EVALUATING ENGLISH TEACHER EVALUATION

How does teacher evaluation policy in England compare to international policy, practice and evidence?

Tom Perry and Paige Johns

Centre for the Use of Research and Evidence in Education (CUREE)

Background

International Review of Teacher Evaluation Systems – 2017-18

- Commissioned by the Inter-American Development Bank (IDB)
- Carried out by the Centre for the Use of Research and Evidence in Education (CUREE), based in Coventry

- **An extensive literature review - teacher evaluation system design**
  - Evaluation evidence
  - Workforce management and professional development systems and practices
  - Implementation and policy contexts
- **Six teacher evaluation system case studies**
  - High-performing **-** Singapore, Ontario (Canada), New Zealand, South Korea and Estonia
  - Mid-performing - England.
- **A comparative analysis** of the six case study countries

Focus Case Study: England

- *Teacher Evaluation Case Study Areas*
  - Aims and Objectives
  - Outcomes and Indicators
  - Key Components: Process, Methods and Stakeholders
    - Frequency
    - Responsibility
    - Process
    - Sources of evidence
  - Wider system
  - Use, Interpretation and Impacts

How does England compare?

Overview

- Here, we consider:
  1. Prescription, Guidance and Regulation
  2. Stakeholder involvement in the evaluation process
  3. Links with accountability, career, pay and professional development practices
  4. Teacher evaluation methods
- Implications/summary
- Questions

Comparison 1: Prescription, Guidance and Regulation

Comparison 1:
Prescription, Guidance and Regulation

Education (School Teachers' Appraisal) (England) Regulations (2012)

- 3. The governing body must adopt and make available… a document which **sets out the appraisal process**.

- 5. The appraisal period is **12 months**.

- 6. The head teacher must:
  - (a) inform the teacher of the **standards** against which performance will be assessed:
    - "Teachers' Standards" - Secretary of State in July 2011
    - any other standards relating to teachers' performance published by the SoS
  - (b) set **objectives** for the teacher which:
    - improve the education of pupils at that school; and
    - Support any plan designed to improve that school's educational provision and performance.

- 7. The head teacher must:
  - **assess the teacher's performance** of their role and responsibilities in each appraisal period against standards and objectives (see 6).
  - assess **professional development needs** and identify any action that should be taken to address them
  - where relevant, include a **recommendation relating to pay**.

- 8. The head teacher must provide the teacher with a **written report** (covering 7)

South Korea – an extreme comparison

◦ Specifies a detailed points system against which teachers are evaluated. This has:

	Points	Proportion (%)
Teaching Experience	70	32.9
Work Performance	100	46.9
Training Outcomes	30	14.1
Additional Points	13	6.1
Total	213	100

Four review areas, each with a specified weighting (%) (e.g. Work Performance)

Appraisal Area	Appraisal sub-areas	Sub-area points
Qualities and attitude	▪ Intrinsic quality as an educator	10 points
	▪ Attitude as a civil servant	10 points
Work performance and ability	▪ Academic instruction	40 points
	▪ Student guidance	20 points
	▪ Education and research (vice-principals, other administrative management functions including educational activity support)	20 points

Each area is broken down into **numerous sub-areas**, each carrying specified points (e.g. Academic instruction, student guidance)

South Korea – an extreme comparison

Evaluation Area	Evaluation Element	Evaluation Criteria
Academic instruction	Lesson preparation	1. Understanding of the curriculum and teaching and learning methods 2. Understanding of student characteristics and subject matter content 3. Plan teaching and learning strategies
	Lesson implementation	4. Lesson introduction 5. Questioning 6. Attitudes 7. Interaction with students 8. Use of lesson materials 9. Lesson progression 10. Lesson Wrap-up
	Assessment and its use	11. Assessment content and methods 12. Use of assessment results
Student advising	Advising of personal aspect	13. Understanding of students' personal issues and advising for creativity and character development 14. Collaboration with parents in advising 15. Career education and support for talents

◦ Each sub-area is judged against given '**elements**' (e.g. Academic instruction = Lesson preparation, implementation, assessment and its use)

◦ Each element has a number of **evaluation criteria** (e.g. Lesson preparation has three e.g. 'Understanding of the curriculum and teaching and learning methods'

There are set **criteria for observations** (by principals and peers) and **standardised surveys** for use with surveys and parents.

New Zealand – a less extreme comparison

◦ New Zealand has a **highly-devolved, self-managing school system**, with schools having some of the highest levels of autonomy across all OECD countries.

◦ Designed to accommodate the **culture and languages** of indigenous and other ethnic groups: European, Māori (with Māori-language immersion schools); Asian; and Pacific peoples (Pasifika).

◦ (Like England) **Aims and principles in an appraisal framework**, evaluation standards (but more tailored for different groups).

◦ And also (unlike England):

  ◦ Standardised **report** format
  ◦ **Minimum requirements** about process/evidence (e.g. 2 observations)
  ◦ **Vision** of what makes effective evaluation – formative/inquiry, drawing on natural-occurring evidence, embedded in CPD
  ◦ **Tools / Guidance** (see next)

New Zealand – a less extreme comparison

Tools/ Guidance

- Goal Recording Examples
- Observations Guidance and Resources
- Templates
- Guidanace: Putting together your evidence package or portfolio
- "A framework for transforming learning in schools: Innovation and the spiral of inquiry"
- Webinars on the appraisal process

Resources and workbooks

Appraisal as a catalyst for learner outcomes: two years on - a joint report with ERO [pdf]

Workbook - Appraisal for Middle Leaders at NZATE Conference, Waitangi 2017 [download the .doc]

Using Inquiry in Leadership Appraisal - NZSTA Forum [download the .doc]

https://educationcouncil.org.nz/content/appraisal

Comparison 2: Stakeholder involvement in the evaluation process

❖Individual **school governing body** is responsible for agreeing school performance management policy to set out how teachers performance is appraised
  ❖May also be delegated to head teachers and senior leaders
❖**Head teachers** primarily responsible for carrying out performance appraisals
  ❖Often delegate to senior leaders/line managers
  ❖Who conduct classroom observations (formal and informal) and review meetings with teachers
  ❖Pupil test results and targets are often used during the appraisal process
❖Highly variable, often informal (see Page, 2015) and subjective approached to evaluation are used.

Ontario – principal-heavy

❖ **School principals** are responsible for scheduling, conducting and regulating appraisals.
  ❖ Can be **delegated to vice-principals** were appropriate
  ❖ Where it is not possible for wither the board must appoint a supervisory officer
❖ Research suggest that **the teacher performance appraisal (TPA) system has a negative effect** on the relationships between teachers and school principals, owing to:
  ❖ Inconsistencies in implementation
  ❖ Lack of subject expertise
  ❖ Lack of support
❖ Concerns regarding negative impacts on teachers and principals wellbeing
❖ Lack of necessary training for many of the school principals to support them in teacher appraisal
❖ But appraisal is only every 5 years...

Singapore – more sophisticated example

❖Teachers are evaluated by their **supervisor in their school**
   ❖Assessment of teacher's Currently Estimated Potential (CEP), is made by a **panel of supervisors** (principal, vice-principal and heads of department), collectively ranking the group of teachers

❖Teacher evaluations allows **a '360' view** of teachers performance
   ❖**Supervisor** (classroom observations and review meeting)
   ❖**Self-assessment** (form of self-evidence and portfolios)
   ❖**Peers** (peer consultation through panel review)
   ❖**Students** (student results)

❖**Relevant training** received system-wide to carry out the evaluation

❖Strong culture of **collaboration** between teachers and leaders

❖South Korea, is also a system which provides a '360' degree view of teacher's performance
   ❖ use of student survey

Comparison 3:
Links with accountability, career, pay and professional development practices

○ **Formative purposes** of appraisal:
  ◦ Provide context and structure for teachers to discuss their professional development needs
  ◦ Discuss future career aspirations
  ◦ Part of the CPD process (e.g. through evidence-collection and enquiry)
○ **Summative purposes** of appraisal:
  ◦ Linked to teachers salary and/or career progression
  ◦ Identifies unsatisfactory performance
○ Performance related pay has been highlighted as problematic:
  ◦ NUT and ATL teacher unions found 15% of respondents were turned down due to funding or budgetary constraints
  ◦ Simon Burgess: "The requirement to differentiate pay by performance does not appear to be something [schools] wanted to embrace – and it seems that they simply did the least they could to comply with the directive." http://theconversation.com/lessons-learned-from-imposing-performance-related-pay-on-teachers-87657

New Zealand

◦ Coherent system with appraisal linked to professional development process

  ◦ Marks the summative end-point or formative inquiry cycle and identifying goals.

◦ The overall system is designed to be coherent with other components and consistent with the aims
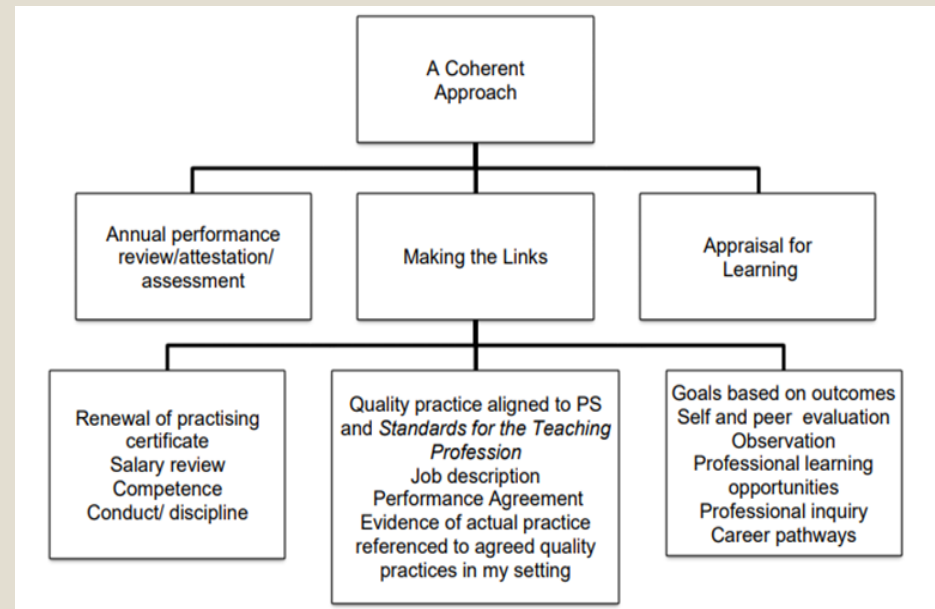
New Zealand – less extreme comparison

- **Annual performance review:** leads to renewal of practicing certificate or competency procedures. Also linked to salary progression

- **Appraisal for learning:** part of goal setting and inquiry process (self- and peer- evaluation, observations)

- **Making the links:** performance review and appraisal for learning all form coherent, joined-up system across contexts and phases (professional standards)

Other country examples

- **Singapore** appraisal links closely to the career tracks present within the system
  - Strong links to CPD
  - Career stages
  - Performance bonus (best example which has made links strong)
  - Recognition
  - Competency/dismissal procedures
- **Ontario –** Annual Learning Plans
- **South Korea** example where links to CPD is seen as a punishment
  - Poor performance linked to mandatory CPD
  - Perception of 'face'
- **Estonia -** linked to their 4 career stages
  - Used as part of professional development
  - Linked to promotion
  - Competency procedures

Comparison 4:
Teacher Evaluation Methods

◦ England – methods/evidence are absent from the standards, regulation, guidance

◦ It is very hard to know what schools are doing! Experience and anecdotal evidence suggests:

  ◦ Methods/evidence used for appraisal are **highly variable**

  ◦ Current practice is heavily reliant on classroom **observations** by head teachers and senior leaders

  ◦ **Ofsted** models (e.g. of observation) have been used by many schools

  ◦ There are many **informal methods** in use (Page, 2015) (learning walks, book scrutiny, informal observations, conversations with colleagues)

    ◦ The DfE model policy mentions 'regular' and 'drop-in' observations

  ◦ **Attainment data** commonly used (e.g. tracking/thresholds/targets)

  ◦ Student feedback is less common (more for for subject/school-level)

Comparison 4:
Teacher Evaluation Methods

- These issues relate to the difficulty of:
  - a) Judging the effectiveness ('causal attribution') of teachers independently of factors which are not, or are only **partially under, the control of teachers**, such as fixed aspects of the learning environment and context, and influences of student background characteristics.
  - b) The impact of **high-stakes consequences** for tests and appraisal.
  - c) The **variation** in teacher performance across time, age-groups and subjects.
  - d) How to approach the **interpretation and use** of imperfect measures.

"I know a good one when I see one..."

Comparison 4:
Teacher Evaluation Methods

Student Test Scores
- Don't use raw scores! (...'Progress' isn't much better!)
- Sophisticated teacher value-added measures (VAMs) can largely remove the influence of non-teacher factors (randomised design studies) but imprecise and highly unstable over time.
- Difficult (and inadvisable) to substantially base teacher evaluation on VAMs in most systems.

Classroom Observations
- Hard to generalise across all pupil groups, teaching contexts and subjects, across range of discrimination (vs. judging the quality of a specific instance of teaching)
- Not currently possible to create highly reliable and context-free estimates of teacher quality (Steinberg and Garrett, 2016)
- Overall ratings - moderate to low year-to-year reliability
- Ratings in specific teacher performance areas have low to very low reliability.
- Observation-based evaluation approaches have strong promise for systems focused on formative aims.
- With:
  - high-quality observation instruments
  - multiple, well-trained observers
  - numerous observations

Comparison 4:
Teacher Evaluation Methods

Multiple Methods and Composite Measures

◦ In general, low to moderate correlations between teacher value-added scores and other sources of evidence (e.g. observations, student surveys, principal ratings).

◦ Combination - The available evidence suggests that spreading weight across indicators tends to produce more stable results and identify components of teaching quality that are common to all indicators (Mihaly et al., 2013).

◦ Questions: moderate, correlations suggest that individual measures capture different aspects of teaching quality and individual measure reliability is moderate at best. This raises questions about high-stakes applications of composite measures and how their components should be weighted to reflect evidence quality and priority.

Comparison 4:
Teacher Evaluation Methods

Portfolios
- Benefits: broader and more reliable judgments; developmental value of collecting, assessing and discussing portfolio evidence
- Challenges: scoring, storage, selectivity in the evidence compiled, (in)ability to capture all areas of practice, suitability for different audiences and the potential to distort practice (Peterson et al., 2001).

Student Surveys
- Can be predictive of student achievement (Bill and Melinda Gates Foundation, 2012) and capture a range of teaching strategies and skills from basic to complex (der Lans *et al.*, 2015).
- Advantages: draw on multiple perspectives cheaply, direct link to student outcomes.
- But: securing teacher acceptance, high-quality survey instrument design, caution with high-stakes or across school (or social group) comparisons

Considering Wider Teacher Outcomes
- Don't forget wider outcomes (attendance, suspensions, academic self-efficacy and intention to pursue further or higher education. (Jackson, 2013).

Implications/Summary

1. Scope for more shaping without prescription – key role for leaders

2. Peer-evaluation, student evaluation? Any use?

3. Coherence, powerful links with CPD?

4. Methods – it appears to be rather difficult...
   ◦ Even sophisticated attempts have low/moderate validity/reliability
   ◦ Practice is highly variable – pot luck?
   ◦ (As per point 1) little attempt to support school leaders do it well
   ◦ Other pressures (retention, recruitment, budgets)

Questions

1. What policy changes would you look into/make?
2. What is the picture on the ground? Good/bad practice?
3. Can/should we make the methods a) more robust, b) more developmental?
4. Can we develop tools for better evaluation?
5. What evaluation CPD do leaders (and teachers?) need?
6. How sophisticated/extensive should links with career progression, pay, school-evaluation, CPD, resource allocation be?
7. Should we be looking at teacher evaluation policy alongside teacher supply and recruitment?

CONTACT

Tom Perry
tom.perry@curee.co.uk
t.w.perry@bham.ac.uk
@TWPerry1

Paige Johns
paige.johns@curee.co.uk
@CUREE_official
www.curee.co.uk