



Title Registration for a Systematic Review: Merit Pay Programs for Improving Teacher Retention, Teacher Satisfaction, and Student Achievement in Primary and Secondary Education: A Systematic Review

Joshua Barnett, Gary Ritter, Nathan Jensen, George Denny

Submitted to the Coordinating Group of:

<input type="checkbox"/>	Crime and Justice
<input checked="" type="checkbox"/>	Education
<input type="checkbox"/>	Disability
<input type="checkbox"/>	International Development
<input type="checkbox"/>	Nutrition
<input type="checkbox"/>	Social Welfare
<input type="checkbox"/>	Other:

Plans to co-register:

<input checked="" type="checkbox"/>	No		
<input type="checkbox"/>	Yes	<input type="checkbox"/> Cochrane	<input type="checkbox"/> Other
<input type="checkbox"/>	Maybe		

Date Submitted: 5 December 2012

Date Revision Submitted: 18 January 2013

Approval Date: 29 January 2013

Publication Date: 1 March 2013

TITLE OF THE REVIEW

Merit Pay Programs for Improving Teacher Retention, Teacher Satisfaction, and Student Achievement in Primary and Secondary Education: A Systematic Review

BACKGROUND

Approximately 95% of K-12 teachers in the United States of America work in a school or district with a salary schedule that provides salary increases largely based on years of experience and number of degrees attained (Podgursky & Springer, 2007). Often referred to as the "single-salary schedule" or "lock-step schedule", this form of compensation was developed to address equity issues and has spread across the globe, where teachers most often receive increased compensation based on qualifications and years of service (inputs) rather than on measures of actual teacher and their students' performance (outputs).

Despite the single-salary schedule's popularity and longevity, a growing number of researchers have argued that this compensation approach has produced problems in schools, including problems recruiting high quality teachers into the classroom, retaining them in the profession, and holding them accountable for student achievement (Goldhaber, 2002; Plucker, Zapf, & McNabb, 2005). Due to these perceived problems, some U.S. states (e.g., Nevada, Texas) and districts (e.g., Charlotte-Mecklenburg, NC; Cincinnati, OH; Denver, CO; Douglas County, CO; Los Angeles, CA; Washoe County, NV) and federal governments (e.g., Australia, New Zealand) are investigating alternatives and moving incrementally away from the single-salary schedule in hopes of recruiting and retaining more qualified teachers who can improve student performance (Kelley, 1998, 2000; Odden & Kelley, 1997). In support of this trend, U.S. President Obama has endorsed salary reforms, including the use of merit pay, through the Race to the Top competition (awarding approximately \$4.5 billion) and the Teacher Incentive Fund competition (awarding approximately \$1 billion).

Yet, as national policymakers and school leaders consider the use of merit pay, they quickly find themselves confronted with a number of core complications inherent in the creation of such plans. These complications can be classified into three key methodological issues that make this type of compensation reform challenging.

Firstly, what evaluation instruments should be used? Most merit pay advocates contend that teacher "merit" should be based, at least in part, on the measured achievement or growth of students in a teacher's classroom. However, the idea of holding teachers accountable for student test score gains is a source of great debate and discord (see Amrein-Beardsley, 2012; Harris, 2011). Nevertheless, most merit pay plans do and will include standardized measures of student test performance. However, school leaders and policymakers need to know more about which instruments have been used and which ones are better suited for this type of compensation approach.

Secondly, should rewards be based on individual or group efforts? Some merit pay plans today are school-based; that is, the school receives some sort of rating for its overall performance over a given time period. Then, all the teachers in that school receive a bonus based on the “merit” of the school as a whole. Other plans rate individual teachers based on their individual classroom performance and allocate different reward levels to different teachers. However, school leaders do not have consistent evidence on the impact of these different approaches.

Thirdly, how should different award levels be determined for different school personnel? In schools or districts with the most limited merit pay programs, only those teachers who teach core subjects (i.e., language/reading; mathematics) with corresponding standardized assessments can participate. In other plans, non-core teachers (e.g., art or music teachers) and school support staff (e.g., custodians, aides) are eligible for awards, although the award levels and rating systems might be different. Also, some merit pay plans include awards for school administrators based on overall performance of the school or district. Therefore, school leaders need information on which types of educators and school employees to include in the bonus pool, and what levels of rewards to offer to each type of employee.

OBJECTIVES

In general, the challenge for compensation reform is that the theory makes intuitive sense—rewarding more effective teachers with additional compensation can retain those effective teachers, encourage others to become more effective, and recruit higher performing individuals into the education field. However, the methodological and practical issues associated with implementing such a system are complicated, in that no single merit pay “plan” exists; instead, there are numerous ways in which school leaders have implemented a merit pay plan for their teachers. The details of these plans are important to understand in determining the overall impact of merit pay as a compensation policy. This review intends to respond to these challenges by examining the following four objectives:

1. To what degree do merit pay programs impact student achievement outcomes?
2. To what degree do merit pay programs impact teacher outcomes (e.g., retention, satisfaction)?
3. What are the distinguishing characteristics of the most successful merit pay programs?
4. For which subgroups of teachers/students/school systems are merit pay programs most or least beneficial?

EXISTING REVIEWS

As noted above in the background section, the evidence on merit pay is unclear with regard to which types of plans are most effective, as numerous decisions must be made in creating a merit pay plan. As a result of the different types of programs and different types of evaluations of those programs, the overall picture of the impact of merit pay is difficult to summarize. Consequently, many of the policy discussions surrounding merit pay rely on a limited amount of evidence, as large-scale and multi-year programs have not historically existed nor been evaluated rigorously.

Harvey-Beavis (2003) reviewed the evidence in international education and determined that the theory of merit pay is strong, but the application in practice has not been realized. Chamberlin et al. (2002) conducted a review of evidence in British schools and reached similar conclusions, noting that a number of programs were implemented, but the evaluations of such programs were limited or poorly performed. Podgursky and Springer (2007) conducted a review of the evidence on merit pay and noted that results from evaluations of these programs were mixed but showed promise for impacting student achievement.

The value and need of a more rigorous systematic review is demonstrated by the fact that, over the previous five years (since Podgursky & Springer, 2007), merit pay programs have become more prominent in education policy circles, particularly in the United States, and more districts, states, and nations are trying different approaches.

INTERVENTION

Eligible interventions to be included are merit pay programs, “where merit pay is defined as rewards for individual teachers, groups of teachers, or schools on any number of factors, including student performance, classroom observations, and teacher portfolios. Merit-based pay is a reward system that hinges on student outcomes attributed to a particular teacher or group of teachers rather than on ‘inputs’ such as skills or knowledge” (Podgursky & Springer, 2007).

The interventions must also meet the following criteria:

- The intervention must be carried out in a K-12 school.
- The intervention must include K-12 teachers. Programs aimed towards staff development or principals/administrators will be included as long as it also includes a teacher component.
- The intervention must be implemented for at least one academic- or school-year (N.B., it may be implemented in schools with quarter, semester, or year-round calendars).

- The intervention must include a specified financial reward (e.g., additional pay, bonus) that is calculated and provided for each school year.
- The comparison groups will receive no treatment or business-as-usual (e.g. traditional salary as indicated in their district or state guidelines).
 - Generally, or at least in the United States, these comparison conditions will use the previously discussed single-salary system, where compensation is based on qualifications and years of experience. Salary decisions are generally based on district-level decisions and will apply to all teachers within a given district except those participating in the treatment condition. Comparison group teachers will continue to receive their salary increase as appropriated within the district policy. An example of this type of increase is \$250 increase in base salary per year of service for a teacher with a Bachelor’s of Education degree and a \$500 increase in base salary per year of service for a teacher with a Master’s of Education degree. Thus, a newly hired teacher with a Bachelor’s degree earning \$35,000 in 2012-13 would earn \$35,250 in 2013-14 regardless of how well the students or school performed.
- Career ladder programs will not be included in the study, where teachers advance on a set performance ladder, which more closely approximates the traditional single salary system than a merit or bonus system.

POPULATION

Target populations for this review include K-12 teachers (or their equivalent outside the U.S.) and their corresponding students. An intervention’s effectiveness could vary by subgroups of students, teachers, or schools. However, whether a study examines effects on subgroups does not affect the inclusion of the study for review. For this review, we will collect information on student characteristics including baseline achievement score, grade, gender, socioeconomic status, racial/ethnic breakdown, second-language status, and “at-risk” status (as defined by study authors) when provided. Reviewers will also collect information on teacher characteristics including individual demographics, tenure in teaching, and educational attainment as provided. School characteristics of interest for this review include location of the schools involved, school type (public, private, religious), school SES (e.g., Title I school), average class size (small, medium, large), school size (small, medium, large), and school community (rural, suburban, urban). The study must be written in English, but the study may be conducted in the United States (including the 50 states, the District of Columbia, territories, and tribal entities) or other nations. We are aware of merit pay programs operating in Australia, New Zealand, India, Israel, and other nations, and we will incorporate these studies if they meet eligibility criteria.

OUTCOMES

The study needs to include at least one measure that involves direct assessment of student achievement or a teacher outcome. For student achievement, standardized test scores or district benchmarks may be used. Additional student outcomes (e.g. attitude toward the subject they learn, motivation, and self-efficacy) are not the focus of this review and do not qualify as relevant outcome measures. Given the nature of most district and state testing standards in the United States under the No Child Left Behind Act of 2001, the authors anticipate student achievement outcomes to be based on achievement in math and reading; however, other standardized tests will be included provided they demonstrate sufficient reliability and validity. Reliability will be assessed using the following standards specified by the What Works Clearinghouse (WWC) Version 2.0 standards: internal consistency (minimum of 0.50), temporal stability/test-retest reliability (minimum of 0.40), and inter-rater reliability (minimum of 0.50). Over-alignment issues will also be considered, as outcome measures should not be too closely linked to the program. As consistent with the WWC standards, if the data are not available to account for the reliability and validity of a measure, and we cannot determine if the measure is standardized, we will exclude the outcome.

For teacher outcomes, retention or satisfaction/attitudes may be examined; however, teaching measures must also demonstrate sufficient reliability and validity as noted under the WWC (v2) standards.

STUDY DESIGNS

The review will include well-designed and well-implemented experimental and quasi-experimental designs. The review will not include single group pre-test/post-test; single-subject; or qualitative-only studies. For an experimental design to be included, it must place study participants (teachers, classrooms, or schools) into each study condition through a random assignment process or functionally random process (i.e., alternating by birth date, alternating by identification number). Quasi-experimental designs must demonstrate that the treatment and comparison groups are equivalent on observable characteristics (e.g., teacher retention, teacher satisfaction, student achievement). We will include studies that use statistical matching (e.g., propensity score matching or covariate matching) and regression adjustment (e.g., differences-in-differences). We expect that most studies will have the teacher or the classroom as the unit of analysis. For studies with the school as the unit of analysis, we will use two approaches. If there are at least five studies we will do separate analyses for school level studies. We will also adapt the variance estimates to the teacher level and analyze teacher level studies with and without the school level studies with modified variances.

The quality of each study (and its reporting) will be assessed according to several

characteristics, including: 1) the transparency of the study; that is, the clarity with which the investigators reported the assignment procedures; 2) the integrity of the assignment design and whether investigators addressed violations of the design; 3) the existence of high levels of attrition (particularly, differential attrition between treatment and control groups) from baseline samples to analysis samples; and 4) baseline equivalence for quasi-experimental designs and for experimental designs with high levels of attrition.

REFERENCES

- Amrein-Beardsley, A. (2012). Recruiting expert teachers into high-needs schools: Leadership, money, and colleagues. *Education Policy Analysis Archives*, 20(27), 1-23.
- Chamberlin, R., Wragg, T., Haynes, G., & Wragg, C. (2002). Performance-related pay and the teaching profession: A review of the literature. *Research Papers in Education*, 17(1), 31–49.
- Goldhaber, D. (2002). The mystery of good teaching. *Education Next*, 1, 50–55.
- Harris, D. N. (2011). *Value-added measures in education: What every educator needs to know*. Cambridge, MA: Harvard Education Press.
- Harvey-Beavis, O. (2003). Performance-related rewards for teachers: A literature review. Distributed at the Organisation for Economic Cooperation and Development (OECD) Attracting, Developing and Retaining Effective Teachers Conference in Athens, Greece. Retrieved 29 January 2013, from <http://www.oecd.org/dataoecd/17/47/34077553.pdf>
- Kelley, C. (1998). The Kentucky school-based performance award program: School-level effects. *Educational Policy*, 12, 305–24.
- Kelley, C. (2000). *Douglas County Colorado performance pay plan*. Madison, WI: Consortium for Policy Research in Education.
- Odden, A., & Kelley, C. (1997). *Paying teachers for what they know and do: New and smarter compensation strategies to improve schools*. Thousand Oaks, CA: Corwin Press, Inc.
- Plucker, J. A., Zapf, J. S., & McNabb, S. A. (2005). Rewarding teachers for students' performance: Improving teaching through alternative teacher compensation programs. *Center for Evaluation & Education Policy: Education Policy Brief*, 3(5), 1-7.
- Podgursky, M. J., & Springer, M. G. (2007). Teacher performance pay: A review. *Journal of Policy Analysis and Management*, 26(4), 909–950.

REVIEW AUTHORS

Lead review author: The lead author is the person who develops and co-ordinates the review team, discusses and assigns roles for individual members of the review team, liaises with the editorial base and takes responsibility for the on-going updates of the review.

Name:	Joshua Barnett
Title:	Assistant Professor
Affiliation:	Arizona State University
Address:	6422 West Irma Lane
City, State, Province or County:	Glendale, Arizona
Postal Code:	85308
Country:	United States of America
Phone:	(479) 387-8973
Email:	joshuahbarnett@gmail.com

Co-authors:

Name:	Gary Ritter
Title:	Professor, Endowed Chair Education Policy
Affiliation:	University of Arkansas
Address:	207 Graduate Education Building
City, State, Province or County:	Fayetteville, Arkansas
Postal Code:	72701
Country:	United States of America
Phone:	(479) 575-4971
Email:	garyr@uark.edu

Name:	Nathan Jensen
Title:	Research Specialist
Affiliation:	Northwest Evaluation Association
Address:	121 NW Everett St
City, State, Province or County:	Portland, OR
Postal Code:	97209
Country:	United States of America
Phone:	(503) 548-5091
Email:	nate.jensen@nwea.org

Name:	George Denny
Title:	Professor
Affiliation:	University of Arkansas
Address:	231 Graduate Education Building
City, State, Province or County:	Fayetteville, Arkansas
Postal Code:	72701
Country:	United States of America
Phone:	(479) 575-7320
Email:	gdenny@uark.edu

ROLES AND RESPONSIBILITIES

The four-person team proposing to conduct this evaluation has extensive content and methodological expertise.

- **Content:** Joshua, Gary, and Nathan have all worked for approximately a decade on various merit pay projects. Joshua Barnett has worked in Arkansas, Arizona, and internationally on merit pay issues from building programs to evaluating existing programs. Gary Ritter has worked on merit pay issues in Arkansas, and has also presented evidence on merit pay issues to the state legislatures of Arkansas, Oklahoma, Florida, and to the federal subcommittee on education. Gary has extensive experience designing and evaluating merit pay programs and is well connected to the other leading researchers on this topic. Nathan Jensen has also worked extensively building and evaluating merit pay programs in three separate school districts in Arkansas, including urban, rural, and charter districts, and helped develop a merit pay program for a charter school in Colorado.
- **Systematic review methods:** Joshua, Gary, Nathan, and George have all worked on at least one systematic review of the literature. Joshua and Nathan conducted systematic reviews on merit pay for their respective dissertation projects (2007, 2012 respectively). Joshua, Gary, and George also previously worked on a Campbell Collaboration review utilizing systematic review methods. Additionally, Gary has taught courses directly aimed at conducting systematic reviews.
- **Statistical analysis:** George Denny has worked for two decades as a methodological expert on research projects and will primarily be involved with the analysis section of the review. George has worked previously on a Campbell Collaboration project, where he served as the statistical analyst.
- **Information retrieval:** Joshua and Gary have worked previously on a Campbell Collaboration review on volunteer tutoring and are familiar with the process and requirements of retrieving information.

POTENTIAL CONFLICTS OF INTEREST

The authors have conducted primary research and published work on this topic area. However, the reviewers maintain that any potential conflicts of interest related to our work in this area will be counter-balanced by the explicit and transparent methods used to conduct the systematic review.

SUPPORT

As we have previously conducted a review, we do not anticipate needing support in methodology and causal inference, systematic searches, coding, or statistical analysis (meta-analysis); however, we do anticipate working with a university librarian for assistance in conducting our review of relevant research. Additionally, when we conducted our previous review, we did have regular contact with the Campbell Collaboration education liaison to ensure that our work was being performed in accordance with the latest standards; we anticipate having such contact again. We would like to be aware of the key contacts with whom we might correspond for each of these steps, and anticipate working closely with the education liaison to ensure our work is completed correctly.

FUNDING

We have received an Education Coordinating Group grant of \$10,500 to help assist with this project.

PRELIMINARY TIMEFRAME

- Date you plan to submit a draft protocol: 15 March 2013
- Date you plan to submit a draft review: 1 September 2013

Date Completed	Milestones/Task
5 December 2012	<ul style="list-style-type: none">• Register title with Campbell Collaboration – initial submission
1 February 2013	<ul style="list-style-type: none">• Submit final title registration, pending revisions from ECG
15 March 2013	<ul style="list-style-type: none">• Submit initial draft of protocol for systematic review, including inclusion and exclusion criteria, search terms and search engines, and decisions on unpublished literature.
15 May 2013	<ul style="list-style-type: none">• Submit revised draft of protocol for systematic review, pending revisions from peer reviewers
15 June 2013	<ul style="list-style-type: none">• Submit final protocol for systematic review
1 September 2013	<ul style="list-style-type: none">• Submit initial draft of review
1 December 2013	<ul style="list-style-type: none">• Submit revised draft of review pending revisions from peer reviewers
1 February 2014	<ul style="list-style-type: none">• Submit final review

DECLARATION

Authors' responsibilities

By completing this form, you accept responsibility for preparing, maintaining, and updating the review in accordance with Campbell Collaboration policy. The Coordinating Group will provide as much support as possible to assist with the preparation of the review.

A draft protocol must be submitted to the Coordinating Group within six months of title acceptance. If drafts are not submitted before the agreed deadlines, or if we are unable to contact you for an extended period, the Coordinating Group has the right to de-register the title or transfer the title to alternative authors. The Coordinating Group also has the right to de-register or transfer the title if it does not meet the standards of the Coordinating Group and/or the Campbell Collaboration.

You accept responsibility for maintaining the review in light of new evidence, comments and criticisms, and other developments, and updating the review at least once every three years, or, if requested, transferring responsibility for maintaining the review to others as agreed with the Coordinating Group.

Publication in the Campbell Library

The support of the Coordinating Group in preparing your review is conditional upon your agreement to publish the protocol, finished review and subsequent updates in the Campbell Library. Concurrent publication in other journals is encouraged. However, a Campbell systematic review should be published either before, or at the same time as, its publication in other journals. Authors should not publish Campbell reviews in journals before they are ready for publication in the Campbell Library. Authors should remember to include the following statement in reviews that are published elsewhere: "This is a version of a Campbell review, which is available in The Campbell Library".

I understand the commitment required to undertake a Campbell review, and agree to publish in the Campbell Library. Signed on behalf of the authors:

Form completed by: Joshua Barnett

Date: 5 December 2012

Joshua H. Barnett
