

Open Research Online

The Open University's repository of research publications and other research outputs

Evaluating classroom practice: a critical analysis of approaches to evaluation in large scale teacher education or education technology programmes, in international development contexts

Conference or Workshop Item

How to cite:

Power, Tom and McCormick, Robert (2014). Evaluating classroom practice: a critical analysis of approaches to evaluation in large scale teacher education or education technology programmes, in international development contexts. In: ICERI2014 Proceedings, IATED, pp. 1269–1275.

For guidance on citations see [FAQs](#).

© 2014 The Authors

Version: Accepted Manuscript

Copyright and Moral Rights for the articles on this site are retained by the individual authors and/or other copyright owners. For more information on Open Research Online's data [policy](#) on reuse of materials please consult the policies page.

EVALUATING CLASSROOM PRACTICE: A CRITICAL ANALYSIS OF APPROACHES TO EVALUATION IN LARGE SCALE TEACHER EDUCATION OR EDUCATION TECHNOLOGY PROGRAMMES, IN INTERNATIONAL DEVELOPMENT CONTEXTS

Tom Power and Robert McCormick

¹*the Open University (UNITED KINGDOM)*

Abstract

This study builds on and contributes to work in teacher education and educational technology, in international development contexts. Recent reviews, funded by the UK Department for International Development (DFID) have examined the characteristics of teacher education programmes (Westbrook et al. 2013) and educational technology programmes (Power et al. 2014), that show evidence of impact on teaching practice or learning outcomes. These both illustrate the importance of a strong focus on improving the quality of classroom practice in programme design, and both indicate some of the key characteristics of effective programme support for teachers. But in both reviews, the studies reviewed present problems of evidence. Such evidential problems arise in relation to reporting changes in: attitudes and understanding; teaching and learning practices; and learning outcomes.

In this article, we draw particular attention to evidence of classroom practice: in terms of extensiveness, of methodology, and of understanding the relationships between the variables considered. As such, the purpose of this article is to provide insight into three inter-related issues: the methodological challenges - of rigour, systematic observation, and extensiveness; the practical challenges - of human capacity for research activity, geographical remoteness, and cost; and the evidence requirements of different audiences - donors, policy makers, practitioners and the academic and research communities. This is done by considering these three issues, through a case study of English in Action, a large scale teacher education programme in Bangladesh, in which Educational Technology plays a central role in supporting both teacher professional development, and new classroom practices.

There are several implications from the recent reviews and the case study, that lead us to argue for greater development of evaluation approaches for classroom practice, based upon rigorous, systematic observation (using standardised observations, of objective behaviors). Such approaches must be capable of deployment at scale, and reliable implementation through relatively inexperienced field researchers, available and affordable in country. This may suggest certain kinds of large scale quantitative observation, that are rare in the global north. Is there an opportunity, for a collective accumulation of data, to deepen our basic understanding of classrooms and the actors within them?

Keywords: technology, research, international development, teacher education, large scale.

1 INTRODUCTION

The result of the review of the Millennium Development Goals in relation to education was a recognition that despite the improvement of children and young people's access to education, there remains a problem with some 58 million children out of school in 2012 (United Nations 2014). Associated with this is the issue of the quality of education - a global challenge driving the policies of many governments and funding agencies (DFID 2010; World Bank 2011). Whilst the indicators of the quality of education are themselves subject to debate, classroom practice is shown to be central (Boissiere 2004), hence our focus.

The scale of the problem is very large where there are a great numbers of poorly qualified or unqualified teachers teaching ever increasing numbers of students (Moon 2014). This has given rise to a need to know the most suitable investments to improve quality; usually through the improvement in teacher quality (i.e. teacher education), in learning materials and increasingly in the use of some form of educational technology. In this article we will confine our discussion to the evidence of teacher education and of educational technology.

To inform such large-scale investments requires a good evidence base. Such evidence should inform both how changes in teaching can bring about improvements in student learning and how teacher

education can bring about the changes in teaching. As the address by Moon (2014) indicates, there are criticisms of the state of the research evidence. We examine the literature available to provide evidence about teacher education and about educational technology in low to middle-income countries. From this a number of problems with the evidence are considered. We will then move to an example of a teacher education project (English in Action (EIA)) that used educational technology to improve classroom practice, to examine its efforts to evaluate its effectiveness. This leads us to a number of issues about evaluation and lessons that we draw from both the EIA experience and that of the research evidence literature.

2 COMMON APPROACHES TO THE EVALUATION OF CLASSROOM PRACTICE

2.1 Teacher education in development contexts

It is now a decade since two of the classic reviews of education in developing countries were conducted (Boissiere 2004; Verspoor 2005). Since then there have been reviews of teacher education in both the development context and from the developed world, though many of these are partial. For example, Avasolos (2011) reviews teacher professional development over ten years of publication in the journal *Teaching and Teacher Education* (2000-2010), but it is evident that only one of the over 100 articles reviewed was from a developing country (Namibia).

We consider more recent reviews that indicate a lack of work in relation to outcomes of teacher development either through changes in teachers' practice or student learning outcome: Westbrook et al. (2013) examines the characteristics of teacher education programmes, in a DFID funded review; Tatto (2008; Cordingley 2013) review of the role of research in international policy and practice in teacher education and continuous professional development (CPD) by the British Educational Research Association. In both cases evidential problems arise in relation to reporting change. Although systematic reviews are criticised (Hammersley 2001) and (Hammersley 2004), the reviews do at least try to capture the range of evidence that is available, whatever bias there may be in the processes of filtering sources.

2.1.1 UK Department for International Development (DFID) review of teacher education programmes

This review covered pedagogy, curriculum, teaching practices and teacher education in developing countries (Westbrook et al. 2013), and started with some 2000 articles, narrowing this down to 489 to produce the thematic overview, of which 62 were examined in depth and used to provide the findings. The studies examined in depth revealed the following about how teacher education can support effective pedagogy:

- professional development aligned with teachers' needs, applied in context with follow-up support;
- teacher peer support;
- support from the head teacher;
- appropriate curriculum in terms of level, accessibility and amount of content, and assessment aligned with content.

Only a handful of studies support the findings of each of these elements with peer support being the best with 8 studies. There were a number of gaps in the research:

...with few exceptions ...investigation of the effectiveness of training did not take a holistic view of pedagogy, that is measuring in a single study changes in teachers' attitudes and beliefs, their knowledge...and their practices. Even here, students learning outcomes as a result of ITE or CPD were often not obtained for reasons of scale and feasibility (Westbrook et al. 2013 p.31).

In fact only three of the robust studies measured student outcomes but without baseline and post-intervention data, and only four studies related classroom observation data to any kind of outcomes (e.g. teacher reports of student learning improvements).

Our analysis of many of the original articles referred to in the review, indicates that they were extremely weak in transparency, with many not providing the instruments used or giving full data. Only a handful of studies were of large-scale implementations of an approach to training and involved large

samples in the studies. Three were specifically of pilots that were to be implemented at a larger scale, but only one of these evidently involved a follow-up study (that we could trace in the literature).

2.1.2 British Educational Research Association (BERA) review of role of research in international policy and practice in teacher education and CPD

This is in fact two separate papers, one focusing implicitly on initial teacher education (Tatto 2013) and the other on CPD (Cordingley 2013). The main thrust of them is to examine the role of research in these areas, and in so doing they reveal the findings, and to a lesser extent, the nature of the evidence and the gaps that exist.

The review of initial teacher education (ITE) found that there were few studies that were large in scale and system-wide. These studies were mostly in the school effectiveness tradition and done by economists that consequently disregarded teacher education (TE) programmes' theory of change with poor outcome measures, revealing a 'field (that) lacked a well-developed research infrastructure to adequately study teacher education trajectories.' (Tatto 2013: 3).

The CPD paper (Cordingley 2013), drawing on a number of systematic reviews, lists the elements of effective CPD as:

- sustained collaboration within professional colleagues, including using specialist expertise and peer support
- an understanding of and commitment to professional learning, including enquiry-oriented learning and learning learn from looking;
- a focus on teaching and learning and the aspirations for specific pupils;
- effective scaffolding and modelling of learning by both teachers and leaders for colleagues and pupils.

Unfortunately for the reader, the nature of the evidence can only be judged by going back to the original systematic reviews upon which it is based. For example, on collaboration Cordingley and her colleagues based their in-depth review on 14 studies of which only 11 were finally used, and of these 6 had data on student outcomes. Most had observational data, but these were mainly qualitative and of such a variety of approaches that comparability was difficult. The evidence is mostly from the USA, with only one study in Namibia representing the developing world. The studies were mainly small-scale with 8 having control groups and 9 studying the intervention before and after implementation programmes

2.2 Educational Technology in Development contexts

Educational technology research has been criticised for a historical over-emphasis on access, rather than how the use of educational technology may improve the quality of teaching and learning (Bebell et al. 2010). For example, a critical overview of the effectiveness of ICT policies and strategies in Central and West Asia finds *'...an emphasis in most systems on hardware provision - and the unfortunate but widespread assumption that provision of hardware by itself is the solution to a range of educational problems.'* (ADB 2012 p.iv).

2.2.1 UK Department for International Development (DFID) review of educational technology programmes

In this context, the DFID topic guide set out to examine evidence on the relationship between educational technology, teaching practice and learning outcomes, from over eighty studies in low to lower-middle income countries (Power et al. 2014). The authors found that: *'...as yet relatively few programme evaluations focus on adequately capturing improvements in the teaching and learning process or measuring improvements in learning outcomes.'* (Power et al. 2014 p.5).

There were many studies (Power et al. 2014 p.7-8) whose findings echoed *'...computers are often not used for teaching and learning purposes and that schools and teachers need to be supported in their use'*. (EdQual 2011, p12, cited in Power et al. (2014 p.8)). Whilst other studies were able to provide evidence of improved learning outcomes, these were often in an economic tradition, and were typically silent regarding changes in attitudes or practices. For example, whilst many Interactive Radio

Instruction (IRI) programmes had evidence of improved learning outcomes, only two studies provided evidence of changes in classroom practice (Power et al. 2014 p.6).

Amongst the studies that did present evidence of changes in practice, the nature of evidence was variable. Classroom practices were often evaluated primarily through teacher self-reporting. Sometimes changes in practice were also evidenced through the perceptions of students or other stakeholders, such as head teachers. Studies reporting direct observation of practice were rare.

Just three studies reported changes in teaching and learning practices, from technology enhanced project work (Power et al. 2014 p.7). All of these used small scale but in-depth qualitative case studies, drawing upon multiple data sources, including classroom observation, video recording, individual and focus-group interviews with teachers, teacher educators, head teachers, students and the wider school community. Three further studies were identified, that provided quantitative evidence of changes in classroom practices (Power et al. 2014 p.15-16). One of these, English in Action, is the focus of the final section of this paper.

2.3 Common problems of evidence relating to classroom practices

Many of the issues of evaluation that have been identified are common both across teacher education programmes and educational technology programmes, in international development contexts.

- [1] **locus of evidence:** Much of what is known about teacher education or educational technology, is derived from studies in developed economies. .
- [2] **problems of scale:** There are few large-scale studies of large-scale programmes. Follow-up studies to see if the results of pilots are repeated at scale are largely absent.
- [3] **weak or absent theories of change:** Studies that examine learning outcomes tend to be in the school effectiveness tradition of economists, focussed on outcomes without a theory of change.
- [4] **limited focus on classroom practice:** Few studies have been identified that examine teaching practice, by observation. Synthesis reports of necessity draw findings based on the relatively modest numbers of studies that do.
- [5] **predominance of qualitative methods:** Most studies that address practice draw primarily upon qualitative methodologies and small-scale case studies. There are very few large-scale, quantitative studies of classroom practice.
- [6] **limited accumulation of data:** Studies vary so much in their method and theoretical stance that they make comparison difficult. Therefore, there is little by way of data accumulation and hence progress in development of the field.

In this context, we consider English in Action, as limited illustration of possible responses to these problems.

3 ILLUSTRATION: ENGLISH IN ACTION

English in Action (EIA) is a £50M 9-year (2008 to 2017) project designed to change the way that children, young people and adults in Bangladesh acquire English. EIA was initiated at the request of the Government of Bangladesh (GoB) and is funded by a grant from UKAid. EIA is managed by BMB Mott McDonald, in partnership with the Open University, BBC Media Action and local Non-Governmental Organisations (NGOs). In this paper we consider only the schools component of EIA, as a case study of a large-scale programme of teacher development, in which educational technology plays a central role.

For many years, English language has been the most commonly failed school subject, in national exams in Bangladesh (Kraft, Ehsan, and Khanam 2009). Baseline studies of classroom practice showed that in the majority of classrooms (90% of lessons observed), there were very few opportunities for students to practice speaking and listening in English: classroom practice was dominated by teacher talk (predominantly in Bangla), teachers marking individual student notebooks, and asked closed-questions to individual students (English in Action (EIA) 2009).

English in Action uses an approach to School Based Teacher Development, that includes the use of mobile technology to provide both video materials as a stimulus for teacher development and classroom materials for use with students. This is supplemented by peer support at school and local

level. The approach is designed as an alternative to 'cascade' or 'centre-based training' (Power et al. 2012).

For programme evaluation, EIA uses three large-scale quantitative cohort studies:

- [1] **teacher and student perceptions:** focussing on: teachers views of English Language Teaching (ELT), their own classroom practices, their students experiences, and their views of the programme; students views of English language and their experiences of English lessons. Data is gathered via self-completed questionnaire for teachers and secondary students, and a structured survey interview for primary students. The most recent study (English in Action 2014a) had a sample of 269 primary teachers, 143 secondary teachers, and over 800 students.
- [2] **classroom practices:** focussing on the extent of: the use of the target language (English) by teachers and students; student talktime; organisation of student talk as an individual, pair, group or choral activity. Data is gathered via a timed observation schedule, with simultaneous sampling at one minute intervals, recording objective features of classroom behaviour. The most recent study (English in Action 2012) had a sample of 401 lesson observations, covering 256 primary teachers, and 145 secondary teachers.
- [3] **student learning outcomes:** focussing on students learning outcomes and teachers English language competence, in speaking and listening. Data is gathered through one-to-one diagnostic interviews by independent assessors from Trinity College, London. The most recent study (English in Action 2014b) had a sample of: 605 teachers; 884 (463 primary, 421 secondary) students.

The findings of the most recent studies, are outlined below.

- [1] **teacher and student perceptions:** Teachers report that they felt they had improved their English Language Competence (96-99% of teachers agreed) and their confidence to use English language (88-89% agreed). Most teachers (63-66%) strongly agreed that the programme had impacted their teaching practice, and almost all (89%) agreed they now focus on student communication and interaction in their lessons. Most students (79-80%) report having more opportunity to speak in the target language in English lessons. However, there remains a strong residual attachment to traditional practices: for example, primary students report enjoying learning grammar rules (95%) and being corrected by teachers (98%). English in Action (2014a)
- [2] **classroom practices:** Teachers talked less (45-48%) but used target language more (76-87%). Student talktime increased substantially (to 27% primary; 24% secondary). Student talk in target language increased substantially (to 91% primary; 87% secondary). There were also substantial increases in observations of student talk in pairs or groups (18% primary; 28% secondary). However, despite these increases, primary student talk remained dominated by choral work (46%), and secondary student talk by individual talk (53%).
- [3] **learning outcomes:** primary students showed large improvement over baseline (34% more Grade 1 or above; 20% more Grade 2 or above), with primary girls (74% pass) performing significantly better than boys (65% pass). Secondary students showed significant improvement over baseline (14% more Grade 2 and above, and 11% more Grade 1 and above).

4 DISCUSSION

In comparison to the literature, it can be seen that this represents an unusual set of studies. Firstly, the simple matter of being a large scale follow-on from a smaller pilot, makes this context quite rare. EIA began with a pilot of some six hundred government teachers and a little over one-hundred thousand teachers, and is currently scaled up to reach twelve thousand five hundred teachers, and approximately two million students. The Next year, EIA will scale up again, to reach thirty eight thousand teachers and over four million students by 2017.

Secondly, EIA has a theory of change that spans attitudes and understanding, classroom pedagogy, and learning outcomes, and a range of studies that try to identify changes in relation to core aspects of these. Whilst the research shows how some of these key elements have changed, the studies do not yet evidence the nature of relationship between elements (for example, do teachers with stronger beliefs about the importance of communication for language learning, have classes in which students talk more?). Further work is required in this regard.

Thirdly, there is a strong emphasis on quantitative evidence of classroom practice, and of learning outcomes. However, the methodology does not allow 'progress' of individual students (or teachers) to

be tracked, as there is no 'pre-intervention' assessment for individual teachers or students. Similarly, there is no control group. Whilst this methodology was favoured at the time of programme design, now, almost a decade later, Randomised Control Trials or quasi-experimental designs are considered more robust (DFID 2014).

Fourthly, the use of simple metrics for classroom practice, although providing a relatively crude proxy for students active participation, provides a mechanism that is quantitative, relatively objective, and might allow for data accumulation and comparison between different teacher development or educational technology interventions. Whilst the approach says little about the quality of student talk or the effectiveness of pair or groupwork activities, in the context of large scale programmes, in low income countries, such relatively crude and simple methods might have some merit, in enabling large scale data gathering and data accumulation between studies.

5 CONCLUSION

It is evident that there remain difficulties in producing robust evaluations of projects that have an element of teacher professional development and/or the use of educational technology. This is particularly so in the context of large-scale projects (and hence large-scale evaluations), that can provide the basis of either further investment in projects or in new development work in other contexts. Any investment decision will draw on the existing evidence to substantiate either the need or the basis for the investment. For example, DFID business cases for such investment require the provision of such evidence for any options considered and its rating as high, medium or low (DFID 2011). In addition any 'theory of change' to identify the evidence for the assumptions of how the preferred option will operate, should also indicate its strength and, if there is a particular weakness that would benefit from additional evidence as part of the investment, to indicate this.¹

The review of literature and studies we considered above has indicated that there are serious failings in the evaluations in teacher education and in particular in developing countries. In particular we lament the lack of large-scale evidence and indeed of quantitative studies. Although we cannot comment on the relative dearth of evidence for teacher education world-wide, our experience in EIA indicates some of the difficulties in evaluation in the developing world. In low to middle-income countries, the predominance of qualitative work is a puzzle given the level of skill required to produce robust evidence by these methods.

We argue for programmes of teacher education and/or educational technology, to more explicitly examine their underlying theories of change, and to seek to evidence the nature and extent of changes observed, through research, monitoring and evaluation. In-depth case studies should be balanced by quantitative methods, with emphasis on approaches that might allow large scale data gathering, and data accumulation between studies.

¹DFID gives guidance on strength of evidence for research (DFID 2014).

REFERENCES

- [1] ADB. 2012. "ICT in Education in Central and West Asia." Manila, Philippines: Asian Development Bank.
- [2] Avasolos, B. 2011. "Teacher Professional Development in *Teaching and Teacher Education* over Ten Years." *Teaching and Teacher Education*, no. 27: 10–20.
- [3] Bebell, Damian, Laura O'Dwyer, Michael Russell, and Tom Hoffmann. 2010. "Concerns, Considerations, and New Ideas for Data Collection and Research in Educational Technology Studies." *Journal of Research in Technology Education* 43 (1): 23–52.
<http://files.eric.ed.gov/fulltext/EJ898527.pdf>.
- [4] Boissiere, M. 2004. "Determinants of Primary Education Outcomes in Developing Countries: Background Paper for the Evaluation of the World Bank's Support to Primary Education." Washington DC: Operations Evaluation Department, World Bank.
- [5] Cordingley, P. 2013. "The Contribution of Research to Teachers' Professional Learning and Development." *Research and Teacher Education: the BERA-RSA Inquiry*. London: BERA.
- [6] DFID. 2010. "Learning for All: DFID's Education Strategy 2010-2015." London. Department for International Development.
- [7] DFID. 2011. "Writing a Business Case." How to Note. A DFID Practice Paper. London: Department for International Development.
- [8] DFID. 2014. "Assessing the Strength of Evidence." How to Note. London: Department for International Development.
- [9] English in Action. 2012. "The Classroom Practices of Primary and Secondary School Teachers Participating in English in Action (Study 2a2)."
- [10] EIA. 2014a. "Perceptions of English Language Learning and Teaching Among Primary and Secondary School Teachers and Students Participating in English in Action: Second Cohort (2013)." Research Report. Dhaka, Bangladesh: English in Action.
- [11] EIA. 2014b. "English Proficiency Assessments of Primary and Secondary Teachers and Students Participating in English in Action: Second Cohort." Research Report. English in Action.
- [12] EIA. 2009. "An Observation Study of English Lessons in Primary and Secondary Schools in Bangladesh." Baseline Study 3. Dhaka, Bangladesh: English In Action.
- [13] Hammersley, Martyn. 2001. "On 'Systematic' Reviews of Research Literatures: a 'Narrative' Response to Evans and Benefield." *British Educational Research Journal* 27 (5): 543–54.
- [14] Hammersley, Martyn. 2004. "Systematic or Unsystematic, Is That the Question? Some Reflections on the Science, Art, and Politics of Reviewing Research Evidence." In *An Evidence-Based Approach to Public Health and Tackling Health Inequalities: Opportunities and Challenges*, edited by M Kelly, A. Kanaris, A. Morgan, B. Naidoo, E Barnett-Page, C. Swann, G. Powell, C. Bannon, A Killoran, and Greenhalgh T. Oxford: Oxford University Press.
- [15] Kraft, R.J., A. Ehsan, and R. Khanam. 2009. "Comprehensive Primary Teacher Education and Continuing Professional Development: A Framework for Reform. Draft Final Report." Dhaka, Bangladesh: UNICEF.
- [16] Moon, Bob. 2014. "Address to the Conference Launching the 2013-14 GMR 'Teaching and Learning: Achieving Quality for All'." London: Institute of Education.
- [17] Power, Tom, Rosalind Gater, Catherine Grant, and Niall Winters. 2014. "Educational Technology." Topic Guide. London, UK: Health; Education Advice; Resource Team, Department for International Development.
- [18] Power, Tom, Robina Shaheen, Mike Solly, Clare Woodward, and Sonia Burton. 2012. "English in Action: School Based Teacher Development in Bangladesh." *The Curriculum Journal* 23 (4): 503–29.
- [19] Tatto, Maria. 2008. "Teacher Policy: a Framework for Comparative Analysis." *Prospects: Quarterly Review of Comparative Education* 38 (4): 487–508.

- [20] Tatto, Maria. 2013. "The Role of Research in International Policy and Practice in Teacher Education." *Research and Teacher Education: the BERA-RSA Inquiry*. BERA.
- [21] United Nations. 2014. "The Millennium Development Goals Report 2014." New York: United Nations.
- [22] Verspoor, M., Adriaan. 2005. *The Challenge of Learning: Improving the Quality of Basic Education in Sub-Saharan Africa*. Paris: Association for Development of Education in Africa.
- [23] Westbrook, J, N Durrani, R Brown, D Orr, J Pryor, J Boddy, and F Salvi. 2013. "Pedagogy, Curriculum, Teaching Practices and Teacher Education in Developing Countries: final Report." *Education Rigorous Literature Review*. EPPI-Centre, Social Science Research Unit, Institute of Education, University of London.
- [24] World Bank. 2011. "Learning for All. Investing in People's Knowledge and Skills to Promote Development." The International Bank for Reconstruction and Development / World Bank.