WHAT WORKS?

OTHER TITLES FROM BLOOMSBURY EDUCATION

Bloomsbury CPD Library: Research-Informed Practice, by Jennifer Ludgate

Celebrating Difference, by Shaun Dellenty

Education: A Manifesto for Change, by Richard Gerver

Just Great Teaching, by Ross Morrison McGill

The Arts in Primary Education, by Ghislaine Kenyon

The Wellbeing Toolkit, by Andrew Cowley

WHAT WORKS?

Research and evidence for successful teaching

Lee Elliot Major and Steve Higgins

BLOOMSBURY EDUCATION
LONDON OXFORD NEW YORK NEW DELHI SYDNEY

BLOOMSBURY EDUCATION
Bloomsbury Publishing Plc
50 Bedford Square, London, WC1B 3DP, UK

BLOOMSBURY, BLOOMSBURY EDUCATION and the Diana logo
are trademarks of Bloomsbury Publishing Plc

First published in the UK by Bloomsbury Education

Text copyright © Lee Elliot Major and Steve Higgins, 2019

Lee Elliot Major and Steve Higgins have asserted their rights under the Copyright, Designs and Patents Act, 1988, to be identified as Authors of this work

Bloomsbury Publishing Plc does not have any control over, or responsibility for, any third-party websites referred to or in this book. All internet addresses given in this book were correct at the time of going to press. The author and publisher regret any inconvenience caused if addresses have changed or sites have ceased to exist, but can accept no responsibility for any such changes

All rights reserved. No part of this publication may be reproduced or transmitted in any form or by any means, electronic or mechanical, including photocopying, recording, or any information storage or retrieval system, without prior permission in writing from the publishers

A catalogue record for this book is available from the British Library

ISBN: PB: 978-1-4729-6563-9; ePDF: 978-1-4729-6562-2;
ePub: 978-1-4729-6561-5

2 4 6 8 10 9 7 5 3 1

Text design by Marcus Duck Design
Typeset by Newgen KnowledgeWorks Pvt. Ltd., Chennai, India
Printed and bound in the UK by CPI Group (UK) Ltd, Croydon, CR0 4YY



All papers used by Bloomsbury Publishing Plc are natural, recyclable products from wood grown in well managed forests. The manufacturing processes conform to the environmental regulations of the country of origin

To find out more about our authors and books visit www.bloomsbury.com
and sign up for our newsletters

CONTENTS

Introduction: Bananarama and best bets for learning vii

How to use this book xx

What works? Or what's worked? Some technical notes xxii

1. Improving classroom teaching 1
2. Effective feedback for learning 13
3. Teaching assistants 25
4. One-to-one tuition 29
5. Thinking for yourself: self-regulation and metacognition 35
6. Speaking and listening skills 41
7. Peer tutoring 45
8. Mastery learning 51
9. Phonics 57
10. Understanding texts: teaching reading comprehension strategies 63
11. Homework 67
12. Learning styles 73
13. Academic setting: education's zero-sum game 77
14. Digital technologies: more than motivating the reluctant learner 83
15. Smaller classes 87
16. Summer-born disadvantage 91
17. Preparing to learn 97
18. Individualised learning 103
19. Collaborative learning: working together effectively 109
20. Sports for health and wider outcomes 115
21. Art for art's sake 119

Afterword: A final principle 125

References 131

Index 145

INTRODUCTION: BANANARAMA AND BEST BETS FOR LEARNING

Teachers always laugh at the joke. The Bananarama Principle is named after Bananarama and Fun Boy Three's 1982 hit single 'It Ain't What You Do'. The original song 'T'aint what you do, (it's the way that cha do it)' was written by jazz musicians Melvin 'Sy' Oliver and James 'Trummy' Young in 1939, and was made famous by Ella Fitzgerald and the 'Shim Sham' tap dance routine. And it's essential advice for today's educators acting on the findings of research. For many teachers the 1980s cultural reference conjures up a bygone era. But the catchy ska riff sticks in their minds. It's the most important thing we have to say.

This guiding law came to life in 2011 when we toured the country to share with teachers the Pupil Premium Toolkit (Higgins et al., 2011). Published by the Sutton Trust charity, the toolkit offered an accessible *Which*-style guide detailing best bets for improving children's attainment. The UK Coalition Government had ring-fenced Pupil Premium funds for schools to help improve the results of their poorest pupils, who continued to lag behind their more privileged peers (Department for Education, 2010a). Our point was a simple one: it wouldn't be what schools spent, but the way they spent it that would get results.

The Pupil Premium Toolkit was born from a meeting of two different minds – a product of compromise. That's its strength and its weakness. One of us (Steve) was obsessed with making meta-synthesis – the summarising of thousands of studies – a practical tool for teachers (Higgins, 2018). This guide would provide the big-picture view, hovering above the research landscape and rising above the contradictory and conflicting conclusions from single studies. And one of us (Lee) was convinced we should spend as much time translating the complex jargon of academe into an accessible language for teachers as producing the findings in the first place. Research summaries had existed before. But like so many publications, they were soon gathering dust on long-forgotten library shelves.

Combining these goals was a balancing act. Syntheses of research are only as good as the assumptions they're built on and the quality of studies they summarise. If you simplify the messages too much, the true story gets lost in crude headlines. The first toolkit (Higgins et al., 2011) was a 20-page report collating findings from thousands of studies corralled into evidence summaries for 20 teaching approaches aimed at improving learning.

A lot of thinking went into the design and presentation of the guide. It presented the costs of implementing the approaches and the strength of evidence underpinning our conclusions of how well it had worked. Brave calls were made translating effect sizes, the comparative impacts of interventions, from the language of statistics into a more meaningful measure for everyday classroom use: extra months gained during an academic year (the assumptions we made in order to do this are explained in the technical notes on page xxii). These estimates are less precise than the underlying statistics, but they are realistic guides, given the limitations of the research findings on which they stand.

Without this, the toolkit wouldn't have caught the imagination of busy classroom professionals. Teachers immediately got it.

Our toolkit talks made for lively debates. They were attended by thousands of teachers and senior school leaders across the country. This too was a respectful coming together of two different perspectives. The best discussion came when the findings gathered by academics under idealised conditions collided with the practical everyday realities of teachers' working lives. Together we would tease out how the big-picture lessons applied on ground level.

The counter-intuitive conclusions caused a stir (Major, 2012). On a grey last day of term in Cambridgeshire we unveiled to hundreds of teaching assistants (TAs) the humbling finding that assistants, on average, added little to pupil achievement. It was a difficult meeting. No caveats or constructive comments could calm the enraged ranks of TAs. All they could see was an attack on their livelihoods. Below this average finding lay the real story. TAs properly managed, prepared and trained, and working alongside teachers, had significant impact on children; those without support, and allocated pupils with severe learning needs, struggled. In other words, it wasn't deploying teaching assistants or not that mattered most, but how they were deployed.

Reducing class sizes had surprisingly limited impact on pupil progress. Smaller classes are a school priority promoted by politicians, parents and teachers alike. But studies have found the benefits of reducing class sizes are not large or clear, until class size is reduced to under 20 or as low as 15 pupils per teacher. Teachers bridled at this conclusion. Many argued smaller classes were effective when serving children with special educational needs or challenging behaviour. The devil was in the research details. Smaller classes work when teachers change the way they teach, catering to individual needs of pupils and receiving more feedback from children on what they need to learn next. It's not reducing class size that matters, but how you adapt your teaching style with fewer pupils.

Grouping children by 'ability' as ever split opinions. For every passionate believer in mixed-ability teaching, there was an equally convinced advocate of academic setting. Evidence accumulated over several decades could not settle these apparently irreconcilable views. The reason was buried in the finer details of the toolkit: the quality of teaching trumps how children are organised. Delivered poorly, neither mixed-ability classes nor sets work; delivered well, both can be effective.

The use of teaching assistants, reducing class sizes and ability grouping are all examples of the Bananarama Principle. The Principle underscores the power but also limitations of evidence in helping a teacher decide what to do in a classroom: how an approach is implemented is vital and just as important as its content. Evidence is necessary, but never sufficient.

Success – a doubled-edged sword

The Pupil Premium Toolkit revealed an insatiable appetite for evidence. Adopted by the Education Endowment Foundation (EEF), it became the *Sutton Trust-EEF Teaching and Learning Toolkit*. When launched by the Government in 2010,

the EEF, funded with an initial £110-million endowment was compared to the 'Race to the Top' competition in the United States (Department for Education, 2010b). In reality the EEF did something different: promoting the use of evidence in schools and funding randomised trials assessing the effectiveness of different approaches and programmes in English schools. The EEF nurtured and developed the toolkit into the 35-strand interactive website you see today at: https://educationendowmentfoundation.org.uk/evidence-summaries/teaching-learning-toolkit. When the school inspectorate in England, Ofsted, gave it a seal of approval, it turned overnight into a must-read for headteachers, who had to report back on how they were using their Pupil Premium funds.

By 2015 two-thirds of school leaders across England reported they had used the toolkit (Cockburn et al., 2015). Versions have been launched in Scotland (https://education.gov.scot/improvement/eefsearch), Australia (https://evidenceforlearning.org.au/the-toolkit), Spain (www.educaixa.com/es/listado-evidencias) and Latin America (www.summaedu.org/plataforma-de-practicas-educativas-efectivas), with other countries waiting in the wings. It has helped to create a more evidence-informed culture in classrooms across the world. Teachers ask questions about research in a way they didn't a decade ago. Hardly a day goes by when one of us doesn't meet a teacher who excitedly encourages us to look at the toolkit they have discovered – not knowing we are its creators!

But the problem with popularity is you lose sight of the reasons that made you successful in the first place. The toolkit's findings were digested directly without the nuances, caveats and ambiguities that can be explained during face-to-face discussion. Research claims were treated as unassailable truths. Classrooms were assumed to be the same homogenous units. Best bets turned into absolute certainties.

The UK Government cited the toolkit as an exemplar for its new network of 'What Works' centres – for policing, early interventions, ageing – alongside those already established for health and education (Cabinet Office, 2013). This drive to spend public money on the most effective practices was laudable, but the 'What Works' label can imply a certainty that oversimplifies reality. Research can only tell us what has worked in the past, not what will work in the future. Indeed, it can only offer indications of what may work under certain conditions.

The toolkit's findings were presented by some journalists as stark, black-and-white decisions. 'Setting and streaming? NO. Teaching assistants? NO,' declared a BBC journalist briefed to write about the Government's new initiative (Easton, 2014). The complexities of converting research into practice had been bastardised into a binary code. Research headlines prompted knee-jerk reactions. We made the case for training teaching assistants – and the teachers working with them – not sacking them.

Meanwhile parallels were made with the movement to create more evidence-based practice in medicine (see www.nice.org.uk). There are several reasons why the pharmaceutical model of research doesn't apply in education, not least of which is that it isn't clear who is supposed to take the tablets in a classroom. The 'medicine' has to be palatable to the teacher yet effective with the pupils. A classroom is a much more complex place than a consulting room. You usually

have one teacher and 30 pupils. It is like finding the prescription to make all your patients healthy, when they are suffering from a range of undiagnosed medical conditions.

It was only a matter of time for a backlash to this drive to make teaching an evidence-led profession to erupt. Prominent education figures argued we were deluded to think teaching could ever be completely research-based (Wiliam, 2015). Others said it was compromising the freedom of teachers to decide what's best to do for the unique make-up of children in their classrooms. Some argued evidence couldn't answer education's most important question: what is it for? (Biesta, 2007)

Surveys of teachers meanwhile revealed greater recognition of evidence, but scant signs of more focus on the approaches shown by the toolkit to offer the highest chance of improving attainment (Sutton Trust, 2013). Only one in 25 teachers had prioritised improving feedback between teachers and pupils, an inexpensive measure with the potential to add eight months to pupils' learning in one year. And only one in 100 said they would use peer-to-peer tutoring schemes, another highly promising approach according to the toolkit.

Publishing the evidence doesn't automatically lead to schools adapting their teaching practices. And some teachers got it plain wrong. We were concerned to hear stories of schools devoting extra hours to the marking of pupils' work under the false belief this would deliver effective feedback. One local authority in Scotland tried to use the toolkit as justification for *increasing* class sizes, failing to consider the overall evidence that pupils in larger classes tend to do a little bit less well, on average. Our argument was that reducing class sizes is only moderately effective and very expensive, so not cost-effective as a strategy to help learners.

England's trials

The inherent uncertainties of applying evidence in the classroom have been confirmed by randomised controlled trials investigating interventions in England's schools. In seven years, the EEF signed up more than 10,000 schools, nurseries and colleges to 150 randomised trials, involving one million children. Most approaches trialled are little better than 'teaching as usual' in other schools: only around a quarter of EEF studies show enough impact to warrant larger scale-up trials involving more schools (EEF, 2019a). The most promising programmes boost students' progress by an extra three months in a school year, yet the effect shrinks when scaled up to hundreds of schools.

Scour through the EEF's evaluation reports and a common thread emerges. Trials of peer tutoring in English schools are a case in point (EEF, 2017a). Peer tutoring, where pupils teach each other, is one of the most promising approaches according to evidence gathered for decades in schools across the world. Yet evaluations in England failed to find any attainment gains on average compared with teaching as usual. How could this be? In part, it was due to inferior past studies that had overestimated the size of effects. But a key aspect was the delivery. Some teachers felt more guidance and time was required to make tutoring work; others felt the programme was too prescriptive. It's tough to get

the right balance between producing detailed instructions and allowing teachers to shape the programme for themselves. We perhaps also forgot that it matters *what* pupils teach each other. We need to get the content right as well as the pedagogy. Every classroom is a unique web of interactions between teachers and pupils. Delivery is difficult.

What works?

Our aim in this book is to capture the spirit of the toolkit and explore the reality of 'what works' in education. On the one hand, that means providing practical tips for teachers and leaders; on the other hand, it means instilling a healthy dose of scepticism for teachers when scrutinising how approaches are working in their own schools and classrooms. As we will explain below, this book allows us to venture into new ground, explaining the recurring principles that apply across many approaches and offering informed but more speculative viewpoints, considering a range of pupil outcomes. We assess the research and evidence for 21 different teaching approaches. We think that there are some broad principles which teachers and schools should bear in mind as they use evidence to help improve outcomes for learners in their care.

'What works?' is a deceptively simple question. The question mark in the book's title is important! Knowing something has worked in the past for some pupils and teachers is no guarantee in education for predicting what will happen in the future for other teachers and pupils. 'What works?' is shorthand for 'What has worked for whom under what conditions?'. Our best attempt at answering this is explaining how the Bananarama Principle applies for each approach.

The principle highlights the statistical pattern we observe in the summaries of effect sizes from different studies. The impact on children varies more within each toolkit strand than it does between strands. It's not choosing to do peer tutoring that matters but how well you do it. The spread of effects for an approach is as important as its average impact: it indicates what you should focus on, and what you should avoid, to maximise the chances of making an impact.

Feedback in the classroom, for example, when delivered well, yields on average higher learning gains compared with most interventions, but studies show some efforts at providing feedback can harm learning. The overall average suggests feedback is a good bet for teachers, but the variation in results raises the risk you could make things worse for your pupils.

Ella Fitzgerald sings, 'It ain't what you do, it's the time that you do it, that's what gets results' in her 1938 song. This is true for feedback, as the timing of the information given and received between teacher and pupil is crucial. Too early and you risk prompting the learner unnecessarily; too late and the moment has passed.

Limits of research

Teachers will always have limited research evidence to draw on. It will never answer every question. Our knowledge from meta-analysis is like a medieval

map of the world, where some areas are better known, such as learning to read (Higgins, 2018). In other areas, the evidence is less secure, but still coherent, such as about collaborative learning or small-group teaching. There are also other areas that resemble mythical lands – fantasies such as learning styles or multiple intelligences.

Much of the educational territory has been explored, but more mapping is needed to increase the odds of successful educational decision-making. There are errors and gaps in what we know. Some of the underlying studies are not as robust as we'd like. Our approach assumes there is no systematic bias in different areas of research (Simpson, 2017). Where bias exists, this may distort some of our conclusions. Our picture isn't precise enough to inform practice as much as we would like, but it's the only approach we have and the best starting point to build upon.

Our compass showing the way ahead will never be perfect. Researchers have found practices observed among effective teachers had little predictive power when applied to a different group of teachers (Coe et al., 2014). Certain types of behaviour led to higher pupil gains, but there were always exceptions. The humbling conclusion is that the behaviour of effective teachers is hard to characterise; much depends on how teachers relate to their particular pupils (Brown et al., 2001). Great teachers are master orchestrators – knowing what to do when for what children. It's not what they do but how they do it that counts.

Coaxing learning amidst the complexities of human interactions feels more like navigating through choppy waters. Classroom feedback has been likened to throwing bottles into the sea – 'no one can be sure that the message they contain will one day find a receiver' (Hargreaves, 2011). In his seminal book *The Hidden Lives of Learners*, Graham Nuthall (2007) revealed the challenge every classroom teacher is up against: 70 per cent of pupils' time was spent pretending to listen; 80 per cent of the feedback pupils' received was from their fellow peers; and 80 per cent of this was wrong!

The education world is also an evolving ecosystem, where what works changes over time. We now know teaching assistants trained to provide specified support for one-to-one or small-group tutoring can boost pupils' learning by up to five extra months in an academic year. This comes from new evidence generated by several trials (Sharples et al., 2018). This territory has been mapped more accurately. Learning benefits were found across several subjects in both primary and secondary schools.

In this book, we have used the best evidence we can – over 200 summaries of 8,000 intervention studies assembled together. It is the best evidence we have at this point in time. If organisations like the EEF and the Institute for Education Sciences in the United States continue to fund evaluations, the quality of this evidence base will improve. We doubt this will make it more prescriptive for teachers. It is more likely to show how difficult it is to improve outcomes for children at scale (Lortie-Forgues and Inglis, 2019). But it may help us understand better what to do for whom and when.