Eight Reasons Merit Pay for Teachers is a Bad Idea

prepared for ETFO
by Ben Levin, OISE
October 2010

There are many ways in which professionals working in public organizations can be paid. Recent discussions in the United States of the idea of ‘merit pay’ have led to some similar discussion in Canada. ‘Merit pay’ here is defined as linking some portion of teachers’ pay to the academic achievement of their students. Many other kinds of pay for performance schemes also exist in other occupations, such as pay based on volume of work (for example doctors or lawyers). Based on an extensive review of research in education and other sectors, this paper argues that linking teachers’ pay to student achievement is not a desirable education policy. Eight reasons are put forward for this view.

1. Very few people anywhere in the labour force are paid on the basis of measured outcomes
2. No other profession is paid on the basis of measured client outcomes
3. Most teachers oppose such schemes
4. Pay based on student achievement is highly likely to lead to displacement of other important education purposes and goals
5. There is no consensus on what the measures of merit should be
6. The measurement of merit in teaching inevitably involves a degree of error
7. The details of merit pay schemes vary widely, yet these details have great impact on how such plans are received and their effects on teachers and schools
8. Merit pay schemes in education have a long record of failure

1 – Very few people anywhere in the labour force are paid on the basis of measured outcomes

It turns out that merit pay, when defined as paying people based on the outcomes of their work, is quite rare. According to Adams and Heywood (2009), various surveys show that only 16 to 30 percent of all workers get any kind of performance pay (which is a much broader category than merit pay as defined above), and only six percent are in ongoing performance pay systems. Further, most of this takes place in sales-related occupations. For example, people in the finance industry are paid on sales results, but that is not necessarily a recommendation, based on recent economic events, for wider application of the idea! Tradespeople are paid on work completed, not the results of that work. Others are paid on how much of a given task they perform. In many fields, earnings have much to do with reputation but reputation is not the same as a measured outcome.

2 – No other professional group is paid based on a measured outcome
The earnings of self-employed professionals may depend on reputation, but as noted that is not a measured outcome. Nurses, lawyers, engineers, architects, or even aircraft pilots are not paid based on outcomes though they may be paid based on volume of work. Even where there is pay for performance, the performance measures are not related primarily to measures of client outcomes. In the corporate world there is no relationship between the pay of corporate CEOs and measured performance, let alone the fact that when a company fails or they are fired they may get huge severance packages.

3 – Most teachers oppose such schemes

Surveys of teachers consistently show strong (70 percent or more) opposition to merit pay schemes. Since improvement in education depends critically on teachers’ commitment, anything that reduces commitment is likely to be unhelpful to better school outcomes.

4 – Pay based on student achievement is highly likely to lead to displacement of other important education purposes and goals

As James March pointed out many years ago, “A system of rewards linked to precise measures is not an incentive to perform well; it is an incentive to obtain a good score” (March, 1984, p. 28). When people have a financial incentive to achieve a score, that incentive may displace other, more desirable traits. Quite a bit of research in psychology shows that extrinsic rewards can act to displace intrinsic motivation, in which case merit pay schemes could reduce some teachers’ desire to do the job well simply because that is their professional responsibility and wish. Teachers, like other public sector workers, are primarily motivated by non-financial factors (though of course pay is also relevant) (Hulleman & Baron, 2010).

Since not all the important goals of education will be measured, goals that are measured in relation to pay are likely to get more attention at the expense of other goals that are not measured, whether those are different subject areas or soft skills or relationships with students.

If merit pay is individual and competitive (as is the case in many but not all plans), teachers will have incentives to cooperate and share less with colleagues, or even reasons to be glad that other colleagues are ineffective in their work.

5 – There is no consensus on what the measures of merit should be

The rationale behind merit pay is to link teachers’ pay to student outcomes. However schooling has many outcomes, so the question of which outcomes to use to determine merit is highly problematic. Academic achievement is not the only important outcome of schooling; we also value students’ ongoing ability to learn, interest in learning, abilities
to work with others, and citizenship skills. Most of these, however, will not be used in any given merit pay scheme because they would make it too complicated.

Even if one were to agree that the appropriate measure is academic achievement, there is the issue of how that should be measured. Does one measure the absolute level of attainment, which is strongly influenced by factors that are outside the teacher’s control such as socio-economic status? Or does one measure the incremental gain in learning, in which case there are problems of regression to the mean (it is very hard, for example, to show large gains if students are already high performing).

Then there is the question of the benchmark. Would teachers’ performance be measured against some norm or benchmark? If so, what benchmark? Or would it be measured against other teachers? If so, would the comparison be to teachers in the same school or district? Or teachers in schools with similar demographics? With others teaching the same course or subject? All of the above?

Another important point is that teacher performance is known to vary from one year or class to another, sometimes quite substantially. Nobody is equally good with all students, all courses and at all times. All human performance, whether teaching or playing an instrument, or athletics, or surgery will vary across time and tasks. This suggests that teacher merit should be judged across multiple years but that raises many further measurement problems (for example people who change roles, schools, grades, or subjects taught).

Then there is the issue of whether a measure of outcome should be a single test, or performance over a full year – but if the latter case what would be a fair assessment measure?

That brings us to the next point, about measurement error.

6 – The measurement of merit in teaching inevitably involves a degree of error

Every human measure involves a degree of error. In the case of merit pay schemes, teachers’ ‘merit’ is being inferred from student performance, but any measure of student performance, whether this is a teacher’s grade or a standardized test, will have some error in it. A good test will report a reliability measure (meant to indicate the correlation in scores of a single student taking the test on two different occasions). A score of 0.9 is a very high reliability for a test, but it means that 10 percent of the average score is potentially due to measurement error - that is, does not give a true picture of the student’s ability. Moreover, different students will be assessed using different measurements (e.g. there will be a different test at different grade levels), compounding the potential error across tests. Where significant amounts of money rest on the measure, a 10 percent error is very significant, especially when combined with the previous point about the uncertainty of which measures to use in the first place. The total measurement error in any merit pay scheme is likely to be quite large, making the results unreliable.
7 – **The details of merit pay schemes vary widely, yet these details have great impact on how such plans are received and their effects on teachers and schools**

The points already made regarding measures are only one of the many decisions to be made in any merit pay scheme. Other choices include:

- Which teachers to include. What happens with teachers who are not teaching a class of students (for example special education teachers or counsellors or librarians), or who are teaching in an area that is not measured (such as music or physical education)? Should teachers who are in their first or second year of teaching, or are teaching a new grade or subject be judged on the same basis? What about classes with high levels of turnover during the year?

- Is the measure applied to each teacher or to groups of teachers? If the latter, is the group an entire school staff or some subset? If groups are used, is the average of each member the right measure? If there is a group reward is it shared equally among all?

- Can all teachers potentially receive the merit amount or is it restricted to a given percentage of teachers? If the latter, what percent would be eligible and on what basis would that limit be set? If merit pay is competitive (i.e. limited to the top x% of performance) it will likely have the same negative effects as have been noted for other competitive measures, such as the potential benefit of seeing colleagues do badly.

- How much money is involved? Is merit pay a small amount (say 1-2% of salary), in which case it might have very little incentive effect on teachers? But if merit pay is much more substantial – say 10 percent of pay – then it will come with a very high overall price tag. If we take Ontario as a case, and posit that in any given year 10 percent of teachers can earn an additional five percent of salary (both quite low numbers), this would mean some 10,000 teachers would each get an additional $4,000, for a total cost of $40 million per year. If 20 percent of teachers could earn 10 percent more, the cost would be $160 million. Yet even in this scheme, 80 percent of teachers would not be involved and so might be demotivated by it. So any scheme that provided significantly more money to a significant number of teachers would require substantial amounts of additional funding.

- Is additional pay one time or is it built into base salary? This would make a big difference to how teachers perceived the program and to its costs over time.

It’s evident that these are very important choices that could have large impacts on how a plan worked and how teachers felt about it. So the concept of ‘merit pay’ can cover many, many very different options with different consequences.

8 – **The evidence for merit pay for teachers is weak, and many schemes have been tried but do not last**

Merit pay is not a new idea. Such plans go back more than 100 years. There has not, however, been a great deal of careful empirical study. Some of the studies currently
cited are from very different contexts, such as India, and may have little applicability to Canada. Studies in developed countries yield equivocal results, but very few have found strong positive effects. Further, though many merit pay schemes have been adopted in various parts of the US in the last 20 years, few of these have lasted more than a few years, suggesting that for one reason or another they were not sustainable. Where evidence is weak and experience is not positive, there are good reasons to be guarded about any policy.

A common conclusion of all the analyses of merit pay is this: “New pay plans can’t be summarily imposed on teachers… teachers need to be part of the process of development, and they need to own the pay plan that emerges” (Koppich, 2010).

References


Other works


