



Evaluating and Rewarding the Quality of Teachers

INTERNATIONAL PRACTICES



Evaluating and Rewarding the Quality of Teachers

INTERNATIONAL PRACTICES



ORGANISATION FOR ECONOMIC CO-OPERATION AND DEVELOPMENT

The OECD is a unique forum where the governments of 30 democracies work together to address the economic, social and environmental challenges of globalisation. The OECD is also at the forefront of efforts to understand and to help governments respond to new developments and concerns, such as corporate governance, the information economy and the challenges of an ageing population. The Organisation provides a setting where governments can compare policy experiences, seek answers to common problems, identify good practice and work to co-ordinate domestic and international policies.

The OECD member countries are: Australia, Austria, Belgium, Canada, the Czech Republic, Denmark, Finland, France, Germany, Greece, Hungary, Iceland, Ireland, Italy, Japan, Korea, Luxembourg, Mexico, the Netherlands, New Zealand, Norway, Poland, Portugal, the Slovak Republic, Spain, Sweden, Switzerland, Turkey, the United Kingdom and the United States. The Commission of the European Communities takes part in the work of the OECD.

OECD Publishing disseminates widely the results of the Organisation's statistics gathering and research on economic, social and environmental issues, as well as the conventions, guidelines and standards agreed by its members.

This work is published on the responsibility of the Secretary-General of the OECD. The opinions expressed and arguments employed herein do not necessarily reflect the official views of the Organisation or of the governments of its member countries.

This publication is a product of the co-operation agreement established between the Government of Mexico and the OECD, which aims to improve the quality of education in Mexico.

ISBN 978-92-64-06198-9 (print)
ISBN 978-92-64-03435-8 (PDF)

Photo credits: ©UNESCO/José Gabriel Ruiz Lembo.

Corrigenda to OECD publications may be found on line at: www.oecd.org/publishing/corrigenda.

© OECD 2009

You can copy, download or print OECD content for your own use, and you can include excerpts from OECD publications, databases and multimedia products in your own documents, presentations, blogs, websites and teaching materials, provided that suitable acknowledgment of OECD as source and copyright owner is given. All requests for public or commercial use and translation rights should be submitted to rights@oecd.org. Requests for permission to photocopy portions of this material for public or commercial use shall be addressed directly to the Copyright Clearance Center (CCC) at info@copyright.com or the Centre français d'exploitation du droit de copie (CFC) at contact@cfcopies.com.

Foreword

Compelling incentives for individuals, economies and societies to raise levels of education have been the driving force for governments to improve the quality of educational services. The prosperity of countries now derives, to a large extent, from their human capital, and to succeed in a rapidly changing world, individuals need to advance their knowledge and skills throughout their lives. On the premise that the quality of education cannot exceed the quality of teachers, countries are devoting increasing efforts towards improving the quality of teachers and teaching.

Many of the best performing education systems now share a commitment to professionalised teaching in ways which imply that teachers are on par with other professions in terms of diagnosis, the application of evidence-based practices, and professional pride. They attract the best graduates to become teachers and put in place incentives and support systems that get the most talented school teachers into the most challenging classrooms. They are reforming inherited, traditional and bureaucratic systems of recruiting and training teachers and leaders, as well as the mechanisms of paying and rewarding them and of shaping their incentives, both short-term and long-term. They also develop teachers into effective instructors, through, for example, coaching classroom practice, moving teacher training to the classroom, developing strong school leaders and enabling teachers to share their knowledge and spread innovation. They also put in place differentiated support systems to ensure that every child is able to benefit from excellent instruction, the image being that of teachers who use data to evaluate the learning needs of their students, and who are consistently expanding their repertoire of pedagogic strategies to address the diversity in students' interests and abilities.

These goals were also at the heart of an Agreement which the Government of Mexico and the OECD signed in 2008: "Improving School Education in Mexican Schools". One of the strands of this agreement involves the development of appropriate policies and practices to evaluate the quality of schools and teachers and to link outcomes to incentives for improvement. The work is being guided by a Steering Group composed of leading experts in the field.

This volume provides first results from the work of this Steering Group. The volume synthesises research and experiences from several countries and perspectives on the relatively new topic of rewarding effective teaching to improve educational outcomes. The preparation of the volume has brought together researchers from both education and the social sciences as active contributors to the debate on teacher evaluation and rewards programmes, guided by Susan Sclafani who designed and edited the volume. Chapter 3 was drafted by Matthew G. Springer and Ryan Balch of the National Center on Performance Incentives, Vanderbilt University; Chapter 4 by John Bangs of the National Union of Teachers, England; Chapter 5 by Dan Goldhaber of the Center on Reinventing Education, University of Washington; Chapter 6 by Cynthia Prince, Patrick Schuermann, James Guthrie and Julie Koppich of Vanderbilt University and Koppich and Associates; Chapter 7 by James Guthrie and Patrick Schuermann of Vanderbilt University and Chapter 8 by Jane Hannaway and Martin Corritore of the Urban Institute.

The volume and the work on evaluation and incentive policies, carried out as part of the agreement with the Mexican government, was produced under the auspices of the Indicators and Analysis Division of the OECD Directorate for Education under the responsibility of Andreas Schleicher, Michael Davidson, Alejandro Gomez Palma, Marlene Gras, Diana Toledo Figueroa, Elisabeth Villoutreix, and Isabelle Moulherat. Editing of the report was done by Janet Looney and layout by Fung Kwan Tam.

Table of Contents

LIST OF ACRONYMS AND ABBREVIATIONS	9
CHAPTER 1 EVALUATING AND REWARDING GOOD TEACHERS	11
The importance of quality teachers	13
Examples of teacher evaluation systems	14
Challenges in linking rewards to evaluation	15
Importance of evaluation and rewards in attracting and retaining quality teachers	16
Salary compensation and links to evaluation	18
• Problems with the single salary schedule	18
Overview and contributions of up-coming chapters	20
References	24
CHAPTER 2 A REVIEW OF INTERNATIONAL PRACTICE	29
Financial incentives	30
• Incentives for teacher education	30
Incentives based on teacher's knowledge and skills	31
• Evaluation of teacher performance	31
Incentives for teaching in shortage subject areas	34
Incentives for teaching in schools in difficult environments	34
Other financial incentives	37
Incentives based on performance in the classroom measured by student results on external examinations	38
Non-financial incentives	41
Evidence of effectiveness	44
Evaluations of national and state/regional programmes	44
• Experimental studies of incentives	49
Annex 2A: Incentives in education	52
References	60
CHAPTER 3 DESIGN COMPONENTS OF INCENTIVE PAY PROGRAMMES IN THE EDUCATION SECTOR	65
Introduction	66
Incentive structure	67
• Rank-order tournaments	67
• Fixed performance contract	68
Unit of accountability	71
• Individual unit of accountability	71
• Group unit of accountability	72
• Hybrid models of accountability	74

Performance measures	75
• Educational inputs	75
• Educational outputs	79
• Multiple measures – linking educational inputs and outputs	82
Performance standards and thresholds	83
• Linear models	83
• Step functions	83
• Limited linear models	84
• Threshold levels	85
Size and distribution of bonus awards	85
• Size of bonus	85
• Distribution of award	87
• Non-monetary incentives	88
Payout frequency	89
Cultural considerations	90
Conclusion	91
References	94
CHAPTER 4 WHOSE INCENTIVES? WHOSE STIMULI? PERFORMANCE MANAGEMENT IN ENGLAND	103
Introduction	104
Incentives and stimuli: performance management and performance-related pay	104
Incentives and stimuli: teachers' continuing professional development	111
Incentives and stimuli: school leadership	115
Evaluation and assessment: the need to differentiate its purposes – a footnote	117
The need for dialogue between teacher organisations and governments on school improvement	118
References	119
CHAPTER 5 EXPLORING THE USE OF INCENTIVES TO INFLUENCE THE QUALITY AND DISTRIBUTION OF TEACHERS	121
Why focus on teachers?	122
Salary schedules and the incentives that exist now	124
Using incentives: technical and practical considerations	127
Evidence on purposeful teacher incentive systems	137
Conclusion: infrastructure to support incentive systems	140
References	144
CHAPTER 6 INCENTIVE SYSTEM IMPLEMENTATION	151
Stakeholder engagement and communication strategies are essential to successful compensation reform initiatives	153
• A range of communication strategies is needed for different purposes and different audiences	153
• Stakeholder engagement helps to ensure programmes reach consensus on core purposes and create a unified vision for the plan	153
• Stakeholder engagement efforts help to ensure the system will be accepted and lead to desired results	154

Potential consequences if educators are not engaged in the formation of policies on compensation	154
• Educators may purposely design pay plans so that rewards are unattainable.....	154
• Teachers may vote to withdraw from the compensation programme.....	155
Desirable practices in communication and stakeholder engagement.....	155
• Create a representative compensation reform committee.....	155
• Build support for the reform effort.....	158
Lessons learnt.....	161
Developing a feedback loop and making mid-course corrections.....	163
How different countries and localities engage teachers and their unions in the development of incentive systems.....	163
Conclusion.....	166
References.....	167

CHAPTER 7 PROFESSIONAL EDUCATORS AND THEIR PAY: POLICY, IMPLEMENTATION AND ALIGNMENT ISSUES.....

Overview of key educator pay issues: stabilising costs and raising student achievement.....	170
Education reform strategies and their results to date.....	173
Educator compensation reform possibilities.....	174
Considerations in the design and implementation of performance pay programmes.....	177
The influence of federal, state, district and school contexts.....	178
Cultivating a culture of continual appraisal and mid-course corrections.....	181
Components of high-quality performance pay programmes.....	182
21 st century educator pay for performance plans: permanent professional practices or ephemeral policy propositions?	183
Iron triangles, high politics, political champions, and transaction costs.....	183
Positioning performance pay on the political continuum.....	184
The political valence of how performance pay is practically portrayed.....	185
The potentially powerful, but unpredictable, impact of research experiments and experience.....	186
The path ahead.....	187
Systemic integration and alignment: the surest way to performance pay survival.....	187
Conclusion.....	191
Annex 7A: Practical preparation for performance pay programme implementation.....	192
References.....	197

CHAPTER 8 THE POLITICS OF PERFORMANCE-BASED INCENTIVES IN EDUCATION: AN INTERNATIONAL LOOK.....

A sharp break: the management of teachers' work.....	201
Unions as political actors.....	202
Internal balancing act.....	208
Research: the devil is in the details.....	209
Conclusion.....	210
References.....	213

CHAPTER 9 PUTTING IT ALL TOGETHER	217
Key design issues	219
• Prerequisite conditions	219
• For evaluation	221
Design decisions	222
Key implementation issues	232
• Design of a pilot programme to test design	232
• Creation of an evaluation plan	233
• Alignment of compensation plan with system improvement plans	234
• Identification of “champion” for the plan	235
• Regular communication with stakeholders	235
• Development of broad political support	235
Training of school leaders and teachers	236
Revisions in plan from pilot results	237
• Full implementation	238
Conclusion	239
References	240

List of Figures and Tables

FIGURES

Figure 1.1	Percentage of students at each proficiency level on the PISA science scale	12
Figure 1.2	Cost of schooling spirals upward, but achievement does not	17
Figure 3.1	Percentage of schools with a minimum of 72% of third grade students scoring proficient in reading and mathematics	69
Figure 3.2	Performance standards and thresholds	84
Figure 3.3	Distribution of actual awards in the Governor Educator Excellence Grant programme	87
Figure 5.1	Distribution of estimated teacher effectiveness, by licensure standard testing	123
Figure 7.1	Current expenditure per pupil in fall enrolment in public elementary and secondary schools: Selected years, 1959-60 through 2005-06	171
Figure 7.2	International comparisons of expenditures for education	172
Figure 7.3	Total public expenditure on education as a percentage of total public expenditure (2000, 2005)	173
Figure 7.4	Examples of education reform strategies	174
Figure 7.5	Components of teacher compensation package	175
Figure 7.6	Comparison between traditional pay raises and performance pay bonuses	176
Figure 7.7	Comparison of compensation systems	176
Figure 7.8	Continuous improvement cycle	181
Figure 7.9	Political and measurement dynamics of performance pay	184
Figure 7.10	Survey results regarding support for added pay for teachers	186
Figure 7.11	Components of the human capital framework	189
Figure 7.12	Contexts and forces of the human capital framework	190

TABLES

Table 3.1	Characteristics of incentive pay programmes in the United States and other countries	76
Table 3.2	Performance measures in incentive pay programmes	86
Table 3.3	Cultural dimensions for selected countries with incentive pay programmes	90
Table 6.1	Communication strategy action plan items for SCTAP (South Carolina Teacher Advancement Programme)	160

List of acronyms and abbreviations

ACE	Alliance for the Quality in Education (<i>Alianza por la Calidad de la Educación</i>)
AEP	Pedagogical Excellence Allowance
AFT	American Federation of Teachers
ANMEB	National Agreement for the Modernisation of Basic Education (<i>Acuerdo Nacional para la Modernización de la Educación Básica</i>)
API	Academic Performance Index
APLE	Assumption Programme of Loans for Education
AST	Advanced Skills Teachers program
AVDI	Variable Allowance of Individual Performance
BBLs	Baccalaureate and Beyond Longitudinal Study
BEST	Better Educated Students and Teachers
BPRS	Best Practice Research Scholarships
CM	<i>Carrera Magisterial</i>
CNN	Cable News Network
CONNECT	Continuity, Experience and Commitment Programme
CPD	Continuous Professional Development
CPRE	Consortium for Policy Research in Education
DATE	District Awards for Teacher Excellence
EPPI	Evidence for Policy and Practice Information
ETS	Educational Testing Service
FCAT	Florida Comprehensive Assessment Test
GCSE	General Certificate of Secondary Education
GDP	Gross Domestic Product
GEEG	Governor's Educator Excellence Grant
GTCE	General Teaching Council for England
ICE	Collective Incentive to Schools (<i>Incentivo Colectivo a Escuelas</i>)
JTPA	Job Training Partnership Act
MAP	Merit Award Program
MTL	Masters level qualification in teaching and learning
NAEP	National Assessment of Educational Progress
NBCTs	National Board Certified Teachers
NBPTS	National Board for Professional Teaching Standards
NCES	National Center for Education Statistics
NCTQ	National Council on Teacher Quality
NEA	National Education Association

NELS	National Education Longitudinal Survey
NFER	National Foundation for Education Research
NIET	National Institute for Effectiveness in Teaching
NPQH	National Professional Qualification for Headship
NUT	National Union of Teachers
NYC	New York City
PISA	Programme of International Student Assessment
PRELAC	Regional Education Project for Latin America and the Caribbean
ProComp	Professional Compensation System for Teachers
Q Comp	Quality Compensation for Teachers
SASS	Schools and Staffing Survey
SAT	Scholastic Aptitude Test
SCTAP	South Carolina Teacher Advancement Program
SEP	Ministry of Public Education (<i>Secretaría de Educación Pública</i>)
SIMCE	Measurement System for the Quality in Education (<i>Sistema de Medición de la Calidad de la Educación</i>)
SNED	National System for Assessing School Performance (<i>Sistema Nacional de Evaluación del Desempeño de los Establecimientos Educativos Subvencionados</i>)
SPBP	School-Wide Performance Bonus Programme
STAR	Special Teachers Are Rewarded
TAP	Teacher Advancement Programme
TEEG	Texas Educator Excellence Grant Programme
TIMSS	Trends in International Math and Science Surveys
TLA	Teacher Learning Academy
TRACS	Toledo Review and Alternative Compensation System
U.S.	United States
VAM	Value-Added Model

CHAPTER 1

Evaluating and Rewarding Good Teachers

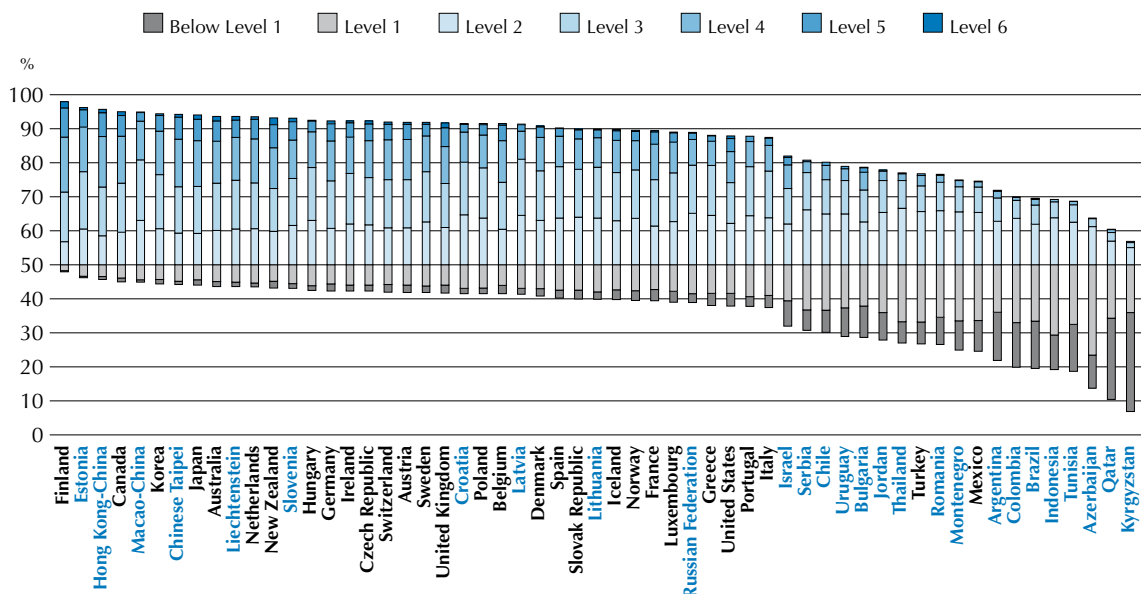
The importance of quality teachers.....	13
Examples of teacher evaluation systems.....	14
Challenges in linking rewards to evaluation.....	15
Importance of evaluation and rewards in attracting and retaining quality teachers.....	16
Salary compensation and links to evaluation.....	18
• Problems with the single salary schedule.....	18
Overview and contributions of up-coming chapters.....	20
Notes	23
References	24

Across the world, nations recognise that their futures are linked to the quality of their education systems. Education economists point out that there is a value to education for both the individual and society (Hanushek, 1996). Not only does education lead to higher skills, increased salaries, and lower unemployment, but also better health, greater social participation and less dependence on social services. Given the changing expectations of knowledge-based economies, there is a greater urgency for nations to address concerns about the quality of their education systems: successful companies demand that employees at every level of the enterprise have a far greater level of knowledge and skills than at any time in the past. To be competitive in attracting and retaining businesses that will sustain productive economies in their countries, nations must develop employee workforces that are able to meet new demands. A recent study (McKinsey & Company, 2009) of the achievement gap in the **United States (U.S.)** suggested that if the United States had closed their achievement gap with countries like **Finland** and **Korea** by 1998, its Gross Domestic Product (GDP) would be 9-16% higher than it is today. That is a greater impact than the level of the negative effect of the recent recession, which was 6.4% at the end of 2008.

Countries are therefore concerned with assessing and evaluating the performance of their educational systems. In 2006, 36 countries (OECD member countries and six partner countries) reported on the implementation of national examinations, national periodical assessments, and school evaluations (either external inspections or self-evaluations of school staff). Ten countries reported the existence of national examinations (**Denmark, France, Iceland, Ireland, Italy, Norway, Portugal, Scotland, Turkey, and Estonia**). In each case, mathematics and language were assessed, as well as other subjects, with the exception of France where science was not part of the assessment and Portugal where only mathematics and language were assessed.¹

Figure 1.1

Percentage of students at each proficiency level on the PISA science scale



Countries are ranked in descending order of percentage of 15-year-olds at Levels 2, 3, 4, 5 and 6.

Source: OECD, 2007.

Education's relevance to international economic competitiveness is also reflected in the use countries make of international assessments and international benchmarking studies to compare the performance of their students to others. As the Programme for International Student Assessment (PISA) and the Trends in International Mathematics and Science Study (TIMSS) demonstrate, some nations' students achieve at much higher levels than others. PISA assesses student knowledge and skills on a scale of 1 to 5, with Level 1 representing the lowest level of achievement. Figure 1.1 provides an example of the contrast among OECD countries, showing that over 80% of **Finland's** students are achieving at or above Level 3 of the science component of the 2006 PISA, while over 80% of **Mexico's** students are achieving at Level 2 or below.² These differences suggest that it is not merely improvement by national standards, but the best performing education systems internationally that set the benchmarks for educational success. Based on the PISA 2003 results and surveys, Fuchs and Wößmann (2004) found that a number of institutional factors were related to higher student achievement, such as external examinations, and school autonomy in the choice of textbooks, hiring of teachers and within-school allocations of school budgets.

THE IMPORTANCE OF QUALITY TEACHERS

If the purpose of educational systems is to ensure quality learning outcomes and higher student achievement, it becomes necessary to identify the elements that contribute most towards this end. Educators and parents have long argued that effective teachers are key to improving student achievement and a number of research studies conducted since 1990 have backed up this belief. Hanushek (1992) found that students whose teachers are at the top of the effectiveness range achieve as much as an additional year of growth in student learning over those with teachers near the bottom of the range – a gain of 1.5 years of academic growth as opposed to 0.5 years of growth in a single year.

Sanders and Rivers (1996) used the large **Tennessee**, student-testing database to identify impacts of teachers on student performance. They found that not only were there large differences between the effectiveness of different teachers, but also that the effects were both additive and cumulative: given two similarly performing students in grade two, if one student had three high performing teachers while the other student had three low-performing teachers, the difference in their performance three years later might be as much as 54 percentile points. That earlier educational experience continued to affect student performances beyond the period studied.

Rivkin *et al.* (2001) compared students of similar socio-economic backgrounds and achievement levels and found that a student moving up to one standard deviation from an average teacher, to a teacher whose effectiveness was rated in the 85th percentile added four percentile points to his/her ranking. That is the equivalent effect of reducing class size by ten students, or nearly 50%. Schacter and Thum (2003) went one step further and looked at the quality of teaching behaviours displayed by high-performing teachers. They found that teachers who scored well against the researchers' evaluation rubrics (representing 12 teaching performance standards) helped students to achieve at higher levels than teachers involved in any other comprehensive school-reform model studied. Teachers in their study produced a 0.91 standard deviation increase in student achievement, greater than the effect of parental income. Their work provides evidence that effective teaching does result in higher student performance and also contributes to identifying the attributes of quality teachers.

Given the importance of teacher quality in determining student learning outcomes, it is inevitable that education systems are increasingly seeking to tie the assessment of teachers' work to student achievement. The Teaching and Learning International Survey (TALIS), conducted by OECD and published in 2009 as *Creating Effective Teaching and Learning Environments: First Results from TALIS*, provides insights from random samples of over 70 000 teachers in 23 countries. TALIS results show that 65% of teachers reported that student test scores were a moderate or highly important criteria in the appraisal or feedback that they received. In Bulgaria, Malaysia, Mexico and Poland this percentage was well over 80%.

While improving student learning outcomes is the central objective of the work of teachers, measures of such outcomes are not the only measure of teacher quality. Recent literature reviews focusing on the different aspects of quality teachers have suggested the following dimensions:³

- **Teacher qualifications:** these include “inputs” such as teacher credentials, years of service, degrees, certifications and other relevant professional development.
- **Teacher characteristics and in-classroom practices:** these include attitudes, expectations, and personal characteristics. Practices are the strategies, methods and actions employed by teachers in the classroom and in their interaction with students.
- **Teacher effectiveness:** an assessment of the degree to which in-service teachers contribute to the learning outcomes of students.

A review of practices in OECD countries shows that different dimensions of teacher quality can be assessed by a variety of instruments, including scores of standardised student assessments, classroom observation, student-generated ratings, peer ratings, school principle and/or administrator ratings, self-evaluation, teacher interviews and portfolios, parental ratings, competence-based tests, and other indirect measures (OECD, 2009b). As discussed in Chapter 5, student learning outcomes are commonly assessed by standardised tests on which statistical analyses of teacher or school contributions, such as Value-Added-Models (VAMs), can be based.

According to teachers’ reports from TALIS, the criteria most commonly used on average among the participating countries are: student-teacher relations (85% of teachers reported this as a criteria with high or moderate importance), knowledge and understanding of the teacher’s subject field (80% of teachers), classroom management (80% of teachers), knowledge and understanding of pedagogical practices (78%), student discipline/behaviour (78% of teachers) and relations with colleagues (78% of teachers). Approaches to teacher evaluation that adequately assess the different dimensions of quality teaching, therefore, are varied and continue to evolve across and within countries.

EXAMPLES OF TEACHER EVALUATION SYSTEMS

In the **United States**, external evaluations include both the National Board for Professional Teaching Standards (NBPTS) (NBPTS, 2009) certification and the PRAXIS III examination. Formed in 1987, the NBPTS created teacher standards and assessment processes, including examinations of teachers’ content and pedagogical knowledge, as well as examples of teaching (e.g. videos of teaching practice and examples of student work). There are currently 64 000 National Board Certified Teachers (NBCTs) in all 50 states and the District of Columbia. A recent National Research Council (2008) meta-analysis of studies found that there was sufficient evidence to conclude that students taught by NBCTs make higher gains on achievement tests than those taught by teachers who have not applied and those who did not achieve certification. However, Chapter 5 cautions that NBPTS certification does not guarantee that every NBCT-certified teacher will outperform teachers without this certification.

PRAXIS III is an assessment process developed by the Educational Testing Service (ETS) (ETS, 2009) that includes direct observation of classroom practice, review of documentation prepared by the teacher, and structured interviews with trained assessors. It is used in some states in the **U.S.** as a final licensing assessment after a teacher’s first year of teaching. Without passing the assessment, a teacher cannot be considered for tenure. **Chile** (Manzi, 2008) provides individual teacher incentives based on the quality of the teacher’s evaluation aligned to national teaching standards, *The Framework for Good Teaching (El Marco para la Buena Enseñanza)*. Some programmes in the **U.S.** establish four to six levels of teaching performance, ranging from novice to master. To move from one level to a higher one, teachers must have a successful performance evaluation that is based on rubrics and teaching standards for the next level. Odden and Wallace (2008) describe several models that have been used by states or districts in the United States, including the BEST system in Connecticut, Charlotte Danielson’s

A Framework for Teaching (2008), and Wallace's *Teacher Excellence Through Compensation*. The purpose of these evaluation systems is to identify teachers' strengths and weaknesses according to the accepted model of teaching, and provide professional development and coaching to assist the teachers in improving their practices. If the only outcome is a rating given to the teacher without feedback and assistance in improving, the evaluation system has done only half its job.

Research in **Cincinnati** (Milanowski, 2004; Milanowski and Kimball, 2005; Milanowski, Kimball and Odden, 2005) on the use of the Danielson model showed that the evaluation system was both reliable across raters and valid, with the group of teachers who scored well on the evaluation system also producing greater student growth on tests. Jacob and Lefgren (2005) compared school leader evaluations of teacher performance and the measurement of gains in student achievement on standardised tests. They found that school leaders were very good at identifying the teachers whose students made the largest and smallest achievement gains on standardised tests, but did not do very well in the middle of the range. However, they found that the school leaders' evaluations were significantly better predictors of student achievement gains than teacher experience, teacher degrees/education or actual salary.

Without better structured evaluation programmes such as the models described above, it is easy to see why teachers do not have faith that their principal is going to be able to make the fine-grained distinctions needed to identify effective teachers and reward them. But such fears can be overcome. As results from TALIS show, some 63% of teachers in TALIS countries agreed and 20% strongly agreed that the appraisal they received was a fair assessment of their work.⁴

Where there is no corresponding incentive or reward to improve practice or to facilitate teachers' development, however, efforts to raise standards in schools will be undermined. Again from TALIS, over three-quarters of teachers report that they would not receive recognition or rewards for improving the quality of their practice or for being more innovative in their teaching. Thus, the most effective teachers are not recognised by their principals, and teachers with poor practice do not receive any less compensation than other teachers. Furthermore, only one-quarter of teachers across the 23 TALIS countries reported that, as a result of the appraisal, there was a moderate or large change in their opportunities for engaging in professional development activities. The effective linking of rewards to evaluation is therefore a crucial consideration in designing the system of teacher evaluation.

CHALLENGES IN LINKING REWARDS TO EVALUATION

Lazear (1996) studied the use of performance-based pay at an auto glass company and found the change to a "pay for piece rate" dramatically improved productivity in two ways. First, employees had an incentive to improve their productivity as they saw their pay rise. Second, those who were successful stayed on the job and had fewer absences, and those who made less money under this system found other employment. This provided more opportunities to hire more productive workers and reduced variation in the workforce. The performance in this case, however, was easily observed and measured. This is not true of the complex role of the teacher, where teachers are required to meet a variety of objectives all at the same time. A teacher must not only teach the specific subject, but must address the needs of the child, the class and the school as well. Monitoring the "outputs" of each teacher would be extremely costly, and yet, as Murnane and Cohen (1985) point out, imperfect monitoring may result in teachers appearing to adopt behaviours the school leader promotes while being observed but then engaging in "opportunistic behaviour" at other times. Opportunistic behaviour refers to unintended behaviours that may increase the likelihood of reward, but go against the organisational goals set for the school. Such practices may include: focusing time on the specific areas of the subject that are most likely to be tested and ignoring the development of more complex conceptual understanding; spending more time on the subjects tested to the detriment of other subjects; paying more attention to students who are more likely to pass a test than other students; and, at the extreme, cheating by giving students answers to questions or changing the students' responses before a test is submitted. These examples highlight the need to have a sufficiently robust and credible mechanism of evaluation on which to base rewards.

Since teachers go into the profession because they want to make a difference in the lives of their students, some worry that intrinsic motivation may indeed be subverted when external financial bonuses are put in place (Rothstein, 2008). Rothstein has two major cautions for educators. First, he explains that Simon's concept of bounded rationality within organisations suggests that education's goals are too complex to easily quantify and measure, and, second, the act of measurement, according to Campbell's law of performance measurement, would corrupt the social processes being monitored. Neal (2008) applies Holmstrom & Milgrom's (1991) multitasking theory to education and suggests that Campbell may be right: because teachers are required to multitask within the classroom as well as in other school responsibilities, they may focus on actions with the greatest impact on the performance statistics which are measured, even if these are not necessarily the actual student outcomes desired, such as long-term learning, problem solving ability or creative thinking.

Eberts, Hollenbeck and Stone (2002) cite three additional constraints relating to the evaluation and rewarding of teachers. The first is the inability to measure desired outputs directly, requiring a reliance on subjective evaluations. This is subject to game-playing on either the evaluator's or the teacher's part; evaluators may also desire to give teachers high scores to keep the peace, or compress scores to appear to be rigorous. The second is the team production problem. School-based incentives promote team efforts where individual efforts are hard to measure. This may lead to some teachers acting as free-riders on the efforts of others. The final constraint is the issue of multiple stakeholders, each with its own set of priorities. If incentives are based on test scores, they may not represent the different and perhaps conflicting priorities of principals, parents, school board members, employers or the community at large. This can lead to poorly aligned goals for the school. These critiques suggest that the basis for rewarding good teaching is a robust and credible mechanism of teacher evaluation.

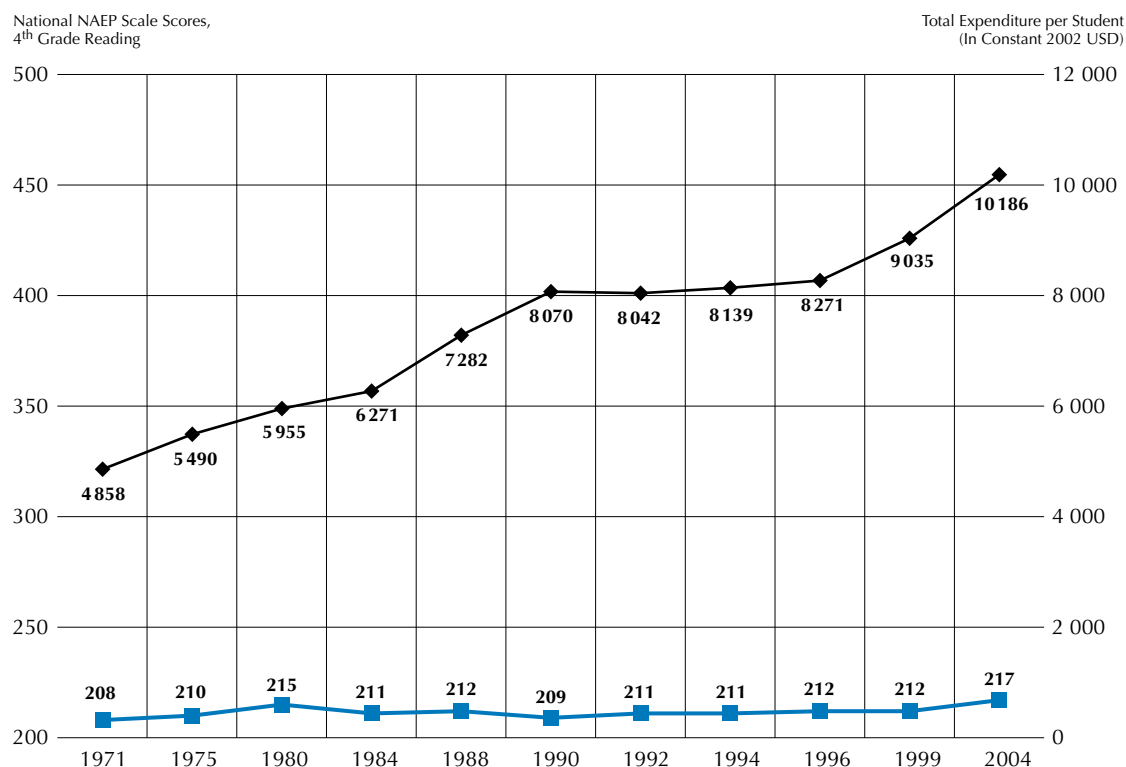
IMPORTANCE OF EVALUATION AND REWARDS IN ATTRACTING AND RETAINING QUALITY TEACHERS

While we know how critical effective teachers are to high student achievement, there are two factors complicating efforts to ensure schools attract and retain them. The first is that, in about half of the OECD countries that participated in the *Teachers Matter* (OECD, 2005) research, there is a shortage of qualified teachers. In some cases, it is an overall shortage, and, in others, it is in specific subjects or specific schools in remote areas or serving a high proportion of disadvantaged students. Lakdawalla (2001) attributes this to the lack of productivity increases in teaching as compared to other fields. Because teachers' knowledge is general rather than specialised, in a society that has placed a premium on specialised knowledge, it loses its productivity over time. As a result, teachers are no longer among the best-educated people in society. Lakdawalla suggests that class-size reductions over the last 20 years were attempts by school districts to make up for lower-quality teachers. In response to shortages, some OECD countries have lowered standards for entry into the profession, exacerbating the problem of quality, particularly given new demands for academic rigour.

The second factor is that teaching appears to be a less popular choice of profession for young people. In most countries, the teaching profession is not as highly respected as it once was, and yet teachers' roles have become more complex as student populations have become more diverse and expectations for their achievement have risen. Teachers are asked to be responsible for both academic and social-emotional development of their students, and to remain up-to-date with emerging knowledge in their subject areas and field (OECD, 2005). As a result, teacher education programmes are attracting students for whom the profession is not a first choice or whose academic backgrounds are weaker. In the **U.S.**, this has resulted in a statistically significant drop in the quality of academic preparation of people entering teaching. Corcoran, Evans and Schwab (2004) found that a female student in the top 10% of her high school class is now half as likely to enter teaching as she would have been in 1964, a decline from 20% to 10%.

Figure 1.2

Cost of schooling spirals upward, but achievement does not



Source: NCES, 2003.

While Scholastic Aptitude Test (SAT)⁵ scores of United States upper-secondary school students entering teacher education programmes are lower than for other students, only 40% of those graduates scoring in the top half in 1982 entered teacher programmes (Hanushek and Pace, 1995). SAT scores of those who graduated and became teachers are somewhat higher, with over 60% of the graduates in the top half of the SAT score distribution, but not in the top quartile where many high-performing countries recruit their teachers. Podgursky, Monroe and Watson (2004) found that both men and women with high scores on the American College Testing programme (ACT)⁶ were less likely to select teaching as a career.

At the same time as the quality and quantity of teacher candidates are declining, countries are facing a higher level of teacher retirements over the next decade than they have in the past (OECD, 2005). Many of the women who chose teaching when there were fewer options are now reaching retirement age. Current attrition rates in most OECD countries and a number of other countries around the world are as high as they have ever been. In the **U.S.**, it is estimated that 50% of all new teachers leave the field of education in the first three to five years of employment. While other countries may have lower attrition rates, only **Italy**,

Japan and **Korea** (OECD, 2005) had rates lower than 3% per year. The cost of replacing each United States teacher (Barnes, Crowe and Schaefer, 2007) is estimated to range from USD 4 400 for a small rural district to nearly USD 18 000 in a large urban district, like the Chicago Public Schools. While not all of the teachers who leave a school leave the profession, teacher exit rates are higher at schools with disadvantaged students (Hanushek, Kain and Rivkin, 2004), resulting in greater disruption to the school and more students served by novice teachers. In addition, male and female teachers of mathematics and science with high ACT scores are more likely to leave teaching, and females with high ACT scores who work with lower ability teachers are more likely to leave (Podgursky, Monroe and Watson, 2004). In addressing the issue of how to attract and retain effective teachers, countries will have to review their compensation systems and incorporate changes that sufficiently reward effectiveness.

SALARY COMPENSATION AND LINKS TO EVALUATION

One of the most radical approaches to compensation systems has been implemented in **Sweden**, where the federal government establishes minimum starting salaries and leaves the decisions about individual teacher salaries to be negotiated annually by the principal and the teacher. If the teacher requests assistance, the teacher union can participate in the negotiation. Many OECD countries, however, rely on a single salary schedule with incentives added for specific objectives, additional work responsibilities or performance.

Problems with the single salary schedule

A number of researchers (Ballou and Podgursky, 1996; Hanushek, 2006; Kowal, Hassel, and Hassel, 2008; Lazear, 2003; Podgursky and Springer, 2006; Umansky, 2005) have documented problems with the single salary schedule. While it appeals to teachers as a fair way to compensate everyone (Heneman, Milanowski and Kimball, 2007) and does not make distinctions that might disturb relationships among teachers, it no longer attracts or retains high-performing teachers. Teachers can predict the high point of their salaries by looking at the salary schedule. If they do not believe that salary is sufficient, they may choose to move into educational administration where salaries are higher or leave the profession.

The single salary schedule ignores the laws of supply and demand. Teachers of fields that are generally in short supply in the economy, such as mathematics and science, have many other options for careers. With a salary schedule that offers small increments for each year of service, these teachers see a greater opportunity cost with each passing year. In addition, there are opportunity costs for high-ability teachers who could find other jobs. Second, if all teachers receive the same pay, there are no incentives to teach in schools that serve disadvantaged students, schools in remote or dangerous locations, or schools in low-wealth districts with lower levels of resources. This situation has led to higher turnover and vacancy rates in these schools (Hanushek, Kain and Rivkin, 2004). In a study of New York City, **New York** (Boyd, Lankford, Loeb and Wyckoff, 2005), high-quartile elementary school teachers were found to be more likely to leave the profession or transfer from low-performing schools after one year than they were to leave schools serving higher-performing students – 34% versus 20%. This compares to 17% of the lowest quartile teachers who left low-performing schools versus 14% who left high-performing schools after one year. Boyd *et al.* also found that over a third of the new teachers had lived outside New York City (NYC) in the **U.S.**, prior to teaching there and were more likely to transfer to another district after their first year of teaching than those who were NYC residents – 10% versus 1%. Such attrition results in schools serving disadvantaged students having higher proportions of new teachers, or hiring teachers without the subject area qualifications, than other schools. Ingersoll (2008) found that over 25% of the core subject classes in high-poverty schools in the United States are taught by teachers who do not have a major or certification in the courses they are teaching, as compared to less than 14% in low-poverty schools.

The third problem with the single salary schedule is that there is no recognition of or reward for effective performance. Equal pay regardless of performance penalises the highly effective teachers who should be earning more and favours the less effective teachers who actually earn more than their efforts deserve. In addition, it is a major factor in fewer high-performing students entering education, and causes many high-performing teachers to leave the field (Podgursky, Monroe and Watson, 2004).

But perhaps the greatest problem with the single salary schedule is that it provides the largest salary increases for the two factors assumed to be proxies for teacher effectiveness: the advanced degrees held and the number of years of experience the teacher has completed. However, researchers (Ballou and Podgursky, 1996; Hanushek, 1996; Hanushek, Kain and Rivkin, 2004; and Hanushek, Kain, O'Brien and Rivkin, 2004) have demonstrated in their studies that masters' degrees and additional credit hours have a weak effect on student outcomes. The exception is the importance of degrees for teachers of mathematics, where Goldhaber and Brewer (2000) documented that **U.S.** teachers of 12th grade mathematics with a bachelor's or master's degree in mathematics are more effective at raising student achievement in mathematics than those without, and that teachers with a bachelor's degree in education have a negative effect on student achievement in mathematics. With respect to experience, beyond the first three to four years, teachers' years of experience do not have a significant impact on student outcomes (Hanushek, 1996; Hanushek, Kain and Rivkin, 2004; and Hanushek, Kain, O'Brien and Rivkin, 2004). Casassus *et al.* (2002), in their Latin American Laboratory Study, found no strong evidence that teacher experience had an impact on student achievement in Latin America. That is not to say that teachers are not important, but that teachers are not compensated based on what makes them effective. In fact, as noted in the beginning of this chapter, an effective teacher can have the greatest of all school effects on student outcomes. A challenge for countries, therefore, is how to manage the transition from the single-salary schedule to a system that can adequately evaluate and reward teachers to foster effective teaching.

One innovation in the design of compensation systems is the use of financial incentives based on student performance on national, state, or other standardised tests. Systems have looked at both school-wide and individual teacher incentives for improved student achievement. Although these are still somewhat controversial, the effectiveness studies have indicated that they can make a difference in student achievement. The challenge is to design the programme so that the possibility of receiving an incentive payment is sufficiently high. Otherwise only those teachers likely to receive one are incentivised. Mizala and Romaguera (2005) found this to be a challenge in the **Chile** SNED programme, where some schools were always in the group to receive incentives while other schools knew they would never be in that group, meaning that the programme was a real incentive only to the schools close to achieving the ranking required. The *Carrera Magisterial* in **Mexico** (Santibáñez *et al.*, 2007; McEwan and Santibáñez, 2005) found the same challenge in that some teachers secured enough points to qualify for the incentives before student achievement was considered while others, even with full credit for high student achievement, could not earn enough points to qualify.

The evaluations of incentive programmes in states and school districts in the **U.S.** are positive in many cases, but the results must be considered in light of the voluntary nature of participation. Do the results mean that the incentives programme has been effective or that effective teachers and schools chose to participate in the incentive programme? The randomised studies conducted in **India** (Muralidharan and Sundararaman, 2008), **Israel** (Lavy, 2002, 2004) and **Kenya** (Glewwe, Ilias and Kremer, 2008) provide better models, and the studies in **India** and **Israel** showed positive effects of the incentive programmes.

In the 1960s and 1970s in the **U.S.**, a number of districts developed merit-pay programmes to supplement the single salary schedule. Teacher merit was primarily based on teacher evaluation, but sometimes using student test scores as well. However, teachers and principals rejected these systems (Murnane and Cohen, 1985) because there were no clear answers to the questions as to why certain teachers received merit pay

while others did not, and what a teacher could do to receive merit pay in the future. Given the complexities of teachers' roles, it is difficult for principals to specify the exact behaviours they are looking for or to determine which teachers have demonstrated those behaviours. Principals were being asked to use the evaluation system to differentiate performance, rather than continue their past practices of giving nearly everyone high ratings that they believed were essential to maintaining collegiality and encouraging teachers to meet the standards of the evaluation. In addition, funds for merit-pay programmes were not stable, so teachers became cynical about the real commitment to support for these programmes, viewing them as another fad that would come and go. Chapter 2 presents some of these challenges and concludes (in the Annex) with a consideration of teachers' attitudes towards rewards and incentives, based on a review of studies conducted in several **U.S.** states and **Australia**. Teachers and teachers' unions, as in the case of the **United Kingdom (U.K.)** reviewed in Chapter 4, often cite the potential negative side-effects of merit pay and financial incentives.

Murnane and Cohen (1985) argued that merit pay could adversely affect the collaborative relationships among teachers, which are critical. The argument continued that merit pay works against intrinsic motivation to act in the social interest and instead causes teachers to focus on the specific skills tested, limit collaboration so that other teachers do not do as well as the individual teacher, or cheat to ensure their students are higher performers. Some teachers respond to merit pay by doing less since they do not see their effort rewarded. Murnane and Cohen (1985) noted that most merit-pay plans lasted less than five years. They studied six plans that had survived up to 20 years to better understand programme characteristics that might account for their longevity. Some of these districts rewarded teachers for taking on additional assignments, others required teachers to develop portfolios documenting why they deserved merit pay, others paid some amount to almost all of their teachers, while still others made the process so demanding that most teachers did not see the value of participation. The most common characteristics were that the plans were developed with significant participation by teachers, were revised as problems arose, and maintained a very low profile. Odden and Wallace (2007) attribute the demise of merit-pay systems in both the private and public education sectors to their lack of alignment to organisational goals, difficulties in evaluation due to poor instruments or lack of evaluator training, low quality of student assessments, poor design and underfunding.⁷

OVERVIEW AND CONTRIBUTIONS OF UP-COMING CHAPTERS

Countries are therefore trying to establish compensation, evaluation, and incentive systems to ensure that effective teachers are recruited, motivated and rewarded. Effective teachers have important effects on the learning outcomes of students, and can furthermore enable economically disadvantaged students to reach achievement levels comparable to their peers (Hanushek, 1992; Sanders and Rivers, 1996; Rivkin et al, 2001). Increased student learning outcomes can lead to economic and social benefits (McKinsey & Company, 2009; Hanushek, 1996). The challenge for countries therefore lies in how to increase learning outcomes of their educational systems in the most effective and resource-efficient manner. The following eight chapters attempt to review experiences from several countries and offer suggestions on how to establish effective mechanisms to identify and reward quality teaching resulting in quality learning outcomes.

Chapter 2 reviews the use of incentives in OECD countries, other countries in Asia and in Latin America and in other parts of the world. While there is widespread use of many types of incentive programmes, rigorous evaluation is only just beginning. This chapter provides descriptions of incentives that offer additional stipends or bonuses to encourage teachers to develop their knowledge and skills, assessed by external examinations or observations by trained evaluators, as well as incentives for teachers who assume additional responsibilities, such as mentoring and coaching other teachers. Many countries have established incentive programmes to meet particular needs over and above their current single salary schedules. These provide incentives to attract students and career changers to preparation programmes, to attract teachers to remote or rural schools, to

schools that are hard to staff because of high percentages of minority or economically disadvantaged students, and to subjects identified as shortage areas, such as mathematics, science, and technology, and, in some places, foreign language and special education.

Chapter 3 discusses the components of compensation systems that include incentives, such as who is to receive the incentives or bonuses and on what basis, what performance measures will be used and how those will be monitored, and how the bonuses will be distributed to individuals or groups of teachers within schools and across the system. There are both advantages and disadvantages to the use of incentives, and this chapter discusses these in the light of implications for the design of incentive systems.

Chapter 4 discusses teacher motivation and provides an argument for the importance of creating incentive systems with the participation of teachers and teacher unions, as opposed to top-down models where governments act alone. Using **England** as a case study, this chapter analyses the last decade of teacher reform efforts. When performance management was linked to school improvement, teachers and head teachers agreed that the appraisal and conference process led to improvements in teacher practice; however when the appraisal was linked to specific student progress objectives and consequences for movement onto the upper salary scales, the response was much more negative. Over 80% of teachers surveyed were opposed to specific student progress targets being used to determine movement to the upper scales of the salary schedule. Based on the National Union of Teachers (NUT) experience with professional development, this chapter suggests that teachers' unions and governments should be able to agree that high-quality, continuing professional development is a major motivator for teachers and can lead to improved teacher practice. The Labour Government's creation of the Union Learning Fund has enabled the NUT to establish peer coaching, best practice research scholarships, learning circles, and study groups that bring teachers together with peers and university tutors to improve practices. The chapter ends by focusing on the importance of effective school leadership to improving school performance – not only the effectiveness of the principal, which is critical, but also the roles of teachers as leaders.

Chapter 5 looks at elements of teacher evaluation that can be used for incentives programmes. While the demand for improved student performance is common across the world, countries have learnt through new data-analysis tools that there are great variations in individual teacher effectiveness. The development and analysis of Value-Added Models is a recent addition to the strategies available to educators seeking to determine the effect individual teachers or schools have on student achievement. As this chapter points out, there are challenges to overcome and the statistical analyses are complex, but it is a method that provides useful information and can be used as one component of evaluation for a compensation system. The question that remains is whether it is an effective method for making high-stakes decisions.

Chapter 6 moves the argument to the need for engaging stakeholders in the design and implementation of a new compensation system. They recommend the formation of a compensation committee that is representative of the various stakeholders and works co-operatively toward a planned design. The committee must be engaged in substantive discussions about the design and must address the areas of disagreement and how they will be handled. The plan should include the development of communication and engagement strategies that will build support for the features of the final design. This chapter includes suggestions for effective communication strategies with a feedback loop that enables the committee to make mid-course corrections. It ends with strategies for evaluating the results of engagement and communications with stakeholders.

Chapter 7 provides a model for aligning a newly-designed incentive system with school improvement initiatives. It discusses the importance of situating the incentive system within a holistic human capital development model and uses the Aspen Institute framework as a model. Thus, incentives become part of the full performance management system, starting with teacher preparation, recruitment, certification, induction,

professional development, evaluation and promotion. This chapter also discusses the influence of classroom, school and district contexts, including the leadership, accountability and support systems, labour relations, human resource programmes, and the level of centralisation in the implementation of incentives. It stresses the importance of making the incentives programme an integral part of school improvement efforts, developing a culture of evaluation, data analysis and data-based decision-making within the schools. Designers should plan on making mid-course corrections to the incentives programme based on the data on student achievement. One final caution offered is the importance of making the incentive programme both programmatically and fiscally sustainable if it is to be part of systemic school reform.

Chapter 8 discusses the politics of teacher incentives, citing the pressures and countervailing pressures on these programmes. These pressures come from the traditional approaches to teacher compensation, norms of teacher autonomy, and the tenets of teacher unions focused on common treatment for all teachers and protection from encroachment of management control. This chapter provides examples of how these pressures have played out differently at different levels of the system – national, state/regional, and local – and in different locations across the world. The chapter ends by discussing political trade-offs that have been made and suggests likely future scenarios for teacher performance incentive systems.

Chapter 9 considers the design of effective evaluation and incentive systems based on the information provided in the previous chapters. It establishes a framework for the design process, identifying first the prerequisite systems relating to evaluation and stakeholder support that must be in place for an incentive programme to be successful. These include stable funding, robust data-management systems, valid and reliable measures of teacher performance, support for improvement of practice, and an effective communications system. To develop the design of the incentive programme, a compensation committee should be formed and should include representatives of both internal and external stakeholders and be led by the leader or a high-profile designee in the education system to demonstrate commitment to the programme. Design decisions will have to be made on a variety of issues from what to reward, to whom to reward, and to the details of designing specific incentive components. The important issue in design is how best to tie financial incentives to practices that improve student learning. This section takes the best of what has been described in the literature and found effective in the evaluation studies available, and applies it to the design issues. The final section of the chapter is devoted to a discussion of implementation issues and the need for evaluation from the start of pilot programmes through all phases of the incentive programme. An effective, on-going evaluation programme will enable the system to monitor impacts of the incentive programme, make mid-course corrections, and continue to improve the quality of teaching and learning in every classroom.

NOTES

1. OECD (2008), *Education at a Glance: OECD Indicators*, OECD, Paris.
2. OECD (2007), p. 20.
3. Adapted from Goe and Stickler, 2008. For a brief description of relevant studies, for example, see Goe, 2007.
4. OECD (2009), p. 158.
5. The Scholastic Aptitude Test is an assessment of high school students' readiness for college in English and mathematics. It is taken in the last 16 months prior to high school graduation and is accepted by most colleges and universities in the United States and abroad as an indicator of likely student success in college.
6. The American College Testing program was started as a competitor to the SAT to focus more attention on the content of the core curriculum (English, mathematics, science and social studies) students should have mastered to be successful in first-year college courses. While it was primarily used by colleges and universities in the Midwest of the United States, it is now used more broadly in the United States and abroad.
7. Odden and Wallace (2008), p. 3. On line at <http://www.freeloadpress.com>.

References

Ballou, D. (2001), "Pay for Performance in Public and Private Schools", *Economics of Education Review*, Vol. 20, No.1, pp. 51-61.

Ballou, D. and **M. Podgursky** (1993), "Teacher Attitudes Toward Merit Pay: Examining Conventional Wisdom", *Industrial and Labor Relations Review*, Vol. 47, No. 1, pp. 50-61.

Ballou, D. and **M. Podgursky** (1996), *Teacher Pay and Teacher Quality*, W.E. Upjohn Institute, Kalamazoo, MI.

Barnes, G., E. Crowe and **B. Schaefer** (2007), "The Cost of Teacher Turnover in Five Districts: A Pilot Study", National Commission on Teaching and America's Future, Washington, D.C., http://www.nctaf.org/resources/demonstration_projects/turnover/documents/CTTFullReportfinal.pdf, accessed March 2009.

Boyd, D., H. Lankford, S. Loeb and **J. Wyckoff** (2005), "Explaining the Short Careers of High Achieving Teachers in Schools with Low-Performing Students", *Understanding Teacher Quality*, Vol. 95, No. 2, pp. 166-171, http://www.teacherpolicyresearch.org/portals/1/pdfs/Explaining_the_short_careers_of_high_achieving_teachers_AER_final.pdf, accessed March 2009.

Casassus, J., S. Cusato, J. Froemel and **J. Palafox** (2002), *First International Comparative Study of Language, Mathematics, and Associated Factors for Students in the third and fourth Years of primary School: Second Report*, UNESCO, Santiago.

Coltham, J. (1972), "Educational Accountability: An English Experiment and its Outcome", *School Review*, Vol. 81, No. 1, pp. 15-34.

Corcoran, S., W. Evans and **R. Schwab** (2004), "Women, the Labour Market and the Declining Relative Quality of Teachers", *Journal of Policy Analysis and Management*, Vol. 23, No. 3, pp. 449-470.

Costrell, R. and **M. Podgursky** (2008), "Peaks, Cliffs, and Valleys: The Peculiar Incentives of Teacher Incentives", *Education Next*, Vol. 8, No. 1, <http://www.hoover.org/publications/ednext/1113017.html>, accessed April 2009.

Danielson, C. (2008), *The Handbook for Enhancing Professional Practice: Using the Framework for Teaching in Your School*, ASCD, Alexandria, VA.

Darling-Hammond, L. (1997), "Doing What Matters Most: Investing in Quality Teaching", National Commission on Teaching and America's Future, New York.

DeArmond, M. and **D. Goldhaber** (2008), "A Leap of Faith: Redesigning Teacher Compensation", School Finance Redesign Project, Center on Reinventing Public Education, University of Washington, Seattle, WA.

Eberts, R., K. Hollenbeck and **J. Stone** (2002), "Teacher Performance Incentives and Student Outcomes", *Journal of Human Resources*, Vol. 37, No. 4, pp. 913-927.

Educational Testing Service (ETS) (2009), *Overview of PRAXIS III*, Educational Testing Service, New York, <http://www.ets.org/portal/site/ets/menuitem.1488512ecfd5b8849a77b13bc3921509/?vgnextoid=f4c20ef8beb4b110VgnVCM10000022f95190RCRD&vgnnextchannel=19b00ef8beb4b110VgnVCM10000022f95190RCRD>, accessed May 2009.

Farkas, S., J. Johnson and **A. Duffett** (2003), "Stand by Me: What Teachers Really Think About Unions, Merit Pay and Other Professional Matters", *Public Agenda Report*, Public Agenda, New York, http://www.publicagenda.org/files/pdf/stand_by_me.pdf, accessed April 2009.

Fuchs, T. and **L. Wößmann** (2004), "What Accounts for International Differences in Student Performance? A Re-examination Using PISA Data", *CESifo Working Paper No. 1235*, CESifo Group, Munich, http://www.cesifo-group.de/pls/guestci/download/CESifo%20Working%20Papers%202004/CESifo%20Working%20Papers%20July%202004/cesifo1_wp1235.pdf, accessed March 2009.

- Glewwe, P., N. Ilias and M. Kremer** (2008), "Teacher Incentives in Developing Countries: Recent Experimental Evidence from Kenya", paper presented at the National Conference on Performance Incentives: Their Growing Impact on American K-12 Education, Nashville, February.
- Goe, Laura** (2007), "The Link Between Teacher Quality and Student Outcomes: A Research Synthesis", National Comprehensive Center for Teacher Quality, Washington, D.C., <http://www.ncctq.org>, accessed October 2009.
- Goe, L. and L. M. Stickler** (2008), "Teacher Quality and Student Achievement: Making the Most of Recent Research", National Comprehensive Center for Teacher Quality, Washington, D.C., <http://www.ncctq.org>, accessed October 2009.
- Goldhaber, D. and D. Brewer** (2000), "Does Teacher Certification Matter? High School Teacher Certification Status and Student Achievement", *Educational Evaluation and Policy Analysis*, Vol. 22, No. 2, pp. 129-145.
- Goldhaber, D., M. DeArmond, A. Liu and D. Player** (2008), "Returns to Skill and Teacher Wage Premiums: What Can We Learn by Comparing the Teacher and Private Sector Labor Markets?", *SFRP Working Paper No. 8*, School Finance Redesign Project, Center on Reinventing Public Education, University of Washington, Seattle, WA.
- Hanushek, E.** (1992), "The Trade-Off Between Child Quantity and Quality.", *Journal of Political Economy* Vol. 100, No. 1, pp. 84-117.
- Hanushek, E.** (1996), "Outcomes, Costs, and Incentives in Schools", in E. Hanushek and D. Jorgenson (eds.), *Improving America's Schools: The Role Of Incentives*, Board on Science, Technology, and Economic Policy, National Academies Press, Washington, D.C., pp. 29-52, <http://www.nap.edu/books/0309054362/html/>, accessed March 2009.
- Hanushek, E.** (2006), "Teacher Compensation" in P. Peterson (ed.), *Reforming Education in Florida*, Hoover Institution Press, pp. 149-163.
- Hanushek, E., J. Kain, D. O'Brien and S. Rivkin** (2004), "The Market for Teacher Quality", *NBER Working Paper No. 11154*, National Bureau of Economic Research, Cambridge, MA.
- Hanushek, E., J. Kain and S. Rivkin** (2004), "Why Public Schools Lose Teachers", *Journal of Human Resources*, Vol. 39, No. 2, pp. 326-354.
- Hanushek, E. and R. Pace** (1995), "Who Chooses to Teach (and Why)?", *Economics of Education Review*, Vol.14, No. 2, pp. 101-117.
- Hanushek, E. and S. Rivkin** (2006), "Teacher Quality", in E. Hanushek and F. Welch (eds.), *Handbook of the Economics of Education, Volume 2*, Elsevier B.V., Amsterdam, pp. 1051-1078, http://edpro.stanford.edu/hanushek/files_det.asp?FileId=188, accessed March 2009.
- Hassel, B.** (2002), "Better Pay for Better Teaching: Making Teacher Compensation Pay Off in an Age of Accountability", *21st Century Schools Project*, Progressive Policy Institute, Washington, D.C.
- Heneman, R.L.** (1992), *Merit Pay: Linking Pay Increases to Performance Ratings*, Addison-Wesley, Reading, MA.
- Heneman, H., A. Milanowski and S. Kimball** (2007), "Teacher Performance Pay: Synthesis of Plans, Research, and Guidelines for Practice", *CPRE Policy Brief RB-46*, University of Pennsylvania, Philadelphia, February.
- Hernani-Limarino, W.** (2005), "Are Teachers Well Paid in Latin America and the Caribbean? Relative Wage and Structure of Returns of Teachers", in E. Vegas (ed.) *Incentives to Improve Teaching: Lessons from Latin America*, World Bank, Washington, D.C., pp.63-96.
- Holmstrom, B. and P. Milgrom** (1991), "Multitask Principal-Agent Analyses: Incentive Contracts, Asset Ownership, and Job Design", *Journal of Law, Economics, & Organization*, 7, pp.24-52. Cited in Neal (2008).
- Hoxby, C.** (1996), "How Teachers' Unions Affect Education Production," *The Quarterly Journal of Economics*, MIT Press, Vol. 111, No. 3, pp. 671-718.
- Hoxby, C.** (2002), "Would School Choice Change the Teaching Profession?", *Journal of Human Resources*, Vol. 37, No. 4, pp. 846-891.

- Ingersoll, R.** (2008), "Core Problems: Out-Of-Field Teaching Persists in Key Academic Courses and High-Poverty Schools", Education Trust, Washington, D.C., <http://www2.edtrust.org/NR/rdonlyres/0D6EB5F1-2A49-4A4D-A01B-881CD2134357/0/SASSreportCoreProblems.pdf>, accessed April 2009.
- Jacob, B.** and **L. Lefgren** (2005), "Principals as Agents: Subjective Performance Measures in Education", *KSG Working Paper No. RWP05-040*, Kennedy School of Government, Harvard University, Cambridge, MA., SSRN: <http://ssrn.com/abstract=757356>, accessed April 2009.
- Jacob, B.** and **M. Springer** (2008), "Teacher Attitudes Toward Pay For Performance: Evidence from Hillsborough County, Florida", paper presented at the National Conference on Performance Incentives: Their Growing Impact on American K-12 Education, Nashville, February.
- Jacobson, S.** (1988), "The Distribution of Salary Increments and its Effect on Teacher Retention", *Educational Administration Quarterly*, Vol. 24, No. 2, pp. 178-199.
- Jacobson, S., E. Hickcox** and **R. Stevenson** (1996), *School Administration: Persistent Dilemmas In Preparation and Practice*, Greenwood Publishing Group, Santa Barbara, CA.
- Kellor, E.** (2005), "Catching Up with the Vaughn Express: Six Years of Standards-Based Teacher Evaluation and Performance Pay", *Education Policy Analysis Archives*, Vol. 13, No. 7 (January 23), <http://epaa.asu.edu/epaa/v13n7/>, accessed March 2009.
- Kershaw, J.** and **R. McKean** (1962), *Teacher Shortages and Salary Schedules*, McGraw-Hill, New York.
- Koppich, J.** (2008), "Toward a More Comprehensive Model of Teacher Pay", paper presented at the National Conference on Performance Incentives: Their Growing Impact on American K-12 Education, Nashville, TN, February.
- Kowal, J., B. Hassel** and **E. Hassel** (2008), "Financial Incentives for Hard-to-Staff Positions: Cross-Sector Lessons for Public Education", Center for American Progress, Washington, D.C.
- Lakdawalla, D.** (2001), "The Declining Quality of Teachers", *NBER Working Paper No. 8263*, National Bureau of Economic Research, Cambridge, MA.
- Lavy, V.** (2002), "Evaluating the Effect of Teachers' Group Performance Incentives on Pupil Achievement", *Journal of Political Economy*, 110 (6), pp.1286-1317.
- Lavy, V.** (2004), "Performance Pay and Teachers' Effort, Productivity and Grading Ethics", *NBER Working Paper No. 10622*, National Bureau of Economic Research, Cambridge, MA.
- Lawler, E.** (1990), *Strategic Pay: Aligning Organizational Strategies and Pay Systems*, Josey-Bass, San Francisco.
- Lazear, E.** (1996), "Performance Pay and Productivity", *NBER Working Paper No. W5672*, National Bureau of Economic Research, Cambridge, MA., SSRN: <http://ssrn.com/abstract=225573>, accessed May 2009.
- Lazear, E.** (2003), "Teacher Incentives", *Swedish Economic Policy Review*, Vol. 10, pp. 179-214.
- Lewis, J.** and **M. Springer** (2008), "Performance Incentives in Texas: Why Schools Chose Not to Participate", paper presented at the National Conference on Performance Incentives: Their Growing Impact on American K-12 Education, Nashville, TN, February.
- Manzi, J.** (2008), "Individual Incentives and Teacher Evaluation: The Chilean Case", paper presented at the International OECD/Mexico Joint Conference, The Quality of Education, Mexico City, December.
- McEwan, P.** and **L. Santibáñez** (2005), "Teacher and Principal Incentives in Mexico", in E. Vegas (ed.) *Incentives to Improve Teaching: Lessons from Latin America*, World Bank, Washington, pp. 213-253.
- McKean, R.** and **J. Kershaw** (1961), *What to Do About Teacher Shortages*, RAND Corporation, Santa Monica, CA.
- McKinsey & Company** (2009), *The Economic Impact of the Achievement Gap in American Schools*, McKinsey Social Sector Office, New York
- Milanowski, A.** (2004), "The Criterion-Related Validity of the Performance Assessment System in Cincinnati", *Peabody Journal of Education*, Vol. 79, No. 4, pp. 33-53.
- Milanowski, A.** (2006), "Performance Pay System Preferences of Students Preparing to Be Teachers", *WCER Working Paper No. 2006-8*, Consortium for Policy Research in Education, Wisconsin Center for Education Research, Madison, WI.

- Milanowski, A.** and **S. Kimball** (2005), "The Relationship between Teacher Expertise and Student Achievement: A Synthesis of Three Years of Data", paper presented at the American Educational Research Association Meeting, Montreal, Quebec.
- Milanowski, A., S. Kimball** and **A. Odden** (2005), "Teacher Accountability Measures and Links to Learning", in L. Stiefel, A. Schwartz, R. Rubenstein and J. Zabel (eds.), *Measuring School Performance and Efficiency: Implications for Practice and Research*, Eye on Education, Larchmont, NY, pp. 137-161.
- Milanowski, A., H. Longwell-Grice, F. Saffold, J. Jones, A. Odden** and **K. Schomisch** (2007), "Recruiting New Teachers to Urban School Districts: What Incentives Will Work", *Working Paper 11*, School Finance Redesign Project, Center on Reinventing Public Education, University of Washington, Seattle, WA.
- Mizala, A.** and **P. Romaguera** (2005), "Teachers' Salary Structure and Incentives in Chile" in E. Vegas (ed.) *Incentives to Improve Teaching: Lessons from Latin America*, World Bank, Washington, D.C., pp. 103-150.
- Mohrman, S.A.** and **E. Lawler** (1996), "Motivation for School Reform", in S. H. Fuhrman and J. O'Day (eds.) *Rewards and Reform: Creating Educational Incentives that Work*, Jossey-Bass, San Francisco, pp. 115-143.
- Murnane, R.** and **D. Cohen** (1985), *Merit Pay and the Evaluation Problem: Understanding Why Most Merit Pay Plans Fail and a Few Survive*, Institute for Research on Education Finance and Governance, Stanford University, Palo Alto, CA., http://www.eric.ed.gov:80/ERICDocs/data/ericdocs2sql/content_storage_01/0000019b/80/2f/4e/ca.pdf, accessed March 2009.
- National Board for Professional Teaching Standards, (NBPTS)** (2009), "About the National Board for Professional Teaching Standards", NBTS, Arlington, VA., http://www.nbpts.org/about_us, accessed April 2009.
- National Center for Education Statistics (NCES)** (2003), "NAEP Trends in Academic Progress Through 1999" in NCES Digest of Education Statistics 2003, NCES, Washington, D.C.
- National Research Council** (2008), *Assessing Accomplished Teaching: Advanced-Level Certification Programs*, in M. Hakel, J. Anderson Koenig and S. Elliott, (eds.) Committee on Evaluation of Teacher Certification by the National Board for Professional Teaching Standards, NRC, National Academies Press, Washington, D.C.
- Neal, D.** (2008), "Designing Incentive Systems for Schools", paper presented at the National Conference on Performance Incentives: Their Growing Impact on American K-12 Education, Nashville, TN, February.
- Odden, A.** and **C. Kelley** (2002), *Paying Teachers for What They Know and Do: New and Smarter Strategies to Improve Schools*, Corwin Press, Thousand Oaks, CA.
- Odden, A.** and **M. Wallace** (2007), *How to Achieve World Class Teacher Compensation*, Freeload Press, St. Paul, MN., <http://www.freeloadpress.com/bookDetail.aspx?bld=1077>, accessed March 2009, free registration required.
- OECD** (2005), *Teachers Matter: Attracting, Developing and Retaining Effective Teachers*, OECD, Paris.
- OECD** (2007), *PISA 2006: Science Competencies for Tomorrow's World*, OECD, Paris.
- OECD** (2008), *Education at a Glance: OECD Indicators*, OECD, Paris.
- OECD** (2009a), *Creating Effective Teaching and Learning Environments: First Results from TALIS*, OECD, Paris.
- OECD** (2009b), "Teacher Evaluation: Current Practices in OECD Countries and a Literature Review", OECD Education Working Papers, No. 23, OECD publishing, OECD.
- Podgursky, M.** (2006), "Teams Versus Bureaucracies: Personnel Policy, Wage-Setting, and Teacher Quality in Traditional Public, Charter, and Private Schools", paper for the National Conference on Charter School Research at Vanderbilt University, Nashville, TN.
- Podgursky, M., R. Monroe** and **D. Watson** (2004), "The Academic Quality of Public School Teachers: An Analysis of Entry and Exit Behavior", *Economics of Education Review*, Vol. 23, No. 5, pp. 507-518
- Podgursky, M.** and **M. Springer** (2006), "K-12 Public School Finance in Missouri: An Overview.", Regional Economic Development. Federal Reserve Bank of St. Louis, Vol. 2, No. 1, pp. 31-50.
- Prendergast, C.** (1999), "The Provision of Incentives in Firms", *Journal of Economic Literature*, Vol. 37, No. 1, pp. 7-63.

- Rivkin, S., E. Hanushek and J. Kain** (2001), "Teachers, Schools, and Academic Achievement", *Working Paper 6691 (revised)*, National Bureau of Economic Research, Cambridge, MA.
- Rothstein, R.** (2008), "Holding Accountability to Account: How Scholarship and Experience in Other Fields Inform Exploration of Performance Incentives in Education", paper presented at the National Conference on Performance Incentives: Their Growing Impact on American K-12 Education, Nashville, TN, February.
- Sanders, W. and J. Rivers** (1996), "Cumulative and Residual Effects of Teachers on Future Student Academic Achievement", *Research Progress Report*, University of Tennessee Value-Added Research and Assessment Center, Knoxville, TN.
- Santibáñez, L., J. Martínez, A. Datar, P. McEwan, C. Setodji, and R. Basuto-Dávila** (2007), "Breaking Ground: Analysis of the Assessment System and Impact of Mexico's Teacher Incentive Program 'Carrera Magisterial'", *RAND Technical Report*, RAND Corporation, Santa Monica, CA.
- Schacter, J. and Y. Thum** (2003), "Paying for High- and Low-Quality Teaching", in *Economics of Education Review*, Vol. 23, Issue 4, pp. 411-430.
- Sclafani, S.** (2008), "Rethinking Human Capital in Education: Singapore as a Model for Teacher Development", paper prepared for The Aspen Institute Education and Society Program, Aspen Institute, Washington, D.C., <http://www.aspeninstitute.org/sites/default/files/content/docs/education%20and%20society%20program/SingaporeEDU.pdf>, accessed May 2009.
- Sclafani, S. and M. Tucker** (2006), "Teacher and Principal Compensation: An International Review", Center for American Progress, Washington, D.C., http://www.americanprogress.org/issues/2006/10/teacher_compensation.html/pdf/education_report.pdf, accessed May 2009.
- Skillbeck, M. and H. Connell** (2004), "Teachers for the Future: The Changing Nature of Society and Related Issues for the Teaching Workforce," Ministerial Council on Education, Employment, Training and Youth Affairs, Canberra, http://www.teachersforthefuture_file.pdf, accessed July 2009.
- Springer, M., M. Podgursky, J. Lewis, M. Ehlert, C. Gardner, B. Ghosh-Dastidar, O. Lopez, C. Patterson and L. Taylor** (2007), *Governor's Educator Excellence Grant (GEEG) Program: Year One Evaluation*, National Center on Performance Incentives, Vanderbilt Peabody College, Nashville, TN, <http://ritter.tea.state.tx.us/opge/progeval/TeacherIncentive/index.html>, accessed April 2009.
- Springer, M., M. Podgursky, J. Lewis, M. Ehlert, B. Ghosh-Dastidar, T. Gronberg, L. Hamilton, D. Janssen, O. Lopez, B. Stecher and L. Taylor** (2008), *Texas Educator Excellence Grant (TEEG) Program: Year Two Evaluation*, National Center on Performance Incentives, Vanderbilt Peabody College, Nashville, TN., http://ritter.tea.state.tx.us/opge/progeval/TeacherIncentive/TEEG_120108.pdf, accessed April 2009.
- Stinebrickner, T., B. Scafidi and D. Sjoquist** (2004), "Do Teachers Really Leave for Higher Paying Jobs in Alternative Occupations?" *Andrew Young School of Policy Studies Research Paper Series No. 06-50*, http://papers.ssrn.com/sol3/papers.cfm?abstract_id=902027, accessed April 2009.
- The System for Teacher and Student Advancement** (2009) "About Us", <http://www.tapsystem.org/about/about.taf?page=niet>, accessed July 2009.
- Umansky, I.** (2005), "A Literature Review of Teacher Quality and Incentives: Theory and Evidence", in E. Vegas (ed.) *Incentives to Improve Teaching: Lessons from Latin America*, World Bank, Washington, D.C., pp. 21-62.
- Vegas, E. and I. Umansky** (2005), "Improving Teaching and Learning Through Effective Incentives", in E. Vegas (ed.) *Incentives to Improve Teaching: Lessons from Latin America*, World Bank, Washington, D.C., pp.1-20.

CHAPTER 2

A Review of International Practice

Financial incentives	30
• Incentives for teacher education	30
Incentives based on teacher's knowledge and skills	31
• Evaluation of teacher performance	31
Incentives for teaching in shortage subject areas	34
Incentives for teaching in schools in difficult environments	34
Other financial incentives	37
Incentives based on performance in the classroom measured by student results on external examinations	38
Non-financial incentives	41
Evidence of effectiveness	44
Evaluations of national and state/regional programmes	44
• Experimental studies of incentives	49
Annex 2A: Incentives in education	52
Notes	59
References	60

The single salary schedule has been the dominant form of compensation for teachers across the world for over half a century. Based on degrees earned and years of experience, the single salary schedule provides a secure salary with small annual increases for remaining on the job, regardless of the quality of the teacher's performance. Differences in salary levels have been negotiated with trade unions at the national, regional and local levels but they have generally affected the level of pay rather than the structure. In many countries, the teacher unions have been instrumental in maintaining the single salary schedule, with a few salary incentives, but no recognition for performance-based pay or incentives for those working in schools serving disadvantaged students. In **Ireland**, for example (Sclafani and Tucker, 2006), assuming additional responsibilities such as teaching the Irish language or working in remote locations, are the only acceptable reasons for teachers to receive additional pay.

In the last 15 years, both developed and developing countries have experienced shortages of qualified teachers in certain subjects or types of schools, such as those serving disadvantaged students, leading them to develop incentive programmes aimed at improving individual teacher performance. The incentive programmes may be based on teachers' knowledge and skills, the amount of extra responsibilities they have assumed or the quality of their work. In these areas, however, the additional salary is above the regular single salary schedule amount, ensuring that teachers are not at risk of earning less than they would have, based on the degrees held and the number of years of experience. **Sweden** is the exception to this as in the mid 1990s it moved to a centrally-developed minimum first-year salary with all increases from that point developed through individual pay negotiations between the school and the individual teacher.¹ This chapter characterises the varied ways in which both developed and developing countries are addressing teacher compensation issues through incentives. To assist the reader, Annex 2A at the end of the chapter provides a description of different incentives, and also presents results from studies on teachers' attitudes toward incentives.

FINANCIAL INCENTIVES

Incentives for teacher education

Many countries see scholarships or stipends for teacher candidates as an incentive to attract new teachers into the profession, but it may be a wasted expense for a government if students use the funds to complete their education but do not actually enter teaching. As a case in point, in the State of **Victoria** in **Australia**, out of every 100 students who applied for teacher education programs in 1999, 56 were accepted, 41 enrolled, 31 were expected to graduate, 23 were available for employment, and only 15 were employed as teachers in schools.² Despite that, a number of countries, including Australia, have chosen to use this strategy to increase the number of young people preparing to become teachers, especially those pursuing majors in teacher shortage subject areas, such as mathematics, science and technology. For example, the **Western Australia** Department of Education and Training (Skillbeck and Connell, 2004) provides teacher education students in subject areas of need (e.g. mathematics, science and technology), a scholarship of AUD 10 000 in their last year of study in return for a two-year commitment to teach in a rural school. It recently added payment of university fees for students studying to be science teachers. The **Northern Territory of Australia** provides a training bursary for students studying to be teachers in shortage subject areas, including information technology, mathematics, science, and special education.

Ross and Hutchings (2003) report that both **England** and **Wales** developed a number of financial incentives for students studying to be teachers, starting in 1986 with grants for students in certain secondary education fields. This evolved into the Shortage Subject Support Scheme that provides up to GBP 5 000 for students facing financial hardships and who are studying a secondary subject identified as a national teacher shortage area. Since 1998, prospective teachers entering a postsecondary training programme receive a GBP 6 000 training bursary with half paid at the completion of the first module and the other half paid once the student has achieved Qualified Teacher Status at the end of the programme. Developing countries such as **Chile** and **Uruguay** (Vegas, 2007) also provide scholarships. In **Chile**, talented students in teacher education can receive

a scholarship covering all of their tuition in return for committing to teach for three years. The priority subject areas are mathematics, science, English, language arts and primary education. **Uruguay** provides free tuition for all teacher education students, but also supplies scholarships for living expenses to attract talented students from disadvantaged backgrounds into teacher education programs.

Singapore (Sclafani, 2008) goes farther than other countries, offering free tuition and fees to prospective teachers, along with a monthly stipend for bachelor degree candidates and a monthly salary equal to a college graduate in a civil service position for those entering the postgraduate programme. To ensure that it does not lose money on students who later choose other professions, Singapore requires that any teacher who does not complete the three to six years of teaching required in the contract for the tuition and stipend support must pay liquidated damages with interest. This is true regardless of whether the individual or the school system determines that the teacher is not a good fit with the system.

The **United States (U.S.)** has established a number of programmes at the federal and state levels that offer scholarships or loan forgiveness to attract high-performing teacher education students to either shortage subject areas or to hard-to-staff schools. Starting in the Sputnik era (from the launch of Sputnik in 1957 through the decade of the 1960s), the National Defense Education Act offered loan forgiveness of USD 1 000 per year for up to five years to teachers of foreign language, mathematics, and science in schools serving low-income students. The Stafford Loan programme, established in 2004, provides up to USD 17 500 in loan forgiveness for teachers of mathematics, science or special education. Recently the U.S. Department of Education established the Teacher Education Assistance for College and Higher Education grant programme, which provides a grant of USD 4 000 per year to teacher candidates who agree to teach in shortage subject areas for four years in schools serving disadvantaged students. If the teachers do not fulfill their commitment within eight years, the grant converts into a loan that must be repaid. According to Johnson (2005), by 2004, 24 states had established loan forgiveness or scholarship programs to attract teachers in shortage subject areas, and 12 states had established them for hard-to-staff schools. The State of **Oklahoma** has a loan forgiveness programme for teachers of mathematics or science who agree to teach for a minimum of five years, while **Mississippi** requires three years of teaching for teachers who have received four years of funding from the state to become teachers in geographic or subject shortage areas (Chait, 2007). The State of **Arkansas** provides up to USD 3 000 to current teachers to go back to school to earn certification in a shortage area.³

INCENTIVES BASED ON TEACHER'S KNOWLEDGE AND SKILLS

A number of countries that wished to reward teachers for increasing their effectiveness have focused on rewarding new knowledge and skills, with programmes setting clear expectations about what teachers must do to improve their skills in priority areas. Odden *et al.* (2001) differentiate knowledge and skills-based pay from merit pay that used school principal evaluations as the method to determine how to divide a fixed sum of funds for teachers. They also differentiate knowledge and skills-based pay from career ladders that required teachers to assume additional responsibilities while remaining in the classroom for part or most of the day. Knowledge and skills-based compensation acknowledges that teachers develop new knowledge and skills throughout their career, and these should be rewarded upon clear documentation of achievements. In some cases, this means that teachers move up to a new category of the salary schedule or that they receive a one-time or annual bonus for demonstrated competencies. In the **U.S.**, this is done through external evaluations, such as the National Board for Professional Teaching Standards or PRAXIS III or school-based evaluations based on rubrics.

Evaluation of teacher performance

Based on the negotiations between the Education Ministry and the Teachers Association of **Chile**, ongoing from 1991 to 2002, Chile created the Pedagogical Excellence Reward to recognise excellent teachers. The reward is

based on individual evaluations using The Framework for Good Teaching (*El Marco para la Buena Enseñanza*), but not on student test scores. Teachers can volunteer for an evaluation twice within each of four levels of years of experience. The evaluation is made up of two parts, a written test of pedagogical and content knowledge, and a portfolio of classroom teaching, including a video observation. Teachers may be awarded additional salary for excellence. A quota is established each year, based on budgets. Once teachers receive the reward they keep it as long as their evaluations are satisfactory and they are in the same strand of experience. The programme includes both primary and secondary teachers.⁴

In some states in **Germany**, teachers are able to move faster through the salary levels, based on their performance evaluation. In **Baden-Wuerttemberg**, each year 10% of the teachers in each school can move up a step on the salary schedule for outstanding performance. In addition, poor evaluations can delay movement to the next step. In **North-Rhine Westphalia**, the step increases are no longer automatic; teachers' step advancement depends upon the evaluation of their performance every two to four years.⁵

The Netherlands has allowed local school boards to establish personnel management systems that include opportunities for a limited number of teachers to move to a higher salary scale or to move more rapidly through the steps in their current scale. School-based personnel budgets give schools greater discretion to award performance-related allowances or bonuses to teachers. Schools decide the conditions under which bonuses or allowances will be granted and how much money they are prepared to spend on them.⁶

Singapore may have the most extensive performance management system, providing bonuses equivalent to one to three months' salary based on their rating on an annual evaluation. Both teachers and evaluators take evaluation very seriously as part of the continuous improvement focus in Singapore. The evaluation process is quite extensive, with a planning meeting at the beginning of the school year to set goals for student achievement, professional development, and contributions to the school and community. It is followed by a review meeting midyear and a final evaluation based on portfolios of work, as well as input from the senior teachers and department/subject area chairs who have worked with the teacher. In addition, teachers may earn bonuses for outstanding individual or team contributions to creativity, cost savings, or peer support, as documented in the evaluation portfolio.

In some regions of **Switzerland**, teachers must successfully complete self and external evaluations before advancing to the next stage of the salary scale. The external evaluations use a broad range of criteria to determine teacher eligibility.⁷ **Zurich** established the "salary-effective qualification system" in 1999, which applies only to teachers in the middle phase of their career. It is not available to teachers at the end of their careers, except in cases of truly outstanding evaluations. In the middle phase, teachers are assessed and, if the evaluation is positive, the teacher may receive a 1-3% salary increase for each of the next four years. A specially trained team of representatives of the school committee conducts the evaluation. The team observes the teacher in the classroom, interviews the teacher, and reviews a report prepared by the teacher describing pedagogical approaches employed in the classroom. If the evaluation decision is negative, the teacher may apply again after a year of work on the areas of identified weakness, but he/she receives no promotion to higher salary until a successful evaluation is achieved.

In **Switzerland** as well, **St. Gallen** initiated the "systematic salary-effective qualification" in 2000. The salary schedule has four grades with steps within each grade. Salary increases from step to step are automatic, but not from grade to grade. When a teacher has completed a grade, he or she is evaluated for promotion to the next grade. Without a successful evaluation, the teacher cannot be promoted to the next grade or receive a salary increase. The teacher meets with an evaluator from the school committee to agree upon the criteria for the evaluation, which includes both self-assessment and external assessment in three areas: "organisation and delivery of lessons; interactions with students, teachers and parents; and participation in in-service training".⁸

The Cincinnati Public Schools in **Ohio** (Odden and Wallace, 2008) established a complex knowledge and skills-based compensation system that included five career levels, with two to three salary steps in each level, based on 16 teaching standards. Teachers moved from Apprentice to Novice levels by passing the PRAXIS III exam with a rating of at least “2” in each area of evaluation. Novice teachers had to move within five years to Career status by receiving a minimum rating of “3” in each evaluation area to remain employed. The programme had a permanent salary increase for teachers with one additional degree or certification, and annual supplements that were valid for one to three years based on demonstration of skills. Although the pilot plan was approved by the teachers’ union and the board of education, it lasted only three years before the teachers voted against the continuation of the plan. Soon after, the superintendent left and school board members in favour of the programme were replaced through elections. One of the major complaints was that teachers at the high end of the previous salary schedule did not have their salary protected under the new system, and those teachers lost up to USD 10 000 each. In addition, teachers did not believe the evaluation system was fair and found it to be too complicated for them to understand.

In the **United Kingdom (U.K.)** (Sclafani and Tucker, 2006), the top of the salary scale can be reached in five steps, considerably less time than in many other countries. In 2000, the School Teachers’ Review Body came out in support of the government’s proposal for performance-related pay for an upper scale of the salary schedule, based upon eight nationally agreed upon teaching standards. Teachers at the top of the main pay scale may apply to be assessed against the standards to pass beyond “the threshold” to the upper pay scale. The standards are divided into five groups: knowledge and understanding (one standard); teaching and assessment (three standards); pupil progress (one standard); wider professional effectiveness (two standards); and professional characteristics (one standard). Principals in **England** and **Wales** were trained to complete the assessment, with the results confirmed by an external assessor. This marked the first use of student performance as part of the evaluation of teachers. Approximately 80% of the eligible teachers applied, and around 95% of them moved forward to step one of the upper scales in 2002.⁹ In the upper pay scales, teachers earn from 12-25% higher salaries. Teachers normally progress along the upper pay scale every two years based on annual evaluations, although progression is not automatic. Teachers who have crossed the threshold must display substantial and sustained performance and contributions to the school before a performance point is awarded.

Additional pay for additional instructional responsibilities. In 1990, the Schools Council for the National Board of Employment, Education and Training in **Australia** recommended the creation of a national classification, Advanced Skills Teacher. Not all states, however, established the Advanced Skills Teachers (AST) programme. **New South Wales** began the programme in 1992, selecting ASTs based on their contributions to professional development of other teachers, curriculum development, student progress, and involvement in improving the educational outcomes of students, and supporting beginning teachers with classroom performance. They were required to operate under the same workload parameters as all other teachers, and to take on a number of added activities and responsibilities. The term of the AST status was one-year and the stipend was approximately AUD 1 200.¹⁰ Other states in Australia have added this category to their salary systems.

In 1998, the governments of **England** and **Wales** also created an Advanced Skills Teacher (AST) programme to provide a career path for excellent teachers who did not want to leave the classroom to take management positions. It was envisioned that 3-5% of teachers would receive this designation. Candidates for AST submit a portfolio demonstrating they have met the standards for the position. The portfolio must be evaluated as successful by an external assessor in order for the teacher to be appointed to an AST post. In that role, they share their skills in teaching practices through outreach work in many schools. Normally they are required to spend 80% of their time teaching in their own classes and 20% sharing good teaching practices with teachers in other schools. AST posts are funded by local authorities in the **U.K.**, or by specific individual schools. There is a separate salary scale for ASTs, and upon appointment as an AST, one can move to the bottom of a five-point

range of pay. Pay also varies according to geographic boundaries, defined as: England and Wales (excluding London), Inner London Area, Outer London Area, and Fringe Area. Teachers at the top of the AST salary scale can earn 43% more than senior teachers at the top of their salary schedule.¹¹

Experienced teachers in the **Netherlands** can take on a different role as a Plus Teacher, replacing ill teachers and also coaching newer teachers while at the school, and receive a bonus for taking on these responsibilities. Such positions provide a variation in routine as well as the challenge of mentoring newer staff.¹² In **New Zealand**, each secondary school can name a specialist teacher who receives an additional NZD 6 300 salary and a four- or eight-hour reduction in teaching duties weekly, depending on the size of the school, for assuming responsibilities for professional development and mentoring other teachers.¹³

INCENTIVES FOR TEACHING IN SHORTAGE SUBJECT AREAS

Students are short-changed when their teachers are not qualified to teach the specific subjects they are teaching. Although the concept of supply and demand is clearly applicable in education, teacher unions in the **U.S.** as well as other countries have remained opposed to the idea of using incentives to attract qualified teachers of specific subjects. They claim that suggesting that mathematics or science teachers are somehow more important than English or social studies teachers will create dissension among teachers, who are a highly democratic group. However, in the U.S., states without collective bargaining have been more active in establishing such programmes. For example, 43% of districts in **Texas** offer incentives for teachers in shortage subject areas. **Utah** offers a USD 5 000 signing bonus for teachers of mathematics and science who stay in the district for at least four years (National Council on Teacher Quality, 2008). Many school districts in the U.S. offer additional stipends of USD 1 500 to USD 3 000 for shortage area teachers. New York City offers housing assistance of up to USD 15 000 in moving expenses, down payments on homes or security deposits on apartments, and USD 400 monthly housing stipends for two years to attract experienced teachers of mathematics, science and special education.

School systems in other OECD countries have developed financial incentives to attract and retain teachers in shortage areas (Sclafani and Tucker, 2006), such as teaching in indigenous languages in **Australia** and **Ireland** or teaching in French in **Brussels**. In addition, **England** and **Wales** have established loan forgiveness programmes for teachers of mathematics, science, special education and technology. The programmes repay all teacher preparation tuition expenses up to GBP 16 000 over ten years of teaching. They also offer a Golden Hello programme that provides GBP 4 000 for successfully completing the first year of teaching in shortage areas. **St. Gallen** is the only canton in **Switzerland** that provides salary allowances for teaching subjects facing shortages. Despite these efforts, in 2003, principals in countries participating in an OECD study of upper secondary schools reported that over 30% of their students attend schools where there are teacher shortages in foreign languages, mathematics, science, and technology.

INCENTIVES FOR TEACHING IN SCHOOLS IN DIFFICULT ENVIRONMENTS

Remote schools. Many countries face challenges in attracting teachers to teach in schools in remote areas. If the salary schedule is uniform across a country, there is little incentive for teachers from metropolitan areas to move to remote and isolated areas. Therefore, some countries have initiated monetary or promotion incentives. Although there has been no evaluation of the effect of providing incentives for teachers to move to remote schools, countries can track the degree to which qualified teachers staff these schools. One must assume that national or provincial/state governments have established the levels of their stipends in response to supply and demand.

In **Australia**, the **Queensland** Remote Area Incentive Scheme (Sclafani and Tucker, 2006) provides cash awards of up to AUD 5 000 for the teacher to travel from their remote area to a metropolitan area for shopping or visiting

family, as well as retention stipends of up to AUD 5 000 per year. **Denmark** provides free accommodations in remote and sparsely populated areas as well as a home personal computer and access to wholesale shopping clubs, while **Ireland** provides an annual stipend of EUR 1 321 to teachers in schools on remote islands. **New Zealand** offers a High Priority Teacher Supply Allowance (HPTSA) of NZD 2 500 for teaching in schools designated as needing special assistance in recruitment and retention. The HPTSA teacher receives transfer and moving allowances to go to the school and to move to another school after having taught there for three years. Countries are also using larger incentives in more remote areas. For example, **Nunavut**, a remote province in **Canada**, provides significant incentives for teachers in its communities, ranging from CAD 12 100-28 350 per year and an additional daily housing stipend of CAD 50-75. Some countries make service in remote schools a prerequisite for promotion, as in **Korea** and **China** where teachers who spend time in schools in remote areas get credit toward promotion to positions of senior teachers or vice-principals.

In Latin America, **Chile**, **Peru** and **Bolivia** have financial incentives to encourage teachers to serve in rural schools. **Chile** (Sclafani and Tucker, 2006) provides geographic incentives of as much as 30% of annual salary for working in areas that are far from urban areas or represent geographic isolation, extreme poverty or difficult access. **Peru** (Crouch, 2005) offers rural bonuses as one of nearly 20 incentives and bonuses that can be added to the basic salary and that can represent 95% of the total salary. Since basic salaries in real dollars are only 50% of what they were in the 1970s, 5% of basic salary for teaching in a rural school does not provide much of an incentive to teachers who would prefer to live in or near metropolitan areas. A further complication is that teachers maintain the bonus even though the school would no longer qualify as rural due to urbanisation, or the teacher may have transferred to a non-rural school.

Bolivia (Urquiola and Vegas, 2005) bases salaries on a complex matrix of school designations of urban, provincial and rural, credentials earned in teacher preparation programmes, and seniority. Teaching in a rural school provides a 20% increase in initial salary placement over teaching in a provincial and urban school at each level of training credentials. The levels of training range from interim (no training) to experienced-certified (no formal training but more than nine years of teaching), ABD (completed training but missing a requirement for certification) and certified (graduated from teacher training and completed two years of service in a rural or provincial school). Entry into the highest salary levels is dependent on being fully certified. When experience is factored into the salary equation, location represents only about 9% of the total salary. There are two additional bonuses for teaching in rural areas: the Zone or Frontier incentives of 20% of monthly base salary for teaching in inaccessible areas or within 50 kilometers of international borders. The IPR (*Incentivo a la Permanencia en el Area Rural Pobre*), a permanent rural incentive, provides an increase of about 10% of monthly base pay for teaching and remaining in poor rural areas. Adding these bonuses and incentives to the maximum incentives and bonuses a teacher can earn for all categories of experience, location, seniority, training, and other behaviors, makes location worth approximately 12.5% of total salary, probably not high enough to really attract teachers to the most remote areas of Bolivia. In 2001 the Ministry abolished the requirement for teaching in provincial or rural schools to attain full certification. While this enabled more teachers to enter the higher salary ranks, it also removed the strongest incentive for teachers to serve in rural schools. As in **Peru**, Bolivia still has schools designated as rural despite urbanisation and teachers designated as rural despite transfers to urban schools, probably due to reluctance to re-categorise schools and teachers and lower the salary levels of the teachers involved.

Hard-to-staff schools. While remote schools pose a significant challenge in attracting teachers, urban and suburban schools serving disadvantaged students face similar challenges. Given choices among positions, teachers often select the schools with the best principals, most engaged parents, and students who come to school ready to learn. **France** has established programmes to recruit experienced teachers to schools serving disadvantaged students in the suburbs of Paris by establishing priority posts that provide additional benefits

regarding placement, training and career progression. These programmes did increase the number of applicants who were fully qualified, but only 40% of them were experienced teachers. The teachers union has been adamant that the current point system of transfers, that favours seniority among other factors, not be changed.¹⁴ In **New Zealand**, a Staffing Incentive Allowance of NZD 1 000 is available for teachers for up to three years for teaching in schools designated by the Secretary of Education as priority schools.¹⁵ London (Ladd, 2007) has raised salaries in its schools by approximately 12% in order to attract teachers, but still has vacancy rates that are higher than other areas of the country. Goldhaber (2008) reports on incentive programmes in the **U.S.**, and identifies 20 states that have established incentive programmes for teachers who are willing to go to schools with staffing challenges. In addition, some states established programmes to attract teachers to low-performing schools, offering annual cash incentives, loan forgiveness programmes and housing allowances. **Arkansas** provides USD 4 000 stipends for the first year, USD 3 000 for the next two years and USD 2 000 per year starting in year four. The state credits this programme with reducing first-year teacher attrition from 19.7% in 2001-02 to 7.1% in 2007-08 (Arkansas Department of Education, 2009).

Many large districts also provide stipends of from USD 2 000 in Los Angeles, **California** and Charlotte, **North Carolina** to USD 3 000 in New York City. Fairfax County, **Virginia** and Miami, **Florida** provided percentage increases of 12% and 20% respectively for working in the lowest-performing schools and required that the teacher workday be extended by one hour (Odden and Wallace, 2007). Guilford County Schools, **North Carolina**, established a programme for its 30 schools with the greatest staffing problems; the programme offered recruitment and retention stipends ranging from USD 2 500 for primary school teachers to USD 10 000 for algebra teachers. In addition, teachers in the state's Mission Possible Schools can earn high student achievement awards of USD 2 500-4 000. Ingersoll (2008) reported that in the first year of the programme, the number of applicants for mathematics positions increased from 7 to 174.

Milanowski *et al.* (2007) conducted focus groups at three teacher training institutions, a large public urban, a small private urban and a large public non-urban, soliciting information as to what would influence prospective teachers' choices of attractive or unattractive schools in which to teach. A survey was developed from the focus group information and administered to over 200 students at the three institutions. Responses indicated that a school with support teachers and a principal known to understand teaching and learning would be the greatest influence on their choice of that school, increasing the probability of taking the job by 17%. In addition, curricular flexibility increased the probability by 10% and a good induction programme by 9%, while a USD 5 000 increase in starting salary raised it less than 4%. The percentage of minority students at the hypothetical school had a very small effect on job choice, equal to about a USD 365 salary increase.

When the analysis was done within each type of institution, large public urban, small private urban and large public non-urban, Milanowski *et al.* (2007) found that responses to the quality of the principal, the induction programme, and curricular flexibility were more important than the increase in salary for respondents across all schools, but the issue of high percentage of minority students was a larger factor for students in the small private urban institution. Overcoming this factor would require a salary increase of nearly USD 10 000. For large non-urban institution students, the differential would be about USD 9 100 for female students and about USD 5 800 for male students. The analysis suggests that improving the quality of principals may be more cost effective than higher beginning salary levels, noting that it would take nearly USD 25 000 to overcome the effect of a poor principal for students from large non-public institutions. However, Kirby, Naftel and Berends (1999) found in a study of data from 1979 to 1995 that in high-risk districts with high numbers of highly disadvantaged students in **Texas**, a USD 1 000 increase in salary was linked to a 6.2% decrease in attrition rates as compared to 1.0% in low-risk districts and 1.6% in medium-risk districts. They also found that African-American and Hispanic teachers were more likely to stay than teachers in general, 5-6% as opposed to 2.9%. The differences were significant.

Findings from Milanowski *et al.* (2007) are congruent with Hanushek, Kain, and Rivkin's (2004) research findings that working conditions were more important to experienced **Texas** teachers than pay differences in their decisions to change schools. Since their findings (Milanowski *et al.*, 2007) differ from those of Clotfelter *et al.* (2008) regarding retention, Milanowski *et al.* suggest that it may take larger incentives to attract new teachers with concerns about their preparation to teach in high-needs schools than to retain those already there. Two other cautions are mentioned; one is that a survey of new teachers may not provide an accurate representation of real attitudes as they may repeat what they have heard from their professors about social justice when they were students, especially concerning willingness to teach in schools with high minority participation. The second is that the survey was done in a state with a surplus of teachers, so students' willingness to take a job that is less than their ideal would be more expected.

OTHER FINANCIAL INCENTIVES

Incentives for high performing teachers. **England** (Ross and Hutchings, 2003) provides a Fast Track Teaching Programme designed to encourage and develop promising new teachers with leadership potential. Fast Track teachers receive an annual retention incentive of GBP 2 000 for each year in the programme, as well as additional coaching, mentoring and participation in leadership development activities. The goal is to encourage the identified teachers to move into leadership positions on a shorter timeline than usual.

Retention bonuses. **Singapore** created the Continuity, Experience and Commitment programme (CONNECT) to counteract attrition from the teaching profession. It provides retention payments upon retirement to teachers, over and above the social security system available to all employees in Singapore. In 2001 in response to the loss of mathematics and science teachers to the private sector, Singapore increased the amounts of the annual payment to SGD 4 200-6 200 per year for teachers with 1 to 15 years of experience and SGD 3 200 per year for more experienced teachers. In addition, it makes some funds available for withdrawal every three to five years to maintain motivation for remaining in the profession. It currently has a teacher retention rate of 97%.

Funds for individual professional development. In addition to 100 hours of professional development available to every teacher, **Singapore** provides reimbursements of SGD 400-700 per year for expenses teachers incur to improve their knowledge and skills. Teachers may purchase software, take foreign language or computer training, join professional organisations, subscribe to journals or participate in activities to enhance their cultural awareness. Teachers may also arrange for full-time or part-time professional development leaves partially funded by the Ministry of Education. They may study or travel abroad, teach in an international school, or work in the private sector to better understand the applications of the subject they are teaching. The goal is to improve their professional skills to better serve Singapore's students.

Additional support in the work environment. In 2003, **England** developed a programme, Raising Standards and Tackling Workload – a National Agreement (Teacher Development Authority, 2003) to improve working conditions for teachers. In a PricewaterhouseCoopers survey, teachers had said that two-thirds of their time was spent on non-teaching activities. Since teacher workload was given as a major reason for teacher retirements and attrition, the new programme reduced workloads by reducing the overall hours in the teacher contract, providing guaranteed planning time, reducing paperwork requirements and adding support staff to provide routine administrative services and help teachers and support students. Support staff including bursars, administrative, technical and classroom support staff were recognised as important members of the school team, and the programme created new career paths in three areas: pedagogical, behavioural/guidance, and administrative/organisational. Hopkins and Ahtaridou (2008) reported that the addition of support staff had a positive impact on teaching, teachers' job satisfaction, stress and workload, and student learning and behaviour. The purpose of the programme was to ensure better learning environments for students and a better working environment for teachers.

Flexible use of teacher time. Flexibility in how teachers use their time can be a major incentive for some. For example (Jacobson, Hickcox and Stevenson, 1996), deferred salary leave plans, available in **Canada** since the 1970s, and more recently in some **U.S.** school districts, enable teachers to defer a portion of their salary each year to allow them to schedule a leave with pay at a later time. This leave can be used for further study or travel to learn about other educational systems. **Singapore** (Sclafani, 2008) allows teachers to accrue a month of full-pay for half-time service or half-pay for full-time study for each year of service, up to 12 months. Teachers are able to use this time in any way that will benefit the Singapore education system. Examples of the activities undertaken include further study in Singapore or abroad, teaching in an independent or foreign school, or working in the private sector to learn more about applications of education knowledge and skills. Another version of flexibility in the use of time is the opportunity for senior teachers to engage in an increased variety of tasks and to reduce their work schedule with or without a small reduction in salary (OECD, 2005). Some local districts in **Norway** allow senior teachers to reduce teaching time and overall duty time to participate in professional development, mentor new teachers, develop curriculum and advise other schools. **Brandenburg, Germany** allows teachers to reduce teaching time by 50% for a 20% reduction in salary, an opportunity accepted by about 10% of the teachers. To retain more experienced teachers, the **Netherlands** allows teachers over the age of 52 to request a reduction in their schedule of duties for a smaller reduction in pay. For example, a teacher who is 52-55 years old can reduce his or her hours by 10% for a salary reduction of 2.5%; for teachers who are 56 or older, the reduction is 20% in time for a reduction in salary of 5%. This has been popular with teachers, with around 40% of eligible teachers participating.¹⁶

INCENTIVES BASED ON PERFORMANCE IN THE CLASSROOM MEASURED BY STUDENT RESULTS ON EXTERNAL EXAMINATIONS

Over the last twenty years, many countries and provinces/states have established large-scale assessment systems that test all students, rather than a representative sample of students, in given grade levels. This has resulted in an opportunity to use the examination results to establish a link between student performance and teacher effectiveness, creating new accountability for teachers and the possibility of linking compensation to teacher effectiveness. This has been the case in **Chile, Mexico** and the **U.S.** Full descriptions of the programmes in **Chile** and **Mexico** are provided in the section of this chapter on effectiveness. All states in the **U.S.** are required under the No Child Left Behind legislation of 2002 to establish state testing systems and publicly report the results for grades three through eight and once in upper-secondary school in reading and mathematics. This has provided a rich dataset that enables more states to connect schools and, in some states, teachers, to student test results. Some states had established large-scale testing programmes in the 1990s and have used those systems to establish school recognition and incentive programmes.

Incentives for high-performing schools. The Secretary of Education in São Paulo, **Brazil** established a performance incentive system in 2007 that rewarded school teams for meeting or exceeding performance targets for student achievement, enrollment, graduation and dropout rates, and teacher attendance and stability. Teachers in schools that meet all of their targets receive up to an additional three months salary, with the pay proportional to the schools' levels of achievement. Schools that exceed their targets receive additional incentives. The programme has not been evaluated, and it continues. **England**, as part of its performance management system, established a School Achievement Awards Scheme that awards staff bonuses shared by all teachers in schools that have demonstrated high achievement or rapid improvement. Approximately 30% of the schools are identified each year for awards.

In the **U.S.**, the State of **Alaska** (Prince *et al.*, 2008) has a Public School Performance Incentive Program that rewards schools that average more than a year's progress in student reading, writing and mathematics as measured on state examinations. Staff members earn awards ranging from USD 2 500 to USD 5 500 for certified staff, and from USD 1 000 to USD 2 500 for non-certified staff, depending upon the amount of additional growth

the school achieved. **California** has had a Distinguished Schools Awards programme since 1986, honouring primary and secondary schools in alternating years. However, the award is honorific, not financial. As of 2009, schools have to show not only progress in improving student achievement, but must also demonstrate that they are closing the achievement gap among students from different subgroups within their schools.

The State of **Florida** has had a School Recognition Program since 1999 that rewards schools receiving an accountability system designation of “A”; schools making exemplary improvement by improving at least one letter grade or sustaining the improvement the following school year; and alternative schools receiving a rating of “Improving” or improving at least one level. The total amount awarded to schools in 2008 was close to USD 150 million, distributed at USD 85 per student enrolled full-time in the eligible school. The funds may be used for staff bonuses, purchases of instructional materials, equipment or the addition of temporary personnel to assist the school in improving performance. It is up to the staff and school advisory committee to decide how to spend the funds, and, if they select staff bonuses, they must also decide which staff will receive them and how large the bonus is for each person. If the committee cannot decide by November 1st, the funds are divided equally among all staff members.

Over the last decade the **Florida** Legislature has developed and eliminated ideas for linking individual teachers’ salaries and bonuses to student performance. The BEST Career Ladder programme was established in 2003 with a pilot funded for four county school districts. While each district developed its own plan, the Teacher Advancement Program (TAP)¹⁷ was recommended as the model. TAP established a career ladder of four steps with additional salary stipends connected to each level. The pilot was cancelled after one year and the programme repealed in 2005 over concerns about the estimated cost of USD 650 million. While the programme was eliminated, districts were still required to develop plans to link teacher salary to performance. Special Teachers Are Rewarded (STAR) was written into the state budget in 2006 to reward 25% of the teachers with a 5% bonus, based on achievement of their students. Because of union and teacher opposition, STAR was replaced in 2007 by the Merit Award Program (MAP). MAP allows districts to develop plans based on group or individual teacher incentives negotiated as part of the collective bargaining process. Districts may develop plans and apply to the state for funding to implement the programme. Only a third of the districts had applied for the 2007-08 school year (Chait, 2007). There has been no evaluation of any of these programmes.

In 2006, the State of **Texas** in the **U.S.** created and funded the largest teacher incentive programme in the country, the Texas Award for Educator Excellence. The programme has three components. The first, the Governor’s Educator Excellence Grant (GEEG), provides three-year grants to 99 schools that are in the top third of schools serving highly disadvantaged students and earning an accountability system rating of exemplary or recognised. Schools design their own plans and are required to use gains in student test scores and measures of teacher collaboration, and they can add optional subjective measures of teacher collaboration, teaching in hard-to-staff positions, and teacher initiative and commitment. The first year evaluation looked only at the plans and the attitudes of the teachers. While the programme recommended minimum awards of USD 3 000, the average awards fell between USD 2 897 and USD 3 726, with actual awards ranging from USD 100 to USD 10 937. The initial evaluation found that teachers were generally positive about the programme, perhaps in part because of their direct involvement in its design at the school level. In the mid-year and spring surveys, teachers indicated that they had changed their practice to focus on those behaviours linked to higher student achievement, and those teachers that received awards were more likely to report they had changed their practice in this way. Teachers did not report that the programme had reduced collaboration or lowered teacher morale. Future evaluations will consider if there are resulting changes in student achievement and teacher turnover.

The second component of the **Texas** plan is the Texas Educator Excellence Grants programme, (TEEG) which awards incentives to teachers in schools serving proportions of disadvantaged students that fall in the top half of the state distribution and that have achieved either high scores or great improvement of scores on the state

accountability tests. It requires that 75% of the funds go to teachers determined by the school through objective measures to have demonstrated impact on student achievement and collaboration. Over 97% of the eligible schools applied for the grants in the first year of the program. Participation in the programme is less stable than GEEG from year to year because of the volatility of the Comparable Improvement ratings and budget constraints. While the programme recommends minimum awards of USD 3 000, maximums are often less than this amount. As in GEEG, evaluations showed that teachers were generally positive about the programme and reported they had implemented more classroom practices linked to higher student achievement, but the teachers did not credit the programme for that change. The evaluators did find a sharp increase in teacher turnover among those who did not receive awards or received small awards and a marked increase in teacher retention among those teachers who received high bonus amounts. This is exactly what Lazear (2003) suggested: one of the great values of incentives is their effect on selection: more effective teachers stay while less effective teachers leave. The question to be answered in future evaluations is whether it is the most effective teachers who are receiving the large awards and remaining in the schools.

The third **Texas** programme component, the District Awards for Teacher Excellence, required districts to develop and implement an award plan district-wide or selected campuses. This is the only plan that focuses on districts and allows districts to volunteer to participate. A little over half of the district plans made all schools eligible for participation, while 40% limited participation to a selected group of schools, and about 7% used the funds to implement the TAP. The review of the initial year of the programme was included in the TEEG evaluation (Springer *et al.*, 2008). At least 60% of the grant award must be used to reward classroom Texas school district teachers who are the most effective at improving student performance. The other 40% can be used on teacher stipends, teacher mentors, principal incentives, awards to other campus employees, and elements of the TAP. Springer *et al.* (2008) reviewed the programme but it has not yet been fully evaluated.

North Carolina and **Texas** have used the data from their state-wide accountability testing programmes as the basis for offering bonuses to schools for high performance or growth. Neither Texas nor Florida has yet evaluated the programmes for impacts on student achievement, although Texas plans to continue its evaluations and will look for evidence of impact. **Arizona** and **Missouri** established teacher career ladder programmes that have recently been evaluated using state testing data. The results of evaluations of the Arizona, Missouri and North Carolina programmes are in the section of this chapter on effectiveness.

Certain school districts in the **U.S.** (Koppich, 2008) have established pay-for-performance programmes although their states may not have done so. Houston (Behrstock and Akerstrom, 2008) has had a series of pay-for-performance programmes since 1976, with the latest, the ASPIRE programme, created in 2006 and now the largest pay-for-performance programme in the **U.S.** Although the current plan encountered major opposition in its first iteration because only a segment of teachers who taught in tested subjects and grades were eligible for awards, it was totally revamped using Sanders' SAS model, a value-added analysis, to make all staff eligible for at least some payments. Teachers are rewarded for value added when their students' progress scores are in one of the top two quartiles in subjects measured by the **Texas** state exams or the Stanford 10, a national norm-referenced exam. All teachers and support staff have access to school-wide bonuses if their school ranks in the top two quartiles of demographically similar schools; if their school ranks in the top two quartiles of progress compared to state education agency-designated comparison schools; if their school earns a ranking of exemplary or recognised; or if 70% or more of the students writing scores meet the college readiness standard in grades 4, 7 or 11 or the top two quartiles in improvement. The total amount of awards per teacher ranges from USD 0-7 800; for instructional support staff from USD 0-1 450; for teacher assistants from USD 0-850; and for operational support staff from USD 0-500. The union and teacher organisations continue to object to the programme and to claims that it has reduced attrition or increased student performance. The largest teacher union representative believes that future work on the system must include the involvement of teachers and

teacher organisations. They also believe that the system is too complicated for teachers to understand what they have done to receive rewards.

The Toledo Review and Alternative Compensation System (TRACS) created in 2002 (Koppich, 2008) provides three tracks: TRACS A providing professional development for teachers requiring remediation; TRACS B for school-based rewards for student performance; and TRACS C for individual teacher performance and eligibility for an additional 15% bonus above base pay. TRACS C requires that teachers have at least five years of successful experience and allows them to move up a career ladder. Career teachers, the first level, focus on an area of student achievement, Accomplished teachers help other teachers as peer reviewers and curriculum developers, and Distinguished teachers serve in high-needs schools for at least three years. The evaluation is underway and will compare the student test scores for TRACS teachers with similar student scores of non-TRACS teachers.

NON-FINANCIAL INCENTIVES

Devolution and greater school autonomy. Nicaragua (Parker, 2005) began decentralising the control of its schools in 1993 to move education closer to the community, increase financial resources available for education, and increase efficiency in schools' use of human and financial resources. Schools had the option to become autonomous, based on a vote of 80% of the staff of a school. Autonomous schools created school councils that included parents, teachers, the school principal and a non-voting student. The councils make hiring and firing decisions for staff and principals. Although many teachers voting for autonomous schools expected to receive bonuses from the increased obligatory fees contributed by parents, these bonuses were greatly reduced when fees were banned in 1995 for primary schools and limited in size for secondary schools. The bonuses were eliminated, by law, for all schools in 2002. In a separate program, primary school teachers received up to 30% of base salaries in World Bank incentives for student attendance and in local bonuses from the school council. For secondary teachers, the range moves up to 50% of base salaries. Despite these two bonus programs, Parker found no differences between autonomous and centralised schools in terms of student achievement, although autonomous schools have younger and wealthier students, as well as better resources and infrastructure. She suggested that further research is required and wondered if the instability of teacher incentives has limited their effect on teacher behaviour.

Positive working conditions. Darling-Hammond (1997) has raised the issue of working conditions as a major source of teacher satisfaction and retention. She found a strong association between teachers planning to stay in teaching and their views of the quality of support provided to teachers by the school's administration, as well as the availability of resources and teachers' ability to have a voice in the school's decisions. Other teacher surveys have shown that it is the lack of such positive work environments that contribute to high attrition rates from schools in high-poverty schools and schools with high-minority student populations (Loeb, Darling-Hammond and Luczak, 2005). Teachers in **England** and **Wales** (OECD, 2005) responded very favourably to the 2003 *Raising Standards and Tackling Workload* agreement. This agreement reduced the amount of administrative/clerical duties assigned to teachers by adding support staff and providing them with better training to assume those responsibilities. It also phased in guaranteed, additional planning, preparation and assessment time for teachers. Over 97% of teachers surveyed for the Department of Education and Skills in 2004 responded that teaching and learning had improved because of the agreement, and about half reported that teacher workloads had decreased overall.

In the **U.S.**, a **California** study using teacher survey data linked to district data on salaries and staffing patterns, Loeb, Darling-Hammond and Luczak (2005) examined the connection between teacher attrition and student demographics, working conditions and salary. They found that salaries and working conditions, such as large classes, facilities problems, school schedules based on multi-tracks, and lack of textbooks, are strong and significant factors in teacher turnover. Once those factors have been accounted for, student demographics

become a smaller factor in a teacher's decision to leave a school. Quartz *et al.* (2004) studying graduates of an elite teacher training programme focused on preparation for teaching in inner-city schools found similar results and suggested that student demographics may appear as a proxy for poor working conditions since high-poverty, high-minority schools so often provide poor working conditions. Although it is suggested that teachers leave high-needs schools because they do not feel prepared to meet the needs of inner-city students, the graduates of the programme still left inner-city schools despite their special training. The teachers reported that they went to schools with better working conditions and more opportunities to learn and grow.

External examinations. PISA and PISA-Extension studies in **Germany** (Juerges, Richter, and Schneider, 2004) that compare states with and without external exit examinations have concluded that accountability in the form of external exit examinations leads to higher student achievement in mathematics and German reading, especially in lower secondary schools when exams are imminent. Using surveys of students and parents, the study concluded that teachers of mathematics in states with external exit examinations put more pressure on their students to perform, have more disciplined classrooms and provide exercises that are more innovative and less repetitious than mathematics teachers in non-exam states. Thus, even without specific connections of teacher pay to student performance, Juerges *et al.* (2004) conclude that teachers value their reputations and increase their effort if the output of their efforts, student exam scores, is measured and published, as they are in states with external exit examinations. Maintaining one's reputation is an incentive to work harder for high performance.

Development of new salary systems. While districts, states and nations have created incentive programmes, only a few have replaced the salary schedule with a new system. **Denver** Public Schools is probably the best known, because it created its programme in collaboration with the teacher union and is funded by long-term tax override placed in a trust administered by the district and union. The programme replaces the single salary schedule with optional components selected by the teacher from increased knowledge and skills, willingness to teach in hard-to-staff schools or shortage subject areas, student test scores and annual evaluations. The programme is voluntary for experienced teachers and required for new teachers entering the district after 2006. Although the evaluation has not been completed, Koppich (2008) cites data showing greatly increased interest in teaching in hard-to-staff and hard-to-serve (high poverty) schools: the number of applicants for transfer into the hard-to-staff schools increased by a factor of eight. In a review of the 16-school pilot, the Community Training and Assistance Center (2004) found that teachers reported greater focus on student achievement and the use of student achievement data. However, teachers still had issues of trust in and their perceptions of the fairness of the programme and believed that principals' evaluations were not consistent. They did not believe that there was sufficient alignment among the systems of instruction, assessment, human resources and professional development. Chapter 7 discusses in depth the issue of alignment.

The Vaughn Next Century Learning Center, a charter school in Los Angeles, **California** (Kellor, 2005) provides an excellent example of a new method of rewarding teachers. Although it is a single school, its strategy could be applied more broadly. The plan was designed by a staff committee and approved by a governance committee with the understanding that it would be refined over time, based on identification of aspects through annual staff surveys that could benefit from change. The school uses standard salary steps for the first 11 years and provides one additional step at 15 years of experience. In addition, it provides supplements of USD 1 000-6 000 to the salary for credentials, such as a master's degree or 30 additional credits in education, National Board Certification, Elementary Teaching Credential (*i.e.* for primary level) or Demonstration Teacher certification. Competency-based supplements up to USD 13 000 can be earned annually for Level One, up to USD 3 500 for increased competence in depth of essentials, such as literacy, technology or classroom management; or Level Two, up to USD 5 000 for increased in breadth and depth, for example, in subjects areas or strategies for English Language learners; or Level Three, up to USD 4 000 for high levels of teaching performance in addition to high performance in Levels One and Two. In 2003-04, the school added a Level Four, Distinguished Teacher, for

those who had held Level Three status for five years. The competency-based supplements are evaluated using detailed rubrics based on the Danielson (2008) *Framework for Teaching* and its four levels of performance. The ratings of a supervisor, a peer and self-evaluation all count equally in the final rating.

The Vaughn incentive programme was required for all new teachers and voluntary for those with less than five years of experience, although at teacher request, eventually all teachers were eligible for participation. In the second year, the school added a School-Based Performance Award based on school-wide student achievement on three tests, with goals set for progress on each one, and on report card grades. If the school met its goals, every teacher and administrator received equal awards. The system was later changed to use the new state system, the Academic Performance Index (API), and the school provides awards of USD 2 000 per teacher using the API criteria. The Consortium for Policy Research in Education, University of Wisconsin, conducted an evaluation of the school using hierarchical linear modeling to compare student test scores and teacher competency evaluation scores. The evaluation found a statistically significant correlation between teachers' literacy scores and student achievement in reading on the Stanford Achievement Test, SAT9, and positive gains for mathematics and language arts. The school has exceeded its goals for the API every year since its start and is one of the highest performing public or charter schools in California, despite its disadvantaged student population.

The State of **Minnesota** developed a voluntary Quality Compensation (Q-Comp) system that enabled school districts, in collaboration with the exclusive teacher collective bargaining organisation, to design new salary schedules based on required components. In return, the district receives USD 260 for each full-time student enrolled in the district to implement the programme. According to state guidelines, the alternative teacher professional pay system agreement must:

- (1) describe how teachers can achieve career advancement and additional compensation;
- (2) describe how the school district, intermediate school district, school site, or charter school will provide teachers with career advancement options that allow teachers to retain primary roles in student instruction and facilitate site-focused professional development that helps other teachers improve their skills;
- (3) reform the "steps and lanes" salary schedule and base at least 60% of any compensation increase on teachers' performance using:
 - (i) school-wide student achievement gains or locally selected standardised assessment outcomes, or both;
 - (ii) measures of student achievement; and
 - (iii) an objective evaluation program that includes:
 - individual teacher evaluation aligned with the educational improvement plan... and the staff development plan...; and
 - objective evaluations using multiple criteria conducted by locally selected and periodically trained evaluation team that understands teaching and learning;
- (4) provide integrated on-going site-based professional development activities to improve instructional skills and learning that are aligned with student needs..., consistent with the staff development plan... and led during the school day by trained teacher leaders such as master or mentor teachers; and
- (5) allow any teacher in a participating school district, intermediate school district, school site, or charter school that implements an alternative pay system to participate in that system without a quota or other limit.¹⁸

In 2002, the Minneapolis Public Schools (Koppich, 2008) developed a new salary system, Professional Pay (Pro-Pay), under the **Minnesota** Q-Comp programme. In 2006 the district and its teachers' union developed the Minneapolis Alternative Teacher Professional Pay System. It created steps, called career increments, based on

experience, participation in professional development and maintenance of a portfolio. The new lanes of the salary schedule, named professional growth credit lanes, provide additional compensation based on earned National Board Certification, masters' degrees in specific subjects and certificates of specialisation in areas needed by the district, such as English as a Second Language or Gifted and Talented certification. As developed under ProPay, teachers can earn supplements above the salary schedule for participating in professional development, using that experience to conduct action research and developing a paper on how they applied the learning to their classroom activities and what results they achieved. Teachers can also receive additional compensation for assuming additional responsibilities such as mentoring, coaching, leading professional development sessions and serving on or leading school or district committees. Schools received Quality Performance Awards prior to 2008 based on student academic achievement measured by test scores, attendance and satisfaction surveys, and a school's score provides its teachers with additional credit toward compensation. Since 60% of increases in teacher salaries have to be based on student achievement under Q-Comp, the QPA programme is being redesigned. Unlike **Denver**, Minneapolis chose to make its ATPPS initially voluntary for all teachers; however, teachers hired after 2006-07 were required to participate. The district has had annual approval by staff to continue the programme, and it revises portions of the plan each year in response to concerns. There has not yet been an evaluation of the programme's impact on student achievement or teacher recruitment and retention.

EVIDENCE OF EFFECTIVENESS

A variety of incentive plans have been developed in the last 20 years. However, most countries do not mandate or do not have the systems in place to collect data from local schools that would provide researchers with clear evidence to determine what has been effective. States in the **U.S.** have been encouraged, through federal grants, to establish robust data systems that collect data on student and teacher demographics, as well as links between teachers and students that enable researchers to connect the performance of teachers to that of their students. However, it is estimated that fewer than 20 states currently have that capability.

There have been sound evaluation studies of some pay-for-performance programmes based on student test scores at the national or state levels, primarily in **Chile**, **Mexico** and the **U.S.**; smaller experimental studies of teacher attendance and student test scores in **India**, **Israel**, and **Kenya**; and analyses of teacher surveys regarding the pay-for-performance programme in **England**; but few of the other types of incentive programmes have been rigorously evaluated.

EVALUATIONS OF NATIONAL AND STATE/REGIONAL PROGRAMMES

Conditional scholarships and loan forgiveness programmes. Most conditional scholarship and loan forgiveness programmes in the **U.S.** have not been rigorously evaluated. Steele, Murnane and Willett (2009) report on a random assignment study of law students done by Field (2009) which found that conditional scholarships increased the likelihood of first job placement in public interest law by 36-45%. Steele *et al.* (2009) studied the State of **California** Governor's Teaching Fellowship programme, which provided a conditional scholarship of USD 20 000, competitively awarded to academically-talented students to attract them to teach in schools in the bottom half of the state's achievement ranking system for at least four years. The programme was in effect for two years before funding problems led to its elimination. Steele *et al.* compared the recipients of that scholarship to others receiving state grants under the Assumption Program of Loans for Education (APLE), which was easier to attain and offered loan forgiveness of USD 11-19 000 for graduates of teacher licensure programmes who taught in shortage subject areas or in hard-to-staff schools for at least four years. Using the APLE database to track recipients for up to four years after earning teacher licenses, the study compared the GTF recipients with the total database of APLE recipients and determined that the GTF scholarship increased by 28% the likelihood that highly talented recipients taught in low performing schools for at least two years. In addition, 85% remained in their schools for at least two years, and 75% completed the four-year commitment.

However, taking the total cost of the programme, approximately USD 14 million, and dividing it by the number of teachers who might not have otherwise taught in low-performing schools, the per-candidate recruitment cost was over USD 19 600 per year per person. Given that individual talent explains only a small portion of teacher effectiveness, Steele *et al.* (2009) suggest that systems explore combining recruitment programmes for highly-talented students with retention incentives based on classroom effectiveness measures.

Retention bonuses. A programme of incentives for teaching in hard-to-staff schools in **North Carolina**, known as the North Carolina Public Schools Bonus Program, was evaluated (Clotfelter *et al.*, 2008). It showed that the USD 1 800 annual incentive paid to experienced teachers of mathematics and science in hard-to-staff or low-performing schools reduced teacher turnover rates by over 10%. The retention programme was discontinued by an incoming governor before the study confirming its effectiveness was published.

Pay-for-performance programmes in Chile and Mexico. The performance bonus plan in **Chile**, the National System of Performance Assessment (*Sistema Nacional de Evaluación de Desempeño de los Establecimientos Educacionales*, SNED), has provided rewards biannually since 1997 to schools in each region whose performance on national examinations places them in the top 25% of performance in the region. Mizala and Romaguera (2005) report that the SNED is not an incentive for additional effort for those schools that always score in the top 25%, nor for those that have never scored in that range. However, for those schools that have a chance of success, it has had a positive effect on student achievement that is significant for individual cycles of the SNED awards and a positive cumulative effect of the different SNED applications over time. In addition, it has affected teachers' attitudes, making them more open to performance evaluation linked to incentive pay. It has also opened the door to greater salary variations in the collective bargaining agreement for 2006, through which the SNED awards double their previous monetary value to 8% of base salary, and the number of schools rewarded has increased to 35%. Even at this level, however, seniority and professional training still provide the greatest incentive for teachers.

A second programme in **Chile** (Manzi, 2008) provides individual teacher incentives based on the teacher's performance evaluation. The system started with the development of national teaching standards, *The Framework for Good Teaching (El Marco para la Buena Enseñanza)*, that outline the criteria for teacher evaluation. The evaluation is based on four components: self evaluation, weighted 10%; evaluation by supervisors, both principal and head of instruction, weighted 10%; a peer evaluation by a teacher from a different school, completing a structured interview based on the criteria of the teaching standards, weighted 20%; and a portfolio, that carries 60% of the evaluation weight. The portfolio gives the teacher the opportunity to showcase and reflect upon how he or she responded to learning goals in the curriculum and contains an hour-long videotape of a lesson that can be analysed by trained observers using a specified rubric based on the performance standards of the teaching framework. Teachers are rated as outstanding, competent, basic or unsatisfactory. The teachers in the top two categories can receive an incentive if they complete an examination of their subject area knowledge. The amounts of the stipend depend upon the score on the examination and range from 5-25% of basic salary. Teachers retain the stipend for up to four years and then must be re-evaluated. Manzi (2008) reports that 60% of the teachers take the test and about half of them receive an incentive award.

Teachers in the bottom two categories of the Chilean teacher evaluation system are offered training to improve their performance. The evaluation provides specific standards within the teaching framework and identifies where the teacher needs to improve so professional development can easily be tied to the specific needs of the individual teacher or of groups of teachers at the school or municipality level. The programme has been in place for six years and consistently evaluates about 60% of the teachers as competent or outstanding, 3% unsatisfactory, and the rest basic. Manzi (2008) suggests that the value of this system is its direct link to the specific criteria in the teaching framework that, if addressed through effective professional development, result

in the greatest impact on the quality of the teacher's performance. This differs from value-added analyses that suggest that the teacher did something to improve student performance, but do not enable the teacher to understand what that was. He cites an interview study by Taut *et al.* (2008) in which stakeholders agreed that the programme had helped to recognise teaching as a valued profession and contributed to a discussion of professional development as a way to improve practice. However, the judgements on the evaluation are strictly subjective despite the use of rubrics, and do not include any use of student performance measures. SIMCE, the national examination system, evaluated the impact of teachers with different evaluation ratings and found that students taught by teachers in the top two evaluation categories significantly outperform their peers who had teachers in the bottom two evaluation categories. Manzi (2008) reports that evaluations by Eisenberg (2008), Leon (2008) and Manzi *et al.* (2008) had findings consistent with these results.

Carrera Magisterial (CM), Mexico's public school teacher incentive programme (Santibáñez *et al.*, 2007; McEwan and Santibáñez, 2005) was one of the earliest programmes to offer economic incentives for teaching performance, having started in 1993. The programme offers additional salary incentives to teachers with permanent contracts and union membership, who are willing to undergo annual evaluations conducted by their school peers, principal and a union representative; participate in professional development activities and be tested on their content; take a general test of subject area knowledge; and have their students tested on national examinations. In addition, teachers get points for seniority and highest degree earned. Teachers can receive a maximum of 100 points, with results of student examinations accounting for a maximum of 20 points. The programme has 5 levels, with teachers earning at least 70 points placed on Level A and earning a bonus of 20% of base salary while those on Level E receiving 150% of base salary. Most teachers eligible for the programme are currently on Level A. Once a teacher receives a salary incentive, it is continued for the rest of the teacher's career.

The evaluation of CM completed for the Mexican Ministry of Education (Santibáñez *et al.*, 2007; McEwan and Santibáñez, 2005) found that although the teacher tests were technically adequate, many items in the teacher content examinations were not cognitively challenging and that the same tests are used no matter how many times the teachers take them. Student examinations were determined to be less technically adequate than the teacher exams, and some examinations at the secondary level were at low levels of cognitive demand, while others were quite high. More to the point, the incentives had no impact on primary student test scores and a modestly positive effect on student test scores for secondary teachers. Once teachers received Level A incentives, there were modest negative effects on their students' test scores. It was also noted that the programme is only a real incentive to improve student performance if teachers have earned at least 50 but fewer than 70 points before student test results are considered. If below 50, the teachers cannot reach the 70 point cut-off and have no incentive to work harder to improve student performance, and if already at 70, there is no need to exert any additional effort. Given the role of educational degrees and seniority in both the base salary and the CM, teachers receive a double incentive for the factors that mean least for teacher effectiveness and student achievement.

Pay-for-performance programmes in U.S. states and school districts. There are a number of evaluations of state and local programmes in the U.S. that provide some evidence of positive impact, but suffer from uncertainties due to the design of the programme, the availability of data, or the design of the evaluation. The Teacher Advancement Program (TAP) described in Chapter 1 is used in 14 states in individual schools. It has been evaluated, (Solomon, White, Cohen and Woo, 2007) using William Sanders' SAS value-added method of analysis. The evaluation was funded by its parent organisation, the National Institute for Excellence in Teaching, and its executive director was the lead author. The results of the evaluation in six states found TAP to be effective in terms of TAP teachers outperforming control group teachers in 63% of the comparisons of student achievement growth. Fewer TAP teachers scored a "1" or "2", indicating lower levels of effectiveness

in raising student levels of achievement than the control group teachers. In addition, more TAP teachers scored a “3” or better than those in the control group, indicating that they were significantly more effective in raising the achievement scores of their students, and more TAP teachers had students making an average of one or more years of growth for a year of instruction. TAP is being evaluated by an external organisation, Mathematica, funded by an independent foundation. The first year report (Glazerman, McKie and Carey, 2009) of a random assignment programme did not find any impacts, but that is not surprising in the first year. The evaluation of the programme will continue and will provide more definitive results.

The State of **Arizona** has had a career ladder programme in place for over 20 years. Districts within the state create plans that include teacher performance and additional responsibilities and may include the use of student achievement in determining teacher pay. Under Arizona state guidelines, all new teachers in Career Ladder districts must be evaluated for placement on the Career Ladder. However, new teachers have the choice of whether to participate in the career ladder programme or remain on the district’s traditional salary schedule. According to the information provided by the State Department of Education:

The program supports and encourages collaboration and teamwork, and provides opportunities for leadership and professional growth, with Career Ladder teachers participating in higher-level instructional responsibilities within their districts. The program also allows districts to apply to implement an additional incentive program for other personnel at the school district level and provides awards based upon group, team, school or district.¹⁹

The **Arizona** Career Ladder programme (Dowling, Murphy and Wang, 2007) has been shown to have a significantly positive impact on student test scores in mathematics, reading and writing after adjusting for differences in students and schools. The effects were stronger for reading and mathematics. There are limitations, however, in the study. Since schools choose to participate in the Career Ladder programme, the differences between schools may not have been totally accounted for by the statistical analysis. In addition, the study did not attempt to provide evidence of what caused the positive impact.

The career ladder programme of the State of **Missouri** (Booker and Glazerman, 2008) uses a combination of teacher tenure, teacher performance and extra responsibilities taken on by teachers to determine salary supplements. While the programme is available to all districts, districts are not required to participate. Districts that have lower per capita income levels receive subsidies from the state so that they can provide the required matching funds. The career ladder has three steps, worth USD 1 500, USD 3 000 or USD 5 000 per year respectively. Moving to each stage requires reviews of the teacher’s career ladder development plan and artifacts regarding performance, as well as scheduled and unscheduled observations by the district career ladder review committee. Teachers must participate in extra work with students and attend professional development activities, for at least an additional 60 to 120 hours outside of their regular responsibilities, depending on the stage. More than 25% of teachers from 64% of the districts in **Missouri** participated in the career ladder programme in 2005-06. Booker and Glazerman (2008) concluded that there is some evidence that the programme has a small positive effect on average mathematics achievement, especially in primary grades, but no significant effect on reading. The evaluation was complicated by the lack of experimental conditions, and, when limiting the review to districts that changed their participation status over time, they found no significant effects. Further analysis, using the more extensive test results recently available, is required before policy changes could be recommended.

North Carolina (Vigdor, 2008) has had a system of state testing since the mid 1990s and, since 1996 has awarded bonuses (USD 1 500 or USD 750) to teachers in schools that have either reached an exemplary or expected level of growth in student performance on the state examinations. This is the longest-running bonus programme in the U.S., and it remains popular in the state. While the statistical analyses are complex, schools understand that they have an incentive to increase the mean scores on reading and mathematics tests.

Vigdor notes that the programme's impact on student achievement has been mixed: achievement has improved on both the state exam and the low-stakes National Assessment of Educational Progress in mathematics, but only on the state exams in reading. He also found that the programme's statistical analyses place schools serving at-risk students at a disadvantage and that there has been a significant increase in the number of teachers transferring out of those schools.

South Carolina had a positively evaluated voluntary school and teacher incentive programme (Boozer, 1999) designed to improve student performance. However, the voluntary nature of the programme makes it unclear if the positive impact was due to the fact that the most effective schools and teachers volunteered to participate rather than to the incentive. The programme no longer exists.

There have been evaluations of school-based incentive programmes in a number of school districts. Smith and Mickelson (2000) found positive impacts on student performance in the Charlotte-Mecklenberg County School District in **North Carolina**. However, because the programme was part of a comprehensive reform programme, it is difficult to attribute the positive effects on teacher motivation and student achievement to the incentive programme. This was not an experimental design with a control group, and it included all schools in the district. However, Smith and Mickelson did compare results to other urban districts in **North Carolina** and found greater improvement in Charlotte-Mecklenberg. In the State of **Texas**, the Dallas programme (Ladd, 1999) used a sophisticated regression analysis to account for differences in student backgrounds. While the study found that the district reasonably accomplished its goal of improved student achievement, teachers and the public did not understand how the programme made its decisions, and some suggested that the programme meant that the district expected less from students coming from disadvantaged economic backgrounds. Compared to other large districts in Texas, Dallas did see greater gains. However, the researchers found that this was true the year before the programme began as well, so it was difficult to attribute the improvement to the incentive programme. Dallas eliminated the programme when Texas initiated school and teacher recognition programmes.

Private foundations and the Little Rock School District, **Arkansas**, funded a pilot programme for teachers in five high poverty, low-achieving primary schools, paying bonuses to teachers if their students raised their test scores. The programme was phased in with the first school, providing bonuses based on the number of students achieving certain levels of growth on nationally standardised tests. Four schools that initiated the programme later received bonuses on the basis of gains in average class scores on the same tests. The payments were graduated, based on the size of the gain, and ranged from USD 50 per student for a 0-4 point gain to USD 400 per student for a gain greater than 15 points. The upper limits for a single teacher were about USD 11 000. Principals and aides were also eligible based on school-wide gains. Ritter *et al.* (2008) evaluated the programme by comparing the five schools to similar schools in Little Rock and found that students in the programme schools outperformed their peers in comparison schools by seven percentage points in mathematics, nine in language and six in reading. The attitudes of teachers regarding the programme were not positive at first, but improved. Teachers' attitudes were more positive about their own effectiveness than teachers in the comparison schools. The study did not meet the criteria for quality research of the What Works Clearinghouse in the U.S. Department of Education, because the schools in the two groups did not have an equivalent level of low-income students. In 2007-2008 the State of Arkansas created a pilot alternative compensation programme that could include incentives similar to this, and two districts are participating in the 2008-09 school year.

New York City, **New York**, initiated a group incentive programme in 2007 to reward teachers and union members in high-needs schools that reach 100% of their targets or reach 75% of their targets on the district's Progress Report Card. The targets include student attendance and student, parent and teacher responses to a survey on the school's learning environment; student performance on high-stakes state examinations in English language, arts and mathematics; and improvements in student test scores on the same examinations. Teachers and other

union members in schools that meet 100% of their targets receive USD 3 000, and in those meeting 75% of their targets, USD 1 500. Developed in collaboration with the United Federation of Teachers, the teachers' union in NYC, the programme is voluntary and requires the approval of 55% of the teachers in a school for it to be eligible. The programme evaluation study (Springer and Winters, 2009) uses a random assignment design, with all eligible schools assigned to either an experimental or control group. This methodology provides an opportunity to identify causal relationships between the incentive programme and Progress Report Card results. Since schools did not enter the programme until mid-year of the first year of implementation, one would not expect to see any gains in the first year. In fact, there were no significant differences between the experimental and control groups in the target areas. However, the opportunity to learn about the impact of the programme will come from evaluations of years two and three.

Figlio and Kenny (2007) used the National Education Longitudinal Survey (NELS), along with personnel practices surveys of 2000 schools within NELS to inquire about the use of incentive payments for teachers and to determine if there was a positive relationship between incentives and improvements in student achievement. While they found a positive association between individual teacher incentives and improvements in student achievement, they were unable to determine if the schools that chose to use incentive programmes differed in other important ways that made them more effective, or if it was the effect of the incentive programme. They did find that private, non-Catholic schools were more likely to use teacher incentives than public schools, and non-unionised schools were more than twice as likely to use teacher incentives than unionised schools. These studies point to the need to conduct carefully designed experiments with control groups if researchers are to be able to say with any surety that incentives are the source of improved student achievement. The next section describes experimental studies that had more definitive outcomes, although not all indicated a positive association with the desired outcomes.

Experimental studies of incentives

Randomised evaluation studies are considered the model most likely to establish a causal relationship between an intervention and the resulting change in behaviour. By establishing experimental and control groups from schools and teachers which have been matched on observable factors, the researchers hope to determine if the intervention applied makes a difference in the experimental group that is not seen in the control group. Much of the educational research in the past established correlations between an action and its results, but that basically means that both have been found, not what the relationship is between them. This section provides descriptions of experimental studies of incentives in education.

A series of experiments in India have demonstrated positive effects for incentives improving both teacher attendance and student achievement on external examinations. Duflo, Hanna and Ryan (2007) did a randomised evaluation of a programme to improve attendance in non-formal education centers in rural areas of **India**, where absenteeism of teachers was estimated at over 40%. By offering bonus funds to teachers whose attendance was verified through the use of camera photographs taken by children at the beginning and the end of each day, the programme decreased teacher absences by 19%. In addition, student test scores at the end of programme in those centers increased by 0.16 to 0.21 standard deviations in each subject tested, and 10% more students graduated to attend government education institutions.

Muralidharan and Sundararaman (2008) conducted a randomised evaluation of teacher incentives in rural government primary schools in **Andhra Pradesh, India**, which tested two approaches to improving education. The first group of experimental schools received either an additional paraprofessional teacher or a cash block grant to the schools to use for additional resources. The second group of experimental schools received either group performance bonuses for increased school test scores or individual teacher performance bonuses for increased test scores of their students. The four sets of experimental schools, that received the equivalent of

3% additional resources, were compared to similar schools acting as controls. They found that the schools receiving performance bonuses had significantly higher test scores than the control groups, equal to a mean treatment effect of 0.15 standard deviations. The incentive schools scored higher on questions of different levels of difficulty in all five grades in all five districts in which the experiment was conducted. Students also did better on tests of subjects that were not part of the incentive programme. There was no difference between group and individual incentives, but the schools had only three teachers each. The two sets of schools that received additional inputs performed significantly higher than the control groups, with student test scores 0.09 standard deviations above the control schools. Both the input programmes and the incentives were found to be five times as cost effective as the status quo, with the incentives achieving higher levels of achievement.

Lavy (2002) reported on a rank-order tournament run by **Israel** in 1996 among teachers of 12th grade English, Hebrew and mathematics. The teachers received cash bonuses for improving the average number of credits earned by students, the share of students who received a matriculation diploma and the school dropout rate. Sixty-two schools were selected to participate and ranked on their annual improvement levels, with the top third of the schools receiving awards. USD 1.5 million rewards were allocated to schools based on their position in the rankings, with awards ranging from USD 105 000 to the top school to USD 13 250 to the last-ranked school that qualified. The awards were divided into 75% for individual teacher salary bonuses shared as a percent of gross salary by all teachers at the schools and 25% for faculty facilities. By surveying teachers in the experimental and non-experimental groups, Lavy was able to determine that the effects of the teachers in the experimental group were the result of changes in teaching strategies as well as additional effort to respond to individual student needs and offer after-school lessons. In a second experiment, Lavy (2004) compared students' scores on matriculation examinations in an individual teacher rank-order tournament programme conducted in experimental and control schools that provided bonuses up to USD 7 500 for a classroom teacher whose base salary was about USD 25 000. He found that student test scores increased significantly in the subjects that were part of the experiment for both the teachers who actually received the bonuses and those eligible teachers who did not receive them, as compared to a control group of teachers who did not participate. He also found that there were minor positive effects on subjects that were not included. In comparing the results on the two experiments, he concluded that the individual bonus programme was more cost effective than the group bonus programme.

Glewwe, Ilias and Kremer (2008) report on a group incentive programme offered to teachers in grades four and eight in randomly selected primary schools in **Kenya**. The prizes for top scoring and most improved schools ranged from 20 to 40% of a teacher's monthly salary. Since all students who started the programme were counted at the end, teachers were encouraged to reduce dropouts and discouraged from only testing the strongest students. While more students tested, the programme did not affect 8th grade graduation rates or dropout rates. Student test scores did increase during the programme, but the student test scores were not significantly different from those in the control schools once the programme ended. They conclude that the teachers focused on short-term gains to get the incentive payments, rather than long-term learning. For example, teacher absentee rates did not go down during the programme, nor did the quantity of homework assignments go up. The major change in teacher behaviour was the additional number of exam preparation sessions during vacations. Further evidence was that test scores did not improve on examinations that differed from the government examination in format and content and that were not included in the incentive programme.

Marsden and Belfield (2005) surveyed teachers in **England** before, during and after the move to performance pay based on movement to new upper salary scales based on teacher evaluations. Teachers were eligible to apply once they reached the top of the regular salary scale, which required at least eight years of experience. Over 80% of the eligible teachers applied and over 95% successfully crossed the threshold to the upper salary scales. The survey indicated that teachers grew less negative over time, though remained concerned that the

funding for the higher salary levels would not be permanently continued. Teachers saw the plan as promoting more effective co-ordination of school priorities, rather than a financial incentive to work harder. Surveys of head teachers indicated that they saw the plan as helping them identify ways the school could support teachers. The greatest benefit they saw was that the plan made teachers more aware of the school's objectives and helped them prioritise their work. However fewer than half of the head teachers thought that it gave teachers greater incentive to focus on student attainment. In schools where teachers perceived that the performance management system was working well, teachers were more likely to look at the education practices of those schools that ranked higher on league tables of student achievement, indicating that the process has made the teachers more interested in investigating ways to improve their practice.

An evaluation of **England's** performance pay programme (Atkinson *et al.*, 2004) found that the eligible teachers did raise student performance on the General Certificate of Secondary Education (GCSE) by about half a grade in lower secondary schools, although not equally for all subject teachers, e.g. scores for teachers of mathematics did not improve. The authors provided two cautions: since the study applied to teachers with eight years or more of experience, the differences between eligible and ineligible teachers could have been a function of experience. The second caution is that the study was done in schools with good performance management and information technology systems, which were able to submit extensive data, and those schools may not be representative of all secondary schools. In addition, the change in value added for a given teacher over time periods was as great as the difference between teachers in value added within a single time period. This concern with stability of value-added scores will be discussed further in Chapter 5.

There is great variety in the programmes developed across the world to address the challenge of attracting and retaining highly effective teachers in schools serving all students. While there is little evaluation of programmes to attract and retain teachers in hard-to-staff schools and shortage subject areas, school systems must see some benefit in continuing the programmes. However, well-designed evaluation programmes that compare different options have the potential to provide insights on how to make incentive programmes more effective. The experimental programmes that were explicit in identifying key issues in their design provide guidance on both design and implementation of incentive programmes. The Teacher Incentive Fund in the **U.S.**, described in greater detail in Chapter 1, is a federal program that provides grants to states or school districts to develop innovative teacher compensation systems. The wide variety of programmes that will be evaluated through the TIF may expand our understanding of what components of an incentive system motivate teachers to work more effectively to raise student achievement. If those can be expanded to more schools, we may create a more effective teacher force by both motivation and selection, as Lazear (2003) suggested.

ANNEX 2A

INCENTIVES IN EDUCATION

A description of the different incentives that are used in the educational sectors of several countries is presented in this annex. International cases show that incentives can take many forms: financial incentives, knowledge and skills-based incentives, incentives for hard-to-staff schools, incentives for shortage fields, pay-for-performance incentives, and others, including models that use a combination of these. The annex ends with a discussion of teacher attitudes toward incentives based on a review of studies conducted in the **U.S.** and **Australia**.

Financial incentives. In analysing financial incentives to attract and retain teachers, Hassel (2002) outlines four main bases for payment: knowledge and skills, the subject taught, the difficulty of the teaching assignment and school environment, and actual performance in the classroom. Determining the appropriate mix of payments to attract and retain teachers and to encourage their continuous improvement of knowledge and skills is the critical issue, and depends upon the conditions the school, school district or state/regional or national government is facing. While there have been evaluations of some programmes as the later chapters in this book will detail, there is no strong research base to guide the design of effective compensation systems.

Knowledge and skills-based incentives. Financial incentives to increase knowledge and skills, goes beyond traditional incentives of degrees and years of experience in the single salary schedule. Instead teachers have incentives to demonstrate their knowledge and skills either through external examinations or teacher evaluations based on a school system-selected model of teacher effectiveness. Knowledge and skills incentive programmes focus on the continuous improvement of teacher competencies likely to lead to improved student outcomes. In knowledge and skills programmes, teachers can receive additional compensation for a single year or multiple years before renewal is required, or for a lifetime, depending on the system.

Incentives for hard-to-staff schools. The third type of incentive is based on the “disamenities” of the school or its location. Providing incentives for teaching in schools that are hard-to-staff creates an opportunity to attract and retain teachers who otherwise might not have been willing to go to a particular school. Hard-to-staff schools include schools in remote areas, schools serving students from disadvantaged neighbourhoods, and schools serving a high proportion of minority or low-performing students. As noted earlier, it is these schools that have the most vacancies. Both new and experienced teachers assigned to the schools tend to leave as opportunities arise in schools with better leadership, greater resources, greater proximity to cities and services, or with students who are thought to be easier to teach. Continuing annual incentives will assist in recruiting and retaining teachers in these schools, but the size and type of the incentive will need to vary by the type and location of the school. In addition, school systems may offer housing assistance, moving expenses, recruitment and retention bonuses, and loan forgiveness.

Incentives for shortage fields. Another area of incentives is additional compensation for teaching certain subjects, often known as shortage subjects, such as foreign languages, mathematics, science, technology, or special education. This incentive is a response to the supply and demand problem in education. Fewer students majoring in technical areas have gone into teaching, in part because of the salary differentials between the public and private sectors. Therefore, shortage subject incentives are offered over and above the annual salary to teachers with special preparation and certification in the subjects needed. It makes sense to offer these incentives on an annual basis so that they can respond to the market, but it is not enough to offer them only as recruitment incentives: unless they are continued for as long as the shortage lasts, they will not continue to address the salary differentials that exist in the market. In addition to technical areas, special education presents a distinctive case because teachers must often earn both a teaching certification and special certification to meet the needs of students with disabilities. In addition, the intensive nature of the work and the amount of

paperwork involved for special education teachers, at least in **U.S.** schools, have led to shortages in teachers willing to teach in this area. Because the single-salary schedule does not recognise the additional time and effort required, some special education teachers move to classes serving students without disabilities. An incentive may draw them back. However, teachers do not support these incentives, as discussed later in this annex of on teacher attitudes toward incentives.

Non-financial incentives. Vegas and Umansky (2005) point out that incentives can range from non-financial – such as recognition and prestige, gratification of intrinsic motivation, job stability, working conditions with adequate resources and support – to financial incentives such as salary differentials and pensions. Clearly any of these may be more important to some teachers than others, especially at different stages of their careers. There is no definitive research on the effectiveness of non-financial incentives, but a study of prospective teachers looking at options for first job placements can provide insights about the relative importance of certain factors. Milanowski *et al.* (2007) found that prospective teachers are more concerned about the quality and supportiveness of the principal, the availability of an induction programme to support them in their first year of teaching, and curricular flexibility at the school than an increase in salary. Other issues that concern teachers about working conditions are the backgrounds and behaviour of students, co-operation from parents, availability of resources and support staff, and opportunities for collaboration and teamwork with other teachers. For teachers who are parents, opportunities to be at work the same hours their children are at school and to have summers free can be important incentives for entering and remaining in teaching.

Pay-for-performance incentives. The fourth type of financial incentive is known as pay for performance. This includes incentives based on a school's performance on specified indicators, such as test scores, improvements in test scores, student attendance, graduation or dropout rates. In the **U.S.**, the development of national and state examination systems, along with the database infrastructure to connect scores to schools, has made the use of test scores more popular as a basis for incentives. However, rewarding high test scores alone does not recognise the efforts of teachers as much as it rewards those schools that have high-achieving students. The question of how much of that achievement is based on teacher or school efforts as opposed to family background and prior education has led to the use of test score gains as an additional incentive metric. Otherwise a system creates another great disincentive to work at schools with lower-achieving students. While the use of test score gains requires sophisticated statistical analyses, known as the value-added model (VAM), to determine school effects as opposed to the effects of student background and experience, it is becoming an accepted practice (Chapter 5 discusses the issues regarding VAM). In school-based programmes, incentive pay can be distributed to teachers; teachers and principals; or teachers, principals and support staff depending on the design. Advocates of collaboration and teamwork at the schools favour such plans as everyone is able to benefit from the group's accomplishments. However, as mentioned earlier, such plans may encourage “free-riders” who do very little and reap the rewards of the hard work of their peers.

Pay-for-performance incentives can also be directed to individual teachers based on test scores or improvement in test scores of their students. As with school-based incentives, teacher incentives can be based on actual test score levels or the gain in test scores identified as occurring because the student spent that year with the teacher, or both. The pay-for-performance incentive is probably the most controversial. As noted in the beginning of this chapter, there can be great differences in the effectiveness of individual teachers, and researchers have found that the within-school differences between teachers are even greater than the between-school differences. While some argue this is exactly the reason to provide performance-based incentives to individual teachers, others argue that establishing competition among individual teachers at the school is the wrong way to move all students forward. Others question whether the VAM are sufficiently sophisticated and stable to make the necessary discrimination. One response has been to require the use of multiple years of data to increase the number of students used in the analysis and to ensure that a teacher is not being unfairly evaluated because

of characteristics of one year's students or other factors the teacher cannot control. Chapter 5 discusses these challenges in greater detail.

A third variety of pay for performance is additional salary for additional responsibilities. In some systems, this has been a traditional part of compensation. Teachers receive stipends for assuming responsibilities for extracurricular activities, chairing a grade/year level or department of the school, or working with a student-teacher. In current pay-for-performance systems, these are still options, but there are also incentives for teachers to work on the school improvement plan, coach peers in specific areas, lead professional development sessions or mentor less experienced teachers. All of these are focused on the continuous improvement of the school, its teachers and students.

The two ideas of incentives for knowledge and skills and for additional responsibilities (Sclafani and Tucker, 2006) come together in the Advanced Skills Teacher designation in **Australia, England and Wales**, the Plus teachers in **the Netherlands** and the Specialist Teachers in **New Zealand**, where the teachers are assessed on their skills in curriculum development, improving student achievement, supporting teacher professional development and mentoring new teachers. In their new roles, they are responsible for assisting other teachers in their own schools or, in the case of **Australia**, spending a portion of their time assisting teachers in other schools as well. The additional responsibilities carry additional compensation, ranging from bonuses and reduced teaching loads in **Australia, the Netherlands and New Zealand** to a higher salary level in **England and Wales**. In **England and Wales**, an AST at the top of the AST salary scale earns 43% more than a senior teacher at the top of the salary schedule.

Additional areas of incentives. There are additional areas of incentives that do not fall easily into Hassel's schema. The first of these is the incentive to enter the profession. Many countries provide incentives to college students and career changers to attract them to education. Scholarships and loan-forgiveness programmes are used to attract students into teacher preparation programmes. The challenge with scholarships is that many young people who attend the preparation programmes do not enter the teaching profession, thus making the incentive costly and less productive. Loan-forgiveness programmes address this issue by requiring repayment of the loan unless the teacher spends a contracted number of years teaching, usually ranging from three to five years. While there are reports of such programmes, there are no strong studies of their effectiveness.

Flexibility in how teachers use their time is a major incentive for some teachers. For example deferred salary leave plans, available in **Canada** since the 1970s and more recently in some **U.S.** school districts, enable teachers to defer a portion of their salary each year to allow them to schedule a leave with pay at a later time (Jacobson, Hickcox and Stevenson, 1996). This leave can be used for further study or travel to learn about other educational systems. Singapore (Sclafani, 2008) allows teachers to accrue a month of full pay for half-time service or half pay for full-time study for each year of service, for up to 12 months. Teachers are able to use this time in any way that will benefit the Singapore education system. Examples of the activities undertaken include further study in Singapore or abroad, teaching in an independent or foreign school, or working in the private sector to learn more about applications of education knowledge and skills.

Another version of flexibility in the use of time is the opportunity for senior teachers to engage in an increased variety of tasks and to reduce their work schedule with or without a small reduction in salary (OECD, 2005). Some local districts in **Norway** allow senior teachers to reduce time in teaching and general duties to participate in professional development, mentor new teachers, develop curriculum and advise other schools. In **Germany**, Brandenburg allows teachers to reduce teaching time by 50% for a 20% reduction in salary, an opportunity accepted by about 10% of the teachers. **The Netherlands** offers teachers who are 52-55 years old the option to reduce their work schedule by 10% for a 2.5% reduction in salary and teachers 56 and older the opportunity to reduce their schedule by 20% for a 5% reduction in pay. These options may enable systems to retain highly effective teachers who no longer wish to work a full schedule.

Positive working conditions are a major incentive for many teachers. In looking at prospective teachers in the **U.S.**, Milanowski *et al.* (2007) found that the quality of the school leader and the amount of choice in how to implement the curriculum were more important than salary levels in choosing a job. Teachers in **England** and **Wales** (OECD, 2005) responded very favorably to the 2003 *Raising Standards and Tackling Workload*, a national agreement among local education authorities and unions. This agreement reduced the amount of administrative/clerical duties assigned to teachers by adding support staff and providing them with better training to assume those responsibilities. It also phased in guaranteed, additional planning, preparation and assessment time for teachers. Over 97% of teachers surveyed responded that teaching and learning had improved because of the agreement. Darling-Hammond (1997) found that teachers who planned to stay in the profession had more positive views of their support from school leaders, the resources available to them, and their opportunity to have a voice in the decision-making at their schools. This is an important reminder to those developing incentives to attract and retain teachers: the incentive may get the teachers to the school, but those who are effective will not remain without the working conditions that enable them to thrive.

Combination incentive models. In the **U.S.**, a number of school districts have experimented with establishing new compensation systems that combine a number of different incentives with modified or totally redesigned salary schedules. To encourage more innovation in this area in high-needs schools, the U.S. Department of Education created the Teacher Incentive Fund in 2006.²⁰

In 2006 and again in 2008, USD 99 million was allocated to fund proposals from school districts and state departments of education, in collaboration with non-profit entities, to develop and implement performance-based teacher and principal compensation systems in high-needs schools. An additional USD 200 million was included in the State Fiscal Stimulus Funding for 2009-2010. The amount of the request from the current U.S. Administration for the 2010 budget is USD 517 million. The programmes must use gains in student test scores in addition to classroom evaluations administered multiple times during the school year, similar to the TAP programme described below. Programmes are also expected to provide incentives for teachers to assume additional responsibilities and leadership roles. Awards ranged from USD 14 million over three years for large school districts to less than USD 1 million for small ones. All grants are provided technical assistance and evaluation through the Center for Educator Compensation Reform, a partnership of educational organisations. The evaluations of the varied programmes should provide valuable information for those who are interested in designing incentive systems.

The Teacher Advancement Program (TAP) is a **U.S.** programme that builds on knowledge and skills incentives with pay for performance to identify, develop, and reward highly effective teachers. It began as a privately-funded programme of the Milken Family Foundation in the 1990s and is now run through the non-profit National Institute for Effectiveness in Teaching (NIET) (The System for Teacher and Student Advancement, 2009). Schools that wish to implement TAP must show that 75% of the teachers approve participation. The programme has four components: multiple career paths, embedded professional development, multiple evaluations of teacher practice, and analysis of student performance. TAP establishes three levels of teacher career positions, starting with career teachers and moving to mentor and master teacher, according to a teacher's interests and accomplishments. As teachers move to higher levels, their compensation increases. The programme provides time during the weekly schedule for teachers to study and collaborate so they are working continuously to improve their practice, and using data and student work to guide their work. What distinguishes the TAP from other programmes is its focus on comprehensive teacher assessment and improving teacher practice. Multiple trained assessors observe teachers four to six times a year and analyse teacher performance using the research-based TAP Teaching Skills, Knowledge and Responsibilities Performance Standards. There are pre- and post-conferences to ensure that teachers are provided feedback that will help them improve their practice. Finally, the evaluation includes an analysis of the test scores of students in the individual teacher's classroom as well as of

students in the school as a whole. The determination of incentives added to the teacher salary are based first on the teacher's position on the career levels, with mentor and master teachers receiving additional compensation for the additional responsibilities they have assumed. In addition, all teachers are eligible for incentive pay, 50% of which is based on the teacher evaluations, 30% on individual classroom achievement growth, and 20% on school-wide achievement growth. A small number of top teachers are selected nationwide each year to receive USD 25 000 awards through the incentive programme. For the most part, teachers in TAP earn incentives in amounts between USD 1 000 and 3 000 annually at the school level. TAP is operational in 14 states, with some districts implementing it district-wide, but many allowing individual schools to opt into participation. It is used in charter schools in three cities. TAP is one of the teacher incentive models being used in some grant programmes designed under the U.S. Department of Education's Teacher Incentive Fund.

Teacher attitudes toward incentives

It is important to know how teachers respond to the new focus on incentives, although there is limited research in this area. In a Public Agenda survey of **U.S.** teachers (Farkas, Johnson and Duffett, 2003), a majority of teachers were in favour of administrators having some discretion in rewarding outstanding teachers. In addition, 70%, supported financial incentives for teachers in low-income neighbourhoods with low-performing schools, 67% supported paying more to teachers who consistently expend more time and effort at school, and 57% supported incentives for National Board Certified teachers. On the other hand, most teachers did not support paying teachers of shortage subjects, such as mathematics or science, more than other teachers, nor did they support paying more to teachers whose students score high on examinations. Additional pay for improving student achievement was supported by about half of the teachers, although the majority of teachers worried that financial incentives in this area would reduce collaboration and positive school cultures. In a later survey of teachers in the State of **Washington**, DeArmond and Goldhaber (2008) found very similar results: 72% were in favour of additional pay for teaching in high-needs schools, but only 41% were in favour of financial incentives for shortage-subject areas. Overall, 82% opposed merit pay, but younger teachers were more supportive than their experienced colleagues, and upper-secondary school teachers were more supportive of both merit- and shortage-subject-area pay than were primary school teachers.

Opinions in the abstract, however, are different from those given when faced with an actual incentive system. Lewis and Springer (2008) surveyed attitudes of teachers and administrators who chose not to participate in the **Texas** Educator Excellence Grant programme, which provides school-wide awards to schools that are in the top half of schools in percentages of disadvantaged students and that have demonstrated academic success or successful improvement on the state examinations. Although 95% of the eligible schools chose to participate, the remaining 5% – a total of 51 schools across the state – did not. Lewis and Springer found their reasons for not participating were based more on the process than on opposition to incentive policies *per se*. Teachers did not think that the process of selecting schools was fair. One group felt that an individual school should not receive the award without acknowledging and sharing the award with feeder or pipeline schools that contributed to the students' achievement. In another case, a school indicated that another school in the district had outperformed them and yet received nothing. Accepting the money might cause conflicts with that school. Teachers were also concerned that decisions about how to allocate the money would be unfair and would cause dissension among their ranks. Finally, particularly in small schools, teachers did not feel the amount of money was worth the time and effort required to implement the programme. Reactions to the system improved after the state revised its procedures and provided greater time for schools to develop their implementation plans.

Ballou and Podgursky (1993) looked at survey data using a limited number of questions on teacher attitudes toward merit pay from the 1987-88 Schools and Staffing Surveys and found that teachers were not negative about merit pay in their districts, even if they did not receive additional funds. In addition, teachers of disadvantaged

students and low-achieving students were generally supportive. This may be the result of their concluding that they had a harder job and deserved additional compensation. Jacob and Springer (2008) surveyed teachers in Hillsborough County, **Florida** about their attitudes toward the Florida incentive programme as well as their general attitudes toward incentives. While over 85% of those surveyed indicated that school districts should place more emphasis on increasing the base salary than on performance incentives, they were more negative about school-wide and group-based incentives than about individual teacher incentives – 59%, 74% and 48% respectively. They were concerned about the impact of incentives on the collaborative culture of teaching. Only about a third of the teachers thought incentives would cause teachers to work harder. When asked on what basis incentives should be rewarded, teachers strongly backed criteria related to additional degrees and years of experience, as well as time spent on professional development. Other areas that over 60% of the teachers agreed were important were: collaboration with peers, efforts to involve parents in their child's education, teaching in hard-to-staff schools or shortage subjects, and the results of independent evaluations of portfolios of teachers' or students' work. Over 50% of the teachers surveyed agreed that National Board Certification and performance evaluations by peers were important. Incentives for student gains and high test scores were supported by 46% and 36% of the teachers respectively – so claims that there is no support for financial incentives based on test scores are exaggerated.

Jacob and Springer (2008) were able to correlate survey responses on incentives to teacher demographics. Again, the younger teachers were more positive about financial incentives than their more experienced peers. Teachers in secondary schools and schools with higher proportions of minority students were more supportive. Three other findings were also important: teachers who believed their principals were good leaders were more positive about incentives, as were teachers with high self-efficacy measures and those who were characterised as risk takers. A survey of teachers in **Australia** in their first ten years of experience found that these younger teachers were more interested in performance-based pay than were more experienced teachers (Skillbeck and Connell, 2004). In a 2006 survey of sophomores and juniors at a large midwestern U.S. university who planned to be classroom teachers, Milanowski (2006) asked the students to rate and then rank four pay scenarios, a) pay for individual performance based on objective measures, such as student achievement; b) pay for developing knowledge and skills, assessed by the principal; c) pay for group performance based on test scores; and d) pay increases for remaining as a teacher with passable performance. The students preferred the first two methods of determining pay, and they responded to the last two in very similar ways, with pay for group performance favoured by only a small percentage of the students. Studies indicating differences in teacher attitudes between less and more experienced teachers could indicate that the financial incentives will become a more accepted practice as more experienced teachers retire.

As discussed earlier, many of the programmes in schools and districts, such as the Vaughn Next Century Learning Center (Kellor, 2005) and the **Denver** Public Schools (Koppich, 2008) required the move to performance-based pay for new teachers, and made it optional for teachers already in the system. This enabled the programmes to offer incentives without fighting a sceptical group of current teachers. Kellor reported that annual attitude surveys at Vaughn showed more positive reactions to the performance pay systems over time, with new teachers expressing more positive agreement than other teachers that the system encouraged them to focus on improving instruction and student achievement and helping one another improve. From a recruitment and retention perspective, 79% of new teachers agreed that the system made Vaughn a more attractive choice, but only 34% agreed that it was an important reason they chose to teach there. Only 49% of more experienced teachers agreed that it made Vaughn a more attractive choice. In Denver, current teachers could use a web-based calculator to determine which salary system is in their best interests and could opt in for up to seven years from the commencement of the programme. About half of the Denver teachers were in the programme during the 2007-2008 school year.

Springer *et al.* (2007) completed an evaluation of the first year implementation of the Governor's Educator Excellence Grant (GEEG) programme in **Texas**. They looked at teacher attitudes toward the programme and various criteria used to identify award recipients. The programme made three-year grants to eligible schools in the top third of those schools serving a larger proportion of disadvantaged students and which had earned an exemplary or recognised rating in the Texas accountability system. Ninety-nine schools developed their plans with the input and approval of their teachers and administrators according to criteria set by the state: 75% of the funds had to go to teachers based on student performance and teacher collaboration, with three other subjective criteria optional. The other 25% could be used for professional development or allocated to bonuses for staff other than teachers. Teachers had generally positive attitudes about the programme, although the authors suggest that may have been in part because of their participation in its development. In attitudes toward objective versus subjective measures, teachers were more positive about the objective measures. Although all schools used the required subjective criterion of collaboration, only half used the optional measure of teacher initiative and commitment, and 15 schools included teaching in hard-to-staff or high-turnover positions as criteria for awards. Teachers did not believe that the programme diminished collaboration among teachers and indicated that they had changed their practice to include more practices believed to positively affect student achievement. However, some teachers noted in their survey responses that they would have changed their practice as a function of their growth and improvement as teachers, even without the GEEG.

Springer *et al.* (2008) evaluated the second and third components of the **Texas** performance-based pay programme, the Texas Teacher Educator Excellence Program (TEEG) and the District Awards for Teacher Excellence (DATE). TEEG provides opportunities for schools in the top half of Texas schools based on the proportion of disadvantaged students enrolled and with high scores or high improvement levels on the Texas state accountability exams. Over 1 000 schools participate in TEEG, although the specific schools eligible changes each year, based in part on the volatility of the Comparable Improvement rankings in the accountability system and the availability of state budget funds. DATE is open to all districts on a voluntary basis and 243 districts participated in the programme in its first year. The evaluators found positive attitudes toward the programmes, especially among the inexperienced staff and those who received awards.

NOTES

1. National Advisory Committee for the Ministry of Education and Science (2003), pp.68-69.
2. Skillbeck, M. and H. Connell (2003), p. 30.
3. Arkansas Department of Education (2009), *Arkansas' Equity Plan*. Retrieved April 2009 from http://www.arkansased.org/nclb/pdf/equity_plan_1008.pdf.
4. Ministry of Education (2003), p.65,
5. Halász *et. al.* (2004), p. 22.
6. Eurydice Database on Education Systems in Europe.
7. OECD (2004), *Education at a Glance: OECD Indicators*, OECD, Paris.
8. Wagner *et. al.* (2004), p. 21.
9. Ross and Hutchings (2003), p. 66.
10. Weppler, M. and S. Bourke (1994), p. 2.
11. Ross and Hutchings (2003), p. 67.
12. Meesters, M. (2003), p. 67.
13. New Zealand Secondary Teacher Collective Agreement, Retrieved May 2009 from <http://www.ppta.org.nz/index.php/collective-agreements/stca/89-part-three?start=10>.
14. Cros, F. and J-P. Obin, (2003), pp. 52-54.
15. New Zealand Secondary Teacher Collective Agreement.
16. Meesters, M. (2003), pp. 63-64.
17. The Teacher Advancement Program is fully described in Chapter 1.
18. Minnesota Guidelines for Q Comp Applications found at <http://education.state.mn.us/mdeprod/groups/QComp/documents/Form/032806.pdf>.
19. Guidelines for Arizona's Career Ladder Program found at <http://www.ade.state.az.us/asd/CareerLadder/>.
20. U.S. Department of Education Teacher Incentive Fund Website: <http://www.ed.gov/programmes/teacherincentive/index.html>.

References

Arizona Department of Education (2009), "Arizona Career Ladder", Little Rock, Arkansas, <http://www.ade.state.az.us/asd/CareerLadder>, accessed April 2009.

Arkansas Department of Education (2009), "Arkansas' Equity Plan", Little Rock, Arkansas, http://www.arkansas.gov/search/gsearch.php?profile=arkedu.state.ar.us&collection=State_Sites&words=Teacher+Incentive+program&sort=&start=10&filetype_filter=, accessed April 2009.

Atkinson, A., S. Burgess, B. Croxson, P. Gregg, C. Propper, H. Slater and D. Wilson (2004), "Evaluating the Impact of Performance-related Pay for Teachers in England", *Working Paper No. 04/113*, Centre for Market and Public Organisation, Bristol, England.

Ballou, D. and M. Podgursky (1993), "Teacher Attitudes Toward Merit Pay: Examining Conventional Wisdom", *Industrial and Labor Relations Review*, Vol. 47, No. 1, pp. 50-61.

Behrstock, E. and J. Akerstrom (2008), "Performance Pay in Houston, A Case Summary", Center for Educator Compensation Reform, Learning Points Associates, Chicago, Illinois.

Booker, K. and S. Glazerman (2008), "Does the Missouri Teacher Career Ladder Program Raise Student Achievement?", paper presented at the National Conference on Performance Incentives: Their Growing Impact on American K-12 Education, Nashville, February.

Boozer, M. (1999), "The Design and Evaluation of Incentive Schemes for Schools: Evidence from South Carolina's Teacher Incentive Pay Project", paper prepared for the National Academy of Sciences Conference on Devising Incentives to Improve Human Capital, Irvine, CA, December.

Chait, R. (2007), "Current State Policies that Reform Teacher Pay: An Examination of Pay-for-Performance Programs in Eight States", Center for American Progress, Washington, D.C.

Clotfelter, C., E. Glennie, H. Ladd and J. Vigdor (2008), "Would Higher Salaries Keep Teachers in High-Poverty Schools? Evidence from a Policy Intervention in North Carolina", *Journal of Public Economics*, 92, 5-6, pp.1352-1370.

Cros, F. and J-P. Obin (2003), "Attracting, Developing and Retaining Effective Teachers: Country Background Report for France", OECD, Paris.

Crouch, L. (2005), "Political Economy, Incentives, and Teachers' Unions: Case Studies in Chile and Peru" in E. Vegas (ed.) *Incentives to Improve Teaching: Lessons from Latin America*, World Bank, Washington, pp. 389-424.

Danielson, C. (2008), *The Handbook for Enhancing Professional Practice: Using the Framework for Teaching in Your School*, ASCD, Alexandria, VA.

Darling-Hammond, L. (1997), *Doing What Matters Most: Investing in Quality Teaching*, National Commission on Teaching and America's Future, New York.

Dowling, J., S. Murphy and B. Wang (2007), "The Effects of the Career Ladder Program on Student Achievement", Evaluation Report, Sheila Murphy Associates, Phoenix, Arizona.

Duflo, E., R. Hanna and S. Ryan (2007), "Monitoring Works: Getting Teachers to Come to School", *NBER Working Paper 11880*, National Bureau of Economic Research, Cambridge, Massachusetts.

Eisenberg, N. (2008), *The Performance of Teachers in Chilean Public Elementary Schools: Exploring Its Relationship with Teacher Backgrounds and Student Achievement, and Its Distribution Across Schools and Municipalities*, Ph. D. Thesis, University of California, Los Angeles. Cited in Manzi (2008).

Eurydice (2009), "Levels of Autonomy and Responsibilities of Teachers in Europe", European Commission, Brussels, http://eacea.ec.europa.eu/ressources/eurydice/pdf/094EN/094EN_001_PRE.pdf, accessed May 2009.

- Farkas, S., J. Johnson and A. Duffett** (2003), "Stand by Me: What Teachers Really Think About Unions, Merit Pay and Other Professional Matters", *Public Agenda Report*, Public Agenda, New York.
- Field, E.** (2009), "Educational Debt Burden and Career Choice: Evidence from a Financial Aid Experiment at NYU Law School", *American Economic Journal: Applied Economics*, Vol I, No., pp. 1-21.
- Figlio, D. and L. Kenny** (2007), "Individual Teacher Incentives and Student Performance", *Working Paper No. 8*, National Center for Analysis of Longitudinal Data in Education Research, Washington, D.C.
- Glazerman, S., A. McKie and N. Carey** (2009), *An Evaluation of the Teacher Advancement Program (TAP) in Chicago: Year One Impact Report*, Mathematica, Washington, D.C.
- Glewwe, P., N. Ilias and M. Kremer** (2008), "Teacher Incentives in Developing Countries: Recent Experimental Evidence from Kenya", paper presented at the National Conference on Performance Incentives: Their Growing Impact on American K-12 Education, Nashville, February.
- Goldhaber, D.** (2008), "The Politics of Teacher Pay Reforms", paper presented at the National Conference on Performance Incentives: Their Growing Impact on American K-12 Education, Nashville, TN, February.
- Hassel, B.** (2002), "Better Pay for Better Teaching: Making Teacher Compensation Pay Off in an Age of Accountability", *21st Century Schools Project*, Progressive Policy Institute, Washington, D.C.
- Halász, G., P. Santiago, M. Ekholm, P. Matthews and P. McKenzie** (2004), "Country Note of Germany for the elaboration of the report 'Attracting, Developing and Retaining Effective Teachers'", OECD, Paris.
- Hanushek, E., J. Kain and S. Rivkin** (2004), "Why Public Schools Lose Teachers", *Journal of Human Resources*, Vol. 39 No. 2, pp.326-354.
- Hopkins, D. and E. Ahtaridou** (2008), "Teacher Incentives and Stimuli: The Case of England", paper presented at the International OECD/Mexico Joint Conference, The Quality of Education, Mexico City, December.
- Ingersoll, R.** (2008), "Core Problems: Out-Of-Field Teaching Persists in Key Academic Courses and High-Poverty Schools", Education Trust, Washington, D.C., <http://www2.edtrust.org/NR/rdonlyres/0D6EB5F1-2A49-4A4D-A01B-881CD2134357/0/SASSreportCoreProblems.pdf>, accessed April 2009.
- Jacob, B. and M. Springer** (2008), "Teacher Attitudes Toward Pay For Performance: Evidence from Hillsborough County, Florida", paper presented at the National Conference on Performance Incentives: Their Growing Impact on American K-12 Education, Nashville, February.
- Jacobson, S., E. Hickcox and R. Stevenson** (1996), *School Administration: Persistent Dilemmas In Preparation and Practice*, Greenwood Publishing Group, Santa Barbara, CA.
- Johnson, J.** (2005), "State Financial Incentive Policies for Recruiting and Retaining Effective New Teachers in Hard-to-Staff Schools", Education Commission of the States Clearinghouse on Teacher Incentives, Denver, <http://www.ecs.org/clearinghouse.61/61/6161.htm>, accessed May 2009.
- Juerges, H., W. Richter and K. Schneider** (2004), "Teacher Quality and Incentives: Theoretical and Empirical Effects of Standards on Teacher Quality" *CESifo Working Paper No. 1296*, CESifo, Munich, Category 1: Public Finance. Available at <http://www.CESifo.de>.
- Kellor, E.** (2005), "Catching Up with the Vaughn Express: Six Years of Standards-Based Teacher Evaluation and Performance Pay", *Education Policy Analysis Archives*, Vol. 13, No. 7 (January 23), <http://epaa.asu.edu/epaa/v13n7/>, accessed March 2009.
- Kirby, S., S. Naftel and M. Berends** (1999), *Staffing At-Risk School Districts in Texas: Problems and Prospects*, RAND Education, Santa Monica, California, http://www.rand.org/pubs/monograph_reports/2007/MR1083.pdf, accessed May 2009.
- Koppich, J.** (2008), "Toward a More Comprehensive Model of Teacher Pay", paper presented at the National Conference on Performance Incentives: Their Growing Impact on American K-12 Education, Nashville, TN, February.
- Ladd, H.** (1999), "The Dallas School Accountability and Incentive Program: An Evaluation of Its Impacts on Student Outcomes", *Economics of Education Review*, 18, No. 1, pp. 1-16.

- Ladd, H.** (2007), "Teacher Labor Markets in Developed Countries" in *The Future of Children*, Vol. 17, No. 1.
- Lavy, V.** (2002), "Evaluating the Effect of Teachers' Group Performance Incentives on Pupil Achievement", *Journal of Political Economy*, Vol. 110, No. 6, pp.1286-1317.
- Lavy, V.** (2004), "Performance Pay and Teachers' Effort, Productivity and Grading Ethics", *NBER Working Paper No. 10622*, National Bureau of Economic Research, Cambridge, MA.
- Lazear, E.** (2003), "Teacher Incentives", *Swedish Economic Policy Review*, Vol. 10, No. 2, pp. 179-214.
- Leon, G.** (2008), *Calidad Docente y Rendimiento Escolar en Chile: Evaluando la Evaluación [Teacher Quality and Student Performance: Evaluating the Evaluation]*, Universidad Católica de Chile, cited in Manzi (2008).
- Lewis, J. and M. Springer** (2008), "Performance Incentives in Texas: Why Schools Chose Not to Participate", paper presented at the National Conference on Performance Incentives: Their Growing Impact on American K-12 Education, Nashville, TN, February.
- Loeb, S., L. Darling-Hammond and J. Luczak** (2005), "How Teaching Conditions Predict Teacher Turnover in California Schools", *Peabody Journal of Education*, Volume 80, Issue 3 July, pp. 44-70.
- Loeb, S., L. Darling-Hammond and J. Luczak** (2006), "Teacher Turnover: The Role of Working Conditions and Salaries in Recruiting and Retaining Teachers", Stanford University School of Education, Palo Alto, CA.
- Manzi, J.** (2008), "Individual Incentives and Teacher Evaluation: The Chilean Case", paper presented at the International OECD/Mexico Joint Conference, The Quality of Education, Mexico City, December.
- Manzi, J., K. Strasser, E. San Martín and D. Contreras** (2008), *Quality of Education in Chile: Final Report of the Interamerican Development Bank Project*, Washington, D.C. Cited in Manzi (2008).
- Marsden, D. and R. Belfield** (2005), "Performance Pay for Teachers: Lining Individual and Organisational-Level Targets", *Discussion Paper No 703*, Centre for Economic Performance, London.
- McEwan, P. and L. Santibáñez** (2005), "Teacher and Principal Incentives in Mexico" in E. Vegas (ed.) *Incentives to Improve Teaching: Lessons from Latin America*, World Bank, Washington, D.C., pp. 213-253.
- Meesters, M.**, (2003), "Attracting, Developing and Retaining Effective Teachers: Country Background Report: Netherlands", OECD, Paris, p. 67.
- Milanowski, A., H. Longwell-Grice, F. Saffold, J. Jones, A. Odden and K. Schomisch** (2007), "Recruiting New Teachers to Urban School Districts: What Incentives Will Work", *Working Paper 11*, School Finance Redesign Project, Center on Reinventing Public Education, University of Washington, Seattle, WA.
- Mizala, A. and P. Romaguera** (2005), "Teachers' Salary Structure and Incentives in Chile" in E. Vegas (ed.) *Incentives to Improve Teaching: Lessons from Latin America*, World Bank, Washington, D.C., pp. 103-150.
- Muralidharan, K. and V. Sundararaman** (2008), "Teacher Incentives in Developing Countries: Experimental Evidence from India", paper presented at the National Conference on Performance Incentives: Their Growing Impact on American K-12 Education, Nashville, February.
- National Advisory Committee for the Ministry of Education and Science** (2003), *Attracting, Developing and Retaining Effective Teachers: Country Background Report for Sweden*, Ministry of Education and Science, Stockholm.
- National Council on Teacher Quality** (2008), "State Teacher Policy Yearbook – National Summary – What States Can Do To Retain Effective New Teachers", Washington, D.C., http://www.nctq.org/stpy08/reports/stpy_national.pdf, accessed 2009.
- New Zealand Ministry of Education** (2004), "New Zealand Secondary Teacher Collective Agreement 2004-2007", Retrieved May 2009 from <http://www.ppta.org.nz/index.php/collective-agreements/stca/89-part-three?start=10>.
- Odden, A., C. Kelley, H. Heneman and A. Milanowski** (2001), "Enhancing Teacher Quality Through Knowledge- and Skills-based Pay", *Consortium for Policy Research in Education Policy Brief, RB-34*, University of Pennsylvania, Philadelphia.
- Odden, A. and M. Wallace** (2007), *Rewarding Teacher Excellence: A Teacher Compensation Handbook for State and Local Policy Makers*, Consortium for Policy Research in Education, University of Wisconsin, Madison, WI., [http://cpre.wceruw.org/publications/TComp%20Handbook%20Feb%2028%2007%20Final%20\(3.05.07\).pdf](http://cpre.wceruw.org/publications/TComp%20Handbook%20Feb%2028%2007%20Final%20(3.05.07).pdf), accessed March 2009.

Odden, A. and M. Wallace (2008), *How to Achieve World Class Teacher Compensation*, Freeload Press, St. Paul, MN., <http://www.freeloadpress.com/bookDetail.aspx?bld=1077>, accessed March 2009.

OECD (2004), *Education at a Glance: OECD Indicators*, OECD, Paris.

OECD (2005), *Teachers Matter: Attracting, Developing and Retaining Effective Teachers*, Organisation for Economic Co-operation and Development (OECD), Paris.

OECD (2007), *PISA 2006: Science Competencies for Tomorrow's World*, OECD, Paris.

Parker, C. (2005), "Teacher Incentives and Student Achievement in Nicaraguan Autonomous Schools", in E. Vegas (ed.) *Incentives to Improve Teaching: Lessons from Latin America*, World Bank, Washington, D.C., pp.359-387.

Prince, C., P. Schuermann, J. Guthrie, P. Witham, A. Milanowski and C. Thorn (2008), "The Other 69 Percent: Rewarding the Performance of Teachers of Non-Tested Subjects and Grades", Center for Educator Compensation Reform, U. S. Department of Education, Office of Elementary and Secondary Education, Washington, D.C., <http://www.cecr.ed.gov/guides/other69Percent.pdf>, accessed April 2009.

Quartz, K., K. Lyons, K. Masyn, B. Olsen, L. Anderson, A. Thomas, J. Goode and E. Horng (2004), "Urban Teacher Retention Policy: A Research Brief", *Urban Teacher Education Collaborative Working Papers*, University of California at Los Angeles, Los Angeles, CA.

Ritter, G., M. Holley, N. Jensen, B. Riffel, M. Winters, J. Barnett and J. Greene (2008), *Year Two Evaluation of the Achievement Challenge Pilot Project in the Little Rock Public School District*, Department of Education Reform, University of Arkansas, Fayetteville, January 22, http://www.uark.edu/ua/der/Research/merit_pay/year_two/Full_Report_without_Appendices.pdf, accessed April 2009.

Ross, A. and M. Hutchings (2003), *Attracting, Developing and Retaining Effective Teachers in the United Kingdom of Great Britain and Northern Ireland*, Institute for Policy Studies in Education, London Metropolitan University, London.

Santibáñez, L., J. Martinez, A. Datar, P. McEwan, C. Setodji and R. Basuto-Dávila (2007), "Breaking Ground: Analysis of the Assessment System and Impact of Mexico's Teacher Incentive Program 'Carrera Magisterial'", *RAND Technical Report*, RAND Corporation, Santa Monica, CA.

Santibáñez, L., G. Vernez and P. Razquin (2005), "Education in Mexico: Challenges and Opportunities", *RAND Documented Briefing*, RAND Corporation, Santa Monica, CA.

Sclafani, S. (2008), "Rethinking Human Capital in Education: Singapore as a Model for Teacher Development", paper prepared for The Aspen Institute Education and Society Program, Aspen Institute, Washington, D.C., <http://www.aspeninstitute.org/sites/default/files/content/docs/education%20and%20society%20program/SingaporeEDU.pdf>, accessed May 2009.

Sclafani, S. and M. Tucker (2006), *Teacher and Principal Compensation: An International Review*, Center for American Progress, Washington, D.C., http://www.americanprogress.org/issues/2006/10/teacher_compensation.html/pdf/education_report.pdf, accessed May 2009.

Skillbeck, M. and H. Connell (2004), "Teachers for the Future: The Changing Nature of Society and Related Issues for the Teaching Workforce," Ministerial Council on Education, Employment, Training and Youth Affairs, Canberra, <http://www.mceetya.edu.au/mceetya/>, accessed April 2009.

Smith, S. and B. Mickelson (2000), "All That Glitters Is Not Gold: School Reform in Charlotte-Mecklenburg", *Education Evaluation and Policy Analysis*, Vol. 22, No. 2, pp.101-127.

Solomon, L., J. White, D. Cohen and D. Woo (2007), *The Effectiveness of the Teacher Advancement Program*, National Institute for Excellence in Teaching, Santa Monica, CA.

Springer, M., M. Podgursky, J. Lewis, M. Ehlert, C. Gardner, B. Ghosh-Dastidar, O. Lopez, C. Patterson and L. Taylor (2007), *Governor's Educator Excellence Grant (GEEG) Program: Year One Evaluation*, National Center on Performance Incentives, Vanderbilt Peabody College, Nashville, <http://ritter.tea.state.tx.us/opge/progeval/TeacherIncentive/index.html>, accessed April 2009.

Springer, M., M. Podgursky, J. Lewis, M. Ehlert, B. Ghosh-Dastidar, T. Gronberg, L. Hamilton, D. Janssen, O. Lopez, B. Stecher and L. Taylor (2008), *Texas Educator Excellence Grant (TEEG) Program: Year Two Evaluation*, National Center on Performance Incentives, Vanderbilt Peabody College, Nashville, TN., http://ritter.tea.state.tx.us/opge/progeval/TeacherIncentive/TEEG_120108.pdf, accessed April 2009.

Steele, J., R. Murnane and J. Willett (2009), "Do Financial Incentives Help Low-Performing Schools Attract and Keep Academically Talented Teachers? Evidence from California", *NBER Working Paper 14780*, National Bureau of Economic Research, Cambridge, MA.

Taut, S., V. Santelices and P. Thibaut (2008), "Explicating Expected Consequences and Underlying Assumptions of a National Teacher Evaluation System", paper presented at the Annual Conference of the American Evaluation Association, Denver, CO, November. Cited in Manzi, (2008).

Teacher Development Authority (2003), "Raising Standards and Tackling Workload --A National Agreement", Government of the United Kingdom, London, http://www.tda.gov.uk/upload/resources/pdf/n/na_standards_workload.pdf, accessed April 2009.

The System for Teacher and Student Advancement (2009) "About Us", <http://www.tapsystem.org/about/about.taf?page=niet>, accessed July 2009.

Urquiola, M. and E. Vegas (2005), "Arbitrary Variation in Teacher Salaries" in E. Vegas (ed.) *Incentives to Improve Teaching: Lessons from Latin America*, World Bank, Washington, D.C., pp.187-213.

Vegas, E. (2007), "Teacher Labor Markets in Developing Countries", *The Future of Children*, Vol. 17, No. 1, pp. 219-232.

Vegas, E. and I. Umansky (2005), "Improving Teaching and Learning Through Effective Incentives", in E. Vegas (ed.) *Incentives to Improve Teaching: Lessons from Latin America*, World Bank, Washington, D.C., pp.1-20.

Wagner, A., P. Santiago, C. Thieme and D. Zay (2004), "Country Note of Switzerland for the elaboration of the report 'Attracting, Developing, and Retaining Effective Teachers'", OECD, Paris.

Webster, E., M. Wooden and G. Marks (2005), "Teaching and the Teacher Labour Market: The Case for Reform", *The Australian Economic Review*, Vol. 38, No. 1, pp. 91-98.

Wepppler, M. and S. Bourke (1994), "Advanced Skills Teachers: Dilemmas and Demands", paper presented at Annual Conference of the Australian Association for Research in Education, Newcastle, NSW.

CHAPTER 3

Design Components of Incentive Pay Programmes in the Education Sector

Introduction	66
Incentive structure	67
• Rank-order tournaments.....	67
• Fixed performance contract.....	68
Unit of accountability	71
• Individual unit of accountability.....	71
• Group unit of accountability.....	72
• Hybrid models of accountability.....	74
Performance measures	75
• Educational inputs.....	75
• Educational outputs.....	79
• Multiple measures – linking educational inputs and outputs.....	82
Performance standards and thresholds	83
• Linear models.....	83
• Step functions.....	83
• Limited linear models.....	84
• Threshold levels.....	85
Size and distribution of bonus awards	85
• Size of bonus.....	85
• Distribution of award.....	87
• Non-monetary incentives.....	88
Payout frequency	89
Cultural considerations	90
Conclusion	91
Notes	92
References	94

INTRODUCTION

Education reform is increasingly focused on human capital formation because high-quality teachers are seen as the most direct and effective pathways to improving student achievement. Based on the argument that prevailing compensation practices provide weak incentives for teachers, and that inefficiencies arise from rigidities in current compensation policies, several national systems of public education have explored teacher compensation reforms (Podgursky and Springer, 2007).¹ Proponents argue that financial incentives can motivate teachers to achieve higher levels of performance, entice more effective teachers to join or remain in the teaching profession, and align teacher behaviours and interests with institutional goals.

In spite of the intuitive appeal incentive pay has for some stakeholders, a sturdy and influential base of individuals and organisations fundamentally oppose its use in education (Eberts, 2007; Goldhaber, 2009; Kingdon and Teal, 2008). Opponents contend that such pay renders schools less effective by crowding out intrinsic rewards; that is, teachers will lose interest in the profession as they are increasingly rewarded for student outcomes. Critics further argue that the education system lacks appropriate measures for evaluating teacher performance, that reward programs will decrease teacher collaboration, and that the current body of evidence on the impact of pay-for-performance programmes is inconclusive.

Recent experimental and quasi-experimental evidence paints a mixed picture of the impact of teacher incentive-pay programmes. Muralidharan and Sundararaman (2008) and Lavy (2002, 2007) found that teacher incentive programs in **India** and **Israel**, respectively, improved student outcomes and promoted positive changes in teacher behaviour and/or classroom pedagogy. Glewwe, Ilias and Kremer (2008) similarly reported that students instructed by teachers eligible to receive a bonus award in **Kenya** demonstrated better scores on high-stakes tests; however, no discernible impact was found on low-stakes tests taken by treatment group students or on the same students when they took high-stakes tests during the post-intervention school year.

The design components of incentive programmes in the education sector have received little attention in the literature, particularly with regard to the relative advantages and limitations of different options for measuring and rewarding teachers. Since the design of an incentive programme can lead to dramatic differences in its impact on students, teachers, and administrators, it is important to learn more about how specific design elements affect teachers' responses.

The following discussion is largely informed by the theoretical literature on incentive systems as well as empirical evidence from evaluations of past and present programmes and policies in the **United States (U.S.)** and other countries. We draw on experiences from both within and outside the education sector to assess specific design components of incentive-pay programmes and policies and their potential impact on student outcomes, teacher attitudes and behaviour, and institutional dynamics. Specifically, we assess:

- incentive structure;
- unit of accountability;
- performance measures;
- performance standards and thresholds;
- size and distribution of bonus awards;
- payout frequency; and
- cultural considerations.

This chapter presents a brief discussion of the relationship between incentive-pay programme characteristics and cultural context before offering concluding remarks.

INCENTIVE STRUCTURE

Incentive structure refers to the scheme or mechanism that guides the allocation of awards in a pay-for-performance system. In some cases only a limited number of employees can earn an award, while in others any employee who meets a predetermined performance standard will receive an award. The two main forms of incentive structure are rank-order tournaments and fixed performance contracts, each of which, under certain conditions, could be an optimal labour contract yielding the first best outcome. Both forms represent very different compensation schemes with distinct advantages and limitations.

Rank-order tournaments

Rank-order tournaments are incentive structures that limit performance-based rewards to a fixed percentage of the relevant work group. The key feature that distinguishes tournaments from other incentive-pay structures is that compensation depends on relative performance rather than absolute performance. For example, an individual level pay-for-performance experiment implemented in **Israel** and evaluated by Lavy (2004) compared the performance of teachers in similar grades and subjects and then awarded individual bonuses to the teachers with positive gains scoring in the top performance quartiles.

A primary strength of tournament incentive structures is lower information costs (Baker, Jensen and Murphy, 1988). For example, if a school district announced that the top ten teachers in a school could win bonuses of USD 5 000, with 20 schools in the district participating in the programme, and each school had 40 teachers, the school district could anticipate that the cost of the bonuses would be USD 1 million a year (10 winning teachers x 20 schools x USD 5 000 per winning teacher). However, if the incentive structure is designed so that all teachers could potentially win an award by exceeding a fixed benchmark, the financial exposure is less certain and potentially much greater. The cost of operating the programme could range from USD 0 (no teachers meet the standard) to USD 4 million (40 winning teachers x 20 schools x USD 5 000 per winning teacher).

Stakeholders wanting to integrate a rank-ordered tournament as part of an incentive-pay programme should explore constructing comparable peer groups to measure and evaluate the performance of individual teachers or teams of teachers. Neal (2009) suggests that school systems construct league-specific tournaments; that is, a school's performance is considered relative to the performance of schools that serve similar types of students and families. Without these comparison sets of homogenous schools that are regularly updated, Neal argued, it is impossible to obtain credible estimates of a school's or teacher's contribution to student achievement outcomes. Conceivably, this notion of league-specific tournaments can be applied at the classroom-level while further specifying the matching criteria on which leagues are formed (*i.e.* subject taught, number of classes taught, average class size), thus minimising within-school variation.

Rank-order tournaments can suffer from one well-recognised defect: promotion of competition among teachers or groups of teachers can lead to a breakdown in the collegiate ethos, thereby reducing rather than increasing performance. This can become a serious threat if teachers in the same school are no longer willing to help one another. For example, the grade five mathematics teacher may opt against sharing a lesson plan with the other grade five mathematics teacher because they are competing for the same pot of money. Indeed, allowing only some teachers, rather than all who met a predetermined standard, to receive a bonus was attributed with the collapse of many merit-pay programmes in the 1980s.

Making high-stakes comparisons of teachers or teams of teachers within a school can encourage teachers to sabotage the performance of their peers (Burgess *et al.*, 2001; Neal, 2009). Murnane and Cohen (1985) suggested sabotage can take place in the form of teachers using friendships with students' parents to spread damaging rumours about a colleague's ability as a teacher. Gonring, Teske and Jupp (2007) did not find any "insider-evidence" of unhealthy competition, however, created by Denver's Professional Compensation System for Teachers (ProComp).

Another prime consideration in developing an incentive-pay programme is that district officials, legislators and other individuals responsible for funding bonuses tend to be risk-averse. Elmore, Abelman and Fuhrman (1996) noted that, “Although legislators and governors frequently ‘talk tough,’ they are very susceptible to pressures by school people to back down and soften policies perceived to be restrictive or punitive” (pp. 85-86). In their recent review of the New Public Management reform movement, which is inextricably linked to the resurgence of interest in performance-related pay policies, Perry, Engbers and Yun (2009, p. 14) similarly observed that, “It is ironic that the same politicians who promote performance-related pay also may vote against appropriations to fund it if they perceive fiscal restraint serves larger political ends.”

Incentive-pay reforms are also vulnerable to shifts in political leadership and public opinion. When supporting legislators, superintendents and/or school board members leave office, the political will to continue educator incentive-pay programmes can disappear rapidly (Ballou, 2001; Ballou and Podgursky, 1997; McCollum, 2001). Given the uncertain and potentially substantial costs, incentive-pay programmes based on fixed performance contracts, for example, are likely to be among the first line items on the political chopping block. Thus it is important to construct an incentive-pay programme that appears financially practical to a wide range of educators, as well as to elected and appointed officials.

Fixed performance contract

The second major form of incentive structure is the fixed performance contract, which defines the performance standard teachers, teams or schools must meet to earn an award. The standard can take on a number of forms, including a single threshold, multiple thresholds (e.g. a step function), or a continuous, linear standard after some threshold has been met. In contrast to a rank-order tournament, any teacher who meets the predetermined performance standard benefits regardless of the performance of other teachers.

The primary rationale for adopting a fixed performance contract is pragmatic – fixed performance contracts avoid competition among teachers. Teaching is characterised by widespread compliance to an implicit or explicit code of professional ethics which can be driven by self-interested reputation, internalised morality, peer esteem and a service ideal (Mallick, unknown). Competition may threaten professional norms, particularly those related to collaboration and teamwork, which some studies have found to be key ingredients in highly effective schools (Bryk and Schneider, 2002; Hallinger and Murphy, 1986).

Many teachers’ unions strongly oppose the introduction of competitive pressures into schools or school systems, irrespective of whether pressures come in the form of an incentive-pay policy or from market competition, such as school choice. In an analysis of multiple-waves of data collected in the U.S. Department of Education’s Schools and Staffing Survey, Ballou (2001) found a strong inverse relationship between the use of pay-for-performance programmes and the degree of union influence. Tomlinson (2000) similarly reported that the National Union of Teachers opposed pay-for-performance proposal in England and Wales because:

The Government’s proposals would profoundly damage the professional culture and teamwork that is at the heart of successfully managed schools. The pressures caused by the performance management structure would generate distrust and counter-productive competition between colleagues. Teachers could be demotivated and deprived of the confidence they need to be successful in the classroom. Stress, anxiety, apprehension and self doubt are not conducive to successful teaching (p. 6).²

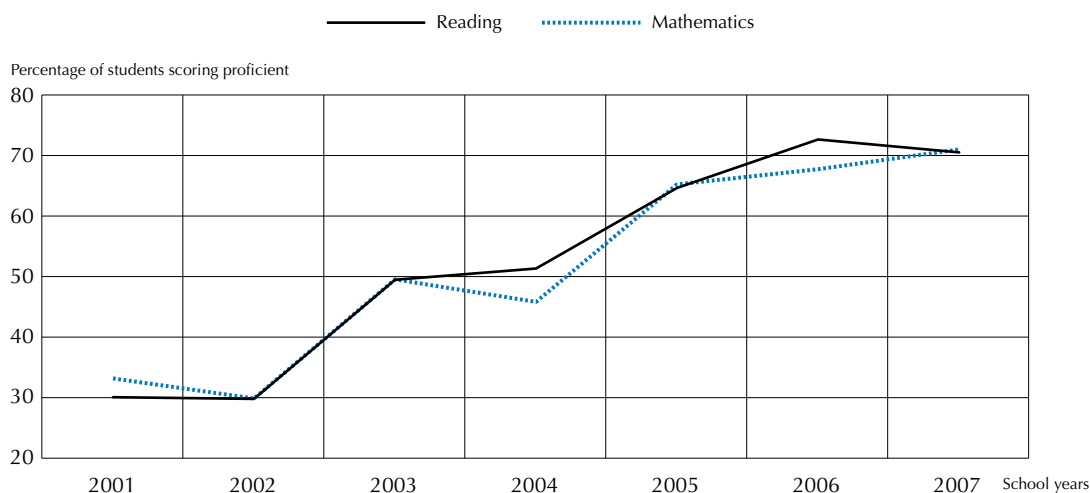
Fixed performance contracts set clear standards which enable teachers to set personal goals in relation to those standards. Goal-setting theory, which refers to the establishment of a purpose in reference to a performance standard, helps direct attention and action, further motivating individuals toward the standard

and, ultimately, leading to better task performance (Locke, 1968; Locke, Latham and Smith, 1990; Locke, Shaw, Saari and Latham, 1981). Recognizing that rank-order tournament incentive schemes are contingent on a “do-your-best” mentality, teachers are more likely to respond to a fixed performance standard because a relative performance standard can be perceived as impossible to attain, or teachers may conclude the evaluation system is biased because teachers do not know what standard they need to achieve in order to earn an award until after the fact.

Conversely, the large financial exposure inherent in most performance contracts has been associated with teachers’ skepticism about incentive-pay policies. In the **United Kingdom’s (U.K.)** Performance Threshold and Upper-Scale Pay System, teachers and principals questioned whether adequate funding existed to implement and sustain the programme given there was no quota on the number of teachers who could earn an award (Adnett, 2003; Cutler and Waine, 2004; Marsden and Belfield, 2006). Skepticism can reduce the likelihood that teachers will increase their effort in pursuit of a bonus award. It can also increase opportunistic behaviour on the part of teachers, and can decrease their level of buy-in (Hamilton, 2005; Kelley, 2002; Lewis and Springer, 2008).

Figure 3.1

Percentage of schools with a minimum of 72% of third grade students scoring proficient in reading and mathematics



Source: OECD.

Fixing in advance the total number of individuals, teams of teachers, or schools that can earn bonuses (*i.e.* a rank-ordered tournament) helps mitigate several complicating factors associated with judging their performance. For example, Figure 3.1 displays the percentage of schools with at least 72% of their third-grade students scoring proficient in reading and mathematics as measured by **Minnesota’s** minimum competency accountability programme. The percentage of schools meeting the 72% standard more than doubled from the 2000-01 to 2006-07 school year, an impressive increase particularly given the relatively short time series.

Now imagine the Minnesota legislature designed and implemented a performance-pay programme that rewards any performance above a minimum threshold, wherein the unit of accountability is the school. Assume the programme also relies on a single performance standard for determining bonus award eligibility – any school with at least 72% of its students scoring proficient receives USD 2 500 per full-time equivalent instructional employee. In the first year of the programme (2001-02 school year), this would have meant that approximately 255 schools, or roughly 30% of eligible schools, met the performance standard, translating to USD 19.2 million in bonus awards (we will assume 30 teachers per school throughout this hypothetical example).

Since the legislature opted for a fixed performance contract incentive structure, the amount of resources necessary to remunerate teachers increases dramatically in a relatively short period of time. In the 2002-03 school year, more than 420 schools met the performance standard, which means bonus awards totaled around USD 32.8 million according to the hypothetical parameters. Two years later, more than 560 schools met the standard, adding an additional USD 10 million to the 2002-03 figure. This would place a huge financial burden on the taxpayer, as well as on the budget of the agency operating the programme.³

Some literature suggests school personnel and bargaining organisations are prone to exploit the fixed performance contract incentive system. Employees “expend time, effort, and ingenuity” to influence others’ decision-making in an attempt to maximise their own benefit (Holmstrom and Milgrom, 1987). Since the design (or reauthorization) of an incentive-pay programme will involve teachers, and/or the bargaining unit protecting the interests of the “median” teacher, these individuals may advocate for a lower performance standard (Indjejikian, Lenk and Nanda, 2000; Merchant and Manzoni, 1989; Murphy, 1999). As Heneman, Milanowski and Kimball (2007) noted:

Over time, teachers exert pressure to lower performance standards, increasing the number of teachers who become eligible for performance pay... while such an eventuality may enhance acceptance of the plan, it also drives up costs, creating long-term funding issues (p.6).

School systems also need to be mindful of a potential ratchet effect when elevating the performance standard (Freixas, Guesnerie and Tirole, 1985; Weitzman, 1980). If a school system increases the rigour of the performance standard that an individual has to meet to earn a bonus award, teachers and other school personnel may perceive the heightened award threshold as a punishment. If so, the recalibration process can potentially reduce productivity, compromise trust and information sharing among school administrators, schools, and teachers, and create sufficient opposition such that the future of the incentive-pay programme is put into question (Lazear, 1995). One strategy to overcome negative reaction to increasing the standard is to increase the size of the bonus or the percentage of employees eligible to earn the bonus at the same time the standards are revised.

Another consideration in determining performance standards is ensuring the standard is not biased in favour of particular schools, teams, or teachers. In their evaluation of **New York City’s** School-Wide Performance Bonus Program (SPBP), Springer and Winters (2009) studied the relationship between the performance target assigned to particular categories of schools and the probability that schools in those categories met at least part of their performance target. Schools were classified according to the previous year’s progress-report data, with Category One schools being the highest performing and Category Five the lowest. According to the authors, the probability of Category Four schools and Category Five schools earning at least part of their performance bonus award was nearly ten times greater than the probability of Category Three and Category Two schools. While the authors’ emphasised these findings came from the first year of the programme, if a similar pattern emerges in subsequent years one might argue the SPBP is rewarding mediocrity given that a percentage of schools in Category Four and Category Five will meet their performance target irrespective of any effort put forward. Moreover, these schools were the lowest performing schools identified by the district’s accountability system.

UNIT OF ACCOUNTABILITY

The next design component of an incentive-pay programme is the unit of accountability. The unit of accountability refers to the entity responsible for a measurable product or service whose performance on that measurable dimension determines bonus eligibility. The unit of accountability can be defined in a myriad of ways, including the individual teacher, a grade-level or departmental team of teachers, all employees within a school, or some combination thereof.

There are distinct advantages and limitations that policy makers need to carefully consider when defining the unit of accountability in incentive-pay programmes. This next section begins with a discussion of the individual as the unit of accountability and then shifts attention to the group, where the group may be defined as a within-school team of teachers or an entire school taken as a single unit. The final section provides a similar review of hybrid models of accountability.

Individual unit of accountability

In an incentive-pay plan that relies on an individual unit of accountability, the performance of the individual teacher determines award eligibility. The individual unit of accountability creates the strongest connection between variation in award size received by teachers and the variation in teacher effectiveness. Since the inability to directly impact one's chances of earning a performance award can lead to decreased effort, the individual unit of accountability is often identified as optimal because the individual assumes sole responsibility for his or her performance (Freeman and Gelber, 2006).

Incentive-pay programmes in the education sector that reward teachers at the individual level of accountability maximise conditions for altering the composition of the teacher labour force. The most effective teachers in the system are more likely to be retained, sending a strong signal to those teachers not receiving a reward to improve or exit. Additionally, prospective teachers most likely to be rewarded under the performance appraisal system are more likely to join the profession. Thus, as Podgursky and Springer (2007) argue, teacher turnover would become part of a virtuous cycle of quality improvement, rather than a problem to be minimised.

The theoretical underpinnings of an incentive-pay programme that rewards individual teacher performance is critically important in light of inefficiencies in current teacher hiring practices and labour market selection (Lazear, 2000). In terms of hiring practices, principals and building administrators must use noisy signals of "true" teacher effectiveness, such as years of experience, highest degree held, or past-employer recommendations. Informational deficiencies in the hiring process are overcome in most professions by employee performance assessments and the close coupling of pay increases and promotion decisions with actual productivity. Without such structures in education, combined with an often insurmountable tenure system that makes contract non-renewal difficult, pay for performance becomes all the more relevant.

In terms of labour market selection, an incentive-pay programme tends to attract and retain individuals who are particularly good at the activity to which incentives are attached, while repelling those who are not. While incentives can raise the productivity of the typical worker, an incentive system can also raise the overall quality of the workforce simply through differential recruitment and retention of more effective workers (Podgursky and Springer, 2007). Lazear's (2000) widely-cited case study of Safelite Glass Corporation, for example, reported that transitioning the company's compensation system from hourly wages to piece rates was associated with a 44% increase in worker productivity, half of which resulted from the sorting of more able workers. In a case study of a **British Columbia** tree-planting firm, Paarsch and Shearer (2000) similarly found increases in productivity attributable to both increased effort and sorting. Of course, the production process and output for teachers and schools is much more complex.

A recent and provocative study of teacher turnover found evidence that the migration of high ability women out of teaching between 1960 and the present primarily resulted from the “push” of teacher pay compression, which took away relatively higher earnings opportunities for teachers, as opposed to the “pull” of more lucrative non-teaching opportunities (Hoxby and Leigh, 2004). Although remunerative opportunities outside of teaching for teachers of high and low ability grew over this period, Hoxby and Leigh (2004) argued it was pay compression within the education system that accelerated the exit of higher ability teachers. Similarly, Corcoran, Evans and Schwab (2004) found the likelihood that a female in the top of her upper-secondary school class will enter teaching decreased dramatically between 1957 and 1992.⁴

However, the literature on optimal incentive structures questions whether the individual as the unit of accountability is the optimal incentive design when outputs cannot easily be attributed to an individual teacher. If work tasks of two or more individuals influence the measured and rewarded activity in an incentive-pay programme, relying on the individual as the unit of accountability may cause some teachers to take exception to high performers because they may affect the standard upon which all other teachers are judged (Baron and Kreps, 1999; Bowles and Gintis, 2002; Deutsch, 1985). Moreover, as discussed previously, rewarding teachers for their individual performance is also contrary to the highly collaborative nature of teaching and may even reduce the incentive to cooperate with others (Mallick, date unknown; Milgrom and Roberts, 1990; Murname and Cohen, 1986).

There are numerous other interdependencies that complicate efforts to isolate the contribution of an individual teacher to student outcomes. If the individual is the unit of accountability, how should the evaluation system take into account the cumulative effects of learning from one school year to the next? Are there practical strategies for taking into consideration depreciation or appreciation in learning over time, and whether the instructional strategies of the prior year's teacher influence learning gain or loss (*i.e.* teaching to the test versus teaching higher order skills)? Identifying the individual contribution of a teacher to student learning is further complicated, for example, when students enroll in a reading and an English language arts course, each of which is taught by a different instructor.⁵ Individual incentive-pay programmes can also create resentment among teachers, as teachers not receiving an award may believe they contributed to the success of students linked with a teacher that receives a bonus award.

Factors influencing student learning also extend outside the school building walls. For example, in their highly influential work on consumer preferences and school choice, Schneider, Teske and Marshall (2000) noted that interactions among students and parents make parents co-producers of educational outcomes; parents undoubtedly shape the educational goals of their children. Social science research has also identified a number of neighbourhood factors that can influence student outcomes, including social disorganisation, community resources, environmental deterioration (*e.g.* abandoned buildings), crime and delinquency (*e.g.* drug dealing and violent crime) (Ginther, Haveman and Wolfe, 2000; Leventhal and Brooks-Gunn, 2004; Wilson, 1987). These and other complicating factors make it difficult to isolate an individual teacher's contribution to student outcomes.

Group unit of accountability

The group unit of accountability refers to award eligibility as the product of aggregated performance among members of a group, where the size of a group can range from as few as two employees to all employees within a firm. Sometimes referred to as profit sharing, gain sharing, team incentives, goal sharing, achievement sharing, win sharing, and results sharing (Gaynor and Pauly, 1990; Holmstrom, 1982), group incentive structures have experienced tremendous growth in popularity. In the **U.S.**, the number of group incentive-pay programmes operating in firms increased from just over 2 000 in 1945 to nearly 500 000 in 1991 (Nalbatian and Schotter, 1997).

It is important to stress that the group unit of accountability is not limited to the school-level, a common mistake encountered among incentive-pay programmes throughout the education sector. Group incentive systems can take on a number of different forms, including monitoring and rewarding performance of grade-level teams of teachers, disciplinary or inter-disciplinary departments of teachers, or any other recognizable unit within the school.

Organisational theory suggests group incentives can promote social cohesion, feelings of fairness, and productivity norms (Lazear, 1998; Pfeffer, 1995; Rosen, 1986). Improved social cohesion among workers can foster knowledge transfer and mutual learning that result in increased productivity in the long run (Che and Yoo, 2001). For example, in the case study of garment plants reported in Berg *et al.* (1996) and Hamilton, Nickerson and Owan (2003), the formation of teams with workers of varying abilities facilitated interactions among high- and low-ability workers so that more able workers taught less effective workers how to better execute tasks and become more productive.

Although productivity gains from knowledge transfer and mutual learning may yield results similar to an individual incentive programme, some empirical research describes a much more nuanced story when firms have adopted team-based incentive programmes. Weiss (1987) and Hansen (1997) found that while the performance of lower-ability workers improved, the productivity among the highest producers decreased. Thus the group incentive-pay programme may not have the same compositional effect on the workforce and, as a consequence, the expected benefit of the group unit of accountability may be all for not.

The free-rider problem and shirking are other concerns designers of incentive-pay programmes need to recognise when considering the group as the unit of accountability. If the school is the unit of accountability, or the size of the within-school team is greater than six to eight teachers, specific individuals on a team may be less likely to shoulder their fair share of the workload. They know the capabilities of teammates can make up for their subpar performance. Because an individual who exerted minimal effort will still receive a bonus award if the performance of the other members of his team offset his mediocre effort, group incentive systems inevitably result in the inefficient allocation of some resources.

Kandel and Lazear (1992) and others have argued that as long as the size of a within-organisation team is not too large, the free-rider problem can be solved through peer pressure. For instance, peer monitoring and the enforcement of social penalties in the form of shame, guilt, empathy, and mutual monitoring can lead to individual team members being accountable for their performance to other members. If a teacher has both monetary and social incentives to not shirk, Kandel and Lazear (1992) contend the motivational forces that would have been “choked off” by the free-rider problem are recovered.

However, even though the theoretical literature has argued that the free-rider problem can be overcome in certain situations, a large-scale random assignment evaluation revealed individual teacher incentives elicited greater gains and productivity as compared to group incentive programmes. Known as the Andhra Pradesh Randomized Evaluation Study (AP RESt), the impact of two output-based incentive systems in **India** (an individual teacher incentive programme and a group-level teacher incentive programme) and two input-based resource interventions (one provided an extra-paraprofessional teacher and another provided block grants) were evaluated. Muralidharan and Sundararaman (2008) found that students enrolled in a class instructed by a teacher selected for the group incentive intervention outperformed students in control condition classes that received only their usual allocations of resources, on both the mathematics and language exams (0.28 and 0.16 standard deviations, respectively).⁶ Students enrolled in schools assigned to the individual incentive condition outperformed students in both the group incentive condition and the control condition after the second year of implementation. In terms of cost effectiveness, moreover, the authors reported both of the incentive-pay interventions were preferable as compared to the input-based resource interventions and the control condition.

An incentive-pay programme relying on the group unit of accountability may also fail to recognise some of the most effective teachers in a school or school system if those individuals are assigned to a team with lower-ability teachers. High-ability teachers who are not rewarded may leave the school system or profession as a consequence of being discouraged with a performance management system that does not recognise their individual performance. Rivkin, Hanushek and Kain (2005) further deduced school-based incentive systems may also be a highly-inefficient approach for rewarding performance given the presence of significant within-school variation in teacher effectiveness. Recognizing that two of the key goals of incentive-pay programmes are to improve the composition of the labour force and organisational productivity, it is critically important that stakeholders designing an incentive system closely monitor and evaluate the programme for any signs of perverse sorting.

A final consideration when adopting a group incentive plan is procedural. How should the composition of within-school teams be formed? The purest approach to assigning teachers to teams is a simple random lottery wherein each teacher has an equal chance of being assigned to a given group (Lazear and Gibbs, 2009). Assuming the number of individuals in the lottery is large enough in relation to the number of teams, randomly assigning teachers to teams means work ability will be balanced across the teams. Another strategy is for a school principal or administrator to assign individual teachers to teams. Most firms rely on supervisors to select the composition of teams, which makes the most sense if the pool of candidates being assigned to teams is new to the company (Lazear and Gibbs, 2009). Lazear and Gibbs (2009) further noted an alternate draw is typically identified as the optimal strategy if schools want to make teachers responsible for the make-up of the within-school teams.⁷

Hybrid models of accountability

Hybrid models offer a third type of accountability system. A hybrid model joins elements of both independent and interdependent work, meaning the incentive structure considers the unique contribution of an individual teacher while also supporting teamwork and collaboration among teachers. Hybrid models are more commonly found in the **U.S.** and, as discussed later in this chapter, may be related to the cultural context in which an incentive system operates.

Initiated in 1999 by the Milken Family Foundation, the Teacher Advancement Program (TAP) is a well-known incentive-pay programme drawing on a hybrid model of accountability. Bonus calculations incorporate standardised tests scores for both individual teacher performance and the performance of the entire school. While each school can ultimately decide the exact percentage to attribute to each unit of accountability, TAP guidelines recommend a breakdown of 20% school level and 30% individual level.⁸

Establishing an appropriate balance among design components can pose a number of challenges. If too little weight is placed on some aspect of the schooling process perceived by teachers and other stakeholders as important, then the influence of the incentive runs the risk of being diluted to the point that the pay programme becomes irrelevant. Such a scenario can also send a strong signal that a particular aspect of the schooling process is not valued, which may lead to discouragement and a breakdown in trust between teachers and administrators. For example, in their study of teacher perceptions to **Florida's** Better Educated Students and Teachers (BEST) Teaching Salary Career Ladder Program, Mullen and Slagle (2007) reported that teachers perceived as highly-flawed the way in which the BEST program conceptualised teaching.

A hybrid model of accountability necessitates that the technical requirements of both group and individual units of accountability be in place. Systems that currently have the capacity for linking records between individual teachers and students will not have significant further requirements if also adopting aggregate units of accountability. Conversely, those school systems starting with a group as the unit of accountability may not have

the data linkage necessary for measuring performance at the individual teacher level, or the technical capacity to manage an incentive-pay programme with a more disaggregated accountability unit. Current capacity and future costs of using a hybrid model could be significant and should be investigated beforehand.⁹

A lack of goal clarity is more likely to affect the motivational effect of an incentive-pay programme relying on a hybrid unit accountability. In a summary report of performance-related pay policies prepared for the U.K.'s National Union of Teachers, Richardson (1999, p. 20) noted, "...a multiplicity of goals is likely to cause problems, because it reduces goal clarity. Being under pressure to meet many goals, especially where some of them are ambiguous, makes it difficult for employees to focus their efforts properly." Recognising the inverse relationship between the number of design components incorporated into incentive-pay programmes and teacher goal clarity, it is important for hybrid models of accountability to clearly specify how the programme works.

Some challenges regarding goal clarity can be mitigated if the programme design team establishes a multi-prong strategy to enhance teachers' and school personnel's understanding of the programme. Stakeholders can train school administrators on the ins-and-outs of the programmes to help facilitate implementation as well as buy-in among teachers. In fact, in their evaluation of school-based performance award programmes in **Kentucky** and **North Carolina**, Kelley (1999) reported programme buy-in is largely conditional on the confidence teachers have about the system functioning as intended and on the responsiveness of administrators to questions from teachers about the programme.¹⁰ Chapter 7 presents a detailed review of implementation and technical assistance issues.

PERFORMANCE MEASURES

Performance measures refer to the evaluation criteria for gauging employee performance. Despite many education systems having a long history of evaluating teachers and schools, there remains considerable variation in stakeholders' perceptions of what should be evaluated, how appraisal criteria should be linked to rewards, and which measures and instruments can reliably and accurately reflect performance.

Table 3.1 displays the primary performance measures implemented in a handful of high-profile incentive-pay programmes. Slightly more than half of the programmes include an input- or a process-oriented performance measure, which can include professional development, job enlargement, student attendance, and teacher attendance.¹¹ Educational outputs cover student achievement levels and gains at the student, teacher, and school-level as well as accumulation of credit hours and student retention rates.

Several programmes also incorporate recruitment and retention incentives or incentives for teaching in a hard-to-staff school or subject. Hard-to-staff school incentives are most often offered for teaching in high-needs schools or districts, typically either high-poverty, low-performing, or geographically remote schools. Like hard-to-staff subject incentives, these incentives are designed specifically to address market factors.¹² As presented in Table 3.1, we classify these market-based compensation reforms under educational inputs and processes.

Educational inputs

Educational inputs refer to measures that reward teachers for activities believed to improve instructional quality. Examples of the types of activities rewarded under an input-focused incentive system include completing an advanced degree, serving as a mentor or master teacher, taking on a leadership role or other special duties, participating in professional development activities, and teacher attendance. We first review more prominent types of input-focused programmes and then discuss perspectives regarding subjective measures of performance.

Table 3.1

Characteristics of incentive pay programmes in the United States and other countries

		Teacher Performance Measures	Size of bonus (USD, normal)		
	Period	Unit of Accountability	Minimum	Maximum	% Monthly Salary
United States					
School Incentive Program (Dallas, Texas)	1992 - 1995	School	\$450	\$1 000	≈ 10% - 22%
Merit Pay Program (State of Michigan)	1996 - 1997	Teacher	\$1 000	≈ \$5,000	≈ 12.5% - 62.5%
Teacher Advancement Program (United States)	1999 - present	Hybrid (Teacher and School)	\$2 500	\$12 000	≈ 45% - 216%
Achievement Challenge Pilot Project (Little Rock, Arkansas)	2005 - 2007	Teacher	\$350	\$7 600	≈ 8% - 174%
Governor Educator Excellence Grant Program (State of Texas)	2005 - 2008	Varies (Teacher, Team, and/or School)	\$75	\$15 000	≈ 1.3% - 270%
Professional Compensation System for Teachers (Denver, Colorado)	2005 - present	Hybrid (Teacher and School)	\$400	\$2 500	≈ 7.2% - 45%
ASPIRE Program (Houston, Texas)	2005 - present	Hybrid (Teacher and School)	\$250	\$8 600	≈ 4.5% - 154.8%
POINT Experiment (Nashville, Tennessee)	2006 - 2009	Individual	\$5 000	\$15 000	≈ 90% - 270%
Texas Educator Excellence Grant Program (State of Texas)	2007 - present	Varies (Teacher, Team, and/or School)	\$20	\$20 462	≈ 0.4% - 365%
Team Incentive Project (Round Rock, Texas)	2008 - present	Team	≈ \$6 000		≈ 108%
School-Wide Performance Bonus Program (New York City, New York)	2008 - present	School	\$1 400	\$3 600	≈ 25.2% - 64.8%
International					
School Performance Program (Israel)	1996 - 1997	School	\$1 000	\$2 500	≈ 30% - 75%
International Christelijk Steunenfonds Incentive Program (Buso and Teso, Kenya)	1998 - 1999	School	\$26	\$51	≈ 21% - 43%
Salario al Merito (Bolivia)	1998 - 1999	Teacher	\$300	\$333	≈ 66%
Teacher-Incentive Experiment (Israel)	2001	Teacher	\$1 750	\$15 000	≈ 10% - 40%
Carrera Magisterial (Mexico)	1993 - present	Teacher	...	...	≈ 27% - 224%
Sistema Nacional de Evaluacion de Desempeno de los Establecimientos Educativos (Chile)	1996 - present	School	\$439	\$439	≈ 4.7% - 7.2%
Performance Threshold and Upper Pay Scale System (England)	1999 - present	Individual	\$2 000		≈ 9%
Incentivo Colectivo a Escuelas (Bolivia)	2001 - present	School	\$281	\$281	≈ 5% - 19%
Randomized Evaluation Project (Andhra Pradesh, India)	2006 - present	Teacher or School	\$2.25	\$450	≈ 33% - 50%
Performance Pay Program (Portugal)	2007 - present	Teacher	...	...	...

Sources: Atkinson *et al.* (2009); Clotfelter and Ladd (1996); Eberts, Hollenbeck, and Stone (2002); Glewwe, Holla, and Kremer (2009); Glewwe, Ilias, and Kremer (2008); Ladd (1999); Lavy (2002, 2004); McEwan and Santibanez (2005); Mizala and Romaguera (2004); Muralidharan and Sundararaman (2008); Podgursky and Springer (2007); Rodriguez (2002); Santibanez *et al.* (2007); Schacter and Thum (2004); Springer, Ballou, and Peng (2008); Springer *et al.* (2008, 2009); Springer and Winters (2009); Taylor, Springer, and Ehlert (2009); Vegas and Umansky (2005); Winters, Ritter, Greene, and Marsh (2009).

Knowledge-based pay and skill-based pay programmes are the most prominent type of incentive-pay policies that focus on rewarding educational inputs. Developed by a team of researchers with the Consortium for Policy Research in Education (CPRE) at the University of Wisconsin, these incentive plans induce teachers to obtain new skills related to classroom management and knowledge of curriculum development and curriculum content in core subjects. Several high-profile initiatives were launched in the U.S. in the late-1990s, including programmes in Cincinnati (**Ohio**), Douglas County (**Colorado**), Vaughn (**California**), and Washoe County (**Nevada**) (Milanowski, 2003).¹³

The acquisition of new knowledge and skills is believed to improve instructional effectiveness of a teacher and consequently student outcomes (Odden and Kelley, 1997). Input-based incentive systems are desirable because they are unlikely to cause “teaching to the test”, or to increase the chance of teachers focusing a disproportionate amount of effort toward those students already near some performance standard in hopes of

maximizing chances of earning a performance award. Firestone (1994, p. 550) further argued that these input-based pay programmes are “compatible with an outcomes focus because the resulting flexibility allows workers to do what is necessary to achieve those outcomes”.

Another appealing feature of input-based systems is rooted in the belief that the opportunity to pursue additional course work or to take on more responsibilities increases teacher interest and motivation (Odden and Kelley, 1997). This feature is very attractive if in fact it does find a way to balance the potential threat of external rewards reducing intrinsic motivation, particularly considering teachers are drawn to the profession by intrinsic factors (Kohn, 1999; Lortie, 1979; Rosenholtz, 1984). As noted by Johnson (2004, p. 46), “In itself, higher pay is unlikely to retain teachers – particularly the most able among them – if they find that they cannot attain the intrinsic rewards for which they initially entered teaching.”

Social psychological literature on public service motivation further argues that individuals seek to contribute to the public good in order to satisfy personal needs, not necessarily in response to incentives offered by organisations for performance or commitment (Courty, Heinrich, and Marschke *et al.*, 2005; Perry and Porter 1982; Rainey 1982; Wise 2004). Wise (2004), in particular, suggested that public-service motives had the potential to empower public servants to “overcome self-serving interests, moral inertia and risk avoidance” (as cited in Courty, Heinrich and Marschke, 2005), an effect to which Crewson (1997) attributed greater organisational commitment and lower employee turnover.

Despite proponents’ contention that the activities rewarded by knowledge- and skill-based pay programmes are associated with student achievement, a growing number of empirical studies indicate teacher effectiveness is, at best, marginally correlated with observable teacher characteristics such as teaching certificate held, level of education, licensing exam scores, and teaching experience beyond a few years (Goldhaber, Brewer and Anderson, 1999). For example, Goldhaber *et al.* (1999) found that observable teacher characteristics explain only about 3% of the differences in student achievement that are attributable to the teacher. In a related study, Goldhaber (2002, p. 50) argued that, “The evidence shows that good teachers make a clear difference in student achievement. The problem is that we really don’t know what makes a good teacher.”

Many stakeholders contend payment for output is likely to be a more efficient and productive way to remunerate teachers given variation in teacher effectiveness cannot be explained by activities rewarded in knowledge- and skill-based pay programmes. In the **U.S.**, for example, 97% of school districts remunerate teachers on criteria largely unrelated to schooling outcomes, while compensation payments for instructional personnel account for approximately 55% of K-12 current expenditures and 90% of instructional expenditures (Springer, 2009).

Input measures have been found to be related to student performance, notably observational evaluations that can be conducted by peers, supervisors, or external evaluators. Studies have reported that a standards-based evaluation can predict student achievement and can therefore provide performance evaluation scores as the basis for a performance-based incentive-pay programme or other personnel decisions (Kimball *et al.*, 2004; Milanowski, 2004). Additionally, Odden, Borman, and Fermanich (2004) reported observational evaluations not only provided a more “comprehensive model” of teacher effects on student achievement, but also have the potential to inform teacher practice by identifying specific areas of improvement.

However, the theoretical and empirical literatures on incentive-pay programmes in education identify a large number of concerns associated with subjective performance measures. Prendergast (1999) argued systematic errors, if frequently observed when employers rank workers, threaten the legitimacy of subjective assessment, often leading to leniency bias, centrality bias, and rent-seeking. Leniency bias is rooted in supervisors’ reluctance to rank the poorest performing workers, thereby distorting the overall distribution of rankings (Prendergast, 1999). Centrality bias, defined as the tendency for rankings to truncate the actual range of performance, introduces further

distortion into this distribution; supervisors may deliberately distort organisational standards and outcomes in order to redistribute rewards on the basis of personal preference (Adnett, 2003; Cutler and Waine, 2004; Marsden and Belfield, 2006). Finally, and as a function of these preferences, subjective assessment invites rent-seeking; workers focus effort away from critical tasks and toward influencing supervisors (Prendergast and Topel, 1993).

Further, raters' knowledge of prior performance appears to affect information processing by framing or anchoring current judgments (Huber, Neale and Northcraft, 1987). The order in which the observer sees good and poor performance therefore affects performance rating, with biased judgments about inconsistent extreme performance moving toward the prior general impression. Teams of raters may also attach greater weights to select elements than is appropriate (Schmitt, Noe and Gottschalk, 1986) or they may distort ratings for political reasons (Longenecker *et al.*, 1987).

The complex nature of teaching makes it difficult to identify specific criteria that should be present during an observational evaluation. While some have developed observational rubrics from supposed "research-identified" effective teaching behaviours (Danielson, 1996; Schacter and Thum, 2004), extensive training for raters is necessary. Observations are also very expensive as they require multiple observations on an individual teacher over the course of a school year.

In a recent study focused entirely on the predictive validity of supervisor evaluations, Jacob and Lefgren (2005) assessed the relationship between teacher performance ratings, as identified on a detailed principal evaluation, and teacher effects, as measured by student achievement gains. In estimating teacher effectiveness measures for 202 teachers of mathematics and reading in grades two through six, Jacob and Lefgren found a statistically significant and positive relationship between value-added measures of teacher productivity and principals' evaluations of teacher performance.

Another interesting dimension of this study was an "out of sample" prediction of 2003 student achievement scores based on principal ratings and teacher value-added estimates from 1998 through 2002. Students had higher average scores in mathematics and science if they had teachers with not only higher measured teacher effectiveness in prior years but also higher principal ratings. Jacob and Lefgren (2005) demonstrated further that the principal evaluation remained a statistically significant predictor of current student achievement even when teacher value-added (in the previous year) was included in the model. This finding suggests that principal evaluations provide an important independent source of information on teacher productivity.

Subjective evaluations also have the potential to serve as important formative assessments for teachers, although some have deemed subjective performance appraisals too fallible for high-stakes personnel decisions. If the teaching standards describe effective instructional strategies, as argued by Odden, Borman, and Fermanich (2004), then when these evaluative tools are implemented in the classroom, student achievement should increase. However, as summarised by Weisberg *et al.* (2009), teacher evaluation systems have failed to produce credible and accurate information about instructional practices of individual teachers.

Even though studies have indicated that principals are relatively adept at identifying above- and below-average teachers, it is unclear whether principal evaluations would persist in a high-stakes performance-pay programme. The fact that a principal identifies a teacher as "inadequate" on an anonymous survey does not mean necessarily that she will do so in a high-stakes environment. Take, for example, two interventions designed to increase teacher attendance rates in parts of **India** and **Kenya**. The first programme offered rather large bonuses if eligible teachers regularly came to work, as monitored by the school principal. An evaluation of the programme did not detect any discernible impacts on a number of teacher and student outcomes (Kremer and Chen, 2001). Analysis further revealed a majority of school principals distributed bonus awards to teachers, irrespective whether that teacher actually met the performance standards defined by the intervention.

In the second teacher attendance intervention, absence patterns were tracked using tamper-proof cameras with time and date functions, and treatment group teachers received a salary commensurate with the total number of days they attended work each month (Duflo, Hanna and Ryan, 2005). Treatment group teachers also received a modest bonus for each day they attended in excess of 21 days in a single month or an equally-sized pay deduction for each of the 21 days they did not show-up for work. Duflo *et al.* (2005) reported an average reduction in teacher absence rate of 19% and an increase in student performance on two standardised assessments.

In their comprehensive review of incentive-pay policies in developing countries, Glewwe, Holla and Kremer (2009) indicated the impact of the intervention evaluated by Duflo and colleagues may have been different from those reported in the Kremer and Chen (2001) study because the human element was removed from the monitoring process. The literature from outside the education sector further lends support in that managers often feel uneasy about monitoring colleagues in contexts with strong interdependencies, which suggests high-stakes performance evaluations conducted by personnel within the same school can possibly be unreliable and subject to perverse behaviours (Murphy and Margulies, 2004).

Educational outputs

Incentive-pay programmes have become increasingly focused on rewarding teachers or groups of teachers using outcome-oriented performance measures. Test scores measured on standardised assessments have gained popularity in large part because assessment instruments are routinely administered by most education systems, and standardised assessments represent an objective measure of performance. In an examination of professional compensation structures that are generally sympathetic to reform, for example, Bok (1993), as quoted in Ballou (2001), noted incentive-pay programmes could not work in the education sector because of the “inherently subjective and contentious process” of evaluating teachers.

Despite the considerable growth in output-focused measures of school, team, and teacher performance, these performance measures have received much criticism (Rothstein, 2009). Unlike sales or the billable hours of a doctor or lawyer, a teacher’s output is not measured readily in a reliable, valid and fair manner (Podgursky and Springer, 2007). Teaching involves a complex array of tasks that vary by the number of students taught, the grade levels of the students, the courses being taught, and the relationship of those courses to the tests (McCaffrey, Han and Lockwood, 2009). These complexities result in uncertainties about how teachers should be evaluated, and the way in which these uncertainties are resolved will ultimately affect the signals sent to teachers and other personnel (McCaffrey, Han and Lockwood, 2009).

Monitoring and rewarding individual teachers or groups of teachers according to student test scores captures only a fraction of the contribution of school personnel as well as the overall mission of a school. In the **U.S.**, for example, approximately two-thirds of teachers do not instruct in a tested grade or subject. Moreover, the typical student engages in a large number of academic and extra-curricular activities beyond subjects tested by most education systems. A sole reliance on objective measures, such as test scores, thus runs the risk of teachers focusing excessively on the metred, rewarded activity to the detriment of important dimensions of schooling (*i.e.* focusing on test-taking skills as opposed to creative or critical thinking skills). Additionally, cultivating and sustaining teacher support and buy-in can be problematic if a teacher believes the incentive-pay programme does not capture a significant aspect of their job or the specific performance objective is unclear (Murnane and Cohen, 1986). As a consequence, this lack of buy-in will weaken behavioural responses to the incentive programme (Chamberlin *et al.*, 2002).

Traditionally high- and low-performing students could be worse off in an incentive-pay programme excessively focused on output measures. Teachers may focus a disproportionate amount of their effort on the students who

are most likely to maximise the teacher's chances of earning an award. In the 1970s, for example, performance contracting experiments in the **U.S.** were found to be associated with teachers focusing excessively on students in the middle of the distribution to the detriment of their higher- and lower-performing peers (Gramlich and Koshel, 1975; Hannaway, 1996; Rapple, 1990). More recently, studies have examined a similar response to the 2002 No Child Left Behind (NCLB) Act (Ballou and Springer, 2009; Booher-Jennings, 2005; Neal and Schanzenbach, forthcoming; Reback, 2008; Springer, 2008).

Numerous other accounts from outside the **U.S.** have documented similar patterns of teacher behaviour in response to incentive-pay programmes. For example, in 1861, the **U.K.'s** New Castle Commission reformed the teacher compensation system as a means "to institute a searching examination... of every child in every school... and to make the prospects and position of the teacher dependent, to a considerable extent, on the results of the examination," (Bourne and MacArthur, 1970, p. 20, as cited in Chamberlin *et al.*, 2002). However, the incentive-pay programme ground to a halt some 30 years later, burdened by a legacy in which "teachers taught to the test, were confined to a narrow, boring curriculum, attempted to arrange the school intake, cheated, ignored bright children and drilled and beat the slower ones until they could satisfy the all-powerful inspectors," (Chamberlin *et al.*, 2002, p.32).

While incentive-pay programmes in the **U.K.** were linked with teachers focusing excessively on the metred, rewarded activity some 150 years ago, a much longer and more storied history of dysfunctional behavioural responses have also been documented. Implemented in 606 and lasting until about 1905, Imperial **China's** Keju examination system was largely used to identify the future employment and social status for each generation of students that completed formal schooling. However, according to Suen and Yu (2006), the Keju system suffered from construct-irrelevance and construct-underrepresentation which included rote memorization, an excessive focus on test-taking skills, and cheating in the form of nepotism, bribery, communicating with outside confederates during the examination, and students and their families hiring substitutes to complete the test. Although the Keju system is certainly an extreme example of potential negative responses to a high-stakes testing policy, the findings are relevant to the current discussion in that many of these negative consequences could have been reduced by broadening the performance measures.

A more recent example of manipulation comes from the International Child Support Incentive Program (ICSIP), a group incentive intervention that randomly assigned 100 schools in rural **Kenya** to either a treatment or control condition. A comprehensive evaluation completed by Glewwe, Ilias and Kremer (2008) revealed that the practices of teachers in the treatment group stayed relatively the same, apart from offering extra test preparation sessions outside of class time. The positive increase in the test scores of students enrolled in treatment group classrooms did not translate to non-tested subjects and were not sustained in subsequent years. An analysis of item-level test data further indicated that treatment condition students were significantly less likely to leave a test question blank.¹⁴

Additional problems are likely to surface when student and teacher interests are not well-aligned. While many standardised tests, such as upper-secondary school exit exams, are associated with high-stakes for students, there are many situations when exam performance has no bearing on the student (Bishop, 1996). In an incentive programme developed and implemented by **Israel's** Ministry of Education, teachers were measured by their students' performance on exams required for a certificate of matriculation, which has large financial implications for the students (Lavy, 2004). Even though teachers eligible for incentives led students to larger gains in achievement, the difference may not have generalised to situations where tests are high-stakes for students in alternating years: if a sixth grade student is required to pass in order to advance to the next grade but seventh grade scores do not carry the same weight, then the seventh grade teacher gains may exhibit bias due to decreased performance stemming from lower personal motivation.

Evaluating teacher performance solely on standardised test scores of students presents a non-trivial risk for teachers. In their seminal study of school accountability measures, Kane and Staiger (2001) noted:

The imprecision of test score measures arises from two sources. The first is sampling variation, which is a particularly striking problem in elementary [(i.e. primary)] schools. With the average elementary school containing only sixty-eight students per grade level, the amount of variation stemming from the idiosyncrasies of the particular sample of students being tested is often large relative to the total amount of variation observed between schools. The second arises from one-time factors that are not sensitive to the size of the sample; for example, a dog barking in the playground on the day of the test, a severe flu season, a disruptive student in a class, or favorable chemistry between a group of students and their teacher. Both small samples and other one-time factors can add considerable volatility to test score measures.

Volatility and other forms of error in the performance measure means teacher earnings are less certain because of factors beyond a teacher's control. Recognizing most teacher salaries are more predictable than in other professions (Liang, 1999), as the number of factors outside the control of teachers' increases, not only will the strength of the incentive be weakened but it could also lead teachers to demand higher levels of compensation. Thus, as summarised in a review of the economic complexities of incentive reforms, Asch (2005) noted that, "The optimum mix of the base and the risk component depends on the degree of risk aversion, earning variability, and the worker's cost of effort," (p. 316).¹⁵

Practitioners and policy makers interested in designing and implementing an incentive-pay programme need to know that even though VAM have gained a reputation for providing fair comparisons of teachers, this does not necessarily mean they can validly support such systems.¹⁶ McCaffrey, Han and Lockwood (2009) noted that VAM research to date has focused on the statistical properties of the measures from the perspective of methodological research rather than from the perspective of an algorithm that translates raw administrative data on students and teachers into information for managing human capital. Furthermore, to illustrate the many complexities and decisions that must be made when designing an incentive-pay programme, McCaffrey and colleagues also developed parameters for a series of hypothetical incentive-pay models and then test the performance of several Value Added modeling (VAM) strategies within this context.¹⁷

As VAM becomes increasingly popular in the education policy arena, it is critical that school systems understand the problems and limitations of their management information systems insofar as they relate to the creation of VAM. Common problems include mismatch of teacher and student data, multiple teachers with responsibility for the same student, unclear team or group composition, a lack of data and capacity to measure teacher value-added, and bonuses awarded to the wrong persons or denied to the right persons. These problems will be an embarrassment to the district and may jeopardise the credibility of the programme. Stakeholders should also bear in mind that even though the use of performance incentives is on the rise, the private sector has moved away from the sole reliance on quantitative measures of individual or group performance and begun to design incentive-pay programmes that incorporate both input- and output-based measures of employee performance (Rothstein, 2009).

Finally, poorly designed incentive-pay programmes can create perverse incentives whereby teachers move away from low-performing schools in order to maximise their chances of earning additional pay (Ladd, 1999; Neal, 2009), potentially exaggerating the already unequal distribution of teachers across historically advantaged and disadvantaged contexts. In a study of **North Carolina's** school accountability system, for example, Clotfelter, Ladd and Vigdor (2004) found the recruitment and retention of high-quality teachers even harder in low-performing schools.

Multiple measures – linking educational inputs and outputs

Schooling is often described as a multi-product, multi-purpose enterprise; the performance of a worker and the mission of a school comprise multiple dimensions. Both empirical and theoretical literature on incentive pay clearly indicate that if there is disconnect between an organisation's mission and the activity to which incentives are attached, employees may shift work toward the metred, rewarded activity, and away from other important activities (Dixit, 2002; Hannaway, 1992; Holmstrom and Milgrom, 1991). Many advocates argue that the use of multiple measures in an incentive-pay programme means the performance of schools, teams of teachers, or individuals can be monitored more effectively at various points throughout the school year, all of which can reduce the chances for system-gaming.

A large number of studies have documented how narrowly-focused performance measures can create greater opportunity in the long run for employee cheating and opportunistic behaviour. In an incentive-pay programme that rewards teacher performance solely on student test scores, dysfunctional behaviours can take the form of changing student responses on answer sheets, providing students with correct answers during the examination, or obtaining copies of test material prior to administration in order to prepare students (Jacob and Levitt, 2003).¹⁸

Although linking educational inputs and outputs can reduce the likelihood of system-gaming, there can be significant costs associated with an incentive-pay programme that evaluates teachers across multiple performance measures. Guthrie and Prince (2009) reported that both Philadelphia (**Pennsylvania**) and Steamboat Springs (**Colorado**) cancelled incentive-pay programmes after learning of the cost of implementing the teacher performance appraisal system. Similarly, several school districts in **Florida** were interested in participating in the state pay-for-performance programme, but abandoned the programme after learning the cost of administering the battery of student assessments required to participate (Guthrie and Prince, 2009).

Each of the stakeholders involved in the design and implementation of an incentive-pay programme may advocate for integrating unique performance measures that satisfy their personal interest or that of their constituents. For example, parents expect schools to “foster creativity, curiosity, self-esteem, tolerance, good citizenship, athletic performance, and a host of other objectives” (Hannaway, 1996, p. 103), which may correspond imperfectly to the goal of schooling as defined by a policy maker. Even though adopting multiple measures can go a long way toward capturing the complexity of the schooling process, stakeholders need to be cognisant of the potential for conflicting goals engendering the misallocation of resources (Baker, 1992, 2002; Banker and Datar, 1989; Feltham and Xie, 1994; Holmstrom and Milgrom, 1991).

Similar to a hybrid unit of accountability, performance evaluations using multiple measures require that weights be assigned to each dimension of the system, an aspect of incentive-pay programmes that has tended to be ignored in the education sector-specific literature. Programmes like **New York City's** School-Wide Performance Bonus Program placed the greatest weight on student test scores (30% student achievement levels and 55% student achievement gains), while **Mexico's** *Carrera Magisterial* adopted a system whereby 70% of a teacher's evaluation relied on input-based measures such as seniority, academic degree, and professional preparation. Stakeholders will certainly request an explanation as to why various criteria are weighted the way they are, and these weights, for better or worse, can send a strong signal to students, parents, educators, and policy makers about the value programme designers placed on specific aspects of the schooling process.

While performance measurement systems that incorporate a myriad of metrics are likely to reduce unintended consequences by offering a more balanced evaluation system, at present, a considerable number of school systems do not have the organisational capacity to implement a programme incorporating multiple measures of teacher, team, or school performance. Even the most basic of incentive-pay programmes requires that a significant amount of personnel time be allocated to monitoring implementation, responding to and resolving

inconsistencies in the design of the programme, and addressing questions and complaints submitted by school personnel and other stakeholders.

In total, contemporary incentive-pay programmes have a lot of room for improvement when it comes to identifying the optimal mix of input- and output-focused performance measures and assigning weights to each of the evaluation criteria adopted. We recommend that school systems need to rely on rigorous, formative evaluations of their incentive programmes as trial and error will need to inform programme development and design in both the short- and long-run (Courty and Marschke, 2003; Podgursky and Springer, 2007). Moreover, rushing implementation before teachers and other school personnel understand how the programme works will in all likelihood set the stage for failure.

PERFORMANCE STANDARDS AND THRESHOLDS

Performance standards and thresholds determine the required level of performance for a school, team of teachers or individual teacher to secure a reward. The standards and thresholds defined by the incentive-pay programme dictate the number of units that can earn a bonus as well as what scale or minimum standards these units must meet. Our discussion reviews three forms of incentive-pay models, including linear models, step functions, and limited linear models. We also discuss the fine balance that must be established between performance thresholds and the number of units projected to earn a bonus award.

Linear models

Linear models, also known as continuous standards, refer to rewards based on the average level of or average gain in, desired outcome (Lazear, 2003). In school settings, an example using change in test scores is outlined by Lazear (2003), where β represents the relationship between earnings and test scores, and K is a constant that sets salaries at their appropriate level.

$$\text{Teacher's salary} = N(\Delta \text{Score})\beta - K$$

Linear models have several attractive features. Linear models are sufficiently robust, do not depend on the timing of the agent's information, and do not require all available information to make a bonus determination (Holmstrom and Milgrom, 1987). Another important feature is that linear models value progress at all parts of the achievement distribution (Lazear, 2003). Instead of focusing on students who are close to a certain cutoff point, a linear model allows for the aggregate growth of all students to determine bonus eligibility.

Countering arguments in favour of linear models, Kole (1997) demonstrated that an incentive-pay programme using a standard broad categorization ignores many important aspects of compensation contracts and can understate incentives. Her findings suggest that exogenous factors that influence performance may bias results if they, for example, impact high- or low-achieving students differentially. Furthermore, linear systems may present more of a logistical challenge for systems due to greater complexity of bonus calculations. Increased differentiation of salary among employees may require bonus determination at an individual level, a requirement that could place undue burden on human resources personnel.

Step functions

Step functions, or non-continuous models, distribute rewards based on meeting a minimum standard or threshold. While multiple standards or steps are possible, step functions establish a cutoff point below which either no bonus is awarded or penalties are incurred. Step functions are optimal when the output is very sensitive to the effort put forward by the unit of accountability in the general vicinity of the performance threshold (Holmstrom, 1979; Mirrlees, 1999). In other words, step functions create strong incentives for schools, teams of teachers, or individual teachers facing a significant reduction in pay if they reduce their effort even marginally.

The literature on incentive contracts has demonstrated that step functions are too easily manipulated, especially when they cover a fixed period of time (Dixit, 2002). In such a context, step functions provide comparatively weak incentives for effort either after the threshold is reached or after it becomes unattainable (Asch, 1990; Holmstrom and Milgrom, 1987; Oyer, 1998). For example, in a recent study of physician medical groups, Mullen, Frank and Rosenthal (2009) found that physicians' groups with baseline performance at or above the performance threshold for receipt of a bonus improved the least, but garnered the largest share of the bonus payments.

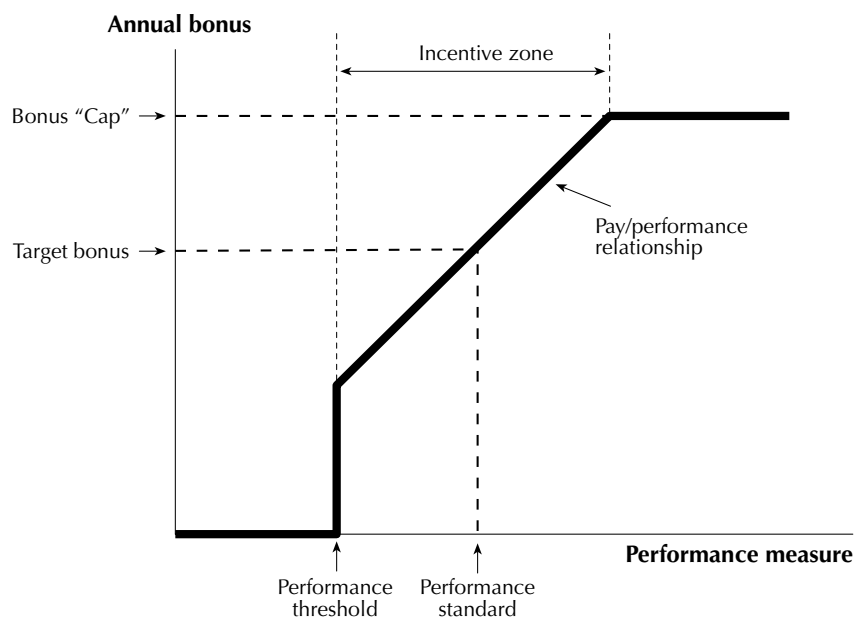
Limited linear models

It is possible to combine elements of both linear and step function models in an incentive programme using a limited linear model framework. As outlined by Murphy (2001), Figure 3.2 displays the theoretical relationship between a performance measure (x-axis) and the annual bonus (y-axis) within the context of a limited linear model. The baseline performance threshold establishes the minimum possible level of performance associated with a bonus award. Beyond this performance threshold floor, there is a positive linear relationship between increasing performance and an annual bonus award.

A limited linear model is further characterised in Figure 3.2 by the fact that at some point along performance continuum the maximum size of a bonus award is capped. Any increase in the performance of the unit of accountability beyond this predetermined level does not earn a school, team of teachers, or individual teacher any more money. The range of performance rewarded under an incentive-pay programme is called the incentive zone, and midway between the lower- and upper-performance thresholds is a predicted level of performance and corollary bonus that can be used in budget projections.

Figure 3.2

Performance standards and thresholds



Note: Adapted from Murphy (2001).

Several aspects of a limited linear model are particularly attractive. The model allows for improvement in the performance of the unit of accountability to increase the size of the annual bonus awarded at any point in the incentive zone. The limited linear model also restricts the bonus size of the lowest performers, thereby allowing for the possibility of the incentive-pay programme to send a strong signal to those at the bottom of the performance continuum. Finally, the financial risk typically imposed by extremely high performance is reduced by the bonus cap.

Threshold levels

Another consideration for thresholds is the targeted number of schools, teams of teachers, or individuals that designers of an incentive-pay programme expect to obtain a bonus award. Inherent in this decision is the balance between attainability and exclusivity, which not only plays a vital role in eliciting behavioural responses, but also dictates whether an incentive-pay programme will realise its full potential.

Reward programmes in **Bolivia** and **England** provide examples at each extreme. On the one hand, in Bolivia's merit wage programme, only 2% of teachers performed sufficiently on content knowledge exams to obtain a reward during the first year of the programme. While public outcry subsequently led to slightly more than 25% of teachers earning a reward in the second year of the merit wage programme, continual resistance on the part of teachers led to dissolution of the programme (Mizala and Romaguera, 2004).

In **England's** Performance Threshold System, on the other hand, about 88% of those teachers eligible to participate in the incentive-pay programme elected to do so. Among participating teachers, an astounding 97% met the predetermined standard established for earning a bonus award. Ultimately, the low level of perceived value and worth of attaining a performance award was attributed with the weak motivational influence elicited by the System (Atkinson *et al.*, 2009).

SIZE AND DISTRIBUTION OF BONUS AWARDS

The size of bonus, or payout level, refers to the amount of the total bonus award a school, team of teachers, or individual can earn. Distribution relates to the share of teachers that received a bonus award and how bonuses vary among employees. While no clear guidance exists on the optimal size of a bonus in a teacher incentive-pay programme, several studies suggested the size of bonus awards for teachers have been so small that the motivational value of most incentive systems have been compromised (Chamberlin, *et al.*, 2002; Heinrich, 2007; Malen 1999; Taylor and Springer, 2009).

In this section, we first discuss various topics found in the theoretical and empirical literature in an effort to offer a broad perspective on the size of bonuses and distribution of bonus awards. We then provide some general information on non-monetary incentives, which can take the form of promotion and advancement, positive feedback, public or private recognition, decreased regulations or requirements, increased status, greater decision-making capacity, access to in-house education or formal training, and so on. While the mechanism may differ from monetary rewards, the underlying principle of influencing behaviour remains constant.

Size of bonus

Table 3.2 displays summary statistics on the size of bonus awards for a number of incentive-pay programmes.¹⁹ In the **U.S.**, the size of the bonus award at the programme level ranges from a low of 0.4% to a high of 365% of a teacher's monthly salary. Both of these estimates come from the **Texas** Educator Excellence Grant programme. Estimates from international incentive-pay programmes are displayed in the bottom half of Table 3.2, where the smallest bonus award was associated with **Chile's** *Sistema Nacional de Evaluación de Desempeño de los Establecimientos Educativos* (SNED) programme. **Mexico's** *Carrera Magisterial* programme generally awarded the largest bonuses, which in some cases exceeded 200% of a teacher's monthly salary.

Table 3.2

Performance measures in incentive pay programmes

		Performance Measures	
	Period	Educational Inputs and Processes	Educational Outputs
United States			
School Incentive Program (Dallas, Texas)	1992 - 1995	Student attendance	Student achievement levels; Student test score gains; Student retention rates
Merit Pay Program (State of Michigan)	1996 - 1997	Student evaluation of teacher	Student retention rates
Achievement Challenge Pilot Project (Little Rock, Arkansas)	2005 - 2007	...	Student test score gains
Teacher Advancement Program (United States)	1999 - present	Professional development; Career ladder (three levels); Supervisor evaluation	Student test score gains (group- and school-level)
Governor Educator Excellence Grant Program (State of Texas)	2005 - 2008	Student attendance; Teacher attendance	Student achievement levels; Student test score gains
Texas Educator Excellence Grant Program (State of Texas)	2007 - present	Student attendance; Teacher attendance	Student achievement levels; Student test score gains
Professional Compensation System for Teachers (Denver, Colorado)	2005 - present	Knowledge and skills; Professional evaluation; High-need schools; Hard-to-staff subjects	Student test score gains; School performance
ASPIRE Program (Houston, Texas)	2005 - present	Teacher attendance	Student test score gains (group- and school-level)
Team Incentive Experiment (Round Rock, Texas)	2008 - present	...	Student test score gains
Project on Incentives in Teaching Experiment (Nashville, Tennessee)	2006 - 2009	...	Student test score gains
Q-Comp (State of Minnesota)	2006 - present	Career advancement; Professional development; Professional evaluation; High-need schools; Hard-to-staff subjects	Student achievement levels; Student test score gains
School-Wide Performance Bonus Program (New York City, New York)	2008- present	Student attendance; Student, parent and teacher perception of school learning environment	Student achievement levels; Student test score gains
International			
School Performance Program (Israel)	1996 - 1997	...	Number of credit units earned per student; Student receiving matriculation certification; School dropout rate
International Christelijk Steuenfonds Incentive Program (Buso and Teso, Kenya)	1998 - 1999	...	Student test score gains; Student achievement levels
Teacher-Incentive Experiment (Israel)	2001	...	Student achievement levels
Randomized Evaluation Project (Andhra Pradesh, India)	2006 - present	...	Student test score gains
Carrera Magisterial (Mexico)	1993 - present	Seniority; Academic degree; Professional level; Professional preparation; Professional performance	Student achievement levels
Salario al Mérito (Bolivia)	1998 - 1999	Teacher content knowledge	...
Incentivo Colectivo a Escuelas (Bolivia)	2001 - present	School organization; Teacher training and retention; Regularity in school management; Number of students per class; Student-teacher ratio; Teaching initiatives developed by teachers; Parent participation	Student pass rates; Student drop-out rates; Student repetition rates
Performance Threshold System (England)	1999 - present	Supervisor evaluation	Student test score gains
Sistema Nacional de Evaluación de Desempeño de los Establecimientos Educacionales (Chile)	1996 - present	School initiative; Improved working conditions	Student test score gains; Student achievement levels; Student pass rates; Student retention rates
Plan de Estímulos a la Labor Educativa Institucional (El Salvador)	2000 - present	School management; Educational management; Institutional planning; Teacher management.	Student drop-out; Student retention rates
Performance Pay Program (Portugal)	2007 - present	Teacher attendance; Attendance at training sessions; Management and pedagogical duties; Involvement in research proects; Parent feedback about teacher	Student test scores

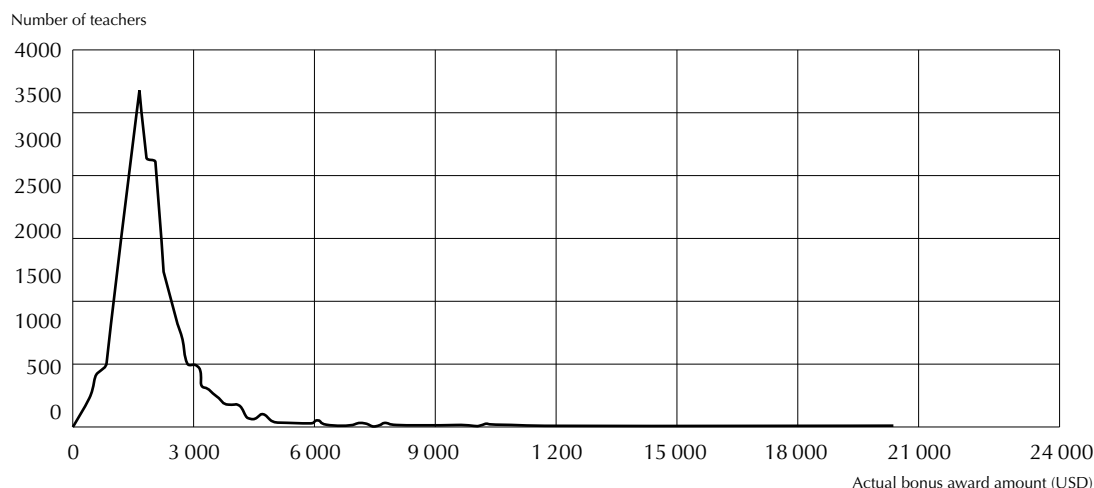
Sources: Atkinson et al. (2009); Clotfelter and Ladd (1996); Eberts, Hollenbeck, and Stone (2002); Glewwe, Holla, and Kremer (2009); Glewwe, Ilias, and Kremer (2008); Ladd (1999); Lavy (2002, 2004); McEwan and Santibanez (2005); Mizala and Romaguera (2004); Muralidharan and Sundararaman (2008); Podgursky and Springer (2007); Rodriguez (2002); Santibanez et al. (2007); Schacter and Thum (2004); Springer, Ballou, and Peng (2008); Springer et al. (2008, 2009); Springer and Winters (2009); Taylor, Springer, and Ehler (2009); Vegas and Umansky (2005); Winters, Ritter, Greene, and Marsh (2009).

Table 3.2 presents information that suggests the magnitude of the maximum bonus award has increased noticeably over time. Lower-secondary school mathematics teachers participating in a pay-for-performance experiment in Nashville, **Tennessee**, conducted by the National Center on Performance Incentives, set bonus awards as large as an additional USD 15 000 per year, or the equivalent of 280% of the average monthly salary. The National Center on Performance Incentives is also evaluating a teacher incentive experiment that awards the top one-third of treatment teams with a USD 6 000 bonus per teacher. Casual empiricism suggests the magnitude of the maximum bonus increased over time not only in response to past reform efforts that provided unappealing award amounts, but also to increased interest in learning whether teachers responded to substantial bonus awards, even if the incentive system lacked a complete array of measures.

Figure 3.3 displays how often teachers participating in the Governor Educator Excellence Grant (GEEG) programme received a bonus award, by bonus award amount (Springer *et al.*, 2008).²⁰ Bonus awards ranged from a low of USD 20 to a high of USD 20 462, with most teachers receiving between USD 1 000 and USD 3 000. What is most striking about the information is that nearly 90% of the teachers that received a bonus award earned less than USD 3 000, despite programme guidelines that encouraged teachers and schools to offer bonuses between USD 3 000 and USD 10 000 (Springer *et al.*, 2008). The pattern in Figure 3.3 suggests most schools approved incentive-pay programmes with a relatively large number of small bonus awards, which, according to Taylor and Springer (2009), did not appear to have induced any significant changes in teacher productivity.

Figure 3.3

Distribution of actual awards in the Governor Educator Excellence Grant programme



Source: Adapted from Springer *et al.* (2008).

Distribution of award

Bonus award distribution systems determine how evenly an incentive-pay programme distributes rewards to eligible employees. An egalitarian distribution plan distributes incentive money widely, in contrast to plans that reward larger sums of money to fewer schools, teams of teachers, or individuals. There is no clear guidance, however, whether an incentive-pay programme should reward a large number of relatively small awards to

teachers in a school, or reward a smaller number of teachers with a relatively large award (Taylor, Springer and Ehlert, 2009). Our discussion regarding the distribution of bonus awards therefore draws heavily on the general personnel incentive literature.

Proponents argue that individualist reward plans help create a meritocracy able to retain an organisation's highest performers, attract similar talent over the long run, send a clear signal to the lowest performers to improve or move elsewhere, and are more cost-effective (Ehrenberg and Smith, 1994; Milgrom and Roberts, 1992; Pfeffer and Langston, 1993; Zenger, 1992). At the same time, a growing body of research suggests egalitarian distributions promote cooperation and group performance, which are critical in participative organisations. Milgrom and Roberts (1992) suggested, moreover, that greater pay dispersion may elevate the performance of the lowest performers.

In a thorough review of the characteristics of incentive-pay programmes that have been adopted in **Bolivia, Chile, El Salvador, and Mexico**, Mizala and Romaguera (2004) reported that Bolivia's *Incentivo Colectivo a Escuelas* (ICE) awarded approximately 5% of eligible teachers a bonus ranging between 5% and 19% of annual income. Chile's SNED distributed a modestly larger number of bonus awards but the amount of those awards was slightly smaller in size, while Mexico's CM awarded the largest sums, with bonuses ranging between 27% and 224% of annual income (Mizala and Romaguera, 2004). In comparison to the ICE and SNED programmes offering bonuses to about 5% and 27% of eligible teachers, respectively, the CM incentive programme provided bonus awards to approximately 75% of eligible primary school teachers (Mizala and Romaguera, 2004).

A study of the Governor's Educator Excellence Grant programme in **Texas** examined the proposed distribution of awards to school personnel within schools participating in the programme. Taylor, Springer and Ehlert (2009) found that, on average, the proposed distribution of awards was more equal than the distribution of disposable income in the **U.S.**, but less equal than the distribution of teacher salaries in the schools. Acknowledging the fact that virtually all schools participating in the GEEG programme adopted highly egalitarian award distribution schemes, as well as the programme not having any discernible association with teacher productivity, a subsequent study of the programme suggested that when given the opportunity, teachers appear to design relatively weak incentive systems (Taylor and Springer, 2009).

Stakeholders interested in designing an incentive-pay programme need to consider the type of educational institution when establishing the size and distribution of bonus awards. For example, in a study of the relationship between wage dispersion and productivity in institutions of higher education, Pfeffer and Langston (1993) found less egalitarian salary structures decreased productivity and increased dissatisfaction among faculty. Yet, the negative effects were not as apparent in private colleges and universities, which the authors attributed to the fact that employee pay in most private institutions is not necessarily a matter of public record.

Newspapers, blogs, and other media outlets can obtain information on the distribution of bonus awards through the Freedom of Information Act (FOIA) in the U.S. Several newspaper outlets have obtained teacher level bonus information and then created venues for the general public to access and search this information via the Internet.²¹ The way in which information about wages is communicated to employees can greatly affect their attitude toward the programme and behaviour (Gerhart and Milkovich, 1990). We recommend stakeholders not only communicate regularly with the media, but also engage and work with teachers and other school personnel so they are well informed of the potential for publication of their bonus award amount.²²

Non-monetary incentives

While this chapter has focused primarily on programmes involving monetary incentives, non-monetary incentives can also function as workplace motivators. Many schools and school systems nominate and recognise certain teachers who perform above expectations, or offer teachers improved working conditions, unpaid leave,

mentoring and induction programmes, and job enlargement. In fact, in a review of reform strategies adopted in response to the problem of rural teacher recruitment among more than 20 developing countries, McEwan (1999) identified special training programmes, reduction in seniority requirements before promotion, and longer vacations as the most prevalent forms of non-monetary incentives.

Stigma and penalties have also been identified as potentially high-powered non-monetary incentives. Several states in **Germany** developed a reporting programme that publicised student learning gains association with schools, which Juerges, Richter and Schneider (2004) reported was a positive influence on the quality of teaching and student outcomes because teachers wanted to preserve and protect their reputation. Stigma and threats of sanction have also been found to have mostly positive effects in **Florida** (Chiang, forthcoming; Figlio and Rouse, 2006; West and Peterson, 2006). At the same time, while some stakeholders argued that information provided by the German reporting system should be used to inform personnel decisions (*i.e.* hiring, firing, and promotion/tenure), two national strikes organised by the teacher union resulted in political leaders in Germany committing to not use the results in this way (Juerges, Richter and Schneider, 2004).

PAYOUT FREQUENCY

Payout frequency refers to the rate of award distribution as well as the time interval between assessment of the incentivised activity and distribution of the performance award. The literature in psychology suggests incentives are most effective when the unit under observation is awarded consistently with minimum time between action and reward (Skinner, 1981). While some delay in gratification is possible, immediacy minimises the difficulty in association between an individual's behaviour and reward. Further, expectancy theory notes more frequent distribution and minimal temporal delay solidify connections between outcome and reward (Vroom, 1964). The delay that now occurs between teacher performance and payment of premium is longer than any behavioural psychologist would suggest for purposes of reinforcing good actions.

Most incentive-pay programmes in the education sector distribute awards on an annual basis, corresponding to the academic year, often due to the timing of results from end-of-year assessments. While practical, minimizing the time interval by tracking progress throughout the year may promote a more transparent link between action and reward. Further, assessing and rewarding behaviours more frequently during the school year will allow for consistent reinforcement and likely result in a more pronounced influence on behaviour (Hollensbe and Guthrie, 2000). This connection is present in the business sector; research conducted by the Consortium for Alternative Reward Strategies found compensation plans with greater payout frequency were linked to business performance improvements (McAdams and Hawk, 1992).

Eberts, Hollenbeck and Stone (2002) studied the effect of an incentive scheme in an alternative upper-secondary school in **Michigan** that rewarded teachers at multiple points throughout the school year. The programme was designed to address a growing dropout rate problem, and introduced a bonus system that paid teachers to raise course completion rates of students at risk of dropping out. Although the programme relied on a rather limited set of performance indicators, the intervention teachers were paid a 12% bonus each quarter if at least 80% of students from the previous quarter were retained.

Increasing the payout frequency may also decrease the likelihood of system gaming since behaviours are assessed on a more regular basis. As found in a study of Navy recruiters, for example, when their performance was evaluated at 12-month intervals, the recruiters' productivity significantly increased as the evaluation date approached. This spike in performance was characterised by a sizeable reduction in output shortly after the performance evaluation. Similarly, Courty and Marschke (1997) reported that training centers under the Job Training Partnership Act focused efforts on more qualified participants entering the programme, and the timing of their graduation, to meet certain standards.

CULTURAL CONSIDERATIONS

The impact of an incentive-pay system may reflect characteristics inherent in cultural contexts. Cultural dimensions, as described in Hofstede (1980, 2001), provide categories in which countries can be compared regarding reward preferences, including power distance, collectivism versus individualism, femininity versus masculinity, and uncertainty avoidance. Power distance refers to how people feel power should be distributed as well as how those with power should be viewed: a high power-distance environment would say power should be limited to fewer individuals whose decisions should be accepted, while lower power distance would have more people in control and would allow the challenging of leaders. A collective culture would place the allegiance to the group above that of the individual with this being reversed in an individualist culture. A feminine culture would value caring for others and establishing social support while a more masculine society would place a higher value on achievement, the accumulation of wealth, and assertiveness. Uncertainty avoidance refers to the extent a culture avoids situations that are ambiguous or have unpredictable outcomes.

Reward preferences can be directly linked to the chance of an incentive-pay programme actually meeting individual needs and cultural values found within particular countries. In their case study of four countries that varied on the dimensions established by Hofstede (1980, 2001), Chiang and Birtch (2005) found that countries with high masculinity scores were more likely to value material rewards as compared to those countries with more feminine scores. They found a similar preference for financial rewards between individualistic and collective cultures, while responses to incentive-pay programmes significantly diverged according to cultural factors. Even though the evidence reported by Chiang and Birtch (2005) is based on observations from a limited number of countries, their findings support the argument that cultural influences are an important aspect for education practitioners and policy makers to consider when designing and implementing an incentive-pay programme.²³

Furthermore, in Table 3.3 as defined by Hofstede (1980, 2001), we illustrate the power dimension scores in several countries with incentive programmes to further assess the unique nature of the cultural context of each incentive-pay programme. Take, for example, the programme designed and implemented by **Israel's** Ministry of Education, which was considered highly competitive because the incentive structure relied on a rank order tournament structure; that is, there were a limited number of people who could earn the reward and teachers eligible for a bonus award were knowingly in direct competition with one another. A plausible explanation for the acceptance of the programme by school personnel, while similar incentive programme designs have typically failed in other contexts, is that Israeli culture has a comparatively high individual score. While more research is definitely needed to determine specific relationships between design features of incentive-pay programmes and the cultural dimensions displayed in Table 3.3, it is important to consider the cultural context when designing an incentive-pay programme.

Table 3.3

Cultural dimensions for selected countries with incentive pay programmes

	Power Distance	Individualism	Masculinity	Uncertainty Avoidance
India	77.0	40	48	56
Israel	13.0	81	54	47
England (Great Britain)	35.0	35	89	66
Kenya (Eastern Africa)	64.0	52	27	41
Chile	63.0	23	28	86
Mexico	81.0	30	69	82
World Average	56.5	43	51	65

Note: Adapted from Hofstede (1980, 2001).

CONCLUSION

In this chapter we reviewed and synthesised many of the theoretical and empirical arguments connected with various design elements found in incentive-pay programmes. We focused on specific design elements including the incentive structure, unit of accountability, performance measures, performance targets and thresholds, size and distribution of bonus awards, and payout frequency. We also discussed the potential relationship between cultural dimensions and the design of incentive-pay programmes.

Incentive structures can impact the number of teachers likely to receive a bonus award as well as the required financial commitment needed to implement and sustain the incentive-pay programme over time. Rank ordered tournaments clearly limit the financial exposure of a fixed performance contract, which is advantageous considering the volatility and noise inherent in many performance measures found in the education sector. School personnel and organisations have historically resisted tournament incentive structures, however, because they can foster competition and other conditions that are incompatible with the norms of the teaching profession.

Incentive-pay programmes have overwhelmingly conceptualised the unit of accountability as being either at the individual- or school-level, or a combination of the two. Yet, this conceptualization misses a potentially important unit found within schools: grade-level or departmental teams of teachers or any other within-school unit. Education stakeholders should consider more closely within-school teams as the unit of accountability, which several theoretical and empirical studies from outside the education sector indicated have the potential to combine the cohesive benefits of group accountability with the individual belief that productivity directly impacts bonus eligibility. The education system should also continue to explore more complex hybrid models that may be more suitable in multi-product, multi-purpose environments.

Careful consideration needs to be given to how the performance of schools, teams of teachers, or individuals is measured. A large literature indicates employee behaviour will tend to shift in the direction of incentivised actions, and both potential positive and negative outcomes should be anticipated. Multiple measures can often minimise gaming behaviour and provide a more valid, reliable, fair and comprehensive assessment of teacher, team and school effectiveness, all of which are crucial to the motivational effect of the incentive-pay programme. Furthermore, there are reports of more and more private sector firms adopting both input- and output-based measures when assessing employee performance.

In regards to performance standards and thresholds, incentive-pay programmes have usually adopted either a linear or step function model. Most researchers have advocated for linear models due to design simplicity and their robust nature, even though some of the behavioural economic literature cautions that a linear model ignores important information about the actual production process. Given the relatively sparse literature on performance standards and thresholds, it is important for future evaluation research and development work to closely examine these structures in an effort to further inform stakeholders of the relative merits and weaknesses of each system.

The size and distribution of incentive-pay awards have tended to vary dramatically among programmes, although a few relatively recent studies have suggested that teachers may be predisposed to adopting highly egalitarian award distribution schemes. When designing an incentive-pay programme, it is critically important to keep in mind that bonus awards need to be sufficiently large to elicit behavioural responses and not so large that gaming or cheating becomes a concern. We further recommend that incentive-pay programmes in the education sector increase the frequency in which performance is evaluated and awards are distributed. To the point that they are feasible, multiple evaluations and performance objectives or benchmarks during the school year strengthen the connection between behaviour and reward. Such arrangements may help teachers, teams of teachers, or schools to monitor and alter behaviour at several points in time instead of basing their performance bonus on a one-time annual assessment.

Cultural norms and preferences toward design elements of incentive-pay programmes are likely important considerations during the design and implementation phase. While a rank-order tournament evaluating individual teacher performance may have elicited support and had a positive impact on student achievement and teacher behaviour in **Israel**, countries predisposed toward more collective cultural processes and norms might be more successful adopting a group-level or hybrid model using a fixed performance contract incentive system. Even within a country, regions may also differ in socio-cultural characteristics as well as the economic situations that influence the ideal design of incentive systems.

It is clear the literature on the design components of incentive-pay programmes in the education sector still needs to be developed before we can define how the first or second-best system should be designed. Educational research, practice, and policy needs to continue to support policy innovation combined with rigorous, independent evaluations of both the short- and long-run impacts on student achievement, teacher attitudes and behaviour, and organisational dynamics.

NOTES

1. Incentive-pay programmes have grown in popularity. High-profile programmes have included **Bolivia's** Merit Wages and ICE program (Vegas and Umansky, 2005; Mizala and Romaguera, 2004); **Chile's** *Sistema Nacional de Evaluación del Desempeño de los Establecimientos Educacionales* (SNED) (Mizala and Romaguera, 2003); **El Salvador's** school awards program (Rodriguez, 2002); **Mexico's** *Carrera Magisterial* (McEwan and Santibanez, 2005; Santibanez *et al.*, 2007); **New York City's** School-Wide Performance Bonus Program (Springer and Winters, 2009); **Texas'** Governor's Educator Excellence Award Programs (Springer *et al.*, 2008; Springer *et al.*, 2009); the **U.S.** Teacher Incentive Fund (Podgursky and Springer, 2007); programmes developed by **Israel's** Ministry of Education (Lavy, 2002; 2007); and experiments in Andhra Pradesh, **India** (Muralidharan and Sundararaman, 2008) and in the Busia and Teso districts of western **Kenya** (Glewwe, Ilias, and Kremer, 2008). In the **U.S.**, pay-for-performance experiments have recently been implemented and are currently being evaluated in Nashville, **TN**, New York City, **NY**, and Round Rock, **TX**.

2. In response to the adoption of **Portugal's** performance-pay policy, there were two national strikes led by teachers and their unions (Martins, 2009).

3. While at first sight there appears to be a significant increase in student learning, the percentage of schools meeting the proficiency standard may not in fact represent a true increase in student achievement. There is a healthy literature on weaknesses of high-stakes assessments and test-based accountability systems, including test scores becoming artificially inflated because a state may have adopted a new standardised assessment, recalibrated proficiency standards, or repeatedly administered a similar test form from one school year to the next (see, for example, Linn, 2000; Koretz, 2003; 2008). In describing what is widely-known as the "Lake Wobegon Effect" [This may need an explanation for the international audience](e.g., scenario where schools or teachers receive overly positive ratings), Cannell (1987; as cited in Linn, 2000) argued that "standardized, nationally normed achievement tests give children, parents, school systems, legislatures, and the press inflated and misleading reports on achievement levels" (p. 7).

4. In the **U.S.**, it has been argued that the increased compression of the federal sector's compensation system relative to that in the private sector has made it more difficult for the federal sector to recruit and retain high-quality employees (Gibbs, 2001; Borjas, 2002; Asch, 2005).

5. Work being conducted by researchers at the National Center on Performance Incentives has begun to explore many of the operational complexities associated with incentive-pay programmes and the effect these complexities have on measures of teacher and team performance.
6. Control condition refers to those classrooms not exposed to one of the four interventions.
7. An alternate draw is a selection process where the leader of the first team selects from a pool of eligible candidates and then the leader of the next team selects from the remaining pool and so on until all eligible candidates are selected.
8. Independent evaluations have been conducted by Springer, Ballou, and Peng (2008) and are currently being conducted by Glazerman, Allison, McKie, and Carey (2009).
9. Levin and McEwan (2001) offer a very thorough and practical overview of cost-effectiveness analysis when evaluating or considering educational programmes and policies.
10. The Center for Educator Compensation Reform offers a number of useful resources regarding the implementation of incentive-pay programmes. For more information, visit <http://www.cecr.ed.gov>
11. Previous research typically codes subjective evaluations of teacher performance as an input measure, output measure, or process measure. The same holds true for student and teacher attendance. In this review, we consider these process measures.
12. Even though the vast majority of states in the **U.S.** have funded incentive plans around teacher recruitment and retention, as observed by Loeb and Miller (2007), most of these programs are not well-aligned with teacher labour market realities, nor is the receipt of an incentive award usually contingent on teacher effectiveness.
13. For more information visit, <http://cppe.wceruw.org/index.php>
14. In his review of the different types of corruption and their potential causes in the education sector, Heyneman (2004) recommends four types of reform to minimise the risk. These include educational structures, the processes of management and adjudication, the mechanism of prevention when wrongdoing occurs, and the system of sanctions.
15. Carnoy, Brodziak, Molina, and Solcias (2007) report on constraints of a school-based incentive programme when using student performance data to reward performance on inter-cohort gains as opposed to intra-cohort test score gains.
16. For a more complete description of issues related to VAM see Chapter 5 by Dan Goldhaber.
17. Ballou (2009) discussed the implications of scale properties in value-added measurements. He analysed several possible methods of formulating value-added estimates from tests that are scored according to Item Response Theory (IRT). He found measuring ability on an interval scale using IRT requires a conjoint structure between students and test items, conditions that are both difficult to meet and verify. Value-added assessments of teacher quality may introduce volatility due to the scale properties of tests that may incorrectly be attributed to teacher effectiveness.
18. Jacob and Levitt's (2003) analysis of cheating within the **Chicago** Public School system provided two methods of detection. The first method involved searching for large-scale uniform patterns in student responses, patterns which would be statistically impossible in normal student responses. Secondly, unusually high student performances in one year that are followed by declines in performance in subsequent years were selected for closer observation. In international programs in **Israel**, **Kenya**, and **India** there were penalties for students who did not take the exam in an effort to prevent schools and teachers from intentionally discouraging low-performing students from test participation. In each of these cases, test participation rates either stayed the same or significantly increased (Lavy, 2004; Glewwe, Ilias, and Kremer, 2003; Muralidharan and Sundararaman, 2008).
19. Table 3.2 builds on evidence reported in Taylor, Springer, and Ehlert (2009).
20. The information displayed in Figure 3.3 does not include those teachers that participated in the TEEG program but did not earn a bonus award during the 2006-07 school year.
21. The **Houston** Chronicle posts information on bonus award amounts delineated by employee name, position, and school (see, for example, <http://www.chron.com/news/houston/bonuses>). The **Austin** Statesman similarly publishes names of teachers earning bonuses under the district's Strategic Compensation Initiative.
22. Burns, Gardner, Muesswen (2009) provide additional discussion within the context of **Austin** Independent School District's Strategic Compensation Initiative.
23. In a related study, Ramamoorthy and Carroll (1998) reported a significant relationship between high levels of individualism and alternative human resource management practices such as merit pay.

References

- Adnett, N.** (2003), "Reforming Teachers' Pay: Incentive Payments, Collegiate Ethos and UK Policy", *Cambridge Journal of Economics*, Vol. 27, No. 1, pp. 145-147.
- Asch, B.** (1990), Navy Recruiter Productivity and the Freeman Plan, *RAND Monograph R-3713-FMP*, Santa Monica, CA.
- Asch, B.** (2005), "The Economic Complexities of Incentive Reforms", in R. Klitgaard and P.C. Light (eds.), *High-Performing Government: Structure, Leadership, Incentives*, RAND Corporation, Santa Monica, CA.
- Atkinson, A, S. Burgess, B. Croxson, P. Gregg, C. Propper, H. Slater and D. Wilson** (2009), "Evaluating the Impact of Performance-Related Pay for Teachers in England", *Labour Economics*, Vol. 16, No. 3, pp. 251-261.
- Baker, G.** (1992), "Incentive Contracts and Performance Measurement", *Journal of Political Economy*, Vol. 100, No. 3, pp. 598-614.
- Baker, G.** (2002), "Distortion and Risk in Optimal Performance Contracts", *The Journal of Human Resources*, Vol. 37, No. 4, pp. 728-751.
- Baker, G., M. Jensen and K. Murphy** (1988), "Compensation and Incentives: Practice vs. Theory", *Journal of Finance*, Vol. 43, No. 3, pp. 593-616.
- Ballou, D.** (2001), "Pay for Performance in Public and Private Schools", *Economics of Education Review*, Vol. 20, No. 1, pp. 51-61.
- Ballou, D. and M. Podgursky** (1997), "Reforming Teacher Training & Recruitment: A Critical Appraisal of the Recommendations of the National Commission on Teaching and America's Future", *Government Union Review*, Vol. 17, No. 4, pp. 1-53.
- Ballou, D. and M. Podgursky** (2009), "Responses to No Child Left Behind Among Traditional Public Schools", *Peabody College of Vanderbilt University Working Paper*, Nashville, TN.
- Banker, R. and S. Datar** (1989), "Sensitivity, Precision, and Linear Aggregation of Signals for Performance Evaluation", *Journal of Accounting Research*, Vol. 27, No. 1, pp. 21-39.
- Baron, J. and D. Kreps** (1999), *Strategic Human Resources: Frameworks for General Managers*, Wiley, New York.
- Berg, P., E. Appelbaum, T. Bailey and A. Kalleberg** (1996), "The Performance Effects of Modular Production in the Apparel Industry", *Industrial Relations*, Vol. 35, No. 3, pp. 356-373.
- Bishop, J.** (1996). "Signaling, Incentives, and School Organization in France, the Netherlands, Britain, and the United States", in E. Hanushek and D. Jorgenson (eds.), *Improving America's Schools: The Role of Incentives*, National Research Council, Washington, D.C.
- Bok, D.** (1993), *The Cost of Talent: How Executives and Professionals are Paid and How It Affects America*, The Free Press, New York.
- Booher-Jennings, J.** (2005), "Below the Bubble: 'Educational Triage' and the Texas Accountability System", *American Educational Research Journal*, Vol. 42, No. 2, pp. 231-268.
- Bourne, R. and B. MacArthur** (1970), *The Struggle for Education: 1870-1970*, School Masters Publishing Company, London.
- Bowles, S. and H. Gintis** (2002), "Schooling in Capitalist America Revisited", *Sociology of Education*, Vol. 75, No. 1, pp. 1-18.
- Bryk, A. and B. Schneider** (2002), *Trust in Schools: A Core Resource for Improvement*, Russell Sage Foundation, New York.
- Burgess, S., B. Croxson, P. Gregg and C. Propper** (2001), "The Intricacies of the Relationship Between Pay and Performance for Teachers: Do Teachers Respond to Performance Related Pay Schemes", *CMPO Working Paper*, Bristol.

- Chamberlin, R., T. Wragg, G. Haynes, and C. Wragg** (2002), "Performance-Related Pay and the Teaching Profession: A Review of the Literature", *Research Papers in Education*, Vol. 17, No. 1, pp. 31-49.
- Che, Y. and S. Yoo** (2001), "Optimal Incentives for Teams", *The American Economic Review*, Vol. 91, No. 3, pp. 525-541.
- Chiang, H.** (forthcoming), "How Accountability Pressure of Failing Schools Affects Student Achievement", *Journal of Public Economics*.
- Chiang, F. and T. Birtch** (2005), "A Taxonomy of Reward Preferences: Examining Country Differences", *Journal of International Management*, Vol. 11, pp. 357-375.
- Clotfelter, C., H. Ladd and J. Vigdor** (2004), "Do School Accountability Systems Make It More Difficult for Low-performing Schools to Attract and Retain High-Quality Teachers?", *Journal of Policy Analysis and Management*, Vol. 23, No. 2, pp. 251-271.
- Clotfelter, C., H. Ladd and J. Vigdor** (2007), "Are Teacher Absences Worth Worrying About in the U.S.?", *NBER Working Paper 13648*, National Bureau of Economic Research, Cambridge, MA.
- Corcoran, S., W. Evans and R. Schwab** (2004), "Women, the Labor Market, and the Declining Relative Quality of Teachers", *Journal of Policy Analysis and Management*, Vol. 23, No. 3, pp. 449-470.
- Courty, P. and G. Marschke** (1997), "Measuring Government Performance: Lessons from a Federal Job-Training Program", *American Economic Review*, Vol. 87, No. 2, pp. 383-388.
- Courty, P. and G. Marschke** (2003), "Dynamics of Performance-Measurement Systems", *Oxford Review of Economic Policy*, Vol. 19, pp. 68-284.
- Courty, P., C. Heinrich and G. Marschke** (2005), "Setting the Standard in Performance Measurement Systems", *International Public Management Journal*, Vol. 8, No. 3, pp. 321-347.
- Crewson, E.** (1997), "Public-Service Motivation: Building Empirical Evidence of Incidence and Effect", *Journal of Public Administration Research and Theory*, Vol. 7, No. 4, pp. 499-518.
- Cutler, T. and B. Waine** (2004), "Performance Without Pay? Managing School Budgets Under Performance Related Pay: Evidence From the First Year of PRP", *Financial Accountability and Management*, Vol. 20, No. 1, pp. 57-75.
- Danielson, C.** (1996), *Enhancing Professional Practice: A Framework for Teaching*. Association for Supervision and Curriculum Development, Alexandria, VA.
- Deutsch, M.** (1985), *Distributive Justice: A Social-Psychological Perspective*, Yale University Press, New Haven, CT.
- Dixit, A.** (2002), "Incentives and Organizations in the Public Sector", *Journal of Human Resources*, Vol. 37, No. 4, pp. 696-727.
- Duflo, E., R. Hanna and S. Ryan** (2005), "Monitoring Works: Getting Teachers to Come to School", *NBER Working Paper 11880*, National Bureau of Economic Research, Cambridge, MA.
- Eberts, R.** (2007), "Teacher Unions and Student Performance: Help or Hindrance?", *The Future of Children*, Vol. 17, No. 1, pp. 175-200.
- Eberts, R., K. Hollenbeck and J. Stone** (2002), "Teacher Performance Incentives and Student Outcomes", *Journal of Human Resources*, Vol. 37, No. 4, 913-927.
- Ehrenberg, R. and R. Smith** (1994), *Modern Labor Economics: Theory and Public Policy*, Fifth Edition, Harper Collins.
- Elmore, R.F., C.H. Abelman, and S.H. Fuhrman**, (1996). "The new accountability in state education reform: From process to performance". In H. Ladd (ed.), *Holding schools accountable: Performance-based reform in education*. (pp. 65±98). Washington, DC: Brookings Institution.
- Feltham, G. and J. Xie** (1994), "Performance Measure Congruity and Diversity in Multi-Task Principal/Agent Relations", *The Accounting Review* Vol. 69, No. 3, pp. 429-453.
- Figlio, D. and C. Rouse** (2007), "Do Accountability and Voucher Threats Improve Low-Performing Schools", *Journal of Public Economics*, Vol. 90, No. 1-2, pp. 239-255.
- Firestone, W.** (1994), "Redesigning Teacher Salary Systems for Educational Reform", *American Educational Research Journal*, Vol. 31, No. 3, pp. 549-574.

- Freeman, R.** and **A. Gelber** (2006), "Optimal Inequality/Optimal Incentive: Evidence from a Tournament", *NBER Working Paper 12588*, National Bureau of Economic Research, Cambridge, MA.
- Freixas, X., R. Guesnerie** and **J. Tirole** (1985), "Planning under Incomplete Information and the Ratchet Effect", *Review of Economic Studies*, Vol. 52, No. 2, pp. 173-191.
- Gaynor, M.** and **M. Pauly** (1990), "Compensation and Productive Efficiency in Partnerships: Evidence from Medical Group Practice", *Journal of Political Economy*, Vol. 98, No. 3, pp. 543-577.
- Gerhart, B.** and **G. Milkovich** (1990), "Organizational Differences in Managerial Compensation and Financial Performance", *Academy of Managerial Journal*, Vol. 33, No. 4, pp. 663-691.
- Ginther, D., R. Haveman** and **B. Wolfe** (2000), "Neighborhood Attributes as Determinants of Children's Outcomes: How Robust are the Relationships", *The Journal of Human Resources*, Vol. 35, No. 4, pp. 603-642.
- Glewwe, P., A. Holla** and **M. Kremer** (2009), "Teacher Incentives in the Developing World", in M.G. Springer (ed.), *Performance Incentives: Their Growing Impact on American K-12 Education*, Brookings Institution Press, Washington, D.C.
- Glewwe, P., N. Ilias** and **M. Kremer** (2008), *Teacher Incentives in the Developing World*, Harvard University, Cambridge, MA.
- Goldhaber, D.** (2002), "The Mystery of Good Teaching," *Education Next*, Vol. 2, No. 1.
- Goldhaber, D.** (2009), "The Politics of Teacher Pay Reform", in Springer, M.G. (ed.), *Performance Incentives: Their Growing Impact on American K-12 Education*, Brookings Institution Press, Washington, D.C.
- Goldhaber, D., D. Brewer** and **D. Anderson** (1999), "A Three-Way Error Components Analysis of Educational Productivity", *Education Economics*, Vol. 7, No. 3, pp. 191-208.
- Gonring, P., P. Teske** and **B. Jupp** (2007), *Pay-for-Performance Teacher Compensation: An Inside View of Denver's ProComp Plan*, Harvard Education Press, Cambridge, MA.
- Gramlich, E.** and **P. Koshel** (1975), *Educational Performance Contracting*, Brookings Institution, Washington, D.C.
- Guthrie, J.** and **C. Prince** (2009), "Paying for and Sustaining a Performance-Based Compensation System", Center for Educator Compensation Reform, Washington, D.C.
- Hallinger, P.** and **J. Murphy** (1986). "The Social Context of Effective Schools", *American Journal of Education*, Vol. 94, No. 3, pp. 328-355.
- Hamilton, L.** (2005), "Lessons from Performance Measurement in Education", in R. Klitgaard and P.C. Light (eds.), *High-Performing Government: Structure, Leadership, Incentives*, RAND Corporation, Santa Monica, CA.
- Hamilton, B., J. Nickerson** and **H. Owan** (2003), "Team Incentives and Worker Heterogeneity: An Empirical Analysis of the Impact of Teams on Productivity and Participation", *Journal of Political Economy*, Vol. 111, No. 3, pp. 465-497.
- Hannaway, J.** (1992), "Higher Order Skills, Job Design, and Incentives: An Analysis and Proposal", *American Educational Research Journal*, Vol. 29, No. 1, pp. 3-21.
- Hannaway, J.** (1996). "Management Decentralization and Performance-Based Incentives: Theoretical Considerations for Schools", in E. Hanushek and D. Jorgenson (eds.), *Improving America's Schools: The Role of Incentives*, National Research Council, Washington, D.C.
- Hansen, D.** (1997), "Worker Performance and Group Incentives: A Case Study," *Industrial Labor Relations Review*, Vol. 51, pp. 37-49.
- Heinrich, C.J.** (2007), "Evidence-Based Policy and Performance Management: Challenges and Prospects in Two Parallel Movements", *The American Review of Public Administration*, Vol. 37, No. 3, pp. 255-277.
- Heneman, H., A. Milanowski** and **S. Kimball** (2007), "Teacher Performance Pay: Synthesis of Plans, Research, and Guidelines for Practice", *CPRE Policy Briefs RB-46*, Madison, WI.
- Hofstede, G.** (1980), *Culture's Consequences: International Differences in Work-Related Values*, Sage Publications, Beverly Hills, CA/ London.

- Hofstede, G.** (2001), *Culture's Consequences: Comparing Values, Behaviors, Institutions, and Organizations Across Nations*, Sage Publications, Thousand Oaks, CA.
- Hollensbe, E. and J. Guthrie** (2000), "Group Pay-for-Performance Plans: The Role of Spontaneous Goal Setting", *The Academy of Management Review*, Vol. 25, No. 4, pp. 864-872.
- Holmstrom, B.** (1979). "Moral Hazard and Observability", *The Bell Journal of Economics*, Vol. 10, No. 1, pp. 74-91.
- Holmstrom, B.** (1982), "Moral Hazard in Teams", *Bell Journal of Economics*, Vol. 13, No. 2, pp. 324-340.
- Holmstrom, B. and P. Milgrom** (1987), "Aggregation and Linearity in the Provision of Intertemporal Incentives", *Econometrica*, Vol. 55, No. 2, pp. 303-328.
- Holmstrom, B. and P. Milgrom** (1991), "Multitask Principal-Agent Analysis: Incentive Contracts, Asset Ownership, and Job Design", *Journal of Law, Economics, and Organization*, Vol. 7, pp. 24-52.
- Hoxby, C. and A. Leigh** (2004), "Pulled Away or Pushed Out? Explaining the Decline of Teacher Aptitude in the United States", *American Economic Review*, Vol. 94, No. 2, pp. 236-240.
- Neale, M., V. Huber and G. Northcraft** (1987), "The Framing of Negotiations: Contextual Versus Task Frames", *Organizational Behavior and Human Decision Processes*, Vol. 39, pp. 228-241.
- Indjejikian, R., P. Lenk and D. Nanda** (2000), "Targets, Standards and Performance Expectations: Evidence from Annual Bonus Plans", *SSRN Working Paper*, http://papers.ssrn.com/sol3/papers.cfm?abstract_id=213628, accessed September 20, 2009.
- Jacob, B. and S. Levitt** (2003), "Rotten Apples: An Investigation of the Prevalence and Predictors of Teacher heating", *Quarterly Journal of Economics*, Vol. 118, No. 3, pp. 843-877.
- Jacob, B. and L. Lefgren** (2005), "Principals as Agents: Subjective Performance Measures in Education", *KSG Working Paper No. RWP05-040*, Cambridge, MA.
- Jacobson, S.** (1992), "Performance-Related Pay for the Teacher: The American Experience", in H. Tomlinson (ed.), *Performance-Related Pay in Education*, Routledge, London.
- Johnson, S.M.** (2004). "Paralysis or Possibility: What Do Teacher Unions and Collective Bargaining Bring", in R. Henderson, W. Urban, and P. Wolman (eds.), *Teacher Unions and Education Policy: Retrenchment or Reform?*, Elsevier, Oxford, U.K.
- Juerges, H., W. Richter and K. Schneider** (2004), "Teacher Quality and Incentives: Theoretical and Empirical Effects of Standards on Teacher Quality", *CESifo Working Paper No. 1296*, CESifo, Munich.
- Kandel, E. and E. Lazear** (1992), "Peer Pressure and Partnerships", *The Journal of Political Economy*, Vol. 100, No. 4, pp. 801-817.
- Kane, T. and D. Staiger** (2001), "Improving School Accountability Measures", *NBER Working Paper 8156*, National Bureau of Economic Research, Cambridge, MA.
- Kelley, C.** (1999), "The Motivational Effects of School-Based Performance Awards", *Journal of Personnel Evaluation in Education*, Vol. 12, No. 4, pp. 54-78.
- Kelley, C.** (2002), "Financial Incentives in State Accountability Systems: Performance Pay for Teachers", Center on Reinventing Public Education and the Washington State Academic Achievement and Accountability Commission, Seattle, WA.
- Kimball, S., B. White, A. Milanowski and G. Borman** (2004), "Examining the Relationship Between Teacher Evaluation and Student Assessment Results in Washoe County", *Peabody Journal of Education*, Vol. 79, No. 4, pp. 54-78.
- Kingdon, G. and F. Teal** (2008), "Teacher Unions, Teacher Pay and Student Achievement in India", *CESifo Working Paper 2428*, CESifo, Munich.
- Kohn, A.** (1999), *Punished by Rewards: The Trouble with Gold Stars, Incentive Plans, A's, Praise, and Other Bribes*, Houghton Mifflin, Boston, MA.
- Kole, S.** (1997), "The Complexity of Compensation Contracts", *Journal of Financial Economics*, Vol. 43, pp. 79-104.

- Kremer, M.** and **D. Chen** (2001), "Interim Report on a Teacher Incentive Program in Kenya", University of Chicago, Chicago, IL <http://home.uchicago.edu/~dlc/papers/Preschool.pdf>.
- Ladd, H.** (1999), "The Dallas School Accountability and Incentive Program: An Evaluation of Its Impacts on Student Outcomes", *Economics of Education Review*, Vol. 18, No. 1, pp. 1-16.
- Lavy, V.** (2002), "Evaluating the Effect of Teachers' Group Performance Incentives on Pupil Achievement", *Journal of Political Economy*, Vol. 110, No. 6, pp. 1286-1317.
- Lavy, V.** (2004), "Performance Pay and Teachers' Effort, Productivity and Grading Ethics", *NBER Working Paper 10622*, National Bureau for Economic Research, Cambridge, MA.
- Lavy, V.** (2007), "Using Performance-Based Pay to Improve the Quality of Teachers", *The Future of Children*, Vol. 17, No. 1, pp. 87-109.
- Lavy, V.** (2008), "Gender Differences in Market Competitiveness in a Real Workplace: Evidence from Performance-Based Tournaments Among Teachers", *NBER Working Paper 14338*, National Bureau of Economic Research, Cambridge, MA.
- Lazear, E.** (1995), *Personnel Economics*, MIT Press, Cambridge, MA.
- Lazear, E.** (1998), *Personnel Economics for Managers*, Wiley, New York.
- Lazear, E.** (2000), "Performance Pay and Productivity", *American Economic Review*, Vol. 90, No. 5, 1346-1361.
- Lazear, E.** (2003), "Teacher Incentives", *Swedish Economic Policy Review*, Vol. 10, pp. 197-213.
- Lazear, E.** and **M. Gibbs** (2009), *Personnel Economics in Practice*, Wiley, Hoboken, NJ.
- Leventhal, T.** and **J. Brooks-Gunn** (2004), "A Randomized Study of Neighborhood Effects on Low-Income Children's Educational Outcomes", *Developmental Psychology*, Vol. 40, No. 4, pp. 488-507.
- Lewis, J.** and **M. Springer** (2008), "Performance Incentives in Texas: Why Schools Chose Not to Participate", *Working Paper #2008-18*, National Center on Performance Incentives, Nashville, TN.
- Liang, X.** (1999), "Teacher Pay in 12 Latin American Countries: How Does Teacher Pay Compare to Other Professions, What Determines Teacher Pay, and Who Are the Teachers?", *LCSHD Paper Series No. 49*, World Bank, Washington, D.C.
- Locke, E.A.** (1968), "Toward a Theory of Task Motivation and Incentives", *Organizational Behavior and Human Performance*, Vol. 3, pp. 157-189.
- Locke, E.** and **G. Latham** (1990), *A Theory of Goal Setting and Task Performance*, Prentice-Hall, Englewood Cliffs, NJ.
- Locke, E., G. Latham** and **K. Smith** (1990). *A Theory of Goal Setting and Task Performance*, Prentice-Hall, Upper Saddle River, NJ.
- Locke, E., L. Saari, K. Shaw, G. Latham** (1981), "Goal Setting and Task Performance: 1968-1980", *Psychological Bulletin*, Vol. 90, No. 1, pp. 125-152.
- Longenecker, C., D. Gioia** and **H. Sims** (1987), "Behind the Mask: The Politics of Employee Appraisal", *Academy of Management Executive*, Vol. 1, No. 3, pp. 183-193.
- Lortie, D.** (1979), *Schoolteacher: A Sociological Study*, University of Chicago Press, Chicago, IL.
- Malen, B. (1999), "On Rewards, Punishments, and Possibilities: Teacher Compensation as an Instrument for Education Reform", *Journal of Personnel Evaluation in Education*, Vol. 12, No. 4, pp. 387-394.
- Marsden, D.** and **R. Belfield** (2006), "Pay for Performance Where Output is Hard to Measure: The Case of Performance Pay for School Teachers", in B. Kaufman, E. Bruce and D. Lewin (eds.), *Advances in Industrial and Labor Relations*, JAI Press, London.
- McAdams, J.** and **E. Hawk** (1992), *Capitalizing on Human Assets*, American Compensation Association, Scottsdale, AZ.
- McCaffrey, D., B. Han** and **J. Lockwood** (2009), "Incentive System Design and Measurement", in M.G. Springer (ed.), *Performance Incentives: Their Growing Impact on American K-12 Education*, Brookings Institution Press, Washington, D.C.
- McCollum, S.** (2001), "How Merit Pay Improves Education", *Educational Leadership*, Vol. 58, No. 5, pp. 21-24.

- McEwan, P.** (1999), "Recruitment of Rural Teachers in Developing Countries: An Economic Analysis", *Teaching and Teacher Education*, Vol. 15, No. 8, pp. 849-859.
- Merchant, K.A. and J. Manzoni** (1989), "The Achievability of Budget Targets in Profit Centers: A Field Study", *The Accounting Review*, Vol. 64, pp. 539-558.
- Milgrom, P. and J. Roberts** (1990), "Rationalizability, Learning, and Equilibrium in Games with Strategic Complementarities", *Econometrica*, Vol. 58, No. 6, pp. 1255-1277.
- Milgrom, P. and J. Roberts** (1992), *Economics, Organization & Management*, Prentice Hall, Englewood Cliffs, CA.
- Milanowski, A.** (2003), "The Varieties of Knowledge and Skill-Based Pay Design: A Comparison of Seven New Pay Systems for K-12 Teachers", *Education Policy Analysis Archives*, Vol. 11, No. 4, <http://epaa.asu.edu/epaa/v11n4/>, accessed April 17, 2009.
- Milanowski, A.** (2004), "The Relationship between Teacher Performance Evaluation Scores and Student Achievement: Evidence from Cincinnati", *Peabody Journal of Education*, Vol. 79, No. 4, pp. 33-53.
- Mirrlees, J.** (1999), "The Theory of Moral Hazard and Unobservable Behaviour: Part I", *Review of Economic Studies*, Vol. 66, No. 1, pp. 3-21.
- Mizala, A. and P. Romaguera** (2003), *Scholastic Performance and Performance Awards*, Centro de Economica Aplicada, Santiago, Chile.
- Mizala, A. and P. Romaguera** (2004), "Schools and Teachers Performance Incentives: The Latin American Experience", *International Journal of Educational Development*, Vol. 24, No. 6, pp. 739-754.
- Mullen, C. and K. Slagle** (2007), "Teacher Perceptions of Systems-Driven Quality Teaching Initiatives", *Academic Leadership*, Vol. 3, No. 1, pp. 1-20.
- Mullen, K., R. Frank and M. Rosenthal** (2009), "Can You Get What You Pay For? Pay-for-Performance and the Quality of Healthcare Providers", *NBER Working Paper 14886*, National Bureau of Economic Research, Cambridge, MA.
- Muralidharan, K. and V. Sundararaman** (2008), "Teacher Incentives in Developing Countries: Experimental Evidence from India", Harvard University, Cambridge, MA.
- Murnane, R. and D. Cohen** (1985), "Merit Pay and the Nature of the Employment Relationship: Lessons from Public School Systems. Education Resources Information Center, Washington, D.C., http://www.eric.ed.gov/ERICDocs/data/ericdocs2sql/content_storage_01/0000019b/80/2f/4e/c7.pdf.
- Murnane, R. and D. Cohen** (1986), "Merit Pay and the Evaluation Problem: Why Most Merit Pay Plans Fail and Few Survive", *Harvard Education Review*, Vol. 56, pp. 1-17.
- Murphy, K.** (1999), "Executive Compensation", in *Handbook of Labor Economics*, Vol. III, Ashenfelter, O. and D. Card (eds.), North Holland, Oxford.
- Murphy, K.** (2001), "Performance Standards in Incentive Contracts", *Journal of Accounting and Economics*, Vol. 30, pp. 245-278.
- Murphy, T. and J. Margulies** (2004), "Performance Appraisals". Paper presented at ABA Employment and Law Section Conference, Equal Opportunity Employment Committee, Needham Heights, MA, March 24-27.
- Nalbantian, H. and A. Schotter** (1997), "Productivity Under Group Incentives: An Experimental Study", *The American Economic Review*, Vol. 87, No. 3, pp. 314-341.
- Neal, D.** (2009), "Designing Incentive Systems for Schools", in M.G. Springer (ed.), *Performance Incentives: Their Growing Impact on American K-12 Education*, Brookings Institution Press, Washington, D.C.
- Neal, D. and D. Schanzenbach** (forthcoming), "Left Behind By Design: Proficiency Counts and Test-Based Accountability", *Review of Economics and Statistics*.
- Odden, A. and C. Kelley** (1997), *Paying Teachers for What They Know and Do: New and Smarter Compensation Strategies to Improve Schools*, Corwin Press, Thousand Oaks, CA.

- Odden, A., G. Borman, and M. Fermanich** (2004), "Assessing Teacher, Classroom, and School Effects, Including Fiscal Effects", *Peabody Journal of Education*, Vol. 79, No. 4, pp. 4-32.
- Oyer, P.** (1998), "Fiscal Year Ends and Nonlinear Incentive Contracts: The Effect on Business Seasonality", *The Quarterly Journal of Economics*, Vol. 113, No. 1, pp. 149-185.
- Paarsch, H. and B. Shearer** (2000), "Piece Rates, Fixed Wages, and Incentive Effects: Statistical Evidence from Payroll Records", *International Economic Review*, Vol. 41, No. 1, pp. 59-92.
- Perry, J., T. Engbers and S. Jun** (2009), "Back to the Future? Performance-Related Pay, Empirical Research, and the Perils of Persistence", *Public Administration Review*, Vol. 69, No. 1, pp. 39-51.
- Perry, J. and L. Porter** (1982), "Factors Affecting the Context for Motivation in Public Organizations", *The Academy of Management Review*, Vol. 7, No. 1, pp. 89-98.
- Pfeffer, J.** (1995), *Competitive Advantage Through People: Unleashing the Power of the Work Force*, Harvard Business School Press, Boston, MA.
- Pfeffer, J. and N. Langston** (1993), "The Effect of Wage Dispersion on Satisfaction, Productivity, and Working Collaboratively: Evidence from College and University Faculty", *Administrative Science Quarterly*, Vol. 38, No. 2, pp. 382-407.
- Podgursky, M. and M. Springer** (2007), "Teacher Performance Pay: A Review", *Journal of Policy Analysis and Management*, Vol. 26, No. 4, pp. 909-949.
- Prendergast, C.** (1999), "The Provision of Incentives in Firms", *Journal of Economic Literature*, Vol. 37, No.1, pp. 7-63.
- Prendergast, C. and R. Topel** (1993), "Discretion and Bias in Performance Evaluation", *European Economic Review*, Vol. 37, No. 2-3, pp. 909-950.
- Rainey, H.** (1982), "Reward Preferences Among Public and Private Managers: In Search of the Service Ethic", *The American Review of Public Administration*, Vol. 16, No. 4, pp. 288-302.
- Rapple, B.** (1990). Payment by Educational Results: An Idea Whose Time Has Gone? Thirty-Five Years of Experimentation with Education Efficiency in England (1862-1897), *MacArthur-Spencer Series Number 14*, Chicago, IL.
- Reback, R.** (2008), "Teaching to the Rating: School Accountability and the Distribution of Student Achievement", *Journal of Public Economics*, Vol. 92, No. 5-6, pp. 1394-1415.
- Richardson, R.** (1999), "Performance Related Pay in Schools: An Assessment of The Green Papers", London School of Economics and Political Science, London.
- Rivkin, S., E. Hanushek and J. Kain** (2005), "Teachers, Schools, and Academic Achievement", *Econometrica*, Vol. 73, No. 2, pp. 417-458.
- Rosen, S.** (1986), "The Theory of Equalizing Differences", in O. Ashenfelter and R. Layard (eds.), *Handbook of Labor Economics*, Vol. 1, North-Holland, Oxford.
- Rosenholtz, S.** (1984), "Myths: Political Myths About Reforming Teaching", National Institute of Education, Washington, D.C.
- Rothstein, R.** (2009), "How Scholarship and Experience in Other Fields Inform Exploration of Performance Incentives in Education", in M.G. Springer (ed.), *Performance Incentives: Their Growing Impact on American K-12 Education*, Brookings Institution Press, Washington, D.C.
- Schacter, J. and Y. Thum** (2004), "Paying for High- and Low-Quality Teaching", *Economics of Education Review*, Vol. 23, No. 4, pp. 411-430.
- Schmitt, N., R. Noe and R. Gottschalk** (1986), "Using the Lens Model to Magnify Raters' Consistency, Matching, and Shared Bias", *Academy of Management Journal*, Vol. 29, pp. 130-139.
- Schneider, M., P. Teske and M. Marshall** (2000), *Choosing Schools: Consumer Choice and the Quality of American Schools*, Princeton University Press, Princeton, NJ.
- Skinner, B.** (1981), "Selection by Consequences", *Science*, Vol. 213, No. 4507, pp. 501-504.

- Springer, M.** (2008), "The Influence of an NCLB Accountability Plan on the Distribution of Student Test Score Gains," *Economics of Education Review*, Vol. 27, No. 5, pp. 556-563.
- Springer, M.** (2009), "Rethinking Teacher Compensation Policies: Why Now, Why Again?", in M.G. Springer (ed.), *Performance Incentives: Their Growing Impact on American K-12 Education*, Brookings Institution Press, Washington, D.C.
- Springer, M., D. Ballou and A. Peng** (2008), *Impact of the Teacher Advancement Program on Student Test Score Gains: Findings from an Independent Appraisal*, National Center on Performance Incentives, Nashville, TN.
- Springer, M., M. Podgursky, J. Lewis, M. Ehlert, T. Gronberg, L. Hamilton, D. Jansen, O. Lopez, A. Peng, B. Stecher and L. Taylor** (2008), *Texas Educator Excellence Grant (TEEG), Program: Year Two Evaluation Report*, National Center on Performance Incentives, Nashville, TN.
- Springer, M., M. Podgursky, J. Lewis, M. Ehlert, T. Gronberg, D. Jansen, O. Lopez, A. Peng, B. Stecher and L. Taylor** (2009), *Governor's Educator Excellence Grant (GEEG), Program: Year Two Evaluation Report*, National Center on Performance Incentives, Nashville, TN.
- Springer, M. and M. Winters** (2009), *New York City's School-Wide Bonus Pay Program: Early Evidence From a Randomized Trial*, National Center on Performance Incentives, Nashville, TN.
- Suen, H. and L. Yu** (2006), "Chronic Consequences of High-Stakes Testing? Lessons from the Chinese Civil Service Exam", *Comparative Education Review*, Vol. 50, No. 1, pp. 46-65.
- Taylor, L. and M. Springer** (2009), "Optimal Incentives for Public Sector Workers: The Case of Teacher-Designed Incentive Pay in Texas", *National Center on Performance Incentives Working Paper*, Nashville, TN.
- Taylor, L., M. Springer and M. Ehlert** (2009), "Characteristics and Determinants of Teacher-Designed Pay for Performance Plans: Evidence From Texas' Governor's Educator Excellence Grant Program", in M.G. Springer (ed.), *Performance Incentives: Their Growing Impact on American K-12 Education*, Brookings Institution Press, Washington, D.C.
- Tomlinson, H.** (2000), "Proposals for Performance Related Pay in English Schools", *School Leadership and Management*, Vol. 20, No. 3, pp. 281-298.
- Vroom, V.** (1964), *Work and Motivation*, Wiley, New York, NY.
- Weiss, A.** (1987), "Incentives and Worker Behavior: Some Evidence", in H. Nalbantian (ed.), *Incentives, Cooperation and Risk Sharing*, Rowan & Littlefield, Latham, MD.
- Weisberg, D., S. Sexton, J. Mulhern and D. Keeling** (2009), *The Widget Effect: Our National Failure to Acknowledge and Act on Differences in Teacher Effectiveness*, The New Teacher Project, Brooklyn, NY.
- Weitzman, M.** (1980), "The Ratchet Principle and Performance Incentives", *Bell Journal of Economics*, Vol. 11, No. 1, pp. 302-308.
- West, M. and P. Peterson** (2006), "The Efficacy of Choice Threats Within School Accountability Systems: Results from Legislatively Induced Experiments", *The Economic Journal*, Vol. 116, No. 510, pp. C46-C62.
- Wilson, W.** (1987), *The Truly Disadvantaged: The Inner City, the Underclass, and Public Policy*, University of Chicago Press, Chicago, IL.
- Wise, S.** (2004), "Work-life Balance and Careers in NHS Nursing and Midwifery", Employment Research Institute, Napier University, Edinburgh.
- Zenger, T.** (1992), "Why Do Employers Only Reward Extreme Performance? Examining the Relationship Among Performance, Pay, and Turnover", *Administrative Science Quarterly*, Vol. 37, No. 2, pp. 198-219.

CHAPTER 4

Whose Incentives? Whose Stimuli? Performance Management in England

Introduction	104
Incentives and stimuli: performance management and performance-related pay	104
Incentives and stimuli: teachers' continuing professional development	111
Incentives and stimuli: school leadership	115
Evaluation and assessment: the need to differentiate its purposes – a footnote	117
The need for dialogue between teacher organisations and governments on school improvement	118
References	119

INTRODUCTION

This chapter explores various forms of incentives and stimuli used by the Westminster Government in England for school and teacher improvement. It describes and evaluates the performance management and performance-related pay system introduced by the Westminster Government in 2000 and developed since then. It also explores other forms of incentives, including the major role played by continuing professional development and effective school leadership.

While individual performance-based incentives for teachers are the main focus of this chapter, there is a range of ways to motivate teachers to improve their practice. Motivation, however, cannot be imposed. It has to emerge from the embedded confidence of teachers themselves. Incentives can provide the conditions for change in motivation, or they can be imposed and have little effect on teacher motivation to improve.

Without teachers and the teaching profession, little or no education is possible. It has been the tendency of some governments to see teachers and their organisations as a problem and as a barrier to achieving high standards. A characteristic of this approach is to criticise the teaching profession as a whole, for example, as being content only with low standards and low expectations; as a profession that can only change through the use of incentives and threats; and as a profession which will only focus on those students who have the capacity to achieve.

Generalisations about teachers are seldom helpful. Such generalisations cast those who are prepared to learn, leading and effective innovators, and at the leading edge of their pedagogy, in the same light as those who resist change. While such teachers in this last category are usually in a small minority, crude generalisations about the teaching profession only serve to alienate those teachers who want change and who are willing to lead it. Change is rarely effective in the long term if it is imposed. Rather, governments should provide the conditions for change if they want reforms to be embedded.

Praise and recognition of achievement and recognition of specialist expertise can lead to significant improvement in teachers' motivation. In addition, opportunities for horizontal career development, based on growing expertise also lead to improved motivation. For example, through its own professional development programme provided for its members, the National Union of Teachers (NUT) in **England** provides opportunities for teachers to become researchers and to investigate the best forms of practice within classrooms. Indeed, high-quality teacher professional development is a significant motivator in itself, and is also a very important factor in the retention of experienced teachers who would otherwise consider leaving the profession.

One particular example of a strategy to recognise experience and quality in teachers is the idea of a chartered-teacher status. The post-McCrone agreement in **Scotland** offers teachers the opportunity of supplementing their knowledge and experience with high-quality professional development and, with the aid of university tuition, with achieving chartered teacher status. Becoming a chartered teacher leads to a significant pay increase, a concept that has been accepted and welcomed by Scottish teachers.

Not all aspects of motivation are covered in this chapter. Being treated fairly and justly and knowing that your opinions and judgements will be responded to with the respect they deserve are prime sources of motivation for teachers, as are having a fair pay/reward system, a fair and open appointments-and-promotion system, and supportive conditions of service. These areas are not covered here. However, they are equally important, and their importance is consistently rehearsed by unions for the benefit of governments and employers in many countries, as is the need for dialogue between unions, employers and governments.

INCENTIVES AND STIMULI: PERFORMANCE MANAGEMENT AND PERFORMANCE-RELATED PAY

The vast majority of teachers enter the profession because they want to make a positive difference to young people's lives. Teaching is a vocation. Mechanisms for performance management and incentives for teachers to

improve have to be informed by this understanding if they are going to be effective. However, the effectiveness of a performance management scheme and its incentives is subordinated to the effectiveness of good leadership, which is explored later in this chapter. This is not to say that the experience of performance management is not worth exploring. Its development in **England** is both informative and instructive.

Performance management in England has gone through several phases. Its initial iteration was described as appraisal, and focused on identifying and securing acknowledgement by the teachers being appraised that there were areas of their teaching which needed developing and improving (DfEE, 2001). New sets of regulations in England, issued in 2000 and in 2006, have progressively formalised appraisal, and it is now known as performance management. The introduction of performance management in schools was not accompanied by additional funding other than for training teachers to conduct the performance management.

The initial performance management framework introduced by the Government had little connection with a parallel scheme, introduced at the same time, for rewarding experienced teachers. The framework was introduced solely as a mechanism for school improvement. The major benefits the Government envisaged for its 2001 scheme were that:

- “[students would] benefit because their teachers [would] have a more sharply focused picture of what, with encouragement, support and high expectations, their [students could] achieve; and
- teachers [would] benefit... teachers [would] have the right to expect that their performance [would] be regularly assessed and that they [would] have a proper opportunity for professional discussion with their team leader about their work and their professional development,” (DfEE, 2001).

Crucial to the performance management framework was a set of reassurances in relation to the purposes of performance management, which the NUT and other unions managed to achieve. The Government characterised existing best practice in schools as:

- a commitment to the attainment and welfare of [students] at their school;
- an appreciation of the crucial role that teachers play;
- an atmosphere of trust between teacher and team leader, which allows them to evaluate strengths and identify areas for development;
- encouragement to share good practice; and
- the integration of performance management with the overall approach to managing the school.

The scheme involved the introduction of annual cycles for performance management. A team leader, usually identified by the principal (referred to as the head teacher in the **U.K.**), was expected to agree on objectives for professional development and for student achievement outcomes within an individual plan for the teacher whose performance was being reviewed. Within the plan, team leaders would discuss and record priorities and objectives with each of the teachers in their team. They would also discuss how progress was to be monitored.

During the annual cycle, monitoring by the teacher and team leader was expected to take place as was any supportive action needed.

At the end of the cycle, both teacher and team leader were expected to review achievements over the year and evaluate the teacher’s overall performance, taking account of progress toward objectives.

In terms of the number of objectives, the Government indicated that teacher reviewers should identify two core objectives for the performance management process. One would focus on student progress based on readily available information about the students. The other objective related to professional development.

At this stage, teacher organisations in **England** were very concerned that principals would simply see the student progress objectives as a way of setting target percentages for students' test and examination results. As the scheme was rolled out, it became clear that very few schools adopted the option of setting objectives for test and examination results. Teacher organisation guidance encouraged the setting of student progress objectives that reflected the broad range of student achievement, rather than examination and test results.

Only a minority of schools fully implemented the Government's initial performance management framework, although the majority adapted elements of it. NUT internal surveys showed that teachers felt that having a performance review which involved a genuine and honest discussion about strengths and weaknesses with their reviewers tended to be helpful. It was only when reviewers attempted to link performance management with pay decisions that the trust and openness in the performance review discussions tended to evaporate. Most principals, the NUT found, did not themselves wish to link the performance review process with pay decisions.

In 1998, just prior to the introduction of performance management arrangements, the Government introduced a scheme, which became known as the "threshold arrangements". This involved the publication of a "Green Paper" (a Government consultative document) setting out Government proposals entitled "Teachers: Meeting the Challenge of Change" (DfEE, 1998).

The Green Paper's diagnosis started from the position that the Government had already put in place a series of changes that raised the prospect of a world-class education service for all children. The changes, the Paper said, would work properly only, "if we attract and motivate teachers and other staff with the ambition, incentives, training and support to exploit this opportunity". This, in turn, was seen to be achievable through a combination of steps involving:

1. strengthening school leadership by rewarding leaders properly;
2. recruiting, retaining and motivating high quality classroom teachers, by paying them more; and
3. providing better support for all teachers and employing teaching resources in a more flexible way.

The Green Paper noted that recruitment to teacher training had become increasingly difficult. It also noted that, "The profession, as a whole, does not attract enough ambitious young people or those in later life looking for a career change" (DfEE, 1998). This was because, "The present reality of teaching too often compares unfavourably with the growing range of alternative careers for successful graduates" (DfEE, 1998).

The strong implication of the Green Paper's analysis was that this situation existed substantially because "Good teacher performance is not sufficiently recognised". In a key paragraph, the Green Paper observed that even though "The present pay system allows for the award of extra pay for excellent performance, less than one per cent of teachers have ever benefited". This was said to be "Because such recognition was not part of their culture Few professions have turned their back on linking pay to performance to the same extent as teaching The tradition in teaching is to treat all teachers as if their performance was similar, even though in every staffroom teachers themselves know this is not true". (DfEE, 1998)

At the heart of the Green Paper, therefore, was a presumption that the existing culture on pay within education seriously inhibited the performance of teachers. It was seen to undercut teachers' motivation and give them insufficient incentives to develop.

The Green Paper proposed, therefore, that a pay structure recognising, in pay terms, the performance of experienced teachers should be developed. It involved a pay structure that maintained the main pay scale for classroom teachers and an upper pay scale for which teachers could apply when they had completed a minimum of five years teaching. The quantum of teachers' pay was increased to take account of the cost of the new arrangements. The pay structure is described below.

When newly-qualified teachers start teaching they are paid on the main pay scale and they are expected to move through a range of six increments, with each increment yielding an annual increase in salary of around GBP 2 000 a year. After the sixth increment, teachers have the opportunity to apply for the upper pay scale. That opportunity was, and still is, known as “crossing the threshold”.

Between 2001 and 2008, teachers were expected to submit a portfolio of evidence to their principals in order to justify their application to move on to the upper scale. Teachers were expected to gather evidence of the positive impact of their work on students’ learning, and that evidence was to be informed by eight national standards as set out in the Government’s legally binding School Teachers’ Pay and Conditions Document (DCSF 2008) – a nationally-determined substitute for a teacher’s contract.

For some years, principals’ decisions on whether or not teachers should cross the threshold were monitored by a Government-appointed external contractor. Teachers themselves decided whether or not to use the results of their performance management reviews in their applications to cross the threshold.

When both the performance management and threshold schemes were first introduced, teacher organisations were highly sceptical about their impact on teachers. It became clear, however, that the vast majority of schools were putting in place arrangements that were acceptable to teachers. Teachers continued to be concerned, however, about the amount of bureaucracy involved in terms of gathering evidence and also concerned about the potential for unlawful discrimination.

The advantage of having an external contractor moderating teacher applications was that the Government was able to acquire data, at a national level, on whether there was unlawful discrimination in the school system. The external contractor did find evidence of unlawful discrimination on the grounds of race. There is currently no central collection of evidence tracking movement of teachers to the upper scale and whether or not discrimination continues to occur since, after five years, the Government decided to drop the external contractor responsible for moderation on the grounds of efficiency and cost. In 2010, it will introduce a national database, which will have the ability to track trends in teacher employment nationally, including any discriminatory trends.

In 2006, the Government introduced a new form of performance management. The purpose of the new performance management scheme, enshrined in statutory regulations, was to merge decisions about moving teachers to the upper pay scale with the performance management process.

Much of the original 2001 scheme was retained. Schools were still required to formulate performance management policies. The scheme was still carried out annually. There were, however, differences between the two systems. The governing body and the principal were given the responsibility of implementing the performance management policy and the policy itself was subject to specific statutory requirements.

Every school, from 2006 onwards, has been expected to have a policy which:

- states what results the policy is intended to achieve and how these will be measured;
- shows how the school’s arrangements for school-teacher performance management links with those for school improvement, school self-evaluation and school development planning;
- shows how the school would seek to achieve consistency of treatment and fairness between those teachers with similar experience or levels of responsibility;
- sets out the timing of the cycle;
- includes a classroom observation protocol;

- provides for performance-management training to be made available as the need arises;
- states the arrangements for monitoring and evaluating the policies; and
- specifies any ancillary or supplementary procedures necessary for the operation of the performance management of teachers at the school (DCSF, 2008).

Principals are responsible for the performance management of every member of staff in the new scheme.

Principals, from 2006 onwards, have been expected to identify teacher reviewers of other staff. Those reviewers were now required to take into account, when reviewing a teacher's performance, that teacher's:

- job description;
- any relevant pay progression criteria;
- any relevant school or team objectives specified in the school improvement plan;
- what could [reasonably be expected of any teacher], given the desirability of the reviewee to ... achieve a satisfactory balance between the time required to discharge his or her professional duties and the time required to pursue his or her personal interests outside work;
- the reviewee's professional aspirations; and
- the relevant professional standards.

The 2006 scheme also introduced a new performance management requirement. Each reviewer was expected to carry out classroom observation of the teacher he or she was reviewing. The regulations made it clear that such classroom observation would not exceed three hours annually and its timing would have to take into account the individual circumstances of the teacher.

In 2006, DCSF (2008) published the School Teachers' Pay and Conditions Document, a framework of professional standards for Excellent Teachers on the upper scale, and for Advanced Skills Teachers (a separate merit-pay scheme for teachers who are expected to help teachers in schools other than their own). The eight generic standards that were used prior to 2006 to assess teachers applying to cross the threshold to the upper scale were replaced with some forty core standards for all teachers, and an additional nine post-threshold standards for teachers wanting to cross the threshold.

A subsequent amendment to the Schoolteachers' Pay and Conditions Document removed the requirement for teachers crossing the threshold and applying for the upper scale to provide new evidence to their principal, but instead noted that principals were to use the results of the teachers' last two annual performance reviews as the basis for decision-making on teacher movement to the upper scale.

It has thus taken schools in England almost a decade to move to individual performance-related pay. What has to be asked is whether performance management and the new pay scale based on classroom teachers' performance have had any effect on teachers' self-efficacy, motivation, skills and knowledge. The NUT, at the beginning of the decade, also had very great concerns that performance-related pay would simply mean that teachers would spend the bulk of their time focusing on trying to secure a narrow range of test and examination results and thus reduce the time needed to foster other areas of student learning and well-being.

In the first half of the decade, it became clear that the original performance management arrangements had had a moderately motivating effect on a minority of teachers. The majority of teachers felt that this was just another process that they had to go through, and in fact the processes became absorbed into the annual routine of school life. The same attitude applied to application for the threshold.

When the Government first decided to introduce its new performance-related pay scale, the NUT commissioned Professor Ray Richardson of the University of London, London School of Economics, to evaluate the Government's proposed scheme. Richardson (1999) evaluated the merits of the proposals contained in the Green Paper, comparing it with other public service schemes. His study included a review of performance-related pay in the Government's Inland Revenue Department, local governments, the National Health Service, the water industry, and the Government's employment service. He examined a range of studies on those areas and concluded that none of them provided strong support for the Government's proposals for performance-related pay for teachers.

To quote Richardson:

None suggest that it would have a dramatic effect on teacher motivation. Expressed enthusiasm amongst public sector employees who have experienced performance related pay tends to be modest. Rarely do more than 20% of the respondents agree that it has given them an incentive to change any specified form of their behaviour at work. Within even that total, the number who "strongly agree" (with performance related pay) is usually very low and is easily swamped by those who express strong disagreement. In this sense, real antipathy amongst staff is more common than real enthusiasm. Nor are our line managers markedly more enthusiastic, except in one study. In addition, respondents are usually of the view that these schemes seriously threaten team working and co-operation. Hence the views of researchers that individual performance related pay in the public sector has been, variously, counter-productive, a "damp squib", occasionally, a very modest success.

This does not mean that performance related pay would necessarily be an outright failure for teachers, although it must be said that, on the basis of Marsden and French's study of head teachers (Marsden and French, 1998) the ground in education looks to be more than usually stony. On the other hand, if performance related pay were to energise even 20% of teachers, it might on balance be considered a success, especially if there was only limited de-motivation amongst the remaining 80%. In other words, it might be deemed, if not a dramatic success, then at least a modest one.

He then turned his attention to the initial Government proposals for performance-related pay:

There will have to be a significant training programme for reviewers.... [T]his will be very important, because performance review is not an easy task, whether in setting the objectives in an appropriate way or in subsequently judging whether those objectives have been met. The process also requires the reviewer to provide useful feedback during the year, which requires frequent monitoring of teacher performance in a way which is not seen as intrusive or unsympathetic. More than training is needed, however. Proper appraisals use resources, most obviously, the time of those engaged in the process. ...This is not a trivial problem. The standing of a performance related pay regime depends in good part on its integrity. This in turn depends not on general principles but on the regime's detailed application. Without proper resourcing, the performance review process will all too often be skimped, and the loss of teacher motivation is likely to be the result.

... There must be a suspicion that the Government is really using performance related pay for a somewhat different purpose than merely recruitment, retention and motivation. Teachers' pay has not grown as fast as the pay of many other employees. It is possible that the new proposals are really a way of rewarding a selective minority of teachers without incurring the wage bill cost of extending the benefits across the board. In other words, it may really be a device to attract in and keep in the profession a number of relative high quality teachers while doing rather little for the large number of average performers.

(Richardson, 1999)

A decade later, has the teachers' performance related pay scheme met the original Government objectives of improving teacher recruitment, retention and motivation? Now that Government intentions introduced in 2006 have been finally realised through new performance management regulations - linking performance to pay - how do teachers feel about this process?

The NUT's approach to both the original, separate performance management arrangements and threshold and the new combined arrangements has been to give advice to its members in schools on the best way of maintaining professional relationships among all teachers in each school while at the same time ensuring equity of access to the upper scale with the fewest possible administrative burdens. The NUT guidance has emphasised that the regulations neither define nor limit the freedom of the school to adapt statutory requirements.

An example of the NUT's approach is that it has sought agreement with schools to combine classroom observations for performance management purposes and for those conducted in line with other statutory requirements, such as visits by external advisers and principal monitoring. It has also sought to establish a reasonable amount of time between each classroom observation.

It is not yet clear how the new phase of performance-related pay is going to work out. The NUT recently conducted a survey of its members on the impact of the new performance management regulations (NUT, 2008). It showed that the majority of schools had adopted sensible performance management policies in line with the NUT's own guidance. The survey indicated teachers' continuing objection to specific student progress objectives being used within teacher performance reviews. Around 80% of all teachers, including principals, opposed the idea of their pay progression to the upper scale depending on students achieving a specific percentage of National Curriculum test results or public examination grades.

Perhaps more importantly, the members' comments made about the whole process revealed a range of reactions (NUT, 2008). For example, a primary classroom teacher was typical among those who were positive about performance management when she said:

Performance management in my school is fair and done with good consultation... During the year we have the opportunity at any time to discuss the progress of our objectives and adjust them if it is necessary.

One secondary teacher's comment was typical of those who were more sceptical:

No one cares if I achieve my professional development objectives – only reviewed or discussed yearly – seems just a paper exercise. Not had any training or help to achieve any of them ever. I have my own targets, which don't seem ever to fit into what the school or the department wants. There is no incentive for me.

Some teachers were concerned about progress that is not easily measurable. A primary teacher said:

You can only be accountable for what and how you teach – not what your pupils learn. I regularly work with pupils who do not make measurable improvements with reading and spelling scores but who make considerable progress with confidence, concentration and improved attitude to learning.

In contrast to this mixed response, teachers were very critical of objectives set by the school or the Government.

A primary teacher said:

One of my objectives is marked as partially achieved because one pupil with emotional and behavioural issues did not make two sub-levels of progress. The fact that other pupils achieved those levels didn't count. This pressure will only encourage teachers to lie about levels in order to pass performance management targets.

Another typical response came from a secondary teacher:

We are told every child must make a two-level progression at each key stage, anything less and we have failed! I feel that it is very unfair that we are judged purely on external examination results! Our partner school, which has a completely different catchment area and ethnic mix, received “outstanding” [from an external inspection] based on their results.

Evidence from NUT members, such as from the survey referred to above, has shown that where space, time and training are made available for teachers, the process of gathering of evidence and its interrogation by both reviewer and reviewee, allow the reviewee to reflect on his or her achievement. The most experienced teachers who apply for pay progression are successful.

However, the NUT’s experience is that the effect of such a process on teacher motivation is short-term and is also confined to key points in the process at the beginning and at the end of a performance management cycle. The reality is that any positive effect of performance management is not continuous and, for many principals, performance management is peripheral.

Effectively implemented, performance management will identify learning and professional development for the teacher under review. The question many principals and teachers ask is, why such a highly formalised set of arrangements is needed to identify professional development needs? The answer may well be to secure consistency, equity and fairness in the process, but the question remains as to whether the performance management process itself has a direct positive impact on teacher quality and, indeed, whether performance linked to pay determination has a similar effect.

It can be argued that performance management and performance-related pay as forms of reward to teachers have largely fulfilled the predictions of NUT’s 1998 London School for Economics-commissioned research (Richardson, 1999) – that there is no evidence of performance-related pay having a powerful, positive impact on teacher recruitment, retention and motivation. It does seem, therefore, that school improvers and reformers need to turn elsewhere if they want to make a deep, sustained and powerful impact on raising the motivation of teachers and on identifying and sharing effective teaching practices.

INCENTIVES AND STIMULI: TEACHERS’ CONTINUING PROFESSIONAL DEVELOPMENT

The evidence available to the NUT points to teachers’ enthusiasm for high quality professional development. Since 2000, the NUT has run its own professional development programme, not only for its members, but for non-members as well. Participating teachers’ evaluations of the NUT’s programme over the last nine years emphasise the motivating impact of high-quality continuing professional development (CPD). The impact of high-quality, interactive and sustained professional development is two-fold. Not only does it encourage teachers’ learning, but it also encourages more experienced teachers to remain in the profession.

The NUT has based its own professional development programme on the findings of a key Evidence for Policy and Practice Information (EPPI) research review (Cordingley *et al.* 2003) which the NUT financed (NUT has also sponsored a number of other EPPI research reviews). The results of the EPPI review show not only the positive motivational effect of CPD on teachers’ motivation, but also its impact on improving children’s and young people’s learning outcomes.

Yet, there is no consistent access to continuing professional development for all teachers in the English education system. The Government, however, along with teacher unions, is currently addressing this issue, debating whether to include a contractual entitlement within the revised professional development strategy scheduled for launch in 2009.

There are a number of impacts arising from inequity of access to CPD. Much CPD takes place outside the school day. Carers for children and those with other important commitments are naturally disadvantaged by this. Teachers on short-term contracts, those who are employed by local authorities and those who have recently returned to teaching all find it difficult to access relevant CPD provision because they are not attached to a particular school.

There remains an imbalance between the emphasis given to national initiatives and teachers' ability to choose their own methods of professional development. There is also an imbalance between the availability of high-cost, post-graduate award-bearing courses and courses for improving classroom practice that focus on solving specific problems.

The NUT believes that for any government, the introduction of a professional development strategy is vital for energising and motivating teachers, and for raising performance against standards. From the evidence of its own programme and from deep involvement in research about effective forms of professional development, the NUT has developed a number of principles which are reflected in its own CPD programme and its wider education policy.

A strategic approach must be taken to embedding CPD in the culture of schools, if it is to be seen as integral to the teaching profession and not as an expendable addition. In order for teachers to maintain the highest standards and meet the expectations of parents and young people, continuing professional development must be part of each teacher's contractual entitlement.

The NUT's proposals and its policy statement, "Bringing Down the Barriers" (NUT, 2004) focused on ensuring that all teachers have a material entitlement to professional development. There are two essential minimum components within the NUT's proposals. Teachers should receive a minimum funded entitlement annually for personal professional development. They should also receive a one-term sabbatical every seven years, to conduct their own research on effective classroom practice.

In this context, it is worth referring to the OECD's recent study, *Teachers Matter: Attracting, Developing and Retaining Effective Teachers* (OECD, 2005). Within that study the OECD identified three broad strategies for integrating professional development throughout teachers' careers. They should be entitlement-based, incentive-based, and school-based.

The OECD's conclusion is important in this context:

A comprehensive approach to professional development would encompass all three strategies. Providing teachers with agreed levels of time release or financial support for professional development is an explicit recognition of its importance in teachers' work and a means of enabling participation. However, it is also important for teachers to see the value of taking part in professional development, to understand that it is an important part of their role and to see the "entitlement" to provision as a minimum extent of their participation rather than the maximum (OECD, 2005).

Incentives for teachers to take-up professional development would, therefore, include both improvements in access and in the quality and relevance of the provision on offer. This would not only make teachers feel valued by their schools and the wider education system, but it would also enable schools to take advantage of a range of professional development opportunities, similar to those found in comparable employment sectors. Any government's approach to CPD should reflect and be guided by the responses of the teaching profession, individually and collectively.

In **England**, the Westminster Government and local authorities have often delegated their own responsibilities to individual officials who take the approach known as "cascade training". This involves one member of staff

attending a course, which is run, by someone following a pre-ordained script. Participants are then expected to deliver the same training, often using the same script, to their colleagues at schools. Such approaches inevitably lead to a superficial understanding of the professional activity concerned, leading at best to teacher cynicism and at worst to a belief that CPD is valueless.

The NUT believes that it would be more beneficial, for both individual teachers and for the education system as a whole, for more attention to be given to identifying and encouraging the sharing of schools' good practice in CPD and disseminating findings to other schools. NUT believes that it is vital that schools be encouraged and given the time to establish professional communities of teachers so that good practice can be shared.

Since 1997, one of the most unnoticed but most positive successes of the Labour Government has been the establishment of the Union Learning Fund and Union Learning Representatives. The Government has allocated sums of money to unions to help provide for their members' learning, and this has led to very many workers, including teachers, experiencing a change in their learning and confidence.

Through the Union Learning Fund, the NUT has been able to organise, in partnership with local providers, basic information communications technology (ICT) training for all teachers. 16 members had expressed a need for such courses. To date, over 14 500 teachers have participated, and the courses have grown in scope to include training on multimedia technology and interactive whiteboards. The success of this programme demonstrates that teachers' professional associations have an important part to play in identifying the kinds of training that will have resonance and credibility with teachers, as well as providing support and help to those applying for grants or conducting research.

Research and evaluation of provision and its impact on teaching and learning are essential to providing evidence on what works in CPD. The NUT's sponsorship of a number of Evidence for Policy and Practice Information (EPPI) research reviews has been mentioned earlier. By drawing on the EPPI findings and other national and international research and through independent evaluation, the NUT has developed peer coaching and mentoring as a professional development tool. Its approaches to teacher-training for literacy and numeracy education have shaped the widespread use of coaching and mentoring in schools across **England** and **Wales**, and have been adopted by the Westminster Government's National Strategies and by the Government's Training and Development Agency.

Such a research-based approach has led to the development of a number of NUT's models for professional learning within its CPD programmes. They include:

1. **Peer Coaching:** Participating teachers, usually in pairs, observe each other teaching and then give each other feedback. The vital element of this approach is that both participating teachers act as coach and coachee.
2. **Best Practice Research Scholarships (BPRS):** This model of professional learning attracted Government funding and provided an opportunity for teachers to investigate an aspect of teaching and learning of their choice. BPRS scholarships were given when teachers applied in pairs and committed themselves to carrying out their learning as partners. The BPRS scholarships were ended, unfortunately, in 2003.
3. **Learning Circles:** In partnership with the University of Cambridge, School of Education, the NUT-CPD Programme has developed a *Learning Circle* approach to professional learning. A group of teachers from a local area meets with the University of Cambridge tutor and learns about good practice in carrying out school/classroom-focused investigations into an aspect of teaching and learning of their choice. Portfolios of their findings are submitted to the University of Cambridge for post-graduate accreditation.
4. **Study Groups:** In partnership with the University of Cumbria, the NUT-CPD Programme has developed a collaborative approach to professional learning, which enables participating teachers to complete a module worth 20 credits towards a Master of Arts in Education.

Recently, the Government has introduced a new master's level qualification in teaching and learning (MTL) for newly-qualified teachers. The MTL is based on the master's degree for teacher training developed in **Finland**. The NUT has given this a qualified welcome, but has expressed concern that newly qualified teachers may not have the time and capacity to take on yet another qualification in their first years of teaching. Rather, it believes that the new master's degree should be a voluntary entitlement for all teachers and should be accompanied by a guaranteed entitlement to a tutor and tutorial time during the school day.

There are other developments in CPD that the NUT believes have potential for raising teachers' motivation. It has supported the development by the General Teaching Council for England (GTCE), at its Teacher Learning Academy (TLA). The TLA provides accreditation to teachers' CPD and serves as an umbrella for particular accreditation arrangements between higher-education institutions and other providers. Unlike a formal master's degree, the TLA can be applied to short courses which do not carry accreditation.

Teachers must have ownership of their CPD, and have a wide range of courses available to meet individual needs. Recently, professional development has been predominantly centrally driven and prescribed. CPD determined by teachers' individual needs is currently dependent, to a large degree, on the culture and circumstances of the school they happen to be working in.

Teachers' prime concerns and motivation for participating in CPD are improving their subject knowledge and pedagogical skills. This makes a continuing equity of focus across different subject areas in the national CPD priorities unsustainable. Even a recent Government report found that:

Teachers of the arts and more generally teachers of foundation subjects take the view that the sustained emphasis on literacy, numeracy and management has in effect cut them off from award bearing courses (Soulsby and Swain, 2003).

The greatest impact on changing teachers' practice comes from learning from and with other teachers, within their own schools and with other schools, through observation, collaborative work and peer coaching and mentoring schemes. Training needs to be focused on specific skill areas with high-quality materials and direct support, so that learning can be applied to classroom situations. The banners between national, local and individual CPD priorities can be addressed through imaginative partnerships between schools, local authorities or school boards, higher-education institutions and other CPD providers, including teachers' professional organisations.

There is every argument for a government to establish a quality assurance code of practice for teachers' continuing professional development. Recently, in **England**, the Training and Development Agency set up a national CPD database, which the NUT has welcomed. Currently, the CPD market is very complex and fragmented. A central database with the ability for users to actually comment on the quality of programmes and courses should contribute to consistency in terms of quality. Much existing CPD provision consists of one-off events or short courses, which are of variable quality and relevance.

Teachers need to have power and autonomy over the direction of their work and this now needs to be recognised and underpinned by changes in the attitudes of policy makers to CPD. The enormous learning potential available to teachers through teacher exchanges, sabbaticals and scholarships, both at home and abroad, has barely been tapped. The NUT believes strongly that a national Professional Development Strategy which, at its core, provides a range of opportunities for teachers to work on projects in developing countries and encourages knowledge, research and skill exchanges between schools in developed and developing countries, has a powerful effect on motivation and retention of more experienced teachers. Teachers have to work at becoming and remaining good teachers and, therefore, need to keep up-to-date with the latest developments in teaching and, increasingly, information and communications technology.

It is the learning involved in effective continuing professional development, not some artificial requirement to add to the initial teacher-training qualification through a further certificate, that is the important thing. In **England**, the resources to enable teachers to achieve this, however, have yet to be put in place:

Nearly all of the teachers interviewed during this research expressed frustration at the limited opportunities they had to reflect on and develop their own practice, which they saw as a result of excessive workload. There was also dissatisfaction with an increasingly centralised, standardised system of training and guidance. (Horne, 2001)

There are strong arguments for developing and identifying pedagogic practice for teachers' learning in the same sense that there are arguments for effective pedagogic practice for children. A study by Cordingley (2008) makes it clear that governments, local authorities and schools are often still at "first base" when it comes to identifying the best ways to teach teachers. The benefits of sophisticated peer coaching and mentoring have still yet to be explored.

Through effective professional development, teachers' knowledge, skills, motivation and self-efficacy can be enhanced for the benefit of children's and young peoples' learning. Too often, a secure evidence-base and learning about what works in the classroom are considered peripheral. Governments may introduce pedagogic programmes for teachers, but teachers understanding of themselves and the importance and relevance of CPD to their own practice, is the prime trigger for motivation and improvement.

INCENTIVES AND STIMULI: SCHOOL LEADERSHIP

One of the most important stimuli for effective teachers and effective teaching comes from school leadership. Much has been written about the nature of school leadership and its vital importance to school improvement, as well as the potential shortage of future leaders. The current and future projected shortage of school leaders in many countries is not inevitable. Virtually every principal is able to spot a future leader or leaders within his or her school. The issue that has to be addressed is whether these potential leaders wish to become principals. Succession planning, which identifies and trains future principals, must, the NUT believes, include the training and identification of middle-and senior-level leaders. Planning for succession has to encompass every leader in the school. Indeed, there is every argument for exploring the concept of "leadership for learning". All teachers should have the capacity to lead innovation in their particular fields of teaching and learning. By encouraging teachers to take pedagogical initiative, the possibility of identifying new, future leaders at the classroom level becomes more likely. The barriers that prevent teachers from aspiring to leadership have to be removed.

The NUT believes that all current and future government initiatives should be subject to an audit, carried out in consultation with principals, of the capacity of schools to implement them. Principals can often feel isolated. It can sometimes seem to principals that they are solely responsible for implementing every single external initiative and they are solely accountable for the success of those initiatives. Principals therefore need the opportunity to develop "critical friendships". That does not mean that such "critical friends" should have a dual loyalty to others responsible for making schools accountable. "Critical friends" should be exactly what the term implies: colleagues of a similar status who not only provide sympathy, empathy and support to principals, but who are also a source of impartial and knowledgeable advice.

In **England**, the role of principal is seen by classroom teachers and middle-level leaders as involving high-risk and high-stakes. The Education and Inspections Act of 2006 gives local authorities new, strict powers of intervention. In addition the English inspection agency, OFSTED, has new powers to intervene on the grounds of parental complaint. Fear of their schools being deemed as failing is a powerful disincentive to teachers who are considering applying for National Professional Qualification for Headship (NPQH), which is mandatory for

those in principal posts. Concern about being responsible for a failing school restricts initiative and innovation as principals may fear making mistakes. It also leads many principals in England to retire early.

Principals need to be confident that leadership is not a form of tightrope walking. Principals should not be expected to leave if their schools are given warnings about the need for school improvements. If such warnings are necessary, then principals should be offered the support they need within their schools, including, where appropriate, peers or advisors of their choice. The provision of supportive professional development, not punishment, should be the first response of external authorities who are responsible for securing that school's improvements.

Removing unreasonable risk associated with the job is only part of the solution to the potential future shortage of principals. It should be recognised that being a principal can involve many decades of service. Principals should be entitled to sabbaticals once every seven years, when they can take up advisory posts or gain experience in administering local authority services. The NUT believes that a pool of advisory principals could also serve as mentors for peers who need them.

The issue of a potential future shortage of school leaders in many countries was the trigger for the OECD's report, *Improving School Leadership* (Pont *et al.*, 2008a, 2008b). As the report notes, English schools are not alone in facing potential school leadership shortages. A key theme of the OECD report is identification of why and how school leadership is vital to teacher motivation and the effective learning and teaching of young people.

No school can be effective if it is led by an ineffective principal. Many teachers can survive ineffective leadership within their schools, but their work will be often carried out in a fragmented way. Effective leadership provides coherence and purpose and is the most important source of incentives, stimuli and motivation for teachers in the classroom and middle leaders.

The 2008 OECD report proposes a set of strategies countries may wish to adopt. They include identifying core responsibilities and redefining leadership. The report argues that principals need time and capacity to promote student learning. Rightly, the report focuses on goal-setting, developing teacher quality, and on developing collaboration between schools. It also proposes treating leadership as a continuum, starting from initial leadership training, for consistency between providers of leadership training, professional recruitment of leaders, and succession planning (*i.e.* identification of future leaders). The OECD report refrains from making any recommendation that countries should experiment with appointing school principals who are not also qualified teachers. For the NUT, evidence shows that the most effective principals are pedagogic leaders with a deep background in teaching, as well as being effective managers.

While the report rightly focuses on distributing school leadership, it does not fully make clear that distributed school leadership should apply to all teachers, as indicated above. The vast majority of teachers have the capacity to lead in their particular field, and leadership for learning should be synonymous with distributed leadership. If school leaders are to be effective they should focus on encouraging classroom teachers' professional self-confidence and motivation to lead particular aspects of learning and curriculum amongst their colleagues.

Recently the NUT embarked on a joint collaboration with the National College for School Leadership in **England**. Both organisations commissioned a report from the University of Manchester to examine the facets of effective leadership for promoting the education of white working-class students. This report arose from the NUT's and the NCSL's concern that a number of student groups in England were showing achievement outcomes that were far lower than those of other groups. Indeed this was the second in a series of such investigations conducted by the NUT. The first was the development of a charter, "Born to be Great", aimed at improving the achievement of black Caribbean boys, another low achieving group (NUT, 2007).

Although the key aims of the research were to focus on working-class achievement, the research review and action research evidence collection also revealed characteristics of effective leadership that apply to all schools in all circumstances. These are listed in a taxonomy of features of effective distributed leadership. Important features included levels of contextual, emotional and social intelligence, and leaders' understanding and development of people.

Teacher organisations across the world have sought to identify the characteristics of effective leadership. One very good example of such an approach was set out by the school leaders' union in Skolelederne, **Denmark** (Balle, 2009). Skolelederne also highlights the importance of leaders needing both to be empathetic (*i.e.* able to understand human relationships) and to inspire confidence and interaction.

All governments need to devote much more attention to developing and enhancing effective school leadership. Since effective and distributed leadership within schools has a powerful effect on teachers' commitment and willingness to improve their practice, the NUT believes that governments must now also focus on developing their leaders in order to develop effective education systems.

EVALUATION AND ASSESSMENT: THE NEED TO DIFFERENTIATE ITS PURPOSES – A FOOTNOTE

The evaluation of schools, as institutions, by governments and/or employers triggers a range of responses from teachers depending upon the approach. The evaluations can be based on "bottom-up" school self-evaluation, as in **Finland**, which is relatively non-controversial, or based on highly controversial "top-down" external evaluations, as in **England**.

The impact of external institutional evaluation on teachers' morale and motivation can be profound. If a school is deemed to be failing, then that sense of failure can affect detrimentally the self-efficacy of teachers. If the evaluation system has "high-stakes" (that is the future of schools and institutional evaluation is based solely on summative judgements), and does not include solutions for school improvement that school communities can see as relevant and effective, then the impact on teacher motivation and, indeed, retention in the profession, can be negative.

The results of a National Foundation for Education Research study (Scanlon, 1999) commissioned by the NUT demonstrates how demoralising the designation of failure or a "special measures" judgement (to use the terminology of the Westminster Government's Inspections Department, OFSTED) can be to teachers' motivation.

The purpose of this chapter is not to address the backwash of institutional evaluations on the morale and motivation of teachers. Nevertheless, it is vital that governments triggering reform programmes for education dovetail the programmes for teacher quality with the development of institutional evaluation systems. The nature of assessment and evaluation must be clear. Countries should distinguish among the evaluation of students, teachers, schools, and the national education system.

With respect to institutions themselves, the NUT believes in the principle of school self-evaluation. In 1996, the NUT commissioned a report entitled, "Schools Speak for Themselves" (MacBeath, 1996). A recent pamphlet issued by the UK's *Times Educational Supplement*, "Great Minds – Education's Most Influential Philosophers", described the impact of the NUT's work in this area.

Learning to learn and pupils' experience formed part of the philosophy of Professor John MacBeath, Chair of Educational Leadership, at Cambridge University, since 2000. His distinctive contribution to contemporary education debate is the movement away from top-down inspection to school self-evaluation. His 1995 report, "Schools Speak for Themselves", commissioned by the National Union of Teachers, helped push Ministers towards the present light touch inspection system. (Judd, 2008)

There is still a powerful debate about the best forms of school evaluation but if they are to be effective, teachers have to take ownership of any type of evaluation system used.

THE NEED FOR DIALOGUE BETWEEN TEACHER ORGANISATIONS AND GOVERNMENTS ON SCHOOL IMPROVEMENT

Nowhere is dialogue more important than in the area of what incentivises teachers in their commitment to enhancing children's learning. This chapter finishes with the "how" of achieving successful reform. One essential element of that "how" is dialogue among teachers, their organisations and governments.

Dialogue is vital between those responsible for the provision of education and teacher organisations. Some describe it as a social dialogue, particularly in the European Union, but such a dialogue is better described as an educational dialogue. A mutual focus on removing barriers to high educational achievement is necessary. Both sides need to have a forum in which issues can be discussed.

Governments and those representing teachers need to agree to be honest about weaknesses and mistakes, proactive at arriving at solutions, and generous about recognising strengths and successes. The bottom line of any discussions should be a commitment to high standards and high expectations of all children.

There are various types of fora for educational dialogue in the **U.K.**, including forums for agreements on pay, conditions and professional development; for improving support for teachers in securing higher standards of student achievement and behaviour; and for improving levels of school security.

A pre-requisite to agreement has to be a recognition of positive pluralism – that both governments and teacher organisations have different roles, responsibilities and audiences. Both have separate, concentric circles of interest and responsibility. Where they overlap, there is the potential for a joint vision for the future of education and for agreement on next steps to achieve it. However, where governments seek to second guess the thinking of unions and unions seek to take on the role of government, confusion and problems will emerge. Each side has to be very clear about its objectives when pursuing an agreed strategy.

The particular forms and features of dialogue between unions and government cannot simply be transplanted from one country to another. Neither can education systems in one country be grafted onto another. Education systems have grown organically in societies. The culture and values of societies inform the curriculum of schools and all forms of education. That does not mean that the education partners in one country cannot learn from partners in another. As the debate in the recent OECD Mexican Government Conference demonstrated, it is quite possible for all groups to learn from and adapt features of other education systems and, indeed, identify and learn from the triggers and processes of education reform.

References

- Balle, A.** (2009), "National Standards: Required Attainment Targets: National Tests", PowerPoint presentation to the Education International Seminar on School Leadership, Helsinki, May, <http://www.ei-ie.org/en/calendar/show.php?id=209&theme=leadership> (Select PowerPoint presentations from related links/ file downloads), accessed September 2009.
- Cordingley, P.** (2008), *Sauce for The Goose: Learning Entitlements that Work for Teachers As Well As for Their Pupil*, CUREE (The Centre for the Use of Research and Evidence in Education), Coventry.
- Cordingley, P., M. Bell, B. Rundell and D. Evans** (2003), "The Impact of Collaborative CPD on Classroom Teaching and Learning" in *Research Evidence in Education Library*, EPPI-Centre, Institute of Education, University of London, London, <http://eppi.ioe.ac.uk/cms/Default.aspx?tabid=132>, accessed September 2009.
- Department for Children, Schools and Families (DCSF)** (2008), "School Teachers' Pay and Conditions Document 2008: Guidance on School Teachers' Pay and Conditions", The Stationery Office (TSO), London, <http://www.teachernet.gov.uk/docbank/index.cfm?id=12833>, accessed September 2009.
- Department for Education and Employment (DfEE)** (1998), "Teachers: Facing the Challenge of Change", The Stationery Office (TSO), London.
- DfEE** (2001) "Performance Management in Schools: Model Performance Management Policy", The Stationery Office (TSO), London, <http://publications.teachernet.gov.uk/eOrderingDownload/DfES%200051%20200MIG1841.pdf>, accessed September 2009.
- Horne, M.** (2001), "Classroom Assistance: Why Teachers Must Transform Teaching", Demos, London, <http://www.demos.co.uk/files/Classroomassistance.pdf?1240939425>, accessed September 2009.
- Judd, J.** (2008), "Thinkers who have Shaped Modern Education" Times Educational Supplement, 28 November, pp. 30-33 ("Great Minds-Education's Most Influential Philosophers: A TES Essential Guide").
- MacBeath, J.** (1999), *Schools Must Speak for Themselves: The Case for School Self-Evaluation*, Routledge, London.
- Marsden, F. and S. French** (1998), "What a performance: Performance Related Pay in the Public Services, Centre for Economic Performance, London School of Economics and Political Science, London, <http://cep.lse.ac.uk/pubs/download/special/performance.pdf>, accessed September 2009.
- National Union of Teachers (NUT)** (2004), "Bringing Down the Barriers: NUT Education Statement", NUT, London, <http://www.teachers.org.uk/resources/pdf/Nut-EducationStatement.FE.pdf>, accessed September 2009.
- NUT** (2007), "Born to be Great: A Charter on Promoting the Achievement of Black Caribbean Boys", NUT, London, <http://www.teachers.org.uk/resources/pdf/Black%20Youngsters%20.pdf>, accessed September 2009.
- NUT** (2008), "National Union of Teachers Survey (of) Performance Management: the Pupil Progress Objectives", NUT, London.
- OECD** (2005), *Teachers Matter: Attracting, Developing and Retaining Effective Teachers*, OECD, Paris.
- Pont, B., D. Nusche and H. Moorman** (2008a), *Improving School Leadership, Vol. 1: Policy and Practice*, OECD, Paris.
- Pont, B., D. Nusche and D. Hopkins** (2008b), *Improving School Leadership, Vol. 2: Case Studies on System Leadership*, OECD, Paris.
- Richardson, R.** (1999), "Performance Related Pay in Schools: An Assessment of the Green Papers: A Report Prepared for the National Union of Teachers", London School of Economics and NUT, London, http://www.teachers.org.uk/resources/pdf/prp_lse.pdf, accessed September 2009.
- Scanlon, M.** (1999), "The Impact of OFSTED Inspections", National Foundation for Educational Research, Slough.
- Soulsby, D. and D. Swain** (2003), "A Report on the Review of the Award Bearing INSET Scheme", DfES (Department for Education and Skills), London, http://www.teachernet.gov.uk/_doc/4129/INSET%20REPORT.doc, accessed September 2009.

CHAPTER 5

Exploring the Use of Incentives to Influence the Quality and Distribution of Teachers

Why focus on teachers?	122
Salary schedules and the incentives that exist now	124
Using incentives: technical and practical considerations	127
Evidence on purposeful teacher incentive systems	137
Conclusion: infrastructure to support incentive systems	140
Notes	141
References	144

WHY FOCUS ON TEACHERS?

A large and growing body of research on schools convincingly shows that teacher quality is the most important *schooling* factor influencing student achievement.¹ The difference between having a very effective versus a very ineffective teacher is estimated to be equivalent to as much as a full year's difference in learning growth for students (Hanushek, 1992). Moreover the impact of differences in teacher quality swamps the impact of other educational investments, such as reductions in class size (Goldhaber *et al.*, 1999; Rivkin *et al.*, 2005).² To put it in perspective, Rivkin *et al.* (2005) and Rockoff (2004) estimate that a one standard deviation increase in teacher quality raises student achievement in reading and mathematics by about 10% of a standard deviation – an achievement effect that is on the same order of magnitude as lowering class size by *10 to 13 students* (Rivkin *et al.*, 2005). Given findings like these, it is no surprise that policy makers are exploring ways to increase the quality of the teacher workforce.

Unfortunately, increasing teacher workforce quality is no simple task, primarily because the link between it and readily quantifiable teacher attributes like licensure status, degree, and experience levels is weak (Goldhaber, 2002).³ While much of the work on teacher quality is based on research in the **United States (U.S.)**, the weak link between quality and teacher credentials is echoed in research linking teacher credentials to student achievement in **Mexico**. Specifically, Santibañez *et al.* (2007) find that teacher experience and degree level are not statistically significant predictors of student achievement, and, in the case of teacher performance on licensure tests, their findings are counterintuitive in the sense that increased teacher test performance is negatively associated with student achievement on national assessments.⁴

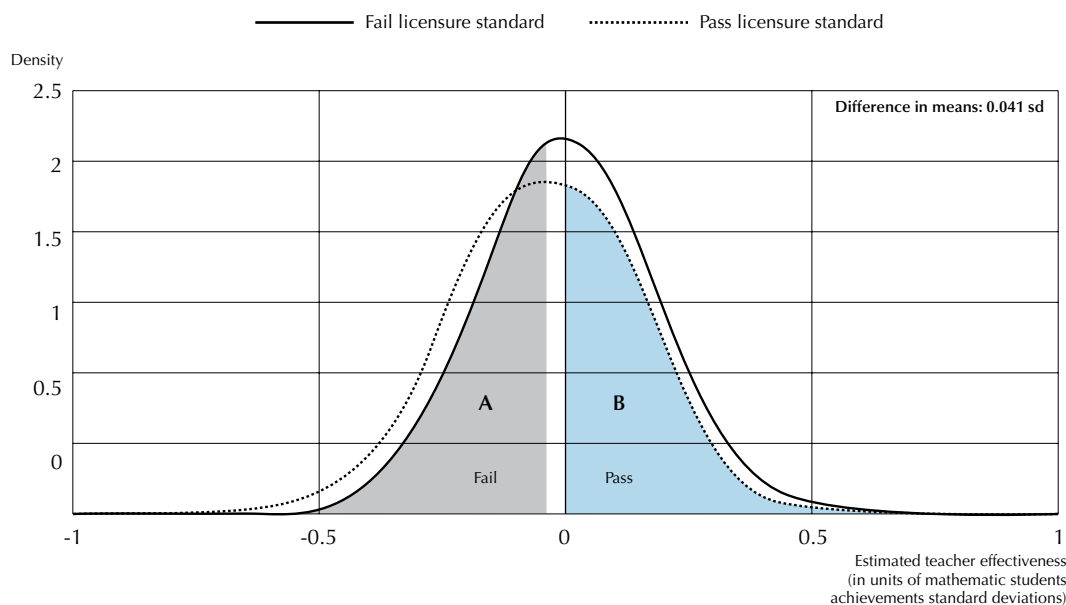
But while the above findings make it challenging to craft policies to increase teacher quality, we know that it is a highly variable commodity: some teachers are simply much better than others. This is one of the most striking findings from recent research focused on teacher effects. Kane *et al.* (2007), for example, find far more variation in observed teacher performance within a particular certification route into the classroom than between routes. They estimate the difference in the average value-added between teachers in the top and bottom quartiles of performance in each certification route to be about 30% of a standard deviation, or about ten times the average difference in value-added for teachers who enter the profession through different teaching routes.

Goldhaber (2007) finds similar results for teacher performance on licensure tests. Figure 5.1 below is derived from his analysis. This figure shows the distribution of estimated effectiveness for teachers who pass the licensure exams required to become fully certified (the solid line) and those who fail to achieve this standard (the dotted line). As one might expect, on average, teachers who achieved the standard are found to be more effective than those who do not, by about 4% of a standard deviation (the horizontal distance between F and P). But, there is also considerable overlap in the effectiveness distributions – over 90%. For example, all of the teachers shown in area A passed the required licensure tests but were less effective than the average teacher who failed to achieve the standard; teachers shown in area B failed the standard, but were more effective than the average teacher who passed.

If the required credentials, like licensure, are only weakly correlated with student achievement, then focusing on “gateway” teacher policies – for instance, requiring that teachers graduate from an approved programme or hold specific educational degrees – will be an inefficient way to influence the quality of the teacher workforce. The reason is that these gateway policies will result in significant numbers of “false positives” and “false negatives” – that is, many applicants who satisfy the criteria for employment eligibility turn out to be ineffective teachers (false positives), and many who do not satisfy the criteria would have been effective in the classroom had they been allowed into the teacher workforce (false negatives). The false negatives may never persevere to become teachers – a loss to the profession – and the false positives may be difficult to remove from the classroom once they have attained the job security that typically exists in public schools in most countries, **Mexico** included.

Figure 5.1

Distribution of estimated teacher effectiveness, by licensure standard testing



Note: Teacher effectiveness estimates are from elementary teachers with valid Praxis II scores in North Carolina public schools from the 1997-2005 school years, given current passing status using the state's standard for the Praxis II Exams. Measurement error has been controlled by using an empirical Bayes adjustment, attenuating estimates toward the grand mean in each school year. See Goldhaber (2007) and Goldhaber and Hasen (2008) for more information.

Source: Goldhaber, D. (2007), "Everyone's Doing It, But What Does Teacher Testing Tell Us About Teacher Effectiveness?", *Journal of Human Resources*, Vol. 42, No. 4, pp. 765-794.

A similar argument holds in the case of teacher compensation. Policies that link teacher pay to teacher credentials will be a crude way to reward performance if that performance is only weakly related to the rewarded credentials, as the empirical evidence suggests.

All of this lends credence to the argument that policy ought to move away from a credentials-based approach and toward one that relies on making decisions about teachers once they have been in the classroom and policy makers can make judgments about their performance. Many contend that a central component of such an approach is adding explicit incentives targeted toward individual teachers, groups of teachers (such as a school), or certain types of teachers (for example, mathematics teachers).

Compensation is, of course, an important management tool for changing incentives. Not surprisingly, there has been considerable new (and renewed) interest in using compensation to reward teacher excellence. There are, for instance, numerous places around the world where policy makers have enacted "pay-for-performance" teacher compensation programmes, although these still tend to be the exception to the normal way in which public school teachers are rewarded. Such programmes are often controversial, but the scant empirical evidence on their effects nevertheless suggests they show promise.

This chapter explores the potential for using incentives, and compensation reform in particular, as a teacher workforce policy. It begins by describing the pay system that is typically used to reward teachers in most countries, how the reward structure differs from private sector labour markets, and what incentives the dichotomy between the way that teachers, and other workers, are paid creates for prospective and existing teachers.

Next is a focus on some of the technical and practical issues that policy makers must confront when designing incentive systems (Chapter 3 provides a more in-depth treatment of the theoretical arguments related to incentive systems). These include, but are not limited to, questions about whether student achievement tests can be used to judge teacher performance, the level at which rewards ought to be targeted (for instance, individual or group), and whether policy makers might wish to use incentives other than just pay to reward teachers. Policy makers focus on many of these issues when thinking about pay-for-performance systems (or any other incentive-based system designed to reward teacher performance). However, as argued below, many of the issues addressed are pertinent to any incentive-based system, whether the incentive targets performance or is designed to achieve some other goal (e.g. a particular distribution of teachers across geographic areas).

Following this, evidence is described for teacher incentive systems, focusing in particular on a number of pay-for-performance-based systems in place around the world and on what available research has shown regarding their effects. This section makes the case that one cannot know *a priori* whether a particular incentive system will be successful, as its success will depend in no small way on the thoughtfulness of the programme implementation. The final paragraphs offer some concluding thoughts on what might be needed to support incentive systems.

SALARY SCHEDULES AND THE INCENTIVES THAT EXIST NOW

Educational policy makers often talk about creating teacher incentives without recognising that labour market incentives already exist, often just not incentives that are purposefully designed.⁵ And, there is good evidence that teachers, like employees in general (Prendergast, 1999), respond to the variety of incentives they face. These incentives influence who enters and remains in the teaching profession, and impact, for instance, the front-end choice of occupation (Hanushek and Pace, 1995; Stinebrickner, 2001a, 2001b), which schools teachers end up in (Lankford *et al.*, 2002), and such back-end choices as when they choose to retire (Costrell and Podgursky, 2008).⁶

A key determinant of whether individuals will opt to teach is the compensation they can command as teachers versus what they might earn in an alternative occupation. Comparing compensation is more challenging than might be guessed at first blush as simple comparisons in say, average salaries across sectors in an economy, may mask important differences in the characteristics and skills of those who work in those sectors, or important non-pecuniary job factors – the level of job security or workplace safety, for instance. Still, salary levels are an important place to start when assessing the desirability of an occupation.

Judging by salary, teachers are actually quite well-paid in many Latin American and Caribbean countries. There is, of course, a significant amount of variation in salaries across countries, and, to some extent, conclusions depend on the comparisons that are made. But, a 2005 review of salaries in 17 Caribbean and Latin American countries by Hernani-Limarino (2005) reveals that teachers are actually well-paid (in a relative sense) in a number of these countries. For example, teachers in **Argentina, Chile, Colombia, El Salvador, Honduras, Panama, Paraguay,** and **Peru** are paid relatively well compared to workers in other occupations, and primary school teachers in **Mexico** receive higher wages than 89% of those outside of education who have 13 to 15 years of schooling.⁷

Having high salaries, however, does not guarantee teacher quality – at least not quality that translates directly into student success. Using the ratio of teacher salaries to Gross Domestic Product per capita as a metric for how

well teachers are paid shows little obvious correlation between salaries and student achievement (Goldhaber, forthcoming a; Ladd, 2007). For example, the ratio for **Mexico** is considerably higher than the average for OECD participating countries, but achievement levels for Mexican students trails most other countries (OECD, 2005).

This is not surprising; while teacher salaries are an important determinant of the incentives associated with choosing to seek employment in schools, they are hardly the only factor that influences the choice.⁸ For example, Goldhaber (forthcoming a) shows that teacher development systems (that is, the training and selection systems that are required to become a teacher) vary widely from one country to the next and these will influence the likelihood that an individual pursuing a job as a teacher can secure one. And, other factors such as workload, working conditions, job security, and the prestige of the profession can vary considerably from one country to the next and will all influence the incentives to become a teacher.⁹

Looking beyond averages, it is likely that the structure (what teacher and job related attributes are rewarded) of teacher compensation (and incentives in general) plays an important role in determining the quality and distribution of teachers. Most OECD and Latin American countries pay teachers according to a salary schedule that determines compensation based almost entirely on a teacher's experience and educational level (Hernani-Limarino, 2005; Ladd, 2007; McEwan and Santibáñez, 2005; Murnane and Steele, 2007: OECD, 2004).¹⁰

While the private sector also often rewards these attributes of individuals, it is far more likely to reward other individual attributes such as performance on standardised tests (Murnane *et al.*, 1995), the selectivity of the college from which the individual graduates (Brewer *et al.*, 1999), and measures of productivity (Bretz and Milkovich, 1989; Hein, 1996; Milkovich and Wigdor, 1991; Podgursky and Springer, 2007).¹¹ Private sector compensation also tends to reflect supply and demand conditions that influence the compensation that individuals with particular skills can command (Grogger and Eide, 1995), or the fact that the working conditions associated with a particular job may require the payment of a "compensating differential" (additional increment of income or wage) in order to entice prospective employees to accept employment in a relatively less-desirable position over one that offers more favorable non-pecuniary working conditions or amenities (Chambers, 1995; Rosen, 1986).

This dichotomy in compensation structures between the private sector and (typically) school systems means that the distribution of earnings in teaching is generally more narrow than in the private sector – a finding reflected in research on teacher salary structure in **Chile** (Mizala and Romaguera, 2005). This in turn creates incentives in the teacher labour market, albeit not incentives that are especially designed to encourage teacher quality. In particular, individuals with attributes or skills that are better rewarded in the private sector will tend to make greater financial sacrifices to teach than others – they face different "opportunity costs". Research on the dichotomy in reward structures between public sector teacher labour markets and the private sector labour market shows the magnitude of the opportunity cost differentials between individuals with varying attributes can be striking: estimates by Goldhaber *et al.* (2008) show that employed college graduates in the **U.S.** who have technical training as a mathematics or science major in college, and who choose to teach, appear to sacrifice far more financially by doing so than do employees who lack training in science or mathematics.¹² This research shows similar findings when it comes to measures of teachers' academic competence (for example, their performance on college entrance exams).

Given these incentives, it is not surprising that research shows school systems often have great difficulty recruiting mathematics and science teachers (Hare *et al.* 2000; Harrington 2001; Shields *et al.* 2001), and that there is faster attrition out of teaching in these specialty areas than other subject areas (Murnane *et al.*, 1991). Similarly, schools with less desirable working conditions are likely to face difficulties in recruitment. In the **U.S.**, research clearly shows that teachers are sorted across schools based on student attributes, with more credentialed and experienced teachers staffing classrooms of more advantaged students. Some of this is based

on the initial match between teachers to schools (Lankford *et al.*, 2002) and because teachers are more likely to leave schools serving poorer and higher minority student populations (Hanushek *et al.*, 2004). One explanation for these findings is that teacher compensation systems typically fail to reflect the difficulty of a job.¹³

It is not unheard of for a teacher compensation package (that is, direct compensation as well as other financial inducements) to account for job difficulties or perceived amenities of a job's location. For example, Latin American countries often have difficulties recruiting teachers into rural areas (Urquiola and Vegas, 2005). Several of these countries have tried using incentives to address this fact. **Costa Rica**, for instance, has provided rural teachers with additional compensation and extra credits toward retirement; **Honduras** has provided rural teachers a salary supplement and quicker advancement up the salary scale; and **Mexico** has provided various forms of housing assistance to teachers in rural schools (McEwan, 1999). These types of incentives are, however, not the norm. Moreover, it is worth noting that salary schedules will, almost by definition, fail to reflect all of the factors that influence the desirability of a particular teaching job (Goldhaber *et al.*, 2007). Competitively determined wages in the private sector, by contrast, are thought to better reflect all factors influencing job desirability, from risk and difficulty to non-financial job amenities (Boyd *et al.*, 2003). This is another reason we would expect a dichotomy between teacher and non-teacher compensation structures.

A major concern that arises with regard to the dichotomy between the incentives associated with teacher salary schedules and the reward structure in the labour market as a whole is the possibility that the structure of compensation dissuades the most talented individuals from pursuing a teaching career. Recent research in the **U.S.** shows that college graduates with higher test scores are less likely to enter the teaching profession than other occupations (Corcoran *et al.*, 2004; Goldhaber and Liu, 2003; Hanushek and Pace, 1995; Henke *et al.*, 1996; Hoxby and Leigh, 2004; Lakdawalla, 2001). This stands in sharp contrast with the situation from several decades ago when many of the most academically capable female graduates became teachers. Specifically, Corcoran *et al.* (2004) find the likelihood that a female teacher will have been in the top 10% (on a standardised achievement test) among high school students dropped by more than half (down from about a 20% probability) in 1964, to about a 10% probability in 2000. The authors argue that the primary explanation for this is that the opportunities for women outside of teaching have expanded dramatically over the last four decades.

Hoxby and Leigh (2004) offer another explanation for the drop in the number of academically talented women entering the teaching profession. They argue that the decline results from changes over time, between teaching and other occupations, in wage compression. They attribute increased compression in teaching to a growing influence of teacher unions, and note that the increased wage compression in teaching means that an individual opting to teach is far more likely to receive a salary that is close to the average than if he or she had opted instead for a non-teaching profession. Consequently teaching has become an increasingly financially attractive occupation for lower-aptitude individuals whose compensation (in teaching) tends to be dragged upward toward the mean, and a less financially desirable occupation to those with higher aptitude whose compensation tends to be dragged downward toward the mean. Their analyses of surveys of recent college graduates from various years between 1961 and 1997 suggests that about 80% of the decline they see in teachers coming from highly selective colleges can be traced to wage compression in the teaching profession.

While the finding here of a decline in the academic calibre of the teacher workforce is restricted to the **U.S.** labour force, there is reason to be concerned that the incentives that created the situation in the **U.S.** could also be at work in Latin American teacher labour markets. In particular, research by Hernani-Limarino (2005) on teacher pay in Latin America and the Caribbean echoes some of the Hoxby and Leigh findings. Specifically, Hernani-Limarino predicts the compensation that individuals could command in and outside the teaching profession using a statistical methodology designed to account not only for observable individual factors, like labour market experience and education levels that influence compensation, but also unobserved factors that

influence an individual's productivity and consequently his or her compensation.¹⁴ This analysis suggests that, in most of the countries examined (**Nicaragua** being a notable exception), individuals who tend to be less productive would earn relatively more as teachers and those with attributes that make them highly productive would tend to earn relatively less.¹⁵ This is perfectly consistent with the Hoxby and Leigh (2004) assertion that wage compression in teaching tends to make it an occupation that is relatively undesirable to talented individuals because they can command greater compensation in an alternative occupation.

In general the research described in this section of the paper strongly suggests that the pay systems in public education that rely strictly on salary schedules do not recognise the labour market reality that some teachers will have more competing and financially attractive opportunities outside of teaching than others. Nor do they reflect the reality that some schools and classrooms will be tougher to staff because of the difficulty of the job or the associated amenities.

One can ignore these labour market incentives or try to design incentive systems that explicitly account for the realities of the labour market. Such systems might, for instance, target teacher performance and/or be designed to redistribute teachers from one school or locality to another. The technical and practical issues that arise in designing incentive systems are explored below.

USING INCENTIVES: TECHNICAL AND PRACTICAL CONSIDERATIONS

One of the issues that arises when considering teacher incentives is what type of incentives are most effective for changing labour market behaviours, e.g. increasing teacher effort, inducing teachers to move from one school to another, or keeping them in particular schools. This chapter will not delve too deeply into the theoretical arguments about different incentive types, as these are covered more extensively in chapters 2 and 3. It is worth noting, however, that a wide variety of potential incentives exist: financial incentives like housing allowances and tuition subsidies; one-time bonuses linked to student outcomes; career ladder credential-based rewards; and a whole variety of non-monetary incentives – from special recognition for performance, to choices to teach particular classes, to more time for preparation or special professional development opportunities. Unfortunately few of these myriad incentives have been rigorously studied to test their efficacy.

Still, the chapter will offer two thoughts on incentive-type. The first is that one should consider the structure of teacher development systems (for example, teacher training institutions and licensure practices) and workforce practices when considering the type of incentives offered. Specifically, incentives that represent large up-front investments may be quite costly because they subsidise individuals who do not end up teaching (or who policy makers might believe did not teach). For example, some countries, like **Singapore**, use tuition subsidies as an incentive to talented individuals to enter teacher-training programmes (Barber and Mourshed, 2007). But one of the distinguishing features of the Singapore system is that the teacher development pipeline is narrowed significantly at the point where individuals enter teacher training, and most of these graduates end up teaching over a long career. Were it the case that the country produced many more individuals from teacher training than there were teaching jobs, a general subsidy of training costs would necessarily be paid out to many individuals who do not teach (or perhaps do not teach for very long).

The second thought is that incentives structured to reward teacher credentials rather than individual teachers, based on assessments of their classroom performance, will necessarily be relatively poorly targeted. As described above, there is ample evidence showing that teachers who appear similar based on paper credentials (such as experience or educational level) vary considerably in terms of effectiveness. Thus, even if the credential that is used as a metric of success is shown, on average, to predict teacher effectiveness, it is quite likely that rewarding teachers based on this will result in significant investments in teachers who are not terribly effective due to false positives.

As an example of this, in the **U.S.**, many states and localities reward teachers who attain an advanced teaching credential, certification by the National Board for Professional Teaching Standards (NBPTS). Some evidence finds statistically significant effects on student achievement associated with having an NBPTS certified teacher (Cavalluzzo, 2004; Goldhaber and Anthony, 2007). Goldhaber (2006a) finds that there is a considerable amount of overlap in the estimated effectiveness distributions (similar to Figure 5.1 in this chapter). Specifically, about 60% of teachers who become certified are more effective than their non-certified counterparts when judging them by students' achievement in mathematics tests, and about 55% when judging them based on students' achievement on reading tests. This is certainly better than what one would expect from a flip of a coin, but hardly a guarantee that an incentive based on this credential will usually reward the more effective teachers.

Several questions arise in the context of incentives targeted to the individual teacher level. While such questions are particularly relevant to pay-for-performance incentives, they really apply to any case where the educational system wishes to target incentives to individual teachers. Policy makers may, for instance, wish to provide greater compensation to teachers with skill sets that are in demand in the private sector, but they might not wish to reward all teachers with these skills given that they will not all be effective teachers. Similarly, policy makers might wish to provide retention incentives to teachers who are in difficult-to-staff schools, but also want these to be paid out only to those who have demonstrated their effectiveness.

Incentives and Teacher Dysfunction. Individualisation of teacher incentives is controversial for a number of reasons. For instance, some suggest that the teaching profession is not well suited toward any individualised rewards because this creates inappropriate incentives. The work that teachers do is multi-dimensional and, perhaps, not all that easy to evaluate. As Dixit (2002) notes in an overview of the theoretical and empirical literature on the use of pay incentives, individualisation of incentives (such as merit pay) is far less likely in jobs that are complex, multi-faceted in objectives, and where it is difficult to accurately quantify employee productivity.

The potential problem with incentives in the case of schools is that teachers may individually exhibit dysfunctional behaviour as a consequence of the incentive. For example, they might focus very narrowly on tasks, subject areas, or content that is rewarded (often referred to as “narrowing of the curriculum” or “teaching to the test”), at the expense of additional goals or tasks valued by the public (for instance, promoting citizenship or reducing drug use or violence).¹⁶ There is also concern that individualisation could have a counterproductive effect on the work place, corroding teacher collegiality and collaboration by introducing competition between teachers (Murnane and Cohen, 1986).

Some empirical work does support the notion that individualised rewards in teaching lead to dysfunctional and counterproductive teacher behaviour. Jacob and Levitt (2003), for instance, find evidence that incentives linked to student achievement on standardised tests leads to cheating (teachers providing students with test answers), and both Cullen and Reback (2006), and Figlio and Getzler (2002) find evidence that teachers employ various mechanisms (such as classification of students into special education) to exclude the test scores of low achieving students used for accountability purposes.

Of course linking incentives to student learning as measured by standardised tests is not the only way to make judgments about teachers. One could, for instance, reward teachers based on assessments from their peers or supervisors, and there are examples of systems that rely on multiple indicators of performance other than just student test scores (a topic discussed more thoroughly in Chapter 3).

It is natural to wonder whether non-student test means of assessment do distinguish effective from ineffective teachers. The relatively thin body of empirical evidence that focuses on the correlation between various means of judging teachers and more direct, “value-added”, measures of teacher effectiveness (value-added is described in more detail below) suggests that there are ways to judge teachers that do line up reasonably well with their contribution to student learning, as measured by student performance on standardised tests.

Joint teacher production and collaboration. Perhaps even more important than the question of whether incentives would lead to dysfunction at the individual teacher level is the larger issue of whether individualisation of incentives leads to school-level dysfunction due to impacts on collaboration. This is a particularly salient issue at the secondary level, where students often switch between classes and teachers, and there are likely to be complementarities across subjects like mathematics and science. Koedel (forthcoming) directly addresses the issue of cross-subject complements and/or their relation to teacher incentives. He estimates value-added reading achievement models and tests whether student's reading achievement is influenced not only by the quality of their English teachers but also by their mathematics, social studies, and science teachers. Koedel finds very little evidence to suggest joint production (across subjects) of students' reading achievement.

Koedel's findings do not definitively show that we should not be worried that an incentive-based system would lead to dysfunction, but they do lessen the concern since problems with collaboration are likely to have greater student achievement consequence if achievement in one subject is strongly related to the quality of teachers in another.¹⁷ Also heartening is evidence that individualised incentives (merit pay in particular) exist in a variety of different work settings that require collaboration, including, for instance, far more extensive use of pay for performance at private schools, at least in the **U.S.** (Ballou and Podgursky, 1997). Finally, direct evidence from an incentive system suggests the concerns about associated school dysfunction may be exaggerated: Lavy (2004) studies an individual-level teacher incentive in Israel and finds no negative impacts on student achievement in subject areas that were not part of the incentive system.

Of course it is possible to directly address the concern that individualised teacher incentives might create within-school collaboration problems by balancing them with group-wide incentives for a department or an entire school. One advantage of school-based incentives is that they help address the problem of how to evaluate teachers in non-tested subjects, at least in the special case of using student tests to judge performance (a subject focused on in more detail below). Student achievement tests generally only cover mathematics and reading (and sometimes science), and are administered in selected grades: in the **U.S.**, for instance, it is typical that less than a quarter of teachers are covered by grades and subjects in which standardised tests are administered to students. This means many teachers could not be directly evaluated were the evaluation to be based on standardised test achievement. Thus, a school-level focus has the virtue of being applicable to all teachers in a school, even those in areas where students are not typically assessed based on tests (e.g. music).

A number of countries have experimented with school-based teacher incentives and they provide some examples of different approaches to comparing schools. In **Chile** for instance, the National System of Performance Assessment (*Sistema Nacional de Evaluación de Desempeño de los Establecimientos Educacionales*, SNED) compares schools within various geographic regions of the country and all teachers in the top performing school in each region receive financial bonuses (Mizala and Romaguera, 2005). It is unclear, however, whether this system is fair to all schools since there may be significant variation in the type of students (and the advantages they have) enrolled in schools within each region of the country.

In the **U.S.**, various states and school districts have also used school-based rewards (Clotfelter and Ladd 1996; Meyer 1996; Odden and Kelley 1997). While some systems rely on comparisons of demographically similar schools, others, unlike the SNED in **Chile**, use some type of statistical adjustments to compare schools. For example, **South Carolina's** group-based performance pay plan awards bonuses to schools based on improvements in test scores, after controlling for differences in socio-economic status and average teacher education level. The **Dallas** Independent School District judged schools based on average performance on the state assessment, controlling for socio-economic and limited-English-proficiency status and race/ethnicity in a more complex, two-stage statistical process.

Regardless of the method, the key when making school-to-school or teacher-to-teacher comparisons is that the comparisons are done in such a way that the contributions of schools or teachers are isolated from other factors that influence students' achievement, such as the background of students. This is fundamental if one wants the rewards associated with an incentive system to be fair in the sense that they are based on what schools or teachers actually influence. So, for instance, one would not want to directly compare the performance at a school serving economically-advantaged students to a school serving disproportionately economically-disadvantaged students.

While there are upsides to utilising school-based teacher incentives – *i.e.* addressing the collaboration problem and accounting for teachers in untested grades or subjects – there are also important downsides to a strictly school-based incentive system. One of these is that the incentives are diffuse, which may result in a tendency for teachers to try to free-ride on the efforts of their colleagues (Hansen, 1997; Prendergast, 1999). Perhaps more importantly, strictly group-based systems will fail to appropriately reward standout teachers within the targeted group. Research on the variation in student test scores suggests this is a significant issue when it comes to teachers, as more of the variation in teacher effectiveness is within schools than between schools (Rivkin *et al.*, 2005).

There is, however, no reason that one has to choose between group and individual incentives. The discussion above suggests that a hybrid system, with a portion of a teacher's incentive tied to group performance and a portion tied to individual performance, might be most efficacious. Given this, it is worth closing this subsection with a practical consideration about the individualisation of teacher incentives. For incentives to have an impact, they have to be perceived to be both attainable and significant enough to induce behavioural changes, for example, large enough monetary awards (Lavy, 2007). As described below, there is considerable variation in both the proportion of teachers who receive incentive payments, and the absolute and relative (to earnings) magnitude of financial inducements across incentive schemes. However, there is no consensus in the empirical literature on what constitutes the right way to structure incentives, in terms of the proportion of teachers to receive them or their size relative to teachers' base compensation.

Value-added methods for using student test scores as a metric to judge teachers. There are often institutional constraints when it comes to subjective evaluations of teachers, and this is one of the key arguments for using statistical methodologies (value-added modeling) to try to isolate the contributions that individual teachers make toward student achievement.

But several potentially important technical issues are associated with using student test scores as a metric to judge teacher effectiveness. These issues are complex, so it can only be provided a cursory overview at best. However, before delving into them, it is worth noting that, for many, the idea of using student achievement on tests to make any kind of judgments about schools is itself controversial. Standardised tests are limited in terms of what they measure. At best tests can only be used to assess a subset of the myriad objectives of schooling (for example, there are unlikely to be tests that cover the complex socialisation behaviour that is taught in schools) so are limited in terms of their ability to identify the extent to which students have learnt important topics covered in the classroom (Koretz, 2008). Moreover, even the objectives that are covered by tests will be imperfectly measured (for instance, a mathematics test will be a sample of what is taught in a mathematics class).

At the deepest level, issues about using tests to evaluate teachers are intimately tied to one's views about whether they are useful in assessing what students have learnt. While it is important to keep the limits of tests in mind when making policy, one must weigh these limits against two other factors. First, in many countries, student tests are typically constructed with input from numerous stakeholder groups that help to drive educational policies (such as school accountability), and thus presumably reflect a country's educational goals. Second, student achievement on standardised tests has been shown to be an important predictor of later life outcomes (Grogger and Eide, 1995; Murnane *et al.*, 1995) and is related to aggregate country growth rates (Hanushek *et al.*, 2008).

Moving beyond the question of whether standardised tests are a good measure of student achievement, it is necessary to confront whether they can be used to assess teacher performance. The statistical models, referred to as “value-added models” (VAMs), that are used to predict the factors that influence student learning rely on a number of strong assumptions. Moreover, these models rely on student tests that themselves have various psychometric properties. Tests may, for instance, based on the specific tests items included in them, mask learning gains for high- or low-achieving students. As a consequence, if used to assess teacher performance, they potentially over or understate teacher effectiveness in different settings.

There are also issues associated with the amount of true information about teacher effectiveness that are conveyed even by the best of student tests. One would not, for instance, want to attach too much meaning to changes in test performance if a teacher only has a small handful number of students. The reason is that tests are always imprecise measures of what students know, a student who takes the same test twice on the same day would likely not get the same score simply due to random factors. Thus, teachers whose measure of effectiveness is based on only a few student tests scores are far more likely to be influenced by these random factors. In statistical parlance, we worry about the amount of true “signal” about teacher effectiveness that is conveyed as opposed to statistical “noise”.

Lastly, there are important questions about what are the right comparisons between teachers. The hope is that statistical models help to adjust for differences in the type of students a teacher has in his or her classroom, but one has to make decisions about whether teachers are compared within a school, across schools in a district, or at some higher level, across a region or the entire country for instance.

Policy interest in using VAMs has led to a growing body of literature that examines all of the above issues (e.g. OECD, 2008). There is in fact a long history of value-added modeling in economics (e.g. Hanushek, 1972) and the term itself covers a broad class of different statistical specifications. It is only recently that VAMs have been used extensively to try to identify the contributions that individual teachers make toward student learning gains.¹⁸

Efforts to identify causal effects of teachers (or any schooling inputs) are complicated by the complexity of the learning process. As Todd and Wolpin (2003) point out, VAMs rely on strong assumptions about the nature of student learning over time and violations of those model assumptions can lead to erroneous conclusions about the impacts of educational resources, such as teachers.

As but a few examples: VAMs typically assume constant decay of learning (the amount of learning that children forget between test administrations), but the amount of decay of student learning from grade-to-grade may not be consistent across grades (a “grade” in the **U.S.** system is equivalent to a year level in other OECD systems). Greater or lesser amounts of decay in one grade or the next would lead statisticians to over- or understate the contributions of teachers in those grades. Teachers might also contribute to student learning beyond the grade they are teaching. VAMs typically assume that the entirety of teacher contributions from one grade to the next is captured by the growth in test scores between those grades (and this then decays at a constant rate over time). Thus, in cases where what teachers teach in grade *g* impacts student learning growth in grade *g*+1 or a higher grade, the impact of that teacher would be unstated and would instead be misattributed to teachers in those higher grades. This might happen if, for instance, teachers in grade *g* teach skills not captured by the grade *g* standardised test.

When estimating teacher effects, statisticians also have to make key decisions about which students (and their test results) are used to calculate teacher value-added estimates. VAMs generally rely on measuring gains in student achievement, so restrict the sample of students to those who are tested at least twice. As a consequence, highly mobile students – those for whom gains are not available – will not be incorporated into estimates of teacher job performance. Even if students have multiple tests, allowing gain estimates, one has to decide what to do with those who are not in a teacher’s classroom for an entire year (e.g. they may have come into a class in October rather than September but still taken a standardised test administered in November).

It is not clear whether the implications of student mobility for assessing individual teacher effects have been empirically examined. An untested but intuitive way to address this issue is to attribute student achievement gains to the teachers in proportion to the time that they are in a teacher's class. This of course is made easier if a state/region or country's data system allows for the tracking of students from one school to the next.

Another reason that students may be excluded from factoring into teacher performance estimates is that data on their backgrounds are missing or limited. If students' background data are needed for accurate assessments of teachers, one would not want to rely on estimates that do not utilise background information. However, if student background information is not important, one would want to include the full complement of students taught by each teacher since that would constitute the most complete data on how teachers impact their students. McCaffrey *et al.* (2003), explore this issue using a series of simulations. They show that models relying on longitudinal data on student achievement produce similar estimates of teacher performance, even in the absence of detailed student background information, but only so long as there is a significant amount of mixing of students across teachers.¹⁹ If there is little overlap of students, then it is necessary to have information on students' backgrounds in order to make valid judgments about teacher performance.

Ballou (2005a) and Ballou *et al.* (2004) focus on the importance of student background controls using empirically-based models in which they estimate teacher effects. They find that the omission of student background controls, such as race and poverty, in value-added teacher effects models seldom leads to substantial errors in estimates of teacher performance when comparisons between teachers are within the same school district. It is important to note, however, that this research is based on a particular district with substantial racial and economic integration (Ballou, 2005b).

Another complicating factor when it comes to using VAM methods to estimate teacher effectiveness is the assignment of students to teachers. Teacher assignment to schools and classes and parental preferences for particular teachers makes it extremely unlikely that teachers and students are randomly matched to one another in the **U.S.**²⁰ In fact, were matching random, we would not need to worry much about including student background controls in statistical models because teachers would not be systematically assigned to different types of students. We also would not worry if we had complete information on the backgrounds of all students and if this information perfectly predicted the contribution to student learning of a student's home and environment (peers, etc.). If this were the case, we could simply use the information in statistical models to adjust the expectations of how a teacher's class should perform for the type of students he or she has in the class. But because statisticians do not have complete background information (they will not, for instance, know the precise amount of time that a student's parents spend encouraging and helping a student with school-work) and test scores are imperfect measures of student ability, a component of any VAM error term will be unmeasured student ability. And, this error is likely to be correlated with the teacher that a student is assigned to given a non-random matching process between teachers and students. This correlation is a violation of a fundamental assumption of statistical models that is necessary to ensure that the findings from the model, including the estimates of teacher effectiveness, are not erroneous.

There is good reason to believe that the student-teacher matching process is more than just a theoretical concern. For example, as mentioned earlier in this chapter, a significant amount of research in the **U.S.** shows that more-advantaged students, in terms of family income and parental education, tend to be assigned to higher-quality teachers, as measured by such characteristics as experience, degree level, and test performance (Lankford *et al.*, 2002; Loeb, 2001). And, research by Clotfelter *et al.* (2006) shows that nonrandom matching is likely to produce upwardly-biased teacher coefficient estimates in statistical models that fail to explicitly account for it.

It may be possible to account for the matching process by estimating more sophisticated statistical models (referred to as "school" or "student fixed-effects" model specifications), but this is often difficult given available

data on schools, teachers and students.²¹ It is for this reason that some researchers (e.g. Rothstein, 2008a) caution against using VAM teacher-effect estimates for high-stakes accountability purposes, like pay for performance. While this is a concern, on balance, it seems worth using VAM since the system typically used today to allocate teacher pay does not look like it is aligned well with teacher productivity. Moreover, there are statistical ways to limit the likely error associated with VAM measures of teacher effectiveness. For example, Rothstein (2008b) notes that VAM models that include a series of prior student achievement are less likely to result in biased estimates of teacher effects. And, Kane and Staiger (2008) directly test the possibility of bias in estimates of teacher effects exploiting an experiment where teachers are randomly matched to classes. They assess various VAM specifications and find that a model specification that includes student test scores from several prior years (as opposed to a single prior year, as is common) produces teacher effects quite similar to those produced under conditions where teachers are randomly matched to their classrooms. Finally, as Koedel and Betts (2009) show, evaluating teachers using multiple years of teacher-student classroom data is also likely to mitigate the potential bias arising from the matching process.

Closely related to the general issues of VAM specification is the fact that VAMs rely on psychometric properties of tests that may or may not exist. Most important among these is the requirement that achievement gains of different students, who may be at different levels of achievement or even grade levels, can be compared with one another. This is possible if the scale on which achievement is measured is of equal units across students, so that one student's gain in achievement can be meaningfully compared to another, even if they do not start at the same point in the achievement distribution. Some tests are designed explicitly to measure growth, typically employing "item response theory" to create "vertically aligned" scaled scores.²² However, many tests are not designed so that they have equal-interval scales within and between all tested grade levels to facilitate the comparisons of students.

Among the potential implications of this is that teachers may not easily be compared. For example, a test may have few high-level items, making it difficult for high-performing students to increase their scores by very much; as a consequence, value-added methods would measure incorrectly (by understating) the contributions of teachers with relatively more of these students in their classes.²³ On the other hand, a test that concentrates too much on high-level items may be inadequate for identifying the needs of low-performing students and/or miss year-to-year learning gains made by these students.²⁴ Recent work (Koedel and Betts, forthcoming) has addressed whether value-added teacher effectiveness estimates may be sensitive to ceilings (such as students scoring near to the very top of a test scale) in the testing instrument, but little other empirical work tackles test scaling and value-added directly. One method used to compensate for a lack of vertical alignment in a test instrument is to place student test results on a separate metric; for instance, standard normalising student achievement scores and comparing where students with particular teachers fall in the distribution from one year to the next. It is unclear whether this type of transformation fully accounts for non-vertical alignment of tests.

Those wishing to make judgments about individual teachers based on student tests are unambiguously better off doing so using tests that are properly scaled. Scaling, however, is not the only psychometric issue that one needs to attend to when it comes to tests used for accountability purposes. As described below, measurement error is a significant issue in assessing teacher performance and measurement error will depend on the quality and reliability of the items in the test or tests that are used for this purpose. In this context it is worth noting that a recent RAND report raises concerns about the psychometric properties of the student tests that are part of **Mexico's Carrera Magisterial** Programme. It notes, for instance, that there is no documentation describing concrete evidence about test item quality, and that test items are often revised informally without regard to the psychometric consequences of changing items, for example the validity of the test (Santibáñez *et al.*, 2007).

Related to the issue of measurement error is the question of how to decide how much student-teacher matched data is necessary before it is appropriate to make judgments about teachers. A student's performance on a test

will reflect both true ability or knowledge as well as random influences (for instance, sickness on testing days) referred to as measurement error or statistical “noise”. When it comes to assessing teacher job performance, this source of test measurement error means a component of any estimated teacher effects is this source of test measurement error. Added to this, in the case of teachers, is the luck of the draw in students that teachers are assigned.

Researchers have explored this issue extensively in the case of school-level performance estimates. Findings from the **U.S.** (Kane and Staiger, 2001) and **Chile** (Chay *et al.*, 2005) show that a substantial amount of “noise” results from test measurement error and, further, this implies that small schools are far more likely than large schools to be classified as being either very effective or ineffective, but this classification is less likely to accurately identify the true effectiveness of these schools.²⁵

In the case of individual teachers, measurement error is of greater concern since VAM estimates of teacher job performance will generally be based on smaller samples of students than are estimates of school-level performance. This means that many teachers will be indistinguishable from one another from a statistical standpoint. For example, in an analysis of individual-teacher-year performance estimates, Ballou (2005b) finds in the overwhelming majority of cases that individual teacher effects cannot be statistically distinguished from the average (at the 90% confidence level) teacher: in reading only 2.5% of primary and 7.6% of lower-secondary-school teacher effects were different from average; in the corresponding figures effects were slightly higher at 17% and 30% respectively for primary- and lower-secondary-school teachers.

Measurement error will also necessarily mean that teachers who are judged based on VAM estimates to be on either side of a cut-point will not be statistically distinguishable. For example, were a country to adopt an incentive policy to reward teachers who fall into the top quintile of performance, teachers at the 79th percentile of estimated performance could make a strong statistical case that they are as effective as teachers who are at the 81st percentile of estimated performance.

The issue of measurement error has recently received a great deal of attention in the context of individual teachers. For example, researchers have used various related methods to decompose the total yearly variance of teacher effectiveness estimates into estimates of “true” teacher effects, “signal”, and sampling error or “noise” (see, for instance, Aaronson *et al.*, 2007; Goldhaber and Hansen, 2008a; Koedel and Betts, 2009a; McCaffrey *et al.*, 2008; Rothstein, 2008a, 2008b), but these find the correlations of teacher value-added job performance estimates from one year to the next that are thought to be “modest” (see Aaronson *et al.*, 2007; Ballou, 2005; Koedel and Betts, 2007; Goldhaber and Hansen, 2008a; McCaffrey *et al.*, 2008), and a non-trivial proportion of the inter-temporal variation in these effects appears to be due to statistical noise. For example, the inter-temporal correlation of teacher VAM effect estimates are in the range of 0.3 to 0.5, and generally higher when using student mathematics achievement as a metric than when using student reading achievement (Goldhaber and Hansen, 2008a). And, ranking teachers based on these estimates shows, at both the primary level (Koedel and Betts, 2007) and the lower-secondary-school level (McCaffrey *et al.*, 2008), that roughly a quarter to a third of teachers who are ranked in the top quintile of performance (based on VAMs) stay in the same quintile from one year to the next; 10 - 15% of teachers move all the way from the bottom quintile to the top; and 10 - 15% move from the top quintile to the bottom.

One cannot definitively distinguish how much of the variability over time in teacher effects is due to noise versus true changes in teacher performance, but one way to improve the stability of these estimates is to use multiple-years to calculate teacher effects and these multi-year estimates will be less prone to be noisy because a lucky draw in one year is offset by an unlucky draw the next (Goldhaber and Hansen, 2008a; Koedel and Betts, 2009). It is also worth putting these issues into context. There is no universal standard as to what constitutes an acceptable job performance stability level, but the inter-temporal correlations reported

above are not very different from estimates of job performance in sectors of the economy that consider them for high-stakes purposes (such as job retention and pay determination), where the year-to-year correlations of performance ranged from 0.26 to 0.55 (see Hoffman *et al.*, 1992, 1993).

Two final technical issues that arise when using test scores to judge teacher effectiveness are related to questions about the comparison group against which teachers are to be judged. Teachers may be compared to other teachers within various geographic or institutional boundaries. For example, teachers may be compared to other teachers within the same school or district or state.²⁶ The downside to less localised models that create between-school teacher comparisons is that they conflate teacher effects with effects that may be important at other levels of the educational system, for example schools or districts (Goldhaber and Hansen, 2008a). All teachers at a school with a highly effective principal will benefit from that principal's influence on student achievement in models that allow for between-school comparisons.

The downside of more localised comparisons is that it may be quite difficult to create appropriate counterfactual comparison groups. This is likely to be particularly problematic, for instance, in the case of smaller school-governance structures (such as small school districts), since these will necessarily have few teachers on whom to base their assessments of relative effectiveness. And, in extreme cases there may be no relevant comparison group – for instance if there is only one mathematics teacher in a particular region of a country. Moreover, teachers who are judged to be highly effective in one school, as in the case of within-school VAM estimates, may actually be relatively ineffective compared to a broader set of teachers, say in a state as a whole.

There is no definitive right answer to the right level of teacher comparisons. In a value-added framework, teacher effects are identified relative to the impact of other teachers, not some objective standard. Thus, a key to obtaining good estimates of a teacher's relative contribution toward student achievement growth is that a given teacher is compared against other teachers who are teaching comparable students in a comparable school setting. Depending on the relative influences of student background versus other school or state or regional factors, one could make the case for either type of comparison, but given the overwhelming research showing the influence of student background on achievement, there is perhaps more of an argument for the broader set of comparisons over those that are more localised.

Alternatives to value-added used for judging teacher performance. The prior subsection deals with the idea of using statistical methods to parcel out the value-added of teachers toward student test performance as a metric for teacher performance. While intuitively appealing, this is not the only means of evaluating teachers. As Toch and Rothman (2008) point out, few teacher evaluation systems appear to be very rigorous or have been tested statistically to see if they correspond to VAM estimates of teacher effectiveness.

There are a few exceptions. For instance, the National Board for Professional Teaching Standards is a U.S.-based organisation specifically designed to assess the skill set of experienced teachers. And, as mentioned above, some research shows that the National Board assessment process does differentiate effective from ineffective teachers, albeit with some error.²⁷ And an investigation of a particular career ladder performance system, where a teacher's advancement (leading to greater compensation) is based largely on subject evaluations of supervisors, shows that teachers who advance through the system are more effective in value-added terms than those who do not advance.²⁸ This study is quite convincing because it is based on an experiment where students are randomly assigned to teachers, eliminating the potential that the findings are a byproduct of the match between teachers and students.²⁹

Finally, Jacob and Lefgren (2008) correlate principal evaluations of teachers to teacher value-added effect estimates and find that principals generally do a good job of identifying very effective and very ineffective teachers, but are less able to discriminate among teachers who fall into the middle of the effectiveness

distribution. Importantly, the authors also find evidence that principal evaluations of teachers predict student achievement and identify teachers' attributes correlated with their success that go beyond what one can detect simply based on statistical measures of teacher effectiveness.³⁰

In sum, the above studies suggest that one can utilise means other than student test scores to assess teachers. However, it is an open question as to whether school systems that rely on other means will actually differentiate amongst teachers. The problem is that, in many places, political and institutional constraints (such as teacher union/school relationships) make it difficult to do rigorous teacher evaluations (Toch and Rothman, 2008), resulting in very few teachers who receive anything other than a top rating (New Teacher Project, 2007). It could be argued that one of the primary reasons that policy makers are intrigued with the idea of using statistical (value-added methods) to identify the effectiveness of individual teachers is to get around the difficulty of getting principals or other evaluators to make honest assessments of teacher performance.

Summary on measuring teacher effectiveness. The research described above has certainly provided valuable lessons about the appropriate ways to measure teacher effectiveness. However, it is an overstatement to suggest that debates about how precisely to measure teacher contributions to student learning are settled. But having described some of the issues that arise in the context of value-added, it seems appropriate to point out that this methodology should not be judged against some imagined nirvana but, rather, against the system that exists. And here the relevant alternative in most countries is basing teacher employment and compensation on credentials that are themselves only weakly, at best, associated with student learning.

In some ways it is appropriate to let one's theory of action for how individual teacher incentives might influence the quality of the workforce be a guide to the ways in which individual teacher effects are measured. There is likely a trade-off between the accuracy of VAM teacher effects estimates and both the transparency to teachers about how their effects are estimated and their view of the incentive. So, if, for instance, the idea is to identify the most effective teachers and hold them out as role models so that other teachers can learn from them, then it is very important that the right teachers be identified. This suggests a complex value-added model. However, if the theory is that teachers respond to incentives for performance, it may be more important that they understand the ways in which they will be judged than that the VAM estimates be as accurate as possible. Teacher performance varies from year to year, so while multiple-year estimates of teacher effects are likely to be more accurate, they are a weaker incentive because high value-added years offset low value-added years and vice versa (Rivkin, forthcoming). It is therefore conceivable that the more efficacious incentive policies might sacrifice some accuracy for sharper incentives.

The size and cost of individual teacher incentives. Two important issues that arise when it comes to connecting measures of teacher effectiveness to teacher incentives for performance are the size of incentives (e.g. what percentage of a teacher's overall compensation package is tied to the incentive) and the cost of an incentive system.

Indeed a criticism of many teacher incentive systems is that the size of the incentive is too small to motivate changes in teacher behaviour. Research on the SNED in **Chile**, for example, shows that only a very small percentage of teachers in the country are eligible for a performance bonus, and the performance component of salaries is small relative to overall compensation and other factors (e.g. experience) that influence them; for example, in 2003, less than one percentage point of a teacher's salary was based on performance (Mizala and Romaguera, 2005).

Economic theory would suggest that even a very small increment of pay that is structured as an incentive should have marginal impacts (Prendergast, 1999), but, in practice, the performance incentive programmes that have been tried in public schools tend to have incentives that are small relative to those outside of public

schools. Ballou (2001) who investigates pay-for-performance incentives in **U.S.** schools finds that the average teacher incentive in private schools is about 10% of overall pay, while it is only about 2% in public schools with incentives. Studies of the private sector suggest that the proportion of pay that is incentive-based varies considerably across different firms and industries (Murphy, 1999; Prendergast, 1999), but, in general the proportion of pay that is tied to the incentive it is far higher than even the 10% observed in private schools.

The next section describes the empirical evidence of the effects of different incentive systems. In general, the evidence presented does not appear to show a clear correspondence between the distribution or size of teacher incentive systems and the effects of those systems.³¹ This is not terribly surprising since there are relatively few studies that assess the impacts of teacher performance incentives.

Another important factor policy makers should consider when designing teacher incentive systems is the cost of the systems. There is research that delves into the cost of different designs, but there are at least two good reasons to believe that these systems should increase educational costs (though hopefully with a more than offsetting increase in student achievement), at least in the short-run. The first is obvious: they will be more costly to administer since they would require assessments of, and the administrative capacity to track payments to, individual teachers. The second is less obvious, but potentially more important. To the degree that incentives represent a significant portion of overall teacher compensation, teachers would have to receive a risk premium since they would have greater uncertainty about their overall compensation than they would under a salary-schedule-based system.

EVIDENCE ON PURPOSEFUL TEACHER INCENTIVE SYSTEMS

There appears to be an upswing in interest in creating teacher incentives to attract talent into teaching, reward high-performing teachers, and redistribute teachers to various geographic areas and schools. This is true in the **U.S.**, where the federal government has recently instituted the Teacher Incentive Fund, which provides grants to states or localities to develop pay alternatives to the traditional salary schedule (Goldhaber, 2006b), as well as in a number of European countries (Sclafani and Tucker, 2006). **Mexico**, as part of the *Carrera Magisterial* programme, was one of the first countries to explicitly link teacher compensation to teacher performance on a countrywide scale, though performance-based teacher incentives are also utilised in other countries like Singapore (Sclafani, 2008).

Despite the interest in using incentives to influence teachers, relatively little is definitively known about the right type of incentives to influence teacher performance or distribution. This is both because few countries widely use teacher incentives and because studying the effects of incentives in a rigorous way requires data that are often not available. The concern when evaluating programmes is that often there are not appropriate counterfactual comparison groups to teachers who are subject to an incentive programme and, under these circumstances, research may conflate the impact of an incentive with the impact of those who receive the incentive. As an example, some pay-for-performance programmes, like that in **Mexico's** *Carrera Magisterial*, are designed such that participation in the incentive system is voluntary (McEwan and Santibáñez, 2005). In this case teachers are self-selecting into programme status and one might worry, for instance, that the more motivated teachers who believe they will perform well will disproportionately select into the programme. Were this the case, a simple comparison between programme participants and non-participants may misattribute the impact of the type of teachers who opt to participate to the impact of the incentive programme itself. Here are described some of the studies on incentives that are designed explicitly to address the issue of self-selection; for a broader discussion of different teacher incentive systems and their impacts, see Chapter 3.

Most of the studies on incentives focus on financial incentives targeting teacher performance. Lavy (2002) and Glewwe *et al.* (2003) focus on school-level performance incentives. Lavy's investigation utilises a "regression-discontinuity" approach to compare student outcomes in 62 schools in **Israel** that are subject to performance

incentives with those of schools that just missed being part of the incentive system (the treatment); these two groups of schools are likely to be quite similar in terms of unobserved teacher (and other) factors in virtue of the fact that the difference between being part of the treatment or control group is likely to be close to random around the cut point that determines eligibility to participate in the incentive programme.³²

Lavy finds evidence that participating in the incentive programme had statistically significant positive impacts on several measured outcomes: student test scores, number of credits taken, and matriculation exam completion. Moreover, the effects were larger for students from more disadvantaged backgrounds. This finding is consistent with research on teacher performance incentives in the U.S. (Figlio and Kenny, 2007) that finds larger impacts of incentives in schools serving more disadvantaged students. It suggests that policy makers who are unsure about the efficacy of such incentives in general might wish to implement them first in the neediest schools since these are the ones that are thought to benefit the most from them.³³

Glewwe *et al.* (2003) focus on an actual experiment in **Kenya** where teachers in 50 schools, randomly selected from 100, were randomly assigned to an experimental group, subject to performance incentives, and were compared to teachers in a non-experimental group who were not part of a performance system. They find that incentivised teachers increased the number of test-preparation sessions they offered after school, but there was only a one-year increase in student pass rates on the state exam. These findings can be interpreted as teachers increasing their effort levels, but the authors also note that one might view them as teachers trying to manipulate the system by focusing narrowly on test preparation. This concern is highlighted by the fact that they did not find an impact on teacher absences or pedagogical approaches and that the student achievement gains did not persist beyond the programme.³⁴

Several rigorous studies also focus on the impacts of performance pay at the individual teacher level. Lavy (2004) uses a regression discontinuity approach to assess a teacher performance incentive programme at the upper-secondary school level in **Israel**.³⁵ He finds statistically significant positive treatment effects (on upper-secondary school matriculation rates), and, moreover, the magnitude of the effects suggests that the performance incentive at the individual teacher level is more cost-effective than school-based bonuses or extra instruction time (and roughly comparable to incentive bonuses paid directly to students for achievement).

Santibáñez *et al.* (2007) investigate the effects of **Mexico's** *Carrera Magisterial*, a far-reaching teacher incentive programme introduced in 1992 with the intent to enhance teacher performance through skill development and explicit financial incentives for student achievement results.³⁶ Like Lavy (2002, 2004), they utilise a regression discontinuity approach to try to isolate the impact of the programme's effect. However, unlike Lavy, this research reports mixed impact findings: the programme was found to have negative (though not statistically significant) effects on students at the primary level and some positive impacts on students at the secondary level (on the order of magnitude of 3-15% of a standard deviation), but these positive results were only found for teachers – about 4% of their sample – who faced very strong incentives. Moreover, they find decreases in student performance subsequent to teachers receiving bonuses or being promoted as part of the programme, perhaps due to mean reversion in test scores or due to the fact that the incentives for teachers to improve student test scores are far weaker after teachers are incorporated into the *Carrera Magisterial* system and/or are promoted (Santibáñez *et al.*, 2007).

While not definitive, the findings from Lavy and Santibáñez *et al.* might be interpreted as showing the importance of the teacher perception of the incentive system. Weak incentives, e.g. those that are small in magnitude or diffuse because they apply to groups, may not have the impact of incentives that are seen to be more significant and based on individual performance. This conclusion is certainly consistent with research on teacher performance incentives in the **U.S.** by Figlio and Kenny (2007), who find that performance incentives do not have an impact on student achievement when they are received by a large fraction of teachers at a school,

but when more selectively administered are positively correlated with student achievement. Thus, as is the case with many educational interventions, the details of implementation may be as important to the success of the intervention as the nature of the intervention itself.

Finally, Muralidharan and Sundararamin (2008) report on the findings from a very large experiment on school resource and teacher-incentive interventions in a representative sample of 500 government-run, rural primary schools in **India**. They find very large effects (0.19 on mathematics tests and 0.12 on language tests, larger effects than in those found in the studies described above) associated with being in an incentive treatment school.³⁷ Moreover, the treatment gains were found on both low-level and more conceptual test items, and students in treatment schools outperformed students in non-treatment schools in subjects that were not the focus of the incentive (science and social studies). These findings suggest that: 1) teachers are not focused on narrow test strategies designed to boost student achievement; and, 2) the incentives, rather than causing teachers to substitute away from subjects not covered by the incentive, resulted in positive spillovers in these subject areas.

Unlike Lavy, Muralidharan and Sundararamin did not find any statistically significant difference between individual and group-based incentive designs, but their findings are consistent with Lavy in that the teacher incentive was found to be more cost-effective than other school resource interventions.³⁸ One particularly interesting aspect of the study is that they analyzed surveys that were administered to teachers, in which they were asked how they would allocate a 15% budget increase between performance pay and an across-the-board increase in salaries. Holding constant expected pay, the researchers find a positive and statistically significant relationship between teachers' preferences for performance pay and the average gains of their students (the survey was administered before teachers knew those gains). This suggests that effective teachers know who they are and wish to be rewarded for their performance, and, further, that performance-based pay systems may yield workforce productivity benefits associated with more sorting, that is, those with teaching talent opting into the teaching profession and the more effective teachers opting to remain in the profession. This idea is consistent with more general research on the labour market; for instance, research by Lazear (2000) finds that about half the gains from performance-pay result from greater efforts among current employees and half from more productive workers sorting into a performance-based system.

There are far fewer rigorous studies of the impact of other types of financial (such as housing assistance or tuition waivers) or non-financial incentives, or studies of incentives designed to attract teachers with particular skill sets (like mathematics and science) or distribute them to particular schools. Much of the empirical literature in this area focuses on inferences that are drawn from observing labour market behaviour in the absence of incentive programmes: what type of people opt to enter and remain in the teacher labour market and where they opt to teach.³⁹ While this research is certainly valuable, one must be wary about drawing strong conclusions from it since estimates of necessary compensating differentials, for instance, may be confounded by unobserved heterogeneity among those who opt to pursue a particular profession or move from one school to another.⁴⁰

Studies that focus on actual incentive-based programmes are more credible. One such study (Clotfelter *et al.*, 2008) focuses on the impact of a financial inducement to teach in hard-to-staff schools. This study analyzes a programme in **North Carolina** designed to retain mathematics and science teachers in economically disadvantaged public schools (in North Carolina). This research finds that a relatively small bonus (USD 1 800, which was worth approximately 5% of an average teachers' salary) reduced turnover rates by just over 15%. More recently, Steele *et al.* (2009) focus on an incentive programme (in **California**) designed to induce academically talented novice teachers to take positions and remain (for at least four years) in low-performing schools.⁴¹ This programme provided far larger incentives than those in North Carolina, USD 5 000 for each year of service (up to four years), which was equivalent to about 15% of average starting salaries. Perhaps not surprising, given

the magnitude of the incentives, the authors find that the incentive increased the probability that incentive recipients opted to teach in low-performing schools by nearly 30 percentage points.⁴²

While there is little empirical work to guide policy makers on the efficacy of non-financial teacher incentives, on the whole, the research described in this section clearly suggests that incentives can have significant effects on teacher behaviour. But, as I briefly describe in the next section, the success of incentive systems are likely to hinge on more than the system itself.

CONCLUSION: INFRASTRUCTURE TO SUPPORT INCENTIVE SYSTEMS

This chapter has described some of the shortcomings of the strict adherence to the single- salary schedule as the primary structure for teacher compensation. One of the reforms that has gained wide-spread attention is tying incentives to value-added estimates of teacher (or school) performance. The research on the testing and data requirements necessary to do this, such that performance estimates are accurate portrayals of true teacher or school performance, is relatively new so there are no clear guidelines for precisely how these effects ought to be measured. At the very least, we know that performance measures, particularly those that focus on individual teachers, are likely to be noisy and will necessarily only reflect the material covered by test assessments. This does not necessarily mean that value-added ought not to be utilised, since alternative ways of assessing teachers have also not been proven. Rather, it is probably a bad idea for policy makers to base incentives solely on value-added measures when they might also rely on indicators such as supervisor or peer evaluations.

Having an incentive programme is not sufficient to guarantee success. One of the notable exceptions to the generally positive findings on teacher performance incentives, reviewed above, is **Mexico's** incentive programme. The research by Santibáñez *et al.* (2007) and Hopkins *et al.*, (2007) finds that, despite the *Carrera Magisterial* programme, the incentives for performance for many Mexican teachers are quite weak. Finally, perhaps more important than the design of a specific teacher incentive is the implementation of and support for incentive programmes. For example, one of the important findings when it comes to incentive reforms is that teachers often do not understand the way that programmes work or even that incentives exist (Clotfelter *et al.*, 2008; Goldhaber, 2009; 2006b; Jacob and Springer, 2007); this is surely not conducive to programmes having their intended effects. That said, while the research on teacher compensation reform is hardly definitive enough to recommend the use of specific pay reforms to reach specific goals, the few quantitative studies that do exist suggest that a more strategic use of teacher compensation could lead to a more equitable allocation of teachers among students and to increased student achievement.

NOTES

1. In this chapter the term “teacher quality” refers to mean the *ability of teachers to contribute in measurable ways to student gains on standardised tests* (the terms “teacher quality,” “teacher performance,” and “teacher effectiveness” are used interchangeably). While helping students to learn so that they perform better on tests is a somewhat narrow conception of the totality of what teachers do, it is useful to define teacher quality in this way because it provides a means of measuring teachers based on measurable student outcomes. As I go on to describe below, other input based definitions of quality (such as a teacher’s credentials) have not been found to be strongly related to student achievement. Moreover, student performance on tests have been shown to be important predictors of later life outcomes (Grogger and Eide, 1995; Murnane *et al.*, 1995), are related to aggregate country growth rates (Hanushek *et al.*, 2008), and, in many countries, help to drive educational policies (for example, school accountability).
2. Not surprisingly, over time the impact of teachers can be substantial. For example, work by Sanders and Rivers (1996) finds that students having a sequence of three consecutive teachers who fall into the highest quintile group score between 52 - 54 percentile points higher in the achievement distribution than those who have three consecutive teachers who fall into the lowest quintile.
3. Hanushek (1986, 1997) makes the case that observable teacher credentials are only weakly correlated with teacher quality. More recently, several studies employ teacher fixed-effects models to estimate the contribution of observable teacher characteristics relative to unobservable teacher factors (Aaronson *et al.*, 2007; Goldhaber *et al.*, 1999; Koedel and Betts, 2007). They typically find, based on the incremental improvement in R^2 , that the vast majority of the total variation in teacher quality (over 95%) is explained by unobservable teacher characteristics rather than observable characteristics such as degree and experience levels.
4. They find somewhat more positive relationships between teacher peer review ratings, participation in professional development, and student achievement.
5. Furthermore, there are many types of incentives, from intrinsic motivation to job security, that are likely to influence labour market choices.
6. For a more thorough review, see Vegas and Umansky (2005).
7. By contrast, depending on the comparison group, teachers in Bolivia, Brazil, the Dominican Republic, Ecuador, and Nicaragua were found to be relatively poorly paid.
8. Moreover, in some countries retirement and health benefits constitute a relatively large share of overall compensation and these benefits are typically not considered when making comparisons in compensation between teaching and other occupations.
9. There are mixed findings on the relationship between average salaries and teacher quality and student achievement. For example, see Ballou and Podgursky (1997), Goldhaber (forthcoming), and Hanushek (1986), all of whom suggest little connection between salary and teacher quality or student outcomes; and Figlio (1997) and Loeb and Page (2000), who find stronger connections between the two.
10. Some countries, such as Korea (Oh, 2008), report having incentive *systems*, but nearly all teachers receive, for instance, a performance award – so there is, de facto, no real incentive. One notable exception to this is Singapore, which provides teachers with large bonuses (equivalent to one to three months of pay) based upon subjective evaluations of both current teacher performance and future potential (Sclafani, 2008).
11. Bretz and Milkovich (1989) estimate that 93 – 99% of private sector firms in the U.S. reward salaried individuals based on some measure of their individual performance.
12. Specifically, four years out of college the estimated *gap* in salary between teachers and non-teachers who have mathematics and science training is USD 13 469, but only USD 6 811 for those who do not have this type of training.
13. In fact, recent evidence (Player, forthcoming) suggests that sought-after teachers are rewarded in the labour market, not by compensation but by assignment to more-advantaged students.
14. To compare individuals with similar levels of unobserved heterogeneity, Hernani-Limarino calculates, based on Least Absolute Deviation estimates, conditional wage differentials for various quantiles along the conditional wage distribution.
15. See Figure 3.12 on page 95 and Table 3.9 on page 96 in Hernani-Limarino (2005).
16. Economists referred to this as the “multi-tasking problem,” whereby the incentivising of activities that are easier to monitor leads employees to focus on these in lieu of other valuable activities that are less easily monitored.
17. Koedel’s findings also suggest that value-added methods might be applied at the secondary level to isolate the contributions of teachers in particular subjects. The isolation of teacher contributions would be far more difficult in the case of strong cross-subject complementarities.

18. For background on this issue, see, for instance, Ballou (2005); Ballou *et al.* (2004); McCaffrey *et al.* (2004); Rothstein (2008a, 2008b); and Todd and Wolpin (2003).
19. For example, to obtain accurate estimates of the effectiveness of teachers A, B, and C, at least some students have to be common to teachers A and B and B and C; or to teachers A and C and B and C. When excluding student covariates, identification of differences across the student population relies on the cross-correlation between students. It is this mixing that ensures the cross-correlation estimates are valid.
20. Matching of teachers and students may happen both formally through ability tracking of students, or more informally, for instance, if “good” teachers are rewarded with choice class assignments or parents exert influence to get their children into choice classrooms (Player, forthcoming).
21. If fixed effects specifications fully control for the student-teacher matching process, then statistical models should produce unbiased estimates of individual teacher effects; however, if the matching process is based on student or teacher attributes that are unobserved, then teacher effect estimates are likely to be biased. Rothstein (2008a), noting the problems of nonrandom teacher-student matches, develops falsification tests of the assumptions undergirding commonly used methods in order to identify the impact of individual teachers using non-experimental data. Specifically, using longitudinal panel data from North Carolina, he shows that 5th grade teachers have statistically significant impacts on 4th grade student achievement in student fixed effects specifications of VAMs. Clearly the 5th grade teacher effect cannot be causal (since students have not yet reached that grade) so his findings imply both that students are not randomly assigned to teachers, and that the inclusion of student fixed effects do not fully account for the teacher-student matching process (note that this falsification test is a test of whether bias exists, not the extent of bias). This can occur if, for instance, the teacher-student matching process is “dynamic,” and unobserved by the statistician.
22. See Ballou (2002) for more information on this methodology.
23. Tests with many lower-level items will more easily detect achievement gains for low-performing students, even if these students are not substantially increasing their learning levels.
24. For a more thorough discussion of the relationship between test scaling and value-added modeling, see Ballou (forthcoming).
25. Regression to the mean is an issue related to measurement error. Specifically, a component of student achievement is luck, but luck is random so students who do unusually well or poorly on tests in one year (or test administration) are likely to do less well the next simply because a component of their performance, be it high or low, was related to luck and that is not likely to persist with repeated tests. Thus, teachers with very high-achieving students in one year are likely to look less effective in value-added terms because their students are not likely to be as lucky when tested again.
26. For example, models that include school-level fixed effects create within-school teacher comparisons.
27. See Goldhaber and Anthony (2007) for a more detailed description of the ways in which the National Board assesses teachers and the findings on teachers who obtain this credential.
28. A discussion is included on value-added teacher effects below, but the basic idea is to use statistical models to isolate the contribution that teachers make toward student achievement gains on standardised tests.
29. A number of studies show this matching process can lead to biased coefficient estimates, see, for instance, Clotfelter *et al.* (2006), Goldhaber (forthcoming b), and Cantrell *et al.* (2008).
30. Specifically, they find that principal evaluations are statistically significant predictors of student achievement even in models that include estimates of the prior value-added of a student’s teacher.
31. Moreover, research is only beginning to assess how teachers perceive and respond to financial incentives versus other types of non-compensation educational investments (for instance, offering teachers a smaller class or special professional development opportunities) that might be used as inducements (Goldhaber *et al.*, 2008a). One study (Figlio and Kenny, 2007) finds that teacher financial incentives have an impact on student performance but non-financial teacher incentives do not.
32. The central idea of regression-discontinuity is that classification into treatment or control status is tantamount to being random around some cut-point, above which is the treatment and below which is the control (or vice versa).
33. Furthermore, given what was said about the resources necessary to successfully implement teacher incentives, in the prior section of the chapter, this would mean that the neediest schools would be targeted with additional educational funds.
34. For other examples of studies on school-wide incentives, see Clotfelter and Ladd (1996) and Ladd (1999).
35. He also confirms his findings based on a propensity score matching methodology.
36. The programme also created salary differentials for work in rural areas of Mexico, but this aspect of the programme has not, to my knowledge, been rigorously assessed.

37. Interestingly, the results were not a consequence of differences in teacher absenteeism (a significant problem in India) or readily observable behaviours; however, teachers in treatment schools report assigning more homework to students and providing more afterschool supplemental classes than teachers in untreated schools.

38. For additional examples of studies on high-stakes teacher incentives systems, see, for instance, Brooke (2008) on Brazil; Mizala and Romaguera (2005) on Chile; and Figlio and Kenny (2007) on the U.S.

39. See, for instance, Boyd *et al.* (2003), Goldhaber *et al.* (2007), Hanushek *et al.* (2004), and Scafidi *et al.* (2007).

40. As an example, research estimates of the cost of inducing teachers to teach or stay in disadvantaged schools, based on observed moves between schools, may be overstated given that a transition a teacher makes from one school to the next will be made only if the benefits of moving to a new school exceed both the benefits of staying in the old school and the transaction cost of the move.

41. A limited number of incentives were available and recipients of these were determined based on a competitive process that judged their academic records.

42. The authors find little difference in retention between incentive recipients and observationally comparable non-recipients, though they are careful to note that these findings are merely descriptive (while they were able to use econometric techniques to try to isolate the impact of the incentive on the likelihood that individuals opted to take positions in low-performing schools, they were unable to utilise these same techniques to determine whether the incentive had a causal impact on retention).

References

- Aaronson, D., L. Barrow and W. Sanders** (2007), "Teachers and Student Achievement in the Chicago Public High Schools", *Journal of Labor Economics*, Vol. 25, No. 1, pp. 95-135.
- Ballou, D.** (2001), "Pay for Performance in Public and Private Schools", *Economics of Education Review*, Vol. 20, Issue 1, February, pp. 51-61.
- Ballou, D.** (2002), "Sizing Up Test Scores", *Education Next*, Vol. 2, No. 2., pp. 10-15.
- Ballou, D.** (2005a), "Value-Added Assessment: Controlling for Context with Misspecified Models", paper presented at the Urban Institute Longitudinal Data Conference, March.
- Ballou, D.** (2005b), "Value-Added Assessments: Lessons from Tennessee", in R. Lissitz (ed.), *Value-Added Models in Education: Theory and Applications*, JAM Press, Maple Grove, MN, pp. 272-297.
- Ballou, D.** (forthcoming), "Test Scaling and Value-Added Measurement", *Education Finance and Policy*.
- Ballou, D. and M. Podgursky** (1997), *Teacher Pay and Teacher Quality*, W. E. Upjohn Institute for Employment Research, Kalamazoo, MI.
- Ballou, D., W. Sanders and P. Wright** (2004), "Controlling for Student Background in Value-Added Assessment of Teachers", *Journal of Educational and Behavioral Statistics*, Vol. 29, No. 1, pp. 37-66.
- Barber, M. and M. Mourshed** (2007), *How the World's Best-Performing School Systems Come Out on Top*, McKinsey & Company, New York, <http://www.closingtheachievementgap.org/cs/ctag/view/resources/111>, accessed January 2009.
- Boyd, D., H. Lankford, S. Loeb and J. Wyckoff** (2003), "Analyzing the Determinants of the Matching Public School Teachers to Jobs: Estimating Compensating Differentials in Imperfect Labor Markets", *NBER Working Paper No. 9878*, National Bureau of Economic Research, Cambridge, MA.
- Bretz, R. and G. Milkovich** (1989), "Performance Appraisal in Large Organisations: Practice and Research Implications", Working Paper No. 89-17, Centre for Advanced Human Resource Studies, Cornell University, Ithaca, NY.
- Brewer, D., E. Eide and R. Ehrenberg** (1999), "Does it Pay to Attend an Elite Private College? Cross-Cohort Evidence on the Effects of College Type On Earnings", *Journal of Human Resources*, Vol. 34, No. 1, pp. 104-123.
- Brooke, N.** (2008), "Educational Accountability in Brazil", *Revista Iberoamericana de Evaluación Educativa*, Vol. 1, No. 1.
- Cantrell, S., J. Fullerton, T. Kane and D. Staiger** (2008), "National Board Certification and Teacher Effectiveness: Evidence from a Random Assignment Experiment", *NBER Working Paper No. 14608*, National Bureau of Economic Research, Cambridge, MA.
- Cavalluzzo, L.** (2004), "Is National Board Certification An Effective Signal Of Teacher Quality?", The CNA Corporation, Washington, D.C.
- Chambers, J. and W. Fowler** (1995), *Public School Teacher Cost Differences Across the United States*, National Center for Education Statistics, NES 95-758, U.S. Department of Education, Washington, D.C.
- Chay, K., P. McEwan and M. Urquiola** (2005), "The Central Role of Noise in Evaluating Interventions that Use Test Scores to Rank Schools", *American Economic Review*, Vol. 95, No. 4, pp. 1237-1258.
- Clotfelter, C., E. Glennie, H. Ladd and J. Vigdor** (2008), "Would Higher Salaries Keep Teachers in High-Poverty Schools? Evidence from a Policy Intervention in North Carolina", *Journal of Public Economics*, Vol. 92, No. 5-6, pp. 1352-1370.
- Clotfelter, C. and H. Ladd** (1996), "Recognizing and Rewarding Success in Public Schools", in H. Ladd (ed.), *Holding Schools Accountable: Performance-based Reform in Education*, The Brookings Institution, Washington, D.C.

- Clotfelter, C., H. Ladd and J. Vigdor** (2006), "Teacher-Student Matching and the Assessment of Teacher Effectiveness", *Journal of Human Resources*, Vol. 41, No. 4, pp. 778–820.
- Corcoran, S., W. Evans and R. Schwab** (2004), "Changing Labor-Market Opportunities for Women and the Quality of Teachers, 1957–2000", *American Economic Review*, Vol. 94, No. 2, pp. 230–235.
- Costrell, R. and M. Podgursky** (2008), "Peaks, Cliffs, and Valleys: The Peculiar Incentives of Teacher Pensions", *Education Next*, Vol. 8, pp. 22–28.
- Cullen, J. and R. Reback** (2006), "Tinkering Toward Accolades: School Gaming Under a Performance Accountability System", *NBER Working Paper No. 12286*, National Bureau of Economic Research, Cambridge, MA.
- Dixit, A.** (2002), "Incentives and Organisations in the Public Sector: An Interpretative Review", *Journal of Human Resources*, Vol. 37, No. 4, pp. 696–727.
- Figlio, D.** (1997), "Teacher Salaries and Teacher Quality", *Economics Letters*, Vol. 55, Issue 2, pp. 267–71.
- Figlio, D. and L. Getzler** (2002), "Accountability, Ability, and Disability: Gaming the System", *NBER Working Paper 9307*, National Bureau of Economic Research, Cambridge, MA.
- Figlio, D. and L. Kenny** (2007), "Individual Teacher Incentives and Student Performance", *Journal of Public Economics*, Vol. 91, Issues 5–6, pp. 901–914.
- Glewwe, P., N. Ilias and M. Kremer** (2003), "Teacher Incentives", *NBER Working Paper 9671*, National Bureau of Economic Research, Cambridge, MA.
- Goldhaber, D.** (2002), "The Mystery of Good Teaching", *Education Next*, Vol. 2, No. 1, pp. 50–55.
- Goldhaber, D.** (2006a), "National Board Teachers Are More Effective, But Are They in the Classrooms Where They're Needed the Most?", *Education Finance and Policy*, Vol. 1, No. 3, pp. 372–382.
- Goldhaber, D.** (2006b), "Teacher Pay Reforms: The Political Implications of Recent Research", Center for American Progress, Washington, D.C.
- Goldhaber, D.** (2007), "Everyone's Doing It, But What Does Teacher Testing Tell Us About Teacher Effectiveness?", *Journal of Human Resources*, Vol. 42, No. 4, pp. 765–794.
- Goldhaber, D.** (2009), "Politics of Teacher Pay Reform", in M. Springer (ed.), *Performance Incentives: Their Growing Impact on American K-12 Education*, Brookings Institution Press, Washington, D.C.
- Goldhaber, D.** (forthcoming a), "Lessons from Abroad: Exploring Cross-Country Differences in Teacher Development Systems and What They Mean for U.S. Policy", in D. Goldhaber and J. Hannaway (eds.), *Creating a New Teaching Profession*, Urban Institute Press, Washington, D.C.
- Goldhaber, D.** (forthcoming b), "Licensure Tests: Their Use and Value for Increasing the Quality of the Teacher Workforce", in M. Kennedy (ed.), *A Handbook on Teacher Assessment and Teacher Quality*, Jossey-Bass, New York.
- Goldhaber, D. and E. Anthony** (2007), "Can Teacher Quality be Effectively Assessed? National Board Certification as a Signal of Effective Teaching", *Review of Economics and Statistics*, Vol. 89, No. 1, pp. 134–150.
- Goldhaber, D., D. Brewer and D. Anderson** (1999), "A Three-Way Error Components Analysis of Educational Productivity", *Education Economics*, Vol. 7, No. 3, pp. 199–208.
- Goldhaber, D., M. DeArmond, A. Liu and D. Player** (2008), "Returns to Skill and Teacher Wage Premiums: What Can We Learn By Comparing the Teacher and Private Sector Labor Markets?", *SFRP Working Paper*, No. 8, Center on Reinventing Public Education, Seattle, WA.
- Goldhaber, D., K. Destler and D. Player** (2007), "Teacher Labor Markets and the Perils of Using Hedonics to Estimate Compensating Differentials in the Public Sector", *SFRP Working Paper*, No. 17, Center on Reinventing Public Education, Seattle, WA.
- Goldhaber D. and M. Hansen** (2008a), "Is It Just a Bad Class? Assessing the Stability of Measured Teacher Performance", *CRPE Working Paper No. 2008-5*, Center on Reinventing Public Education, Seattle, WA.

- Goldhaber D.** and **M. Hansen** (2008b), "Assessing the Potential of Using Value-Added Estimates of Teacher Job Performance for Making Tenure Decisions", *CRPE Research Brief*. Center on Reinventing Public Education, Seattle, WA.
- Goldhaber, D.** and **A. Liu** (2003), "Occupational Choices and the Academic Proficiency of the Teacher Workforce", in W. Fowler (ed.), *Developments in School Finance 2001–02*, National Center for Education Statistics, Washington, D.C., pp. 53-75.
- Grogger, J.** and **E. Eide** (1995), "Changes in College Skills and the Rise in the College Wage Premium", *Journal of Human Resources*, Vol. 30, No. 2, pp. 280–310.
- Hansen, D.** (1997), "Worker Performance and Group Incentives: A Case Study", *Industrial and Labor Relations Review*, Vol. 51, No. 1, pp. 37-49.
- Hanushek, E.** (1972), *Education and Race: An Analysis of the Educational Production Process*, D.C. Heath and Company, Lexington, MA.
- Hanushek, E.** (1986), "The Economics of Schooling: Production and Efficiency in Public Schools", *Journal of Economic Literature*, Vol. 24, No. 3, pp. 1141-77.
- Hanushek, E.** (1992), "The Trade-off Between Child Quantity and Quality", *Journal of Political Economy*, Vol. 100, No. 1, pp. 84-117.
- Hanushek, E.** (1997), "Assessing the Effects of School Resources on Student Performance: An Update", *Educational Evaluation and Policy Analysis*, Vol. 19, No. 2, pp. 141-164.
- Hanushek, E., D. Jamison, E. Jamison and L. Wößmann** (2008), "Education and Economic Growth: It's Not Just Going to School But Learning That Matters", *Education Next*, Vol. 8, No. 2, pp. 62-70.
- Hanushek, E., J. Kain and S. Rivkin** (2004), "Why Public Schools Lose Teachers", *Journal of Human Resources*, Vol. 39, No. 2, pp. 326-354.
- Hanushek, E.** and **R. Pace** (1995), "Who Chooses to Teach (and Why?)", *Economics of Education Review*, Vol. 14, No. 2, pp. 101-17.
- Harrington, P.** (2001), "Attracting New Teachers Requires Changing Old Rules", *College Board Review*, Vol. 192, No. 6, pp. 6-11.
- Hare, D., J. Nathan, J. Darland and S. Laine** (2000), "Teacher Shortages in the Midwest: Current Trends and Future Issues", North Central Regional Educational Laboratory, Oak Brook, IL.
- Hein, K.** (1996), "Raises Fail, But Incentives Save the Day", *Incentive*, Vol. 170, p. 11.
- Henke, R., S. Geis, J. Giambattista and P. Knepper** (1996), "Out of the Lecture Hall and into the Classroom: 1992-1993 College Graduates and Elementary/Secondary School Teaching: With an Essay on Undergraduate Academic Experiences", U.S. Department of Education, Office of Educational Research and Improvement, Washington, D.C.
- Hernani-Limarino, W.** (2005), "Are Teachers Well Paid in Latin America and the Caribbean? Relative Wage and Structure of Returns of Teachers", in E. Vegas (ed.), *Incentives to Improve Teaching: Lessons from Latin America*, The World Bank, Washington, D.C., pp. 63-96.
- Hoffman, D., R. Jacobs and S. Gerras** (1992), "Mapping Individual Performance Over Time", *American Psychological Association*, Vol. 77, No. 2, pp. 185-195.
- Hoffman, D., R. Jacobs and J. Baratta** (1993), "Dynamic Criteria and the Measurement of Change", *Journal of Applied Psychology*, Vol. 78, No. 2, pp. 194-204.
- Hopkins, D., E. Ahtaridou, P. Matthews and C. Posner** (2007), "An Analysis of the Mexican School System in Light of PISA 2006", London Centre for Leadership in Learning, Institute of Education, University of London.
- Hoxby, C.** and **A. Leigh** (2004), "Pulled Away or Pushed Out? Explaining the Decline of Teacher Aptitude in the United States", *American Economic Review*, Vol. 94, Issue 2, pp. 236-240.
- Jacob, B.** and **L. Lefgren** (2008), "Can Principals Identify Effective Teachers? Evidence on Subjective Performance Evaluation in Education", *Journal of Labor Economics*, Vol. 26, No. 1, pp. 101- 136.

- Jacob, B.** and **S. Levitt** (2003), "Rotten Apples: An Investigation of the Prevalence and Predictors of Teacher Cheating", *Quarterly Journal of Economics* Vol.118, No. 3, pp. 843-78.
- Jacob, B.** and **M. Springer** (2007), "Teacher Attitudes on Pay For Performance: A Pilot Study", *Working Paper 2007-06*, National Center on Performance Incentives, Vanderbilt University, Nashville, TN.
- Kane, T., J. Rockoff** and **D. Staiger** (2007), "Photo Finish: Certification Doesn't Guarantee a Winner", *Education Next*, Vol. 7, No. 1, pp. 60-67.
- Kane, T.** and **D. Staiger** (2001), *Improving School Accountability Measures*, National Bureau of Economic Research, Cambridge, MA.
- Kane, T.** and **D. Staiger** (2008), "Are Teacher-Level Value-Added Estimates Biased? An Experimental Validation of Non-Experimental Estimates", paper presented at National Conference on Value-Added Modeling, University of Wisconsin, Madison, WI.
- Koedel, C.** (forthcoming), "An Empirical Analysis of Teacher Spillover Effects in Secondary School", *Economics of Education Review*.
- Koedel, C.** and **J. Betts** (2007), "Re-Examining the Role of Teacher Quality in the Educational Production Function", *Working Paper WP 07-08*, University of Missouri-Columbia, Columbia, MO.
- Koedel, C.** and **J. Betts** (2009), "Does Student Sorting Invalidate Value-Added Models of Teacher Effectiveness? An Extended Analysis of the Rothstein Critique", *Working Paper WP 09-02*, University of Missouri-Columbia, Columbia, MO.
- Koedel, C.** and **J. Betts** (forthcoming), "Value-added to What? How a Ceiling in the Testing Instrument Influences Value-Added Estimation", *Education Finance and Policy*.
- Koretz, D.** (2008), *Measuring Up: What Educational Testing Really Tells Us*, Harvard University Press, Cambridge, MA.
- Ladd, H.** (1999), "The Dallas School Accountability and Incentive Program: An Evaluation of its Impacts on Student Outcomes", *Economics of Education Review*, Vol. 18, No. 1, pp. 1-16.
- Ladd, H.** (2007), "Teacher Labor Markets in Developed Countries", *The Future of Children*, Vol. 17, No. 1, pp. 201-217.
- Lakdawalla, D.** (2001), "The Declining Quality of Teachers", *NBER Working Paper No. 8263*, National Bureau of Economic Research, Cambridge, MA.
- Lankford, H., S. Loeb** and **J. Wyckoff** (2002), "Teacher Sorting and the Plight of Urban Schools: A Descriptive Analysis", *Educational Evaluation and Policy Analysis*, Vol. 24, No. 1, pp. 37-62.
- Lavy, V.** (2002), "Evaluating the Effect of Teachers' Group Performance Incentives on Pupil Achievement", *Journal of Political Economy*, Vol. 110, No. 6, pp. 1286-317.
- Lavy, V.** (2004), "Performance Pay and Teachers' Effort, Productivity and Grading Ethics", *NBER Working Paper No. 10622*, National Bureau of Economic Research, Cambridge, MA.
- Lavy, V.** (2007), "Using Performance-Based Pay to Improve the Quality of Teachers", *The Future of Children*, Vol. 17, No. 1, pp. 87-110.
- Lazear, E.** (2000), "Performance Pay and Productivity", *American Economic Review*, Vol. 90, No. 5, pp. 1346-1361.
- Loeb, S.** (2001), "Teacher Quality: Its Enhancement and Potential for Improving Pupil Achievement", in D. Monk, H. Walberg and M. Wang (eds.), *Improving Educational Productivity*, Information Age Publishing, Greenwich, CT.
- Loeb, S.** and **M. Page** (2000), "Examining the Link between Teacher Wages and Student Outcomes: The Importance of Alternative Labor Market Opportunities and Non-pecuniary Variation", *The Review of Economics and Statistics*, Vol. 82, Issue 3, pp. 393-408.
- McCaffrey, D., D. Koretz, J. Lockwood** and **L. Hamilton** (2003), "Evaluating Value-Added Models for Teacher Accountability", RAND Corporation, Santa Monica, CA.
- McCaffrey, D., T. Sass** and **J. Lockwood** (2008), "The Intertemporal Stability of Teacher Effect Estimates", *NCPI Working Paper No. 2008-22*, National Center on Performance Incentives, Nashville, TN.

- McEwan, P.J.** (1999), "Recruitment of Rural Teachers in Developing Countries: An Economic Analysis", *Teaching and Teacher Education*, Vol. 15, No. 8, pp. 849-859.
- McEwan, P.** and **L. Santibáñez** (2005), "Teacher and Principal Incentives in Mexico", in E. Vegas (ed.), *Incentives to Improve Teaching: Lessons from Latin America*, The World Bank, Washington, D.C., pp. 213-253.
- Meyer, R.H.** (1996), "Comments on Chapters Two, Three, and Four", in H. Ladd (ed.), *Holding Schools Accountable: Performance-Based Reform in Education*, The Brookings Institution, Washington, D.C.
- Milkovich, G.** and **A. Wigdor** (eds.) (1991), *Pay for Performance: Evaluating Performance Appraisal and Merit Pay*, National Academies Press, Washington, D.C.
- Mizala, A.** and **P. Romaguera** (2005), "Teachers' Salary Structure and Incentives in Chile", in E. Vegas (ed.), *Incentives to Improve Teaching: Lessons from Latin America*, The World Bank, Washington, D.C., pp. 103-150.
- Muralidharan, K.** and **V. Sundaraman** (2008), "Teacher Incentives in Developing Countries: Experimental Evidence from India", *NCPI Working Paper Number 2008-13*, National Center on Performance Incentives, Nashville, TN.
- Murnane, R.J.** and **D. Cohen** (1986), "Merit Pay and the Evaluation Problem: Why Most Merit Pay Plans Fail and a Few Survive", *Harvard Educational Review*, Vol. 56, pp. 1-17.
- Murnane, R., J. Singer, J. Willett, J. Kemple** and **R. Olsen** (1991), *Who Will Teach? Policies That Matter*, Harvard University Press, Cambridge, MA.
- Murnane, R.** and **J. Steele** (2007), "What is the Problem? The Challenge of Providing Effective Teachers for All Children", *The Future of Children*, Vol. 17, No. 1, pp. 15-44.
- Murnane, R., J. Willett** and **F. Levy** (1995), "The Growing Importance Of Cognitive Skills In Wage Determination", *The Review of Economics and Statistics*, Vol. 77, No. 2, pp.251-66.
- Murphy, K.** (1999), "Executive Compensation", in O. Ashenfelter and D. Card (eds.), *Handbook of Labor Economics*, Vol. 3, North Holland, Amsterdam, pp. 2485-2563.
- New Teacher Project** (2007), "Hiring, Assignment, and Transfers in Chicago Public Schools," <http://www.tntp.org/files/TNTPAnalysis-Chicago.pdf>, accessed April 2009.
- Odden, A.** and **C. Kelley** (1997), *Paying Teachers for What They Know and Do: New and Smarter Compensation Strategies to Improve Schools*, Corwin Press, Thousand Oaks, CA.
- OECD** (2004), "The Quality of the Teaching Workforce", *Policy Brief* (February), OECD, Paris.
- OECD** (2005), *Teachers Matter: Attracting, Developing, and Retaining Effective Teachers*, OECD, Paris.
- OECD** (2008), *Measuring Improvements in Learning Outcomes: Best Practices to Assess the Value-Added in Schools*, OECD, Paris.
- Oh, H.** (2008), "The Teacher Training System and Teaching Profession in Korea", paper prepared for the Urban Institute's Gates Human Capital Reform Project, Washington, D.C.
- Player, D.** (forthcoming), "Nonmonetary Returns to Quality in the Public Teacher Labor Market", *Education Finance and Policy*.
- Podgursky, M.** and **M. Springer** (2007), "Teacher Performance Pay: A Review", *Journal of Policy Analysis and Management*, Vol. 26, No. 4, pp. 909-950.
- Prendergast, C.** (1999), "The Provision of Incentives in Firms", *Journal of Economic Literature*, Vol. 37, No. 1, pp. 7-63.
- Rivkin, S.** (forthcoming), "The Estimation of Teacher Value-added as a Determinant of Performance Pay", in D. Goldhaber and J. Hannaway (eds.), *Creating a New Teaching Profession*, Urban Institute Press, Washington, D.C.
- Rivkin, S., E. Hanushek** and **J. Kain** (2005), "Teachers, Schools, and Academic Achievement", *Econometrica*, Vol. 73, No. 2, pp. 417-458.
- Rockoff, J.** (2004), "The Impact of Individual Teachers on Students' Achievement: Evidence from Panel Data", *American Economic Review*, Vol. 94, Issue 2, pp. 247-252.

- Rosen, S.** (1986), "The Theory of Equalizing Differences", in O. Ashenfelter and R. Layard (eds.), *Handbook of Labor Economics*, Vol. 1, North Holland, Amsterdam.
- Rothstein, J.** (2008a), "Teacher Quality in Educational Production: Tracking, Decay, and Student Achievement", *NBER Working Paper No. 14442*, National Bureau for Economic Research, Cambridge, MA.
- Rothstein, J.** (2008b), "Student Sorting and Bias in Value Added Estimation: Selection on Observables and Unobservables", *NBER Working Paper No. 14666*, National Bureau for Economic Research, Cambridge, MA.
- Sanders, W. and J. Rivers** (1996), "Cumulative and Residual Effects of Teachers on Future Student Academic Achievement", University of Tennessee Value-Added Research and Assessment Center, Knoxville, TN.
- Santibáñez, L., J. Martínez, A. Datar, P. McEwan, C. Setodji and R. Basurto-Dávila** (2007), "Breaking Ground Analysis of the Assessment System and Impact of Mexico's Teacher Incentive Program 'Carrera Magisterial'", *RAND Technical Report*, RAND Corporation, Santa Monica, CA.
- Scafidi, B., D. Sjoquist and T. Stinebrickner** (2007), "Race, Poverty, and Teacher Mobility", *Economics of Education Review*, Vol. 26, No. 2, pp. 145-159.
- Sclafani, S.** (2008), "Rethinking Human Capital in Education: Singapore As A Model for Teacher Development", The Aspen Institute Education and Society Program, Washington, D.C.
- Sclafani, S. and M. Tucker** (2006), "Teacher and Principal Compensation: An International Review", Center for American Progress, Washington, D.C.
- Shields, P., D. Humphrey, M. Wechsler, L. Riehl, J. Tiffany-Morales, K. Woodworth, V. Young and T. Price** (2001), "The Status of the Teaching Profession 2001", The Center for the Future of Teaching and Learning, Santa Cruz, CA.
- Steele, J., R. Murnane and J. Willet** (2009), "Do Financial Incentives Help Low-Performing Schools Attract and Keep Academically Talented Teachers? Evidence from California", *NBER Working Paper No. 14780*, National Bureau for Economic Research, Cambridge, MA.
- Stinebrickner, T.** (2001a), "Compensation Policies and Teacher Decisions", *International Economic Review*, Vol. 42, Issue 3, pp. 751-779.
- Stinebrickner, T.** (2001b), "A Dynamic Model of Teacher Labor Supply", *Journal of Labor Economics*, Vol. 19, Issue 1, pp. 196-230.
- Toch, T. and R. Rothman** (2008), "Rush to Judgment: Teacher Evaluation in Public Education", *Education Sector Report*, Washington, D.C.
- Todd, P. and K. Wolpin** (2003), "On the Specification and Estimation of the Production Function for Cognitive Achievement", *Economic Journal*, Vol. 113, pp. F3-F33.
- Urquiola, M. and E. Vegas** (2005), "Arbitrary Variation in Teacher Salaries: An Analysis of Teacher Pay in Bolivia", in E. Vegas (ed.), *Incentives to Improve Teaching: Lessons from Latin America*, The World Bank, Washington, D.C., pp. 187-212.
- Vegas, E. and I. Umansky** (2005), "Improving Teaching and Learning through Effective Incentives", in E. Vegas (ed.), *Incentives to Improve Teaching: Lessons from Latin America*, The World Bank, Washington, D.C., pp. 1-20.

CHAPTER 6

Incentive System Implementation

Stakeholder engagement and communication strategies are essential to successful compensation reform initiatives	153
• A range of communication strategies is needed for different purposes and different audiences	153
• Stakeholder engagement helps to ensure programmes reach consensus on core purposes and create a unified vision for the plan	153
• Stakeholder engagement efforts help to ensure the system will be accepted and lead to desired results	154
Potential consequences if educators are not engaged in the formation of policies on compensation	154
• Educators may purposely design pay plans so that rewards are unattainable	154
• Teachers may vote to withdraw from the compensation programme	155
Desirable practices in communication and stakeholder engagement	155
• Create a representative compensation reform committee	155
• Build support for the reform effort	158
Lessons learnt	161
Developing a feedback loop and making mid-course corrections	163
How different countries and localities engage teachers and their unions in the development of incentive systems	163
Conclusion	166
Notes	166
References	167

As the preceding chapters in this book have shown, the design and implementation of performance-pay systems for educators are complex endeavours that affect key stakeholders in variable and nuanced ways. No matter how carefully designed and thoughtfully implemented, a new compensation system for teachers and school leaders can be controversial, and in any case, is bound to produce questions and concerns. As such, no discussion of the attributes of an effective teacher incentive system would be complete without considering ways to engage educators themselves in its design and implementation or without suggesting strategies for ongoing communication about the system.

Experience has shown that securing teacher buy-in and commitment to a new compensation system is essential to its success as well as its long-term sustainability. Further, ongoing and targeted communication is critical to ensure programmatic success. However, those responsible for developing new teacher-pay plans do not always seek teacher and administrator involvement or keep educators fully informed about ways in which changes to the compensation system will affect them. Teachers, school leaders, and their associations and unions are not likely to support new pay plans under these conditions, and may actively work to dismantle them.

Increasing educator participation in school reform initiatives, including the development of new policies on compensation and incentives, is an issue that has gained worldwide interest. In Latin America, education leaders recently affirmed that improving the educational systems in their countries would require an expansion of teachers' roles so that they were active participants in all aspects of education, including the definition of local and national education policies:

Without the participation of teachers, changes in education are impossible. Time and again, the teacher question is cited as one of the most important in order for changes to take place and for them to be expressed in improved learning of children and young people, better school management, and improved effectiveness of education systems. These are the trusts of the declarations of the Regional Education Project for Latin America and the Caribbean (PRELAC), approved by the ministers of education of the region in 2002. The project defines as one of its five strategic focuses that action must be based on strengthening the active participation of teachers in order to meet student learning needs, fostering their involvement in changes, and contributing toward transforming education systems (Robalino Campos, 2005).

Teacher participation in the formulation of local and national policies on compensation varies widely between and within countries. The purposes of this chapter are to describe desirable practices in communication and stakeholder engagement and to provide guidance as to how policies on teacher compensation should be developed to ensure stakeholder involvement and commitment. The chapter begins with a brief discussion of why it is important to build stakeholder buy-in and commitment and the consequences that can ensue when education and policy leaders overlook this critical step. The chapter then provides a summary of desirable practices in communication and stakeholder engagement that have previously been described by Koppich and Prince (2007). This is followed by an overview of the ways in which various countries and localities engage educators and their associations in the development of incentive systems. Throughout the chapter, examples are drawn from teacher incentive programmes in the **United States (U.S.)** as well as other countries that participated in the OECD-Mexico workshop in December 2008, such as **Mexico** and **Chile**. The chapter concludes with recommendations for designing and implementing incentives systems that teachers and school leaders will be likely to support and value, and which are likely to be programmatically and fiscally sustainable.

STAKEHOLDER ENGAGEMENT AND COMMUNICATION STRATEGIES ARE ESSENTIAL TO SUCCESSFUL COMPENSATION REFORM INITIATIVES

Appropriately engaging and systematically communicating with stakeholders is important to the successful design and implementation of a new educator incentive plan. The participation and acceptance of key groups are crucial to developing a performance-pay system that is supported by the larger community and likely to be sustained within the school environment. Recognising this at the start of an anticipated reform effort is important; planning to incorporate key stakeholders from the outset is also a most reasonable action.

As there is a considerable range of stakeholders to include and keep apprised of progress, a variety of communication strategies must be considered. The inclusion of multiple stakeholder groups helps to create a holistic and unified vision of the purposes of the plan. Further, engagement and communication efforts help to solidify acceptance of the plan and ensure it leads to desired results. Strategies and international examples related to each of these are provided below.

A range of communication strategies is needed for different purposes and different audiences

A comprehensive stakeholder engagement and communications plan will include specific communication strategies to address such important items as:

- Core purposes of the new compensation system (e.g. to reward most effective teachers, to identify low-performing teachers in order to target professional development and encourage them to change practice in ways that will improve student achievement, to attract and retain teachers in high-need schools, and to attract potential teachers of subjects that are difficult to fill).
- Expectations (e.g. what good teaching looks like and standards for what teachers should know and be able to do).
- Details of the reward system (e.g. who is eligible, level of performance that must be attained, how performance will be measured, how big awards will be, etc.).
- How teacher performance will be evaluated (e.g. combination of student outcomes, teacher competencies, additional roles and responsibilities).
- Desired and achieved results (e.g. communicating evaluation results to teachers and providing feedback on ways to improve, communicating school performance to public for accountability purposes and building support for programme sustainability).

Stakeholder engagement helps to ensure programmes reach consensus on core purposes and create a unified vision for the plan

Reaching consensus on core purposes for educator compensation reform initiatives is a challenging task. Romaguera (2008) describes the situation in Latin America by stating that:

Reforming salary structures is difficult to achieve because the education sector involves multiple stakeholders with significant political weight. There may also be strong controversies over which style of educational policy should be adopted. These discussions are often marked, at least in Latin America, by strong ideological positions, which make agreements very difficult. Implementing educational reforms and plans to introduce incentives frequently requires union agreements and even political agreements. Therefore, governments need both a high negotiation capacity to move forward with the reforms, and a strong political will to maintain the priority of the educational agenda.

Highlighting the challenge of reaching consensus around politically charged school reform initiatives such as performance pay, Santibáñez *et al.* (2007) offer the following insight:

Mexico was one of the first countries to explicitly tie the salary of all public school teachers to teacher performance. *Carrera Magisterial* is a national, centralised program, operated jointly by the Ministry of Education (SEP) and the teachers' union (SNTE). Perhaps due to the tension between these two actors and their differing visions, CM's nature is ambiguous. It lies somewhere between a merit-pay system and a new salary schedule.

According to Robalino Campos (2005), across Latin America, education systems have operated using planning teams that determine what their schools, communities and societies need, based on their academic and social experiences. These planning teams are an example of how engaging key stakeholder groups in the design of national policies can ultimately lead to local policies likely to gain acceptance among all of those involved in education.¹ Robalino Campos contends that “these processes are not without their difficulties, and although ‘difficult’ doesn’t mean ‘impossible,’ school leaders must recognise that democratic consultation and decision-making processes shake the very foundations of systems built upon the principle of verticality”.

Highlighting the need to consider more inclusive models of collaboration and communication, Robalino Campos (2005) argues that:

Collective consultation and decisions require efforts that bring discussion and decision closer together at everyday levels in order to make possible the participation of all those involved by integrating schools, teachers, and families. This will entail providing opportunities at local, provincial, and state levels so that national policies have meaning for those who will be responsible for carrying them out within diverse conditions and contexts.

Stakeholder engagement efforts help to ensure the system will be accepted and lead to desired results

Milanowski *et al.* (2007) find that it is important for evaluation and compensation systems to incorporate tools and feedback for the educators being evaluated. Across all OECD nations, a primary purpose of incentive systems is to motivate people to improve performance. Milanowski *et al.* emphasise that teachers must be able to use data linked to the evaluation and compensation systems to identify needs for future professional development. As they point out, improving classroom performance is frequently a matter of skill and focus, not just will and effort. Confirming this practical knowledge, behavioural science research suggests that “performers need to know what the specific performance expectations are, how well they are doing, and how to get better. Research on teacher evaluation also suggests that receiving high-quality feedback and coaching improves the acceptance of the evaluation process by those being evaluated” (Heneman and Milanowski, 2003).

POTENTIAL CONSEQUENCES IF EDUCATORS ARE NOT ENGAGED IN THE FORMATION OF POLICIES ON COMPENSATION

As the previous section illustrated, there are many reasons why it is important to include a variety of stakeholders in the design process and to incorporate communication strategies into all stages of programme implementation. This section focuses briefly on the consequences of failing to engage key stakeholders in the design of an incentive system.

Educators may purposely design pay plans so that rewards are unattainable

In the **U.S.**, an early iteration of **Florida's** performance-pay programme, Special Teachers are Rewarded (STAR), provides an example of how lack of stakeholder engagement early on leads to trouble later. Studying the

implementation of STAR, Hegarty (2003) found that union negotiations were a primary factor leading to the development of district plans that awarded few teachers. Quoting a union leader who was brought in late in the game, the goal of the union representative was “to make it nearly impossible to earn a bonus”.

Teachers may vote to withdraw from the compensation programme

Also in the **U.S.**, in the **Minnesota’s** state-wide performance-pay programme, Q-Comp, experienced attrition problems when teachers and school leaders were not kept apprised of key fiscal issues. For example, the union that represents teachers in one school system voted to withdraw from the Q-Comp programme when district officials sought to lower the amount that teachers would be eligible to earn (Patterson, 2009).

In other instances across the globe, educators have rejected or returned bonuses or have gone on strike when they were not fully included in plans to develop and implement alternative compensation systems. With a more purposeful approach to stakeholder engagement and the implementation of a quality communication plan, these negative consequences can be substantially minimised. Having made a case for the importance of engaging and communicating with stakeholders, the next section now turns to specific methods for carrying out this important work.

DESIRABLE PRACTICES IN COMMUNICATION AND STAKEHOLDER ENGAGEMENT

This section presents a set of desirable practices in communication and stakeholder engagement that are drawn from a guidebook chapter developed by Koppich and Prince (2007) for the U.S. Department of Education’s Center for Educator Compensation Reform. The work was primarily intended for educators and policy makers in the **U.S.** who were interested or actively engaged in developing performance-based compensation plans for local school districts. As such, adjustments would need to be made to the suggestions that follow if a performance-pay plan is being developed at the state/regional or national level.

Appropriate adjustments would also need to be made to these suggestions based on country-to-country differences, such as the degree to which educational decision making is centralised or decentralised, and the goals and objectives of the country’s specific teacher incentives and stimulus system. As Romaguera (2008) notes, there is no one best approach to follow when introducing educator incentives as an element of education policy, and assuredly experiences will differ from country to country. On the whole, however, we believe that the three considerations provided here are relevant to most countries and governments weighing the adoption of teacher incentives tied to performance outcomes.

Create a representative compensation reform committee

A first consideration is to assemble a compensation committee or working group that includes representatives of all major groups of stakeholders. The formation of such a group will enable a performance-pay plan to have a better chance of succeeding by cooperatively engaging school and district leaders from the outset.

Much of the hard work at the beginning of compensation reform efforts involves building a coalition for success. Some districts have made the mistake of attempting to develop and implement new teacher-compensation systems without the involvement of teachers. In other words, administrative officials, operating alone, designed the programme and then announced it to the teachers who would be affected by it. This is an approach that is nearly guaranteed to doom any compensation plan before it even gets off the ground, as witnessed, for example, by fierce opposition to previous district performance-pay proposals in **Los Angeles** and **New York City** (see Johnston, 2000; Sahagún, 2000). At the state level, governors have also made the mistake of announcing intentions to create state-wide performance-pay plans in their State of the State addresses without first consulting union leaders, state boards of education, or even their chief state school officers (see Borg, 2005; DiMassa and Rubin, 2005).

When an issue such as employees' pay is being examined, employee representatives may rightly expect to be part of the conversation from the beginning.

To be sure, involving major stakeholders in the discussion may at times be somewhat messy and rocky. But it is necessary and sometimes legally required. In nations or states/regions with collective bargaining, salaries (or any change to the salary structure) must be negotiated between the district and local teachers' union or association. Yet even in jurisdictions in which collective bargaining is not a statutory obligation, involving teachers in the design and implementation of their wage construct may prove to be a productive precursor to teacher "buy-in" to the ultimate plan.

The compensation reform committee should include a mix of policy makers, administrators and the practitioners who will be directly affected by the new pay plan. For example, a school district should consider including the district superintendent, other staff from the school district's central office such as the budget director, representatives of the local teachers' union or association, additional practicing classroom teachers, one or more school leaders, a school board member and other representatives from the community, such as business leaders.² A rationale for the inclusion of specific stakeholder groups, as well as considerations regarding committee size, leadership, and purpose, are provided below.

The superintendent. The presence of a district's chief executive officer sends a message to others on the committee and throughout the district that this issue is of considerable importance. Particularly at the beginning of the process, when the idea of alternative compensation is being tried out and tried on, the superintendent should be at the table.

Other central office personnel. Consideration should be given to other central administration personnel who might be included on the pay committee. If the superintendent intends to have a designee represent him or her at meetings following the initial one(s), that individual should be a member of the committee from the outset. In addition, consideration should be given to other district officials who might have information that would be useful to the development and implementation of the pay plan and, therefore, should be committee members. For example, it might make sense to include the director of human resources, the chief budget officer, the administrator in charge of teacher professional development, and so on. As an alternative to including these individuals on the committee, of course, each can be brought in on an "as-needed" basis.

Principals. Principals need a clear understanding of the plan's components and operation and some involvement in its design, even if the pay plan is intended to affect only teacher pay. Toward that end, the compensation committee should include at least one school principal.

Teacher union/association representatives. Even in locales in which collective bargaining is not legally required, there often is a teachers' union or association with which the district does business, formally or informally. One or more representatives of this organisation should be part of the committee. Teacher pay plans must be negotiated between the local school board and the local teachers' union in those countries and states that authorise collective bargaining. In such cases, the compensation committee may be asked to make recommendations to a labour-management contract bargaining team, which would negotiate and then submit a plan to the local school board.

Teacher approval could be obtained in a number of different ways. Teachers could be asked to vote on the proposed pay plan if it is part of a comprehensive contract package. As an alternative, teachers might be asked to approve the proposed pay plan through a special vote separate from a full contract. In some cases, the district and the teachers' union or association may prefer to establish a trust agreement or memorandum of understanding instead of incorporating the new pay plan immediately into a formal contract. These types of

alternative labour-management agreements are usually easier to modify than long-term, binding contracts. The district and the teachers' union or association could then decide when to schedule a formal teacher vote.

In the American states that do not authorise collective bargaining, the committee's decision-making authority is defined by the local school board. The committee could be empowered to make decisions or it could be authorised to provide input and recommendations to the superintendent (*i.e.* the school district leader) and the school board. Even though a formal teacher vote is not legally required in non-collective bargaining states, it is still a good idea to seek teacher opinion informally to build support for any new pay plan.

Additional teachers. Consideration should be given to having practicing classroom teachers as a part of the committee. A (small) number of these teachers should either be appointed by the teachers' union or association, or agreed upon in a collaborative way by the district and union or association.

School board members. Including a member of the governing board if there is one on the compensation committee can be helpful. Board members will need to approve any new compensation arrangement, so the more a representative of this body understands the specifics and the discussions that led to the outline of the final plan, the better.

Other community representatives. Sometimes it is useful to include representatives of other kinds of organisations on the pay plan committee. Is there an active parents' organisation whose support might be helpful? Does the community have a business organisation whose representation, or skills, on the committee might be useful to securing wider support for the plan? For example, a representative of an influential local business group may help with fundraising efforts. It can be helpful to include a human resources specialist from a local firm who is knowledgeable about compensation options. Some locations have included executives associated with various media and news organisations.

Committee size. The size of the committee is a local decision and depends, in part, on local context and circumstances. The committee should be sufficiently large to be appropriately representative of stakeholders and constituents and sufficiently small to allow deliberations to take place. As a rule of thumb, committees should generally be kept to 10-15 members so that conversations and a free exchange of ideas are possible.

Committee leadership. The leadership of the compensation committee is symbolically important. Who then should be, at least nominally, "in charge" of the compensation work? If the pay committee is a joint labour-management undertaking, the committee might have co-chairs: one district representative and one teacher representative. This is the structure that was used to develop the alternative compensation systems in settings such as **Denver, Toledo, Minneapolis, Columbus, and Douglas County (Colorado)** in the **U.S.** Alternatively, the committee might agree on a single chair, but that individual must be able to reflect the multiple views that are likely to emerge in the course of the compensation conversation.

Committee purpose. The purpose of the compensation committee and the scope of its authority should be clearly defined at the outset. In some cases, a school district may grant the committee the authority to make decisions about the pay plan. In other cases, the committee's role may be strictly advisory. In **Guilford County (North Carolina)** for example, a full-time senior director and data manager oversee the performance-pay programme with guidance from two advisory teams. The school system's executive advisory team is made up of district and school leaders, while the teacher advisory team is made up of mathematics, science and language arts teachers from across the district (Rowland, 2008). In comparison to this distributed model, educational decision making in **Mexico** is highly centralised in areas pertinent to performance pay, such as curricular adoption and pre-service education (Santibáñez, 2008).

Build support for the reform effort

As was described at the beginning of this article, crafting a new teacher compensation system is likely to produce some tension-filled conversations. Not everyone on the committee is likely to be in agreement – on the reasons for developing such a plan, on the kind of plan that should be designed, perhaps even on data or fiscal needs. It will be essential for the committee to arrive at agreement on a plan. This requires that committee members respect differing points of view and try, as best they can, to view the work from other committee members' perspectives.

In part, this means accepting that there is no single model, no one right way to structure teacher compensation. Experiences of others who have gone before can provide guidance. But research cannot yet say definitively that one approach (e.g. using only test scores or using multiple measures, or, for that matter, which measures) produces superior results. Again, selecting the appropriate mix of salary incentives is dependent on determining what is to be accomplished and making a judgment about the best means to accomplish this.

If an impasse is reached, it may be useful to bring in an outside, neutral third party facilitator to help the conversation move forward. The facilitator, ideally, should be someone who is trusted (or is able to gain the trust) of principal players and is able to get the conversation moving again. It may simply be someone on whom the union and the district can agree, who is quite knowledgeable about alternative compensation, or just someone with the skills to listen well, hear what everyone is saying, and help to find common ground.

As a committee works through these almost inevitable and certainly important issues, they are able to move on to the central task of generating support for the reform effort. Among the most important tasks for the compensation committee is developing and implementing communication strategies. Keeping interested parties informed as to the work and its progress will be important to ensuring smoother implementation when that time comes.

Thoughtful and careful planning. A compensation committee must exhibit and encourage thoughtful and careful planning of the performance incentive programme. Too often, this has not been the case with alternate compensation reform strategies. For example, in the State of **Florida**, districts had about four months to develop a plan for implementing E-Comp, negotiate it with the teachers' union, and submit it for approval to the state. The **Broward County** superintendent warned that offering such a short timeline to negotiate a plan could create tension between the district and the local teachers' union (Pinzur, 2006). A representative of the state superintendent association raised a similar concern that rushing the development of a pay plan could "create more concerns and errors than it's going to resolve" (Winchester, 2006).

The type of pay plan that the committee selects should be aligned with the reasons why the school district, state/region, or country is interested in changing the way that teachers are paid. Pay plans generally fall into one of two categories. The first type bases teacher rewards only on student achievement outcomes, while the second type bases teacher rewards on a combination of factors, such as student achievement, teacher knowledge and skills, and assumption of additional professional roles and responsibilities. As Santibáñez *et al.* (2007) noted, **Mexico's Carrera Magisterial** teacher incentive programme has five specific goals:

- (1) strengthen the social appreciation of teachers; (2) encourage teachers to improve student achievement;
- (3) retain teachers; (4) reward and recognise teachers who work in low development areas and who teach special-needs students; and (5) strengthen interest in professional development courses (*Comisión Nacional SEP-SNTE*, 1998).

Once the committee reaches agreement on the purpose of the teacher incentive system and selects the type of pay plan; ensures that rewards are aligned with the educational priorities of the school district, state/region, or country; has reached agreement on a number of details about the new compensation system, such as who

should be included; how big the rewards should be; whether to reward individuals, groups, or both; which indicators will make up the pay plan; and how performance will be measured, the committee is situated to build momentum for the plan by spreading the word.

Spreading the word. In order to build support and momentum for the performance-incentive plan, a number of strategies can be employed to keep stakeholders informed. These include the following:

- **Written materials** – A set of clear, concise, but comprehensive written materials should be designed as early as possible. In some instances, it will make sense to tailor written materials for specific audiences – teachers, principals, parents and community members. At the beginning of the project, it may also be helpful to distribute a letter signed jointly by the superintendent/district leader and the union or association president explaining the work about to be undertaken. Thereafter, regular updates are key.

For example, **Denver** developed a ProComp handbook that is available to all staff online and distributes a special ProComp newsletter to all district personnel.

- **Meetings** – School-based meetings are helpful for informing teachers and principals about the compensation work in an in-person format. One strategy that has been used (especially in small districts) is for the superintendent or some other high-level district official and a union or association official to conduct a “road show”, bringing the conversation and a question and answer period to schools so that teachers and principals have an opportunity to learn firsthand about the compensation work.

Here again, Denver holds “voluntary information” sessions after school hours in different parts of the city so that teachers can drop in at their convenience and ask any questions they might have about the pay system and how it will affect them personally. Denver has also established a hotline that teachers can call and an e-mail address that they can use to submit written questions and requests.

- **Electronic communication** – An electronic newsletter or a website dedicated to compensation is another way of ensuring widespread communication.

Denver, Minneapolis, and Guilford County (North Carolina) are examples of districts that have developed extensive websites to explain how their alternative teacher pay systems work. **Guilford County** posts an electronic newsletter to its website each month. And **Denver** has developed an online video about ProComp and an online calculator that teachers can use to project their earnings under the ProComp system over the course of their teaching careers.

Communication plan elements. All performance-pay programmes are strongly encouraged to develop a written communications plan outlining how they intend to communicate details of the new compensation plan to various stakeholders, the methods they will use, who will be responsible for developing and communicating information, and timelines for implementation. An additional consideration in the formulation and implementation of a communication plan includes the incorporation of time-sensitive goals. Performance-pay plans unfold in a series of stages throughout the course of a school year, each marked with specific targets for success. Initial stages are marked by outreach efforts aimed at ensuring that all key stakeholders are aware of important programme elements, specifically the dimensions upon which teacher and principal performance will be measured. Further along in implementation, all attention will be focused on the payout process. During this time it is imperative to establish targets for effectively communicating how all of the performance- measurement data were verified and how payout determinations were made. By linking up specific communication strategies with unique programmatic activities and by establishing success targets for each stage of communication, the compensation programme leaders enhance stakeholder awareness and bolster support. Of course, a note of caution is in order. By virtue of establishing time-sensitive targets, one runs the risk of failure. However, the benefits to be gained by increased stakeholder awareness and support should mitigate this fear and encourage compensation programme leaders to infuse time-relevant targets into the communication plan.

A sample communications plan that was developed for the **South Carolina** Department of Education's state-wide performance-pay plan is included as a model below. First, the goals of the communication plan are listed, then the action plan is provided and finally the methods for evaluating the effectiveness of the plan are discussed.

The goals of the community relations plan include:

- Establishing a regular and timely communication process between the SCTAP (South Carolina Teacher Advancement Program) office and each participating school.
- Raising the awareness of SCTAP among educators and non-educators (parents and community members) across the state.
- Using positive publicity to increase funding.
- Garnering support for the programme from state-wide stakeholders.
- Using media outlets effectively to promote SCTAP.

An effective communication plan includes multiple vehicles and methods for delivering communications. Table 6.1 illustrates the type, audience, person responsible, and the timeframe of various action plan elements.

Table 6.1

Communication strategy action plan items for SCTAP (South Carolina Teacher Advancement Programme)

Type	Audience	Person Responsible	Timeframe	Goal
Provide monthly newsletters to participating schools	Internal	Administrative Assistant	Monthly	1
Provide quarterly reports to USDE and private foundations	External	Project Director	Quarterly	3, 4
Air SCTAP updates on SCETV (South Carolina Educational Television)	Internal and external	Project Director	Annually	2, 3, 4, 5
Present SCTAP at various conferences around the state	External	Project Director/Program Specialist	Monthly	2
Present SCTAP at Rotary Clubs and other non-educational meetings	External	Project Director	Monthly	2
Sponsor SCTAP State Conference	External and internal	ALL SCTAP Staff	Annually	2, 3, 4
Establish positive relationships with journalists across the state who cover education	External	Program Specialist	Weekly	2, 3, 4, 5
Create pamphlets for local libraries and conferences	External	Administrative Assistant	Annually	2
Provide current and interactive Web site	External and internal	Administrative Assistant	Weekly	1, 2, 5
Provide training and guidance to principals on media relations	External	Program Specialist	Annually	5
Hold "Legislators in Schools Day" to increase knowledge about SCTAP	External	Program Specialist	Annually	4, 5
Train at least one career teacher from each school on calculating payouts	Internal	Project Director	Annually	4
Establish e-mail address (hotline) for questions	Internal and external	Project Director		1, 2, 4
Hold after-school Q/A meeting at each participating school	Internal		Monthly	1, 2
Conduct regional (town hall-type) meetings	External	Project Director	Monthly	2, 4
Establish a focus group of stakeholders to expose concerns in the communication plan	External	Project Director	Annually	2, 4, 5

Source: South Carolina Teacher Advancement Program Office, in Koppich and Prince, 2007.

Listed below are the goals of the South Carolina Teacher Advancement Program (SCTAP) communication plan and the methods proposed for evaluating the goals.

- Establish a regular and timely communication process between the SCTAP office and each participating school.

How evaluated: Feedback forms will be provided to all school leadership personnel and will solicit specific information on whether they perceive the communication is regular and timely. In addition, all SCTAP staff will complete a similar Likert-style survey about each participating school. The survey results will be aggregated for the SCTAP office and disaggregated for each participating school. Results from the surveys will be used to guide changes in communication policy.

- Raise the awareness of SCTAP among educators and non-educators across the state.

How evaluated: In partnership with external researchers, conduct a baseline (Y1) survey of educators and non-educators across South Carolina via random sampling to rate the awareness of SCTAP. Administer the survey each subsequent year to note if the awareness levels are increasing.

- Use positive publicity to increase funding.

How evaluated: Conduct annual budget reviews and survey those involved (or their assistants/aides) with allocating money for SCTAP to determine their motivation.

- Garner support for the programme from state-wide stakeholders.

How evaluated: Sponsor a SCTAP annual conference with a specific list of invitees; administer a Likert-style survey at the conference to gauge the support for the programme. Compare the data longitudinally.

- Use media outlets effectively to promote SCTAP.

How evaluated: Monitor the news outlets across the state, looking for a) the number of stories/pieces on SCTAP each year; and b) the percentage of positive, neutral, or negative stories/pieces. These data will be compared longitudinally.

The data sources mentioned above are analysed annually and the community relations plan changed accordingly. The Project Director conducts the data analysis and meets with the compensation committee for input on appropriate changes.³

A second example of a communication plan is outlined below for the Mission Possible incentive programme in **Guilford County**. The Mission Possible communications plan details four goals:

- Communicating about goals, policies and procedures.
- Raising awareness about goals, activities and achievements.
- Using programme success as a marketing strategy.
- Using media outlets to garner awareness and support for the programme.

The following excerpt appears in Rowland's (2008) description of the Mission Possible programme:

In an effort to share information about program components and processes (e.g., when payouts occur) with teachers and other stakeholders, the Mission Possible program staff focus a great deal of attention on communication. The program website⁴ serves as the hub for all communication. The site includes links to the Mission Possible communications plan, lists of Mission Possible schools by level, program criteria, frequently asked questions, monthly newsletters, and a glossary of terms.... The district also relies on face-to-face communication with staff at Mission Possible schools to share program information as well as to acquire first-hand knowledge about the implementation of the program at each school. Two opportunities for face-to-face communication include site-based orientations and "walk-throughs" at the Mission Possible schools. Every year, the project director and other program staff conduct site-based orientations at each Mission Possible school to review the program with school staff. The program director and staff also conduct periodic walk-throughs of the schools to identify school climate or instructional quality issues. Walk-throughs take place on an as-needed basis to carry out status-checks or to follow-up on a specific issue of concern".

LESSONS LEARNT

Through extensive study of stakeholder engagement strategies and communication plans in performance-pay sites across the **U.S.**, several important lessons emerge. These lessons learnt area summarised below.

- Choose language carefully – The language used to describe the pay plan and its intent is important. For example, the term “merit pay” has an unfortunate recent history in education. The term itself, rightly or wrongly, can raise red flags with teachers. Avoiding this term just makes sense. In addition to avoiding “hot-button terms”, it is important that the terms used be explained clearly. For example, if the proposed compensation plan uses a “value-added” calculation of test scores, this term and its application need to be adequately and accurately explained. (Readers may refer to Chapter 5 for a more detailed discussion of “value-added”).
- Give the plan a name and an identity – Many districts have selected their own names for their pay programmes. Naming the programme provides it with a local identity and, often, a shorthand way of referring to it. In the **U.S.**, **Denver’s** is called ProComp; **Minneapolis’** is the Alternative Teacher Professional Pay System; **Toledo’s** plan is called TRACS (Toledo Review and Alternative Compensation System); **Guilford County’s** is called Mission Possible; and **Columbus’** is the Performance Advancement System. Several states have also selected names for their state-wide pay programmes. **Minnesota’s** programme is Q Comp (Quality Compensation for Teachers) and **Texas’** programme is DATE (the District Award for Teacher Excellence).
- Anticipate the critics and sceptics – No matter how well planned, how clearly articulated, how thoughtfully described, there will be critics and sceptics of any new compensation system. One of the lessons from districts and states that have developed and implemented these systems is to anticipate the questions and criticisms and be ready with a response. **Minneapolis** made sure a group of well-informed teachers was ready at a moment’s notice to answer teachers’ questions and respond to critiques. **Douglas County** put a vocal teacher critic of the plan on the compensation committee (and eventually won him over).
- Communicate with the media – Developing a good working relationship with the local media is critical. Teacher pay is bound to generate a newspaper, radio, or TV story, and stakeholders will want the paper or broadcaster to write an accurate portrayal of the work. It may be useful to select a single district spokesperson, or a spokesperson each from the district and the teachers’ union or association, to be the point persons for media contact. When the pay project begins, consider a meeting of district and union or association officials with the newspaper and broadcast media editorial boards and education writers and reporters to explain the intent of the emerging work. Thereafter, periodic news releases (even if they do not result in a news story) can help to keep the press aware and informed.

By engaging the media early and often, districts and states may be able to ward off the type of negative headlines generated in **North Carolina** in 2005 when the state distributed performance bonuses to teachers in improving (but still very low-performing) schools (see Associated Press, 2005; Buchanan, 2005; Silberman, 2005). Under the state’s ABC programme, teachers whose students show one year of academic growth on state achievement tests are eligible for USD 750 bonuses. If their students exceed the target, they are eligible for USD 1 500 bonuses. However, it is possible for schools to meet or exceed growth targets and qualify for performance bonuses while still scoring at very low levels. This was the case in seven **North Carolina** schools in 2005. Fewer than 60% of the students in each of the schools had passed state tests, but their gains had been large enough so that teachers qualified for USD 0.2 million in state performance bonuses. Rather than focusing on the achievement gains that had been made, the headlines focused on rewards for low levels of attainment in “failing” schools.

The opportunities and challenges associated with engaging the media are common across the globe. Considering the situation in **Brazil**, researchers at the World Bank (2001) contend that in order to create incentives for schools, one of the dimensions by which school leaders could be recruited would be to rely upon key performance indicators such as teachers’ value-added contributions to student learning. They also suggest that states in Brazil might consider – with appropriate precautions – publishing the results of school report cards, as is the practice in the Netherlands, the **United Kingdom (U.K.)** and the **U.S.** The rationale for such action is the view that “greater transparency and performance feedback to parents can be a powerful stimulant for academic growth” (World Bank, 2001).

DEVELOPING A FEEDBACK LOOP AND MAKING MID-COURSE CORRECTIONS

An essential element in any school reform initiative, and one that must be done in collaboration with key stakeholders, is to continually appraise progress and make necessary adjustments to practice. Once major stakeholders agree to a compensation plan, a process needs to be put in place to assess regularly the extent to which the plan is being successfully communicated and implemented. Focus groups and Internet-based satisfaction surveys among teachers, and perhaps principals, are relatively easy and inexpensive ways to take the temperature of the new compensation plan.

It may be that periodic mid-course corrections are necessary in the early phases of implementation. Overwhelming demand for mid-course corrections to **Florida's** state-wide performance-pay plan led the state legislature effectively to replace the original programme in March 2007 so that it now gives local districts greater flexibility, relies less on state assessment (FCAT) scores to determine teacher effectiveness, rewards a larger pool of teachers, and increases the maximum size of the bonuses that teachers can earn (Max, 2007).

A number of modifications to **Houston's** performance-pay programme are also underway, based on lessons learnt during the district's first payouts to teachers and principals. Changes include allowing teachers to opt out of the new pay plan; increasing the size of the bonuses that teachers of pre-kindergarten to Grade 2 can earn; basing upper secondary school teacher awards on departmental student gains, not just the gains of students in their individual classrooms; and moving to a value-added measurement system that bases teacher awards on more stable, multi-year patterns of student achievement gains (see Radcliffe, 2007a; 2007b).

One of the most challenging elements of performance pay implementation involves measuring teacher and school effectiveness. Few, if any, states or districts will get this challenging task completely correct the first time around. The following excerpt (Behrstock and Akerstrom, 2008) chronicles the manner in which leaders in Houston have adapted to this challenge by creating training materials to educate stakeholders about the nuances of value-added measurement:

The details of the value-added model may be complex; however, Houston has worked hard to assist schools, teachers and principals in gaining a deeper understanding of the value-added model. Houston officials believe that schools and the public now have confidence in the fairness of the system. The district developed a series of four levels of value-added training. The district also is working on a credentialing process that will track the levels of training that staff complete, and the district will recognise schools that have a high percentage of staff who have completed all four levels of training.

In the following chapter, more will be said about the essential task of engaging in data-driven decision making to guide the ongoing implementation of performance incentive systems. Having discussed desirable characteristics of communication strategies, we next turn our attention to examples of how different countries have engaged educators and associations in the development of performance incentive systems.

HOW DIFFERENT COUNTRIES AND LOCALITIES ENGAGE TEACHERS AND THEIR UNIONS IN THE DEVELOPMENT OF INCENTIVE SYSTEMS

The historical, economic, social and political contexts of each country influence the design and implementation of education reform initiatives. Further, within education institutions, micro contexts impact the way various stakeholder groups relate to one another. **Mexico** was one of the first countries to implement an alternate compensation system on a large scale, and the evolution of this reform initiative is continually being impacted by societal contexts and stakeholder interactions. As Santibáñez *et al.* (2007) note, "to understand the climate for reform that brought about the implementation of *Carrera Magisterial* in 1993, it is important to recognise the role played by the teachers' union (SNTE) and the particular political climate of that time".

Santibáñez *et al.* explain that declining wages and deteriorating living and working conditions for teachers during the 1980s gave rise to teacher strikes and protests, and ultimately changes in leadership of the national union, the National Union of Education Workers (*Sindicato Nacional de Trabajadores de la Educación, SNTE*) (Ornelas, 2002; Ávila and Martínez, 1990). These transformations were taking place at the same time that the President of Mexico was promoting educational decentralisation (Ornelas, 2002). The new teacher compensation system that developed was the result of the SNTE's desire to increase wages as well as the Ministry of Education's (SEP's) desire to modernise education.

Mexico's new teacher compensation system was created after federal and state officials and teachers' union representatives signed the National Agreement for the Modernisation of Basic Education (*Acuerdo Nacional para la Modernización de la Educación Básica, ANMEB*) in 1992. Citing Ornelas (2002), Santibáñez *et al.* report that negotiations were long and difficult, and the Ministry of Education and the teacher's union differed on several key points. The Ministry of Education wanted to include only classroom teachers in the new compensation system and favoured a true merit-pay system designed to reward the best teachers. The teachers' union, on the other hand, wanted the new pay plan to include all of its members, and favoured a salary schedule that would reward all teachers equally. The final version of the agreement included portions of each proposal, but tended to include more of the teachers' union's demands (Ornelas, 2002).

Mexican law allowed only one teachers' union, and the SNTE was a key player in Mexican education and politics since it was established in 1943. Membership in the union had grown to about one million educators by 1990, and teachers contributed roughly 1% of their wages in union dues. According to Santibáñez *et al.* (2007), SNTE enjoyed both political and economic power because of its large membership base and the dues that these teachers contributed. A new national reform under consideration in Mexico, named Alliance for the Quality in Education (*Alianza para la Calidad Educativa, ACE*), proposes to make some changes to CM, such as creating a new incentive programme based entirely on student achievement test results. This is currently under development and is on which the OECD is providing advice as part of its project to support quality improvements in the Mexican education system. It will be worth noting how new reform strategies build on prior efforts and how policy makers will engage the SNTE moving forward.

This challenge is certainly not isolated to Mexico. In **Chile**, for example, negotiations with teachers' unions have resulted in a number of agreements intended to improve the quality of teaching, although implementing these political reforms has not always been easy. According to Romaguera (2008):

The union negotiation processes that have taken place since the return to democracy and particularly since 1995 have allowed for a substantial improvement in teachers' salaries. They also have provided some space for discussing and introducing educational policy modifications with greater reach and, in particular, a limited incorporation of incentives policies within the teachers' salary structure.

For example, both the level and the structure of wages were modified as part of the negotiations between the teachers' union and the government. Moreover, both individual- and school-level incentives have been introduced in the last few years, along with a system of teacher evaluation, although the incentives are still relatively small and few teachers receive them.

Romaguera concludes that:

With regard to school group incentives, the National System for Assessing Public School Performance (*Sistema Nacional de Evaluación del Desempeño Docente, SNED*) shows the importance of all aspects of design and implementation. Schools and teachers must understand the evaluation process itself, how to go about improving the topics under assessment, and must consider the comparisons between schools to be fair. In the case of SNED, a methodology of homogenous groups is applied; other countries use regression techniques. In terms of the indicators, it is important to always remember

the complex and multi-sided nature of education, though there should be a strong emphasis on the results of student learning. In this regard, the SNED combines indicators of improvements over time (schools with high results present more difficulties progressing) with indicators of level of achievement. Finally, the process must be transparent and free of any discretionary elements, particularly given that there are monetary stimuli involved.

Each of these design and implementation elements highlights the importance of communication between the teachers' trade union and the government.

Brazil's attempts to establish a performance-based bonus system have resulted in a host of public relations challenges. As Brooke (2008) points out:

Opposition to the annual bonus has been strong. The most vociferous opposition, from the official teachers' union, has expressed a rejection of the "abusive transfer of market logic" to education (Cavaliere, 2003, p. 30) and the competition between schools and school principals that the policy is seen to foster (CDEP, 2006). Other criticisms have had more to do with the problems of cheating, there being opportunity for the teacher to help pupils fill in the answer sheets or to doctor pupil answers given that teachers are asked to apply and correct the tests. These cases are considered "well known" among state school teachers (Dissidência, 2006) as are the bureaucratic problems associated with making the individual teacher calculations and payments. A further argument, apparently well justified, concerns the use of average pupil performance results without any attempt to control for socioeconomic differences between schools or for the impact of prior learning levels.

Brooke (2008) concludes that despite good intentions, **Brazil's** bonus programme never fully achieved its intended purpose of rewarding staff for exemplary school performance. Among the criticisms leveled at the programme was that:

The Secretariat has been accused of a lack of transparency over the way in which the bonus is calculated (Dissidência, 2006) which would also contradict the idea that it is intent on establishing performance as a clear-cut component of teacher pay. It would seem, therefore, that without any statement to this end or any explanation as to whether the results were deemed satisfactory or not, the accountability program that was started in 2001 has effectively ended. The continued use of minimum teacher attendance rates or the use of teacher attendance at training courses, as included among the 2005 criteria, are a reflection of ongoing problems of teacher absenteeism rather than an expression of policy to attribute responsibility for school results to teachers or to other members of school staff (Brooke, 2008).

According to the World Bank, only a systemic approach to teacher policy can achieve change of the breadth and depth needed by **Brazil**. To that end, members of the World Bank have argued for greater incentives for innovation and excellence in the classroom and stronger links between teacher salaries and performance in Brazil:

This means developing a full, multi-pronged strategy but progressing in alignment with capacity realities, which vary from state to state. Success will depend on creating horizontal synergy between teacher development, micro and macro management, and incentives, while keeping a single-minded focus on the key goal of "learning for all children." This synergy will depend on alignment on a set of professional standards, developed in a consultation with the stakeholders (World Bank, 2001).

Each of these international examples highlights the enormous complexity involved in designing, negotiating and implementing a performance-incentive system for teachers and school leaders. These real life challenges underscore the need for ongoing collaboration between stakeholder groups and endorse the benefits of transparency, both of which are at the heart of strategic communication planning efforts.

CONCLUSION

The ideas and strategies presented in this chapter will help establish a strong foundation for engaging teachers and others in the development and implementation of new teacher compensation systems and communicating information about those systems to educators and the public. The ideas are derived largely from the experiences of districts, states/regions and countries that have paved the way. It is important to emphasise that reconceptualising the way in which teachers are paid is just one prong of a more comprehensive strategy to improve the quality of teaching and levels of student learning. Taken together with other improvement strategies, new forms of teacher compensation have the prospect of providing powerful incentives to attract and retain high-quality teachers and encourage them to take on more challenging assignments. The success of these plans will be enhanced to the extent that teachers are involved from the very beginning and are continually engaged throughout programme implementation.

NOTES

1. The Social Contract in Ecuador is a citizen-based movement that involves wide sectors of society in the development of public policy. This movement contributed to Ecuador's decision to provide universal access to the first year of basic education (preschool or kindergarten elsewhere). Participants include the Ministry of Education, local governments, the media, political leaders, and business and labour representatives.
2. If this is a state-level/regional compensation plan, the state-level/regional compensation committee should be structured much as the local one would be: with the state superintendent/regional educational leader, one or more local superintendents, representatives of the state/regional associations or unions (for both teachers and principals), and a member of the state/regional board or ministry of education.
3. For additional information about South Carolina's communication plan, contact: Jason Culbertson, TIF Project Director, JCulbertson@scteachers.org, (864) 200-0171.
4. The website can be accessed at http://www.gcsnc.com/depts/mission_possible/index.htm.

References

Associated Press (2005), "Teachers at 37 Low-performing [North Carolina] High Schools Get Bonuses", *The News & Observer*, Raleigh, NC, September 23. [In N&O archives as "Bonuses for some teachers disputed"]

Ávila, C.E. and B.H. Martínez (1990), *Historia del Movimiento Magisterial (1910–1989): Democracia y Salario*, Ediciones Quinto Sol, Mexico City.

Behrstock, E. and J. Akerstrom (2008), *Performance Pay in Houston*, Center for Educator Compensation Reform, U.S. Department of Education, Office of Elementary and Secondary Education, Washington, D.C.

Borg, L. (2005), "Union Leaders Fault Merit-pay Proposal", *The Providence Journal*, Providence, RI, January 25.

Brooke, N. (2008), "Educational Accountability in Brazil", *Revista Iberoamericana de Evaluación Educativa*, Vol. 1, No. 1, pp. 94-108.

Buchanan, B. (2005), "Low-performing Schools Still Get Bonuses", *The News-Record*, Greensboro, NC, September 24.

Cavaliere, A.M. (2003), "Entre o Pioneirismo e o Impasse: A reforma paulista de 1920", *Educação e Pesquisa*, Vol. 29, No. 1, pp.27-44.

Comisión Nacional SEP-SNTE (1998). *Lineamientos Generales de Carrera Magisterial*. México : Autor.

Comitê em Defesa da Escola Pública (CDEP) (2006), reprint of **Fórum Estadual em Defesa da Escola Pública** (2003), *Plano Estadual de Educação: Proposta da Sociedade Paulista. Fórum Estadual em Defesa da Escola Pública*, 14 de outubro, São Paulo, <http://www.adusp.org.br/PEE/PEE.rtf>, accessed December 2006.

DiMassa, C.M. and J. Rubin (2005), "CA Teachers Unions Blast Governor's Merit Pay Plan", *Los Angeles Times*, January 10.

Hegarty, S. (2003), "Teachers aren't Buying Bonuses", *St. Petersburg Times*, April 3, p. 1B, <http://pqasb.pqarchiver.com/sptimes/access/321106061.html?dids=321106061:321106061&FMT=FT&FMTS=ABS:FT>, accessed December 2007.

Heneman, H.G., III and A.T. Milanowski (2003), "Continuing Assessment of Teacher Reactions to a Standards-based Teacher Evaluation System", *Journal of Personnel Evaluation in Education*, Vol. 17, No. 2, pp. 171-195.

Johnston, R. (2000), "Merit Pay Fight, Politics Alter N.Y.C. Summer Program", *Education Week*, April 19.

Koppich, J. and C.D. Prince (2007), *Stakeholder Engagement and Communication*, Center for Educator Compensation Reform, U.S. Department of Education, Office of Elementary and Secondary Education, Washington, D.C.

Max, J. (2007), "The Evolution of Performance Pay in Florida", Center for Educator Compensation Reform, U.S. Department of Education, Office of Elementary and Secondary Education, Washington, D.C.

Milanowski, T., J. Koppich and C.D. Prince (2007), *Classroom Observations of Teacher Performance*, Center for Educator Compensation Reform, U.S. Department of Education, Office of Elementary and Secondary Education, Washington, D.C.

Ornelas, C. (2002), "Incentivos a los Maestros: La Paradoja Mexicana", in C. Ornelas (ed.), *Valores, Calidad y Educación*, Santillana/Aula XXI, Mexico City.

Patterson, G.A. (2009), "Forest Lake teachers reject merit pay change," *Minneapolis Star Tribune*, April 24.

Pinzur, M. (2006), "Teachers' Merit Pay Up for Vote Today", *The Miami Herald*, Miami, FL, February 21.

Radcliffe, J. (2007a), "Houston Teachers Could Opt Out of New Incentive Pay Plan", *Houston Chronicle*, Houston, TX, August 23.

Radcliffe, J. (2007b), "Houston Tying Teacher Bonuses to Teamwork", *Houston Chronicle*, Houston, TX, August 24.

Robalino Campos, M. (2005), "Passive Bystanders or Active Participants? The Dilemmas and Social Responsibilities of Teachers", in F. Reimers et al. (eds.), *Teacher Involvement in Educational Change*, PRELAC, Vol. 1, No. 1, pp. 7-23.

- Romaguera, P.** (2008), "Incentives and Stimuli for Teachers: The Case of Chile", paper prepared for the International OECD Mexico Joint Conference, The Quality of Education, Mexico City, December.
- Rowland, C.** (2008), *Mission Possible: A Comprehensive Teacher Incentive Program in Guilford County, North Carolina*, Center for Educator Compensation Reform, U.S. Department of Education, Office of Elementary and Secondary Education, Washington, D.C., April.
- Sahagún, L.** (2000), "Teachers Protest Merit Pay Proposal", *Los Angeles Times*, Los Angeles, CA, March 29.
- Santibáñez, L.** (2008), "Teacher Incentive Programs in Mexico: Evidence and Prospects for Reform", paper prepared for the International OECD Mexico Joint Conference, The Quality of Education, Mexico City, December.
- Santibáñez, L., J. Martínez, A. Datar, P. McEwan, C. Setodji, and R. Basuto-Dávila** (2007), "Breaking Ground: Analysis of the Assessment System and Impact of Mexico's Teacher Incentive Program 'Carrera Magisterial'", *RAND Technical Report*, RAND Corporation, Santa Monica, CA.
- Silberman, T.** (2005), "Bonuses for Some Teachers Disputed; Gains Rewarded in Failing Schools", *The News & Observer*, Raleigh, NC, September 23.
- Winchester, D.** (2006), "FCAT Bonus Plan for Teachers Wins Approval," *St. Petersburg Times*, Saint Petersburg, FL, February 22.
- World Bank** (2001), "Brazil Teachers Development and Incentives: A Strategic Framework", *Report No. 20408 BR*, Human Development Department, Brazil Country Management Unit, Latin America and Caribbean Regional Office, The World Bank, Washington, D.C., December.

CHAPTER 7

Professional Educators and Their Pay: Policy, Implementation and Alignment Issues

Overview of key educator pay issues: stabilising costs and raising student achievement	170
Education reform strategies and their results to date	173
Educator compensation reform possibilities	174
Considerations in the design and implementation of performance pay programmes	177
The influence of federal, state, district and school contexts	178
Cultivating a culture of continual appraisal and mid-course corrections	181
Components of high-quality performance pay programmes	182
21 st century educator pay for performance plans: permanent professional practices or ephemeral policy propositions?	183
Iron triangles, high politics, political champions, and transaction costs	183
Positioning performance pay on the political continuum	184
The political valence of how performance pay is practically portrayed	185
The potentially powerful, but unpredictable, impact of research experiments and experience	186
The path ahead	187
Systemic integration and alignment: the surest way to performance pay survival	187
Conclusion	191
Annex 7A: Practical preparation for performance pay programme implementation	192
Notes	196
References	197

What nobler employment, or more valuable to the state, than that of the one who instructs the rising generation?

Cicero, 1st BC

Across the globe, both the numbers of elementary and secondary schoolteachers and their inflation-adjusted salaries have substantially increased during the past half-century. In many nations, however, student achievement has remained stagnant. Two increasingly voiced policy goals, therefore, are to stabilise spending and raise performance. A central issue to consider when striving for these goals is the optimal means by which educator pay can be reshaped to contribute more forcefully to higher levels of student achievement.

This chapter addresses this global concern by analysing educator financial reward structures, describing their individual and organisational consequences, and proposing means for productive reform. This chapter addresses key policy, implementation and alignment issues inherent in incentive programmes for teachers and school leaders. It considers “pay” as a broad concept, encompassing salaries, pensions, and fringe benefits. It also considers pay as a lever for shaping individual career choices (e.g. enhancing job status and career mobility) and improving organisational effectiveness.

As the various contributors to this book argue, educator compensation is important both for individual teachers and for society. If professionals’ pay is widely perceived as poor, then societies runs the risk of attracting too few able individuals into the teaching field. At the same time, those already in the field are most likely to pursue goals that schools reward. Thus, if the reward structure is insufficiently aligned with or even antagonistic to a school’s core purposes, the risk is high that resources will be wasted and goals unfulfilled. As this chapter argues, it is important that educator pay and incentive programmes be aligned to holistic approaches to developing human capital and elevating student performance.

OVERVIEW OF KEY EDUCATOR PAY ISSUES: STABILISING COSTS AND RAISING STUDENT ACHIEVEMENT

If the objective is improving student academic achievement, there is no substitute for policies that directly relate to student outcomes.

Eric Hanushek (2007)

Incentives and rewards, including pay, shape individual choices as to whether to become an educator, how long to stay in the field, the frequency of relocation, the commitment to becoming an effective professional, career aspirations, and when to resign or retire. This chapter addresses these dimensions, offering, where appropriate, historical explanations based on experiences in the **United States (U.S.)** and international comparisons. The chapter suggests an integrated approach to educator compensation and related organisational reforms as a mechanism for ensuring both programmatic and fiscal sustainability to teacher incentive programmes. Before turning to these specifics, however, there are important contextual conditions worth considering. Two evolving conditions are forcefully propelling educator compensation as a policy priority in industrialised nations. Both conditions are related to productivity. They are: (1) demands for higher levels of student achievement and (2) the need to control rapidly rising education labour costs.

Industrial nations now face intensified levels of global commercial competition; a highly educated workforce and citizenry is vital for success. This situation is motivating nations to attempt what has never occurred before in history: education of entire populations to high academic standards. To be sure, small homogenous populations, traditionally typified by Scandinavian nations, can claim high levels of student achievement, but for larger and more diverse economies, such as in North America, Latin America, Asia and larger European nations,

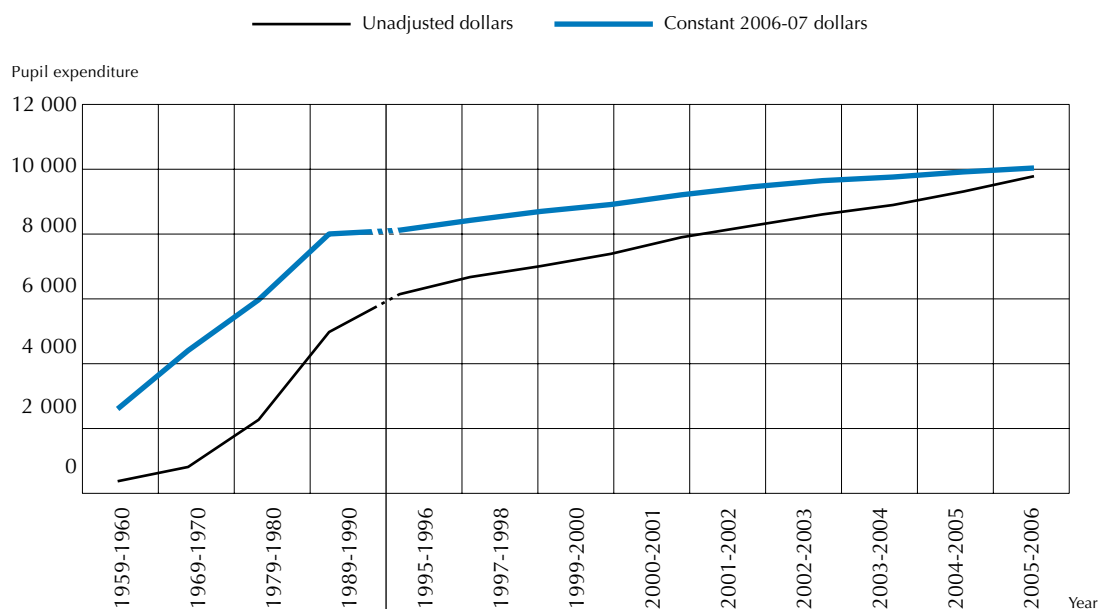
this is a new expectation. Even so, countries such as **Korea** and **Japan** have had high levels of student achievement. Certainly, across the globe, the bar for school productivity is higher than ever before and will likely be set at an even higher level in the future.

This chapter suggests that educator performance pay be conceptualised and implemented as an integral element of a holistic human capital development strategy. If approached in this manner, teacher incentive systems can serve as a vehicle to propel student learning and render school systems more fiscally efficient. In the private sector, human capital is generally defined as the accumulated value of an individual's intellect, knowledge, experience, competencies, and commitment contributing to the achievement of an organisation's vision and business objectives (OECD, 2001). Within the context of elementary and secondary education systems, the bottom line or 'business objective' is student achievement. Thus, in public education, human capital encompasses the knowledge and skill sets of teachers and school leaders that result in increased levels of learning for students.

The second condition, one confronting the full spectrum of OECD nations, is ever-rising school costs. As an example of this global trend, Figure 7.1 displays a five-decade-long view of ever-upward, inflation-adjusted, per-pupil spending in the **U.S.** In fact, with the exception of a two-year period in the midst of the Great Depression and a repeated downturn in the midst of World War II, the U.S. has for six decades never seen a time in which year-over-year adjusted per-pupil spending declined. Similar trends in rising education costs can be seen in Latin America, Europe and Asia.

Figure 7.1

**Current expenditure per pupil in fall enrolment in public elementary and secondary schools:
Selected years, 1959-60 through 2005-06**

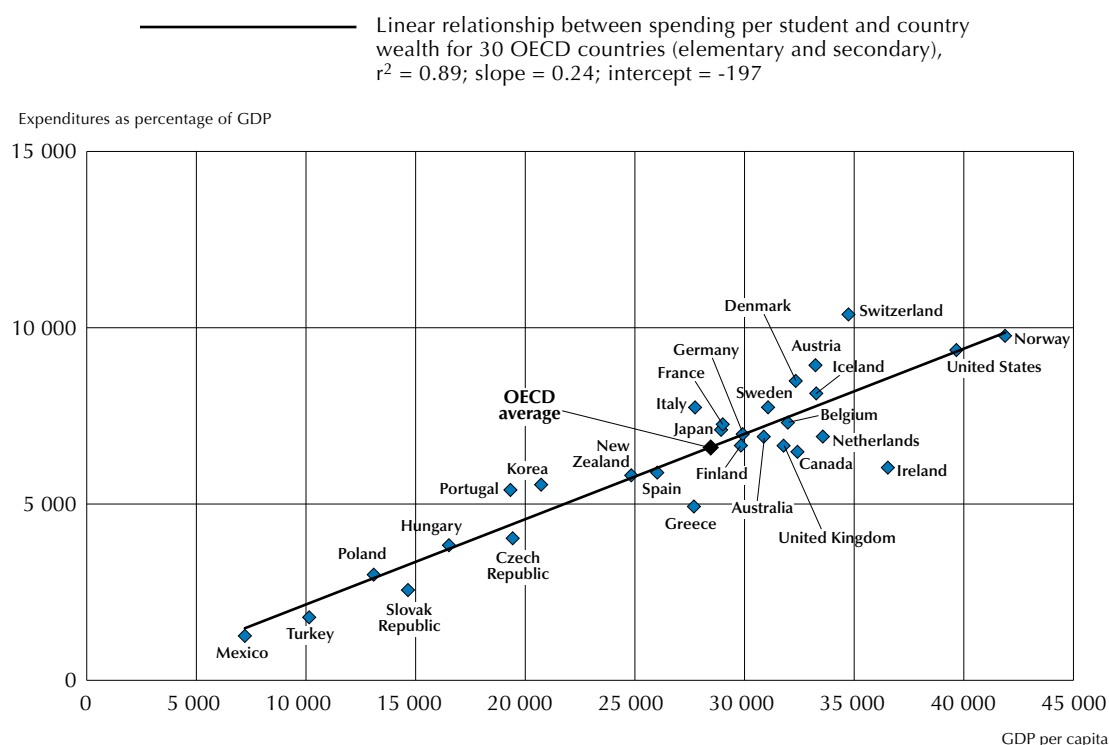


Source: National Center for Education Statistics, 2006.

Additionally, Figure 7.2 illuminates the linear relationship between per student spending and the per capita wealth of OECD nations. That is, the more money per person a nation generates, the more it spends per pupil on education. Figure 7.3 highlights the percentage of total public expenditures allocated to education in selected OECD countries. Here we notice some interesting trends. Most notable is the case of **Mexico**, which allocates the highest proportion of its public expenditures to education (twice the OECD average), yet has the lowest per pupil spending of the OECD countries, reflecting Mexico's low GDP per capita (the lowest in the OECD). In addition, we see eight nations allocated a lower percentage of all public spending to education in 2005 than in 2000: **Chile, Korea, the U.S., Switzerland, Sweden, Portugal, Hungary and France**.

Figure 7.2

International comparisons of expenditures for education



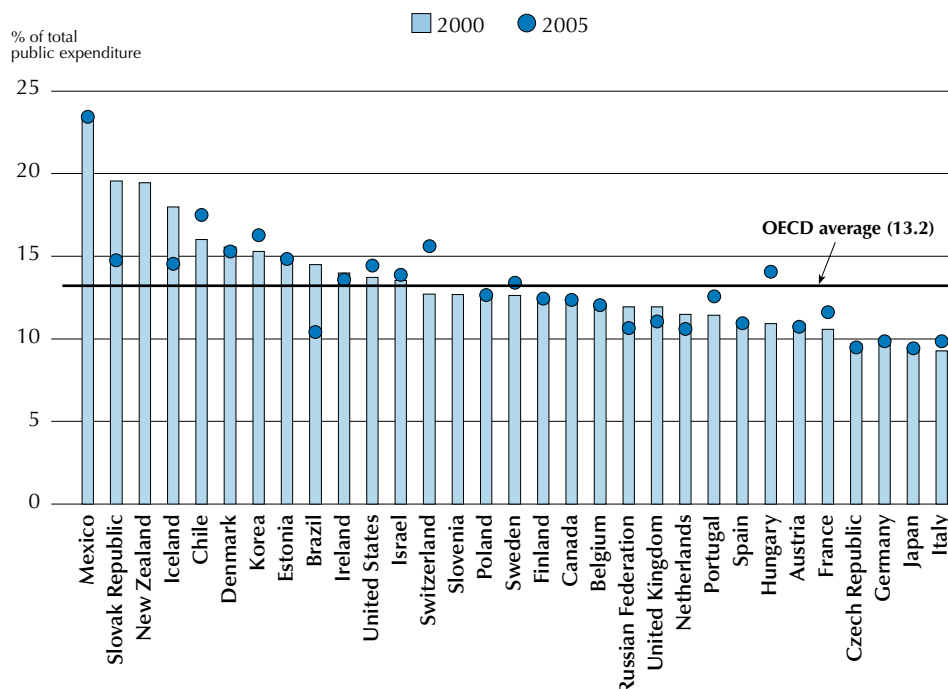
Source: National Center for Education Statistics, 2009.

The principal reason costs continue to escalate is that education is remarkably labour intensive and becoming more so; whereas other sectors, such as communication, finance, agriculture, retailing and manufacturing, long ago began to supplement labour with less expensive forms of capital. For example, bolstered by enhanced machinery and technology, each American farmer now feeds approximately 75 people, a number that continues to climb every year. Such reliance upon new methods of production has not yet occurred in the field of education. This is not for lack of attention. Across OECD countries, the education sector has been the target of many reform strategies. None has so far succeeded, however, at enhancing the efficiency of instructional delivery and few have directly targeted the issue of labour productivity.

Figure 7.3

Total public expenditure on education as a percentage of total public expenditure (2000, 2005)

Direct public expenditure on educational institutions plus public subsidies to households (which include subsidies for living costs), and other private entities as a percentage of total public expenditure, for all levels of education, combined and by year



Countries are ranked in descending order of total public expenditure on education at all levels of education as a percentage of total public expenditure in 2005.

Source: OECD, 2008.

EDUCATION REFORM STRATEGIES AND THEIR RESULTS TO DATE

A multitude of reform strategies can be found across the education landscape of every nation. Figure 7.4 describes several of these. Each nation has experimented with its own versions of these reform strategies as well as other efforts for rendering schools more effective. The strategies on this list are not mutually exclusive, and a proponent of any one of them might proclaim that the full strength of the treatment is unknown because it has never been widely or fully implemented. This is true for reform efforts such as school choice and competition or accountability where there has never been much of the former and few consequences attached to the latter. Perhaps the most costly of these common reform strategies, has been a continual reduction in pupil/teacher ratios. This could be called a personnel saturation strategy. Schooling's labour intensity is of necessity at the heart of any effort to elevate academic achievement and increase educator productivity. The following sections deal with reform efforts targeting these important issues.

Figure 7.4

Examples of education reform strategies

-
- **Intensification of academic requirements** (high school graduation, college admission)
 - **Education finance** (intra-state and intra-district equal spending, smaller classes)
 - **Governance changes** (mayoral takeover or decentralisation)
 - **Curricular and instructional alignment** (goals, textbooks, curriculum, and tests aligned)
 - **“Professionalisation” of teachers** (more pre-service preparation, career ladders)
 - **Accountability** (sanctions related to student achievement results)
 - **Market solutions** (vouchers, charter schools, outsourcing of services)
 - **School-based solutions** (small learning communities, schools within schools)
 - **Out-of-school aid to students** (health, housing, nutrition, supplemental services)
 - **Technology** (laptop programmes, online materials, distance learning)
-

Few education policies or operational practices are as complex or as controversial as those concerned with personnel, particularly compensation. Why is this so? Why are educator compensation issues often so contentious? Might personnel remuneration policies be changed to better fulfill goals, including meeting employees’ career aspirations, schools’ organisational goals, and societal priorities?

The answers to these questions are complicated. From an economist’s standpoint, if individuals are seeking, accepting, and remaining in teaching positions, then they are being paid sufficiently. If they were not, or perceived themselves as better off in another endeavour, presumably they would leave and find other employment. Of course, this is an overly simplified answer. Labour markets are not all that perfect in their operation. There are myriad other considerations that influence individuals to persist in teaching, such as job location, family ties, security, affection for children, the reward of teaching in and of itself, etc. Given the substantial employment security in teaching (few layoffs), a nine-or ten-month working schedule, generous fringe and personal benefits, and the protections afforded by the single salary schedule, teachers would appear to be paid well in the U.S.

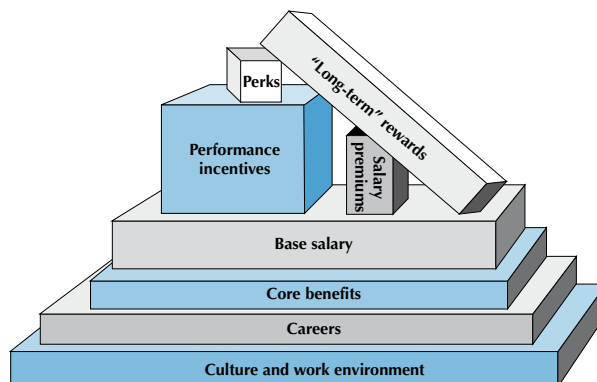
EDUCATOR COMPENSATION REFORM POSSIBILITIES

As illustrated in Figure 7.5, there are multiple aspects of employee compensation. The complexity and interactions of pay, working conditions, long and short-run benefits, and overall organisational culture or contexts are depicted here. This suggests that salary, while certainly significant, is not by itself the only important piece of the remuneration mosaic. The following discussion addresses a range of significant teacher compensation components, in addition to base salary.

Working conditions. Schools, almost no matter how challenging the working conditions, are not coal mines, steel mills, or even product assembly lines. On the other hand, they do not have the prestige and comfort of expansive corporate offices. Also, despite the periodic media blitzes to the contrary, schools are quite safe from personal violence and physical injury. Finally, teachers are responsible for a smaller number of students than in the past, due to consistent class-size reductions and the expansion of teaching staff over the past half-century. What then is there not to like about the working environment in a public school?

Figure 7.5

Components of teacher compensation package



A primary problem is work isolation. The independence of schools and autonomy of classrooms means that classroom teachers are separated from other adults and teaching colleagues. Most other professionals operate in teams: physicians with nurses, lawyers with colleagues, engineers with technicians and contractors, etc. Efforts at structuring teaching around teams have episodically been tried, but seldom persist. At least striving to schedule teaching teams around the same preparation period or lunch period would seem to assist in breaking down the work isolation problem.

Professional benefits. Small measures may have a large impact on teacher satisfaction. Progressive districts provide teachers with professional benefits such as business cards. Each month a selected teacher may be provided with a particularly convenient parking space. More effective, although not necessarily more senior teachers, may receive their pick of classrooms or more favourable teaching times. Clerical support in communicating with parents can be helpful for teachers, particularly teachers attempting to gain parent engagement in a student's learning. Recognition in district or school newsletters or other communiqués of outstanding teacher actions or unusual efforts at professional improvement are worthy of comment. Provision of an Internet service and an email address may also reinforce a sense of professionalism. Finally, it is an unusual teacher who has an office, other than his or her classroom. Similarly, many teachers do not have regular access to a business telephone. All of these are relatively low cost investments that hold the potential to reinforce a sense of professionalism among classroom teachers.

Performance incentives and salary premiums. This is a dimension on which most public school employee salaries in the U.S. are notably deficient. There is little ability for an individual educator in a public school system to earn more based on their teaching skills or the improved performance of their students, however the latter is measured. The absence of performance pay and salary premiums is one of the principal reasons why annual teacher pay is relatively low as compared to other professionals, particularly those in the private sector (Ballou and Podgursky, 1997). Figure 7.6 provides a comparison of across-the-board raises and performance-based-pay raises in the education sector. This comparison highlights the manner in which performance pay promotes multiple activities linked to a core goal of OECD and other nations' education systems, namely the academic growth and achievement of students.

Figure 7.6

Comparison between traditional pay raises and performance pay bonuses

Across-the-board raises	Performance-pay bonuses
Funds are not linked to the most important outcomes of schooling	Funds are directly linked to the most important outcomes of schooling
Do not provide motivation or rewards for elevated levels of school or teacher effectiveness	Do provide motivation or rewards for elevated levels of school or teacher effectiveness
Do not encourage the continued professional development of teachers and principals	Do encourage the continued professional development of teachers and principals
Do not provide impetus for schools and districts to align their resources with their core goals	Do provide impetus for schools and districts to align their resources with their core goals
Endorse the status quo	Challenge the status quo
Do not help schools and districts attract and retain highly effective teachers and administrators	Do help schools and districts attract and retain highly effective teachers and administrators
Do not encourage schools to consider how to effectively assess student learning across multiple grades and subject areas.	Do encourage schools to consider how to effectively assess student learning across multiple grades and subject areas.

There are numerous proposals for overcoming the disparities noted above. Figure 7.7 summarises the various means by which pay-for-performance premiums can operate. Note that rewards can be for an entire school, a team of teachers within a school, or for individual teachers. Of course, these arrangements could extend to an entire district, region or an entire nation. Note also that rewards can be for a mix of conditions including raising levels of student performance, instructing or administering in hard-to-staff schools or subjects, or meeting school or district goals for individual or group professional development. Finally, these rewards can apply to classroom teachers alone or can also be aimed at non-instructional staff, including administrators and other members of a school's staff.

Figure 7.7

Comparison of compensation systems

Alternative educator remuneration strategies	Target: individual or group	Illustrative performance measure(s)	Possible form of reward	Strengths	Weaknesses
Whole-School Reward (inclusion of classified employees optional)	Group	Student test scores Student attendance Teacher attendance	Annual bonus	Reinforces collaborative effort	Free-rider problem
Specialists/Teaching-Team Reward (e.g. all math teachers in a school, a district, a region, or a state)	Group	Student test scores Student attendance Teacher attendance	Annual bonus	Reinforces collaborative effort Reduces free rider problem	The larger the group the more likely there is a free-rider issue
Teacher Value-Added Reward	Individual	Student test scores	Annual bonus	Possibly enhances instructor motivation	Limited empirical measures could result in narrowing of curriculum, could foster dysfunctional competition
Teacher Knowledge and Skills Reward	Individual	Acquisition of attributes specified as abetting district or school pursuit of higher student achievement	Bonus, base salary addition or pay scale acceleration	Diminishes dysfunctional consequences of exclusive test score reliance	Lacks uniformity across school districts Not easily linked to empirically verified attributes
Teacher Appraisal-Based Reward	Individual	Peer review and superior appraisals of teacher performance and (possibly) knowledge and skills	Bonus	Diminishes dysfunctional consequences of exclusive test score reliance	Few empirically validated appraisal dimensions Risk or fear of favouritism and cronyism
Hard to Staff/Hard to Serve Schools	Individual	Market factors applied to specified teacher shortage definitions	Bonus, base salary addition or pay scale acceleration	Applies market incentives to solve shortages	Targets select schools and teachers
Teacher Career Ladder	Individual	Peer review and superior appraisals of teacher performance Student test scores Student attendance Teacher attendance	Bonus, base salary addition or pay scale acceleration	Rewards instructions Retains teacher talent	Never has lasted long in past experiments

CONSIDERATIONS IN THE DESIGN AND IMPLEMENTATION OF PERFORMANCE PAY PROGRAMMES

Experience with modern performance pay in schools across the globe suggests that the most crucial conditions associated with success are (1) high-level leadership; (2) adequate planning; (3) accurate performance measurement; (4) financial responsibility; and (5) employee and public engagement.¹

Leadership. If a district superintendent is insufficiently committed to the idea of performance pay, and unwilling to place the prestige of his or her CEO position on the line in support of change, the likelihood is great that whatever plan emerges will have a short shelf life.

Performance pay is often a wrenching change for the organisational culture of school districts and its professional educators. For change to be accepted, and to persist, persuasion and financial resources are necessary that typically only a high-level leader can command. Ambivalence – less than full leader commitment – will curtail the effort of likely supporters and be seized upon by opponents.

In addition to district, state, regional or even national support for performance pay, the school-level principal is essential for successful performance-pay programme implementation. Schools are the implementation unit of performance pay, and as such school-level leaders have significant ability to shape perceptions regarding the importance of the programme and to provide the ongoing levels of communication and coordination necessary for performance-pay programmes to succeed. Without a high degree of commitment from school-level leaders, performance-pay programmes will never reach high levels of operational transparency and functionality.

Planning. A performance-pay plan must be conceived from the outset, and continually reinforced thereafter, as a major change in a school or school system. The reward structure for employees cuts to the core of any organisation's culture. Hence, leaders must understand the need for comprehensive planning for changes to the reward system, and make the time for this process.

One cannot turn simply to a district's human resources department and request that they unilaterally alter the salary schedule. Rather, an expansive view of the operation must include the data processing department, budget office, professional development efforts, testing and measurement, collective bargaining, public information, and plans for future hiring of employees.

Resources. Few issues will undo a performance-pay initiative more quickly than a district's inability to be unable fully to honour its financial obligations to those eligible for performance awards. To ensure that there are sufficient resources to cover possible financial exposure, district budget officials must develop financial projections with generous payout assumptions to ensure that the district can cover the proffered rewards. Furthermore, efforts at the outset must be undertaken to ensure programmatic and fiscal sustainability of the incentive programme. If a long-range plan for sustainability is not communicated to stakeholders, they will likely lump performance-pay programmes into the category of fleeting fads. The demise of the merit-pay era of past decades was directly attributable to a lack of sustainable funding and inconsistent resource allocation for the programmes.

Measurement. Appraisal of student achievement, or teacher or administrator progress on other reward dimensions, is a fundamental underpinning of a performance-pay programme. Participants must have confidence that the activities they are undertaking are subject to accurate and objective measurement, if they are to develop any commitment to the objectives of the programme.

Engagement. Surprises are usually not good in a large organisation. Pay-for-performance programme success crucially depends upon teachers, administrators, and other employees understanding fully that which is being proposed, having an opportunity to shape important components, and understanding that every effort will be made to correct problems, should they emerge.

This dictum applies to members of the public as well. Hence, the planning stages of a pay-for-performance initiative should include external stakeholders, and there should additionally be an intense outreach effort to ensure that the public is informed of changes.

THE INFLUENCE OF FEDERAL, STATE, DISTRICT AND SCHOOL CONTEXTS

No school reform effort exists in isolation from important contextual influences. Certainly state, regional, district and local influences greatly impact the manner in which education policies are implemented. Performance-pay programmes for educators are no exception to this rule. This section of the chapter will discuss the influence of the following factors upon the design and implementation of performance pay programmes: (1) accountability and support systems (2) labour relations and contracts, and (3) human-resource and data-management infrastructures. This section will also discuss how these contextual factors impact design decisions such as central or local implementation, and which of these approaches are more prevalent and effective.

Accountability and support systems. Over the past decade, more and more OECD countries are holding education institutions accountable for improving student outcomes. While the specific nature and goals of heightened accountability expectations vary across nations, the governing bodies that set and monitor progress towards meeting the goals are primarily national, state or regional departments of education. Accountability systems articulate to local education agencies, such as districts and schools, what outcomes are expected for students in the system. By sending messages about what is most important to accomplish, accountability systems define what educational processes and outcomes are measured. As in the private sector, the activities and results that get measured are the ones that organisations focus resources upon, whether financial, programmatic or personnel. Since accountability systems signal what is most important for the school or district to accomplish, the system will necessarily influence any programmes that are directly linked to those goals, including performance-award programmes.

For example, the No Child Left Behind legislation (2001) in the **U.S.** has heightened the country's emphasis on educating all students to high standards. One result of this legislation has been for school systems to target resources to close the "achievement gap" between populations with different levels of socio-economic status (SES). It is no surprise that the federal government's first initiative to fund pilot programmes in performance pay² was targeted to hard-to-staff schools with elevated levels of students on free- and reduced-price lunch programmes (i.e. for low SES students).

The influence of heightened accountability expectations has been felt across OECD nations as well. For example, the Every Child Matters Agenda in **England** is leading towards increasingly personalised learning experiences for students and often requires collaboration between a range of professionals, both within and outside of schools (Hopkins and Ahtaridou, 2008). Several countries in Central America have introduced school-based management reforms to encourage greater local accountability. In **El Salvador**, the Education with Community Participation Programme brings important decisions regarding teacher quality and student performance to school-level discussions where parents and students have a voice. The programme has generated learning gains for schools in important dimensions including student academic achievement. Similarly, a school-based management reform project in **Honduras**, Community Education Project, has resulted in elevated student performance in mathematics, science and Spanish-language (Vegas, 2007).

These examples highlight the manner in which accountability expectations and systems can influence the design and implementation of performance-pay programmes. Support systems also influence the manner in which incentive systems are designed and implemented. One widespread reaction to elevated accountability requirements is for professionals to seek supports, such as professional development, in order to meet heightened expectations. A well-designed performance-pay programme will take into consideration the support systems

already in place for teachers and school leaders, determine the efficacy of these systems and, where necessary, create additional supports to which all professionals have access and that are directly aligned to the goals of the performance-pay plan. As most support systems for educators are implemented at the district and school level, their capacity to administer the necessary support will influence the effectiveness of the performance-pay programme.

Labour relations and contracts. Performance-pay programmes cut to the core of employee relations and the labour expectations outlined in teacher contracts. For decades, teacher pay has been based upon criteria that reward longevity and privilege seniority. New models of incentive pay challenge traditional bases of teacher pay and seek to differentiate teachers along a continuum of effectiveness. Due to this reality, performance-pay programme leaders must engage association and unions to negotiate elements of the performance-award programme. For example, several states in the **U.S.** have collective-bargaining statutes that require teacher contracts to be negotiated with local union leaders. Collective bargaining shapes many day-to-day operations in a school, such as the way schools are organised, financed, and staffed.

In order to design and implement reforms that will successfully improve teacher quality and raise student achievement, it is important for school and district leaders to understand the role of unions and collective-bargaining bodies. In the U.S., there have been variable responses to state and district attempts at performance pay from local unions and associations to date. The process always involves give and take. For example, one of the most well-known alternate-compensation systems is the ProComp plan in Denver, **Colorado**. This plan was negotiated between the local board of education and the Denver Classroom Teacher Association. While the association consented to levels of differentiation in teacher pay, it required that the dimensions of this approach be developed by teachers and approved by principals. A similar arrangement has emerged in Austin, **Texas**. Here, the district superintendent and teacher association president are both present when making any formal remarks about the district's Strategic Compensation initiative. Through these joint appearances, district leaders are assuring teachers and principals that their association leaders are at the table and a key part of any critical decisions about the performance-pay programme.

In contrast to the autonomy and decentralisation evident in **El Salvador** and **Honduras**, the educational system in **Chile** emphasises a national curriculum and standardised learning measures. Within this more centrally-controlled system, perhaps it is not surprising that the conflict and resistance between the Chilean Teachers Union and the Government leads to many compromises. For example, while the teacher union annually secures raises for all teachers, the government has installed a series of incentive systems related to performance, first, with the National System of School Performance Assessment (*Sistema Nacional de Evaluación del Desempeño*, SNED) in 1996, later with the creation of the Pedagogical Excellence Allowance (*La Asignación de Excelencia Pedagógica*, AEP) in 2002, and finally with the Variable Allowance of Individual Performance (*Asignación Variable por Desempeño Individual*, AVDI) in 2006. Due to a lack of collaboration with teachers in creating the incentive programmes, only a small percentage of teachers are financially recognised and the system has created competition between teachers. Researchers have found that this approach neither promotes professional development nor improves the quality of education; on the contrary, it perpetuates conflict between the centralised government and the teachers' union (Pavez Urrutia, 2008). From these examples, we see how critical these external forces are to the creation of a teacher-incentive system that will advance teacher quality and student performance while having a positive impact on the culture and climate of schools.

Human-resource and data-management infrastructures. Once accountability demands have been aligned with the performance-pay plan, and agreements with local unions and associations have been made, district and school leaders must still contend with the potential constraints of existing human-resource and data-management infrastructures. In order to effectively implement a performance-pay plan, a district must have

a data-management system that can facilitate desired linkages between teacher and administrator human-resources data, student course scheduling and demographic data, and the payroll system. Few school systems have a data-management infrastructure that can currently support these linkages on the level of individual students and teachers. For example, consider if your school, district or state or region's data system is able to provide you with a very high level of accuracy, for every teacher in a particular school, every student they teach during every class period. Is the data system able to account for student mobility? Does the data system account for team teaching? Can the data system specify individual student and teacher attendance rates during the year? Is the data system electronic, or would key teacher or student level data need to be verified by hand? Answers to all of these questions will influence the design of a performance-pay programme that can be reliably implemented and provide a high degree of accurate data by a school or school system.

As an example, many of the performance-pay programmes funded by the **U.S.** Department of Education's Teacher Incentive Fund were designed to award individual teacher payouts based on student-achievement data. However, when it came time to calculate teacher payouts, the state, district and school data systems were not able to make the necessary linkages between individual teachers and groups of students. In some cases, small districts were able to do this data verification without the help of a data management system, but in large, urban districts there were simply too many teachers and students to feasibly do this hand-checking with a high degree of accuracy. As a result, many of the programmes had to resort to awarding payouts to individual teachers based on school-wide levels of student performance.

Contextual factors impact design decisions. As described above, increasing accountability requirements for improved education outcomes are exerting tremendous pressure on schools to develop reform efforts that will enhance the efficiency and effectiveness of teachers and school leaders. Local bargaining units also influence the way programmes are designed and implemented. Additionally, the current reality of human-resource and data-management infrastructures influences the feasibility of implementing desired design elements. These contextual realities shape performance-pay programmes in multiple ways, including the locus of control for programmes – that is, the degree to which programmes are centrally or locally developed, and to what extent programmes are more homogenous in approach or individually tailored by an entrepreneurial school and leadership team.

The performance-pay movement in **Texas** over the past three years provides a picture of how these dynamics interact. In the fall of 2006, the Governor's Education Excellence Grant programme made available non-competitive, three-year grants to 100 schools ranging from USD 60 000 to USD 220 000 per year. Grants were distributed to schools that were rated as high-performing with high proportions of economically-disadvantaged students. As there was no statewide precedent for performance pay at the time, the state legislature determined that the appropriate place to start was to allow entrepreneurial leaders and schools a high degree of autonomy in implementing programmes. Since the programmes did not require district-wide participation, individual schools could tailor a programme to meet local needs while complying with broad programme guidelines. While this approach launched the statewide movement by funding 100 pilot programmes, subsequent models for allocation of funds for performance pay have moved from a campus-based approach to the most recent District Award for Teacher Excellence programme (which allocated USD 147.5 million in funding for 200 districts to implement district-wide performance-award programmes during the 2008-09 school year). This transition was made in part to take advantage of economies of scale, support systems, and human-resource and data-infrastructure elements at the district level. Additionally, calculations of teacher and school effectiveness gain a degree of statistical accuracy within a larger sample of demographically similar schools afforded by district-based programmes.

In response to the contextual variables described above, several performance-pay programmes have begun with a group of pilot schools within a district that have self-selected into the programme. The strategic compensation programme in Austin, **Texas** is an example of this approach. Having begun with a pilot cohort of nine schools in

2007-08, the programme anticipates adding several schools each year until the programme is implemented at scale in the district's 200 schools. In other cases where performance pay was initially implemented district-wide, such as **Denver's** ProComp programme, existing teachers were given the choice of opting into the programme or staying with the traditional salary schedule. In both of these cases, we see how contextual influences, such as negotiations with the local teacher association, impacted programme design and implementation.

An incentive programme in **Mexico** provides a broader international perspective on the influence of contextual variables upon incentive system design and implementation. The National Agreement for the Modernization of Basic Education reform of 1993 resulted in a significant decentralisation of Mexico's public schools from a federal system to one driven more autonomously by its 32 federal entities. One element of this reform was the establishment of *Carrera Magisterial* (teachers' career ladder), a nation-wide teacher incentive programme that is jointly managed by the Ministry of Education and the teacher union (Santibáñez, 2009). While many international performance-pay programmes have group and whole-school elements, the foundation of the *Carrera Magisterial* incentive programme is based on awards to individual teachers, signaling a commitment to the enhanced autonomy reflected in the National Agreement.

Due to the inevitable impact of external factors on performance-pay design and implementation, it is important that district and school leaders continually evaluate and revise programmes using data-driven decision-making processes. This topic is the focus of the next section of the chapter.

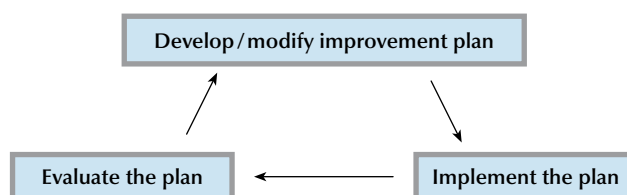
CULTIVATING A CULTURE OF CONTINUAL APPRAISAL AND MID-COURSE CORRECTIONS

The systematic collection, analysis and utilisation of data can serve as a catalyst for organisational learning. Education leaders can harness the regular information flow from data to sustain a culture of continuous improvement within their districts and schools. Data can provide leaders with continual feedback to support individual and collective learning in educational organisations. In the case of performance-pay programmes, which are complex and involve many moving parts, it is essential to continually engage in a cycle that involves the development, implementation, evaluation and re-appraisal of the plan. These activities are illustrated in Figure 7.8 below.

Once an incentive programme has been developed in collaboration with key stakeholder groups and is being implemented, it is imperative that a full spectrum of programmatic elements be evaluated to determine the impact of the programme on key processes and outcomes. These include teacher perceptions of the programme, influences of the plan on the school culture, ability of the plan to enhance retention and attraction of teachers, and improvements in student performance.

Figure 7.8

Continuous improvement cycle



Another way districts have infused data-driven decision making into the performance-pay process is to incorporate programme elements that rely on student data. For example, in order to include a broader spectrum of teachers in performance-pay programmes than those who teach subjects and grade levels in which there are standardised tests, many districts are incorporating Student Learning Objectives into their programmes. In this approach, each teacher must analyse data specific to students in his or her grade and subject area. Referring to these data in collaboration with principals or mentors (such as master teachers or academic coaches), teachers will establish two to three goals for the year. Benchmarks for success are established and specific evaluation instruments are created, if necessary. Performance-award payouts are linked to the degree to which teachers are successful in meeting the data-based goals. Districts that have incorporated Student Learning Objectives into a performance-award programme have seen substantial improvements related to student data and instructional planning.

COMPONENTS OF HIGH-QUALITY PERFORMANCE PAY PROGRAMMES

The following list displays what school, district, state, and national/federal education leaders identify as components of high-quality, performance-pay programmes. These are essential elements to consider in the construction and implementation of a performance-pay plan. Further, they provide an experience-based endorsement that there is no one-size-fits-all model for teacher-incentive systems. Rather, they indicate that with collaborative planning, local customisation, and systemic integration of a variety of models can be successful.

- The performance-pay plan is framed as a school-improvement strategy, part of the district's larger human-capital-development system, and directly aligned with core instructional goals at the district and school levels.
- Multiple, valid and reliable assessments of student learning are used to measure teacher, team, and school effectiveness.
- Orientation toward value-added measures of teacher effectiveness, coupled with use of these data beyond pay determination, such as the identification of effective practice, domains for potential improvement, and evaluations of professional development experiences.
- Inclusion of rewards other than financial, such as a supportive environment, mentoring, and strong, consistent leadership.
- Continual programme appraisal and improvement, and sustained collaboration with key stakeholders, including union and association members, supported by a comprehensive and accurate data system that ensures stakeholders that the information used to make performance-award decisions is accurate, valid and reliable.
- Multiple award levels (individual, team, school) and opportunities for choice (a resource teacher may collaborate with a core teacher to establish student learning objectives, or receive school-based award).
- Support for each performance goal through targeted professional development and ongoing training for performance assessors.
- Ongoing communication with all stakeholder groups in multiple formats targeted to specific stages of programme implementation (such as the initial programme overview, verification of measurements and test data, and the payout process).

Appendix A contains a checklist of specific questions that may be useful to schools, districts, states/regions or nations designing performance-pay programmes. The first level of decisions focuses on: who should be involved in design and implementation, how decisions will be made, where necessary resources may be obtained in the short term and for sustained implementation, when major project milestones are to be completed and how programme effectiveness is to be determined. The checklist identifies questions that should be asked at a more detailed level as specific components of the plan are developed.

21ST CENTURY EDUCATOR PAY FOR PERFORMANCE PLANS: PERMANENT PROFESSIONAL PRACTICES OR EPHEMERAL POLICY PROPOSITIONS?

For a variety of reasons, including increased accountability demands and struggling economies, teacher pay has now come to centre stage in the education policy arena. Nevertheless, given the pervasive and persistent nature of the conventional salary schedule, and negative precedents set by prior single-salary replacement efforts, presently operating compensation practices may not give way quickly. The discussion in this section focuses primarily on **U.S.** examples with implications for an international audience.

The principal question for this section of the chapter is: will contemporary educator pay-for-performance reforms persist and be adopted widely as a productive policy innovation for public education? Historical and political perspectives based on experiences in the **U.S.** suggest that the likely answer to this question has much to do with fluid factors such as what is meant by “pay-for-performance,” support of elected officials, forthcoming validation from the research community on the impact of performance pay, and the history of lessons learned from practical implementation.

Political perspectives. Pay-for-performance, given its controversial past and high visibility among the general public presence, is a potentially volatile political topic. Moreover, the degree and nature of the political dynamics that surround it may determine its future persistence.

Since 2000, performance pay has enjoyed high visibility in the **U.S.** and has had champions from the executive branch (mayors and superintendents, governors, and the United States President), as well as state and federal legislative advocates. Prominent members of both major political parties have been supportive. This support has resulted in state performance-pay regulatory mandates, government and philanthropic financial incentives, and a great deal of positive rhetoric from highly-placed elected officials. The current U.S. Administration has signaled clear support for the idea, or at least selected parts of the idea. This high degree of political support has enabled performance pay to spread among the states, mitigate several pockets of resistance, and so far, to be sustained.

IRON TRIANGLES, HIGH POLITICS, POLITICAL CHAMPIONS, AND TRANSACTION COSTS

Political processes are fluid and may be played at varying levels of intensity, engagement, and visibility. Political stages may be viewed as a continuum with scenarios involving relatively low citizen participation levels and restrained media attention and visibility on the left and high visibility and widespread public engagement matters on the right. In such a scenario, “iron triangle” dynamics are anchored on the left and so-called “high politics” reside on the right, as depicted in Figure 7.9.

Iron triangles. The “iron triangle” of politics refers to the almost unstoppable coalition of elected officials, bureaucratic managers, and interest group beneficiaries who associate in support of or in opposition to government programmes. A critical mass of interest-group representatives cutting across these three sectors can virtually assure passage and sustained funding, or failure, of a programme, as long as the issues involved are not highly visible, overly expensive, or controversial. The “iron triangle” thrives in the absence of widespread public scrutiny and outside the glare of intense media coverage. “Iron-Triangle” politics are insider politics, the politics of micro change, incrementalism, omissions, or adjustments to the status quo (Koppich and Guthrie, 1990).

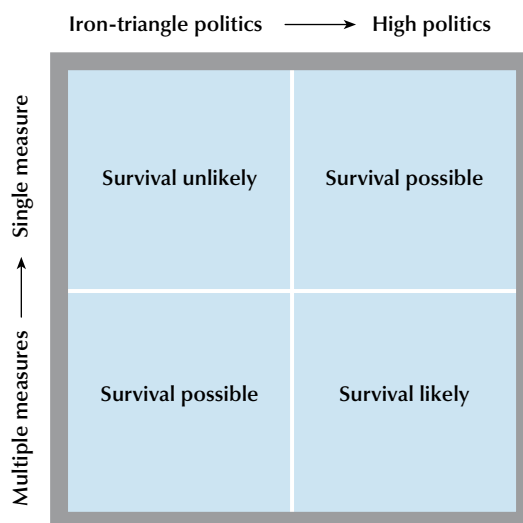
The influence of the “iron triangle” is often reinforced by the high transaction costs attached to controversial political issues, performance pay included. The perceived returns, possibly attracting larger numbers of more able teachers at a later point, and the unknown odds of obtaining higher levels of student academic performance are seldom perceived by local citizens and parents as justifying the effort and informational opportunity costs involved in advocating for and sustaining performance pay. Performance pay programmes are sustained only if there are high-level political champions who can galvanise and sustain widespread political support.

Conversely, the risks perceived by local-level, performance-pay opponents are sufficiently intense as to justify their sustained and organised political opposition to performance pay. While this discussion refers to politics within the **U.S.**, many features are applicable to other OECD countries as well.

High politics and political champions. A select few issues gain extraordinary political attention and become affixed to the agendas of major political parties, legislative bodies, and executive branch officials. These issues come to attract unusually high media attention and public awareness. This is what Kingdon (1984) describes as “high politics”. When this happens with issues such as economic growth, ecological threats, military actions, or even education policy, engagement of multiple constituencies can become sufficiently intense as to overwhelm whatever position, supportive or oppositional, of iron triangle participants. Influential political players own big policy issues, and the iron triangle can be swamped as a consequence. This is particularly so if a major league politician, be it an individual or organisation, emerges to champion a cause. In high politics, political parties or government agencies assume the high transaction costs that work against widespread local citizen participation in iron triangle issues.

Figure 7.9

Political and measurement dynamics of performance pay



POSITIONING PERFORMANCE PAY ON THE POLITICAL CONTINUUM

In the **U.S.**, present day pockets of opposition to performance-pay include professional educators and their local, state, and national political action and organisational allies and for elected allies. Political proponents are, commonly higher-level government officials, fiscal watchdogs and taxpayer advocates, and the state and national business communities.

If performance-pay political dynamics are contained within the local, state, or national interest-group-dominated conventional iron triangles of politics, then its future may be at risk. Within this sphere, narrow interests dominate, and if the narrow interests are negative, then they may prevail.

However, if for whatever reason, performance pay persists as a topic of interest and is on the action agenda of those in the higher echelons of politics (major party officials, mayors, governors, city councils, state legislatures, congress, and the President), its shelf life within the policy system may be extended.

THE POLITICAL VALENCE OF HOW PERFORMANCE PAY IS PRACTICALLY PORTRAYED

Performance pay is not articulated as a precise policy. Multiple teacher compensation arrangements may cluster under this banner. These various performance-pay arrangements, however, can be arrayed on a continuum, with their dependence upon the number and kind of performance measures and reliance on student academic achievement measures on the scale as the trigger for higher pay. This continuum forms the second axis of Figure 7.9. One end of the continuum is anchored by the example of Little Rock, **Arkansas**, in the U.S. In Little Rock's Meadow Cliff School, each student in a class carries a bounty, e.g. a dollar reward for which the individual classroom teacher is eligible if the student achieves a predetermined goal for improved test scores in a given subject matter. Teachers know the premium value of each student's scores. Their prospective performance pay reward is linked to little else. This would be considered the 'hard' end of the continuum.

Denver, Colorado's celebrated ProComp plan illustrates the other end of the hard/soft measure continuum. In Denver, a teacher's potential salary premium is only modestly influenced by direct measures of student achievement and is more a consequence of an individual teacher and her school colleagues achieving various collective goals. Below is the manner in which the 13 February, 2008, issue of TIME magazine described the Denver plan:

For Taylor Betz, the program is a no-brainer. A highly regarded 15-year veteran who teaches math in the city's struggling Bruce Randolph School, Betz can rack up an additional USD 4 268 this school year if she and her school meet all their goals. That includes USD 1 067 for working in a high-needs school, another USD 1 067 if students in her school exceed expectations on the state exams, USD 356 if she meets professional academic objectives she helped set in the beginning of the year, USD 1 067 if she earns a good evaluation from her principal [*i.e.*, school leader] and USD 711 if her school is judged to be a 'distinguished school', on the basis of a mix of criteria that includes parent satisfaction. (Wallis, TIME Magazine, 2008)

The Milken Family Foundation's Teacher Advancement Programme (TAP) occupies a midpoint on the hard/soft, simple/complex continuum. Its four teacher-related reward dimensions take into account both professional processes and student achievement outcomes. The greater a performance-reward plan's dependency upon a single teacher measure, particularly if that measure is student standardised test scores alone, the higher the opportunity for measurement error, goal displacement, and other dysfunctional conditions such as those described by Rothstein in his paper "Holding Accountability to Account". (Rothstein, 2008).

Most importantly for these analyses, the higher a performance-pay plan's reliance upon a single teacher-appraisal criterion, particularly if that criterion is a student-achievement measure, the greater the likelihood of educator hostility to the programme. Conversely, the greater a programme's reliance upon multiple performance measures, and the less weight given to standardised test scores in the equation, the higher the likelihood of educator acceptance.

Added complexity stems from conditions that render performance pay more acceptable to professional educators: multiple measures of process and minimal reliance upon measures of student outcomes. This likely dampens the enthusiasm of non-educators. Non-educator performance-pay proponents frequently desire simpler measures and almost always want to include students' standardised test scores.

THE POTENTIALLY POWERFUL, BUT UNPREDICTABLE, IMPACT OF RESEARCH EXPERIMENTS AND EXPERIENCE

Several conditions complicate efforts to calculate the potential persistence of performance pay, among which are the implementation experiences of performance pay programmes and the potential consequences of scientifically conducted research.

Operational experiences. Whereas success is difficult, failure is easy. If pay for performance is to trigger successes, such as elevated interest in teaching as a profession or higher student academic achievement, it will take time. Significant short-run successes should not be expected from such a fundamental and complicated change. Ironically, however, while the challenges to successful performance pay are many, the road to performance-pay failure is easy.

The changes in school district operation needed to support and sustain performance pay are varied. These include attention to measurement, data infrastructure, financial and accounting accuracy, public relations, and teacher engagement. Failure on any of these dimensions can severely jeopardise the success of a new remuneration system. There have already been notable failures, instances where teacher financial rewards were badly miscalculated. Indeed, there is even a school district that not only miscalculated teacher performance pay premium payments, but also, upon learning of overpayments, mandated that teachers pay the unearned money back. A critical mass of such visible miscalculations could spell doom for performance pay.

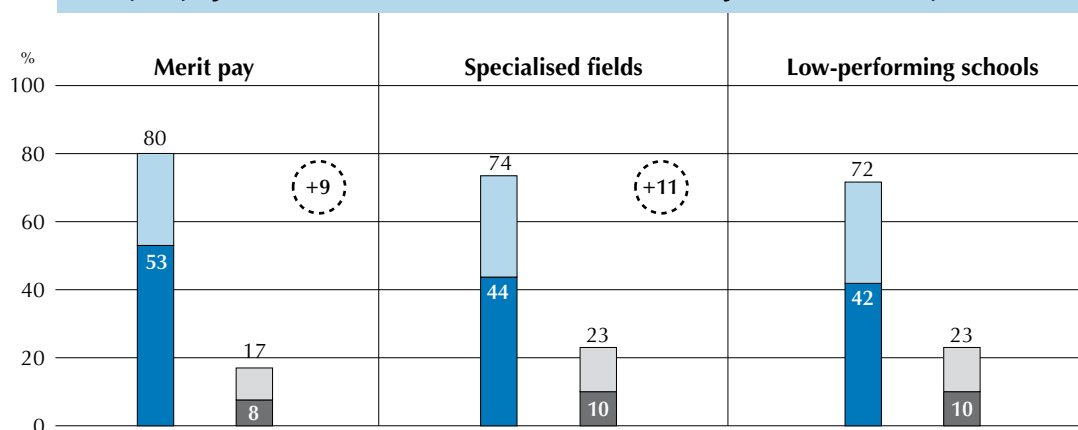
Figure 7.10

Survey results regarding support for added pay for teachers

■ Strongly support ■ Somewhat support ■ Strongly oppose ■ Somewhat oppose

Q.26-28

Do you support or oppose: paying teachers who work in specialised fields, such as math and sciences more money; paying teachers who demonstrate proven success in the classroom more money; paying teachers who choose to work in low-performing schools more money?



Source: Global Strategy Group, 2008.

Research results. Present day political support for performance pay is perched precariously in the U.S. Scientifically conducted research, conducted objectively and to high standards, could influence the persistence of performance pay. This is true whether the research results are positive or negative. In the U.S., the **Tennessee STAR** study, one of the few true experiments in education, has had enormous policy influence; in some ways too much influence as advocates have used these experimental results to gain support for smaller classes even when doing so stretched the generalisability of STAR findings (Zaharias, 1999). Regardless of the research outcome, such studies will be subjected to the most intense scrutiny imaginable as performance-pay protagonists and opponents undertake analyses. This interest, however, is testimony to the potential significance such research can play in the policy realm.

Finally, there is the matter of **popular support**. Figure 7.10 displays results of a 2008 poll displaying that the majority of the public in the U.S. is favourably disposed to performance pay. The figure highlights greater support for paying teachers who demonstrate success over those who work in specialised fields or in low-performing schools.

THE PATH AHEAD

The future of early 21st century performance-pay plans is subject to the volatile interaction of fluid political dynamics and the kind of performance pay involved. It is possible to conceive of scenarios in which performance pay survives and flourishes and other scenarios where it disappear.

Scenario One: Secure Survival (Unlikely)

If performance pay has sustained support from those at the highest levels of the policy system, then it will persist, regardless of design type or research results.

Scenario Two: Immediate Curtailment or Dramatic Modification (Unlikely)

If high-level political support diminishes or disappears, and iron-triangle politics prevail, then performance-pay opponents gain a powerful upper hand. Opponents assuredly will strive to eliminate simplified, single-measure, individualistic plans, as in Little Rock, Arkansas. This is particularly true if scientific research results unequivocally disprove the utility of pay rewards. Professional educators might possibly tolerate more complex evaluation and collective reward strategies, such as TAP and Procomp, if risks to individual teachers were perceived as being sufficiently mitigated.

Scenario Three: Modified Survival (Likely)

A more likely scenario is that political support for performance pay remains muddled, and in such circumstances the creditability of scientific research results and the type of pay plan will greatly influence the shelf life of the idea. The more variables involved in the pay calculation, the more that the plan is collective rather than individualistic, and the smaller role that student academic-achievement measures play, the more likely the strategy will persist.

SYSTEMIC INTEGRATION AND ALIGNMENT: THE SUREST WAY TO PERFORMANCE PAY SURVIVAL

This final section of the chapter addresses interconnectedness between teacher-incentive programmes and systemic human-capital-development strategies, necessary to sustain performance-pay programmes. Returning to the definition of human capital provided early in the chapter, a holistic human-capital-management system in a school district refers to how teachers and school leaders are recruited and selected, hired and placed, offered induction and ongoing professional development; how they are evaluated, compensated, provided with career-advancement opportunities and, when necessary, terminated. Considering the scope of activities listed above, and the interrelated manner in which they impact the professional lives of teachers, one realises that component pieces cannot continue to be viewed in isolation. Instead, policy makers should adopt a coordinated comprehensive approach to identifying, nurturing and sustaining talented educators for a country's schools.

When these activities are not thoughtfully coordinated, as is the case in many district offices in the **U.S.**, any number of undesirable situations can occur. Sigler and Kashyap (2008) illuminate potential negative consequences of such isolation through several realistic scenarios. For example, a district has an excellent recruitment and marketing campaign in human resources, paired with a high level of service for applicants and new-hires. At the same time, this district's office of professional development has inconsistent and poor-quality mentoring and a lack of quality professional-development options for teachers. The result for this imaginary school system, just as it is for most school systems with similar circumstances, is predictable: high turnover. Today's high-quality new-hires quickly become tomorrow's attrition statistics.

In another case, a district could have an effective office of professional development that coordinates high-quality skill-building and training options, but with no connection to the district's teacher-evaluation process. In this case, while evaluations may identify areas for growth, there is no guarantee that teachers will be connected to the district resources that might help them in those areas and, therefore, an opportunity to improve teacher quality and the level of instruction in the district is missed.

Yet another example might be a district that has no problem recruiting elementary-school teachers, but cannot attract enough lower secondary school subject teachers to meet its needs. This same district has human resources doing recruitment, while the office of teaching and learning handles teacher training and relationships with local teacher education programmes. A district like this *must* work with those teacher preparation programmes to address the inadequate supply of middle school teachers. Options include encouraging current and incoming teacher candidates to consider coursework for a middle-school certificate or creating streamlined coursework options for current district elementary teachers to become certified in lower secondary school subjects. Both of these options should be coordinated with incentives that the superintendent, teachers' union, and budget office would need to approve.

While coordination of internal district personnel across various programme areas is important, it is also imperative that solid partnerships with external stakeholder groups are developed by programme leaders as part of a comprehensive human-capital strategy. This is important for a variety of reasons related to the fact that districts have finite capacity and domains of expertise and influence. Along the human-capital-development continuum, there are activities at which district personnel may excel and should therefore concentrate their efforts in those domains. They may find it is helpful to work with quality service providers to supplement these efforts. Another way to forge partnerships around human-capital development issues is to engage local teacher association or union leaders in activities such as coaching and the provision of professional development. The partnerships help develop the trust that will enable districts to move forward with innovative, and at times controversial, school-reform initiatives such as performance pay.

As there are many moving parts to a comprehensive human-capital-development strategy, it is helpful to have a conceptual framework to guide this work. Scholars at the Aspen Institute (Wurtzel and Curtis, 2008) have developed a framework for human-capital management in K-12 education that recognises the importance of state and federal policy forces; labour-market conditions and demographic shifts; the teacher-preparation pipeline; district, school and classroom contexts; and incorporates eight essential components:

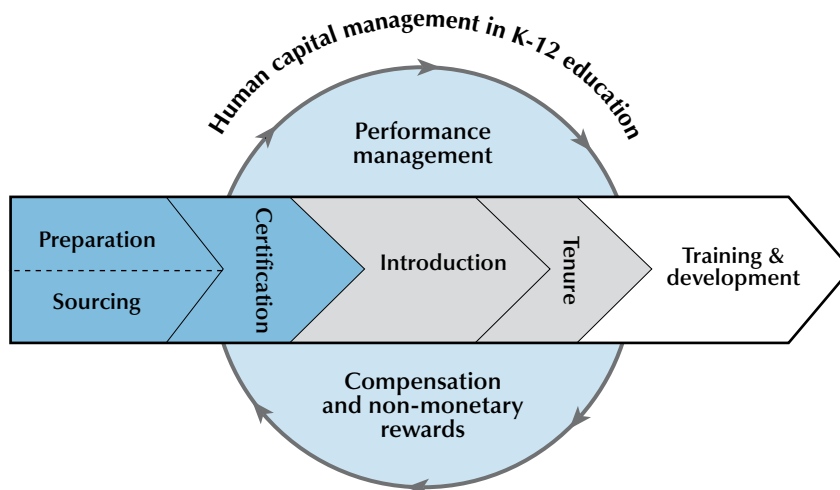
- Preparation.
- Sourcing.
- Certification.
- Induction.
- Tenure.
- Training and development.
- Performance management.
- Compensation and non-monetary rewards.

To implement a system-wide approach to developing the teaching workforce and to elevating student achievement, school systems need to be clear about their mission and goals, and identify and help prioritise the strategies that will help them achieve them. They will also need to align their district and school-support structures with their key stakeholder groups, and form partnership with outside organisations. Certainly, every nation, and each educational organisation will address these challenges in a unique manner. Those using a systemic and strategic approach will address each component individually and collectively to create a comprehensive system.

The Aspen Institute contends that while the human capital framework depicted in Figure 7.11 is organised by components, each component should be considered in relation to the others and in relation to the district's mission and goals. For example, a performance-pay system that recognises high-performing teachers and leaders requires rethinking of evaluation, approaches to assessing key outcome variables, compensation and non-monetary rewards for performance, career-development opportunities, and the creation of a professional culture that celebrates excellence and continuous improvement.

Figure 7.11

Components of the human capital framework



COMPONENTS

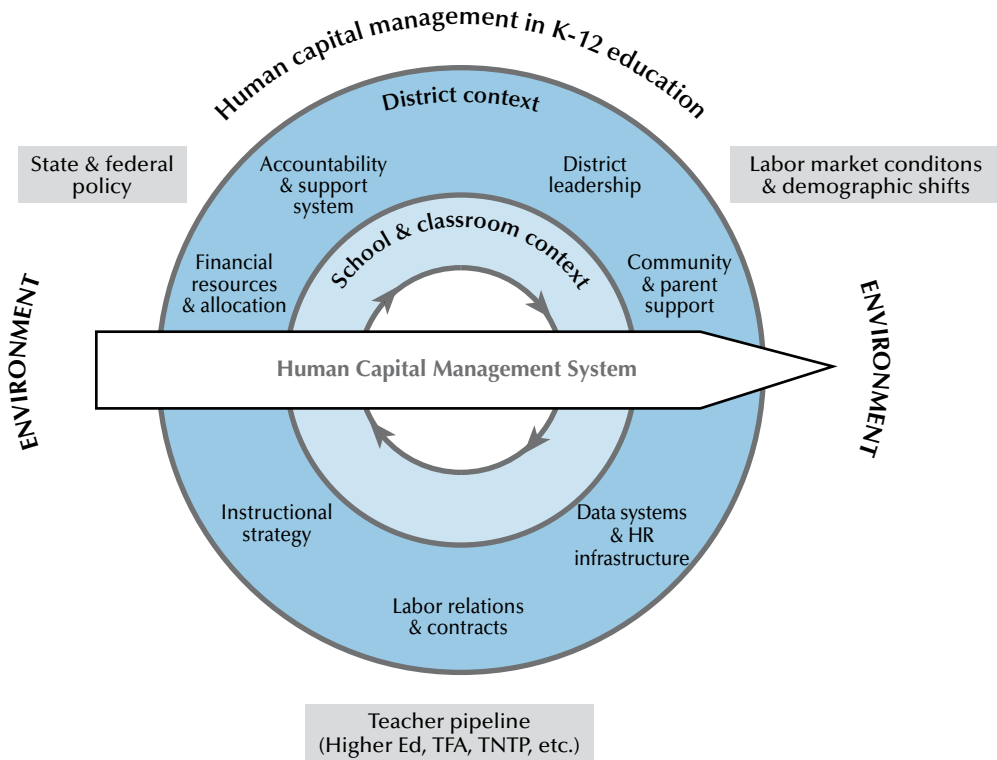
Preparation - Traditional higher education - Alternative certification - District-based preparation	Induction - Orientation - Mentoring - Reduced teaching load - Differentiated professional development	Training and development - Professional development - Identification of high-potential employees - Career management - Career pathways	Performance management - Expectation setting - Assessment - Calibration - Feedback - Outcomes	Compensation and non-monetary rewards - Annual salary and benefits - Recognition, growth opportunities, and working conditions
Source - Marketing - Recruitment - Screening - Hiring - On-boarding - Deployment	Certification Process managed by state department of education to approve teacher candidates	Tenure Point at which a district commits permanent employment to a teacher		

Source: Wurtzel and Curtis, 2008.

Thinking holistically also requires that human-capital development and management be considered in the larger contexts and forces that can either support or impede high-quality teaching and school leadership. For example, a district could address every component of the framework without realising significant improvements if it did not address student learning needs, school working conditions, allocation of resources or the development of robust data systems that track and inform improvement efforts. While districts must prioritise and should not address all of the components and contextual issues simultaneously, they must keep the entire terrain in mind as they identify their high-leverage strategies. These forces are depicted in Figure 7.12.

Figure 7.12

Contexts and forces of the human capital framework



Source: Wurtzel and Curtis, 2008.

By laying out a picture of the full terrain of human-capital development and management of teachers, the framework can promote fundamentally different dialogue about this issue. The framework is intended as a tool to help educators and vested stakeholders organise their thinking about the full scope of the issue, assess the work they are currently doing, determine where and how to first focus their efforts, and develop a long-term strategy that ultimately addresses all elements.

Wurtzel and Curtis (2008) contend that a comprehensive human-capital-development and management system will require a dramatic shift from a system that currently measures its performance by how many people apply for positions to one that values demonstrated success and is explicitly organised to recruit, develop and retain effective teachers and school leaders. Across the spectrum of OECD nations, the way this shift is managed will be informed by local and regional contexts, be oriented to different priorities, and employ culturally and situationally appropriate strategies. While the approaches may vary, the fundamental goal of an effective teacher in every classroom is universal.

Hopkins and Ahtaridou (2008) echo the importance of alignment in their report on the implications of **England's** teacher incentives for **Mexico**, concluding that it is essential that any policy regarding teacher incentives be well integrated with other policy and reform initiatives. Policies on teacher incentives need to be well coordinated with other aspects of the overall reform agenda. Policies on teacher incentives, however good, cannot by themselves have a direct impact on student achievement and social equality; they have to form part of a more comprehensive reform programme (Hopkins and Ahtaridou, 2008).

CONCLUSION

This chapter ends with a return to the start, emphasising once again the tremendous importance of teacher effectiveness – for every student and for societies. We have sought to emphasise that if a reward structure for educators is insufficiently aligned with or even antagonistic to an organisation's core purposes, the risk is high that resources will be wasted and goals unfulfilled. Financial resources are too limited, and the task of educating youth is too important for schools to be plagued by inefficiency and ineffectiveness. As such, it is imperative that educator pay and incentive programmes be aligned with holistic approaches to building human capital and elevating student performance; that they are responsive to influential contextual forces, and that compensation systems and human resource management continue to be utilised as valuable tools for enhancing teacher effectiveness and elevating student achievement.

ANNEX 7A**PRACTICAL PREPARATION FOR PERFORMANCE PAY PROGRAMME IMPLEMENTATION**

Each item in the checklist³ below may assist districts and schools with essential elements of planning, design and implementation of a performance-pay programme. This checklist identifies the major components of an alternative educator-compensation plan and the steps of planning, design, and implementation. Before launching a new compensation system at the school or district level, we recommend that nations, states or regions and districts use this checklist to discuss and decide:

- who should be involved on the planning, design and implementation teams;
- how decisions will be made;
- where necessary resources will be obtained, both immediately and for long-term sustainability;
- when major project milestones will be completed; and
- how programme effectiveness will be determined.

Once these key decisions are made, the appropriate teams should consider the following questions relating to both general and specific components of performance-pay programmes:

A. Overall plan components

- ☐ 1. Have programme developers identified a project director? Does he or she have sufficient authority to make decisions?
- ☐ 2. Are individuals at the upper levels of district or institutional leadership informed, engaged, and committed?
- ☐ 3. Have programme developers agreed upon the type and size of performance rewards?
- ☐ 4. Have programme developers identified and agreed upon the measures you will use to assess group or individual performance?
- ☐ 5. Have programme developers determined how much the new reward system will cost, overall and per year? Have you identified sources of funding and secured the necessary commitments?
- ☐ 6. Have programme developers created a representative compensation committee to ensure that you have the necessary buy-in and support from teachers and principals?
- ☐ 7. Do programme developers have a comprehensive public information and media communication plan in place?
- ☐ 8. Do programme developers have sufficient capacity in the areas of information technology, testing, and research to implement and evaluate a new compensation system?
- ☐ 9. Are other relevant organisational components aligned, such as Human Resources, Payroll, Research and Evaluation, Curriculum, and Testing departments?
- ☐ 10. Have programme developers developed an appropriate and comprehensive evaluation design?
- ☐ 11. Have project staff and leadership agreed to incorporate feedback from the evaluation and undertake mid-course corrections?
- ☐ 12. Have programme developers established benchmarks to assess project progress?

B. Stakeholder approvals and engagement

- ☐ 1. Have programme developers assembled a representative compensation committee that includes school district officials as well as the teachers and/or principals whose salaries will be affected by the new plan?

- ☐ 2. Have programme developers invited the following individuals and groups to serve on the committee so that they are active participants in discussions, planning, and decisions from the beginning?
 - ☐ a. Superintendent
 - ☐ b. Teacher union/association representatives
 - ☐ c. Additional teachers
 - ☐ d. Principal(s)
 - ☐ e. Other central office personnel
 - ☐ f. School board members
 - ☐ g. Other community representatives
 - ☐ h. State or municipal officials, if necessary
 - ☐ i. Other _____

C. Communication plan

- ☐ 1. As part of your communication plan, have programme developers developed information that clearly explains to teachers and principals the criteria you are using to determine which educators are eligible for a performance award, and what they must do to earn one?
- ☐ 2. Have programme developers developed information that clearly explains the professional development opportunities you are providing to help teachers and principals improve their performance so that they can earn a performance award?
- ☐ 3. As part of the communication plan, have programme developers developed information specifically for parents that explains how the new educator compensation works and why you are implementing it?
- ☐ 4. Does your communication plan include multiple means of distributing information to educators and the public (e.g. brochures, pamphlets, newsletters, website)?
- ☐ 5. At a minimum, does your communication plan include regular meetings with teachers and principals so that they can ask questions and raise concerns? Does it also provide other ways for educators to gather information quickly and easily (e.g. confidential telephone hotline, convenient after-school drop-in sessions, trained individuals at each school site who can answer questions)?
- ☐ 6. Have programme developers developed a strategy to explain the new compensation system to the media and to explain how awards were determined when you are ready to distribute them?
- ☐ 7. Does your communication plan include strategies to sustain the new compensation system by building support for it among policy makers, the business community, foundations, the public, and other key stakeholders?

D. Reward structure

- ☐ 1. Have programme developers decided which and how many educator positions will be included (e.g. all classroom teachers, only teachers of core academic subjects, paraprofessionals as well as teachers, assistant principals as well as principals)?
- ☐ 2. Have programme developers decided whether you will reward individuals, groups, or both? If you are planning to reward groups of teachers, have you decided which groups (e.g. all teachers in the school, all mathematics teachers in the school, all 4th year mathematics teachers in the school)?
- ☐ 3. If you intend to reward all teachers or groups of teachers, have you determined how you will evaluate the performance of those who teach non-tested subjects and year levels (e.g. preschool, art, music, physical education, 5th grade science)?
- ☐ 4. If your compensation plan will exclude any individuals or groups, have you developed a clear justification that explains why?

- ☐ 5. Have programme developers agreed upon the size of the rewards that will be offered?
- ☐ 6. Have programme developers decided to offer any non-cash awards? If so, what will they be (e.g. housing incentives, tuition assistance, tax incentives, additional credit toward retirement)?
- ☐ 7. Have programme developers decided whether the new compensation system will be voluntary or mandatory? Have you decided to phase in the new compensation system as new teachers are hired, or switch all teachers to the new system at the same time? Will current teachers be allowed to opt out if they want to remain under the present pay system?

E. Financial consequences

- ☐ 1. Have programme developers identified the sources of funding that you will use to pay for the new compensation system?
- ☐ 2. Have programme developers calculated the maximum cost of the new compensation system, year by year?
- ☐ 3. Have programme developers decided how the rewards will be paid (e.g. one-time bonus; permanent increase to base salary; premium for teachers of hard-to-fill subjects in addition to their regular salary; in-kind payment made in the form of goods and services, rather than cash)?
- ☐ 4. Have programme developers decided whether rewards will count toward individual pensions?
- ☐ 5. Have programme developers determined how close to the period of performance the rewards will be paid?
- ☐ 6. Have programme developers identified the agency that will actually pay the rewards (e.g. state or region, school district, independent or community foundation)?

F. Performance measures

- ☐ 1. Have programme developers determined whether student achievement will be measured as absolute or as relative growth (e.g. percentage of students who score at or above proficient in reading vs. percentage of students who exceed expected growth in reading)?
- ☐ 2. Have programme developers decided upon the method that you will use to measure student achievement (e.g. value-added, gain scores)?
- ☐ 3. Have programme developers identified the tests that you will use to measure student academic performance?
- ☐ 4. Have programme developers agreed upon the other sources of information that you will use to assess educator performance over time?

Examples:

- ☐ a. Supervisors' judgments (e.g. principals, mentor teachers)
- ☐ b. Peers' judgments
- ☐ c. Other _____
- ☐ 5. Have programme developers identified the teacher or principal evaluation instrument(s) that will be used?
- ☐ 6. Have programme developers developed a plan for training the individuals who will be using these instruments to evaluate teacher and principal performance?
- ☐ 7. Have programme developers determined whether other teacher and administrator actions will be rewarded, and if so, how you will weight them?

Examples:

- ☐ a. Completes specific professional development activities
- ☐ b. Assumes additional roles and responsibilities (e.g. master teacher, mentor teacher)
- ☐ c. Works in a hard-to-staff school
- ☐ d. Teaches hard-to-fill subject or specialisation (e.g. mathematics, science, special education, bilingual education)
- ☐ e. Other _____

G. Information technology considerations

1. Decision support needs

- ☐ a. Does your accountability system provide timely data for effective data-based decision-making? For example, are diagnostic test results available in time and in the right subject areas to allow school staff to identify students who need additional support to pass high-stakes tests?
- ☐ b. Are data available at the appropriate level of specificity? For example, does your data system enable you to link students to teachers to subjects taught so that you can identify which teachers to reward when student performance improves? Can you do this at both the elementary and secondary grades?
- ☐ c. Are multiple forms of evidence or data included in any evaluation of performance?

Examples:

- ☐ Value-added measures
- ☐ Observational/evaluative ratings of teachers and school leaders
- ☐ Adequacy of classroom resources
- ☐ Portfolios or other examples of student work
- ☐ School and classroom climate

2. Reporting applications

- ☐ a. Does your information-technology system allow you to go beyond basic compliance reporting and evaluate curriculum effectiveness, success of teacher induction practices, etc.?
- ☐ b. Are formal or informal processes in place to identify the data and reporting needs of individuals at varying levels of the organisation (e.g. district administrators, school principals, individual classroom teachers)?
- ☐ c. Does the data collection design process include capacity for future growth? Does your district have the ability to collect either additional data or existing data more frequently? Does the system development process allow users to define new information needs?

3. Technical considerations

- ☐ a. Have programme developers determined who needs information or data access and from where? Do you have a governance system in place to respond to changing needs?
- ☐ b. Are the systems used for creating and viewing reports available at the school and classroom levels?

4. Organisational considerations

- ☐ a. Have information-system users been an integral part of the design process?
- ☐ b. Have decisions been made regarding school versus year versus classroom value-added information?
- ☐ c. Do data-systems managers have performance and use targets that will inform them of the effectiveness of the system?
- ☐ d. Do programme developers have a data-dictionary system in place that contains information about the stored data, including details of its meaning, its relationship to other data, and its origin, usage, and format? Is a process in place to capture and incorporate change?
- ☐ e. Does the group charged with developing and deploying decision support resources have access to senior decision makers in the district?
- ☐ f. Do senior district leaders support the creation of cross-functional teams (for example, groups that include curricular specialists with information technology and accountability staff)?

NOTES

1. These and related design and implementation components are explained in greater detail in the Guidebook section of the Center for Educator Compensation Reform website. www.cecr.ed.gov.
2. The Teacher Incentive Fund is a USD 500 million allocation of funds from the U.S. Department of Education that established 34 unique 5-year performance-pay pilot programs across the U.S. The program began in 2006 and is being implemented in 20 different states in a range of districts (from large urban to small rural) and in select individual charter schools.
3. The checklist was created by a trans-institutional team of experts that comprise the Center for Educator Compensation Reform.

References

- Ballou, D.** and **M. Podgursky** (1997), "Reforming Teacher Training & Recruitment: A Critical Appraisal of the Recommendations of the National Commission on Teaching and America's Future", *Government Union Review*, Vol. 17, No. 4, pp. 1-53.
- Hanushek, E.** (2007), "The Single Salary Schedule and Other Issues of Teacher Pay", *Peabody Journal of Education*, Vol. 82, No. 4, pp. 574-586.
- Hopkins, D.** and **E. Ahtaridou** (2008), "Teacher Incentives and Stimuli: the Case of England", paper presented at the International OECD Mexico Joint Conference, The Quality of Education: Workshop on Teacher Incentives and Stimuli, Mexico City, December.
- Kingdon, J.** (1984), *Agendas, Alternatives, and Public Policies*, Little Brown, Boston.
- Koppich, J.** and **J. Guthrie** (1990), "Examining Contemporary Education Reform Efforts in the United States", in H. Beare and W. Boyd (eds.), *Restructuring Schools*, Falmer, Bristol.
- National Center for Education Statistics** (2006). "Projections of education statistics to 2015", U.S. Department of Education, Washington, DC, http://nces.ed.gov/programs/projections/projections2015/tables/table_28.asp, accessed 2009.
- National Center for Education Statistics** (2009), "The Condition of Education", U.S. Department of Education, Washington, DC, <http://nces.ed.gov/pubs2009/2009081.pdf>, accessed 2009.
- OECD** (2001), *The Well-Being of Nations: The Role of Human and Social Capital*, OECD, Paris.
- OECD** (2008), *Education at a Glance: OECD Indicators*, OECD, Paris.
- Pavez Urrutia, J.** (2008), "Incentives and Motivation for the Professors: The Chile Case", paper presented at the International OECD Mexico Joint Conference, The Quality of Education: Workshop on Teacher Incentives and Stimuli, Mexico City, December.
- Rothstein, R.** (2008), "Holding Accountability to Account". *Research Brief for the National Center on Performance Incentives*, <http://www.performanceincentives.org>, accessed May 2009.
- Santibáñez, L.** (2009), "Teacher Incentive Programs in Mexico: Evidence and Prospects for Reform", paper presented at the International OECD Mexico Joint Conference, The Quality of Education: Workshop on Teacher Incentives and Stimuli, Mexico City, December.
- Sigler, D.** and **M. Kashyap** (2008), "Human Capital Management: A New Approach for Districts", *Human Capital*, Vol. 20, No. 1, pp. 5-12.
- Vegas, E.** (2007), "Teacher Labor Markets in Developing Countries", *Excellence in the Classroom*, Vol. 17, No. 1, pp. 219-232.
- Wallis, C.** (2008), "How to Make Great Teachers" *TIME Magazine*, 13 February 2008, <http://www.time.com/time/nation/article/0,8599,1713174-4,00.html>.
- Wurtzel, J.** and **R. Curtis** (2008), "Human Capital Framework for K-12 Urban Education: Organizing for Success, A Working Draft", Aspen Institute, Washington, D.C. Retrieved May 2009 from, <http://www.aspeninstitute.org/publications/human-capital-management-k-12-education>.
- Zaharias, J.** (1999), "Project STAR: The Story of the Tennessee Class-size Study", *American Educator*.

CHAPTER 8

The Politics of Performance-Based Incentives in Education: An International Look

A sharp break: the management of teachers' work	201
Unions as political actors	202
Internal balancing act	208
Research: the devil is in the details	209
Conclusion	210
Notes	212
References	213

Countries around the world increasingly recognise the importance of education for national economic well-being, and a growing cadre of reformers are looking for ways to increase education productivity. Reformers are also searching for ways to reduce achievement gaps between disadvantaged students and their more fortunate peers for social and moral reasons, as well as economic ones. Teachers are at the centre of reform efforts. An impressive body of research indicates that teachers are the most important school-level factor affecting student learning, and that the variation in effectiveness among teachers is large. But the relationship between standard indicators of teacher quality, such as certification status, years of experience beyond the first few years, and graduate degrees, are only weakly related to teacher performance, at best. And these factors weigh heavily in current teacher pay structures. The underlying findings are strong. They have been replicated in different jurisdictions, by different researchers, with different data based on different tests.¹

To outsiders, the case for change in human resource management strategies in education must seem compelling. Significant resources are being invested in ways that lead teachers to pursue ends that do not matter much for productivity. Reformers are attempting to change the focus by structuring policy instruments that reward schools and teachers for actual performance in the classroom, especially the learning gains of their students. Traditionally, calls for performance incentives have come from the political right, but the calls today are coming from the right, left and centre. Indeed, in the **United States (U.S.)**, President Obama and Secretary of Education Duncan have publicly stated their support for the use of teacher and school performance incentives to improve student performance and teacher location incentives to ensure that the most challenging schools have the most talented teachers.

The policy shift to performance-based systems is a major departure from the policies of the past that were focused, for the most part, on putting more resources into education. In the last 40 years, the U.S., for example, quadrupled spending per student (Hanushek and Lindseth, 2009). Other countries, especially developing countries, have also dramatically expanded their education systems by increasing the number of students enrolled. “More is better” has been the education policy mantra. Expansionary policies are naturally supported by both beneficiaries (students and families) and providers (teachers) and no doubt have had benefits for society. Politicians promoting such policies have also enjoyed popular support. Presumably everyone was better off. But research is showing that more is not necessarily better; expansion alone is insufficient. At least some significant fraction of resources for education is wasted. Increasingly, available information about performance is disappointing. The U.S. is a good case in point. While education resources have quadrupled, student performance has been relatively flat. The story in many OECD countries is similar (McKinsey, 2007). Increased resources do not translate easily into higher performance. Policies designed to improve quality, however, run into a political thicket far different from that of policies that simply provide “more”.

Performance incentives face strong opposition, first and foremost, from teachers and their unions. Today’s reformers commonly benefit from a broad political base, but resistance from parties with direct interests is also strong. Some of the reasons for opposition are legitimate. For example, there is concern that, if incentives are located at the individual teacher level, teamwork in schools will be sacrificed. But as discussed in earlier chapters, if it is based at the school level, there is a risk of “free riding”, with less productive teachers simply coasting on the results of those who are more productive. Another concern is that performance incentives will distort instruction because those aspects of performance that can be most readily measured, say particular skills in tested subjects, will be emphasised at the cost of other valued outcomes that are more difficult to measure. Also, incentives pose a risk that schools will attempt to exclude low-performing students from testing in order to show better results. Then there are thorny questions about the reliability, and hence fairness, of various performance measures. Still other opponents claim that performance incentives are unlikely to work in any case, because the problem is not one of teacher effort, but rather one of teacher capacity and know-how. That is, greater investment in teacher development is needed.

Proponents argue that, apart from any intrinsic rewards teachers might receive from their work, education systems typically offer little or no incentive for them to strive for high performance. In many developing countries, incentives for teachers even to show up for work are weak. One common argument is that with performance incentives the efficiency and productivity of the system would be promoted for at least two reasons.

First, teachers would be more likely to focus greater effort on what the system identifies as most important for students to learn. In short, the goals of the system would be clarified. Second, performance-based strategies might affect the workings of teacher labour markets. The risk premium associated with the profession would shift, possibly resulting in a flow of different types of individuals into the profession, presumably individuals who consider themselves to be high performers (although it is not yet clear that these individuals would, indeed, be high performers or what level of compensation they might require for taking on a riskier compensation arrangement). Performance-based strategies might also result in different teacher-sorting mechanisms, with possibly more retention of the strongest teachers and increased exiting of the weakest, as well as more targeted assignment of the strongest teachers to the most challenging posts. To a large extent, questions regarding the consequences of performance-based systems – teacher behavioural outcomes and student learning outcomes – can be framed empirically and addressed systematically.

The fact of the matter, however, is that while a growing number of reformers are calling for performance-based strategies for human capital management in education, the state of our knowledge about *how* to best design such strategies is limited. The case for change is strong, but the lack of knowledge about the efficacy of these strategies supports the position of those opposing reform. As Goldhaber has written, “...even though the existing pay system does not appear to work well, the politics of pay reform generally make the standard of evidence for change higher than what applies to the *status quo*” (Goldhaber, 2009). For this reason the implementation of performance incentives across different countries is discussed in this chapter both in terms of the available evidence about the effects of incentives and in terms of the interests and changing political dynamics that shape, and are shaped by, these efforts. Looking across countries may be particularly useful for developing new insights since the institutional and political constraints on the structure and implementation of such efforts differ across countries, allowing us see a wide range of possibilities. This chapter, however, devotes particular attention to the U.S., where broad calls for teacher compensation reform are being met with the implementation of the newest merit-pay programmes in local jurisdictions.

This chapter is organised into four sections. First, we discuss how management strategies based on performance incentives are fundamentally different from the traditional management of education work. Teachers are understandably wary and their unions are understandably protective. The second section describes the role that unions have played in shaping policies. In addition to any effect on teachers themselves, however, performance-based compensation policies could also affect the power base of unions, which we discuss in the third section. The final section focuses on the important role that research is playing, and should continue to play, in the interchange between politics and policy surrounding the reform of teacher pay systems.

A SHARP BREAK: THE MANAGEMENT OF TEACHERS' WORK

Traditionally teachers have carried out their work with little formal external control over what they do. The reason was simple: the usual mechanisms by which organisational work is managed do not fit well with teachers' work (Hannaway and Rotherham, 2008). For example, many types of work are managed through monitoring and supervising activities directly, but this mechanism is not well suited for teachers' work, which largely takes place in relatively independent classrooms that are not easily observable, at least not without great cost. Another common way work is managed is through rules and procedures. Doctors, for example, have established procedures for diagnosing and treating different ailments, and airline pilots have established protocols that direct much of their activity. Such strategies increase performance reliability when the best ways to proceed

are reasonably well known and the timing or sequencing of their application is clear. These management tools, however, do not fit well in education, where teachers, especially the good ones, continuously adapt what they do to the specific needs and characteristics of their students and the dynamics of their classroom. To be sure, recent standards and accountability policies have shaped teachers' work in measurable ways,² but surveys of teachers, at least in the **U.S.**, show that teachers still have considerable discretion over their work. In the Schools and Staffing Survey 2003-2004 (SASS), over 90% of teachers reported they had at least moderate control over teaching techniques, evaluation of students, discipline of students and homework assignments. A majority of teachers also reported similar levels of control over selection of instructional materials and course content. Reports by novice and experienced teachers were similar, suggesting that discretion is inherent to the job, not something that results from expertise developed on the job (Hannaway and Rotherham, 2008).³

Examining output or product is a third way to manage work, such as in sales or in counts of publications by academics. Until recently, this last means was not possible in education because measures of a teacher's product, in particular student performance gains, were not available. The development of data systems, specifically longitudinal data systems that link students with their teachers, has made it possible to examine the effects of individual teachers on their students' learning.⁴ The genie is out of the bottle.

Policy contexts are changing, as a consequence of sharply increased demand for more effective education systems and sharply increased capacity of data systems that measure performance. The possibility of dramatic change is no doubt unsettling for teachers. The rules of the game change fundamentally for teachers when the system is managed by output or performance. The boundaries of discretion that teachers have exercised in their work, arguably an implicit part of the employment contract they entered into when they took their job,⁵ would be challenged.

Guarding employee-management boundaries is a key role of employees' unions. In the case of teachers' unions, collective bargaining agreements institutionally formalise many of the discretionary boundaries associated with teachers' work. For example, rules specifying the number of times and the length of time that principals observe teachers' classroom practice are common, at least in the **U.S.** An analysis of collective bargaining agreements by the National Council on Teacher Quality (NCTQ) found that only 42% of the 50 largest districts in the country required observation of tenured teachers once a year and 20% required it only every 3 years.⁶ Even observations of non-tenured teachers are low.⁷ The actual amount of required time for observation is miniscule. Hannaway and Rotherham (2008) estimate it to be, on average, 0.08% of the work time of tenured teachers and 0.23% of the time of untenured teachers, hardly enough to get a reliable picture of performance given the complex and variable nature of teaching tasks. But the arrangement has protected teachers from supervisors reaching very far into their day-to-day work in the classroom. Performance data would open a big window into the classroom for administrators and no doubt invite closer inspection and control where performance appears to be sub-par.

UNIONS AS POLITICAL ACTORS

The *raison d'être* of unions is to advance the common interests of their members, and to protect individual workers from arbitrary treatment by management. Prior to collective bargaining in the 1960s, history suggests that in the **U.S.**, teachers were not well treated. Even apart from any capriciousness on the part of management, teachers often had no breaks for lunch, had to have a doctor's note if they were out sick, and "...female teachers [had to] leave the classroom as soon as they became pregnant" (Kahlenberg, 2006). In addition, pay and benefits were low.

Wages and other benefits are central issues in both public and private sector unions. Public sector unions, however, differ in at least one basic way. While private sector unions tend to influence the wages and benefits of private sector employees almost solely through the collective bargaining process, at least in the **U.S.**, public sector unions tend to involve themselves directly in the political process. As Farber (1984) explains:

Unions in the public sector have additional incentives and functions... Allocation of funds that can be used to pay public employees is in the hands of local and state government officials. Lobbying and working for the re-election of union-friendly officials can have a direct payoff in contract terms. A strong public-sector union can increase the amount of funds available for union members. The resulting increase in membership benefits will, other things being equal, make unions relatively more attractive to workers in the public sector (p. 37).

In the private sector, labour negotiations are disciplined by the market; in the public sector, they are disciplined by the political process. The notion is that the stronger the political power of the union, the greater the resources available for members and, hence, the greater the attractiveness of public sector unions to their members relative to the private sector (Farber, 2006).

The political power of unions is enhanced to the extent that members see unions as bargaining for the collective, “the interests their members hold in common” (Casey, 2007). A focus on the collective naturally conflicts with policies that differentiate staff; performance incentives, by their very definition, lead to differentiating staff by pay. Hence, such policies are treated carefully by unions. Drawing distinctions among staff, at least theoretically, could lead to a splintering of the membership and, as a consequence, a possible dilution of union power.

The two teachers’ unions in the **U.S.** – the National Education Association (NEA) and the American Federation of Teachers (AFT) – approach the issue of performance incentives somewhat differently. The NEA, for example, strongly supports the single salary schedule that rewards teachers similarly on the basis of objective measures of education and experience, and it opposes “merit pay” in no uncertain terms. NEA 2009 Resolution (F-9) states “performance pay schedules, such as merit pay or any other system of compensation based on an evaluation of an education employee’s performance, are inappropriate.” It also opposes providing additional compensation to attract or retain education employees in hard-to-staff schools.

The AFT is more flexible. It supports pay differentials for teachers with different skills and knowledge as well as for teachers who are willing to work in hard-to-staff schools, such as high-poverty schools. Both unions agree that all pay possibilities should be “universal in scope, open in principle to all teachers who want to participate, and based on objective standards” (Casey, 2007). In short, they oppose policies that can have winners and losers, such as pay incentives that allow additional pay only for the top performers. They also oppose policies that rely on judgments or on other bases that may not be objective for making distinctions among teaching staff. Both unions also agree that all elements of salary compensation should be negotiated through the local collective bargaining process, a key difference from many other countries where bargaining is at the national level.

The NEA convened its members in July 2009 for its annual Representative Assembly (RA), and the AFT recently held its Professional Issues Conference in Washington, D.C.. The differing tones of these gatherings are telling, and useful for assessing where each union stands on performance incentives and on education reform proposals more broadly.

At the NEA Representative Assembly, outgoing NEA General Counsel Bob Chanin proclaimed, “NEA and affiliates must never lose sight of the fact that they are unions, and unions first and foremost represent their members” (Sawchuk, 2009a). Commentator Sawchuk notes that Chanin’s words were “a fitting capstone that nicely summed up the flavour of this year’s RA” (Sawchuk, 2009a). Teachers’ unions are often regarded as unique in the labour arena because they commonly espouse two causes: 1) ensuring the welfare of member teachers, and 2) promoting the achievement of the students those members teach. Chanin’s comment suggests that the NEA leans, however slightly, towards prioritising the former.

While the NEA may be more prone to act reflexively, “digging in its heels” in opposition to performance-incentive proposals that may differentiate member interests, the AFT is taking a decidedly more proactive approach. At the AFT professional issues conference, also in July 2009, AFT President Randi Weingarten vowed that her union would remain receptive to being challenged by uncomfortable ideas such as performance incentives. Her only demand was that the AFT have a seat at the table and a lead role in crafting any reforms regarding the way teachers are paid (Sawchuk, 2009b). By all indications, the pressure for districts to implement teacher performance incentives will grow over time, and Weingarten sees active collaboration with the official powers as the best way to proceed to ensure the interests of her members. As she recently wrote, “No issue should be off the table, provided it is good for students and fair to teachers” (Weingarten, 2009).

Weingarten seems to be enjoying convincing reformers that the most effective changes will come in collaboration with teachers’ unions. At the AFT’s July 2009 professional-issues conference, Weingarten unveiled her “collaboration meter,” a plug-in planned for the union’s website to provide visitors with an assessment of the degree teachers are involved with planning reforms. At the conference, Weingarten said, “When it comes to making the changes that will make our schools better, do it with us – not to us.” Some 2 500 union members in attendance were adorned with buttons that read “with us, not to us” (Quaid, 2009).

The AFT has made its position clear. The union will not just capitulate in the face of oncoming teacher compensation reforms, nor will it refuse to consider the implementation of some performance-pay incentives proven effective at improving student outcomes. Instead, the AFT will take its seat at the table and play a role in shaping future teacher-compensation policies.

Weingarten and the AFT appear to be attempting to redefine the traditional role of teachers’ unions in the debate over performance incentives and other reform proposals. The union is publicly pursuing collaboration with President Obama as he promotes his agenda to reform teacher compensation. Weingarten recently announced the launch of the AFT Innovation Fund, a USD 2.8 million fund commissioned to support state and local efforts to improve public education. The Innovation Fund is the first union-led effort supported by private foundations to provide grant monies to AFT unions for the development and implementation of “bold education innovations in public schools”, according to Weingarten. It is being supported by some prominent funders of education initiatives, including the Broad Foundation and the Bill and Melinda Gates Foundation (Mathews, 2009). The AFT’s Innovation Fund will support a variety of proposals, including “differentiated pay...where union involvement established fair measures and incentivised co-operation” (AFT, 2009). This initiative has the potential to make substantive investments in districts seeking to implement alternative compensation policies.

Unlike many countries, education in the **U.S.** is decentralised: states, rather than the federal government, have constitutional responsibility for education. Collective bargaining laws differ across states. Indeed, in some states it is prohibited. As noted earlier, however, unions are politically active and influence policies through state legislation as well as through collective bargaining. As a consequence, compensation policies in states that prohibit collective bargaining tend not to differ much from policies in states where it is not prohibited (Hess and Kelly, 2006). Where performance-incentive programmes have been instituted in the U.S., the bargaining has taken place at the district level where the local unions have played a large role in shaping the programmes. The programmes tend to be voluntary;⁸ teachers and/or schools opt in. The programmes also tend to be pilot efforts that receive support from outside sources, e.g. private foundations or special tax levies, so the basic salary structure is protected for teachers who choose not to participate. Multiple bases for inferring and rewarding performance are used, including the indicators where the link to student performance is unclear, such as participation in professional development activities.

Voluntary performance-pay programmes may be more politically viable, but they can lead to low teacher and/or school participation, undermining proponents’ claims of their cost-effectiveness and fairness. In **Florida**, the state-level Merit Awards Program (MAP) was instituted in 2007 as a voluntary programme. Only 7 of 67 county school districts participated (Stein, 2008).

Low participation in voluntary performance-incentive programmes is likely caused in part by political pressure exerted by teachers' unions on local elected officials and superintendents. Indeed, Goldhaber suggests that unions' political power is most effectively exerted at the district level. It is there that teachers' unions can tip the balance in a close school board election by endorsing a candidate, or for a district leader (in the U.S., "superintendent") considering a change in teacher compensation (Goldhaber, 2008).

Concerns about low participation and the marginal cost of instituting performance-pay plans have sometimes subjected such initiatives to budget cuts. In the 2009-10 Florida state budget, although Governor Charlie Crist used his veto power to preserve merit pay for teachers who are nationally board certified (Caputo, 2009), he did not include MAP programme funding in his budget proposal (Florida Governor's Office, 2009).

In **Texas**, Governor Rick Perry originally ordered the Texas Educator Excellence Grant (TEEG) programme as a pilot programme for the 2005-06 school year, and it was expanded state-wide by the state legislature in 2006-07 with USD 100 million in funding (Stutz, 2009). But in June 2008, it was reported that fewer than one in five districts in Texas had elected to participate in the programme. Most districts cited either teacher opposition or "financial concerns, including the requirement that participants put up some of their own money – a local match of 15 percent", as reasons they elected not to participate (Stutz, 2008). In 2009, the Texas State Legislature discontinued the TEEG merit-pay plan; part of its funding was shifted to DATE (District Awards for Teacher Excellence), another performance-pay initiative. TEEG was criticised because it was only available in low-income districts, and those districts dropped in and out of the programme. Sixty per cent of schools that had participated in TEEG in 2007-08 were dropped from the programme in 2008-09 because they failed to meet student performance targets, among other criteria (Stutz, 2009). Altogether, Governor Perry only received USD 390 million of the USD 623 million he requested for state-wide merit-pay programmes in 2009-10 (Garrett, 2009). In both the **Florida** and Texas examples, the budget cuts were at least partially attributable to a lack of broad-based support for these merit-pay programmes among the public.

Nevertheless, a promising and robust political dialogue about teacher compensation policies has emerged in many states in the last few years. According to a review of 'state of the state' addresses, 20 governors declared in 2005 that teacher compensation was one of their major education issues; while nine specifically mentioned some type of merit pay (Azordegan *et al.*, 2005).

Despite teacher unions' political influence, there is little evidence that governors are suffering politically for proposing or instituting teacher performance-pay plans. In 2005 alone, 14 state governors made proposals to reform teacher compensation in some way (Azordegan *et al.*, 2005). As of October 2007, eight states were administering a state-wide performance-pay plan (Vu, 2007). Sweetland sees the emergence of state-wide performance-pay plans as a sign that the standards-based movement is being surpassed by the teacher-quality movement. At every governance level, "a consensus is emerging that teacher-quality policies are something that we have to get right" (Sweetland, 2008).

Reforms in the **U.S.** are emerging on a piecemeal basis, given the decentralised nature of the system. In other countries, where the education system is national, reforms are generally negotiated between the national union and the national Ministry of Education. Under these circumstances unions are particularly powerful. Teachers' unions are often the largest union in a country; indeed, the teachers' union in **Mexico** is the largest union in Latin America (Grindle, 2001).

Unions in many countries have an advantage over the government in terms of stability and bargaining power. Changes in government administrations, economic circumstances that prevent the government from honouring past wage commitments, for example, and the shorter-term focus of elected officials who come and go, all strengthen the position of unions. Unions can also call a strike, an action viewed as resembling all veto power.

Under such conditions, education systems run the risk of “capture by providers” (Stein *et al.*, 2006). Basic principles defining negotiations tend to be similar across unions and across countries – a focus on the collective compensation policies that promote across-the-board pay increases for all members.

A common criticism of negotiations is that the beneficiaries of education, students and their families, are usually not included in the process, even though they have the most to lose from a poorly-run education system and the most to gain from a well-run one. Indeed, the general public is often unaware of negotiations and their consequences (Hannaway and Rotherham, 2006). It remains to be seen, then, how President Obama’s support for performance incentives will affect **U.S.** public opinion and awareness, as well as the level of political will available to jumpstart broader reforms. Several days after President Obama called for the use of performance incentives, a CNN poll revealed a divided public: 50% of respondents liked and 48% disliked the idea that teachers should get more money if their students score higher on tests (Steinhauser, 2009). This result is similar to a 2004 Gallup poll, in which 49% of those surveyed said student standardised test performance should be used as a measure of teacher quality, while 47% disagreed (Crabtree, 2004). Still, 65% of respondents in a CNN March 2009 poll said they back President Obama on education reform (CNN, 2009).

These poll numbers, however, provide little insight into the intensity of views on teacher performance incentives in the U.S., and little on reveal how much citizens are aware of the issue. In fact, there is reason to question if enough political will exists among the general public to pressure elected officials to institute performance incentives on a wide scale in the face of focused union opposition. A 2008 Gallup poll revealed a divide among parents between how they rate the quality of their children’s education versus the quality of the educational system as a whole. Seventy-seven per cent of parents surveyed were satisfied with their own children’s education, compared with only 44% of U.S. adults who were satisfied with primary and secondary education in general (Saad, 2008). With so many parents satisfied with their children’s experiences in school, one has to question how willing they may be to actively engage in a broad-based coalition pushing for reforms with few perceived direct benefits.

When the primary actors – public authorities and unions – overplay their hand, however, the public can become more aware. **Peru** provides an example. Peru came in last out of 43 countries in reading, math and science in rankings of the 2000 PISA (Programme for International Student Assessment) (OECD, 2003). The results received considerable press attention. While Peru had done well in increasing enrolments in schools at both the primary and secondary levels, not much learning was occurring, at least as indicated by PISA results (OECD, 2003; The Economist, 2007).

In an effort to raise education standards, the Peruvian government attempted to improve the teaching workforce by requiring teacher tests. Getting the right people into teaching jobs is a first step in improving performance. Such efforts were met with teacher strikes. In late 2006, the government pushed forward mandatory teacher testing, despite union opposition. Approximately 60% of the teachers took the test with dismal results; reportedly, only a quarter had adequate levels of reading comprehension (Puryear, 2007). In 2007, the Peruvian Congress passed a bill requiring teachers to take regular competency tests; three failures would result in dismissal. In response, the teachers, joined by other groups, called a strike that lasted for 15 days with several incidents of violence. Public opinion turned against the teachers (Puryear, 2007), and the government stood its ground. Teacher tests were given in March 2008, and only 150 out of over 180 000 teachers attained a minimum score on the test. The government allowed those relatively close to the threshold score to retake the test and about 50% of this group passed. The government also moved to close down universities and teacher training institutes that did not meet standards. At the same time, probably as part of an agreement with the union, the government invested in teacher training programmes, at least in the more well-regarded universities. If the public had not become aware of, and concerned about, teacher quality, and if the unions had not responded publicly in ways considered inappropriate, the government may not have been able to sustain the policy.

Chile, by contrast, has had an easier time introducing performance incentives to education. In 1996, the Chilean government established SNED – a nationwide incentive programme (*Sistema Nacional de Evaluación del Desempeño Docente*). SNED is a school-level, performance-based incentive system that uses a type of tournament model, where not everyone is rewarded.⁹ Based on tested student performance and within designated groups of similar schools, the top-performing schools are rewarded with financial bonuses that are then distributed among teachers of the rewarded schools. As a general rule, 90% of the school's award goes to teachers. In addition, a voluntary bonus programme was established in 2002 for primary school teachers who agreed to take an examination and submit classroom materials for evaluation. This programme became mandatory in 2003. The SNED programme was expanded in 2004. The changes doubled the size of the award for teachers, amounting to about 10% of a teacher's annual salary. They also increased the percentage of schools receiving the award from 20% to 35%. While seniority still accounts for most of the variation in teacher salaries over time, with the agreement of the teachers' union, a larger fraction of teachers' salaries was made more flexible and was more closely tied to performance.

The history of Chile may help explain the development of its policies. The Pinochet government had eliminated collective bargaining for teachers in the 1970s and also decentralised the education system, giving authority to municipalities (Carnoy, 1998). When a new democratic order was established in 1990, the government made an effort to re-establish trust and promote the stability of the education policy-making process by reinstating a teacher statute that allowed the union nationwide bargaining power in salary negotiations (Stein *et al.*, 2006). The teacher statute was an important concession to the union, particularly in regard to national negotiations, but the interrupted history of union power had weakened its influence and the government was able to establish performance incentives to an extent greater than countries with either a powerful union or a weak state.

Reform in **Mexico** provides a different example. As noted earlier, the teachers' union in Mexico is the largest in Latin America. Mexico's teacher incentive programme – *Carrera Magisterial* (CM) – became national policy as part of the country's education reform agenda in 1992.¹⁰ It is one of the first national efforts to tie the salary of all teachers to teacher performance. The programme is the result of bargaining between the national teachers' union, SNTE (*Sindicato Nacional de Trabajadores de la Educación*), which wanted to increase salaries across the board following a decade of declines in teacher wages, and the Mexican Ministry of Education (*Secretaría de Educación Pública*, SEP), which wanted to reform the system by introducing merit-pay to reward the best teachers.

The programme rewards teachers in a large variety of ways, including for seniority, highest earned degree, teacher test performance, ratings by the school principal and other teachers, subject matter knowledge, and student test performance.¹¹ Each of these components is given points which are combined in a complicated process.

The CM programme differs in fundamental ways from most conceptions of performance-based incentive programmes, even apart from the fact that rewards are based heavily on multiple factors beyond student achievement. Most importantly, the increased payments that teachers receive are not one-time bonus payments, but are factored into base salary. The promotions that result from the programme propel teachers into different pay lanes based on their overall performance level – from levels A to E. The levels must be pursued sequentially. Teachers are never demoted to lower levels, and the programme is voluntary. Teachers choose whether they want to enter the programme and whether to apply for promotion to the next level. The salary increases are substantial. Entry into Level A, a low hurdle, yields a 20% increase over base salary and Level E can result in a 200% increase over base salary.

The majority of teachers in Mexico are part of the CM programme.¹² Given the participation rates, the programme should have resulted, at least, in a more selective teaching force. However, about two-thirds of participating

teachers joined in the first two years of the programme, before the assessment procedures were fully in place. In effect, most teachers in the programme were never formally evaluated, and the salary “bonus” is actually an across-the-board national salary increase.

Apart from any programme design issues, efforts to evaluate the effects of the programme have been disappointing (Santibáñez *et al.*, 2007), partly because of the quality of the student test and other measurement instruments used in the programme. Considerable discussion is underway now in Mexico about the re-design of both the programme and the various ways in which teachers and students are assessed. The extent to which performance levels that differentiate among teachers can be effectively established and the extent to which they are based on student performance remain to be seen.

INTERNAL BALANCING ACT

The relationship between teachers and their unions is symbiotic. Teachers benefit off when the unions that represent their interests are strong; and unions are strong to the extent they have the support of their members. During times of change the relationship can become unsettled. If the preferences of teachers and the actions of unions are not well aligned, for example, or if teachers’ preferences are splintered, the political power of unions is weakened. Ideally, unions tend towards democratic practices. The leadership of unions cannot be seen to be “too far in front” or “too far behind” its members in terms of the stands they take. Otherwise, leadership runs the risk of being challenged and possibly voted out of office. Like in other organisations, the balance can be tenuous during times of flux.

There is indication that some splintering and misalignment of teacher preferences and union positions may be emerging. A number of studies in the **U.S.**, for example, have suggested that the preferences of younger, early career teachers may be different from the preferences of veterans, especially about issues surrounding teacher compensation and human resource management strategies more generally.

More than 15 years ago, Ballou and Podgursky (1993) examined data from the U.S. 1987-88 Schools and Staffing Survey (SASS) and found that teachers with more experience are more likely to oppose merit pay. Since then, the same pattern has turned up in a number of different surveys. A Public Agenda national survey of teachers reported that 55% of “newcomers”, compared to 33% of “veterans”, think districts should be able to use criteria other than seniority and educational qualifications to reward teachers financially. Moreover, 39% of newcomers and 23% of veterans think merit pay could be effective at attracting “the best and the brightest” into teaching (Farkas, Johnson and Duffet, 2003). More recently, an Education Sector Report, also based on a national sample of K-12 teachers in the U.S., found that 65% of teachers with fewer than five years experience support the union taking “the lead on negotiating a way to add teacher performance as a consideration when deciding an individual teacher’s salary”, compared to 45% of teachers with more than 20 years of experience (Duffett, Farkas, Rotherham and Silva, 2008). The two latter surveys, while not directly comparable, show a repeated pattern of attitude differences between new and veteran teachers. At the same time, they also suggest increased warming to the idea of performance incentives among both new and experienced teachers over the last five years.

Evidence from state-specific performance-incentive programmes in the U.S. and from controlled studies in other countries also suggests that teachers who have had experience with performance-based systems view them favourably. So the wariness of teachers overall may diminish as teachers have more experience and learn more about others’ experience with performance-based strategies. For example, a large survey of all instructional personnel who participated in the TEEG Program found that 75% agreed that performance pay for school performance is a “positive change”, and 64% also agreed that performance pay for individual performance is positive (Springer *et al.*, 2007).

Positive teacher attitudes are also evident in some of the controlled experiments that have taken place in different countries, even when the reform under study involved a tournament model where only the top reformers received rewards. For example, Lavy (2008) examined the effect of monetary incentives at the individual teacher level using data from **Israeli** upper secondary schools. Individual teachers received pay bonuses based on the performance of their students on upper-secondary-school matriculation exams after taking into account a number of factors, including student background characteristics and previous performance. Teachers were placed into four ranks within their school on the basis of both student scores and whether students passed the exam. More highly-ranked teachers received larger bonuses than teachers in the lower ranks. The bonuses were substantial, ranging from 6% to 25% of the average annual teacher salary. Teachers' responses to a survey suggest an openness to such an incentive plan as a strategy for education improvement. Two-thirds of the teachers thought the programme would lead to higher student performance – which it did.

A careful random assignment study, conducted by Muralidharan and Sundararaman (2008) in the state of Andhra Pradesh in **India**, also provides evidence of positive teacher views of performance incentives. Unlike a tournament model, where only top teachers or schools receive bonuses, the Andhra Pradesh programme rewarded all teachers and schools that improved student performance by at least 5%. The teachers were overwhelmingly (85%) in favour of the incentive programme. The better teachers (before they knew their results) were especially positive about the programme, suggesting that over time more effective teachers may be attracted to, and retained by, schools using performance bonuses.

In the case of **Chile**, Mizala and Romaguera (2005) studied the country's programme of teacher compensation and suggested that more accepting attitudes of teachers toward performance incentives may explain the union's willingness to agree to the expanded incentive terms in the collective bargaining agreement of 2004.

RESEARCH: THE DEVIL IS IN THE DETAILS

The politics of performance-pay systems will rest largely on two conditions. The first is the effectiveness of these systems for promoting student performance. The details of the system are likely to matter, such as the balance between individual and school or group incentives, the performance measures used, the size of the rewards, and whether it is a tournament or contract model, for example. As different models are implemented, researchers will be able to sort out these different factors and provide guidance on how best to design such policies. If effective models can be verified, government officials and the public will have an advantage in negotiations with teachers and their unions, should they present resistance. The second condition is the perceived fairness of the system. Teachers and their unions are more likely to support such systems, and make them work, if all teachers have a fair chance to receive awards. They will certainly strongly oppose these systems if fairness cannot be established.

A good part of the evidence we have about teachers' views of performance-based compensation systems comes from studies providing important information about the effects of performance incentives under controlled circumstances. For the most part, the studies show that performance incentives, both at the individual and the school level, lead to higher student performance. They also show that performance incentives do not lead, for the most part, to some of the undesired consequences that opponents of performance incentives often expect. It is important to stress, however, that these controlled programmes had design elements that discouraged the inclusion of some of these undesired effects, and this may be the reason why teachers' views of the programmes were favourable.

For example, Lavy (2008) reports that there were policies in place in **Israel** to assess and sanction efforts by teachers to game the system by manipulating student test scores. In Andhra Pradesh, **India**, (Muralidharan and Sundararaman, 2008), external teams conducted the student testing, and scores were calculated at a

supervised central location. Another common concern in incentive programmes is that teachers will direct attention to students likely to make the greatest gains. In the Lavy (2002) study of school-level incentives, the weakest students were found to gain the most, perhaps because the incentive was designed in a way that encouraged schools to direct more effort to those students.¹³ Similarly, performance incentives could lead teachers to encourage low-performing students to drop out or to not take the test, but pilot programmes guarded against this possibility. For example, in the programme in **Kenya** (Glewwe, Ilias and Kremer, 2003) students who dropped out or did not take the test were assigned very low scores, presumably lower than what many of them might have actually achieved. Building such safeguards into policies may foster trust in the implementation of performance-based policies and lead to better cooperation with teachers.

CONCLUSION

The returns of a well-educated population to national economic competitiveness are only likely to increase. At the same time, the capabilities of countries to measure the effectiveness of their education systems, and the contributions of teachers, are also expected to increase. In short, the likelihood of further development of performance-based policies is high. Teachers and their unions, of course, have special concerns. They want to ensure that whatever reforms are established are fair to all teachers. With any major policy shift there are likely to be disruptions for at least some of the current participants in the system. But both teachers' unions and governments have a common interest in making the education system more productive and should join together to design and test the viability of different performance-based policies. What is clear from our brief review of certain international examples is that the politics of performance-based incentives in education are changing, and so too should our conceptions of the teachers, their unions, and the elected officials involved in the debate.

An international focus on the use of performance-based incentives in education proves particularly informative; the politics of this policy debate depend greatly on whether a country's education system is centralised nationally. This chapter has attempted to describe the nuances of various merit-pay programmes recently implemented in individual states in the **U.S.** The decentralised education system in the U.S. provides a lens through which to examine the myriad details of programme design that affect the prospect of collaboration between unions and government. In contrast, nationwide experiments with performance-based systems in **Peru, Chile, and Mexico** demonstrate how these policies have been shaped by stark political realities, one case resulting in violent protests and others in compromise on the extent to which incentives are tied to student and teacher performance. Looking across countries, different political climates have produced many variations of performance-based incentives.

And yet some broad trends can be identified. There is near consensus that teacher quality has a sizeable effect on student achievement – a host of new policy interventions that aim to improve teacher performance should then be expected. A well-developed body of research that casts doubt on levers traditionally thought to raise teacher quality, including rewards for teacher certification and graduate degrees, is permeating popular dialogue. These are indications of a broader use of performance-based incentives in the near future to improve teacher output.

This emergent ethic of data collection, performance measurement, and accountability contrasts sharply with the old paradigm of teacher discretion over work and accountability based only on inputs. This is a sea change for teachers, especially veterans, and policy makers designing performance-based incentives have aware of this.

The chapter is premised on the idea that unions and elected officials play prominent roles in shaping performance-incentive policies. While teachers' unions have historically been staunch opponents of performance-based incentive plans of all varieties, some have recently displayed a tentative willingness to

collaborate with government in the design of incentives. Moreover, generational differences among union members could change the face of collective bargaining related to performance-based incentives. Numerous surveys, as described in earlier chapters, indicate that newer teachers support performance-based incentives more than veterans. Further research should examine how this phenomenon will affect collective bargaining and how teachers' views may change as they gain more experience working under incentive systems.

Elected officials from across the political spectrum have also shown willingness to support performance-based incentive plans, and have not suffered much politically when those plans prompted controversy or required revision. However, there is still reason to question how much the public supports such plans and how it would respond to any widespread reform.

Politics will continue to have positive and negative effects on the science of policy design, and evidence-based research should play a larger role in the development of incentive systems. Researchers, for their part, should also make their findings more accessible to both politicians and educators. We know that performance-based incentive policies are more than just technical components – buy-in and trust in the fairness of these systems among teachers are key to their long-term effectiveness. Clearly presented policies, practices and evaluation findings will help to alleviate anxiety among teachers faced with navigating the complexities inherent in performance-based incentive systems.

If one conclusion warrants highlighting, it is that collaboration between unions and government will be needed to produce scalable and sustainable evaluation and performance-based incentive systems in education. This point is relevant whether talking about **Peru's** centralised education system or the thousands of school districts across the **U.S.** considering incentive programmes. What remains to be seen is how this need for political compromise will affect policy design as schools enter an era of data, performance measurement and accountability.

NOTES

1. See, for example, www.caldercenter.org for multiple independent studies.
2. See, for example, Shepard and Dougherty, 1991; Romberg, Zarinia and Williams, 1989; Hamilton, Berends and Stecher, 2005; Koretz and Hamilton, 2003, and Hannaway and Hamilton, 2008.
3. Also see Kardos and Johnson, 2007.
4. See www.dataqualitycampaign.org.
5. Chester Barnard (1938) and Herbert Simon (1960) developed these ideas recognising that one of the reasons organisational change is difficult is that organisational actors defend their boundaries of discretion. Barnard used the term “zone of acceptance” to define the area in which employees agree to accept organisational authority. What is outside the zone is unacceptable and not part of the employment bargain.
6. Principals could, of course, observe more often, if teachers agreed.
7. About one-quarter of districts required observations two to three times a year while over 20% of districts had no requirement. For further discussion see Olson, 2008; Toch and Rothman, 2008; Hannaway and Rotherham, 2008.
8. The Denver *ProComp* plan is mandatory for new hires only.
9. This section relies heavily on Mizala and Romaguera (2004).
10. This section draws heavily from Santibañez *et al.*, (2007).
11. In non-tested subjects and grades a score is computed on the basis of the teacher test.
12. In 2002, 58% of teachers were members of the CM, ten years after the programme’s inception. Two-thirds of these teachers joined in the first two years of the programme’s operation. Not all teachers are eligible to join the programme. For example, teachers must have at least two years of service before joining.
13. In particular, the school mean for credit accumulation was calculated using only the student’s first 20 credits.

References

- American Federation of Teachers (AFT)** (2009), "AFT Innovation Fund", AFT, New York, http://www.aft.org/innovate/download/IF_Overview.pdf, accessed July 2009.
- Azordegan, J., P. Byrnett, K. Campbell, J. Greenman and T. Coulter** (2005), "Diversifying Teacher Compensation", *Education Commission of the States Issue Paper*, Education Commission of the States, Denver, CO.
- Ballou, D. and M.J. Podgursky** (1993), "Teachers' Attitudes toward Merit Pay: Examining Conventional Wisdom", *Industrial and Labor Relations Review*, Vol. 47, No. 1, pp. 50-61.
- Barnard, C.I.** (1938), *The Functions of the Executive*, Harvard University Press, Cambridge, MA.
- Carnoy, M.** (1998), "National Voucher Plans in Chile and Sweden: Did Privatization Reforms Make for Better Education?", *Comparative Education Review*, Vol. 42, No. 3, pp. 309-337.
- Caputo, M.** (2009), "Gov. Crist's Budget Cut Vetoes Irk Republicans", *The Miami Herald*, Miami, FL, January 28.
- Casey, L.** (2007), "The Quest for Professional Voice", *American Educator*, Summer 2007.
- CNN** (2009), "Poll: Most Americans Say Obama is Doing Too Much", CNN Report, March 18, <http://www.cnn.com/2009/POLITICS/03/18/obama.poll/>, accessed July 2009.
- Crabtree, S.** (2004), "Can Merit Plans Raise Teacher Pay Without Raising Eyebrows?", Gallup.com, October 19.
- Duffett, A., S. Farkas, A.J. Rotherham and E. Silva** (2008), "Waiting to Be Won Over: Teachers Speak on the Profession, Unions, and Reform", *Education Sector Report*, Education Sector, Washington, D.C.
- Farber, H.S.** (1984), "The Analysis of Union Behavior", *NBER Working Paper No. W1502*, National Bureau of Economic Research, Inc., Cambridge, MA.
- Farber, H.S.** (2006), "Union Membership in the United States: The Divergence between the Public and Private Sectors", in Hannaway, J. and A. J. Rotherham (eds), *Collective Bargaining in Education*, Harvard Education Press, Cambridge, MA.
- Farkas, S., J. Johnson and A. Duffett** (2003), "Stand by Me: What Teachers Really Think About Unions, Merit Pay and Other Professional Matters", *Public Agenda Report*, Public Agenda, New York.
- Florida Governor's Office** (2009), "Fiscal Year 2009-2010: Governor's Recommended Budget", February, http://blogs.tampabay.com/schools/files/govfefp_full_run_proposed_fefp_run.pdf, accessed July 2009.
- Garrett, R.T.** (2009), "Perry, Dewhurst See Pet Projects Trimmed By Republican Budget Negotiators", *Dallas Morning News*, Dallas, TX, May 20.
- Glewwe, Paul, N. Ilias and M. Kremer** (2003), "Teacher Incentives", *NBER Working Paper No. 9671*, National Bureau of Economic Research, Inc., Cambridge, MA.
- Glewwe, P., N. Ilias and M. Kremer** (2008), "Teacher Incentives in Developing Countries: Recent Experimental Evidence from Kenya", *Working Paper 2008-09*, National Center on Performance Incentives, Nashville, TN.
- Goldhaber, D.** (2008), "The Politics of Teacher Pay Reforms", *Working Paper 2008-01*, National Center on Performance Incentives, Nashville, TN.
- Goldhaber, D.** (2009), "Politics of Teacher Pay Reform", in Springer, M.G. (ed.), *Performance Incentives: Their Growing Impact on American K-12 Education*, Brookings Institution Press, Washington, D.C.
- Goldhaber, D., M. DeArmond and S. DeBurgomaster** (2007), "Teacher Attitudes About Compensation Reform: Implications for Reform Implementation", *Working Paper No. 20*, School Finance Redesign Project, Center on Reinventing Public Education, Seattle, WA.

- Grindle, Merilee S.** (2001), "Education Reform in Mexico: Where are the Parents?", *ReVista: Harvard Review of Latin America*, Vol.1, No. 1, pp. 16 – 18.
- Hamilton, L.S., M. Berends and B.M. Stecher** (2005), "Teachers' Responses to Standards-Based Accountability", *Working Paper WR-259-EDU*, RAND Education, RAND Corporation, Santa Monica, CA.
- Hannaway, J. and L.S. Hamilton** (2008), "Performance-Based Accountability Policies: Implications for School and Classroom Practices", The Urban Institute, Washington, D.C., http://www.urban.org/UploadedPDF/411779_accountability_policies.pdf, accessed July 2009.
- Hannaway, J. and A.J. Rotherham** (2006), *Collective Bargaining in Education: Negotiating Change in Today's Schools*, Harvard Education Press, Cambridge, MA.
- Hannaway, J. and A.J. Rotherham** (2008), "Collective Bargaining in Education and Pay for Performance", *Working Paper 2008-11*, National Center on Performance Incentives, Nashville, TN.
- Hanushek, E.A. and A.A. Lindseth** (2009), *Schoolhouses, Courthouses, and Statehouses: Solving the Funding-Achievement Puzzle in America's Public Schools*, Princeton University Press, Princeton, NJ.
- Hess, F.M. and A.P. Kelly** (2006), "Scapegoat, Albatross, or What? The Status Quo in Teacher Collective Bargaining", in Hannaway, J. and A. J. Rotherham (eds), *Collective Bargaining in Education*, Harvard Education Press, Cambridge, MA.
- Kahlenberg, R.D.** (2006), "The History of Collective Bargaining Among Teachers", in Hannaway, J. and A. J. Rotherham (eds.), *Collective Bargaining in Education*, Harvard Education Press, Cambridge, MA.
- Kardos, S.M. and S.M. Johnson** (2007), "On Their Own and Presumed Expert: New Teacher Experiences with Their Colleagues", *Teachers College Record*, Vol. 109, No. 9, pp. 2083-2016.
- Koretz, D.M. and L.S. Hamilton** (2003), *Teachers' Responses to High-Stakes Testing and the Validity of Gains: A Pilot Study, CSE Technical Report 610*, Center for the Study of Evaluation, University of California, Los Angeles, CA.
- Lavy, V.** (2002), "Evaluating the Effect of Teachers' Group Performance Incentives on Pupil Achievement ", *Journal of Political Economy*, Vol, 110, No. 6, pp. 1286-1317.
- Lavy, V.** (2008), "Performance Pay and Teachers' Effort, Productivity and Grading Ethics", (forthcoming, 2009), *American Economic Review*, http://economics.huji.ac.il/faculty/lavy/Teachers%20incentives_Second%20Revision_April_16_2008.pdf, accessed July 2009.
- Mathews, J.** (2009), "Rare Alliance May Signal Ebb in Union's Charter Opposition", *The Washington Post*, Washington, D.C., May 4.
- Mizala, A. and P. Romaguera** (2005), "Teachers' Salary Structure and Incentives in Chile", in E. Vegas (ed.), *Incentives to Improve Teaching: Lessons from Latin America*, The World Bank, Washington, DC, pp. 103-150.
- McKinsey & Company** (2007), "How the world's best-performing school systems come out on top", McKinsey & Company, London.
- Muralidharan, K. and V. Sundararaman** (2008), "Teacher Incentives in Developing Countries: Experimental Evidence from India", *Working Paper 2008-13*, National Center on Performance Incentives, Nashville, TN.
- National Education Association** (2009), *NEA Handbook 2008–2009*, National Education Association of the United States, Washington, D.C.
- OECD** (2003), *Literacy Skills for the World of Tomorrow - Further Results from PISA 2000* OECD, Paris, France.
- Olson, L.** (2008), "Human Resources a Weak Spot", *Quality Counts, Education Week*, Vol. 27, No.18, p. 14.
- Puryear, J.** (2007), "Peruvian Teachers' Strike Highlights Struggle for Reform", *Worldfund*, July 20, <http://www.worldfund.org/index.php?q=peruvian-teachers-strike-highlights-struggle-for-reform.html>, accessed July 2009.
- Quaid, L.** (2009), "Teachers Urge Obama to Collaborate to Fix Schools", *The Associated Press*, July 13.
- Romberg, T.A., A.E. Zarinnia and S.R. Williams** (1989), *The Influence of Mandated Testing on Mathematics Instruction: Grade 8 Teachers' Perceptions*, University of Wisconsin, Center for Educational Research, Madison, WI.

- Saad, L.** (2008), "U.S. Education System Garners Split Reviews", Gallup.com, September 2, <http://www.gallup.com/poll/109945/us-education-system-garners-split-reviews.aspx>, accessed July 2009.
- Santibáñez, L., J. Martínez, A. Datar, P. McEwan, C. Setodji, and R. Basuto-Dávila** (2007), "Breaking Ground: Analysis of the Assessment System and Impact of Mexico's Teacher Incentive Program 'Carrera Magisterial'", *RAND Technical Report*. RAND Corporation, Santa Monica, CA.
- Shepard, L.A. and K.C. Dougherty** (1991), "Effects of High-Stakes Testing on Instruction", paper presented at the Annual Meeting of the American Educational Research Association and National Council on Measurement in Education, Chicago.
- Sawchuk, S.** (2009a), "Bob Chanin Has the Last Word", *Education Week Teacher Beat Blog*, July 6, <http://blogs.edweek.org/edweek/teacherbeat/2009/07/>, accessed July 2009.
- Sawchuk, S.** (2009b), "Weingarten Wants Administration to Stick With Commitment to Collaborate", *Education Week Teacher Beat Blog*, July 13, <http://blogs.edweek.org/edweek/teacherbeat/2009/07/>, accessed July 2009.
- Simon, H.** (1960), *The New Science of Management Decision*, Harper and Row, New York.
- Springer, M., M. Podgursky, J. Lewis, M. Ehlert, C. Gardner, B. Ghosh-Dastidar, O. Lopez, C. Patterson and L. Taylor** (2007), *Governor's Educator Excellence Grant (GEEG) Program: Year One Evaluation*, National Center on Performance Incentives, Vanderbilt Peabody College, Nashville, TN.
- Springer, M., M. Podgursky, J. Lewis, M. Ehlert, B. Ghosh-Dastidar, T. Gronberg, L. Hamilton, D. Janssen, O. Lopez, B. Stecher and L. Taylor** (2008), *Texas Educator Excellence Grant (TEEG) Program: Year Two Evaluation*, National Center on Performance Incentives, Vanderbilt Peabody College, Nashville, TN.
- Stein, E., M. Tommasi, K. Echebarría, E. Lora and M. Payne** (2006), *The Politics of Policies: Economic and Social Progress in Latin-America*, 2006 Report, Inter-American Development Bank, Washington, D.C.
- Stein, L.** (2008), "Merit Pay Flows To Affluent Schools", *St. Petersburg Times*, February 24, p. 1A.
- Steinhauser, P.** (2009), "Poll: Most Americans Would Trade Some School Control For Funding", CNN Report, March 20, <http://www.cnn.com/2009/POLITICS/03/20/education.poll/index.html>, accessed July 2009.
- Stutz, T.** (2008), "Texas Schools Get Merit Pay Windfall", *Dallas Morning News*, June 21, Dallas, TX.
- Stutz, T.** (2009), "State Halts One Merit Pay Plan For Teachers, Keeps Newer Program", *Dallas Morning News*, May 18, Dallas, TX.
- Sweetland, J.** (2008), "Seeking – and Finding – Good Teaching", *Education Week*, Vol. 28, No. 9, October 22, pp. 29-30.
- The Economist** (2007), "When teacher is a dunce: Education in Peru". Retrieved from <http://www.economist.com>.
- Toch, T. and R. Rothman** (2008), "Rush to Judgment: Teacher Evaluation in Public Education", *Education Sector Report*, Education Sector, Washington, D.C.
- Vu, P.** (2007), "States Venture Into Performance Pay", Stateline.org., Washington, D.C., October 9.
- Weingarten, R.** (2009), "The time is ripe for real education reform", *American Teacher*, Vol. 93, No. 8, May/June, pp. 1.

CHAPTER 9

Putting It All Together

Key design issues	219
• Prerequisite conditions	219
• For evaluation	221
Design decisions	222
Key implementation issues	232
• Design of a pilot programme to test design	232
• Creation of an evaluation plan	233
• Alignment of compensation plan with system improvement plans	234
• Identification of “champion” for the plan	235
• Regular communication with stakeholders	235
• Development of broad political support	235
Training of school leaders and teachers	236
Revisions in plan from pilot results	237
• Full implementation	238
Conclusion	239
References	240

As the chapters in this book have shown, countries across the world are trying to establish compensation and incentive systems that will ensure the successful recruitment, evaluation, and retention of effective teachers in their education systems. An effective teacher in every classroom would dramatically increase the learning of our students and would enable economically disadvantaged students to achieve at levels comparable to their peers (Hanushek, 1992; Sanders and Rivers, 1996; Rivkin *et al.*, 2005). As McKinsey & Company (2009) has shown, this would lead to higher Gross Domestic Product (GDP) levels and a society with fewer health, welfare and crime problems (Hanushek, 1996). A recent report (Hanushek and Wößmann, 2009) suggested that the low cognitive demand of education in many Latin American countries may explain why there was less economic growth over the last 30 years in the Latin American economies than that experienced in East Asian countries which started at similar levels.

Creating a new evaluation and rewards system requires a great deal of thought and must include participation from the stakeholder groups that will be affected by the results. There is no one system that has been developed that provides an evidence-based answer to the most frequently asked question: will this system encourage teachers to improve their practice and focus their energies, both individually and as a team, on improving student achievement in ways that prepare students for successful lives in a competitive international economy and society? Empirical research is starting to provide evidence about what is needed, but one must consider whether systems created in one culture are appropriate in others. Chapter 3 notes that cultural considerations have to be taken into account in the development of new compensation structures. Chiang and Birtch (2005) found that countries with high masculinity scores and high individualism scores were more likely to value material rewards as compared to those countries with more feminine or collective scores.

Singapore (Sclafani, 2008) was cited in Chapter 2 as one of the most coherent and comprehensive systems because it uses a combination of incentives from initial recruitment into teacher preparation through to retirement, all of which are aligned to the goals of the system and enable it to select and sustain effective teachers in the system. The system has been developed over time and refinements have been added as new issues or conditions have arisen. Its first incentive is the selectivity of its admission process into teacher education programmes: teaching is a greatly honored profession, in part because the standards for selection are so high. These standards include not just the top quartile academic performance of prospective teachers, but their attitudes, dispositions and personalities, as well. So the first incentive is reputational. The second incentive is financial: prospective teachers receive full tuition as well as a salary or stipend for the entire training period. Because it is linked to the selectivity of the candidates for teacher training, it does not waste resources on people who are unlikely to become successful teachers. This funding must be paid back as liquidated damages with interest if the individual does not complete the required number of years in the classroom. This is a major incentive to work hard and succeed at one's studies, student teaching practice and initial years in the profession. It dramatically reduces early attrition in the first three years, as experienced in many other countries.

Once in the Singapore teaching corps, annual evaluations carry the possibility of performance bonuses of 10-30% of base salary based on teaching practice, student success and contributions to the school and community. The evaluations also pinpoint areas of needed growth that form the basis of the personal professional development plan for the following year. All teachers have access to 100 hours of professional development each year, at no cost to the teacher, that they can use to make progress on their personal development plans. Poor-performing teachers are provided assistance to improve and are dismissed if they do not improve their practice. In addition, teachers receive annual reimbursements for improving their knowledge and skills through professional development, subscriptions to professional journals, language learning, or technology training. Teachers move along a series of career steps that include greater compensation for greater responsibility and contributions to the profession and the school. And to keep effective teachers in the profession, there are

retirement payments through the CONNECT system that add approximately USD 84 000-112 000 (nominal 2008 values) to the national pension system over a career, and are available for partial withdrawal every four to five years.

While the **Singapore** system includes many components that could be copied, it is the coherence of the whole system that is important. To replicate it in other countries, designers would have to build a culture in which teachers are committed to developing every child in the system to his or her fullest potential and to working together to identify and implement best practices to make that goal a reality. This is certainly possible in charter schools and school systems. Because of Singapore's unique situation, however, the system does not include all aspects of compensation in public and charter schools in other countries that previous chapters have identified. Because Singapore has a single system and teachers are centrally assigned, market factors within the system are not the issue they are in other countries. For example, there are no major excesses in supply of teachers of some subjects nor in demand in other subjects because Singapore carefully projects its manpower needs in every sector and allows only the number of slots in teacher preparation programmes that will be needed in the future teaching force. In 2001 when it found a possible shortage problem as mathematics and science teachers saw more lucrative career opportunities in the private sector, Singapore added the CONNECT programme, described above, as an incentive to remain in teaching. There are no "hard-to-staff" schools because teachers are assigned where they are most needed, resulting in a mix of less and more experienced teachers in every school.

Countries need to identify the goals for their school systems that they want to reinforce and then design their compensation systems appropriately. The rationale for considering a performance-based compensation system is to improve student achievement through the increased and more effective efforts of the teachers. Chapter 3 supports the view that the use of financial incentives can motivate teachers to higher levels of performance, by encouraging more effective teachers to enter or remain in the teaching profession, and can align teacher behaviours and interests with institutional goals. Within compensation systems, there needs to be some flexibility for local educational systems and individual schools to select incentive options that best fit the culture and needs of their schools. Allowing schools to develop their own plans for distributing incentives within the school develops commitment to the plan, as seen in the Vaughn Next Century Learning Center and the **Florida** and **Texas** incentive programmes described in Chapter 2. If plans are to improve over time, they must include an evaluation component that periodically describes the impact of the plan and suggests improvements.

Neal (2008) suggested that systems might wish to establish a market system in which private providers run schools and compete for teachers who can help them reach their performance targets. In this way, the providers would have an incentive to develop efficient compensation policies and practices that made their schools more competitive. They would seek to attract and retain teachers who make the greatest contribution to the quality of the school and pay them according to their contribution. Systems that encourage states or regions, districts, or individual schools to develop their own personnel practices and improve them over time will be able to do the same thing and help identify which practices are most effective and efficient. This should be the goal of every school system, but it is difficult for some to leave behind past practices that have been negotiated with unions and employee organisations.

KEY DESIGN ISSUES

Prerequisite conditions

From past experiences in school systems across the world we can identify key design issues for the development and maintenance of an effective compensation system. The first set of issues regards the prerequisite systems and decisions (Heneman, Milanowski and Kimball, 2007) that should be in place prior to designing a new compensation system.

Funding. The first issue is funding. The basic teacher compensation system must be competitive with neighboring school systems before a system decides to institute a performance incentive system. The organisation will have to decide if it wishes to pursue financial, non-financial or a combination of the two types of incentives. Non-financial incentives include: public recognition at school/state or regional/national levels; gratification of intrinsic motivation; additional autonomy for teacher/school; positive working conditions, including supportive leadership and curricular flexibility; and job stability. While these may have some costs attached to them, they are not as costly as adding financial incentives.

If an organisation decides to use financial incentives in addition to non-financial ones, these typically add from 2-15% of base salary, but the additions will not attract or retain teachers to compensation systems that are not already competitive. In an effective incentive system, teachers would receive a competitive salary but would understand that future salary increases would be the result of their effort and success as teachers. The second part of the funding issue is the cost of initiating and sustaining an incentive system. If incentives are to be included in the salary, rather than considered as bonuses, the organisation will have to take into account the tax and retirement funds required for implementation. If the incentives do become part of salary, the system must decide if they will become permanent parts of the salary or will have to be earned annually but counted as salary for that year. This is a major question as continuing to receive an incentive, regardless of one's future performance, means that it is a raise, not an incentive to future productivity.

If the incentives are based on evaluations of teacher performance, the cost of developing an evaluation system based on rubrics, a rating system that assigns points to narrative descriptions of the presence or absence of desired behaviours; of hiring and training evaluators; and collecting, analysing and reporting data may be new expenses for the district that need to be considered. If incentives are based on student performance on standardised assessments, additional costs may result from the selection of assessments in more content areas or at more grades or the need for pre-tests as well as annual assessments. All of these costs would have to become part of the budget considerations for the plan.

Guthrie and Prince (2009) noted that it is especially important to consider the sustained costs of implementing the programme at the district, state/region or nation levels when designing the incentive programme and certainly before embarking on a pilot. Guthrie and Prince suggested that systems could reprioritise current funds, seek additional funding from taxpayers, or solicit additional funds from foundations. A part of reprioritising current funds could be to base all future salary increases on the criteria of the incentive programme and not on experience. While this may not be popular with teachers, it does make clear that productivity is valued, not just years on the job. The designers of the incentive programme would have to make their design decisions with the amount of available funds in mind, and each year the system must include a careful analysis of the maximum rewards possible and where the funds to pay those will be found in the unlikely case that they exceed the budget.

The identification of a stable source of funding of sufficient size to fund incentives on top of ongoing compensation for the system is essential if stakeholders are to have confidence that a new plan will last beyond the current year. The history of merit-pay plans in the **United States (U.S.)** in the 1970s-1990s, when continued funding for incentive programmes was often vulnerable, has made educators cynical about discussions of new systems. If there is not enough money to pay the incentives teachers have worked toward, there are only three options (Odden and Wallace, 2007). The system can add more money to the incentive budget, reduce the amount of payments to teachers, or raise the performance standards for receiving an award. While the first is the best option for the long-term teacher acceptance of the programme, it may be difficult to impossible to find sufficient funds. The other two options reduce the motivation of teachers to work toward the incentive goals because they are not convinced their efforts will be rewarded. The **Denver** ProComp system described in

Chapter 2 developed the trust of teachers and union leaders only after voters agreed to an annual tax increase of USD 25 million that would be placed in a trust administered by the district and the union. Of course, stability is a challenge that cannot be fully overcome, as it is difficult to maintain commitment to past decisions in new administrations. For example, in a recent legislative discussion in **Texas**, the Teacher Educator Excellence Grant was proposed for elimination, and the programme was eliminated, with its funding moving to the district incentive programme DATE discussed in Chapter 1.

For evaluation

Robust data management system. A second prerequisite is that a country, state/region or district puts in place a robust data-management system that is able to connect student and teacher data. The Data Quality Campaign in the U.S. has identified ten features that must be present to engage in effective academic research evaluation and to design effective compensation systems:

- A unique statewide student identifier that connects student data across key databases across years.
- Student-level enrollment, demographic and program participation information.
- The ability to match individual students' test records from year to year to measure academic growth.
- Information on untested students and the reasons they were not tested.
- A teacher identifier system with the ability to match teachers to students.
- Student-level transcript information, including information on courses completed and grades earned.
- Student-level college-readiness test scores.
- Student-level graduation and dropout data.
- The ability to match student records between the Pre-K-12 and higher education systems.
- A state data audit system assessing data quality, validity and reliability.

(Data Quality Campaign, 2009, <http://www.dataqualitycampaign.org/survey/elements>)

A database that includes all of these elements creates a foundation for data-based decision-making that can be used in school improvement efforts as well as in a teacher-incentive programme. It enables a system to connect teacher and student performance over time, which is critical to the use of Value Added Model (VAM) analysis for school-based or individual teacher-based incentives. (See the discussion of VAM in Chapter 5.)

Valid performance measures. Along with the data system, a school system must have valid and reliable measures of teacher performance that teachers agree are fair and accurate. This may include assessments of teacher performance that are based on multiple observations by trained evaluators using a standards-based rubric that teachers believe reflects good instructional practices. Standards-based rubrics have been developed in **Chile** (Manzi, 2008) in The Framework for Good Teaching (*El Marco para la Buena Enseñanza*) that is used for evaluations that can lead to additional salary bonuses. In the **U.S.**, the Danielson Framework for Teaching (Danielson, 2008) and Wallace's Teacher Excellence Through Compensation Teacher Evaluation System (Odden and Wallace, 2008) have been used both for evaluation and professional development. External examinations such as Praxis III (Educational Testing Service, 2009) or the certification programme of the National Board of Professional Teaching Standards can also be used for this purpose. Other measures of teacher performance may include contributions to the school-improvement efforts or performance in specific areas based on external certifications.

Performance measures may also include student performance and/or student progress on high-quality assessments that are aligned to learning standards teachers are required to use. If the tests are to be used to determine value-added progress that students have made (*i.e.* using statistical methods to analyze the student's current scores in light of past performance to get an accurate reading of the effect of the school/teacher on the

student's performance), the test should be designed to enable analyses of year-to-year gains in performance at the individual student level. Such assessments should allow for a range of performance that differentiates between students at all levels of knowledge and proficiency, from lowest to highest performers. This is not the case with many state/regional or national tests used across the world that are competency and content-based, but are not designed to measure student performance at defined levels of proficiency. Most have low ceilings for student performance or represent a small band of competence around the grade level expectations for student performance (Koedel and Betts, 2009). High-performing students will not be able to show growth if they, based on past knowledge and skills, start the year achieving at the highest level required on the test. In the same way, tests that include too few questions at the low range of competence will not allow students who are working below the expected standards to demonstrate their growth from one year to the next.

Other objective measures of student performance include enrollment in advanced courses, student attendance, student graduation rates, and student dropout rates. Analysis of student work can provide another measure of student performance, but it requires time and funding for a group of trained assessors to evaluate portfolios of student work and determine evidence of growth. Student and parent responses to well-constructed surveys are another way to measure the performance of the school.

Communications system. To build teacher acceptance of the plan, designers must communicate their interest in establishing a differentiated compensation plan that might include different salary levels based on performance, knowledge and skills, disadvantaged students, shortage subject areas or additional responsibilities undertaken by the teacher. As the attitude surveys in Chapter 1 indicated, some teachers are opposed to all forms of differentiation, while many are interested in only some types of incentives. However one designs the plan, it is important to communicate early and often about the design and get input from stakeholders, including teacher unions. Odden and Wallace (2008) suggested that communication must be the rule and referred to a case where school leaders thought everyone had been included in the process, but many teachers saw the plan as a surprise when it was formally placed before the faculty. It may be wise to enlist the services of professional media professionals to serve on the committee and to assist in the design of appropriate language to describe the plan from the first open communication about selecting stakeholders through the launch of the programme. This will ensure that the internal and external stakeholders have accurate information that avoids the use of terminology that incites distrust, as discussed in Chapter 6.

Access to feedback and improvement. The last prerequisite would be to ensure that teachers have access to feedback, support and professional development to improve their knowledge and skills. Incentive programmes are designed to motivate greater efforts in areas that improve student achievement. As part of "expectancy theory" (Heneman, 1992) (discussed in Chapter 1), teachers must believe they are capable of the performance required to earn the incentive. If teachers are doing the best they know how to do, the system must provide feedback on areas of strength and weakness as well as support to overcome weaknesses. This includes access to professional development on content knowledge and effective content strategies shown to improve student performance; analyses of student work; and student performance on formative, interim assessments done in the classroom or schoolwide during the school term, and summative assessments done at the end of the year to evaluate progress on the year's curriculum; as well as mentors and coaches who can assist the teacher in improving his or her performance through modeling and observations of master teachers. It must also be clear that teachers who are unwilling or unable to improve their performance should not remain in the system.

DESIGN DECISIONS

Once these prerequisite systems are in place, it is time to start designing a performance-based compensation system. The design can be undertaken at any level: local, state/regional or national. In this section, "system" will be used to describe the organization that has undertaken the design of a compensation system. This is

not a quick process, as the decisions must be made in ways that develop support among stakeholders for the programme. The creation of a representative compensation committee will be key, and the system must gather information and identify experts who can work with the committee to help members understand the nuances and interrelationships of the components they decide to include in the incentive programme. The committee members have to stand behind the programme they design, and they can only do that if they are confident they were aware of the possible consequences, both negative and positive, of the components they include. The system will need to budget money and staff time for the programme development process. Successfully done, the work of the compensation committee will result not only in an incentive programme plan, but also have forged a strong advocacy team that will work together to sell the programme to their constituency groups.

Compensation committee. As suggested in Chapter 6, the creation of a compensation committee representing stakeholders is the first step in design. Generally a committee should be limited to about 15 members, but inclusiveness of internal and external stakeholders is critical to success. The educational leader of the system should be involved in determining the membership of the committee and should be present at key meetings. If attendance at all meetings is not possible, a high level designee will be critical to signal the commitment of the chief education officer, whether the Minister of Education or school district superintendent. Teachers and teacher unions should also be represented on the committee; the perspectives of the most effective teachers and the union responsible for considerations of salaries and working conditions are needed. Heneman, Milanowski and Kimball (2007) pointed out that principals and district administrators who will play a major role in implementation should be included on the committee. Without their support and deep understanding of the plan, implementation is likely to be more difficult. In addition, their participation will facilitate alignment of the compensation plan to the other human resources and instructional systems. It will be critical to include two departments that will be heavily involved in implementation of the plan: the research and evaluation department that will be responsible for collecting and analysing the data to determine awards, and the finance department that will plan and budget for the incentives and pay the compensation to appropriate recipients. Representatives of these departments can provide the compensation committee with important information and data. For example, members of the committee need to know what data are currently collected and what the costs of collecting new data would be. The research and evaluation staff should also be analysing components of the evaluation plan that will be needed to ensure long-term success of the incentive programme. The finance representative will be able to assist in determining implications of specific committee recommendations for current and future budgets so that the plan does not include components that would be too costly for the organisation to afford.

The membership of the compensation committee should include external stakeholders, as well. Incentive programmes are highly political in nature, and systems need to develop and maintain support of a wide range of elected officials and business leaders. The plan's success may depend on the future incumbents in these positions, not just those currently holding leadership positions, so wide support is essential to continued implementation. The committee should include representatives from the external governing board of the organisation, whether that is a legislator leading the education committee at the national or state or regional level, a member of the state or regional board of education, or at the district level, a member of the local board of education. Business leaders should be represented as well, and may provide good contrasts and lessons from the business sector that can assist educators on the committee in making their recommendations. All members of the committee will be assigned responsibility for specific roles in communication and building support for the plan in addition to their design responsibilities. In fact, the committee should develop a communications strategy as one of its first activities.

The work of the committee should start with representatives of the system discussing the goals of the plan and the role of the committee. It is critical to establish at the outset whether the committee is there to provide input

to the system's design of a plan or to design a plan to recommend to the system for approval by the governance board. The latter will develop greater buy-in, but if the former is the role the system wants the committee to play, it should be sure there is no confusion about that. The system should also clarify if it wants the committee members to maintain open communication with colleagues about the work during the design process. This is often a good way to start developing support for the programme.

To ensure that all members of the committee are using a common set of references, the system should provide background information on what an incentive programme is and what components of an incentive plan have been implemented in similar entities in other geographic areas. Evidence should be shared on the success or challenges these plans have encountered and the modifications that have been made to them over time. Experts in the field are perhaps best able to establish an objective grounding for these discussions. From this point, the committee will focus on the specific objectives of the plan they are designing and decisions they will need to make among the options for inclusion into the plan. It is to be expected that the committee will not agree on everything, and it is important for the leader of the system to clarify the basis for the work. Will the committee include only those items on which they can reach consensus or will a majority or supermajority vote be required to recommend a component? Max and Koppich (2007) suggested that the committee may need a neutral facilitator who can help the committee address and overcome any impasses. This might be the expert who provided the overview of options at the beginning of the committee's work.

Development of a communications strategy. As soon as a compensation committee is appointed, people within the system will begin to offer opinions and concerns about an incentive programme. To provide valid and up-to-date information on the purpose of the committee and its progress, the committee must establish a communications strategy for keeping stakeholders informed. Chapter 7 includes a checklist for communications that would be useful to review. Odden and Wallace (2008, p131) suggested a series of questions the committee should explore:

- What do we need to communicate? What do people need to know about the compensation programme at this point in our progress?
- Who is the audience? Who needs to know?
- What is the best timing for the communication?
- What are the key points to communicate?
- What is the best medium for communicating the message?
- How can formal communications (e.g. reports, minutes, formal meetings) be used?
- Possible uses of informal media (e.g. one-on-one communication, informal networks)?
- Possibilities for Web-based or Intranet media?
- Can parallel media be combined?

It may be wise to bring in professionals to assist in this work. The choice of words to describe the options may be critical to avoiding unnecessary dissent. In fact, the definitions of the terms to be used should be carefully thought through and reviewed for sensitivity. A union official on the design team in Denver noted that the term "pay for performance" would not be well received by teachers, and so the district used different words to convey its goal of rewarding teachers for growth in student achievement (Max and Koppich, 2007). Koppich and Prince (2007) suggested that certain terms such as "merit pay" that recall past failed attempts to link compensation to performance not be used to describe the new plan.

The written materials should be clear and transparent so that stakeholders understand what the committee is doing. It may be useful to customise materials to address the specific interests of different stakeholder groups. Parents will have different interests than teachers, who will want to understand the details of the plan.

Dissemination of the materials should take advantage of existing newsletters to parents, administrators or teachers; information pages on the system's website; or a section on its Intranet, reserved to system employees. Denver created a ProComp handbook and made it available on its website, and sent newsletters to all district employees (Koppich and Prince, 2007).

The committee may choose to schedule meetings at convenient locations for stakeholder groups to explain the goals of the plan and the progress they are making. This makes it clear that the work is an open process that seeks input and feedback from constituent groups. Regular updates at governance board meetings add to the transparency of the process and provide all board members an opportunity to hear and make comments. Odden and Wallace (2008) suggested that each member of the compensation committee establish a personal network of 10-15 contacts whom they individually keep informed, and from whom they solicit feedback. Since the committee members represent different constituencies, the committee can be assured of reaching a wide breadth of constituents in this manner.

External media should also be part of the information-sharing process. The media will be alerted to the development of the programme, and it is better for the system to contact them at the beginning of the development process so that they can become partners in sharing correct information. Koppich and Prince (2007) noted that early engagement of the media may prevent negative headlines based on misunderstandings of the incentive programme process when the plan is implemented. While it may not be possible to avoid negative headlines, they are less likely if the media has been kept apprised of the programme and its components from its early development.

What to reward? Hassel (2002) provided a list of questions that must be answered by the designers. The first decisions include the “what, who, and how” of incentives. Compensation should be based on more than credentials and years of experience since these are weak indicators of teacher effectiveness. If the goal of changing the compensation system is to improve student achievement, then the system rewards must be connected to practices that promote student learning. Schacter and Thum (2004) developed and tested a synthesis of standards of effective teaching and found that teachers who scored high on evaluations against those standards also had students who scored higher on standardised assessments. That may encourage systems to include both evaluations of teacher performance as well as evaluations of student performance in their compensation system. Manzi (2008) suggested that **Chile's** identification of a framework for good instruction, The Framework for Good Teaching (*El Marco para la Buena Enseñanza*), ensured acceptance and effectiveness of its incentive system.

As with merit pay, teachers will want to know the basis on which incentive pay awards are being granted, and what they can do to earn the awards. Unless the expectations for student performance and practices being rewarded are clearly articulated, these questions will not be answered to the satisfaction of teachers in the system. These practices should also be used throughout the human resources system, from selection of new teachers, to evaluation, professional development, and coaching. It may be difficult for systems to develop support, however, if they completely neglect credit for credentials and experience. The Cincinnati, **Ohio**, plan discussed in Chapter 2 lost support of the teachers because teachers could receive less compensation in the new plan than they had when experience and credentials were the determinants of salary (Delisio, 2003). Holding past teacher salaries unaffected as teachers move into the new system may require more money, but it will add to the likelihood of teacher support for the programme.

Who should receive rewards? The next set of questions concerns which staff to reward. As many of the examples in previous chapters demonstrated, eligibility for rewards should be open to all teachers if teachers are to view them as incentives to work harder, although actual rewards for performance should probably be limited to the top quarter or third of teachers. There has not been sufficient research in education settings on the percentage

of teachers who should receive incentive payments to make definitive statements in this regard. Educators are left to trial and error and additional experimentation to determine what is enough to motivate most teachers to work more productively or too small to motivate anyone.

A second issue is whether to reward others at the school. Principals and assistant principals may not serve in the classroom, but their work is critical to establishing an environment that is conducive to improving student achievement. School administrators are accountable for the performance of the school and can be considered as members of the team receiving school-wide incentives if the school meets or exceeds its targets. The importance of the principal was described in Chapter 1, where teachers identified the quality of the principal's leadership and the school's culture and working conditions as critical components to their decisions to select and elect to remain in a school (Milanowski, 2006). If the school is recognised for the quality of its performance then the administrative team should be included in the rewards. Other professional staff at the schools should be considered as well. In **Denver's** ProComp plan, "teachers" include counselors, social workers, psychologists and nurses. Each professional has to establish, and receive approval for, measurable goals for student learning and the kind of evidence they will use to evaluate progress.

In the same way, non-professional school staff makes contributions to the school environment that enable teaching and learning to take place. Without them, the school would not succeed. If school-wide incentive programmes recognise that success is a team effort, then all staff should be considered to share in the rewards. In many programmes, such as the Governor's Educator Excellence Grant in **Texas** (Springer *et al.*, 2007 and 2008), the amounts for teacher incentives make up 75% of the money granted to the school, but the other 25% is available for incentive payments to principals and other staff at the school. In **Alaska** (Max and Koppich, 2007), schools receive awards for improving student performance on state tests. The principal and teachers can receive bonuses of up to USD 5 500, while non-certified support staff can receive up to USD 2 500.

Market-based incentives. Supply and demand are issues in education as well as in other fields. If the system is having difficulty in attracting a sufficient number of teachers to specific fields or schools, it should consider incentives for those fields. Making distinctions among teachers by teaching field is not a popular basis for incentives among teachers who feel teachers in all fields make important contributions to the education of children. The committee's decision is whether to include incentives for shortage-area teachers to address the wage differentials available in the market place for them. If staffing certain schools is difficult, either because of location or student demographics and achievement levels, the committee may wish to add incentives to teach in such schools. The **Denver** ProComp programme (Koppich, 2008) described in Chapter 2, resulted in applications to hard-to-staff schools rising eightfold after the introduction of a 3% of base salary incentive for teaching there.

In both shortage areas and hard-to-staff schools, one must consider making the incentive available to new teachers going to the school and to teachers who already teach that subject or teach in that school. Since the incentive will most likely be needed to both attract and retain, it will be important to establish criteria in advance to avoid rewarding current shortage teachers who do not meet the same criteria as incoming teachers or teachers who demonstrate through evaluations or measures of student achievement that they are not effective teachers of those fields or in those schools. Not all teachers are effective with all types of students. This points to the interconnectedness of all of the incentives and the importance of continuous attention to their alignment to the goal of improved student learning.

Unit of reward. A major issue is whether the awards are targeted to individual teachers, groups of teachers or the whole school. As discussed in Chapter 3, there are advantages and disadvantages to each approach. Lazear (1996) concluded that individual incentives serve a dual purpose, both motivation and selection. They reward individuals who are most effective at meeting the specific goals measured, and those individuals are motivated

by the additional compensation they receive. As a result, the highly-rewarded individuals are more likely to stay in the field, while those who are not successful are more likely to leave, providing room for the hiring of more effective employees. Chapter 3 also notes that individual rewards have two advantages: they most directly link teacher performance to the performance of students in their classes, and the individual teachers know that they have direct control over their chances of receiving a reward. Only a third of the teachers, however, in most systems are teaching in grades or subjects that have external, objective measures of student performance. If a system wishes to include teacher results, it will have to determine whether to add external objective measures of performance or use more subjective measures such as portfolios of student work that require trained assessors and funding to provide them time in which to conduct the assessments.

If the system decides to use standardised tests, it may be hard to differentiate the impact of the current teacher from that of past teachers, both effective and ineffective. On the one hand, current teachers may benefit from the learning provided by past teachers who went beyond their grade level's requirements. On the other hand, current teachers may have to work much harder to get students up to grade level when they enter far below because of ineffective teachers students may have had in the past. As mentioned earlier in the prerequisites section of this chapter, the nature of assessments may contribute to the difficulty in identifying the impact of individual teachers. It is also difficult to isolate teacher effects on student learning from all other effects in a student's environment. The most often-mentioned concern about individual teacher incentives is that they will destroy teacher cooperation and collegiality by establishing a competition for achieving limited awards (Murnane and Cohen, 1985). As described in chapter 3, this can be avoided by carefully designing the plan to include both individual and group components to the incentive programme. Lavy (2002) found that individual teacher incentives resulted in greater student achievement than group incentives in his study of a tournament incentive programme experiment in **Israel**.

Incentive programmes may consider a group as a unit of accountability. For example, the programme may evaluate the performance of grade-level teams, disciplinary departments, or other groupings that fit a school's structure and mission. Group incentives are found to promote social cohesion, feelings of fairness, and productivity norms (Lazear, 1998; Pfeffer, 1995), and they may foster the transfer of knowledge and mutual learning among teachers that can lead to improved results (Che and Yoo, 2001). They may also foster competition between groups, however, that could reduce collaboration among teachers. School-wide incentives encourage collaboration among teachers to ensure the school meets the criteria for rewards, but they may have disadvantages as well. Previous chapters have noted the problem of "free riders" in group programmes, including school-wide programmes, but Chapter 3 notes that these may be overcome by keeping groups small or by establishing programmes to exert social pressure as well as to monitor peer contributions to the group's performance. In one of the few experimental studies, in **Andhra Pradesh, India**, Muralidharan and Sundaraman (2008) evaluated the impacts of two output-based incentive systems (an individual teacher incentive programme and a school-level teacher incentive programme) and two input-based resource interventions (one provided an extra paraprofessional teacher and another provided block grants to purchase desired resources). They found that individual incentives produced greater student learning than either the school-wide programme or the two inputs-added programmes.

In considering group or school-wide incentives, the committee has to consider the impact on individual teachers who are highly effective and know that if they were rewarded for their individual impact on the students they teach, they would receive a reward. If working in a group or a school that has significant numbers of teachers who either are not working hard to improve student performance or are ineffective despite how hard they work, the effective teacher may decide to transfer to another school that has more productive or more effective teachers. The **Texas** Educator Excellence Programme described in Chapter 2 found a sharp increase in teacher turnover among those who did not receive awards or received small awards and a marked increase

in teacher retention among those teachers that received high bonus amounts. A hybrid system that combines elements of individual and group incentives is recommended by Hanushek and Lindseth (2009) as a way to ensure that high-performing teachers are rewarded for their efforts even if the whole school's performance has not improved. Low-performing schools cannot improve without having high-performing teachers on the staff. Chapter 3 warns against a lack of goal clarity, however, that can arise when an incentive system has multiple components.

Payout structure. The next question is whether to structure the payout of the incentive programme as a rank-order tournament or a fixed performance system as discussed in Chapter 3. The first has the advantage of establishing at the outset the maximum amount of money that a district or country will be obligated to spend. If the decision is made to reward the top quarter of all schools or teachers, the required funding will be known in advance and can be allocated. If schools or teachers know that only the top quarter will be rewarded, however, they may only be motivated to work harder or more productively if they believe that they have a chance of reaching that level. Grouping schools or teachers by the demographics of the students they teach, and then rewarding the top quarter, will help to ensure the comparisons across school are fair and the incentive more productive. Going a step further and using a value-added method of analysis will provide schools and teachers assurances that they can join the rewarded group by taking their students to *higher* levels of performance, not just to the *highest*.

A fixed performance system establishes a goal for student performance or for teacher evaluation and rewards all schools or teachers that meet that performance level. The system must very clearly describe what the performance is that teachers are expected to reach so that teachers understand what they must do to pursue this opportunity. It truly opens the possibility of earning a reward to more teachers and encourages teachers to develop their skills and work more effectively, but it raises the amount of money that must be set aside to fund the incentives. While a system might expect that a certain percentage would reach the goal, it must have funds set aside to ensure that the incentive fund is sufficiently large to pay all teachers who meet the performance goal. To renege on the payment of incentive rewards to teachers will doom an incentive programme, as teachers question the commitment to improvement that it represents. The decision between rank-order tournaments and fixed performance systems may depend on the resources available, although a system may choose to use a combination of the two when it wishes to focus on teachers or students mastering a specific skill or knowledge area.

Input versus output measures. The next major issue is what performance measures will be used for the incentive system. There are two types of measures to choose from: input measures and output measures. Previous chapters provide examples of the types of activities rewarded as input measures, including completing an advanced degree, serving as a mentor or master teacher, taking on a leadership role or other special duties, participating in professional development activities, and teacher attendance. Odden and Kelley (2002) suggested that opportunities to learn new skills or mentor younger teachers provided both interest and motivation for teachers to improve their practice.

Knowledge-based and skill-based pay programmes are the most prominent type of incentive-pay policies that focus on rewarding educational inputs. Odden and Wallace (2008) described five types of knowledge-based measures that are used in the **U.S.**: earning a professional teaching licence, earning National Board Certification, receiving a high score on a performance evaluation, participation in professional development activities specified by the district or an individual development plan, or completing an advanced degree in a subject area or credits for an endorsement in a specific area, such as educational technology or English as a Second Language. The professional licence is a step beyond initial certification that requires demonstration of knowledge and skills on an observation system, such as Praxis III (Education Testing Service, 2009), or performance evaluations. such as the **Connecticut** Beginning Educator Support and Training (BEST) system, that lead to advanced levels

of professional licensure (Odden and Wallace 2007). As discussed in Chapter 2, both **Chile** and **Mexico** provide incentives for teacher performance evaluations, and Mexico includes professional development as an additional criterion for its incentive system. Santibáñez *et al.* (2007) noted that the evaluations done by peers, the principal and a union representative in the Mexican *Carrera Magisterial* programme were uniformly highly rated. That does not help to distinguish high performers from others. Chapter 4 noted that the evidence available to the National Union of Teachers points to teachers' enthusiasm for high quality professional development. He suggests that high quality, interactive and sustained professional development encourages teachers' learning and also encourages more experienced teachers to remain in the profession.

Input-based incentive systems are well connected to teachers' beliefs that improving their knowledge and skills will improve their practice. In addition, Chapter 3 makes the point that rewarding inputs reduces the focus on "teaching to the test," or focusing a disproportionate amount of effort toward those students near some performance standard in hopes of maximising chances of earning a performance award. The Chapter also notes, however, that since teacher inputs show very low correlations in empirical studies with teacher effectiveness, there is a greater interest in using output measures, especially increases in student achievement.

Output measures carry their own advantages and challenges. Systems have used a variety of measures including teacher evaluations, portfolios of student work, student and parent surveys, and student-achievement measures such as standardised state/regional or national examinations. The first three examples require subjective evaluations and are susceptible to the challenges described in Chapter 1 as the rationale for adopting the single salary schedule: gender and racial biases and rewarding friends and allies for personal reasons rather than teaching expertise. While teacher evaluations are inherently subjective, they can be better designed and implemented to include trained evaluators using rubrics for evaluation based on models of teaching standards accepted by teachers as valid standards for their practice. Chapter 4 indicated that teachers in England felt that evaluations could provide meaningful feedback on their performance, but they were not pleased when their evaluations were tied to decisions about their salary levels. Because of their reliance on the subjective decisions of individual principals, many teachers felt that the evaluations based on individual performance plans were not accurate or helpful measures of their performance. Hassel (2002) suggested that random audits of the evaluations reduce the tendency of peers and local assessors to overrate the performance of local teachers. Odden and Wallace (2008) reported that research conducted by the Consortium for Policy Research in Education validated the effectiveness of standards-based evaluation measures as compared to student learning gains. Teachers who performed well on the evaluations had students with high performance on the standardised assessments. Some **U.S.** systems provide incentives for earning external certifications that have been validated as useful measures of teacher effectiveness. These certifications overcome the issue of local subjectivity, but they are measures of what a teacher *can* do, not necessarily what they *do* on a day-to-day basis.

As student performance data have become more available, many systems have started to rely on standardised tests to provide more objective data. As discussed earlier in this chapter's section on prerequisites to design, the quality of the assessment is critically important in this decision. Chapter 5 describes the advantages and challenges associated with VAM analysis, and notes that VAM analyses rely on psychometric properties of tests that may or may not exist in a given test. Specifically, Chapter 5 cites the requirement that tests are constructed so that gains of different students, who may be at different levels of achievement or even grade levels, can be meaningfully compared with one another. This is not always the case, making comparisons of teacher effectiveness questionable. Chapter 5 also mentions the difficulty in accurately measuring individual teachers' contributions to student learning because students are most often not randomly assigned to teachers, external events may affect student performance, student background data needed in the calculation are missing, and the small size of an individual teacher's student group adds volatility to the measures. This volatility is a problem with measures of school-wide student performance in small schools as well, where small groups of students

with uneven performance have a larger relative impact on schools' performance from year to year. VAMs are becoming more sophisticated, however, and their worth lies in their ability to address the issues of comparing performances of teachers with students from different socio-economic backgrounds and prior achievement levels, making them a more fair evaluation of teacher contributions to student growth than simple comparison of last year's and this year's tests. Systems deciding to use growth comparisons need to carefully select the methodology that has the greatest evidence of validity and reliability.

A second consideration in using standardised tests to measure individual teacher effectiveness is that only about a third of the teachers are engaged in grades or subject classes that are assessed by standardised tests, leaving the rest of the teachers without a valid means for demonstrating their effectiveness. As pointed out in Chapter 3, it can be very costly to add examinations in other subject areas, especially when one is required to have both pre- and post-tests to provide evidence of growth in student achievement. The chapter also notes that **U.S.** school districts in Philadelphia, **Pennsylvania** and Steamboat Springs, **Colorado** canceled incentive programmes and districts in **Florida** decided not to participate in the state incentive programme after they estimated the costs of creating the necessary measurement systems. Chapter 7 describes the use of Student Learning Objectives at schools, which requires teachers to set one to three learning objectives for their students that are measured by specially designed evaluations that identify the level to which the goals have been met.

While these issues may argue against incentive systems that solely reward individual teachers for growth in test scores, it does not preclude school-wide programmes that reward all teachers for school-wide improvements in student achievement in areas assessed by standardised tests or incentive plans that combine a number of input and output measures. Such plans can encourage teachers to engage in collaboration and teamwork to improve practice among all members of the staff. In fact, the **Texas** Educator Excellence Grant (Springer *et al.*, 2007) required schools to include a collaboration objective in their incentive plans to measure whether all teachers were contributing to the school's performance. That addresses the problem of "free riders" who add very little to the group's effort yet still receive the incentive award. Some systems also provide incentives for measures that work to prevent bad practices. These may include attendance, dropout, promotion and graduation rates or participation-in-testing rates to prevent pushing out students who might test poorly on the day of testing or out of school altogether. To avoid teaching to lower cognitive standards often measured on standardised tests, some systems use scores on assessments requiring higher order thinking skills as well.

Size of the incentive payment. Chapter 3 stated that the size of the maximum bonuses to teachers has increased over time. This is in part a reaction to past merit-pay programmes that were deemed ineffective because the amounts of the incentive awards were too low. Table 3.2 in Chapter 3 provides examples of **U.S.**, and international programmes that offer performance incentives ranging from less than 1% to over 360% of a teacher's monthly salary. There are a few high-paying programmes such as the Teacher Advancement Program and individual school plans in the **Texas** Educator Excellence and Governor Education Excellence Program, as well as experimental programmes in **India**, **Israel**, **Kenya** and the **U.S.** (described further in Chapters 1, 2 and 3). In a pay-for-performance experiment in the Nashville, **Tennessee** schools in the **U.S.**, lower secondary school mathematics teachers could earn an additional USD 15 000 per year, an amount representing 280% of the average monthly salary in the district. Another experiment conducted by the National Center on Performance Incentives offered bonuses of USD 6 000 to members of teams in grades 6-8. Through these experiments, analysts are attempting to learn whether teachers respond to substantial bonus awards. Odden and Wallace (2007) suggested that performance incentives should be between 4% and 8% of the annual salary to provide meaningful incentives to teachers to change behaviours. Heneman, Milanowski and Kimball (2007) agreed that teachers had to value the incentive and suggested that districts should not start performance incentive programmes if the amount of the incentives were not of sufficient size to motivate teachers.

For shortage subjects and hard-to-staff schools, the amounts of rewards may need to be larger. A study of math, science and engineering students (Milanowski, 2003) suggested that, to become a teacher, students would accept a salary bonus equal to half the difference between the teacher's salary and the private-sector salary. At that time, that was an additional USD 5 000, but a system would have to keep the amount competitive. Odden and Wallace (2008) suggest that systems start with USD 5 000 and determine whether that has had the desired impact on recruitment and retention. Hanushek, Kain and Rivkin (2004) found in their Texas research that it would take from 20-40% of annual salary to attract effective teachers to hard-to-staff schools. As an incentive for going to a hard-to-staff school, the above-mentioned **Denver** ProComp programme (Odden and Wallace, 2008) offered only a 3% of base salary bonus and got great results – an eightfold increase in applicants for open positions. Other districts have offered significantly higher bonuses: Fairfax County, **Virginia** offers a bonus of 7% of base salary, while Miami-Dade School District in **Florida** offers 20% of base salary. Fairfax has seen significant improvements in school achievement levels since adopting the incentive. Again, there has not been sufficient research to determine the optimal incentive size for either shortage subjects or hard-to-staff schools.

In the meantime, educators are learning from one another and from the private sector. Varadarajan and Futrell (1984) reviewed surveys in the private sector indicating that incentives would have to range from 3.5-11.5% of annual salary to be considered meaningful by employees and found in their study that 7.5% was the smallest meaningful incentive. Looking at recruitment incentives in the private sector in the **U.S.**, Kowal, Hassel and Hassel (2008) found that the range varied from 35% to 86% of base salary to recruit and retain employees in hard-to-staff positions. In the U.S. military, over 7% of the total salary budget is spent on special and incentive pays to recruit and retain personnel, with specific incentives for assignments to remote or dangerous locations. Individuals could receive up to USD 200 000 for knowledge and skills needed in dangerous locations. Of course those were life-threatening environments, not just less desirable locations.

Choices in the programme. There are two aspects of choice that the compensation committee should consider. The first is whether participation in the incentive programme will be voluntary or mandatory. The choice can be made at the school level as in the Teacher Advancement Programme described in Chapter 1, where 75% of the teachers at a school must agree to participate before the programme can be implemented. Or the choice can be made at the system level that participation is mandated for all new employees, but voluntary for previously employed teachers, as in the **Denver** ProComp programme. Or the decision can be left to the state/region to determine if it wishes to participate in a national programme, as in **Florida** and **Texas**. Allowing voluntary participation is a way to implement an incentive programme without fighting the minority of teachers who are categorically opposed to the concept of incentives. As the programme proves its value over time, many of the resisters will join the group. If the majority of new teacher-compensation funding goes into the incentive programme, rather than increases to the salary schedule, these teachers will have an explicit incentive to join the programme.

The second aspect of choice is whether to dictate the specific components of the incentive programme for all participants or set guidelines and allow individual states/regions, districts, or schools to develop their own plans within those guidelines. The higher level can require approval of the plans prior to implementation to ensure that the intents of the programme are carried out. This option allows the entity accountable for results to design its own plan within the given parameters that will best address the areas of need for its students and staff. While this is a more complex undertaking, it may also present an ideal opportunity to experiment and evaluate the best combination of incentive components. It is absolutely critical that external evaluations are required and funded if the system is to learn from the experiments taking place.

Supplement or replace the single salary schedule? By the time the committee has discussed all of these options, it may be ready to propose a compensation system that supplements the current salary schedule or one that

replaces it. What is critical is that the incentive programme encourages teachers to focus on improving the achievement of all students, to focus on higher order thinking skills as well as knowledge that can be easily remembered, and to work with other teachers to improve the knowledge and skills of the teaching staff and the learning of the students. There are still major issues, however, to decide. If the programme supplements the existing single salary schedule, the system commits to continuing funding the steps and lanes of the salary schedule. This has two disadvantages: it continues to fund the two areas shown to have weak correlations with student achievement: experience and additional degrees; and it reduces funds available to finance incentive awards in future years. That will require the system to find new money each year to continue the incentive programme, and it will leave the system vulnerable to critics who wonder why teachers should receive salary increases based on experience and additional incentives for doing their jobs.

The choice to replace the salary schedule, while logical and fiscally sound, will raise much more opposition than simply supplementing it. Teachers who know that they are either not good enough or are not working hard enough to earn incentives will fight to maintain the current salary schedule. They will rightfully argue that this was the contract under which they chose to enter teaching and it should not be changed during their career. This makes the argument for making the system mandatory for new teachers and voluntary for those who believe they made a bargain for life. It is easier to design a new system that provides some continued credit for experience or performance levels as the Vaughan Next Century Learning Center, described in Chapter 2. While teachers get credit for years of experience through year 11 and once again at year 15, the rest of the compensation comes from reaching higher levels of competence, evaluated using detailed rubrics; receiving one external certification or degree; and contributing to school-wide gains in student achievement. The **Cincinnati** programme also replaced the salary schedule, but it was overturned because it did not protect teachers' previous salary levels when they moved into the new system and did not perform at the required level to maintain their previous level of compensation (Max and Koppich, 2007). It would be wise to heed that warning. A system will not gain acceptance if teacher compensation can fluctuate below previously established levels of guaranteed salaries. Teachers understand if they do not receive the same amount of incentive awards each year, but their previous salary must remain stable.

KEY IMPLEMENTATION ISSUES

Odden and Wallace (2008) point out that the failure of most new pay plans is due to implementation flaws rather than design flaws. This was certainly the case in the **Cincinnati** plan discussed in Chapter 2. In Hillsborough, **Florida**, Jacob and Springer (2008) found that most teachers did not understand the pay-for-performance systems the state had established. Designers must engage stakeholders in the design and then communicate the plan in detail to teachers and other educators in the system if they want to ensure longevity of the plan. Incentive programmes can hardly act as motivators if teachers do not understand what behaviours are rewarded under what conditions. Chapter 7 provides a checklist that can be used to ensure that all key areas have been addressed in the implementation plan.

Design of a pilot programme to test design

While the compensation committee may have designed the perfect plan, the plan should be tested in a pilot to identify any challenges with the plan itself or its implementation before rolling it out to the entire system. If the data-management system cannot actually access and analyse the data in a timely and effective manner or if the training for the new evaluation system did not prepare teachers and principals for effective implementation, the system can make appropriate changes prior to the full programme launch. A pilot also enables the system to try variations of the plan in different settings to see which has the most beneficial effect. For example (Rogers, 2009), a system may wish to test individual teacher, group and school-wide incentive programmes in different schools in similar settings to determine which is most effective. Alternatively, the system may wish to allow each state/

region to design its own programme within system parameters to enable states/regions to select those options that best meet the needs of their schools. Or the national government could put out a Request for Proposals with clear parameters to schools to enable each school-level committee to design an incentive plan that meets its needs.

A pilot also allows the system to implement the specific learning measures and performance measures within a small number of representative schools within the state/region or district and analyse the results to ensure they are what the committee anticipated. This can enable the system to identify teachers' and administrators' responses to the plan or to specific components. Careful monitoring of the responses of teachers, administrators, unions, and external stakeholders to the plan from announcement of the programme throughout the pilot period will provide important information. If valid concerns are expressed, the system can work to accommodate appropriate changes before the plan goes system-wide. It also allows the system to try out its communication strategies on a small scale and learn which are most effective. Piloting is an essential component in a successful implementation of a programme that challenges the status quo in compensation. The length of the pilot will vary depending on the complexity of the design of the plan and the desire to test different options in different states/regions, districts or schools. Odden and Wallace (2008) suggest a minimum of one year, but **Denver** piloted ProComp for four years prior to full implementation. It depends on the size and variety of the system how long the pilot should last.

Creation of an evaluation plan

Throughout the chapters, it has been noted that there is little evaluation of current incentive programmes. Without good evaluation, individual systems cannot improve their plans over time, determine the most effective options, or identify the optimal size of incentive rewards for the different categories of incentives. If the programme is not contributing to the improvement of student achievement, the system needs to make adjustments to its plan. Even if student achievement is improving, the system will not know what caused the improvement without an evaluation that isolates the effects of the incentive programme from other influences and activities. It is important the the evaluation be contracted to an objective third party. While the system may have the expertise to evaluate the programme internally, programmes that affect the compensation of employees are best done outside of the organisation. This will remove concerns that the system is confounding data or withholding results that indicate a problem with the system. The evaluation will have both formative and summative components to catch the challenges identified in the pilots and ensure they are addressed before full implementation. Good evaluation practices are being developed and collected by the Center for Educator Compensation Reform of the **U.S.** Department of Education. Review of these practices will inform systems what they might include in the evaluation of their own programmes. Annual surveys of the attitudes of teachers, administrators, students and parents starting at the beginning of the development process provide excellent feedback on the impact of the programme on attitudes and practices. The results should be used to determine if and when changes should be made to the incentive programme. As participants become more comfortable with the programme, they will be able to suggest refinements that could improve the outcomes for students.

Rogers (2009) suggested that there are several ways to structure the pilot programme to provide insights into the impact of its components: random-assignment trial, index-based selection of schools, or matched pilot and control schools. The first structure provides the clearest evidence of impact and requires the system to randomly select a group of schools to implement the pilot and a comparison group of randomly-selected control schools. The second looks at schools ranked according to an index of poverty and selects schools just below a given cutoff for implementation of the pilot and just above the cutoff point for the control group. The schools will be very similar so the system can trust that differences in results are due to the implementation of the incentive system pilot. The third structure calls for the system to pilot the incentive programme in a selected group of schools and create a control group by identifying schools with characteristics matched to the pilot schools.

The system needs to develop a long-term evaluation plan that follows the schools and individuals receiving incentives to determine changes in their behaviours and results. The evaluation should measure student achievement as defined in the system's strategic plan or improvement plan. This may be based on test scores, analysis of student work, reduction in student repetitions in grade, improvements in graduation rates, and/or success in postsecondary education. Multiple measures of student achievement are essential to ensure that improvements in student performance on high-stakes tests are accurate evidence of student learning and not the results of teaching to the test or other corruptions of the learning programme. It should also look at teacher behaviours, such as teacher attendance; teacher transfers within the district; teacher attrition from the district; results of reviews of teacher classroom behaviours using trained observers and a standards-based evaluation system; tests/certification of knowledge and skills in specific areas identified as critical to student success; or surveys of principals, students and parents about the quality of teaching at the school. In addition, the system will want to hold focus groups and administer attitude surveys annually to teachers to learn their perceptions of and reactions to the incentive programme. Through these measures, the system will be able to determine if the incentive programme is having the desired effects and avoiding the unintended negative consequences sometimes reported with incentive programmes. These evaluations will enable the system to make appropriate changes to the compensation system based on evidence, not just anecdotal reports.

Alignment of compensation plan with system improvement plans

For the compensation plan to be successful, teachers must have an opportunity to engage in evaluation, feedback and improvement activities that identify the areas in which teachers need to improve and provide assistance in making those improvements. Thus the compensation plan must be part of the instructional improvement plan for the system. Specifically, the goals of the system's compensation plan must be aligned to the improvement plan. The specific areas of incentives for performance should be the same as those in the objectives of the system's improvement plan. If the system has set out specific areas for professional development based on student achievement results in its improvement plan, then professional development in those areas should be identified for incentive awards, if incentives for knowledge and skills are included in the compensation plan design. The professional development programme must include ongoing programmes for content and pedagogy, with follow-up support and coaching available to individual teachers who need to improve their practice. In addition, the instructional support plan needs to ensure that principals and teachers are well trained in the evaluation system on which the compensation plan depends.

The plan must also align to activities of other departments of the system. The increased demand for data collection and analysis must be accommodated in the annual plans of the research and evaluation department and the budgets of the finance department. If the evaluation system requires multiple observations by trained evaluators, the Human Resources (HR) Department must ensure that trained evaluators are available and funds for their payment are part of the department's budget.

The HR Department should use the compensation plan in its recruitment and hiring programme, focusing on the subjects and schools requiring additional teachers, and recruiting teachers who are interested in a performance-based compensation system. It should be part of the human-capital-management system. Chapter 7 argues that the whole human-resources function must be aligned within the system, including the incentive programme and revised compensation plan. The Chapter also includes a figure developed by Wurtzel and Curtis (2008) to represent the necessity of creating a complete human capital management system. While some of the functions, such as induction support and professional development, may lie in other departments, the functions must be coordinated to be effective. If the work of the compensation committee leads to a new performance-management system, including standards-based evaluations for all instructional personnel, then the compensation system must reward the practices advocated in the performance-management system. In addition, teachers must see

how the professional development and coaching they receive will enable them to improve their evaluations, but, more importantly, improve the achievement of their students. The incentives need to be tied to measures of student and teacher performance that will lead to continuous improvement. A discussion of the performance-management system would be a great introduction to the work of the compensation committee.

If the system has decentralised authority to schools, then some or all of the decisions on incentives should be made by a compensation committee at the school, not just one at the district, state/regional or national level. Individual schools may not have exactly the same needs as other schools do. For example, one school may need teachers trained to focus on the needs of English language learners and would like to use market-based incentives to attract those teachers, not teachers of mathematics and science as needed in the school system as a whole. Criteria can be established to ensure that schools are following the major tenets of the system's plan design, and schools can be required to provide evidence of reliable measures for the incentive award areas.

Identification of “champion” for the plan

Throughout the period from initial marketing through full implementation of the pay-for-performance plan, the system needs a champion. Chapter 7 stresses the importance of having the chief educational leader on board, but also sharing the role with the head of the teacher union. In Austin, **Texas**, both the superintendent and union president were present at all major announcements and visited schools together to reinforce their partnership. In this way, it was clear that the incentive programme was designed *with* teachers and was not a top-down mandate done *to* teachers. Chapter 8 describes how the president of the American Federation of Teachers, Randy Weingarten, used that phrase as her 2009 slogan: “with us, not to us” (Quaid, 2009). The issue is not just a **U.S.** issue: Pavez Urrutia (2008) suggested that the lack of collaboration of the teachers’ union in the design of the programme was one of the problems with the **Chilean** incentive programme.

The work of the leader of the system must be reinforced throughout the implementation by the involvement of top management staff. The incentive programme cannot be undertaken as just a compensation programme, but rather as the key to achieving the goal of the system: improvement of student achievement through the coherent focus on instruction delivered by effective teachers.

Regular communication with stakeholders

The communications strategy that the compensation committee develops must remain a high priority throughout the development and implementation of the pilot programme as well as the initial programme launch. It is easy for systems to assume that such an explicit communications strategy is no longer needed, but the constituent groups may interpret silence as trying to hide something about implementation. When communication remains open and regular, stakeholders have greater trust in the system.

Development of broad political support

Chapter 7 states that incentive programmes have received positive support from high-level political figures in the **U.S.** Romaguera (2008) pointed out that it is difficult to reform salary structures in Latin America because the education sector involves multiple stakeholders, each of which has significant political weight. They also have very strong ideological positions about issues of performance pay, working conditions and the individual rights of teachers. The challenge in implementing an incentive programme in any system is to identify high-level political supporters who are able and willing to galvanise broad political support for the programme. There will clearly be opponents who will work to sway political and public support to their side as well. Chapters 6 and 8 discuss examples of teachers unions that play a significant role in the design and acceptance of any change in compensation systems. This is a reason why systems will want to include broad representation on the committee.

Not only will discussions within the committee identify the sticking points for certain stakeholder groups, but the committee will also be able to address them as it develops the plan. That will make it more difficult for opponents to claim that their views were not included in the deliberations.

In addition to specific stakeholder groups, the public should be included in communications from the start so that they understand what is happening and why. Educators tend to keep information restricted to their inner circle of decision makers, but it is the public that will pay for the plan and send children to the schools and therefore has a great stake in its development. Highlighting the need to consider more inclusive models of collaboration and communication, Chapter 6 includes a discussion of a study by Vaillant (2005) of Latin American countries and the need to include participation of families and educators at the local, state/regional levels in the discussions of education reform. This collaboration not only identifies issues that pertain to the specific circumstances of their locales but also develops support.

With frequent and regular information going out about the programme, the system can work to maintain support while it develops the programme. With implementation, the opponents may be more active, raising objections to specific issues or providing examples of what appear to be unfair payments or lack of payments. However, if there is broad support for the programme and the system takes the concerns seriously and acts in good faith, the opposition will lessen over time.

TRAINING OF SCHOOL LEADERS AND TEACHERS

Training on the compensation system. Prior to implementation of the pilot, it is essential to prepare teachers and principals for the impact the plan will have on current compensation and the opportunities for added compensation through incentive payments. Assuming that communication has been ongoing throughout the development process, training should include review of the purpose, development process, and decisions that were made by the compensation committee and the administration on options for incentives. It should then move on to the details of the plan, including the ways in which teachers can earn incentives and what measures will be used to determine who receives them. Handbooks, such as those developed in the **Denver** ProComp programme (Denver Public Schools, 2006), can be useful in providing the level of detail teachers and principals will want. Members of the compensation committee, especially teachers and principals, can be helpful in answering questions that may arise.

The training would start with the schools engaged in the pilot, but it should continue into the other schools in the system, as well. It is better for teachers to hear directly what the new compensation system is and why it was developed, rather than hear it from teachers or others who are less knowledgeable or against the concept of incentives altogether. As the pilot proceeds, teachers in the pilot schools should be asked whether the initial training was sufficient for them to understand and take advantage of the incentives offered. If not, the changes they recommend should be put in place immediately to ensure other teachers in the system are better informed.

Training and provision of support services. The second type of training should make information on the support services available to teachers and principals as they implement the new compensation system. Again, all of the training should include surveys of teacher and principal attitudes and reports of what they have learnt, in what areas they require more training, and in what areas further explanation of available support services is needed. Training should be provided on the analysis of data from student assessments in all schools, both summative standardised assessments and formative assessments created and used at the school and classroom levels. Teachers need to understand how they can use the data from assessments to improve their individual and collective practices.

In addition, training is needed on what the system considers effective teacher practice, and it should be provided to teachers and principals in the pilot schools before the pilot programme is implemented. If the system has adopted a standards-based evaluation model, the training should be provided to both teachers and principals on what the expectations are and how the evaluation will be implemented. Principals will need more specific and intensive training on how to assess teacher practices in the context of the standards-based model so that they become proficient. Leaders of the system may wish to establish a certification program to identify principals who have demonstrated proficiency in reviewing instruction using the evaluation model, so that teachers can be assured that the evaluation is completed with fidelity to the model. This is especially important if incentives are tied to the evaluation results, but it is also important in assuring teachers that they are receiving valid feedback on their performance.

With the leadership of the principal, the school can create a collaborative community of learners. Teachers with great strengths in areas in which the school needs to improve can act as coaches and mentors to other teachers. Teachers should receive opportunities to participate in professional development on their areas of weakness. With the ongoing support of the principal and mentor teachers, teachers will learn how they need to change their practices to improve the achievement of their students. This is a critical part of the motivation equation: teachers must believe that they can improve their practice (Heneman, 1992). Without training and support, that will not happen. Chapter 6 cites research on teacher evaluation that suggests that receiving high-quality feedback and coaching improves the acceptance of the evaluation process by those being evaluated (Heneman and Milanowski, 2003).

REVISIONS IN PLAN FROM PILOT RESULTS

As indicated in an earlier section of this chapter, the evaluation plan needs to include a rigorous evaluation of the pilot programme so that the system's leaders can review results and determine what changes, if any, are needed prior to full implementation. Systems have implemented pilots of different length; the purpose is to identify any problems or issues that had not been identified in the design phase so they can be addressed prior to full implementation. That may take a single year or multiple years. If the system implemented different models in the pilot, the evaluations provide evidence for deciding which model will be adopted or, if all models appeared to be equally effective, if states/regions, districts or schools can choose the model that best fits their individual contexts. Developing public support may be easier with flexibility in the choice of models. Most countries still battle over "top-down" mandates versus local control of policy.

Even when the models are effective, there can be other problems that should be addressed prior to full implementation. If teacher surveys indicate that pilot teachers do not understand the components of the plan, it may be helpful to engage teachers in the development of new materials for training and dissemination. If the problems identified are administrative – e.g. not getting the data to the research department in a timely manner or not getting the necessary information to accounting in time to make the scheduled payments, the system must determine if a second pilot year is needed to fix those problems. If the technology system is not programmed to provide the right payments to the right teachers, the problem must be addressed immediately, and the programme should not move forward until it is corrected.

It would be unwise to move into full implementation before the pilot is considered successful in the eyes of the teachers, administrators, teacher union leaders and civic leaders who have been involved in its design, implementation and support. Even with a successful pilot, full implementation may place new strains on the administrative systems. If the system is unsure it is ready for full implementation, expanding the pilot to a larger share of schools might be the best next step.

Full implementation

Transitions from pilot. To ensure the success of the full-scale implementation, the system will have to review its efforts prior to the pilot programme in the areas of communications, training and support of teachers and principals, and in readying the administrative systems in human resources, finance, and computer services. If the pilot indicated weaknesses in any of those areas, they must be addressed before implementation moves forward. Once the system is confident that it has addressed these issues in a satisfactory manner, making adjustments where indicated by the evaluation of the pilot, it is ready to begin full implementation. Ongoing evaluation of the programme will provide critical feedback for system improvement.

The first step is to launch the system-wide communications campaigns, along with a public information campaign involving print and technological media. Members of the compensation committee can play a key role in this effort, from giving interviews on regional and local television stations, to writing opinion editorials in newspapers and on websites, to making speeches at meetings of civic organisations, educator conferences, and conferences with parents and taxpayers. Written material as described in the communications plan should be revised and disseminated. Again, it will be an important show of support to have both the leader of the education system and the leader of the teacher union/association standing together in person, in print, and in the media. The communications need to occur at regular intervals and continue throughout the life of the programme.

Training of teachers and principals in the details of the compensation system and the evaluation and professional support systems should be completed in the semester before full implementation. If the system has been scheduling training throughout the pilot period, this round of training should focus on lessons learnt and changes made to the system as a result of the pilot experiences. In this way, teachers can have greater confidence that the concerns raised in the pilot were heard and acted upon. Chapter 6 provides examples of systems that have had to revise their systems and communicate with teachers about those changes. Because of challenges from teachers and their teacher associations about the fairness of its incentive system (Behrstock and Akerstrom, 2008), Houston developed four levels of training on the value-added method. While the training is not required, the system will recognise schools with a high percentage of staff that complete the four levels of training. Training on the compensation system must also be provided to new employees as they enter the system.

Understanding the system is necessary, but systems must also provide professional development support to teachers and principals to enable them to change and improve their practices. This will only be accomplished in sustained professional development programmes that link to classroom practices and are reinforced through coaching in the classroom and at the school. Evaluation of the initial programmes during the pilot programme will provide suggestions for improvement prior to full implementation. Teacher surveys and principals' evaluations can be used to evaluate the effectiveness of the professional development programme on an ongoing basis. Effective professional development needs to become a part of the system's operations, and regularly assessed and improved, if the system is to sustain improved instructional practices.

Finally the administrative services need to be refined to ensure accuracy and timeliness of incentive payments. The more complicated the incentive programme, the more opportunities for missing data, delivery delays, analysis problems and technology bugs to occur. The system may wish to conduct a mock implementation using data from previous years to ensure the data management system can handle the size of data runs needed for full implementation. The key is to ensure the administrative system is working, and full implementation should be delayed until the system is ready to make that happen.

Political challenges to full implementation. Chapter 8 notes that the movement from experimental programmes to public policies is political and typically involves union negotiation; there are likely to be clear tensions—one side focused on performance and the other on across-the-board and greater pay for inputs. The movement to

public policies also means less controlled programmes, and once every school is involved, there are greater risks. Although the plan may be designed to provide both positive incentives for productive behaviours and disincentives for negative behaviours, monitoring strategies and regulations can help to ensure that the programme is well implemented, fair to teachers and administrators, and productive for students in the districts. If not, public acceptance of the programme will suffer. That is why the evaluation plan must include strategies to monitor and assess behaviours and attitudes along the way. In addition, teachers and principals should have access to a hotline or web link to identify questions and problems, and the system must respond quickly to problems. Systems can work with the third-party evaluator or local universities to conduct controlled studies to inform changes. Teachers are often more willing to be honest with an external interviewer than with the system's representatives.

CONCLUSION

As this chapter and, indeed, all of the chapters in this volume indicate, designing and implementing a compensation system that includes teacher evaluation and performance incentives will not be an easy task. Given what we know about the impact of effective teachers on student achievement, however, systems need to do whatever it takes to establish environments that will attract and retain effective teachers. System design can take place at the national, state /regional level, or with the local education authority. If done at the national or state/regional levels, the components of the plan should be designed with input of local education authorities so that its incentives match their needs. To maximise that possibility, the system might include a set of options from which the local education authorities can select. For example, the plan might offer an incentive for subject-shortage areas, but leave the specific areas to the school to determine. Or it might include incentives for hard-to-staff schools, without specifying why the school is hard-to-staff. The local education authority could be required to provide evidence, such as teacher attrition rates to justify this classification. In the same way, the plan might require an evaluation system with incentives for specific levels of knowledge and skills, but leave the specific model for teacher evaluation to be selected by the local education authority from a list of approved models. In all of these examples, the important point is that the plan be designed with educators at all levels and that they understand the plan and its focus on improving student achievement. Teachers also need to believe that the system is fair, transparent and supported by opportunities for teachers to improve their practice.

A well-designed evaluation and reward plan can be vital to student success, accompanied by positive working environments in which teachers can flourish, and with efficient systems of administration and support. With the collaboration of all interested stakeholders in its design, a new performance-based compensation system can recognise and reward teachers who make the greatest contributions to students' growth and success, and thus attract and retain high performing teachers to its classrooms. Since the future of each country's economy and society depends on the quality of education of its children, no education system can afford to do any less.

References

Behrstock, E. and J. Akerstrom (2008), *Performance Pay in Houston*, Center for Educator Compensation Reform, U.S. Department of Education, Office of Elementary and Secondary Education, Washington, D.C. Retrieved August 2009 from <http://cecr.ed.gov/guides/HoustonCaseSummary.pdf>.

Che, Y. and S. Yoo (2001), "Optimal Incentives for Teams", *The American Economic Review*, Vol. 91, No. 3, pp. 525-541.

Chiang, F. and T. Birtch (2005), "A Taxonomy of Reward Preferences: Examining Country Differences", *Journal of International Management*, Vol. 11, Issue 3, pp. 357-375.

Danielson, C. (2008), *The Handbook for Enhancing Professional Practice: Using the Framework for Teaching in Your School*, Association for Supervision and Curriculum Development, Alexandria, VA.

Data Quality Campaign (2009) *10 Essential Elements of a State Longitudinal Data System*, Austin, TX, Data Quality Campaign Website, <http://www.dataqualitycampaign.org/survey/elements>, accessed May 2009.

Delisio, E. (2003), "Pay for Performance: What Went Wrong in Cincinnati?", *Education World*, January 28, http://www.educationworld.com/a_issues/issues/issues374b.shtml, accessed May 2009.

Denver Public Schools (2006), *ProComp at a Glance: A Quick Reference Handbook*, Denver, CO., <http://www.dpsk12.org/manila/programs/denverprocomp/ProCompHandbook.pdf>, accessed June 2009.

Educational Testing Service (2009), *Overview of PRAXIS III*, Educational Testing Service, New York, <http://www.ets.org/portal/site/ets/menuitem.1488512ecfd5b8849a77b13bc3921509/?vgnnextoid=f4c20ef8beb4b110VgnVCM10000022f95190RCRD&vgnnextchannel=19b00ef8beb4b110VgnVCM10000022f95190RCRD>, accessed May 2009.

Guthrie, J. and C. Prince (2009), "Paying for and Sustaining a Performance-Based Compensation System", Center for Educator Compensation Reform, Washington, D.C.

Hanushek, E. (1992), "The Trade-off Between Child Quantity and Quality", *Journal of Political Economy*, Vol. 100, No. 1, pp. 84-117.

Hanushek, E. (1996), "Outcomes, Costs, and Incentives in Schools", in E. Hanushek and D. Jorgenson (eds.), *Improving America's Schools: The Role Of Incentives*, Board on Science, Technology, and Economic Policy, National Research Council, National Academies Press, Washington, D.C., pp. 29-52, <http://www.nap.edu/books/0309054362/html/>, accessed March 2009.

Hanushek, E., J. Kain and S. Rivkin (2004), "Why Public Schools Lose Teachers", *Journal of Human Resources*, Vol. 39, No. 2, pp.326-354.

Hanushek, E. and A. Lindseth (2009), "Performance-Based Funding", *Education Week*, Vol. 28, Issue 33, pp. 28-30.

Hanushek, E. and L. Wößmann (2009), "Schooling, Cognitive Skills, and the Latin American Growth Puzzle", *Working Paper 15066*, National Bureau of Economic Research, Cambridge, MA.

Hassel, B. (2002), "Better Pay for Better Teaching: Making Teacher Compensation Pay Off in an Age of Accountability", *21st Century Schools Project*, Progressive Policy Institute, Washington, D.C.

Heneman, R.L. (1992), *Merit Pay: Linking Pay Increases to Performance Ratings*, Addison-Wesley, Reading, MA.

Heneman, H., and A. Milanowski (2003), "Continuing Assessment of Teacher Reactions to a Standards-based Teacher Evaluation System", *Journal of Personnel Evaluation in Education*, Vol. 17, No. 2, pp. 173-195.

Heneman, H., A. Milanowski and S. Kimball (2007), "Teacher Performance Pay: Synthesis of Plans, Research, and Guidelines for Practice", *CPRE Policy Brief RB-46*, University of Pennsylvania, Philadelphia, PA, February.

Jacob, B. and M. Springer (2008), "Teacher Attitudes Toward Pay for Performance: Evidence from Hillsborough County, Florida", paper presented at the National Conference on Performance Incentives: Their Growing Impact on American K-12 Education, Nashville, TN, February.

Koedel, C. and J. Betts (2009), "Value-Added to What? How a Ceiling in the Testing Instrument Influences Value-Added Estimation", *NBER Working Paper 14778*, National Bureau of Economic Research, Cambridge, MA., <http://www.nber.org/papers/w14778>, accessed June 2009.

Koppich, J. (2008), "Toward a More Comprehensive Model of Teacher Pay", paper presented at the National Conference on Performance Incentives: Their Growing Impact on American K-12 Education, Nashville, TN, February.

Koppich, J. and C. Prince (2007), "Stakeholder Engagement and Communication", Center for Educator Compensation Reform, U.S. Department of Education, Office of Elementary and Secondary Education, Washington, D.C.

Kowal, J., B. Hassel and E. Hassel (2008), "Financial Incentives for Hard-to-Staff Positions: Cross-Sector Lessons for Public Education", Center for American Progress, Washington, D.C.

Lavy, V. (2002), "Evaluating the Effect of Teachers' Group Performance Incentives on Pupil Achievement", *Journal of Political Economy*, Vol.110, No.6, pp.1286-1317.

Lazear, E. (1996), "Performance Pay and Productivity", *NBER Working Paper No.W5672* National Bureau of Economic Research, Cambridge, MA., SSRN: <http://ssrn.com/abstract=225573>, accessed May 2009.

Lazear, E. (1998), *Personnel Economics for Managers*, Wiley, New York.

Manzi, J. (2008), "Individual Incentives and Teacher Evaluation: The Chilean Case", paper presented at the International OECD/Mexico Joint Conference, The Quality of Education, Mexico City, December.

Max, J. and J. Koppich (2007), "Engaging Stakeholders in Teacher Pay Reform", Emerging Issues, Report No. 1, Center for Educator Compensation Reform, U.S. Department of Education, Office of Elementary and Secondary Education, Washington, D.C.

Milanowski, A. (2003), "An Exploration of the Pay Levels Needed to Attract Students with Mathematics, Science and Technology Skills to a Career in K-12 Teaching", *Education Policy Analysis Archives*, Vol. 11, No. 50, <http://www.epaa.asu.edu/epaa/v11n50.pdf>, accessed May 2009.

Milanowski, A. (2006), "Performance Pay System Preferences of Students Preparing to Be Teachers", *WCER Working Paper No. 2006-9*, Consortium for Policy Research in Education, Wisconsin Center for Education Research, Madison, WI.

Muralidharan, K. and V. Sundararaman (2008), "Teacher Incentives in Developing Countries: Experimental Evidence from India", paper presented at the National Conference on Performance Incentives: Their Growing Impact on American K-12 Education, Nashville, TN, February.

Murnane, R. and D. Cohen (1985), "Merit Pay and the Evaluation Problem: Understanding Why Most Merit Pay Plans Fail and a Few Survive", Institute for Research on Education Finance and Governance, Stanford University, Palo Alto, California, http://www.eric.ed.gov:80/ERICDocs/data/ericdocs2sql/content_storage_01/0000019b/80/2f/4e/ca.pdf, accessed March 2009.

Neal, D. (2008), "Designing Incentive Systems for Schools", paper presented at the National Conference on Performance Incentives: Their Growing Impact on American K-12 Education, Nashville, TN, February.

Odden, A. and C. Kelley (2002), *Paying Teachers for What They Know and Do: New and Smarter Strategies to Improve Schools*, Corwin Press, Thousand Oaks, CA.

Odden, A. and M. Wallace (2007), "Rewarding Teacher Excellence: A Teacher Compensation Handbook for State and Local Policy Makers", Consortium for Policy Research in Education, University of Wisconsin, Madison, WI, [http://cpre.wceruw.org/publications/TComp%20Handbook%20Feb%2028%2007%20Final%20\(3.05.07\).pdf](http://cpre.wceruw.org/publications/TComp%20Handbook%20Feb%2028%2007%20Final%20(3.05.07).pdf), accessed March 2009.

Odden, A. and M. Wallace (2008), *How to Achieve World Class Teacher Compensation*, Freeload Press, St. Paul, MN., <http://www.freeloadpress.com/>, accessed March 2009, free registration required.

Pavez Urrutia, J. (2008), "Incentives and Motivation for the Professors: The Chile Case", paper presented at the International OECD Mexico Joint Conference, The Quality of Education: Workshop on Teacher Incentives and Stimuli, Mexico City, December.

Pfeffer, J. (1995), *Competitive Advantage Through People: Unleashing the Power of the Work Force*, Harvard Business School Press, Boston, MA.

Quaid, L. (2009), "Teachers Urge Obama to Collaborate to Fix Schools", *The Associated Press*, July 13.

Rivkin, S., E. Hanushek and J. Kain (2005), "Teachers, Schools, and Academic Achievement," *Econometrica*, Econometric Society, Vol. 73(2), pp. 417-458, 03.

Rogers, F.H. (2009), "Piloting and Evaluating Teacher Incentive Programs", PowerPoint presented at the OECD Mexico Joint Workshop on Education Quality Standards and Assessment, Mexico City, July.

Romaguera, P. (2008), "Incentives and Stimuli for Teachers: The Case of Chile", paper presented at the International OECD Mexico Joint Conference, The Quality of Education, Mexico City, December.

Sanders, W. and J. Rivers (1996), "Cumulative and Residual Effects of Teachers on Future Student Academic Achievement", *Research Progress Report*, University of Tennessee Value-Added Research and Assessment Center, Knoxville, TN.

Santibáñez, L., J. Martínez, A. Datar, P. McEwan, C. Setodji, and R. Basuto-Dávila (2007), *Breaking Ground: Analysis of the Assessment System and Impact of Mexico's Teacher Incentive Program "Carrera Magisterial"*, RAND Corporation, Santa Monica, CA.

Schacter, J. and Y.M. Thum (2004). "Paying for high and low-quality teaching", *Economics of Education Review*, 23, 411-430.

Sclafani, S. (2008), "Rethinking Human Capital in Education: Singapore as a Model for Teacher Development", paper prepared for The Aspen Institute Education and Society Program, Aspen Institute, Washington, D.C., <http://www.aspeninstitute.org/sites/default/files/content/docs/education%20and%20society%20program/SingaporeEDU.pdf>, accessed May 2009.

Springer, M., M. Podgursky, J. Lewis, M. Ehlert, C. Gardner, B. Ghosh-Dastidar, O. Lopez, C. Patterson and L. Taylor (2007), "Governor's Educator Excellence Grant (GEEG) Program: Year One Evaluation, National Center on Performance Incentives, Vanderbilt Peabody College], Nashville, TN, <http://ritter.tea.state.tx.us/opge/progeval/TeacherIncentive/index.html>, accessed April 2009.

Springer, M., M. Podgursky, J. Lewis, M. Ehlert, B. Ghosh-Dastidar, T. Gronberg, L. Hamilton, D. Janssen, O. Lopez, B. Stecher and L. Taylor (2008), "Texas Educator Excellence Grant (TEEG) Program: Year Two Evaluation", National Center on Performance Incentives, Vanderbilt Peabody College, Nashville, TN., http://ritter.tea.state.tx.us/opge/progeval/TeacherIncentive/TEEG_120108.pdf, accessed April 2009.

Varadarajan, P. and C. Futrell (1984), "Factors Affecting Perceptions of Smallest Meaningful Pay Increases", *Industrial Relations*, Vol. 23, No. 2, pp. 278-286.

Vaillant, D. (2005). *Formación de formadores. Estado de la Práctica*. PREAL Documents No 25. Santiago, Chile: PREAL.

Wurtzel, J. and R. Curtis (2008), "Human Capital Framework for K-12 Urban Education: Organizing for Success, A Working Draft", Aspen Institute, Washington, D.C., <http://www.aspeninstitute.org/publications/human-capital-management-k-12-education>, accessed May 2009.

OECD PUBLICATIONS, 2, rue André-Pascal, 75775 PARIS CEDEX 16
PRINTED IN FRANCE
(87 2009 03 1 P) ISBN 978-92-64-06198-9 – No. 57037 2009

Evaluating and Rewarding the Quality of Teachers

INTERNATIONAL PRACTICES

An educational system can only be as good as its teachers. This book identifies good practices in the design and implementation of evaluation and teacher incentive systems from various perspectives through formulation, stakeholder negotiation, implementation, monitoring and follow-up. It guides the reader through the challenges and lessons learned from different educational environments, drawing on experiences from 24 countries (including India, Mexico, Singapore, the United Kingdom and the United States).

The book serves as general reference for policy makers and educational authorities charged with designing, implementing and reforming teacher evaluation and incentive policies. It provides answers to the following questions:

- What aspects of performance in education should be rewarded and how should they be measured in evaluation systems?
- Who should be rewarded? Individual teachers, groups of teachers or schools?
- What mechanisms have effectively been used to reward teachers?
- How should such policies be developed and implemented to ensure stakeholder engagement and commitment?

This book has been edited by Susan Sclafani, former advisor of the government of the United States and one of the architects of its national educational policy, drawing together contributions from leading experts in the area of teacher evaluation, incentives and stimuli.

This publication is a product of the co-operation agreement established between the Government of Mexico and the OECD, which aims to improve the quality of education in Mexico.

The full text of this book is available on line via this link:

www.sourceoecd.org/education/9789264061989

Those with access to all OECD books on line should use this link:

www.sourceoecd.org/9789264061989

SourceOECD is the OECD's online library of books, periodicals and statistical databases.

For more information about this award-winning service and free trials ask your librarian, or write to us at **SourceOECD@oecd.org**.