

Information provided by longitudinal assessment data helps governments and schools improve learning achievements

Longitudinal datasets that match individual students' assessment data from two or more points in time allow schools and governments to analyse the extent to which they have effectively raised pupil achievement. By taking students' prior attainment into account they provide 'value added' measures.

Although value added measures are now commonly used to evaluate school quality in UK and some other high income countries, many low income countries do not collect longitudinal data. Therefore raw measures of pupils' academic outcomes and transition rates to the next educational level are used as the *only* key indicators of school quality. Consequently, schools with disadvantaged intakes tend to be judged unfairly leading to demoralization amongst teachers and students. It is also difficult for educational planners to identify best practice and the characteristics of effective schools and administration systems.

At the national/sub-national level, value added approaches can be used to identify inputs and processes associated with effective schools and groups of students (e.g. by income level, ethnicity, parental occupation) that perform above or below average. They also allow schools to identify groups of learners (e.g. high/middle/low ability) that tend to under achieve in the school and the effect of a change in processes or inputs on learning achievement.

Recent research conducted in China and Zanzibar highlights the need for longitudinal datasets and hence improved education evaluation. There is a need to look for system level as well as school effects and to evaluate the effects within school classes and regions. However methods need to be adapted and tailored to the specific educational objectives, priorities and context of the country concerned.

Different assessment outcomes and different explanatory variables are likely to be useful in different countries. The level of sophistication in analysis that is appropriate may also differ (Scheerens, Glass & Thomas, 2003; Yu & Thomas, 2008). Overall the evidence calls for the design of longitudinal datasets and evaluation strategies suited to the particular socio-cultural contexts and economic realities of low and middle income countries. These would support educational improvement initiatives to enable young people to acquire the skills, competencies, values, knowledge, and experience they need for lifelong learning and to be active and productive citizens.

EdQual RPC and education quality

EdQual's research and the **The Improving Educational Evaluation and Quality in China (IEEQC) Project** focus on how to improve the quality of basic education. Raising education quality is one of the six Education for All (EFA) goals and is key to achieving the second Millennium Development Goal (MDG) of every child completing primary education.

Key Policy Messages

Low and middle income countries need to design and collect longitudinal data on educational achievement.

Such data sets provide essential evidence for:

- identifying best practices;
- improving schools and education systems;
- identifying groups of children and localities not accessing education of adequate quality.

However ...

Longitudinal datasets need to be tailored to national (and sometimes sub-national) contexts.

Longitudinal datasets include ...

- (1) **Unique Identifying Codes** for each student, class and school.
- (2) **Student Outcome Measures** - individual student assessment results for different academic subjects.
- (3) **Student Prior attainment Measures** from one or more time points *before* the collection of the outcome measures.
- (4) **Student Background Measures** - individual data on student background characteristics such as gender, age, ethnicity, family income.
- (5) **School/Class Context Measures** - individual data on student characteristics aggregated to the school/class level to estimate, for example mean family income. Any other type of contextual variable outside the control of the school.
- (6) **School/Class Input or Process Measures** - quantitative measures relating to any input or process aspects of schooling (e.g. teacher qualifications or length of school day). This data indicates potential explanations for observed differences in schools' value added effects.

China—Improving Educational Evaluation and Quality

About the research

The Improving Educational Evaluation and Quality in China project used value added methods and innovative quantitative methodology (multilevel modeling) to investigate school effectiveness in China. It also explored the local application of new school evaluation methods to educational policy and practice in rural and urban senior secondary schools. The research sought to promote the development of innovation in school evaluation and guidelines for implementation via bottom-up and top-down dialogues involving key stakeholders such as policy makers, Local Education Authority officers, teachers and students. A key rationale underpinning the project is that understanding education quality and evaluation processes will also assist in achieving wider goals including social justice, social cohesion and equal opportunities.

Methodology

Longitudinal data from 90,000+ students in 120+ schools was analysed to estimate the size and extent of school effectiveness in China. Multilevel modeling was used to create contextualized value added measures from the datasets - which comprised for each student their entrance examination to Higher Education scores matched to their entrance examination for senior high school scores as well as other relevant student background and school context factors. Interviews were also conducted with key stakeholders.

The research sought to promote the development of innovation in school evaluation

Key Policy Pointers from China Research

Value added measures provide a different picture of school quality

Significant differences were found both between and within schools in terms of value added measures. Crucially, these measures of school performance provided a different picture of educational quality in comparison to the raw examination scores. Stakeholders reported that value added measures would provide an important and welcome addition to current school evaluation systems in China but they need to reflect local priorities. It is also important to raise awareness of the limitations such as measurement error.

A range of value added measures are required for different subject groups and groups of students

Many Chinese stakeholders consider it crucial to extend the evaluation of student outcomes to include all-round development and non-academic outcomes such as vocational and attitude measures. This is in line with Western concerns but seems to go much further, due in part to a new national policy emphasis on the importance of students' all-round development. Cultural issues (e.g. Confucius thinking), the intense competition between students given limited Higher Education opportunities and the emphasis on improving equity, especially in terms of the rural-urban/East-West gap, also seem to play a very critical role in how educational quality and evaluation systems are perceived in China.

3. Data quality is crucial

Rigorous and systematic longitudinal data collection procedures are required to ensure data quality—especially the use of a unique student identifier to match data records over time. Also crucial are explicit agreements between schools, administration and research organisations collecting data.

Consider creating regional student databases in China

Differences in examination systems between provinces and cities (particularly at lower educational levels) means that creating a national value added system would be very difficult. However value added evaluation systems are feasible for regions or cities so consider the possibilities for regional student databases within a nationally agreed framework.

Value added measures would provide an important addition to current school evaluation systems in China

Focus on school self-evaluation and school improvement

Stakeholders reported a new government focus on school self evaluation and school improvement would be welcomed, as well as reform of Higher Education entrance requirements to reduce the focus on raw examination scores. This would involve widespread and comprehensive training to support teachers' use of new evaluation concepts and methods.

MEASURING EDUCATION QUALITY

Some Considerations

- From whose perspective is quality judged?
- Which area of activity within an organisation determines quality?
- At what level of the organisation is quality analysed?
- How is quality defined in terms of time?
- What data are used to form an opinion of quality?
- What standards or measures are used in order to make quality judgements?

(adapted from Scheerens, 1992)

Zanzibar—School Effectiveness and School Self Evaluation

About the research

As in China, the Zanzibar project used value added methods involving innovative longitudinal datasets and quantitative methodology (multilevel modeling) to investigate secondary school effectiveness in Zanzibar. The aim was to examine the applicability of value added measures in this low income country context and investigate for the first time in the East Africa region the range and extent of school effectiveness using a longitudinal methodological approach. The possible explanations of any differences observed in school effects and the relevance and potential for using value added measures to enhance evaluation processes in Zanzibar were also explored.

Methodology

Survey and interview data were collected from 7,356 students and 110 head teachers; including student background characteristics and school context information, which was subsequently matched to students' national assessment outcome data (Form 4 examination results taken at age 18 years) and their prior attainment (Form 2 examination results taken at age 16 years). Multilevel modeling techniques were used to calculate contextualised value added measures of relative student progress, which were additionally tested against head teachers' views on educational processes and other school factors. Interviews were also conducted with education stakeholders.

Key Policy Pointers from Zanzibar Research

Value added methods can be applied in Zanzibar

The findings indicated that statistically significant differences do appear to exist between Zanzibar secondary schools in terms of effectiveness, typically somewhat greater than countries such as UK and China (Thomas & Peng, 2011) and somewhat smaller than some sub-Saharan African countries such as South Africa, Kenya, Namibia and Uganda (Lee, Zuze and Ross, 2005). Within individual schools there is also evidence of differential effectiveness for different student groups (by prior attainment and gender) as well as across curriculum subject areas. Importantly head teachers' views on school process factors were found to explain some of the differences in schools' value added results.

Value added school evaluation would improve education quality in Zanzibar

A value added school evaluation framework could promote systematic record keeping of student attainment and progress and support less effective schools in achieving equitable learning outcomes. This is because school self-evaluation motivates teachers to improve their practices. Moreover, value added measures could feed into efforts to improve accountability and assist policymakers in identifying where best to focus improvement efforts and limited resources.

Data quality, especially of assessment data, is crucial

Stakeholders raised some pertinent student assessment issues, including in relation to Form 2 examinations: quality of assessment instruments and teacher assessments; adequacy of assessment record keeping; match between curriculum and assessment design and purpose; teacher competence in student assessment; language of instruction and assessment; and political manipulation of pass marks. Given student examination results are key indicators of student outcomes used in national educational evaluation systems, the findings highlight that fit-for-purpose student assessment is a critical area for development and improvement in Zanzibar.

School self evaluation motivates teachers to improve their practices

Research Limitations (China & Zanzibar)

The quality of assessment data depends on the reliability and validity of the student assessment systems, which provide baseline and outcome measures. In particular, there were concerns about the quality of national assessment in Zanzibar.

There were technical data issues, such as measurement and other possible errors in the data. The statistical significance of results was always considered (Goldstein, 1997).

Research Strengths (China & Zanzibar)

- Size of the datasets - a complete or representative national or regional cohort of students was used.
- The detailed and longitudinal nature of data collected - students' prior and outcome attainments were collected in two or more curriculum subjects, as well as additional student background characteristics.
- The collection of school process variables, which were used to explore possible explanations of differences observed in school effects.
- The use of mixed methodology and evidence from both quantitative and qualitative research methods to provide complimentary sources of explanations and illustrations of "what works" in promoting educational quality in different contexts.

Research Sponsors and teams

The **Improving Educational Evaluation and Quality in China (IEEQC) Project** research was funded by the UK Economic and Social Science Research Council (ESRC) and UKAID. It was conducted by the University of Bristol, UK in collaboration with China National Institute for Educational Research (CNIER) (Thomas & Peng, 2011; IEEQC, 2011)

The **School effectiveness and school self evaluation in Zanzibar (SESEZ) Project** was funded by UKAID and conducted as part of the EdQual research programme consortium as PhD research by Massoud Salim, from the Ministry of Education and Vocational Training, Zanzibar (Salim, 2011).

Further Reading

Salim, M. (2011) Summary of School effectiveness and school self evaluation in Zanzibar (SESEZ) Project. Research Report prepared for EdQual.

Thomas S. M. (2010), Assessment and the Evaluation of Institutional Effectiveness. In: Penelope Peterson, Eva Baker, Barry McGaw, (Editors) *International Encyclopedia of Education*, volume 3, pp. 172-180. Oxford: Elsevier.

Thomas, S.M., Salim, M., Munoz-Chereau, B. and Peng, W.-J (2012, in press) Educational Quality, Effectiveness and Evaluation: Perspectives from China, South America and Africa. In Chapman, C. et al (Editors.) *Challenging the Orthodoxy of School Effectiveness and School Improvement: Towards new theoretical perspectives*. London: Routledge.

References

Goldstein, H. (1997). Methods in School Effectiveness Research. *School Effectiveness and School Improvement*, 8 (4): 369—395.

IEEQC (2011) Improving Educational Evaluation and Quality in China Website. Accessed 1st July 2011: <http://ieeqc.bristol.ac.uk>

Lee, V. E., Zuze, T. L. and Ross, K. N. (2005). School effectiveness in 14 sub-Saharan African countries: Links with 6th graders' reading achievement. *Studies in Educational Evaluation*, 31(1):207—246.

Salim, M. (2011) Exploring Issues of School Effectiveness and Self-Evaluation at the System and School Levels in the Context of Zanzibar. Unpublished PhD thesis. University of Bristol.

Scheerens, J. (1992) *Effective schooling: Research, Theory and Practice*. London: Cassell.

Scheerens, J., Glas, C. & Thomas, S. (2003) *Educational Evaluation, Assessment and Monitoring. A systemic approach*. Swets & Zeitlinger Publishers.

Thomas, S. & Peng, W. J. (2011). Improving Educational Evaluation and Quality in China. Final report to ESRC.

Yu, G. & Thomas, S. (2008) Exploring school effects across Southern and East African school systems and in Tanzania. *Assessment in Education* 15(3): 283-306.

This policy brief was written by:

Prof. Sally Thomas, University of Bristol

Dr. Massoud Salim, Ministry of Education and Vocational Training, Zanzibar.

Contact: S.Thomas@bristol.ac.uk

For more EdQual Policy Briefs or further information please see www.edqual.org or contact EdQual UK:

Tel: +44 (0)117 331 4284

Email: angeline.barrett@bristol.ac.uk



The Research Team

Prof. Sally M. Thomas, University of Bristol

Dr. Massoud Salim, Ministry of Education and Vocational Training, Zanzibar

Dr. Wen-Jung Peng, University of Bristol

Dr. Guoxing Yu, University of Bristol



www.edqual.org

EdQual is a consortium of six higher education institutions in the UK and Africa funded by the UK Department for International Development (DFID) to carry out a five year (2005-2010) programme of research on education quality mainly in low income countries, focusing on sub-Saharan Africa.

EdQual partner institutions are:
 University of Bristol, UK (lead)
 University of Bath, UK
 University of Cape Coast, Ghana
 University of Dar es Salaam, Tanzania
 Kigali Institute of Education, Rwanda
 University of the Witwatersrand, South Africa