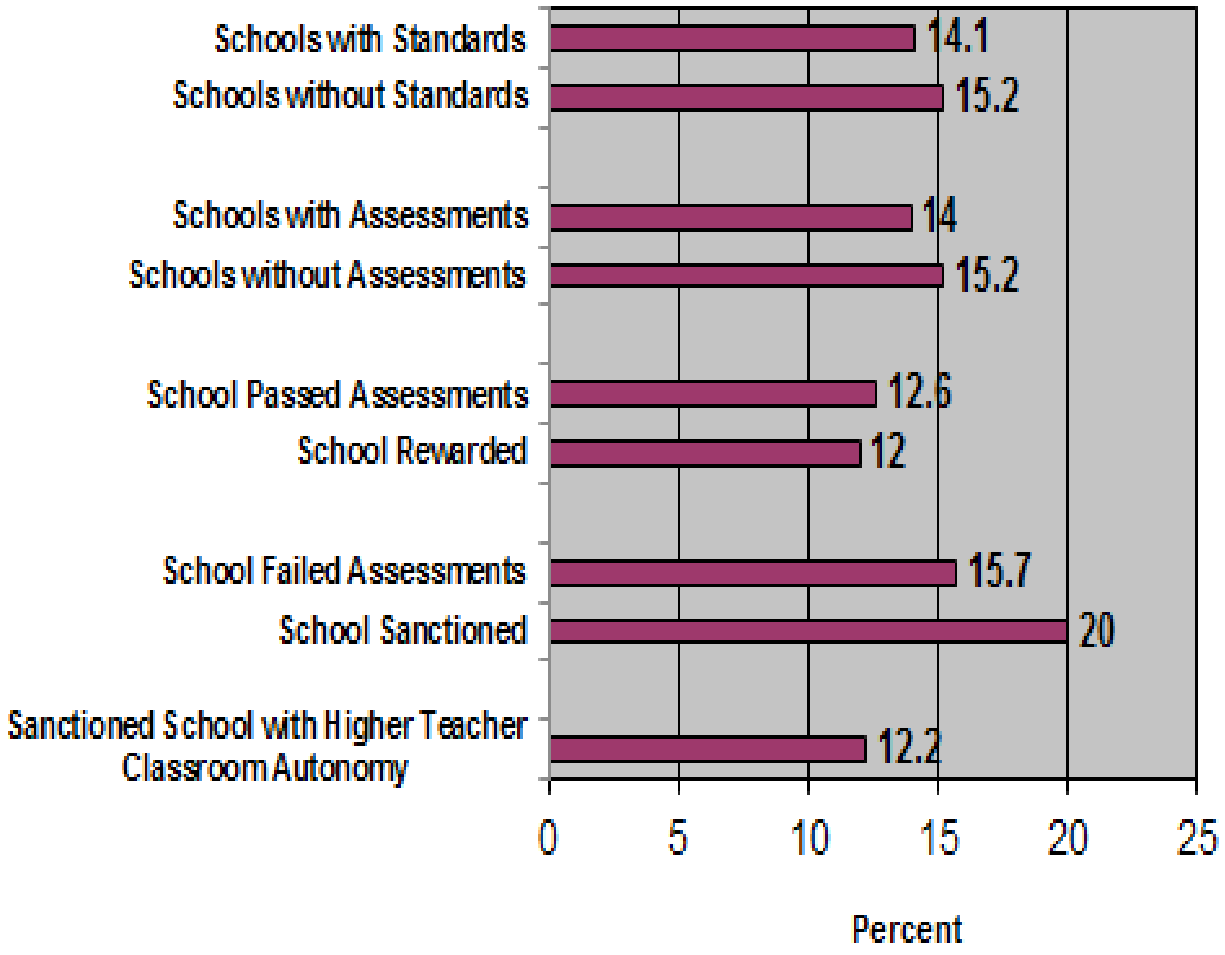
The Effects of Accountability on Teacher Turnover

Our multilevel multiple regression analyses also show that, after controlling for the background characteristics of both the teachers (gender, teaching experience, race-ethnicity, age, the selectivity of their undergraduate college or university) and their schools (poverty, size, urbanicity, level), some of the steps and components in school accountability theory were significantly associated with teacher retention and turnover, while some were not.

Once we controlled for background characteristics, teachers in schools for which performance standards had been established did not significantly differ in their likelihood of

departing than the small minority of teachers employed in schools without standards (7% of all public schools) (see Figure 4). The same held for those schools that both had performance standards set up, and were also assessed in regard to those standards. Teachers in those schools that were assessed in regard to district or state standards were not significantly more, or less, likely to depart, than were teachers who were employed in schools without state or district assessments (another 7% of all public schools).

Figure 4: Predicted Probabilities of Teacher Turnover, by Different Accountability Steps: 2004-05



In contrast, how well a school fared on the performance assessments was strongly related to the turnover of their teachers. Not surprisingly, successful schools – that passed all of the state/district assessments – had significantly less turnover, after controlling for the background characteristics of both the teachers and their schools. The opposite held for unsuccessful schools – those that did not pass some, or all, of the standard evaluations. Teachers in the lower-performing schools had significantly higher turnover.

Furthermore, the subsequent consequences resulting from a schools' performance were also often related to teacher turnover. But, this depended on whether a school received rewards for their higher performance, or sanctions for their lower performance. Among schools that passed all of their performance assessments, only one of the three kinds of "carrots" – cash bonuses/additional resources to support school-wide activities – was individually related to decreases in turnover and this effect was only of borderline statistical significance (p-value = 89%).

In contrast, among schools that failed some or all of the state/district assessments, those receiving some kind of penalty or sanction – "sticks" – had distinctly higher turnover than other lower-performing schools that received no sanctions. As mentioned above, we examined four types of school-level sanctions: required to write a school improvement plan; put on an evaluation cycle and timetable with required improvements by specific dates; penalized with a reduction in resources to the school; subject to takeover or reconstitution regulations. Our analyses show that of the four types of sanctions, being put on an evaluation cycle/timetable had the most consistent and strongest association with increased teacher turnover.

These sanctions did not necessarily exist in isolation as independent variables; schools often received more than one. To capture the cumulative effects of receiving different sets of sanctions, we created measures that combined these sanctions into "packages." We limited our analysis to six possible packages:

1.) No Sanctions: low-performing schools that were not subject to any of the four sanctions: not required to both write a school improvement plan, nor to undertake an evaluation cycle with required improvements by specific dates, nor subject to takeover or reconstitution, nor penalized with a reduction in resources. This non-package represented 41 percent of the teachers employed in low-performing public schools.

2.) Improvement Plan: low-performing schools that were only required to write a school improvement plan. This package represented 26 percent of the teachers employed in low-performing public schools.

3.) Evaluation Cycle: low-performing schools that were only required to undertake an evaluation cycle with required improvements by specific dates. This package represented 8 percent of the teachers employed in low-performing public schools.

4.) Improvement Plan and Evaluation Cycle: low-performing schools that were required to both write a school improvement plan and undertake an evaluation cycle with required

improvements by specific dates. This package represented 21 percent of the teachers employed in low-performing public schools.

5.) Improvement Plan and Evaluation Cycle and Subject to Takeover or Reconstitution: low-performing schools that were required to both write a school improvement plan and undertake an evaluation cycle with required improvements by specific dates and also subject to takeover or reconstitution. This package represented 2 percent of teachers employed in low-performing public schools.

6.) Improvement Plan and Evaluation Cycle and Penalized by a Reduction in Resources: this package was the same as #5 above, except instead of the school being subject to takeover or reconstitution, it was penalized with a reduction in resources. This package represented 2 percent of teachers employed in low-performing public schools.

We included all of the sanctions packages (#2 through #6) in regression models, allowing us to determine the effect of each package, compared to low-performing schools without sanctions (#1), while holding the other packages constant and after controlling for the background characteristics of the teachers and schools..

Each of the sanctions packages was directly associated with higher turnover. Teachers in low-performing schools only required to write a school improvement plan (package #2), had slightly higher turnover than those in low-performing schools without any of the four sanctions. Teachers in low-performing schools only required to undertake an evaluation cycle/timetable with required improvements by specific dates (package #3) had far higher turnover. Teachers in low-performing schools required to both write a school improvement plan and undertake an evaluation cycle with required improvements by specific dates (package #4) also had significantly higher turnover. In turn, those in schools with package #5 had far higher turnover. Schools with a reduction of resources, instead of reconstitution (package #6), showed a positive association with turnover in our bivariate analysis, but once controls for the other packages were added, this package did not differ from the others at a statistically significant level. This latter finding is not surprising given the very small number of schools receiving this package.

In sum, schools subject to performance standards and assessments did not have significantly different teacher turnover than those without standards or assessments, but schools that performed better had lower turnover and schools that performed poorly on the assessments had more turnover. While rewards to higher performing schools did little to improve the already high retention, the losses of teachers in lower-performing schools were exacerbated when sanctions were applied.

Our third research question now arises: is it inevitable that low-performing schools, especially those subject to penalties, lose more teachers? Do some low-performing schools have far better teacher retention than others and is it due to the way they are managed and organized? Is there anything that lower performing schools, especially those subject to sanctions, can do to ameliorate their losses of teachers? In short, is it possible to implement accountability in a way that does not exacerbate the difficulties in lower-performing schools by driving out teachers?

The Role of School Organization in the Impact of Accountability on Teacher Turnover

The objective of this last portion of our analysis was to ascertain if the impact of performance and sanctions on teacher turnover depended on several key organizational conditions, long associated with effective schools:

1.) *School Leadership Support*: the school mean of teachers' reports for four items: principal communicates expectations; administration is supportive; principal enforces rules for student discipline; principal communicates objectives; staff are recognized for job well done.

2.) Classroom Resources: the school mean of teachers' reports for one item: necessary materials such as textbooks, supplies and copy machines are available as needed by the staff.

3.) School-wide Faculty Influence: the school mean of collective faculty influence over seven areas: student performance standards; curriculum; content of in-service programs; evaluating teachers; hiring teachers; school discipline policy; deciding spending of budget.

4.) *Classroom Teacher Autonomy:* the school mean of individual teacher's control over six areas: selecting textbooks and other instructional materials; selecting content, topics and skills to be taught; selecting teaching techniques; evaluating and grading students; determining the amount of homework to be assigned; disciplining students.

The data show that lower-performing schools had significantly lower levels of these four organizational conditions than higher-performing schools. But, the data also showed large school-to-school differences in these conditions and, moreover, that these differences in organizational conditions were significantly related to differences in turnover; teachers in schools with higher levels of leadership support, or classroom resources, or classroom autonomy, or school-wide influence all had lower turnover, after controlling for the background characteristics of the teachers and schools and after controlling for school performance, rewards or sanctions.

Moreover, these organizational conditions appeared to interact with the effects of sanctions in lower-performing schools. That is, levels of turnover in low-performing schools with sanctions, depended on levels of these organizational conditions. However, one of these organizational variables stood out over the others as strongly related and consistently statistically significant – classroom teacher autonomy (see bottom of Figure 4). That is, the effects of sanctions on turnover strongly depended on the level of classroom autonomy allowed to teachers. Low-performing schools subject to sanctions have higher turnover, but this was significantly ameliorated if their teachers had higher levels of classroom autonomy. In other words, lower-performing schools with sanctions, had far higher turnover if their teachers were allowed less classroom autonomy and had far lower turnover if their teachers were allowed more autonomy.

This latter finding is especially important for highly tested subjects such as math. The data do not show that math or science teachers are more, or less, affected by low-performance or

sanctions than other teachers. But, the data do show that the degree of classroom teacher autonomy has an especially strong independent impact on the turnover of math teachers; several times higher than for other teachers. Hence, the data show that autonomy is associated with decreases in the high turnover in sanctioned lower-performing schools and this is especially pertinent for math teachers because autonomy is strongly tied to their turnover, regardless of the type of school.

Conclusion

Performance-based school accountability is a reform whose time has come. Almost all schools are now subject to performance standards established by their district or state, and are subject to evaluations assessing their performance in regard to the standards. Some of the schools that perform successfully receive rewards and many of those who perform less successfully are subject to sanctions. Moreover, not only is accountability widespread, there is no question these reforms have had an impact on teachers and their teaching practices.

Does accountability have an impact on the retention or turnover of teachers? Having performance standards and assessments themselves in a school did not have a negative impact on teacher retention. That is, a school having accountability standards and having state or district testing did not, in and of themselves, drive out teachers. But, how schools performed did impact retention and turnover. Successful schools had better retention. Less successful schools had worse retention. Rewards given to higher performing schools did little to improve the already higher retention. Sanctions applied to lower-performing schools did a lot to worsen their already lower retention. In short, the losses of teachers in lower-performing schools were exacerbated when sanctions were applied. Among the most consequential of sanctions for teacher turnover was a school being put on an evaluation cycle with required improvement by specific dates. Thirty percent of the lower-performing schools were subject to this sanction, either alone, or in combination with other sanctions. These schools had significantly lower retention than lower-performing schools that were not subject to sanctions.

But, our analyses also show that it is not inevitable that sanctioned lower-performing schools lose more teachers. Our analyses indicate that there is an important role for the leadership, management and organizational conditions in these schools. Poor performance and the application of sanctions did make it more difficult to retain qualified teachers. But, this depended on how much autonomy teachers were allowed in their own classrooms over key issues such as selecting textbooks and other instructional materials, selecting content, topics and skills to be taught, selecting teaching techniques, evaluating and grading students, determining the amount of homework to be assigned and disciplining students.

These findings are especially relevant for math teachers. The data show the following: Lower-performing schools allow less classroom autonomy for teachers. Lower-performing schools, especially those schools subject to sanctions, have more teacher turnover. Providing classroom autonomy to teachers in sanctioned lower-performing schools stems the outflows of teachers. Autonomy is especially consequential for math teachers; autonomy is independently and very strongly tied to their turnover, regardless of the type of school. Hence, if lower-performing schools are to better ensure that all their classrooms are staffed with qualified

mathematics teachers, the data suggest the importance of ensuring their staff has sufficient autonomy.

These results have large implications for reform and policy. Experts on organizational management and leadership have long advocated a balanced approach – employee accountability and employee autonomy must go hand in hand in workplaces, and increases in one must be accompanied by increases in the other (for a fuller discussion see Ingersoll, 2007; 2012). Imbalances between the two can result in problems for both the employee and for the organization. Delegating discretion and autonomy to employees without commensurate responsibility is irresponsible and can even be dangerous and harmful. That is, giving teachers more autonomy alone is not the answer. Likewise, accountability without commensurate autonomy is unfair and can also be harmful. In other words, it does not make sense to hold somebody accountable for something they do not control, nor does it make sense to give someone control over something for which they are not then held accountable. Both of these changes are necessary, but neither alone is sufficient to accomplish the larger systemic goal – ensuring quality teachers in every classroom.

Further Reading

Ingersoll, R. (2007). "Short on Power, Long on Responsibility." *Educational Leadership* 65:1:20-25

Ingersoll, R. (2011). "Do We Produce Enough Mathematics and Science Teachers?" *Phi Delta Kappan*, 92(6) 37-41.

Ingersoll, R. (2012). "Power, Accountability and the Teacher Quality Problem" Chapter 5, pp 97-109 in *Assessing Teacher Quality: Understanding Teacher Effects on Instruction and Achievement*. Edited by Sean Kelly. New York: Teachers' College Press.

Ingersoll, R. & Perda, D. (2010). Is The Supply of Mathematics and Science Teachers Sufficient? *American Educational Research Journal.* 47(3): 563-594.

Ingersoll, R. & May, H. (2012). "The Magnitude, Destinations and Determinants of Mathematics and Science Teacher Turnover." *Education Evaluation and Policy Analysi*s. 34 (4) 435 - 464 (Dec 2012)