

Good Intentions Are Not Enough: Evaluation is Essential

Stephen Tall (www.educationendowmentfoundation.org.uk)

This article addresses programme evaluation utilising rigorous methodological approaches including randomised control trials to evaluate the effectiveness of their interventions.

Other approaches to measuring impact can be seen in the following examples (this is not a comprehensive list, just a sampling):

- Social Impact Analysts Association www.siaassociation.org. They have free online resource signposting current information about the measurement, analysis, assessment and evaluation of social impact and social value worldwide.
- The SROI Network www.thesroinetwork.org, is a social enterprise that encourages and advocates for measuring social return on investment – a monetary evaluation of the value of an organisation's interventions.
- New economics Foundation www.nef-consulting.co.uk/?s=evaluation has reports on its site covering outcome evaluation and impact assessment.
- New Philanthropy Capital www.thinknpc.org/ studies on impact on its web site.
- EVPA report on social impact strategies in banks – evpa.eu.com/blog/2014/03/new-evpa-publication-social-impact-strategies-for-banks/
- Inspiring Scotland VP organisation www.inspiringscotland.org.uk/media/11731/An-Independent-Research-Report-Inspiring-Scotland.pdf
- LSE paper (the paper mentioned in relation to the distinction between social investment vs impact investing) files.lsecities.net/files/2013/10/Measuring_Impact-full-length-Oct-20131.pdf
- Pro Bono Economics www.proboneconomics.com evaluation of services and programme including 'An assessment of the potential savings from Barnardo's interventions for young people who have been sexually exploited.'
- Bridges Ventures <http://bridgesventures.co/category/news/> see Bridges web site for a description of the social investment funds and reports including Bridges Impact report A Spotlight on our Methodology.
- Nesta's standards of evidence paper www.nesta.org.uk/publications/nesta-standards-evidence

The EEF was founded in 2011 by lead charity The Sutton Trust, in partnership with Impetus Trust, with a £125m grant from the Department for Education. To date, more than £9.5m has been pledged by a range of organisations to help extend the reach of the work the EEF is funding.



Stephen Tall

Good intentions are not enough. Let me give you an example. A programme called ‘Scared Straight’ was developed in the USA in the 1970s to deter juvenile delinquents and at-risk children from criminal behaviour by bringing them into contact with adult inmates to make them aware of the grim realities of life in prison.



Early studies showed astonishingly high success rates, as much as 94 per cent, and the programme were readily adopted in the UK and other countries. However, none of these evaluations had a ‘comparison group’ showing what would have happened to the participants if they had not taken part. When tested through Randomised Controlled Trials it was discovered participation in ‘Scared Straight’ resulted in higher rates of offending behaviour than non-participation: “doing nothing would have been better than exposing juveniles to the program”.¹ Yet it continues to be championed by some British police forces despite the clear evidence it actively increases crime.

What this illustrates is the importance of ‘the counter-factual’ - i.e., what would have happened otherwise? This is a crucial question for philanthropists, all of whom will have greater calls on their generosity than they can possibly meet. Inevitably this means there is an opportunity cost in making a donation: whatever money you give to one charity is, of necessity, money denied to another.

All philanthropists are acutely aware of this responsibility. But how many can confidently say their decisions to fund one charity over another are always based on sound evidence? And how many, when making their donation, also seek to ensure the work they are supporting is being robustly evaluated to ensure it is doing the good everyone hopes it will? Put bluntly, how do you know your money is not being used to fund another ‘Scared Straight’, a programme developed with the best of intentions, but which inadvertently did harm to the young people it aimed to help?

At the Education Endowment Foundation (EEF) we begin with the existing evidence. In our first three years, we have awarded grants for 87 different projects – often co-funded with partners – working in some 2,400 schools and involving more than 500,000

pupils. Our grant-making is informed by the evidence in the Sutton Trust-EEF Teaching and Learning Toolkit, a synthesis of more than 10,000 high-quality research reports, that what we are trialling will raise the attainment of the pupils involved, and that it will make a particular difference for those from low-income backgrounds.

For example, the evidence in this Toolkit is that ‘feedback’ (how children’s effort and activity can best be focused to achieve their goal) can deliver high impact for low-cost. We have, therefore, funded eight projects that will give us a much better understanding of what effective feedback might look like in the classroom.

Though the EEF backs only those projects we think have the best evidence of promise that they will raise children’s attainment and narrow the gap between rich and poor, it is inevitable that not all will work out as well as we hope. We appoint independent evaluators to make sure that neither we (as the funders) nor the delivery organisation (as the grantee) are conflicted. Working collaboratively, we design trials which aim to give the project we are funding the best chance of success in the ‘real world’ environment of English primary and secondary schools; but, crucially, which will also subject the project to a robust test so we find out if its good intentions are matched by pupils’ progress.

Too often, impact evaluations are little more than ‘before and after’ studies which will make claims such as “children’s performance increased by 67% as a result of our work”. The statistic might sound impressive, but it doesn’t tell us whether the improvements would have happened in any case: it doesn’t answer the counter-factual. After all, it’s quite possible the attainment of those children might have improved more under business-as-usual conditions or if a different intervention had been tried instead. We just don’t know. In our heads we accept that ‘correlation

does not imply causation', but it's amazing how often we are willing to suspend scepticism and follow our hearts when offered such false confidence, even if it isn't justified by the evidence.

The independent evaluations the EEF funds aim to build the evidence – both quantitative, mostly Randomised Controlled Trials, as well as qualitative – of 'what works' in improving educational attainment. All will be reported in full and in public so that schools and policy-makers can make use of the findings in their own work.

We hope the EEF's work will have widespread relevance. For example, we are currently helping design and fund four trials which will test within 8,000 schools how evidence can best be used to improve teaching. Which works best: face-to-face instruction or access to websites? Twitter chats or posting information booklets to schools? Professional development sessions or research conferences aimed at teachers? The trials will provide some answers to these questions, bringing us closer to building a system that can cost-effectively keep teachers informed about research and help them achieve the best possible outcomes for students. There are, we think, implications here for others involved in sharing effective practice in many other areas of social policy.

By no means everything the EEF does is about large-scale Randomised Controlled Trials. With Durham University, we have written an online DIY Evaluation Guide for teachers and schools. This introduces the key principles of educational evaluation – in particular the use of comparison groups – and provides practical advice on designing and carrying out small-scale evaluations in schools. It is intended to help teachers and schools understand whether the interventions they are developing are effective within their own school context.

This gets to the heart of the EEF's mission. Our role is to support schools testing new ways of boosting the attainment of their pupils, especially the most disadvantaged. But this comes with two important professional responsibilities: for us, as funders, but also for our grantees, as practitioners. First, that this should be 'informed innovation', innovation that builds on what we already understand from existing evidence. And secondly, that these new approaches are robustly evaluated so we find out if what we hoped to see happening is what is actually happening. In other words, that our good intentions are leading to good outcomes for children.

A fair test to find out if Chess in Schools raises attainment

Can learning to play chess improve children's ability to develop thinking skills and boost their attainment? That's the question being asked by one of the 87 trials the Education Endowment Foundation (EEF) is funding.

Delivered by the charity Chess in Schools and Communities, the programme involves children in Year 5 (ie, 9-10 year-olds) being taught chess by accredited coaches for one hour a week over 30 weeks during normal school time.

There is good evidence to suggest this might make a difference to attainment: a Randomised Controlled Trial (RCT) in Italy found that learning chess can have a positive effect on pupils' progress in Mathematics. However, we cannot simply assume the same gains will automatically apply within the English school context.

The EEF has, therefore, appointed academics from the Institute of Education, University of London, to carry out an RCT – one of 74 RCTs we are funding – designed to estimate intervention impacts by creating equivalent groups, one of which will receive the intervention and the other of which will not.

The charity has recruited 100 primary schools from a range of locations: 50 will receive chess coaching during the evaluation and the other 50 (who will act as the 'comparison group') will receive it two years later. In this way, all the children will receive coaching in chess, but the evaluation will be able to estimate the difference the programme has made to pupils' academic progress as measured by their performance in Key Stage 2 tests. An online survey, in-class observations and interviews with teachers will be used to test the feasibility of the Chess in Schools programme.

The evaluation report will be published in 2016.

¹ See Laura Haynes, Owain Service, Ben Goldacre and David Torgerson: 'Test, Learn, Adapt: Developing Public Policy with Randomised Controlled Trials' (Cabinet Office - Behavioural Insights Team, 2012), p.17.

https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/62529/TLA-1906126.pdf