Test and Learn: A global revolution in teacher-led research
Introduction

What is Test and Learn and how does it work?

Empowering practitioners as creators of new evidence and learning ................................. 08
How can practitioner-led research strengthen school systems and improve outcomes for children? ................................. 10
How does it work? ....................................................................................................................... 14

Case studies

How to understand the results section in an individual teacher poster ................................................. 18
Case study 1, Miriam Mason-Sesay, Sierra Leone ........................................................................ 19
Case study 2, Luis Miguel Bermúdez Gutiérrez, Colombia ................................................................. 23
Case study 3, Kenneth Silburn, Australia ....................................................................................... 27
Case study 4, Souad Belcaid, USA ............................................................................................... 31
Case study 5, Francis Jim Tuscano, The Philippines ....................................................................... 35

What’s the prize?

Appendix and references

Appendix 1: Meta-analysis of teacher research effect sizes .......................................................... 44
References ........................................................................................................................................ 47
About Education Development Trust ....................................................................................... 49
About the authors ......................................................................................................................... 50
In this report, we explore the impact and potential power of Education Development Trust’s Test and Learn approach to identifying what works (and what doesn’t), in what contexts and for which pupils. We share the personal experiences of five of the Varkey Foundation’s Global Teacher Ambassadors in conducting experimental research in their classrooms around the world. With thanks to the Varkey Foundation, these teacher-practitioners, alongside 13 others, were the first to carry out simultaneous experimental research studies, including randomised controlled trials, in their classrooms around the world. We believe that their stories and the emerging impact of their work marks the start of the global revolution in teacher-led research.

The teacher-researchers
- Akash Patel
- Alexander Rubio Alvarez
- Andrews Nhessie
- Eligio Salamanca
- Frances Jim Tuscano
- Hiba Ballout
- Itodo Samuel Anthony
- Katerine Franco Cárdenas
- Kavita Sanghvi
- Kenneth Silburn
- Luis Miguel Bermúdez Gutiérrez
- Marcella Henriquez
- Marj Brown
- Miriam Mason
- Morris Oates
- Odeogbola Ayodele
- Sarah Mathews
- Souad Belcaid
- Thomas J. Rusk

Acknowledgements
We would like to express our deepest gratitude to the Varkey Foundation for sharing our enthusiasm for the power of teacher-researchers and funding the work highlighted in this report. Thank you to the Varkey Teacher Ambassadors listed above who embraced the challenge and worked hard to use our methods and generate the findings.
Introduction

Test and Learn: Building scientific literacy among education practitioners

Millions of pounds are spent on education research every year, yet compared to the role that research plays in medicine and healthcare, education lags dramatically behind. Education receives a fraction of the funding that medicine and healthcare receives. But perhaps more fundamentally, funding tends to go to researchers who no longer practice, or in some cases have never practised in the classroom. This disconnect between serving teachers as consumers rather than producers of the key research evidence is unsustainable. In many cases, it leads to research foci that are distant from the practical day-to-day concerns of teachers. In contrast, in medicine and healthcare, serving clinical practitioners do much of the work translating evidence from 'bench to bedside'. This publication focuses on one approach we have created to bridge that gap and support teachers to not only engage in research but produce evidence that has the potential to generate real change in education.

Test and Learn is a practical, accessible approach to education research that builds scientific literacy among practitioners, enabling them to test and evaluate theories, approaches, pedagogies and policies, and to interpret, challenge and critique the wider evidence base. It unleashes the voice and agency of teachers, empowering them to make evidence-based decisions and improve their practice. It was developed in response to requests from teachers themselves who participated in a large-scale Randomised Controlled Trial (RCT) on a Department for Education project in England, Closing the Gap: Test and Learn. Teachers discovered the potential of carrying out research in their classrooms and sparked the idea for us to develop a method and analysis tools to make it accessible to all practitioners without a background in research or statistics.

Taking the Test and Learn innovation to scale

During ‘Closing the Gap-Test and Learn’ we supported teachers to carry out 50 teacher-led studies, some of which were published in ‘Evidence that counts: A meta-analysis on teacher-led randomised controlled trials and other forms of experimental research.’ This programme made a major contribution to the government’s drive towards a self-improving school system where teaching is an evidence-based profession, paralleling evidence-based practice in medicine. The programme also built research capacity in teaching schools by providing training in experimental research methods for use in teacher/practitioner research and for school improvement. In 2018, we worked with the Wellcome Trust, Eleanor Dommett and Ian Devonshire to use Test and Learn with teachers to investigate the impact of neuroscience-informed teaching practice on learning outcomes, leading to 31 trials and replications (34 findings). More information can be found in a forthcoming article in the US Journal Mind, Brain and Education: ‘Translating laboratory evidence into classroom practice with teacher-led randomised controlled trials – a perspective and meta-analysis.’ The meta-analysis showed overall, positive effects were found over short periods (one to six weeks).

Infinite Learning, a private teacher professional development organisation in Dubai, and Pickwick Teaching School Alliance in the UK have contracted us for three consecutive years to deliver teacher-led research training and support teachers to carry out research. In total these engagements have alone produced around 100 teacher-led RCTs. Introductory workshops have also been delivered in Kenya, Dubai and Jakarta and teacher research has been presented at academic conferences around the world.

Building the capacity of the world’s best teachers in Test and Learn

The Varkey Foundation saw great potential in Test and Learn to build the capacity of some of the best teachers in the world to carry out classroom-based research and investigate ‘what works’ in different contexts. The project featured in this report was a collaboration with nineteen Varkey Foundation Teacher Ambassadors, teachers who feature in the top 50 finalists of the Global Teacher Prize, a highly prestigious award, presented annually to an exceptional teacher who has made an outstanding contribution to the profession. They were all trained in the Test and Learn method, to design and deliver practitioner-led randomised controlled trials in classrooms and collaborating schools. They learned about the concepts of ‘effect size’, ‘confidence interval’ and ‘significance’, and were introduced to a range of research methods and our bespoke data analysis tool, StatsWizard. They were guided through a scaffolded approach to define their hypotheses and design their protocols. All follow-up support was provided remotely in groups and one-to-one sessions to refine the methods and analyse results.

The Varkey Foundation is a family foundation established to improve standards of education and raise the status and capacity of teachers throughout the world.
Nineteen trials were completed in twelve countries
• Australia
• Chile
• Colombia
• India
• Lebanon
• Malawi
• Nigeria
• Philippines
• Sierra Leone
• South Africa
• Sri Lanka
• United States

Teachers evaluated the impact of their interventions on a total of thirty-four measures covering the following outcome areas:
1. progress/attainment;
2. attitude/confidence; and
3. personal/emotional.

The project culminated in a series of academic-quality research posters, illustrating the work and a conference poster event at the Global Teacher Prize in Dubai. The Test and Learn approach allows for the combination of teacher results in a single analysis (known as a meta-analysis). Overall there were very positive effects on a range of different measures (see Appendix 1).

What is Test and Learn and how does it work?
Empowering practitioners as creators of new evidence and learning

At Education Development Trust we believe that equipping teachers and education practitioners with scientific literacy is critical to building a more empowered profession, where individuals are valued for the evidence-based professional judgements they make. We are not alone in acknowledging the importance of teachers and practitioners being able to access and draw on existing evidence to inform their practice or indeed carry out their own qualitative research to test ‘what works.’ Where we differ is in our belief in the power and potential of tightly controlled, practitioner-led, classroom-based quantitative research, usually the domain of academics and think tanks, to deliver robust, reliable data with high levels of both internal validity (through the use of tightly controlled research over shorter timescales) and external validity (through replication and the meta-analysis of results across such trials).

Five of the Varkey Foundation teachers’ posters can be found in an open out format towards the end of this publication.

Using techniques that are common in science, medicine and healthcare, teachers learn about the use of statistics in assessment, while making substantial contributions to the common understanding of ‘what works’ in education. In science, medicine and healthcare, it is often serving clinicians and professionals that publish studies on clinical practice. However, in education few practitioner studies are widely disseminated or reach journals, and those studies that are published are by those who no longer teach. The ‘democratic deficit’ in education research is further compounded by the fact that most teacher education has, until recently, limited itself to teaching teachers to do small-scale qualitative action research, rather than controlled quantitative evaluation.

With the aim of directly exploring how this gap in the education system could be filled, we sought to place teachers directly in the ‘clinical researcher space.’

Test and Learn differs from more traditional approaches in that it enables collection of robust quantitative evidence by practitioners themselves.

Teachers are trained to write up their research in a conference poster format following a typical scientific research format:

- Purpose
- Research Design (including hypotheses)
- Method, consisting of:
  - a) Participants, sample size and randomisation;
  - b) Procedures;
  - c) Materials (and apparatus)
- Results
- Limitations
- Conclusions and recommendations for future research

This gives their research a better platform for dissemination and empowers them to accurately research and critique the wider evidence base produced elsewhere.

‘When you do something different in the classroom it is often difficult to assess the effect. And prove to others that the new method is better than the old.’

Kenneth Silburn, Sydney, Australia

‘Think there is a lot of observational or non-scientific research that we already do and that is possible to improve and make more scientific. It is also possible to create support networks with other teachers to facilitate and enrich this work.’

Eligio Salamanca Navarrete, Pucon, Chile

‘Test and Learn helped me to evaluate my pedagogical practice and motivate my colleagues to set new strategies to improve our education quality. It’s important to see the classroom as a laboratory of new learnings.’

Katerine Franco Cardenas, Medellin, Colombia

12 Biesta (2007)
How can practitioner-led research strengthen school systems and improve outcomes for children?

Evidence from our projects, gathered using an evidence-based practice questionnaire (adapted from medicine and healthcare)\(^{13}\) supports the theory that teachers and practitioners who participate in Test and Learn and carry out their own quantitative, experimental research, build their scientific literacy, enabling them to access and use evidence as a matter of routine to significantly improve their practice, in turn enhancing outcomes for students.\(^{14}\) This is further substantiated by qualitative evidence gathered from Varkey Teacher Ambassadors in the process of writing this report, illustrated in the case studies.

At Education Development Trust we believe in six school system accelerators that are required to effect long term, sustainable change at the system level. Test and Learn has the power to significantly enhance every one of these accelerators as illustrated in the diagram below.

93% of the interventions designed by the Varkey Teacher Ambassadors had positive effects on children’s outcomes

---

\(^{13}\) Upton and Upton (2006); Upton, Upton and Scurlock-Evans L (2014)  \(^{14}\) Churches, Hall and Higgins (2017b); Education Development Trust (2018)
1. Vision and leadership

Test and Learn enables teachers and leaders to develop a shared vision of what makes great teaching and learning in their schools in their contexts.

Souad Belcaid was able to demonstrate the positive impact of an online, interactive, subject-specific website and galvanise interest and support from her leadership team. ‘I am now able to use it as a supplement and in some grades as an entire curriculum. By conducting the research in my own classroom and with my students, it allowed me to see whether and why the teaching is effective. It also clarified my own assumption about the online interactive program that I have been using for the last five years.’

Luis Miguel Bermúdez Gutiérrez introduced a new, contextually relevant sex education curriculum in his school in one of the poorest areas in the city of Bogota in Colombia. Luis overcame political sensitivities by presenting objective data demonstrating the impact of the new curriculum on reducing adolescent pregnancy risk situations. Luis has obtained agreement from his leadership team to roll out the new sex education curriculum throughout the school.

2. Coalitions for change

Test and Learn builds teacher and practitioner agency to be the voice of practice, harnessing their capacity to be change agents who shape the way education is planned and delivered.

Miriam Mason-Sesay’s objective when running her RCTs was to demonstrate to her network of underappreciated and noncommittal teachers that their work has an impact on children’s learning and unleashes their agency and voice. Miriam’s network of teachers have taken on more responsibility, demonstrating ownership of learning outcomes for their students.

Test and Learn ‘gave me autonomy and leadership in the teaching process. It allowed me to validate my own teaching practices, as well as obtaining better information in the performance of my students.’

Marcela Henríquez, Santiago, Chile

3. Delivery architecture including school collaboration

Test and Learn requires and enables collaboration between teachers and practitioners at all levels of the system to deliver planned replications of protocols and support each other in determining what works, for whom and in what contexts.

The collaboration between the Varkey Teacher Ambassadors clearly demonstrates the power of teachers working together to promote positive change in the way practitioners perceive, use and generate evidence to inform their practice. The planned replications of protocols will further build on the evidence, giving an indication of how the same intervention might work differently in different contexts and with different cohorts.

‘Working with the best teachers in the world showed that regardless of where you are you can still promote positive change.’

Ken Silburn, Australia Varkey Teacher Ambassador

4. Data for collective accountability and improvement

Test and Learn helps to put data and evidence on ‘what works’ straight into the hands of practitioners. It helps schools to analyse and understand results and pupil attainment better, leading to enhanced practitioner accountability for student outcomes. It also empowers practitioners to challenge leaders and policymakers on accepted wisdom with real, robust data. Francis Jim Tuscano had not encountered this way of measuring the impact of interventions before Test and Learn. He believes that teachers have a responsibility to carry out their own research in their own classrooms to provide data on what works for student outcomes. This data, in Jim’s opinion, should influence the continuous learning and improvement cycle for teachers. Luis Miguel Bermúdez Gutiérrez and Souad Belcaid both used their research data to influence the leadership and vision of their schools.

‘In my context, problems are many. The knowledge is assisting me to better understand issues and deploy appropriate strategies.’

Andrews Nicholson, Malawi

‘My participation in this project gave me priceless tools to better my teaching methods, as well as taking better decisions. Thus [optimising] my students’ working time as well as their learning result(s).’

Marcela Henríquez, Santiago, Chile

5. Teachers and school leadership effectiveness

Test and Learn supports a constant cycle of research and development – paralleling what happens in medicine and healthcare – in line with the concept of continuous school improvement even where schools have reached and maintained high levels of performance.

Francis Jim Tuscano’s school has received the highest accreditation available from the Filipino government. Nonetheless, Jim strives for continuous improvement and found that ‘Test and Learn is a great way to share his ambition with teachers: ‘I lead a number of studies, in collaboration with other teachers in my school, that helped the school community to reflect on the worth and impact of different school programs, such as the use of technology tools in the learning.’

‘I was talking with some [colleagues] about this experience and with my Principle. They consider [that it] is important to train teachers in [these kind[s] of topics] because sometimes teachers forget... [what] we could discover through our students while they are participating in the classroom... the discoveries can help to improve... our work as teachers.’

Katerine Franco Cardenas, Medellin, Colombia

6. Evidence informed policy and learning

The culminating impact of Test and Learn on all education system accelerators is that it creates a cultural, behavioural and mindset shift towards a profession and a sector that is evidence-informed and continuously learning.

All teachers included in our case studies were excited about the potential of collecting small-scale, grassroots data which is synthesised to inform policy, at school and at government level. They all believe that Test and Learn could transform the status of teachers across the world.
How does it work?

Anyone can participate and learn to carry out a piece of experimental research with a little bit of training and support. You do not need a background in research or statistics.

The training is designed to scaffold participants’ understanding. We introduce three essential statistical concepts:

1. **Effect size** (the strength and direction of any change);
2. **Confidence intervals** (showing the generalisability of the effect); and
3. **Significance** (the probability that the result may be misleading).

The training also covers a wide range of potential research designs such as: between-subject design and within-subject design as well as more advanced design approaches (such as case-matching and different forms of randomisation).

The basic approach follows these steps:

- Teachers are given some pre-reading including high level overview of methods and some example studies.
- They attend day 1 of the training when they learn about the key statistical concepts and experimental research methods and begin the process of designing their own protocols.
- Teachers then run their trials with access to remote feedback from Education Development Trust throughout the process.
- They then come back together for day 2 to analyse their results and write them up.
- Teachers are given access to StatsWizard, the Excel tool that makes the approach accessible to anyone without needing to understand or use complicated data software tools. StatsWizard asks a series of questions about the protocol to determine which statistical test to use and guides teachers to the correct results.

- At the end of the project, all results from the teachers are synthesised in a meta-analysis. Meta-analysis is critical to the understanding of the overall effect of different studies. ‘When the effect or impact varies from one study to the next, meta-analysis can be used to identify features of the various studies that explain the variation.’ In addition, an overall effect and its significance can be calculated across all the studies that have been entered into the analysis.

More information on meta-analysis and the results from the project with the Varkey Teacher Ambassadors can be found in Appendix 1.

With the Varkey Foundation, we provided all follow up support remotely using email, telephone and Skype and worked with simultaneous translation during the process to support teachers with English as a foreign language.

**FIGURE 1**

The simplest form of randomised controlled trials (RCTs)

*Randomised controlled trials (RCTs)* are the ‘gold standard’ research method in many sciences. The method uses a control group to compare the results of an intervention that is given to another group. In education, the control ‘condition’ will usually be existing best practice – as in surgery, where you might compare a new operation to a current one. There would be no point not treating people or not teaching children at all. Randomisation helps to allocate participants to the conditions (control or intervention), removing researcher bias (where a researcher influences the outcome by allocating participants unequally between control and intervention groups). The conditions are tested (or ‘tried’) using an appropriate measure of the desirable outcome.

Other forms of randomised controlled trials

- **Between-participant** – pupils divided into two groups that each experience a different teaching approach (or condition).
- **Within-participant** – all pupils experience all approaches but in different orders (counterbalancing). It is usual to counterbalance the order in which things happen in this type of design. In order to balance out effects that might transfer from one condition to another (carry over or order effects).
- **Matched-pair (or case matched)** – similar pupils are paired and each member of the pair randomly allocated.
- **2x2 factorial design** – a design in which four things can be compared against each other, simultaneously.
Case studies
How to understand the results section in an individual teacher poster

Sample results
Gain scores were first calculated from pre- and post-test scores (Figure 2). A one-tailed independent samples t-test showed that the intervention had a significant \( p = .01 \) positive effect \( (d = 0.63) \) compared to the control condition \((CI (95\%) = 0.40 - 0.83)\).

Pre- and post-test design
In a pre- and post-test design one way to analyse the results is to calculate the change in each pupil’s score and use this to assess progress (other methods are possible).

Effect size
Effect size is the strength and direction of the change. This can be positive or negative. Different effect sizes are used for different data \( (d, r, np^2 \text{ etc}) \).

Confidence interval
The confidence interval \( (CI) \) is an estimate of the range of effect size that you might expect in 95 out of 100 replications of the study.

Probability
Different tests are used to calculate the probability that the result may be misleading \( (0.01 \text{ is a one in a hundred probability}) \). The test used depends on the design and types of data you have.

One-tailed means the researcher was predicting the result in their hypothesis, two-tailed that they were not \( (\text{this affects the final } p\text{-value}) \). Where three things are compared an extra test is first done on the overall pattern to deal with what is known as family-wise error.

Significant findings
A finding is significant if it crosses a certain threshold \((\text{alpha})\) – this is usually \( 0.05 \) but can be set smaller, if the researcher wants to or needs to be more rigorous \( (\text{e.g. when comparing more than three things at once}) \).

Researchers
Miriam Mason-Sesay

Country
Sierra Leone

Working title
Reading to a doll is less effective than reading alone or reading in pairs – findings from a small-scale randomised controlled trial
Reading to a doll is less effective than reading alone or reading in pairs – findings from a small-scale randomised controlled trial

In Sierra Leone, is there any impact on middle-primary age children’s reading fluency and their competence when reading to a doll or to a partner, as opposed to reading alone?

Miriam Mason – EducAid Rolal (Port Loko District) & EducAid Freetown (Western Area)

Research purpose

The purpose of the research was to investigate the impact of reading to a doll, or reading to a partner, versus reading alone. Specifically, it aimed to establish if there were any effects on children’s reading fluency and competence. The research took place over a four week period for 30 minutes every day.

EducAid works with young people who are usually the first generation in their family acquiring an education and who do not have books at home. Accessing reading material is difficult and children are learning to read in their second or third language. There is not much of a reading culture in the communities in which we work. EducAid wanted to investigate strategies that have been found useful in supporting the development of reading fluency and the enjoyment of reading. There is some evidence, in more westernised and first language English contexts that reading out loud and sometimes to a pet (le Roux, Swartz, & Swart, 2014) or doll, can strengthen children’s reading competence. EducAid wanted to investigate the possibility of using similar strategies within a resource constrained, ESL context. With a view to using these ideas across its teacher-training programme with teachers from over 70 primary schools, EducAid wanted to gather evidence of the usefulness of the different strategies in its own primary schools first.

The research design

A between-participant design with pre- and post-tests was used. The independent variable (IV) was operationally defined by creating 3 conditions (Figure 1):

- Independent Variable Level 1 (Control condition) reading alone
- Independent Variable Level 2 (Experimental condition A) reading to a doll or soft toy
- Independent Variable Level 3 (Experimental condition B) reading to a younger child

Procedures

A literacy test and a reading fluency test were administered to all participating children in the week before, and during the week after the four-week intervention period. Each class had a group of children that had been randomly allocated to either read alone, read to a doll, or read to a younger child. The children were allocated to the read to a doll group made their dolls from basic materials and named them. Each afternoon for a four-week period (i.e. 20 school days), during the last half an hour of the day before their end of day registration, the groups undertook their reading. Reading books were selected and it was ensured that the mixture of reading materials in each class were the same, or nearly the same.

Materials (and apparatus)

ASER test and a basic literacy test were used for pre- and post-test assessments. Both tests assess development along a literacy development scale going from sounds and letters to words to paragraphs to stories.

- Boxes of assorted children’s books.
- Standardised instructions to teachers for each of the groups.
- Materials to make dolls: plastic bags, sticks and markers etc.

Limitations

While everything was done to standardise the procedures and limit variation between the schools, inevitably, with different teachers, class sizes and locations there were extraneous variables that were hard to eliminate completely (e.g. fidelity to the 30 minutes reading time, classroom environment being conducive to reading, the types of books that were available). In addition, the children made their own dolls which meant that there were differences between dolls. This may have resulted in different engagement and reactions. The sample size was small so the findings can only be understood as indicative of what might be found in a larger study.

Research participants

100 children in 4 Class (4 classes from 4 primary schools (Table 1)) were randomly allocated to either the control or doll, or partner reading groups. For each class, the names of each girl and then each boy were pulled out of a hat in order: control, doll, younger child, until all children had been allocated (so that, for each class there were groups of approximately the same size for each IV level). The girls and boys were selected in equal numbers, as far as possible, to each category in order to minimise any bias due to gender.

Dependent variables were used to assess reading fluency and competence (pre- and post-intervention):

- DV1 Basic literacy test, which while testing at the sound, word and paragraph level, also provides a standardised aggregate score that can be compared easily pre- and post-intervention.
- DV2 ASER style test which provides categorical results. Children are classified as being at one of three proficiency levels: letter/sound; word; paragraph or story.

Methods

Participants, sample size and randomisation

100 children in 4 Class (4 classes from 4 primary schools (Table 1)) were randomly allocated to either the control doll, or partner reading groups. For each class, the names of each girl and then each boy were pulled out of a hat in order: control, doll, younger child, until all children had been allocated (so that, for each class there were groups of approximately the same size for each IV level). The girls and boys were selected in equal numbers, as far as possible, to each category in order to minimise any bias due to gender.

A literacy test and a reading fluency test were administered to all participating children in the week before, and during the week after the four-week intervention period. Each class had a group of children that had been randomly allocated to either read alone, read to a doll or read to a younger child. The children allocated to the read to a doll group made their dolls from basic materials and named them. Each afternoon for a four-week period (i.e. 20 school days), during the last half an hour of the day before their end of day registration, the groups undertook their reading. Reading books were selected and it was ensured that the mixture of reading materials in each class were the same, or nearly the same.

Materials (and apparatus)

ASER tests and a basic literacy test were used for pre- and post-test assessments. Both tests assess development along a literacy development scale going from sounds and letters to words to paragraphs to stories.

- Boxes of assorted children’s books.
- Standardised instructions to teachers for each of the groups.
- Materials to make dolls: plastic bags, sticks and markers etc.

Limitations

While everything was done to standardise the procedures and limit variation between the schools, inevitably, with different teachers, class sizes and locations there were extraneous variables that were hard to eliminate completely (e.g. fidelity to the 30 minutes reading time, classroom environment being conducive to reading, the types of books that were available). In addition, the children made their own dolls which meant that there were differences between dolls. This may have resulted in different engagement and reactions. The sample size was small so the findings can only be understood as indicative of what might be found in a larger study.

After the controlled trial, focus group discussions were undertaken to understand more about how the students felt about the different ways of working and to more fully understand the differences identified.

Results

An initial two-tailed Kruskal-Wallis ANOVA indicated that the effect (w = 0.36) across all three conditions was significant suggesting that it was unlikely to have been caused by family-wise error (p = .009). This was followed by separate one-tailed Mann-Whitney U tests to compare the different conditions with each other (Table 2).

Conclusions and recommendations for future research

Existing practice (reading alone) was notably better than reading to a doll (Intervention A) (a significant moderately large effect size difference). Reading in pairs (Intervention B) was also significantly more effective than reading to a doll. These effect sizes were large enough for the results to be significant, despite the small sample size and a Bonferroni adjusted alpha. A small non-significant positive effect was found for reading in pairs compared to reading to a doll. From this small sample trial, we found that reading to a partner results in better progress in literacy than reading alone. It was also better than reading to a doll. Further replication will be needed to establish the findings in this study and increase sample sizes. However, if these findings were to be replicated in a range of different contexts with larger sample sizes, the implications could be that teachers should regularly provide opportunities for children to read in pairs.
About Miriam

Miriam Mason-Sesay is Country Director of EducAid Sierra Leone, which owns and manages three schools and supports a network of schools, delivering teacher professional development and school improvement programming. Miriam lives in the North West of Sierra Leone and dedicates her life to working with teachers to improve their practice. Miriam was driven to participate in the Test and Learn programme as a way of introducing the concept of evidence-based practice in a context where the profession is yet unfamiliar with the idea that there may be more than one way to teach!

Context and challenge

Sierra Leone’s education system was left devastated by the 11-year war as educational institutions were targeted by the Revolutionary United Front rebel army, resulting in the destruction of many schools and others abandoned (Turrent, 2012). As a result of post-war re-staffing of rural schools with unqualified teachers, the teaching force in Sierra Leone is seen as an unattractive and underappreciated profession.

Miriam’s research

Miriam and her network of teachers ran three different research experiments. For all studies, she believed that her key to running a successful experiment was to keep things simple. Simple interventions and measures (or dependent variables) were more likely to be well implemented and accurately measured. This, in turn would be more likely to spark curiosity and build the intrinsic motivation of teachers to engage with research and evidence-informed teaching materials to improve their practice.

The impact of Miriam’s research

Reading alone was more effective than reading to a doll, however reading in pairs was both more effective than reading to a doll and reading alone. By far more interesting was the impact on participating teachers who realised through the process that there are different ways of teaching which may lead to different outcomes. Teachers began taking greater responsibility for their pupils’ outcomes such as establishing reading clubs for children to read to each other.

'I thought that pairing children up would make a difference [to their reading proficiency], and I thought that it would be simple to implement if the trial demonstrated a positive effect.'

Miriam discovered a whole new way to measure the impact of her work as a teacher-trainer and middle-tier professional within the Sierra Leonean education system. She is currently planning to incorporate experimental research and quasi-experimental research as an evaluation tool to measure the impact of her school improvement interventions.

The power of teacher-led research

As an education organisation seeking to effect change at scale, we were most excited by Miriam’s account of the way in which the research positively influenced teachers’ intrinsic motivation and enthusiasm for the profession. Miriam believes that Test and Learn has the potential to dramatically change the way education is delivered in Sierra Leone. By building a sense of agency and purpose within the profession, Test and Learn could improve teachers’ engagement with teaching materials and research, helping them de-bunk myths and stereotypes (particularly around gender and expectations for boys and girls) and move beyond rote learning as the only way to teach. Miriam believes...
that teacher-led research has the potential to transform the status of teachers in Sierra Leone:

‘As teachers are currently disregarded in Sierra Leone, professionally engaged teachers that are doing a professional job will start to get noticed.’

Miriam’s experience illustrates the potential of Test and Learn to influence culture and mindsets which have far reaching positive effects on the quality of teaching and learning.

Miriam’s medium-term vision for Sierra Leone is for all teachers in EducAid’s network to participate in teacher-led research every year and expand EducAid’s network to improve teaching standards in rural communities. In a few years, Miriam envisions a nationwide annual teacher-led research conference, where teachers can present their findings to peers, the government, and the national press.

‘Motivation is only truly intrinsic when it is driven by feelings of enjoyment or fulfilment. For example, a love of the process of teaching. In teachers, motivation is a key factor in classroom effectiveness and student motivation, which both enable beneficial outcomes for students.’

Han and Yin, 2016

**Miriam’s top tips for teacher-led research:**

**1. Think about the things that might go wrong!**
What could contaminate the findings? Can you reverse the effects of a treatment in a within-participant design? Is there a school trip in the middle of the treatment window?

**2. Make sure that you understand the experiment to ensure clarity.**
What are you testing? What is the independent variable? Is the IV being measured under equal conditions? For example: what difference will applying the ‘treatment’ have on a Tuesday morning, compared to a Friday afternoon? Could extra-curricula activities be influencing the result of the trial?

**3. Make sure that your measures [dependent variables] are appropriate!**
Is the dependent variable linked to the independent variable? Is it measuring what you think it is measuring? What proxies can you put in place?

---

**Researcher**
Luis Miguel Bermúdez Gutiérrez

**Country**
Colombia

**Working title**
Effectiveness of the sex education curriculum in decreasing adolescent pregnancy risk situations
Effectiveness of the sex education curriculum in decreasing adolescent pregnancy risk situations

Luis Miguel Bermúdez Gutiérrez, Colegio Distital Gerardo Paredes

Research purpose

In a study carried out by UNESCO (2017) it is suggested that gender violence has a direct relationship with adolescent pregnancy, and in another study carried out by the same entity (UNESCO 2018), it is affirmed that a comprehensive sex education decreases exposure to the risk of this situation in adolescents. However, this correlation occurs differently depending on the socio-cultural context in which it occurs. For example, while in sub-Saharan Africa and Southeast Asia early pregnancy occurs within marriage, in Latin American countries and the Caribbean, pregnancy in adolescents between 12 and 19 years manifests mostly outside the marriage. Therefore, the success of a curricular plan in sex education will depend on how gender-based violence in a particular cultural and school context is identified, interpreted and transformed.

Thus, the purpose of this research was to investigate through a randomized controlled trial, the effectiveness of the sex education curriculum based on the prevention of gender violence and the promotion of human, sexual and reproductive rights, to reduce exposure to the risk of early pregnancy in 10th grade students.

The research design

A between-participant design was used with a pre-test and a post-test (Figure 1). The independent variable (IV) ‘sex education curriculum based on the prevention of gender violence and the promotion of human, sexual and reproductive rights’ was defined by creating two conditions.

- IV Level 1 (Control): Traditional sex education classes based on abstinence and biology.
- IV Level 2 (Intervention): Comprehensive sex education based on prevention of gender violence and the promotion of human, sexual and reproductive rights.

Dependent Variables (DVs)

- DV1 Standard test of 50 questions that measures levels of exposure to early pregnancy.
- DV2 Percentage of students accessing methods of fertility regulation, student counselling for gender violence and sexual and reproductive health.

Methods

Participants, sample size and randomisation

The participants were 164 10th grade students from the Gerardo Paredes School in Bogotá Colombia, who were randomly assigned to control and intervention groups through the simple randomization process. The characteristics of the students are varied, including 6 learners with special educational needs.

Procedures

At the beginning of the learning period, all the students from control and intervention groups completed a pre-test that measured the exposure to risk situations of early pregnancy. Then 10 didactic sequences or classes were applied to the intervention group using comprehensive sex education based on the prevention of gender violence and the promotion of human, sexual and reproductive rights. Meanwhile, the control group received 10 traditional classes based on abstinence and biology.

Afterwards, a 50-question post-test was applied, which again measured the exposure to situations of pregnancy risk. In addition, the percentage of students who autonomously requested counselling services on sexual and reproductive health at the school, such as access to methods of fertility regulation, was quantified.

Materials (and apparatus)

- Pre-test (50 questions).
- Post-test (10 questions).
- Register form counselling on sexual and reproductive health.
- Register form school assistance to victims of violence based on gender.
- General plan of subject.
- Didactic units.

Results

Gain scores were first calculated using pre- and post-test results (Figure 2).

A one-tailed Mann-Whitney U test indicated that the intervention group had a significant (p = .001) large positive effect (r = 0.70) (CI (95%) = 0.28 – 1.12) on reducing the exposure to early pregnancy situations among students.

In the pre-test, students from control and intervention groups were asked for the number of times they had attended counselling services on sexual and reproductive health at the school. Then the number of students of each condition that requested these services in the 10 classes of the experiment was measured (Figure 3).

A chi-squared test of independence indicated that the intervention was associated with a significant moderate positive effect (φ = 0.37) compared to the control (X² = 7.25, p = 0.0499).

Conclusions and recommendations for future research

This research suggests that the implementation of a sex education curriculum (based on the prevention of gender-based violence (GBV), the promotion and exercise of human, sexual and reproductive rights, and the recognition of the socioaffective dimension of sexuality) improves a student’s ability to recognize the risk situations of early pregnancy.

Future studies could replicate the design with larger sample sizes in other schools and adapting the curriculum to the specific context.
Luis Miguel Bermúdez Gutiérrez is a secondary school teacher in Bogotá, Colombia. Luis was shortlisted in the top 10 finalists for the Global Teacher Prize in 2018 as a result of his dedication to addressing early pregnancy, sexual and gender-based violence, and homophobia in his classroom.

Context and challenge
As in many countries, sex education is a highly controversial issue in Colombia. Luis teaches in a school located in one of the poorest areas of Bogotá, which has been faced with early pregnancies, bullying, and violence around gender and sexuality. He believes that addressing gender-based violence at home starts in school.

Research and investigation doesn’t have to be something that is far away and requires being a scientist or a researcher, but it can be something more approachable. Research can be used to measure daily life facts and can be shared in class on a computer. Thanks to this some of my students are less afraid of investigation and research, and have started to show interest in research."

Luis’ research
In 2014, Luis introduced a new sex education curriculum in his school that was informed by human rights, pregnancy prevention guidelines and sexual citizenship. As a result, teenage pregnancies in his school have dramatically decreased. He sought to challenge gender stereotypes, including violence in masculinity and submission in femininity.

He believed that objective data on the impact of the new curriculum could be collected by combining experimental research and education ethnography. This would help him to provide students with what was needed for a successful education by preventing adolescent pregnancy and reducing gender-based violence.

The impact of Luis’ research
Luis’ sex education curriculum had a significant positive effect on decreasing adolescent pregnancy risk situations. Anecdotally, Luis found that both male and female students were more open about their sexuality, their body, and puberty. As a result, he believes that gender equality is improving by normalising sexuality in the classroom.

Luis has become more reflective on his pedagogical practice, which he believes has improved his performance at school. For example, his practice is more systematic, which allows his students to better follow, understand and engage with content.

Luis has been contacted by universities, NGOs and medical centres to understand the practicalities of implementing the new sex education curriculum and the impact on adolescent sexual health. The project and research highlighted the multisectoral ‘need to work together’ to ensure the best outcomes for young people.

For Luis, the most interesting aspect of the research was how the outcome addressed a series of beliefs on sex education, including his own. Test and Learn provided Luis the opportunity to create a shared value for teaching content that improves student outcomes. Luis’ school leaders have confirmed that sex and gender education will be integrated into the school’s curriculum. They have approved for each student to attend sex education classes for two hours per week.
The power of teacher-led research

Luis’ project has helped to transform the curriculum, and it has started to change the culture within his school and in his community. The curriculum for other subjects is being updated to ensure that the content is contextually relevant for their students, to enhance their prospects.

He is confident that Test and Learn presents the opportunity to deliver robust data to demonstrate results and create more trust in the system on what is being done, and therefore improve student outcomes.

Luis envisions a culture of investigation, evaluation and improvement across the Colombian teaching force. If teachers are encouraged to participate in and carry out their own research, there is potential to empower them to engage critically with their pedagogy.

As Luis has demonstrated by attracting attention from health and medicine professionals, teacher-led research can create respect for teachers from the wider society.

Luis’ top tips for teacher-led research:

1. Don’t be overwhelmed by the work involved and be patient
   You may encounter an increase in workload to ensure that teachers understand the intervention, and you might feel at one point that you won’t be able to produce beneficial data. However, in the long-term, the time that you invested will be worth it as the process will teach you a lot, and it will make your pedagogic practice a lot more efficient. It’s worth it, and you might have the opportunity to share your findings, and give speeches in national and international events! A teacher is more likely to follow something that another teacher has created, rather than from another professional that isn’t aware of teaching practice, so you have the opportunity to help others.

2. Share your findings with your students!
   They will be happy to see the results of your investigation, and will use the outcome to improve their personal lives.

Researcher
Kenneth Silburn

Country
Australia

Working title
A study on the effect of a multi school student engagement program (iSTEM) on student attitudes towards science
A study on the effect of a multi school student engagement program (iSTEM) on student attitudes towards science

Kenneth Silburn

Research purpose

The iSTEM program provides enrichment and engagement activities for High School Science Students in the South West of Sydney. The aim is to ultimately increase the educational outcomes of students from socioeconomically disadvantaged areas and to increase student awareness in the opportunities available to move into STEM related careers.

The program has received numerous awards and anecdotal evidence suggests that the program is successful. Students participating in the program were invited to attend activities including visits to museums, science organisations, universities and field trips.

This study sought to evaluate if participation in the program increases students overall attitudes to science.

The research design

A mixed-method was used. This involved: a) a student semi-structured questionnaire evaluating the effectiveness of the program; and, b) non-randomised controlled trial with a post-test only, using the Test of Science Related Attitudes (TOSRA) to measure students' attitudes toward science in seven categories: Social Implications of Science, Normality of Scientists, Attitude toward Scientific Inquiry, Adoption of Scientific Attitudes, Enjoyment of Science Lessons, Leisure Interest in Science, and Career Interest in Science with the hypothesis that the program would increase positive attitudes to science compared to a baseline of TOSRA scores.

Dependent Variables (DVsf)

- DV1 Student perceived effect
- DV2 TOSRA score

In addition, student qualitative unstructured responses were also analysis and summarised.

Methods

Participants and sample size

35 students (23 females; 12 males) intervention (participation in the program and normal population TOSRA scores) were compared to a control group of 350 pupils (approx. equal males and females). Randomisation was not possible.

Procedures

Students who had completed the program were contacted by email or phone and invited to participate in the study. An online questionnaire was made using google forms to facilitate data collection. There was an added complexity in locating the students as many had matriculated from school.

Materials (and apparatus)

Online survey using google forms.

Limitations

As no pre-test was completed for the research it may be that the participants already possessed positive attitudes towards science due to the fact that they self-nominated to participate in the program. However, participants identified an increase in their awareness and enjoyment of science as a result of the program.

Results

A) Qualitative findings

Overwhelmingly participants reported that participation in the iSTEM program had been a positive influence on their time at school. Furthermore, that it had given them the opportunity to mix with other students who enjoyed science and had given them an insight into careers in science and engineering.

A clear point reported by all participants was that they enjoyed science at school more after participating in the program. Responses to open ended questions in the questionnaire were also very positive.

New friends, wider experience in the science field

"Participating in the iSTEM program benefited my growth mindset. It helped me to mature and understand the significance of STEM careers, especially for the future. My interest for a career in STEM has grown drastically and I genuinely consider it as a career path."

Confidence and a greater range of experience

"The experience, extra knowledge I gained, being with peers with similar interests and the friendships gained. I gained a broader understanding of many iSTEM topics, was given a great chance to socialize with like-minded peers and greatly improved my teamwork skills."

B) Quantitative analysis of TOSRA results (Single sample t-test)

<table>
<thead>
<tr>
<th></th>
<th>Effect size (d)</th>
<th>CI (95%)</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Social implications of science</td>
<td>1.33</td>
<td>0.99 - 1.77</td>
<td>&lt; .0005</td>
</tr>
<tr>
<td>Normality of scientists</td>
<td>0.67</td>
<td>0.55 - 0.79</td>
<td>&lt; .0005</td>
</tr>
<tr>
<td>Attitude toward scientific inquiry</td>
<td>0.62</td>
<td>0.50 - 0.74</td>
<td>&lt; .0082</td>
</tr>
<tr>
<td>Adoption of scientific attitudes</td>
<td>1.00</td>
<td>0.87 - 1.13</td>
<td>&lt; .0005</td>
</tr>
<tr>
<td>Enjoyment of science lessons</td>
<td>1.33</td>
<td>1.18 - 1.48</td>
<td>&lt; .0005</td>
</tr>
<tr>
<td>Leisure interest in science</td>
<td>1.47</td>
<td>1.32 - 1.63</td>
<td>&lt; .0005</td>
</tr>
<tr>
<td>Career interest in science</td>
<td>1.90</td>
<td>1.72 - 2.09</td>
<td>&lt; .0005</td>
</tr>
</tbody>
</table>

Conclusions and recommendations for future research

The study indicates that participation in the iSTEM program had a significant positive effect on the attitudes of students to science and science related careers - demonstrated by Cohen’s effect size values greater than 1 for five of the categories investigated as part of the TOSRA questionnaire. The highest being student responses to career choices in Science (effect size = 1.90) and leisure interests in Science (effect size = 1.47). This is further supported by responses to open ended questions in the questionnaire and to positive responses to the seven-point Likert scale scoring greater than 6, highlighting positive components and attitudes towards the program.

Future research may explore the effect of the program by gender. Initial analysis indicates that the effect size was greater for boys in all areas of the TOSRA analysis with the exception of questions relating to the normality of scientists of which the effect size for girls (0.71) was slightly more than for boys (0.56).
About Ken

Kenneth Silburn is a high-school science teacher located south west of Sydney. In Australia, teachers tend to be highly regarded as professionals. Ken was shortlisted in the top 10 Global Teacher Prize winners in 2017 due to his vision and commitment to providing quality teaching and learning for science students.

Context and challenge

Although he has recently moved schools, the school at which Ken carried out his research was his previous school, a medium-sized high school with children from deprived socio-economic backgrounds.

Back in 2009/2010, Ken received a scholarship to attend an educators’ course in a space camp in Huntsville in the US. The space camp also ran programmes for students, which Ken thought would benefit his students and boost their love for science. Ken and his colleagues took 42 students to Huntsville.

Given the socio-economic status of the students in Ken’s locality, he decided to make science learning more accessible to all children through the programme iSTEM (invigorate Science, Technology, Engineering and Maths), which Ken founded and developed. Rather than continued...
trips to space camp. Ken arranged visits on the weekend to museums and universities to ensure that all students have the opportunity to learn and love science.

Ken’s research
Before Test and Learn, Ken used to evaluate the iSTEM programme after students had completed it. He had no control group to know whether or not the programme was making a difference.

The impact of Ken’s research
The impact of iSTEM on students’ confidence and interest in science is incredible. Ken expected to see an increased effect size for the impact on the pupils’ leisure interest and career interest in science, given that these students were already keen science students that had given up their free time to participate in the programme. However, to have found an effect size of 1.47 and 1.9 respectively was ‘absolutely amazing’.

Ken has become more systematic in collecting data now that he has discovered how powerful it is to have robust, reliable results on the impact of interventions. For example, before he starts a new activity (for example, watching a film), he will ask his students to answer questions on a sheet. After the activity, he asks his students to answer the same set of questions to compare the difference. As a result, the students understand their own learning better mirroring what we know about the power of metacognition in teaching and learning.

The power of teacher-led research
Teacher-led research has the potential to improve the quality of teaching and learning because it provides teachers with a focus on what they do in the classroom, which is real-time and personal to them.

‘If they can see [the impact of what they do] it makes a big difference. Being able to present that information and have a dialogue with other teachers about what they’ve done, [that has] to be positive.’

Ken believes that teacher-led research should be expected of all teachers in Australia. This could be encouraged by including the research outcomes as mandatory criteria for awards such as the National Teaching Award to demonstrate the effectiveness of teaching.

Ken’s top tips for teacher-led research:
Talk to somebody else about what it is you want to test, and then decide how you will test it
It’s worthwhile putting in the effort and getting the feedback at the start so that the study produces reliable data. Ken spoke with former science teachers who are now working in academia to help him ensure that his study was appropriate to collect the right data.

Researcher
Souad Belcaid

Country
USA

Working title
Using MysteryScience.com can improve elementary students’ progress in science
Using MysteryScience.com can improve elementary students’ progress in science

Souad Belcaid

Research purpose

Despite the emphasis on improving science and math, American students still lag globally behind their peers. The 2015 PISA results placed the U.S. 38th out of 71 countries in math and 24th in science (Louis Sentario, 2017). The purpose of this study was to ascertain whether using MysteryScience.com would improve the students’ understanding of new scientific concepts at the elementary level. Mystery Science is a new approach to elementary science education. It consists of open-and-go lessons that inspire students to be more engaged and motivated in learning new science concepts. Lessons are aligned with Common Core and Next Generation Science Standards and designed to supplement the existing curriculum.

The research design

A between-participant design was used with a pre-test and a post-test (Figure 1). The independent variable (IV) (MysteryScience.com) was defined operationally by creating two conditions:

• IV Level 1 (Control condition) normal science lesson using a textbook
• IV Level 2 (Experimental condition A) the same science lesson using MysteryScience.com only

Methods

Participants, sample size and randomisation
Participants were all students of Fessenden School. The Fessenden School (PRE-K to Grade 9) is an independent day and boarding school for boys. There were 52 participants in total: 16 in Grade 3A and 16 in Grade 3C. The existing classes were already stratified for equal numbers of boys and abilities. These whole classes were randomly allocated to the control or the intervention.

Procedures
Both classes were taught one science lesson on the formation of clouds. This subject was chosen because it is not part of the current curriculum and so could be taught as a stand-alone lesson. The participants were taught by the same teacher and the lesson took place in their own classroom. The control class was taught through a traditional method using the textbook, while the intervention group was taught using MysteryScience.com. Both lessons took 45 minutes each.

Materials (and apparatus)
The teacher was supplied with the lesson plan and worksheets for the control group. For the intervention group the same teacher was supplied with a computer, head projector and the MysteryScience open-and-go lesson. For both the control and intervention group, the teacher was supplied with the same 10-question pre- and post-test comprehension question sheet for the children (based on the lesson content) and an answer sheet for the TA to mark from. The classroom layout was as it normally is in lessons.

Limitations
The trial was limited by its small sample size and same gender grouping, since only boys were used in this trial. A larger sample and inclusion of both genders could be useful in subsequent trials. In addition, some extraneous variables were not controlled for (e.g. timing of lessons, prior knowledge, and the amount of time given to answer each question), which may have affected the results.

Results

DV1: Gain scores were first calculated from the pre- and post-test results (Figure 3). A one-tailed Mann-Whitney U test indicated that the intervention (online interactive learning site) had a significant (p = .001) very large positive effect (r = 0.73, CI (95%) = 0.61 – 0.81) on students’ progress compared to the control condition (normal classroom practice) [d = 2.23].

FIGURE 3: PRE- AND POST-TEST RESULTS FOR CONTROL AND INTERVENTION

Conclusions and recommendations for future research

Use of the online MysteryScience interactive site (intervention group) had a large positive effect. Randomisation resulted in lower attainment in the intervention group at pre-test, the intervention not only reversed this lower attainment but also ensured that students attained more than the control group at post-test.

In addition, a survey on how often they use the interactive online MysteryScience site in their classroom was given to teachers across the country. Out of 56 teachers, 76.3% use it all the time, 23.2% use it sometimes and 1.8% once in a while.

The study had a small sample size and therefore will need replication with larger number of participants. Future studies may also wish to replicate the design across grades, and with students from different socioeconomic backgrounds.
Souad Belcaid has taught in schools in Morocco, Egypt and the United States. Souad currently teaches in a private boys’ school in West Newton, Massachusetts, which caters for K-9.

Context and challenge
Souad’s school doesn’t experience the same challenges as public schools in the United States, most notably around budget cuts and frequent changes to the curriculum. However, as a private, prestigious school, the teachers face very high expectations around high quality, innovative and creative teaching from their international body of parents and the school’s governing body.

Souad’s research
Souad had been using an interactive, online site called Mystery Science for five years, which provides online resources for elementary teachers and school leaders that support the K-5 science curriculum. In her experience as a teacher, Souad had come across various initiatives, but had found that Mystery Science was the best at engaging and motivating her students, allowing pupils to reach conclusions on their own.

However, Souad’s supervisor and head of department were concerned that the teaching would not meet their high expectations by using a website developed by someone outside of the school. Souad therefore sought to demonstrate the benefit of using the programme as a tool to supplement her own teaching and improve student outcomes.

The impact of Souad’s research
Souad found that the use of Mystery Science improved elementary students’ understanding of new scientific concepts. She is now able to explain to her supervisor how and why she is using this approach to teaching and has the data to back this up. Her colleagues who were once sceptical about the use of the website are now enthusiastic.

By far the most significant impact however was Souad’s realisation that she could be the master of her own research. She feels more skilful as a teacher and can demonstrate that she has done her job well.

The power of teacher-led research
Without doubt, the power of teacher-led research for Souad is the potential to transform and energise the teaching force in the United States and beyond. Souad strongly believes that ‘it’s about time that teachers have control of their own teaching’ by conducting their own research in their classrooms.

‘Teachers will be more confident and will have something tangible to demonstrate that their teaching is doing good to their students. We know our students the best, not a researcher outside of the classroom... Doing research in the classroom makes much more sense.’

‘Teachers in the US don’t currently command the respect that they deserve as they are not considered highly professional. When practitioner-researchers talk in the medical field, people listen. When teachers talk about their practice, people don’t listen. If teachers use and collect research like medical doctors, people will listen. There will be more respect for teachers.’

‘For teachers, we’ve always had people telling us what to do and how to do it. Research is always communicated to teachers, not done by teachers to measure the effect of their work.’
Souad’s top tips for teacher-led research:

1. Just do it!
The results can help you prove your impact to your peers, seniors and wider community if the administration, curriculum or policy changes and it can be empowering.

2. Be in control of your classroom and your teaching practice
It’s a skill that is needed to develop our voice as teachers.

‘To be honest with you, the first time I heard of the research, I thought, ‘no, not for us, not for teachers’. Research is done by doctors and PhD students and other researchers, not teachers. But then, when Richard presented this idea, I was like, ‘Oh, I can do this?’ It’s fabulous and I was the first to apply! I wanted to do something that shows that my teaching is worthwhile and it’s not something that is imposed on me or done by someone who has never been in my classroom or in my school.’

Researcher
Francis Jim Tuscano

Country
The Philippines

Working title
The use of design-thinking process improved student performance in understanding a Science unit in Grade 4
The use of design-thinking process improves student performance in understanding a Science unit in Grade 4

Francis Jim Tuscano

Research purpose

Research literature on design-thinking suggests that the approach, a solution-centered framework, helps in fostering soft skills such as creativity and critical-thinking needed for solving difficult and complex problems. Most of the research on design-thinking is qualitative and often centers on the qualities of the designers or of the design process. Experimental evidence on how design-thinking affects student achievement in specific knowledge domains or student learning outcomes is lacking. This is an important area to explore using a randomised controlled trial because a number of educators have started to use design-thinking in various learning experiences without the needed experimental evidence. Hence, specifically, the study sought to evaluate if the design-thinking approach can improve student performance in understanding a Science unit in Grade 4 students.

The research design

A between-participant design with a pre- and post-test was used. The independent variable (IV) ‘design-thinking approach’ was operationally defined by creating two conditions:

- Level 1 (Control) delivery of lesson per normal practice
- Level 2 (Intervention) use of design-thinking approach

Dependent Variables (DVs)

Pre- and post-test scores on assessments related to the content of the unit (student performance).

Methods

Participants, sample size and randomisation

Ten Grade 4 Science classes were randomly allocated to the control or intervention group. 342 male participants took part in the study. Whole class simple randomization was used, ensuring an equal number of Grade 4 classes in each condition.

Procedures

The Science unit on ‘Healthy Diet’ was selected. Xavier School - San Juan employs the Singaporean Science curriculum, a spiralled curriculum, which means that the students have encountered the unit in the previous grade level but will be presented with it in the current year in a more complex manner.

The Science teachers were trained to design and facilitate unit plans that integrate the design-thinking approach.

A pre-test on the unit was done after the randomization process and before the start of the trial. Within the same week or schedule, both groups or conditions went through the same unit. The teachers handling the control group delivered the lesson using the normal unit plan while the teachers of the intervention group used design-thinking approach in the learning experiences. At the end of the unit, both control and intervention groups were given the same post-test.

Materials (and apparatus)

Identical lesson plans on the ‘Healthy Diet’ unit planned by the Science Grade 4 teacher team. Students exposed to the intervention made use of the design-thinking activity worksheet.

The pre- and post-tests designed in-house were based on the expected Grade 4 Science competencies following the Singaporean Science curriculum as integrated and modified for Xavier School’s use.

Hypothesis (one-tailed)

The use of a design-thinking approach improves student performance in understanding a Science unit in Grade 4 students.

Results

Gain scores were first calculated from pre- and post-test results (Figure 2). A one-tailed independent samples t-test indicated that the intervention had a significant (p = .001) moderate positive effect on progress (d = 0.38, CI (95%) = 0.03 - 0.73) compared to the control condition.

Conclusions and recommendations for future research

The research was effective in showing that students in the intervention group significantly experienced higher progress in their tests than those in the control group. The results suggested that the intervention group benefited from the use of design-thinking in teaching and learning, more specifically in applying learned ideas and skills to concrete and realistic problems. Some of the benefits may have included the presence of more student-centered activities, such as empathizing with the persons in the real-life problems being solved and the opportunity to prototype solutions and receive feedback. The intervention produced a moderate positive effect on student progress.

This study may be one of the few experimental studies investigating the effect of design-thinking on student outcomes. Hence, the recommendation is for more experimental research that would investigate the effect of the use of design-thinking, not just in Science but in other academic subjects. 
About Jim

Francis Jim Tuscano (Jim) is Education Technology coordinator and coach in Xavier School-San Juan, a progressive school in the Philippines. Jim’s school has received the highest accreditation available for a K-12 school, from the Philippine education accreditation organisation. This means that the school has the freedom to implement its own teaching policies and innovations, informed by the experiences of its teachers.

Context and challenge

Before Jim’s participation in Test and Learn, his school had already been working on numerous ways to measure the impact of different teaching initiatives on their pupils’ learning. However, these had been qualitative studies.

The picture in the rest of the Philippines is a different story. Jim explained that teacher-led research is very rare, even though a number of teachers have completed post-graduate degrees and doctorates whilst teaching.

Jim’s research

Participation in Test and Learn couldn’t have come at a better time for Jim and his school because it provides a tool to collect ‘hard data’ and measure impact on students. Jim was inspired by the world of business to embed design thinking through project-based learning in the classroom. He had also identified a gap in research on the effectiveness of this approach which drove him to investigate.

Jim believed that the solutions-focused mindset that is encouraged through the use of design thinking process would improve student outcomes. The results of Jim’s trial supported this belief.

The impact of Jim’s research

Jim was empowered by collecting his own research on an initiative that he introduced to his school. As a result, he feels more confident in his own pedagogy and ability to identify and address problems in the classroom.

Jim’s school leadership have embraced this method of evaluation to ensure the best possible outcomes for their pupils. As a result, Jim has supported 15 of his colleagues to run a further four randomised controlled trials within his school since completing his own on design thinking.

‘Being part of a global team, the first time that there were simultaneous teacher-led RCTs around the world... was an amazing experience to be part of.’
Jim believes that teachers in the Philippines are not achieving their full potential as educators, if they do not gather data to help inform practice and changes in their classroom. He considered himself a member of the global education research community by contributing to and critically engaging with the research base.

Jim suspects that teachers in the Philippines and across the world imagine having to write lengthy university-style dissertations when thinking about conducting teacher-led, classroom based research, which he believes puts them off. Test and Learn provides teachers with a better way to look at research, seeing it as part of a teacher’s toolkit, collecting data on their craft, sharing the effect of the intervention, learning, and continuously improving.

Jim was excited about the potential of the large-scale impact of Test and Learn and teacher-led research to inform education policy:

‘If we want data from the classroom, then we must teach teachers to gather this in an effective way. One way is through this approach. It gives teachers the opportunity to collect real-time data from the classroom and provide a picture to the leadership and academics of what is happening. What kind of education system do we want? Something that’s extremely driven and animated by teachers at the grassroots.’

What’s the prize?

Jim’s top tips for teacher-led research:

1. Have an open mind!
   Before venturing into your research, be clear on what you are looking into and measuring. Be ready before you conduct your research, and be open to feedback and discussion on the results of your study. It will help the next one!

2. Persevere! Don’t give up
   The research study is easy to prepare, but it can get messy (we are talking about the classroom with real people and real students). Have grit and patience when doing research.
Collaborative, multiple planned replication by teachers could not only resolve many of the challenges that remain in education reform but could, in addition, help to ensure the buy-in of the teaching profession.

Practitioner-led experimental research has greater potential to control extraneous variables (variations in implementation) compared to larger scale trials – which face even greater challenges, in this regard, compared to laboratory studies. They also have the potential for higher levels of mundane realism (‘everyday-ness’) – reducing the way that participants may react to being in the trial. Of course, they themselves face the challenge that at an individual trial level they have a smaller sample size and therefore it is harder to generalise. The same is true for areas such as surgery. The answer is to conduct multiple planned replications building large samples with specific meta-analysis.

Practitioner-led experimental research offers up the possibility of breaking complex multifaceted pedagogical interventions down into their component parts; and, by extension, it becomes more feasible to study the effects of interventions on different children in different contexts and with subtle variations in protocol.

As in the type of research conducted by serving clinical practitioners in medicine and healthcare, serving teachers are perhaps better placed to come up with interventions and variants on current practice than educationists who no longer practice the art of teaching on a daily basis, or commercial organisations driven by the goal of finding a standardised cost-effective product, rather than the sort of context-specific outcomes that are demanded in evidence-based practice.

It is important to remember that it was teachers themselves who started producing trial results after attending training – the aim of which was simply to develop their understanding of the large-scale trials that they had been recruited to support. Looking back, we should perhaps not have been surprised at the enthusiasm of the teachers who have been involved. Teacher are constantly ‘experimenting’ informally, allocating children to different tasks, testing, and reporting what they see to improve pupil outcomes. Experimental research design gives them a structure to hang these activities on.

The findings from our projects suggest that with systematic organisation teacher-led experimental research could become a national and global means for teachers to create, communicate, and share findings.

There is also the possibility of teachers beginning to collaborate at scale in a similar way to that shown to work in the emerging field of citizen science. For example, teachers could be given access to well-designed protocols that they then

---

"This practice should be taken to the UNESCO’s domain to influence every country’s education Ministers to effect this in teacher education curriculum and in practice in the field as a standard practice across the globe."

Ayodele Odeogbola, Lagos, Nigeria

"I feel the research should not end here. There should be opportunities to conduct more research projects and present them. The world needs more in-depth research conducted over longer periods of time and varied conditions to implement the takeaways in the education sector."

Kavita Sanghe, Mumbai, India

---

"My sense of agency is huge... as [Test and Learn] has encouraged me to collaborate with other teachers around the globe, in my lessons."

Marj Brown, Johannesburg, South Africa

---

Practitioner-led experimental research has greater potential to control extraneous variables (variations in implementation) compared to larger scale trials – which face even greater challenges, in this regard, compared to laboratory studies. They also have the potential for higher levels of mundane realism (‘everyday-ness’) – reducing the way that participants may react to being in the trial. Of course, they themselves face the challenge that at an individual trial level they have a smaller sample size and therefore it is harder to generalise. The same is true for areas such as surgery. The answer is to conduct multiple planned replications building large samples with specific meta-analysis.

Practitioner-led experimental research offers up the possibility of breaking complex multifaceted pedagogical interventions down into their component parts; and, by extension, it becomes more feasible to study the effects of interventions on different children in different contexts and with subtle variations in protocol.

As in the type of research conducted by serving clinical practitioners in medicine and healthcare, serving teachers are perhaps better placed to come up with interventions and variants on current practice than educationists who no longer practice the art of teaching on a daily basis, or commercial organisations driven by the goal of finding a standardised cost-effective product, rather than the sort of context-specific outcomes that are demanded in evidence-based practice.

It is important to remember that it was teachers themselves who started producing trial results after attending training – the aim of which was simply to develop their understanding of the large-scale trials that they had been recruited to support. Looking back, we should perhaps not have been surprised at the enthusiasm of the teachers who have been involved. Teacher are constantly ‘experimenting’ informally, allocating children to different tasks, testing, and reporting what they see to improve pupil outcomes. Experimental research design gives them a structure to hang these activities on.

The findings from our projects suggest that with systematic organisation teacher-led experimental research could become a national and global means for teachers to create, communicate, and share findings.

There is also the possibility of teachers beginning to collaborate at scale in a similar way to that shown to work in the emerging field of citizen science. For example, teachers could be given access to well-designed protocols that they then deliver in their local context contributing to large-scale data sets. Finally, multiple planned teacher-led trials exploring a single intervention (with the synthesis of findings in a meta-analysis as the outcome), could have much potential in adaptive programming environments – where testing, learning and iteration are required to find solutions.

Successive adaptations, trialled in short pre- and post-test designs, could be entered into the meta-analysis creating the opportunity to assess the relative effectiveness of adaptation over time, as well as comparing the effects on different pupil groups and contexts.

Scaling will inevitably have its challenges. Effective change of this nature needs national buy-in, collaboration, and investment. Turning the world of education research on its head so that teachers become the producers of the research rather than just the passive consumers of it, requires changes in behaviour at all levels of the system and within a wide range of stakeholder groups (including academia, government, and school leadership teams). At the same time, maintaining quality at scale will require robust systems of accountability (including links to accreditation and peer review) that ensure scrutiny of research methods and results to the same standard expected in academic research.

---

Appendix and references
Appendix 1: Meta-analysis of teacher research effect sizes

How to read the meta-analysis

Figure 5 (below) gives a 'forest plot' for the 35 results from the 19 trials completed by the Varkey Foundation Global Teacher Ambassadors across 11 countries. The total sample size across all the trials was 2,605. Because many of the studies produced data that was not normally distributed the effect sizes reported in the main forest plot are Rosenthal’s r rather than Cohen’s d. Although many popular education meta-analyses report d for all studies this would normally be considered a break with the convention of using other effect sizes where the data is non-parametric (e.g. not normally distributed). In this meta-analysis, other effect sizes were converted to r, so that the trial results were presented on the same scale. Because the trial results measured the effect of different interventions in different ways a random effects model was used to combine the results and calculate the overall effect size and significance.

The effect sizes produced by the teacher trials are grouped into three types of intervention. These represent the broad areas of interest chosen by the teachers:

- Progress/attainment
- Attitude/confidence
- Personal/emotional

Each dot [•] represents the effect size. The error bars [±] (either side of the dot) represent 95% confidence intervals. Such intervals are estimates of the range of results that one might expect in 95 out of 100 replications and therefore can be considered a measure of the generalisability of the teachers’ results. The relative size of the dot indicates the weighting that was assigned to that effect size within the random effects model.

Positive effect sizes, those on the right of central vertical line (r > 0.00), show that the treatment produced improved pupil outcomes compared to the control condition (normal classroom practice).

Negative effect sizes, to the left of the central vertical line (r < 0.00) show that the control group performed better than the intervention. For completeness, and because much education research has defaulted to the use of Cohen’s d, this effect size has been given in the final right hand column along with an indication of the statistical significance of each trial (using p-values). The left-hand side of the forest plot gives a narrative summary of the areas that were being explored in the trials; while on the opposing side information is given about the school year the children were in and the subject area that was being investigated.

Finally, information is given about the type of research design deployed by the teachers:

- Between-participant (no marking)
- Within-participant
- Matched-pair/case-controlled (I)
- Non-randomised controlled trial (II)

Meta-analysis of the results from the Varkey Foundation project

<table>
<thead>
<tr>
<th>Effect size (r)</th>
<th>United States</th>
<th>Chile</th>
<th>Colombia</th>
<th>Philippines</th>
<th>Nigeria</th>
<th>Australia</th>
<th>India</th>
<th>South Africa</th>
<th>Colombia</th>
<th>Lebanon</th>
</tr>
</thead>
<tbody>
<tr>
<td>-0.60</td>
<td>-0.40</td>
<td>-0.20</td>
<td>0.00</td>
<td>0.20</td>
<td>0.40</td>
<td>0.60</td>
<td>0.80</td>
<td>1.00</td>
<td>1.20</td>
<td></td>
</tr>
<tr>
<td>Full-time science (Belgard)</td>
<td>Science test (Gr 8)</td>
<td>2.14 ***</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Reading comprehension training (Henriques)</td>
<td>Standardised reading (He)</td>
<td>1.35 **</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Augmented reality learning (Samhita)</td>
<td>Science test (PT)</td>
<td>1.19 **</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Proficiency centred instruction, foreign language (Patel)</td>
<td>Reading comp. (Gr 7-8)</td>
<td>1.59 **</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>In-service training on terms (Salamanca)</td>
<td>Length measurement (Gr 6)</td>
<td>1.20</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Higher level thinking opportunities (Mason A)</td>
<td>End of unit test (BS)</td>
<td>1.05</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Electronic book (France Cartoons)</td>
<td>Reading comp. (Gr 4-9)</td>
<td>0.35</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Read a child to read another child (Mason B)</td>
<td>Reading test scores (P)</td>
<td>0.87</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Design-thinking (Tuscans)</td>
<td>Critical thinking (Gr 4)</td>
<td>0.82</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Classroom seating arrangement of peer collaboration (Ayodele)</td>
<td>Science understanding (Gr 4)</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Reading a child to read alone (Mason B)</td>
<td>Reading test scores (BS)</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>R2: mathematics (Mathews and O’Cearra)</td>
<td>Multitask test scores (TB)</td>
<td>0.49</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Monitoring by an industry expert (Samhita)</td>
<td>Science attainment GCSE (HS)</td>
<td>0.41</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Reading alone vs reading to another child (Mason B)</td>
<td>Reading test scores (ES)</td>
<td>0.20</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>R2: mathematics lessons (Mathews and O’Cearra)</td>
<td>Multitask test scores (TB)</td>
<td>0.10</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mental health of the teacher (Anasth)</td>
<td>Multitask test scores (TF)</td>
<td>0.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>After school science (Anthony)</td>
<td>Critical thinking (Gr 9-11)</td>
<td>0.10</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Proficiency centred instruction, foreign language (Patel)</td>
<td>Motivation (Gr 7-8)</td>
<td>4.15 ***</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Proficiency centred instruction, foreign language (Patel)</td>
<td>Use everyday life (Gr 7-8)</td>
<td>3.71 ***</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Proficiency centred instruction, foreign language (Patel)</td>
<td>Engagement (Gr 7-8)</td>
<td>3.51 ***</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>School student engagement program (ISTEM) (Silharta)</td>
<td>Attitude – career (HS)</td>
<td>3.91 ***</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Proficiency centred instruction, foreign language (Patel)</td>
<td>Understanding (Gr 7-8)</td>
<td>2.57 ***</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>School student engagement program (ISTEM) (Silharta)</td>
<td>Attitude – leisure interest (HS)</td>
<td>3.20 ***</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>School student engagement program (ISTEM) (Silharta)</td>
<td>Attitude – enjoyment (HS)</td>
<td>3.12 ***</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>School student engagement program (ISTEM) (Silharta)</td>
<td>Attitude – implications (HS)</td>
<td>3.82 ***</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>School student engagement program (ISTEM) (Silharta)</td>
<td>Attitude – adoption (HS)</td>
<td>3.11 ***</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>School student engagement program (ISTEM) (Silharta)</td>
<td>Attitude – normality (HS)</td>
<td>0.68 **</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>School student engagement program (ISTEM) (Silharta)</td>
<td>Attitude – inquiry (HS)</td>
<td>0.63 **</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intervention by an industry expert (Samhita)</td>
<td>Science confidence</td>
<td>0.39</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Race, science and pseudo-science lessons (Brownell)</td>
<td>Race / identity scores (Gr 10)</td>
<td>0.30</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sea Education Curriculum (Bernardo Galvez)</td>
<td>Pregnancy risk situations (Gr 10)</td>
<td>0.96</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Somatic techniques of mindfulness (Rubin Aronson)</td>
<td>Co-existence w/ls violence (Gr 11)</td>
<td>0.82</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sea Education Curriculum (Bernardo Galvez)</td>
<td>Asked for health services (Gr 12)</td>
<td>0.45</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Home grown reading programs (Buchanan)</td>
<td>Attendance (Gr 7)</td>
<td>0.60</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Addition of BrainGym to Ruler Method (Ballout)</td>
<td>End of term exam (Gr 7-8)</td>
<td>-0.10</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

 appendix and references

Appendix 1: Meta-analysis of teacher research effect sizes

‘Meta-analysis is a statistical method that allows for the synthesis of quantitative evidence from related research in a way that can summarise that body of evidence.’

Two forms of meta-analysis may be undertaken using either a ‘fixed effects’, or a ‘random effects’ model. When a fixed effect model is applied the assumption is that all results are an estimate of the same ‘fixed’ treatment effect typically all the effects are related to each other, both in terms of the intervention and the form of measurement). By contrast, a random effects model allows for difference in treatment effect between studies (i.e. the combination of varied interventions and measures). The use of RCs in education has been criticised. However, it is suggested that multiple planned replication by teachers (leading to creation of robust meta-analyses) could mitigate many of these criticisms.

How to read the meta-analysis

Figure 5 (below) gives a ‘forest plot’ for the 35 results from the 19 trials completed by the Varkey Foundation Global Teacher Ambassadors across 11 countries. The total sample size across all the trials was 2,605. Because many of the studies produced data that was not normally distributed the effect sizes reported in the main forest plot are Rosenthal’s r rather than Cohen’s d. Although many popular education meta-analyses report d for all studies this would normally be considered a break with the convention of using other effect sizes where the data is non-parametric (e.g. not normally distributed). In this meta-analysis, other effect sizes were converted to r, so that the trial results were presented on the same scale. Because the trial results measured the effect of different interventions in different ways a random effects model was used to combine the results and calculate the overall effect size and significance.

The effect sizes produced by the teacher trials are grouped into three types of intervention. These represent the broad areas of interest chosen by the teachers:

- Progress/attainment
- Attitude/confidence
- Personal/emotional

Each dot [•] represents the effect size. The error bars [±] (either side of the dot) represent 95% confidence intervals. Such intervals are estimates of the range of results that one might expect in 95 out of 100 replications and therefore can be considered a measure of the generalisability of the teachers’ results. The relative size of the dot indicates the weighting that was assigned to that effect size within the random effects model.

Positive effect sizes, those on the right of central vertical line (r > 0.00), show that the treatment produced improved pupil outcomes compared to the control condition (normal classroom practice).

Negative effect sizes, to the left of the central vertical line (r < 0.00) show that the control group performed better than the intervention. For completeness, and because much education research has defaulted to the use of Cohen’s d, this effect size has been given in the final right hand column along with an indication of the statistical significance of each trial (using p-values). The left-hand side of the forest plot gives a narrative summary of the areas that were being explored in the trials; while on the opposing side information is given about the school year the children were in and the subject area that was being investigated.

Finally, information is given about the type of research design deployed by the teachers:

- Between-participant (no marking)
- Within-participant
- Matched-pair/case-controlled (I)
- Non-randomised controlled trial (II)
What the meta-analysis tells us

Overall the teacher designed (or selected) interventions had a large positive effect.30 All but one effect size was positive. In the case of the one negative effect the teacher’s intention was to explore whether the use of brain gym might have an effect (without necessarily assuming it to be a positive treatment). The range of positive effect sizes was d = 0.02 to 4.13, with a substantial proportion of large² effect sizes (65.7%). Such findings represent treatment effects that are substantially greater than those normally found in education research generally and specifically in randomised controlled trials.32

In relation to interpreting the individual effect sizes (d), ‘0.2 is considered a small effect, 0.5 a moderate effect . . . 0.8 and above a large effect. The EEF, Sutton Trust effect, 0.5 a moderate effect . . . 0.8 and effect sizes (d), ‘0.2 is considered a small effect, 0.5 a moderate effect . . . 0.8 and above a large effect. The EEF, Sutton Trust effect, 0.5 a moderate effect . . . 0.8 and effect sizes (d), ‘0.2 is considered a small effect, 0.5 a moderate effect . . . 0.8 and above a large effect. The EEF, Sutton Trust effect, 0.5 a moderate effect . . . 0.8 and effect sizes (d)

APPENDIX AND REFERENCES

References


Mason, M. (2019a). Does providing higher-level thinking opportunities have an impact on student learning outcomes? An experiment in the use of enhanced learning materials promoting higher order thinking skills in a group of schools in Sierra Leone. Conference Poster, presented at the Global Education Skills Forum, Dubai, 22nd March.

Mason, M. (2019b). Reading to a doll is less effective than reading alone or reading in pairs – findings from a small-scale randomised controlled trial. Is there any impact on middle primary age children’s reading fluency and competence of them reading to a doll or to a partner as opposed to reading alone in Sierra Leone? Conference Poster, presented at the Global Education Skills Forum, Dubai, 22nd March.


Salamanca, E. (2019). A preliminary study on activities associated with the farming practices context in a school greenhouse, and its positive impact in length measurement learning, compared to traditional classes, in rural primary school students from 4th to 6th grade. Conference Poster, presented at the Global Education Skills Forum, Dubai, 22nd March.


About Education Development Trust

At Education Development Trust, we have been improving education around the world for over 50 years. We design and implement improvement programmes for school systems and provide consultancy services deploying specialists internationally.

Our work is informed by our continually refreshed body of research which focuses on the bright spots in education, from education authorities as diverse as those in Vietnam, Kenya, England, New York and Dubai.

Bringing about real change that alters the aspects of a national system that, for many reasons, aren’t working so well at the time, requires knowledge and ability to design and implement changes to any of the levers that can impede great educational outcomes. The ability to affect policy, practices, pedagogy, behaviour, funding, attitudes and more is a prerequisite for an organisation that can truly claim to transform lives through improving education.

As highly informed agents of change operating in low- to high-income countries with their varying contexts, we design, show and enable, so when working with us, everyone involved, from policymakers to school leaders and teachers, is able to apply their new knowledge to drive sustainable system reform.

Our expert knowledge, programme design and implementation expertise is also deployed in delivering Ofsted-rated outstanding careers services in England, and in owning and managing a family of independent schools.

We are a not-for-profit and we are driven by our values of integrity, accountability, excellence and collaboration.

For more information, visit www.educationdevelopmenttrust.com

The ability to affect policy, practices, pedagogy, behaviour, funding, attitudes and more is a prerequisite for an organisation that can truly claim to transform lives through improving education.
About the authors

Dr Richard Churches

Richard has sixteen years’ experience translating research evidence into government policy and practical teacher resources. Richard is co-author of Teacher-led research: designing and implementing randomised controlled trials and other forms of experimental research and Neuroscience for teachers: applying brain science in the classroom (foreword by Susan Greenfield) with Eleanor Dommett and Ian Devonshire. Their other book, Teachers’ Pocketbook Learning and the brain was a finalist in the 2012 Education Resources Awards. Richard has extensive experience leading a wide range of teacher-practitioner research. He completed a Master’s degree in education whilst an Advanced Skills Teachers, majoring on research methods and conducted qualitative participant observational research for his dissertation. His PhD, looked at the effects of charismatic oratory on the structure of consciousness and was quantitative laboratory experimental. Whilst a doctoral student at the University of Surrey, he was awarded the Pearson Education Prize for the accessibility and rigour of his research, won the Vitae Three Minute Thesis Competition and was nominated as departmental postgraduate of the year.

Astrid Korin

Astrid has fourteen years’ experience in policy formulation and multi-sector strategic planning across health, education and social development. Since joining Education Development Trust in 2017, Astrid has been leading a team of consultants and senior consultants whilst delivering on projects in the UK and overseas focussing on system leadership, evidence-based practice, collaborative school improvement and accountability in the early years. She is focussing on harnessing and building expertise in teacher-led randomised controlled trials (RCTs) as a way of increasing the evidence-base on what works and as a powerful tool to unleash voice and agency for teachers.

Kate Sims

Kate is an Education Consultant at Education Development Trust. Kate joined Education Development Trust in May 2017 after finishing her MA in International Relations and quickly developed a passion for working with inspirational teachers. Her role includes working closely with Dr Richard Churches to project manage and co-facilitate professional development courses for teachers in practitioner-led, classroom-based research.

Kate Sims

Kate is an Education Consultant at Education Development Trust. Kate joined Education Development Trust in May 2017 after finishing her MA in International Relations and quickly developed a passion for working with inspirational teachers. Her role includes working closely with Dr Richard Churches to project manage and co-facilitate professional development courses for teachers in practitioner-led, classroom-based research.
Education Development Trust’s large scale delivery is underpinned by evidence and research. As a registered charity, we reinvest a percentage of our surplus into a publicly available programme of educational research. We also invest in an ongoing Research and Development cycle to rigorously review the impact of our own programmes, generate new evidence on what works, and share insights to support broader debate and policy.

The evidence and insights shared in this handbook are part of this ongoing commitment to Research and Development.