Main title	What is the evidence of the impact of increasing salaries on improving the performance of public servants, including teachers, nurses and judges?
Sub title	None
Review group	DFID/Self-organized
Section	PROCOTOL
Authors IN ORDER OF CREDIT (Please include first and surnames, institutions. Include titles – Dr, Prof – if you want them to be used.)	Professor Stuart C. Carr, Massey University, New Zealand
	Professor Malcolm MacLachlan, Trinity College, Dublin, Ireland
	Professor Michael Clarke, UK Cochrane Centre and University of Oxford, UK
	Professor T. S. Papola, Dr. Jesim Pais, Institute for Studies in Industrial Development, New Delhi, India
	Professor Charles Normand, Trinity College, Dublin, Ireland
	Dr. Steve Thomas, Trinity College, Dublin, Ireland
	Professor Eilish McAuliffe, Trinity College, Dublin, Ireland
	Dr. Chez Leggatt-Cook, Massey University, New Zealand
EPPI-Centre reference number	[To be completed by EPPI-Centre]
Month/year of publication	[To be completed by EPPI-Centre]
This report should be cited as…	Carr, S. C., MacLachlan, M., Clarke, M., Papola, T. S., Normand, C., Thomas, S., McAuliffe, E., & Leggatt-Cook, C. (2010). *What is the evidence of the impact of increasing aid salaries on improving the performance of public servants, including teachers, nurses and judges?* Massey University: New Zealand.
Contact details (address, phone number, email)	Professor Stuart C. Carr, *Poverty Research Group*, Massey University, Private Bag 102 904, North Shore MSC, NZ. +64 9 414 0800 x41228. s.c.carr@massey.ac.nz
Institutional base	Massey University, New Zealand.
Review Group (with institutions)	By arrangement with the UK Cochrane Centre
Advisory group (with institutions)	Peter Evans, DFID Senior Governance Advisor, India (Policy Lead), with additional Country Advisors during the life of the project.
Conflicts of interest (if any)	None
Acknowledgements	We wish to acknowledge the input of Dr. Chez Leggatt-Cook to this protocol. Dr. Leggatt-Cook is our Postdoctoral Assistant for the project. We are very grateful to our eight peer-reviewers, who have consistently provided a wealth of useful, constructive, and innovative feedback for this protocol document. This feedback we hope will significantly improve this particular systematic review, both process and outcome. It has sometimes been included in this protocol verbatim, or in adapted form.

1. Background

1.1. Aims and rationale for review

This applied research and policy question sits at an intersection between multiple literatures, not only in health, education and correctional services, but also Development Economics, Human Resource Management, Organisational Psychology and Sociology. Inter-disciplinary issues like this require a fresh, integrated approach to theory, method, analysis, and synthesis (MacLachlan et al, 2008). The topic is also imbued with myth and taboo, necessitating an integrated evidence-based review to disentangle fact from fiction (Whitehead et al, 2001).

The project is set against a general policy background of the Millennium Development Goals; in particular their focus on decent work and the importance of civil service sectors in lowering human poverty by 2015 (United Nations, 2010). In lower-income settings where "Need" is greatest (Figure 1), basic salary increases may logically be a salient means of reducing poverty, not only directly through salary itself but also indirectly by improving the delivery of civil services. Yet this important possibility has not to our knowledge been systematically tested using standardised procedures and analytical techniques nowadays available to guide evidence-based policy and practice (Oliver et al, 2005).

This project combines interdisciplinary review approaches to explore the evidence for the effects of salary increases on improving the performance of public servants. Central to the MDGs, these include (i) teachers from education, and (ii) nurses from health. We also include (iii) judges from correctional services, for two major reasons: (1) their positions are extraordinarily powerful, influential for governance and dependent on trust; and yet (2) they are less –researched than other sectors.

1.2. Definitional and Conceptual issues

Here we discuss the relationships between (A) pay rises/cuts (variation), (B) motivation and (C) performance. Theoretical linkages are summarised in Figure 1.
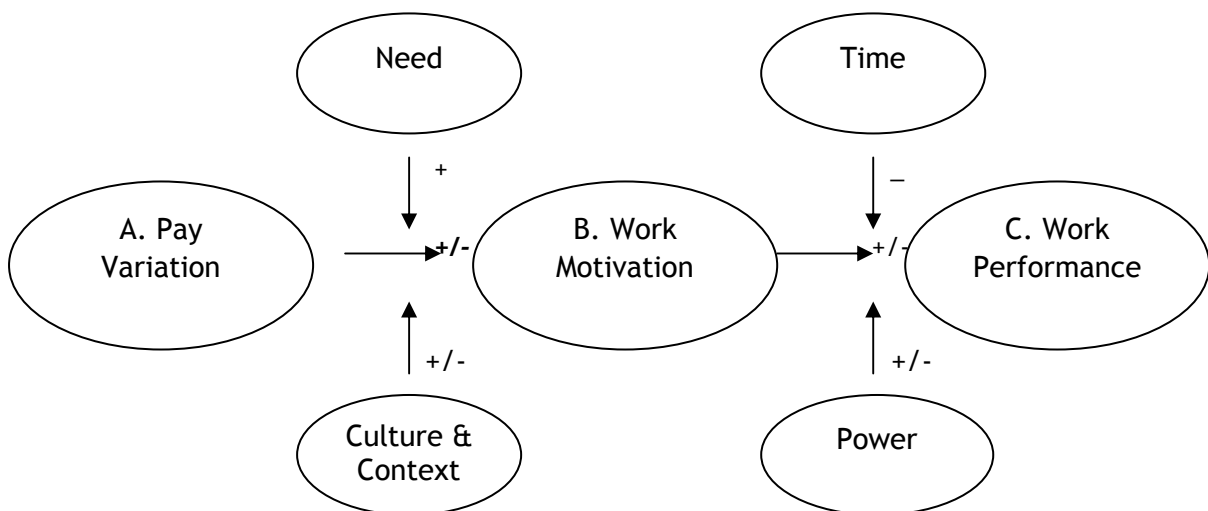

Figure 1 – A Model of the Links between Salary Change Context and Performance

A. In figure 1, pay changes are antecedents in a causal model, ranging from fixed salaries (and increments) to performance-based systems. The latter are contingent on individual performance on the job, and usually operate over-and-above a more fundamental concern to most civil service workers – base (or fixed) salary (Furnham, 2005; OECD, 2005). "Performance-based" pay is not the norm in many public sectors, for example in India which has 65% of formal employment in the public sector, so they are not the focus in this review. "Salaries" are fixed and group/job-based, *not* variable or directly 'contingent 'on changes in individual or organisational "performance." Hence there is no obvious "tournament" to be played, or utility for an individual worker to raise their performance (Lazear & Rosen, 1981). Rather, individual and aggregated individual (organisational) performance change is quite likely (though not necessarily) *discretionary*, i.e.

B. An intervening variable, broadly classified in Figure 1 as "Work Motivation," theoretically mediates between A. salary increase and corresponding increases in

C. Work Performance (Latham, 2007), which in Figure 1 includes both individual and (aggregate) organizational output.

In higher-income settings, financial incentives have been linked to increases in affective and cognitive motivation, e.g., respectively, enhanced mood and work interest (Eisenberger, Rhoades & Cameron, 1999). Pay is the most obvious and widely recognised basic means of motivating people both "to" work, and while "at" work (Paulsen, 2008). The ambit includes health and other services in the public sector (Das, Hammer & Leonard, 2008; Wright, 2001). Theories in support of pay's motivating properties include classical utility/behaviourist models, that stress the "instrumentality" aspect of rewards like money and wages (e.g., Latham & Pinder, 2005) Principal-Agent theory, which stresses how salaries can help align the motivation of a civil servant "Agent" to the civilian "Principal" (Laffont & Martimort, 2001) and Efficiency Wages, which predict that paying one group more than the market rate will for example reduce "moral hazard" (Akerlof & Yellen, 1986); to complementary theories in sociological social science, for example Social Equity, which adds workers' comparisons with the pay of others, and negative saliency for a "perceived injustice" (Bloom, 2007; MacLachlan et al, 2010; Marai et al, 2010; Munthali et al, 2010; Zhou et al, 2010).

In meta-reviews of utility, pure reinforcement models have yielded positive effect sizes on the criterion, performance, in the region of 0.51-0.95 (Stajkovic & Luthans, 1997). With corrections for sample size, measurement error and restriction of range, pay incentives have been found to have a meta-effect size on performance = 0.34 (Jenkins, Mitra, Gyupta & Shaw, 1998). Explaining 11.56% (r^2-statistic)) of variability in performance, such an effect size would normally be considered "small," although there is no accounting, often, for effects of error that have not been statistically controlled (Doucouliagos, 2010), for example through structural equation modelling. In the Jenkins et al (ibid) meta-review, and in Grindle (1997) there was no effect for task type/occupation. Additional effects for work justice, including equity, have been more incremental but still statistically significant, ranging from 0.13-0.16 (Cohen-Charash & Spector, 2001).

To our knowledge, the above research has not been experimental or inclusive of multi-level structural equation modelling or multi-level mediated models, leaving us reluctant to draw any firm inferences about either direction or nature of causality. Statistical tests of mediation (Figure 1 is a mediation model) do not establish causation unless there is a longitudinal design ("Time" in Figure 1), with

appropriate control groups (Stone-Romero & Rosopa, 2008). For example, student talent might help salary-raises rather than vice-versa, whilst a third factor, e.g., injections of educational resources, could be causing each (Pugh, Mangan, & Gray, 2010). In this project therefore, we will be looking in particular for research designs that are preferably either (i) experimental (unlikely) and (ii) longitudinal (i.e., that include Time in Figure 1).

Casting fundamental (theory-based) doubt on a simple 1:1 "Elasticity" (Doucouliagos, 2010) between pay on the one hand and performance on the other, counter-intuitively perhaps some theories actually predict that pay increases may backfire, particularly perhaps on work quality (OECD, 2005). Some of the contrary theories include for example Cognitive Evaluation models, which predict that emphasising extrinsic rewards at best appeases a sense of entitlement/natural progression, or at worst undermines intrinsic motivation and organisational citizenship behaviour (Ambrose & Kulik, 1999). Relative Deprivation Theory suggests that pay rises can create unrealistic expectations that they will keep on rising, and because they cannot, lead to frustration and deflate performance (Zhou et al, 2010). In a similar vein perhaps, extinction and habituation theories in learning, and the Backwardly Bending Labour Supply Curve in economics, each suggest that any effects of salary rises will be short-lived, attenuating and even reversing in the course of organizational developmental time (Zhou et al, 2010). Hence, Time in Figure 1 is a negative factor (-), as time continues, gains will fade.

Recently, theories of motivation and performance have stressed the crucial implications (for workplace performance) of having structured goals, career trajectories, and meaningful feedback in performance management systems (Das et al, 2008; Raabe, Frese & Beehr, 2007). These help to build worker confidence, commitment and efficacy (Stajkovic & Luthans, 1998). With effect sizes in the region of 0.42-0.80 (Locke & Latham, 2002), available research on structured goal-setting processes (mostly from high-income, OECD countries) suggests that salary in itself may be a necessary (but not sufficient) condition to enhance workplace performance (OECD, 2005). Indeed, in some cases, relying solely on pecuniary incentives can undermine intrinsic motivation and spur counter-productive work behaviour (Kluger & DeNisi, 1996; Kohn, 1995).

Adding to the issue of 'whether' increasing pay improves work performance or not, is the possibility that any linkage is significantly moderated, principally by contextual factors that are socio-cultural and socio-political, as well as socio-economic. Motivational factors are likely to have different impacts in different contexts (Fischer & Smith, 2003). Low-income countries present from Figure 1 a "Culture & Context" that is not only particularly challenging for performance but also quite distinct socio-culturally and -politically from contexts where motivational theories were originally developed - and have mainly been tested (Bolino & Turnley, 2008). Even if pay improves motivation and a worker wants to perform better, s/he cannot because of a basic lack of resources, equipment, etc (Easterly, 2006). This review will attempt to code for Macro-economic considerations such as the strength of the labour market. Economic crisis generally could have important moderating effects including on the outcomes of particular studies. 'Time since pay-change' is also a key factor (Figure 1). Such variables are very important not only as moderators per se, but also because of their potential to broaden the range of salary changes on both motivation and performance, specifically to include salary *cuts* (Chew, 1990).

Whilst this review will focus on pay increases, it will also extend to pay *variation*, to include 'negative' values on the antecedent ("Variation" in Figure 1). Figure 1 also includes the moderators "culture," because the motivational consequences of a salary change depends in part on values like materialism, individualism, and collectivism; and "power," because the *expression* of individual and group motives is part controlled by political considerations like status, seniority, caste, and social dominance (MacLachlan et al, 2010). Moderators like this are multi-faceted and complex, but they are also vital to recognise, and represent, when achieving a useful and reasonably robust evidence base.

For instance, many settings will be relatively high on collectivistic rather than individualistic work norms, and hierarchical rather than egalitarian beliefs about the distribution of status and income (Erez, 2000). Cultural and political norms like that carry the potential to interact with the impact of salary changes on motivation and performance, at individual and organisational as well as societal levels (Roe et al, 2000). Socio-cultural contingencies include for example diverse norms and values about procedural and distributive justice, including valuing need over equity rather than vice versa (Greenberg, 2008). Thus from Figure 1, an incentives strategy that is seen as appropriate and motivating in one country (+) may not be seen in the same way in another country (-), or even among ethnic groups within the same country (Ambrose & Kulik, 1999).

To summarise, what the above analysis and our provisional working model in Figure 1 warn is that the question for this review is both complex and open, with multiple types of intervention and possible moderators of motivational consequences, from job (dis)satisfaction and (dis)engagement to perceived (in)justice and lowered or raised self-determination (Eisenberger, Rhoades, & Cameron, 1999). Yet under such conditions, finding an overall level of effect for salary increases will precisely provide valuable "baseline information against which to asses the utility of specific financial treatments in work settings" (Jenkins et al, 1998, p. 777).

Crucially, complexity and openness apply equally to how we think about "performance" (Colquitt et al, 2001). Basically it is multi-dimensional. A major development in the literature is meta-factor analytic work on types of workplace performance, resulting for instance in an eight-dimensional ("Great Eight") model of workplace performance (Bartram, 2005). Today there is growing recognition that professional work *performance*, like its antecedents in motivation and motivators, is invariably multi-faceted (Meyers & Houssemand, 2006). Although the newer innovatory understandings may not yet be fully reflected in the literature from lower-income settings, this project remains open to a wide range of definitions of "performance," whilst staying focused on the overall question: do salary increases generally improve performance, whatever its particular facet?

Last but not least, and beyond the immediate question for this review, we will also consider the wider context of "consequential validity" (longer-term social implications of an intervention when evaluating it) in the evidence reviewed), for example the possibility that performance-based pay may have negative effects on worker job security (Ihle-Helledy, Zytowski, & Fouada, 2004).

1.3. Policy and Practice background

Historically, salaries of civil servants in lower-income countries fell during the structural reforms of aid and government systems during the 1980s and 1990s (Chew, 1990; Terrell, 1993). During that period, aid and development focused on

"governance" (Hjertholm & White, 2000). This is a domain in which the behaviour of civil servants at all organisational levels continues to feature prominently and centrally (Easterly, 2006).

To give one example, gaps between civil service and private sector-incomes in lower-income countries, combined with economic necessity could in theory motivate a range of potentially "counter-productive workplace behaviours" (Latham, 2007).   Examples identified by a World Bank research group include "moonlighting", de-motivation and "predatory corruption", to more positive forms of workplace motivation and performance (Van Lerberghe et al, 2002; although these authors are critical of such pejorative terms, preferring to see them as partly motivated at least by a wish for restorative work justice). Civil service salaries are certainly linked, via Utility theory in the literature, to international mobility, and "brain drain" (Brown & Connell, 2004). These are secondary forms of reduced 'performance' with which the literature is already imbued.

Today therefore, there is widespread policy interest in the evidential base concerning the re-structuring of basic civil service salaries and related incentives, its potential to motivate civil service work performance, thereby enhancing poverty reduction initiatives. At the present time, other forms of incentive such as performance-based pay are regarded as peripheral, not only by policy-makers but also by many civil servants themselves (OECD, 2005). Hence they are not included in the ambit of the present review.

1.4. Research background

There appears to be plentiful evidence in the research literature that poor pay has a negative effect on the performance of public servants (Dieleman et al, 2003 and 2006; Kingma, 2001 & 2007), and is a core factor in worker de-motivation (Willis-Shattuck et al., 2008). This is especially the case when salaries are not adequate for obtaining the basic necessities of daily life (Agyepong et al, 2004), i.e. "need" in Figure 1. The opportunity to earn higher salaries is one of several major push factors for the migration of health care workers (Vujicic et al, 2004).

However, a narrow focus on financial incentives may not be adequate for improving performance and motivation. For example, Vujicic et al (2004) have shown that wage increases in source countries are not likely to decrease the migration of health care workers. This is partly because wage differentials between source and *destination* countries are so large that small increases in wages *locally* are not likely to have much effect. Vujicic et al suggest that non-wage instruments might have more effectiveness with respect to slowing down migration and brain drain, although this does not necessarily extend to in-country local civil service.

One particular challenge for this review, identified above, is that while financial incentives are very important for the performance (especially motivation and satisfaction) of public-service workers, over-reliance on financial incentives can create problems. Crewson (1997) argues that understandings of human motivation have rested on neo-classical notions of the importance of external, economic rewards. This fails to account for a moral dimension (linked to intrinsic rewards such as feelings of pride; appreciation from the community) that could be particularly prominent in public servants (e.g., Macq et al, 2001; Van Lerberghe et al, 2002). Public servants need to be adequately compensated for their work, but increased financial rewards can lead them to start viewing these as more important, creating a conflict between their values and the messages they receive

about working for financial gain (Crewson, 1997; Franco et al., 2002). Furthermore, worker motivation is highly complex, and a number of non-financial factors have been found to increase worker motivation, particularly recognition, responsibility and training (Bradley & McAuliffe, 2009; Chomitz et al, 1998; Crewson, 1997; Dieleman et al, 2003; Dieleman et al, 2006; Franco et al, 2002; Franco et al, 2004).

Perhaps this means that increasing pay should not be attempted in isolation from the improvement of acknowledgement and recognition for work, management practices, job characteristics/working conditions (including adequate resources and infrastructure), the provision of job and career training, and opportunities for career advancement. In other words, "bundles" of financial and non-financial interventions may ultimately be more effective for attracting and retaining staff in remote rural areas in low- and middle-income countries (Lehmann et al, 2008).

Ultimately however, in an environment where resources are scarce it is important that they are deployed to maximum effect. Outside of higher-income/OECD countries like the United States and the United Kingdom, as we have noted above (Figure 1), the salience of basic material rewards may take priority over arguable 'luxuries' such as "work-life balance," "career plans" (OECD, 2005), and even "job satisfaction" (Carr et al, 2010). Increasing the pay of public servants is one of the most costly ways of attempting to improve performance. Hence it has become critical for governments to have evidence on the potential effects of such increases on performance.

1.5. Objectives

1.5.1. Primary objective

This project seeks to systematically review the available evidence for the impact of increasing salaries on improving the performance of selected public servants. These comprise Public sector employees in low- and middle-income countries.

1.5.2. Secondary objectives

The review will assess the effectiveness of different types of salary change. These span (i) pecuniary, (ii) non-pecuniary changes, and (iii) bundles of financial with non-financial interventions (as discussed above).

Following discussion with our Policy Lead (above), the review will focus on health workers (nurses, doctors and alternative cadres), education (teachers) and the judiciary (judges).

The review will also assess any moderating impact of contextual factors on salary change (such as cultural norms, sector norms, socio-economic status, and reference groups).

The preamble for the review will also need to include references to the impact of civil service salary reforms in higher-income countries (OECD, 2005). This will enable us to highlight any differences in relative effects between advantaged versus less-advantaged groups.

2. Methods used in the review

2.2 User involvement

2.1.1 Approach and rationale

This is a Rapid Evidence Assessment project, with little time to consult widely with end-users and key stakeholders before the data-gathering phase. Our approach has been to carefully select a cross-disciplinary team, with extensive contacts and everyday practice experience in the domain, for instance the Institute for studies in Industrial Development, India. In addition, our Policy Lead on the project from DFID India, Dr. Peter Evans, is a Senior Governance Advisor within the Governance Group, and former Equity Advisor; with extensive experience of the topic and its potential to inform development and organisational policy and practice.

In India, the state (public sector) continues to be the largest employer providing formal employment accounting for well over 65 percent of formal employment. And within this, the employees in general administration (revenue, police, tax, agriculture .....) is substantially large. Thus besides reviewing salary increases and their effects on public sector workers in health (nurses), education (teachers) and judges, in the India case, our team members there will give some importance to employees in 'general administration' or the general 'civil servant'. In other words, state employees providing civil services continue to form a substantial part of organised employment in India. Any change in the conditions of work of these employees is likely to have an impact on the provision of these services.

2.2 Identifying and describing studies

2.2.1 Defining relevant studies: inclusion and exclusion criteria

Types of study: We will seek empirical research that used qualitative or quantitative methodologies to assess the effects of a change in salary. Our aim is not to Meta-analyse studies comparing salary packages against each other – although we will need to review some of these to set the broader theoretical context for the data we draw. A comparative eye on the different occupations and sectors, from teachers in education to judges in corrections, will perform a similar (comparative) function.

Our estimate of the size and quality of the evidence base is that it is largely (a) multi- rather than inter-disciplinary, (b) single- rather than multi-level, and (c) methodologically (and theoretically) diverse, ranging from qualitative case studies and ethnographic approaches to quantitative country-level surveys and regressions. Furthermore, there is (d) insufficient attention to the inherently nested nature of the data, principally within sectors, organisations, and countries. Overall, therefore, the proposed evaluation needs to be multi-method rather than simply meta-analytic, exploratory rather than confirmatory, and inductive insofar as choosing a specific method, or methods, of data analysis. We have to remain open to the possibility of finding incompatible research paradigms, and in that event using a reliable content-analysis of distinctive and divergent meta-themes (Greenhalgh et al, 2005).

Types of study to be categorised range from case studies to within-country, multiple regression survey designs. These will each be longitudinal, with a premeasured of performance prior to salary rise, and preferably include evidence of sustainable rather than shorter-term changes. Specifically, we are interested in designs that are experimental (randomised control trials) and quasi-experimental (premeasured performance - pay rise – post measure performance), in organisational surveys (with hard measures of performance, not soft self-reports which are subject to smile factors and Hawthorne effects), and single and multiple case studies (where demonstrably sound measures of changes in both pay and performance have been taken). We will also check that in any quantitative study above, adequate statistical assumptions regarding statistical power, independence, distribution etc, are adequately met. At present, we expect to find very little published that meets these criteria in health, however they may exist in the other sectors. If this data is not available, the study will not be included. For qualitative studies, we will be focusing on evidence of reliability between coders, and on the possibility of Hawthorne-type effects.

Types of participant: Public sector employees in low- and middle-income countries.   These will comprise health (nurses, including midwives who perform a key role in MDGs 4 and 5 that focus on Maternal and child health, doctors and alternative cadres), education (teachers) and judiciary (judges).

Types of intervention: Interventions involving upward or downward changes in fixed salary (i.e. a wage that is paid on a regular basis, in a fixed amount).

Studies will be excluded if:
1. They do not include salary increases as the sole intervention, or as part of a remuneration package.
2. They explore the effects of particular remuneration packages rather than the effects of *changes* in remuneration package.
3. The salary increase is introduced for reasons other than improving performance (for instance to restore equity between genders in line with human rights legislation).
4. Study designs do not include at least one repeated measure (before and after pay variation), or a control group with no salary increase against which the treatment group can be compared cross-sectionally or in a lagged-control design.
5. Study designs do not control for nesting and non-independence of observations across levels (e.g., country, sector, organisation, individual). This may eliminate many studies from the large initial database.
6. Studies do not report or conduct appropriate tests for common method variance (where applicable).
7. The study does not include a hard (objective) measure of output, such as individual performance appraisal data or archival records of actual output by an organization or department within it.

Types of outcome measure: Measures of performance will include quantity (for instance, absenteeism) and/or quality (user satisfaction) of work; primary (work performance) and secondary (turnover). Outcome measures will not form part of the search or inclusion criteria for the review. However we expect the two main outcome measures to derive from (1) individual performance appraisals and (2) indicators of organisational productivity or efficiency, such as Value for Money (VfM). Each type of outcome can be used in private and public sectors, although (2) may be modified to include evaluative data from consumers of the civil services

in question (for students grade point averages or exam results and other forms of educational attainment; by patient satisfaction surveys and 'harder' indexes like actual health outcomes; and by externally-derived governance/corruption indicators for the judiciary, including any available indicators of public trust/confidence). In general, the project will use inductive rather than deductive methods to help ensure that due consideration is given to the relevant of the outcome measures across the alternative occupations that are being reviewed. Following Jenkins et al (1998), the outcome measures will be presented in a summary descriptive table.

Examples of included studies:

(1) Muralidharan, K., & Sundararaman, V. (2009). *Teacher performance pay: Experimental evidence from India*. Program of Education Policy and Governance Working Paper Series/NBER Working Paper No. 15323.
(2) Sabry, M. A. El-Fattah. (2010). Longitudinal effects of pay increase on teachers' job satisfaction: A motivational perspective. *The Journal of International Social Research, 3*, 1-21.
(3) Vujicic, M. (2009). *How you pay health workers matters: A primer on health worker remuneration methods*. Technical Brief. Washington, DC: The World Bank (useful also for forward and backward search snowballing).

Examples of excluded studies:

(1) Anyangwe, S. C. E., & Mtonga, C. (2007). Inequities in global health workforces: The greatest impediment to health in sub Saharan Africa. *International Journal of Environmental Research and Public Health, 4*, 93-100.
(2) Buscaglia, E. (1999). *Judicial corruption in developing countries: Its causes and economic consequences.* Berkely Program in Law and Economics, Working Paper Series.
(3) Chaudury, W., Hammer, J., Kremer, M. Muralidheran, K., & Rogers, F. H. (2006). Missing in action: Teacher and health worker absence in developing countries. *Journal of Economic Perspectives, 20*, 91-116.

2.2.2 Identification of potential studies: Search strategy

At the moment, we have no plans to apply publication year limits to our searches. However, as discussed below, this might be necessary for feasibility and efficiency. If we do apply time limits, our focus will be on the more recent literature. Our preliminary work, which informed the discussion provided above, has identified that although the question might appear to be narrowly focused on the surface, the literature itself is both complex and open. In systematic reviews of this nature, it is wise to rely on a combination of both formal (protocol-driven) and informal search strategies, such as "snowballing," i.e., backward reference-tracking and forward citation-tracking (Greenhalgh & Peacock, 2005). The starting point for the search however will be a structured search which will be undertaken by the UK Cochrane Centre.

Search terms: We will use the following elements to conduct initial scans of abstracts and full-text articles in Business Source Complete, PsycINFO etc. A search plan will be finalised in consultation with our team information specialist at the UK Cochrane Centre, and will be adapted for other databases:

Antecedent terms	Mediator terms (Post hoc coding only)	Performance terms
Pay* OR	Motivation	Task performance
Remuneration OR	Satisfaction	Work performance
Salar*	Commitment	Performance
Benefits OR	Efficacy	
Incentive* OR	Engagement	
Financial OR	Citizenship	
Money OR	Initiative	
Monetary OR		
Reward* OR		
Wage* OR		

AND

Change* OR
Increase* OR
Rise* OR
Augmentation* OR
Growth*

AND

Low* income
Middle income
Developing countr*
Developing nation*
Third World

AND

Teacher OR
Doctor OR
Nurse OR
Cadre OR
Judge

Related Thesaurus terms, e.g., "least developed countr*"

_____

Meta-search term: Antecedents AND Performance terms, all possible combinations
_____

The volume of literature will be identified initially from the following digital sources:

Bibliographic databases:

Our preliminary scoping work has identified an extensive list of databases which might include material of relevance to this systematic review. However, within the time and resources available for this project, it is unlikely that it will be possible to search all of them. We will use this base to work with the information specialist at the UK Cochrane Centre to focus the search the grounds of feasibility and efficiency, which may include the use of publication year limits for some of the searches.

      A+ Education
      Academic OneFile
      Academic Search Elite
      Annual Reviews
      APAIS
      AnthroSource
      Anthropological Index Online
      Berkeley Electronic Resource Journals
      BMJ Journals Online
      Business Source Complete
      Business Source Premier
      Campbell Register
      CINAHL
      Cochrane Library
      Current Contents Connect
      DFID Res4Dev
      EBSCO Open Access French Collection (PI is French-Speaking)
      EBSCO Open Access Journals
      EBSCO Open Access Portugese Collection
      EBSCO Open Access Spanish Collection
      EBCO host Platform
      Econlit
      Full-text EconLit
      EdResearch Online
      EMBASE
      ERIC
      Health Business Elite
      Health Databases on EBSCOhost
      Health Source: Nursing/Academic Edition
      IBSS (International Bibliography of the Social Sciences)
      ILO Archive
      ISID online reference <http://58.68.105.146/login3.asp
      [This bibiographic database is similar to the econlit having an Index of journal articles appearing in 125 Social Science journals published in India

(basically Indian journals). This may be a rich source of information on India. Unfortunately, the abstracts of the articles (most of which has to be hand-searched) are not available with this facility].

      ISI Proceedings
      Journal Citation Reports
      Journals@Ovid
      JSTOR
      MasterFILE Premier
      MEDLINE
      Mental Measurements Yearbook (to check any questionable measures)
      PapersFirst
      Pre-MEDLINE
      ProceedingsFirst
      Professional Development Collection
      Project MUSE
      ProQuest (theses)
      PsychBooks
      PsycINFO
      PubMed Central Open Access
      Sage Reference Online
      Science Citation Index
      Social Science Citation Index
      Social Services Abstracts
      Sociological Abstracts
      Source OECD
      SpringerLINK
      Teacher Reference Center
      Wiley InterScience Platform
      Web of Knowledge
      Web of Science
      World Bank Archive

The above list is exceedingly long and extensive. Based on preliminary screening work utilising the above long list of databases, their accessibility to the people who will do the searching and the need to keep within the limited resources available for the review, we have decided to limit the searching to the following databases:

Annual reviews
Business Source Complete
C2-SPECTR (the study register of The Campbell Collaboration)
CINAHL
The Cochrane Library (including NHS EED)
Econlit
EMBASE
ERIC
IBSS
JSTOR
MEDLINE
Project MUSE
ProQuest
PsycINFO
Science Citation Index
Social Science Citation Index

This is still a long list and it is likely to yield approximately 10,000 records to check, but we think it will be manageable. Furthermore, we have found several additional sources which might also be worth searching:

LILACS (Latin American and Caribbean Health Sciences Literature)
INDMed (indmed.nic.in)
http://repec/org/

Iff time allows:
http://www.revues.org/
http://www/persee.fr/
http://www.cairn.info/
A database system will be set up to keep track of and extract data from the studies found during the review, as well as keeping a structured log of studies that are excluded. We have opted to use the new system developed by the EPPI-Centre reviewer software for information management. This has generously been funded by DFID following our recent workshop meeting in London.

2.2.3 Screening studies: applying inclusion and exclusion criteria

The UK Cochrane Centre's information specialist will conduct an initial search across the appropriate databases, using the search tactics chosen above. At this stage, for speed, some of the screening of records should be done just using the titles (and not the abstracts even if these are available - providing they can be distracting). This runs the risk of missing some potentially relevant articles but the trade off is that it will allow us to look at many more records in the first place and we are expecting that the titles will be a very strong indicator of a major study of relevance to the review. We might also want to reduce some of the screening to one person. For example, for some of the online databases it might be easier to look at things on-screen and online, rather than downloading records. One reviewer is willing to do this for some of the databases where there might not be many hits, but we will have two people working independently on the bulk of the material.

Two reviewers (SCC and CLC) will separately check the abstracts and/or titles of the citations identified by the search to determine whether each paper meets the pre-determined criteria. We will use the appraisal tool (minus coding section, at this stage) in Appendix 2.4. In case of doubt or disagreement, the full article will be obtained for inspection.   The full text of each potentially relevant study will be obtained and will be independently assessed to determine whether it meets the inclusion criteria. In the event of a disagreement, a third reviewer (from the team, with expertise in the particular domain) will be asked to give her or his opinion to resolve the issue. Identified studies will be tracked down and entered into the review-specific database, using the pre-purchased EPPI-Centre software.

2.2.4 Characterising included studies

The chief limitations likely to be faced in this systematic review are the relative paucity of high-quality, inter-disciplinary research; soft measures of performance; and of well-controlled experimental studies for determining causation rather than correlation to help balance generalisation with respect for particularity.

2.2.5 Identifying and describing studies: quality assurance process

Throughout the appraisal and coding process two team members (SCC and CLC) will independently assess the methodological quality of, and variables in, the eligible studies. In the event of a disagreement, a third reviewer (from an appropriate section of the research team, given the methodology and focus of the particular study) will be consulted.

The basic tool we will use to structure the individual scans of quality in each paper is found in Public Health Resource Unit (2006). Anchored in human services, this document gives separate structured tools for use with quantitative studies (randomised control trials), qualitative research, and prior reviews in the field. Hence it is well-suited to the scope of this project.

2.3 Methods for synthesis

2.3.1 Assessing quality of studies

In addition to the CASP tools above, rigour of the research methods used in studies will be evaluated using the appropriate, established procedures and indices from organisations such as The Cochrane Centre. We use the protocol in Appendix 2.4. This is designed to address specific statistical challenges to research in this field. They include an inherent nesting of data within levels (countries, organisations, sectors, departments). Nesting is a multi-level design factor – and as such a contextual variable that technically inflates the risks of Type I error. Treating the higher-levels as dummy variables does not allow for the quantification of variation among the higher-level units (Stride, 2008). Studies that have not adequately controlled for nesting, including non-independence of observations, will be removed from the review. In some cases, the precision of estimates can be addressed by appropriate modifications of the corresponding test statistics (Scariano & Davenport, 1987). If studies display awareness of the issues and take steps to adjust prevision measures accordingly, then we will deem them suitable for inclusion.

2.3.2 Overall approach to and process of synthesis

Initially at least, we will construct a descriptive summary table, resembling a matrix (Jenkins et al, 1998). This will have author studies down the vertical, grouped firstly into *levels of analysis* (individual, organisational, country, followed by the various cross-level combinations analysed) and secondarily, by chronological order. Across the horizontal are included type of performance measure (categorised into objective work quantity, work quality, absenteeism/"not-on-seat'/"temporary out," turnover, etc), country site, and other key potential moderators identified from the theories discussed already, including occupation (teachers, health workers, judges) followed by study *N* and key quantitative (statistical) or qualitative (textual) indicators reported. Each categorisation undertaken will have a reliability coefficient for the coding process on which it is based, using Cohen's Kappa (Robson, 2002, p. 222)[1].

With respect to quantitative studies, if our sample size in this category allows, our statistical approach to heterogeneity and pooling of data is to make appropriate

---

[1] At this stage we do not anticipate using more than two coders at any one time. If we do, a more appropriate index of reliability would be for example the Kappa test in Fleiss (1971).

corrections for three major artefacts: Sampling error; measurement error; and range restriction. Together, these can account for a large proportion of variance between studies. According to Jenkins et al (1998), the best approach to estimating the true population mean of the effect size for salary increase would be a sample size-weighted mean correlation corrected for measurement error and restriction of range. According to Hunter and Schmidt (1990), treatment effects can be converted to d (effect size) or r (correlation) statistics, but because the former is an algebraic transformation of *r*-statistics, treatment effects can be converted to point-serial correlations using a formula in Glass et al (1981). We rely on percentages of variance explained and credibility intervals (width, including zero) to indicate significant moderators.

2.3.2.1 Selection of studies for synthesis (if not all studies are included)

We will use a "q-squared" method of synthesis involving all studies. If we manage to extract a sufficient number of experimental studies (randomised control trial and repeated measures designs, combined), we will conduct a separate meta-analysis for these studies, using meta-analysis (Hunter & Schmidt, 1990).

2.3.2.2 Selection of outcome data for synthesis

Performance can be parsed into individual, group and organisational levels, and is multi-faceted, including job performance, turnover behaviour and global mobility (Carr et al, 2010). We also need to keep the individual research studies separated into the different sector occupations, namely teachers, nurses-doctors-cadres, and judges. It might be possible to test for the impact of pay, and pay/benefits bundles, separately, and to calibrate the magnitude of pay variation, using a percentage-increase scale (this would be included as a covariate in subsequent analysis). However we anticipate that most studies will examine only changes, mostly but not necessarily upward, in pay.

2.3.2.3 Process used to combine and synthesise data

Methodological approaches range from case studies to macro-level regressions, which ideologically do not sit easily alongside each other. In principle however, "Q-squared" (qualitative/quantitative combined) methods (Fanelli, 2007; Hulme, 2007), Ecological Triangulations (Barnett-Page & Thomas, 2009) or Qualitative Comparative Analysis (QCA), might be appropriate for analysing and estimating the different combinations of variables (such as culture) that interact with type of pay increase (e.g., fixed salary versus performance-based increase) to help determine any identified likely performance shift (Rihoux, 2007). This approach would include all studies selected, regardless of their methodological bent (qual or quant). A disadvantage in QCA however is that it requires both antecedents and criteria to be collapsed into a qualitative category, +, 0 change, or -, losing valuable quantitative information.

A less restrictive and more inclusive alternative is reported in Oliver at al (2005). This framework allows both quantitative and qualitative studies to be analysed separately, and later synthesized by an interdisciplinary research team. The approach furthermore respects levels of analysis, and moderators, which are core features in our particular topic and model (Figure 1). Oliver et al describe how a meta-analysis and meta-narrative analysis can be combined by *juxtaposing* qual-quant studies that focus on related constructs. For example, if at-risk youth tell us

that *parental* support is important, do government programs that target *parenting* skills make a difference to youth safety (2005, p. 437)? This technique thus identifies cross-level and q-squared, *convergent validation*.

In our review, to give one hypothetical example, we might ask, are studies highlighting the salience of *individual* equity with the private sector matched by studies that highlight the role of *market* equity? Are salary rises most effective when both individuals *and* culture value status and rank? In this way, the inclusive and pluralistic technique described in detail in Oliver et al (2005) may "move beyond 'what works' to questions concerning what works for whom, in what contexts, and why" (ibid, p. 441).

2.4 Deriving conclusions and implications

Using the techniques above, we do *not* expect to be able to have the requisite data to be able to run a full test of the model in Figure 1 (which is multi-level mediation-moderation). However we *do* expect to be able to respect both quantitative and qualitative traditions, and to position the analysis to draw meaningful conclusions, on (i) 'whether' pay increases in general improve performance (ii) for 'whom' (in what type of work), and 'why' (theory)?

At the dissemination end of the project, our team is well-placed to convey findings through applied development-focused conferences and journal forums. We would submit the paper to a highly-reputable peer-reviewed journal, for example the Journal of Applied Psychology for a quantitative meta-analysis if this proves practicable on the strength of the data, and World Development for a more general, "q-squared" type of report, and would take findings to academics working in the field, senior line-managers in civil service organisations, policy-makers and most importantly of all, stakeholders in lower-income communities (Aguinis et al, 2008). We aim to publish the review in a peer-reviewed international journal with impact and outreach into lower-income settings (e.g., through an Open Access journal).

References

Aguinis, H., & Pierce, C.A. (2008). Enhancing the relevance of organisational behaviour by embracing performance management research. *Journal of Organisational Behaviour, 29,* 139-45.

Agyepong, I.A., Anafi, P., Asiamah, E., Ansah, E., Ashon, D., & Narh-Dometey, C. (2004). Health worker (internal customer) satisfaction and motivation in the public sector in Ghana. *International Journal of Health Planning and Management, 19,* 319–336.

Akerlof, G., & Yellen, J. (1986). *Efficiency wage models of the labor market.* Cambridge, MA: Cambridge University Press.

Ambrose, M. L., & Kulik, C. T. 91999). Old friends, new faces: Motivation research in the 1990s. *Journal of Management, 25,* 231-92.

Barnett-Page, E., & Thomas, J. (2009). Methods for the synthesis of qualitative research: A critical review. BMC Medical Research Methodology, 9, http://www.biomedcentral.com/1471-2288/9/59.

Bartram, D. (2005). The Great Eight competencies: A criterion-centric approach to validation. *Journal of Applied Psychology, 90,* 1185-1203.

Bloom, M. 2004. The Ethics of Compensation Systems, *Journal of Business Ethics*, *52*, 149 – 152.

Bolino, M. C., & Turnley, W. H. (2008). Old faces, new places: Equity theory in cross-cultural contexts. *Journal of Organizational Behavior, 29,* 29-50.

Bradley, S. & McAuliffe, E. (2009). Mid-level providers in emergency obstetric and newborn health care: Factors affecting their performance and retention within the Malawian health system. *Human Resources for Health, 7*, 14.

Brown, R.P.C., & Connell, J. (2004). The migration of doctors and nurses from South Pacific Island nations. *Social Science & Medicine, 58*, 2193-2210.

Carr, S.C., McWha, I., MacLachlan, M., & Furnham, A. (2010). International- Local remuneration differences across six countries: Do they undermine poverty reduction work? *International Journal of Psychology*, in press.

Chew, D.C.E. (1990). Internal adjustments to falling civil service salaries: Insights from Uganda. *World Development, 18* (7), 1003-14.

Chomitz, K., Setiadi, G., Azwar, A., Ismail, N. (1998). What do doctors want? Developing incentives for doctors to serve in Indonesia's rural and remote area. *Policy Research Working Paper (1888)*. Washington DC: Development Research Group, The World Bank.

Cohen-Charash, Y., & Spector, P. E. (2001). The role of justice in organizations: A meta-analysis. *Organizational Behavior and Human Decision Processes, 86*, 278-321.

Colquitt, J. A., Conlon, D. E., Wesson, M. J., Porter, C.O.L.H., & K. Yee Ng. (2001). Justice at the millennium: A meta-analytic review of 25 years of organizational justice research. *Journal of Applied Psychology, 86*, 425-45.

Crewson, P. (1997) Public-service motivation: Building empirical evidence of incidence and effect. *Journal of Public Administration Research an Theory, 7*, 499-518.

Das, J., Hammer, J., & Leonard, K. (2008). The quality of medical eadvice in low-income countries. *Journal of Economic Perspectives, 22*, 93-114.

Dieleman, M., Toonen, J., Touré, H., & Martineau T. (2006). The match between motivation and performance management of health sector workers in Mali. *Human Resources for Health, 4.*

Dieleman, M., Viet Cuong P., Vu, Anh, L, & Martineau T. (2003). Identifying factors for job motivation of rural health workers in North Vietnam. *Human Resources for Health, 1.*

Doucouliagos, H. C. (2010). How large is large? *Preliminary guidelines for interpreting partial correlations in economics*. Melbourne: Deakin University.

Easterly, W. (2006). *The white man's burden: Why the West's efforts to aid the rest have done so much ill and so little good*. London: Penguin.

Eisenberger, R., Rhoades, L., & Cameron, J. (1999). Does pay for performance increase or decrease perceived self-determination and intrinsic motivation? *Journal of Personality and Social Psychology, 77*, 1026-40.

Erez, M. (2000). Make management practice fit the national culture. In E. A. Locke (Ed.), *Handbook of principles of organisational behaviour* (pp. 418-34). Oxford, UK: Blackwell.

Fanelli, J. M. (2007). *Taking stock of global research projects' methodology: The experience with "Understanding reform."* New Delhi: Global Development Network.

Ferrinho, P., & Van Lerberghe, W. (2002). *Managing health professionals in the context of limited resources: A fine line between corruption and the need for moonlighting*. Washington, DC: World Bank.

Fischer, R., & Smith, P. B. (2003). Reward allocation and culture: A meta-analysis. *Journal of Cross-Cultural Psychology, 34,* 251-68.

Fleiss, J. L. (1971) Measuring nominal scale agreement among many raters. *Psychological Bulletin*, 76, 378–382.

Franco, L.M., Bennett, S., & Kanfer R. (2002). Health sector reform and public sector health worker motivation: a conceptual framework. *Social Science and Medicine, 54,* 1255–1266.

Franco, L.M., Bennett, S, Kanfer, R., & Stubblebine, P. (2004). Determinants and consequences of health worker motivation in hospitals in Jordan and Georgia. *Soc Sci Med., 58,* 343–355.

Furnham, A. (2005). *The psychology of behaviour at work: The individual in the organisation.* London: Taylor & Francis.

Greenberg, J. (2008). *Managing behavior in organisations* (5th international edition). New York, NJ: Pearson.

Greenhalgh, T., & Peacock, R. (2005). Effectiveness and efficiency of search methods in systematic reviews of complex evidence: Audit of primary sources. *British Medical Journal, 331, November,* 331-2.

Greenhalgh, T., Robert, G., Macfarlane, F., Bate, P., Kyriakidou, O., & Peacock, R. (2005). Storylines of research in diffusion of innovation: A meta-narrative approach to systematic review. *Social Science & Medicine, 61,* 417-30.

Hjertholm, P., & White, H. (2000). Foreign aid in historical perspective: Background and trends. In F. Tarp (Ed.), *Foreign aid and development: lessons learnt and directions for the future* (pp. 80-102). London: Routledge.

Hofstede, G. (2001). *Culture's consequences: Comparing values, behaviors, institutions and organizations across nations.* London: Sage.

Hulme, D. (2007). *Integrating quantitative and qualitative research for country case studies of development.* New Delhi: Global Development Network.

Hunter, J. E., & Schmidt, F. L. (1990). *Methods of meta-analysis: Correcting error and bias in research findings.* Newbury Park, CA: Sage.

Ihle-Helledy, K., Zytowski, D. G., & Fouada, N. A. (2004). Kuder Career SDearch: Test-Retest reliability and Consequential validity. *Journal of Career Assessment, 12,* 285-97.

Jenkins, G. D., Mitra, A., Gupta, N., & Shaw, J. D. (1998). Are financial incentives related to performance: A meta-analytic review of empirical research. *Journal of Applied Psychology, 83,* 777-87.

Kingma, M. (2007). Nurses on the move: A global overview, *Health Service Research, 42,* 1281-1298.

Kingma M. (2001). Nursing migration: global treasure hunt or disaster in the making. *Nursing Inquiry, 8,* 205–212.

Kluger, A. N., & DeNisi, A. (1996). Effects of feedback intervention on performance: A historical review, a meta-analysis, and a preliminary feedback intervention theory. *Psychological Bulletin, 119,* 254-84.

Kohn, A. (1995). Why incentive plans cannot work. *Harvard Business Review, September/October,* 1-7.

Laffont, J. J., & Martimort, D. (2001). *The theory of incentives: The Principal-Agent model.* Princeton, NJ: Princeton University Press.

Latham, G. (2007). *Work motivation: History, theory, research and practice.* New Delhi: Sage.

Latham, G. P., & Pinder, C. C. (2005). Work motivation theory and research at the dawn of the twenty-first century. *Annual Review of Psychology, 56,* 485-516.

Lazear, E. P., & Rosen, S. (1981). Rank-order tournaments as optimum labor contracts. *Journal of Political Economy, 89,* 841-64.

Lehmann, U., Dieleman, M. & Martineau, T. (2008). Staffing remote areas in middle- and low-income countries: A literature review of attraction and retention. *BMC Health Services Research, 8*, 19.

Locke, E. A., & Latham, G. P. (2002). Building a practically useful theory of goal setting and task motivation. *American Psychologist, 57*, 705-17.

MacLachlan, M., Carr, S. C. & McAuliffe, E. (2010). *The aid triangle: Recognising the Human dynamics of dominance, justice and identity*. New York: Zed.

MacLachlan, M., Carr, S.C., & McWha, I. (Eds.). (2008). *Interdisciplinary research for development: A workbook on content and process challenges*. New Delhi: Global Development Network/GDN Books.

Macq, J., Ferrinho, P., De Brouwere, V., & Van Lerberghe, W. (2001). Managing health services in developing countries: Between the ethics of the civil servant and the need for moonlighting: Managing and moonlighting. *Human Resources for Health Development Journal, 5,* 17-24.

Marai, L., Kewibu, V., Kinkin, E., Peniop, J. P., Salini, C., & Kofana, G. (2010). Remuneration disparities in Oceania: Papua New Guinea and Solomon Islands. *International Journal of Psychology*, in press.

Meyers, R., & Houssemand, C. (2006). Comment évaluer les competences clés dans le domaine professionnel? *Revue Européenne de psychologie appliqué, 56*, 123-38.

Munthali, A., Matagi, L., & Tumwebaze, C. (2010). Remuneration discrepancies in the landlocked economies of Malaŵi and Uganda. *International Journal of Psychology*, in press.

OECD. (Organisation for Economic Co-operation and Development). (2005). Paying for performance: Policies for government employees. *Policy Brief - OECD Observer*, May, 1-8.

Oliver, S., Harden, A., Rees, R, Shepherd, J., Brunton, G., Garcia, J., & Oakley, A. (2005). An emerging framework for including different types of evidence in systematic reviews for public policy. *Evaluation, 11*, 428-46.

Paulsen, R. (2008). Economically forced to work: A critical reconsideration of the Lottery Question. *Basic Income Studies, 3*, 1-20.

Public Health Resource Unit (2006). *Critical Appraisal Skills Programme (CASP): Making sense of evidence*. Oxford, UK: CASP. Available online from: *http://www.phru.nhs.uk/pages/phd/resources.htm* (accessed 11.09.10).

Pugh, G., Mangan, J., & Gray, J. (2010). Do increased resources increase educational attainment during a period of rising expenditure? Evidence from English secondary schools using a dynamic panel analysis. *British Educational Research Journal, 1*, 1-27.

Raabe, B., Frese, M., & Beehr, T. A. (2007). Action regulation theory and career self-management. *Journal of Vocational Behavior, 70*, 297-311.

Robson, C. (2002). *Real world research*. Thousand Oaks, CA: Sage.

Roe, R.A., Zinovieva, I.L., Diebes, E., & Tenhorn, L.A. (2000). A comparison of work motivation in Bulgaria, Hungary, and the Nethelrands: Test of a model. *Applied Psychology: An International Review, 49*, 658-87.

Rihoux, B. (2007). *Qualitative Comparative Analysis (QCA): Getting more out of systematic comparison of cases for the GDN projects*. New Delhi: Global Development Network.

Scariano, S. M., & Davenport, J. M. (1987). The effects of violations of independence assumptions in the one-way ANOVA. *American Statistician, 41*, 123-9.

Stajkovic, A. D., & Luthans, F. (1997). A meta-analysis of the effects of organizational behaviour modification on task performance, 1975-95. *Academy of Management Journal, 40*, 1122-49.

Stajkovic, A. D., & Luthans, F. (1998). Self-efficacy and work-related performance: A meta-analysis. *Psychological Bulletin, 124,* 240-61.

Stone-Romero, E. F., & Rosopa, P. J. (2008). The relative validity of inferences about mediation as a function of research design characteristics. *Organizational Research Methods, 11,* 326-52.

Stride, C. (2008). *Figure it out: multi-level modelling using SPSS.* Sheffield, UK: Institute of Work Psychology, The University of Sheffield.

Terrell, K. (1993). Public-private wage differentials in Haiti: Do public servants earn a rent? *Journal of Development Economics, 42,* 293-314.

United Nations. (2010). *Report of the Secretary General: Keeping the promise – a forward looking review to promote an agreed action agenda to achieve the Millennium Development Goals by 2015.* New York: United Nations.

Van Lerberghe, W., Conceição, Van Damme, W., & Ferrinho, P. (2002). When staff is underpaid: Dealing with the individual coping strategies of health personnel. *Bulletin of the World Health Organization,* 80 (7), 581-4.

Vujicic, M., Zurn, P., Diallo, K., Adams, O, & Dal-Poz, M.R. (2004). The role of wages in the migration of health care professionals from developing countries. *Human Resources for Health, 2.*

Whitehead, M., Dahlgren, & Evans, T. (2001). Equity and health sector reforms: Can low-income countries escape the medical poverty trap? *The Lancet, 358,* September, 833-36.

Willis-Shattrick, M., Bidwell, P., Thomas, S., Wyness, L., Blaauw, D., & Ditlopo, P. (2008). Motivation and retention of health workers in developing countries: A systematic review. *BMC Health Service Research, 8,* 247.

Wright, B. E. (2001). Public-Sector work motivation: A review of the current literature and a revised conceptual model. *Journal of Public Administration Research and Theory, 4,* 559-586.

Zhou, E. I., Li, T., Lu, Z., Li, X., Papola, T. S., Pais, J., & Sahu, P. P. (2010). Remuneration differences in the emerging economies of China and India. *International Journal of Psychology,* in press.

Appendices

Appendix 1.1: Authorship of this report

Professor Stuart C. Carr, Massey University, New Zealand

Professor Malcolm MacLachlan, Trinity College, Dublin, Ireland

Professor Michael Clarke, UK Cochrane Centre and University of Oxford, UK

Professor T. S. Papola, Dr. Jesim Pais, Institute for Studies in Industrial Development, New Delhi, India

Professor Charles Normand, Trinity College, Dublin, Ireland

Professor Steve Thomas, Trinity College, Dublin, Ireland

Professor Eilish McAuliffe, Trinity College, Dublin, Ireland

Dr Chez Leggatt-Cook, Massey University, New Zealand

Details of Advisory Group membership: Dr. Peter Evans, DFID, India (Policy Lead)

Details of Review Group membership: UK Cochrane Centre.

Appendix 2.1: Inclusion and exclusion criteria

Please refer to section 2.2.1 in the Protocol.

Appendix 2.2: Initial Search strategy for electronic databases

At the outset, we have no plans to apply publication year limits to our searches. However, this might be necessary for feasibility and efficiency. If we do apply time limits, our focus will be on the more recent literature (Jenkins et al, 1998 reports that pre-1960 studies lacked rigor), with the post-1960 'year-of-study' recorded as a precautionary potential moderator. Our preliminary work, which informed the discussion provided above, has identified that although the question might appear to be narrowly focused on the surface, the literature itself is both complex and open. In systematic reviews of this nature, it is wise to rely on a combination of both formal (protocol-driven) and informal search strategies, such as "snowballing," i.e., backward reference-tracking and forward citation-tracking (Greenhalgh & Peacock, 2005). The starting point for the search however will be a structured search which will be undertaken by the UK Cochrane Centre.

Search terms: We will use the following elements to conduct abstract and full-text searches of Business Source Complete, and PsycINFO (below). This will be finalised in consultation with the information specialist at the UK Cochrane Centre, and will be adapted for other databases:

Antecedent terms	Mediator terms (Post hoc coding only)	Performance terms
Pay* OR	Motivation	Task performance
Remuneration OR	Satisfaction	Work performance
Salar*	Commitment	Performance
Benefits OR	Efficacy	
Incentive* OR	Engagement	
Financial OR	Citizenship	
Money OR	Initiative	
Monetary OR		
Reward* OR		
Wage* OR		

AND

Change* OR
Increase* OR
Rise* OR
Augmentation* OR
Growth*

AND

Low* income
Middle income
Developing countr*
Developing nation*
Third World

AND

Teacher OR
Doctor OR
Nurse OR
Cadre OR
Judge

Related Thesaurus terms, e.g., "least developed countr*"

_____

Meta-search term: Antecedents AND Performance terms, all possible combinations

_____

Reports will be identified and garnered initially from the following digital sources:

Bibliographic databases:
Our preliminary scoping work has identified an extensive list of databases which might include material of relevance to this systematic review. However, within the time and resources available for this project, it is unlikely that it will be possible to search all of them. We will work with the information specialist at the UK Cochrane Centre to focus the search the grounds of feasibility and efficiency, which may include the use of publication year limits for some of the searches.

Appendix 2.4: Appraisal/Coding tool(s)

APPRAISAL Phase

1. Does the paper report on findings from qualitative or quantitative research and did that work involve data and their analysis?

2. Is the research relevant to the topic?

3. Is the paper anchored in the/a literature?

4. Is the conceptualisation culturally competent?

5. Is the sampling appropriate? (People, organisations, sectors, countries)

6. Is the sample design clearly specified? Are there sampling biases?

7. Are the measures reliable and valid?

8. Do they meet International Test Commission Guidelines concerning construct bias, item bias, and method bias?

9. Are there adequate statistical controls on the measures, e.g., for common method variance and nesting in levels?

10. Is the procedure replicable?

11. Are any statistical assumptions checked?

12. Are appropriate forms of data analysis employed?

13. Overall, is the study of sufficient conceptual, methodological and analytical merit to be included in the review?

CODING phase

14. Code inductively for type of design (repeated measure, randomised control trial), mediators (motives) and outcomes (performance types/dimensions, including level of analysis, either individual or organisational or "ecological" [country/national], or combinations of these). We will also code for the following moderators suggested by or actually in Figure 1:

15. Magnitude of increase (or decrease) in salary, level of poverty (HDR ranking at time of study), time since salary variation, country site (possibly grouped by cross-cultural category indicators in Hofstede, 2001), and governance system (possibly grouped by indicators in the appropriate HDR for the study's year of publication). We will also code for

16. Study author(s), Level(s) of analysis, performance measure, and any core summary and test statistics, ranging from N/n to F and r.