Nesta...

PEARSON

FROM GOOD INTENTIONS TO REAL IMPACT

Rethinking the role of evidence in education businesses

Geoff Mulgan, Ruth Puttick and Simon Breakspear
Foreword by **Sir Michael Barber**

February 2014

Acknowledgements

The authors would like to thank Francesca Sanders for helping us research the paper in the early stages, and Joe Ludlow for sharing Nesta Impact Investments' experience of using the Standards of Evidence. Thanks also to our Nesta colleagues Mark Griffiths, Kathleen Stokes and Amy Solder for providing useful comments and insights.

We are also grateful for the assistance provided by Jacqueline Cheng and Rachel Brujis from Pearson for helping finalise the paper.

We would also like to thank the external reviewers, Robbie Coleman from the Education Endowment Foundation and Pamela Sammons, University of Oxford, for providing such helpful insight and feedback.

PEARSON

Pearson is the world's leading learning company.

Our education business combines 150 years of experience in publishing with the latest learning technology and online support. We serve learners of all ages around the globe, employing 45,000 people in more than 70 countries, helping people to learn whatever, whenever and however they choose. Whether it's designing qualifications in the UK, supporting colleges in the US, training school leaders in the Middle East or helping students in China learn English, we aim to help people make progress in their lives through learning.

Nesta...

Nesta is the UK's innovation foundation.

An independent charity, we help people and organisations bring great ideas to life. We do this by providing investments and grants and mobilising research, networks and skills.

Nesta is a registered charity in England and Wales with company number 7706036 and charity number 1144091. Registered as a charity in Scotland number SCO42833. Registered office: 1 Plough Place, London, EC4A 1DE.

www.nesta.org.uk

FROM GOOD INTENTIONS TO REAL IMPACT

Rethinking the role of evidence in education businesses

CONTENTS

FOREWORD
SIR MICHAEL BARBER, PEARSON

Today we regularly see front-page articles bearing the tides of a new age, where big data will provide the insights to improve outcomes in both public and private systems. There is an increasing recognition of the power of data, or put another way, the insights and evidence we can draw from data, to improve performance. We might date the popular trend from the unlikely story of *Moneyball,* made famous by Michael Lewis, where the scrappy Oakland Athletics used evidence and data analytics to identify the best players and dominate Major League Baseball. Despite spending two-thirds less money on salaries than competitor teams, Oakland Athletics won a league-leading 103 games and made the playoffs.

Similarly, in the UK government Prime Minister's Delivery Unit, which I led from 2001–2005, we used regular analysis and persistent interrogation of data to improve the performance of public services and increase value for money for the taxpayer. As the public increasingly demands outcomes, the demand for evidence about what works has gone up in tandem

Given evidence and data's strong track in improving outcomes, the demand for evidence is understandably growing. Equally, there is an ever-accelerating effort from public and private industry to meet this demand. In education, for example, there has already been great progress in recent years with the growing quality of education researchers, and notable contributions to the data from the OECD.

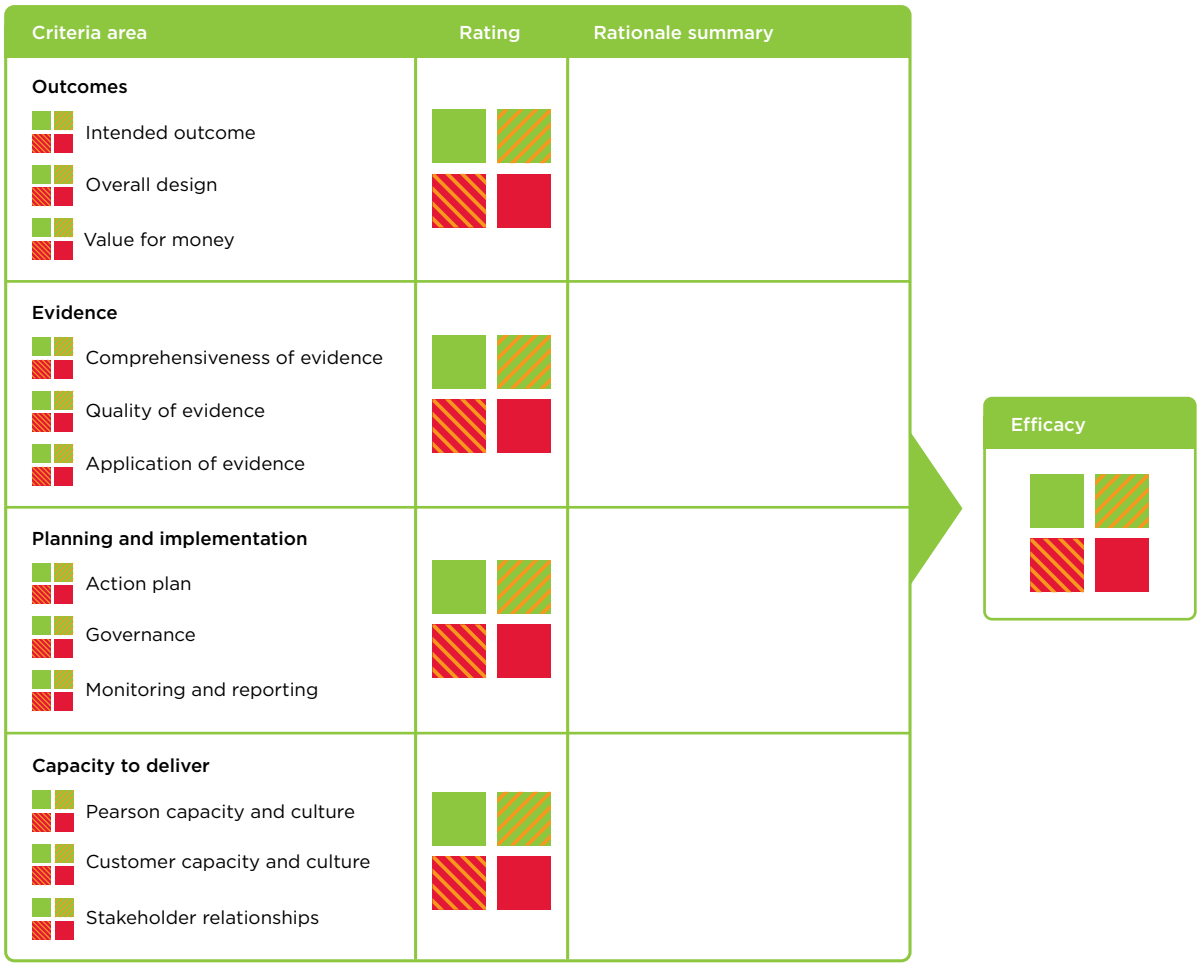
As evidence proliferates, it is important that we all become savvy interrogators of the available research. As we see in the standards set out here – not all evidence is created equal. It is the responsibility of every business leader to both increase the quality of the evidence created, and to elevate the rigor with which we examine the existing evidence.

This is less straightforward than we might like. Often there are difficult decisions for business leaders to make about the scale of impact and the rigor of testing that is required. The Lean Start-Up[1] or Jugaad Innovation[2] would tell us to innovate quickly, fail fast and learn from our errors, but by its nature this method will have a lower standard of evidence. Similarly, the time required to do randomised control trials will demand that we often make decisions with less than perfect knowledge of what works. This is why a holistic, systemic understanding of your intended outcomes is critical to underpin effective evidence gathering and decision making.

At Pearson, we are therefore implementing the efficacy framework, which takes our broad aspiration of improving people's lives through learning and turns it into a practical reality. We ask every product team to answer four key questions about their products to understand the likelihood of impact:

1. What outcomes are you trying to achieve?
2. What evidence do you have that you can achieve it?
3. What is your plan to deliver?
4. Do you and your partners (teachers, districts, students) have the capacity to delivery?

Figure 1



Learn more at: **efficacy.pearson.com**

Asking robust questions about outcomes and the evidence of social impact is radical for a business. It is more complicated and less well defined then measuring financial outcomes, and requires a deep understanding of the standards of evidence.

As Nesta lay out in this paper, every business will need to tackle these problems over time. They must increase the quality of the evidence used to inform the development of new products, and improve the processes to measure the impact of existing products. We see this already in the growth of R&D around the globe, including, incidentally, as part of my work in Pearson. We also see this in public education systems, as they increasingly measure individual student level data to monitor and improve every learner's individual performance. Helping leaders to define what data to collect, and how to evaluate the evidence, will be a meaningful impact of publishing these standards of evidence.

To achieve this impact, we require entirely new paradigms to understand the collection, use and analysis of evidence in order to draw meaningful insights. To continue the debate started here by Nesta, Pearson will publish a paper, *Digital Ocean,* which sets out the increasingly close link between collection, analysis, insight and action in education. Importantly, the revolution requires not just an increase in the volume of analysis, but rather entirely new ways of thinking about our education system and pedagogical approaches to take advantage of the newly available information.

It also requires the radical transformation of entire private organisations. At Pearson, we are transforming the company around a single idea – our ability to have a measurable impact on learners' lives. I, and my colleague Saad Rizvi, have set out the transformation journey for one business in *The Incomplete Guide to Delivering Learning Outcomes.*

While we have started this journey towards improving and measuring learner outcomes, we know that we don't yet have all the answers. We are committed to the path to efficacy and to reporting publicly on learner outcomes, in addition to financial results, by 2018. We know, though, that we won't be able to do this alone – and that there are many people thinking about how to improve efficacy and measure their impact. That is why we are collaborating with Nesta – to provide us, as well as other businesses the world over, with clarity about how to measure social impact with credibility. This has helped to advance our thinking already.

This is just one story though, and we hope that sharing this paper will inspire others to take up this journey and share their experiences applying the standards of evidence in business. We learn from this paper what every business can do to improve the quality of their evidence, and how this can meaningfully improve people's lives. We all now have a responsibility to carefully apply what works in order to consistently improve quality, value and equity. We hope that this paper will make a significant contribution to making this a reality.

Sir Michael Barber, Chief Education Advisor, Pearson

INTRODUCTION

The starting point for this paper was growing recognition that we should rethink the role of evidence in education businesses.

Over the past few decades a movement calling for better use of evidence in decision making has taken shape, with a focus on the need for 'evidence–based' policy and practice. Yet the role the private sector could and should play has largely been ignored.

We believe this is a mistake. The private sector plays an increasingly important role in education,[3] driving innovation through taking risks and investing new capital, providing books, technologies, courses, as well as running schools and colleges. This means the quality of what the sector provides, and evidence about its efficacy matters, and is increasingly important for learners themselves, as well as for public organisations that commission or purchase goods and services.[4]

Ensuring educational businesses and startups use evidence effectively has the potential for the new wave of educational innovation to have a consistently positive impact on learning outcomes. This means there is both a business imperative for taking evidence seriously and a moral imperative for businesses endowed with power and resources to do all they can to ensure that what they sell works.

Yet to achieve this, there is a need for new approaches, frameworks and ways of working to reframe current understandings of what evidence is, how it is captured, and to develop strategies and systems to ensure its better use. Rather than seeing success only in terms of sales, growth and profitability, we need evidence that encapsulates the impact that products and services have on the overall functioning of different sectors of the education sector including benefits for users.

In response to this challenge, this paper presents Standards of Evidence, an approach used by Nesta to guide the evaluation of their innovation programmes and investments.[5] Assessing the strength of evidence is a subcomponent of Pearson's Efficacy Framework (see Figure 1). Taken together we hope that they will enable education businesses to ensure they are having the biggest impact possible.

This paper is intended to start a discussion. At Nesta we recognise that many organisations already possess excellent evidence of the impact of at least some of their products and services, yet we have also come to realise this is not consistent across the field. We hope this paper will help stimulate a debate about the role evidence plays, what it could and should be, and the responsibility of businesses to commit to producing and applying it.

Definition of key terms[6]

We define:

1. An output as a measurable unit of a product or a defined episode of service delivery.
2. An outcome as an observable, and measurable, change for an individual or organisation. Pearson set the bar high for this, defining a positive outcome as an education product having a measurable impact on an individual's life, such as through preparing for work or college. Outcomes will then be clarified to confirm the target people to be reached and the timelines for achieving impact.
3. An education impact as the effect on educational outcomes attributable to the output, which may be positive or negative, and will be identified through high quality evaluation.
4. Efficacy is defined by Pearson as when education products[7] have *"a measurable impact on improving lives though learning."*[8]

SECTION 1: WHY WE NEED EVIDENCE OF IMPACT IN EDUCATION BUSINESSES

1. Growing interest in evidence

Education should be about learning, drawing on the world's accumulated knowledge, leading to new developments and advances. Yet a remarkable amount of education practice doesn't apply basic principles of good education to itself. It isn't continually reflecting on what is known and incorporating this evidence into ways of working.

Why evidence matters for decision making should be obvious.[9,10] Without evidence, even the best intentions could cause harm. The Scared Straight programme is a good example of this. A lot of resources, time and money were invested in taking children to visit prisons with the aim of deterring criminal behaviour. Yet it had the opposite impact; children who participated were in fact more likely to commit crime.[11] Analysis by the Washington State Institute for Public Policy revealed that for every $80 spent on such programmes, there were additional costs of $14,000 associated with the youths' recurring contact with the criminal justice system.[12] Unfortunately though, robust evidence of impact of this kind is too often absent in public policy, with many programmes growing without any acknowledgement of whether they are making a difference.

There are plenty of examples in education of fashionable policies and programmes that have spread, sometimes with little supporting evidence. For example, the push for smaller class sizes[13] or giving a laptop to every child[14] have become increasingly popular and advocated as a universal solution, when the impacts of these initiatives is often doubtful.

Yet change is happening. Over the past decade there has been a growing interest across the world in making evidence more used, and more useful, increasing the supply of research and evidence, as well as a growing demand for it, from across government, academia, education providers, users and others alongside.

In 2011, Nesta launched the Alliance for Useful Evidence,[15] in partnership with the Big Lottery Fund and the Economic and Social Research Council (ESRC), to accelerate this movement and advance the evidence agenda. The Alliance for Useful Evidence is a network of over 1,000 individuals and organisations, all committed to improving the use of evidence. Through a programme of events and publications, the Alliance brings together different fields, such as education, medicine, criminal justice and others, to act as a focal point for promoting evidence, enabling issues to be debated, and ideas and practice to be shared.

More recently, the UK government committed to creating the What Works Network, comprised of new evidence centres to bridge the evidence–practice divide in a number of policy areas,[16] including the Education Endowment Foundation (EEF) which is tasked with raising the bar for the field. In addition, there are centres focussing on evidence in education, such as the Evidence for Policy and Practice Information and Co-ordinating Centre (EPPI-centre),[17] the Institute for Effective Education at the University of York,[18] and the Centre for Effective Education at Queen's University Belfast,[19] all good examples of the drive to provide better evidence and more accessible information to policymakers and practitioners. The USA also has a long history of evaluating education policy with institutions like the Northwest Regional Education Laboratory[20] and the federal government-led What Works Clearing House.[21]

In 2011 Pearson partnered with the Economist Intelligence Unit to create The Learning Curve,[22] a collation of internationally–comparable data on education. They then worked with the EIU to analyse more than 2,500 data points covering 50 countries and two decades, and to draw out the lessons for policymakers and educationalists.

Global bodies like the World Bank and OECD have been pioneers in leading evidence–focussed projects, such as the Development Impact Evaluation Initiative,[23] the Systems Approach for Better Education Results[24] and the Centre for Educational Research and Innovation.[25] Across other areas of public policy, the use of rigorous evaluation techniques such as Randomised Controlled Trials (RCTs) are spreading further than just medicine.

At the same time, tools like Randomise Me,[26] a free–to–use online trials generator supported by Nesta, are making it easier for a wider population to carry out research on different interventions. Teachers can use Randomise Me to test different interventions or methods of teaching, such as new apps or timetable structures, and to help more teachers easily assess which best improves outcomes.

Funds are also cropping up that seek to support research and evaluation generation. The World Bank launched the Strategic Impact Evaluation Fund in 2012 to generate evidence of what works in various policy areas, including education. Launched in 2010, the Spanish Impact Evaluation Fund provided $14,000,000 of funding and support towards the evaluation of innovative programmes in health, education and social care. In France, the Fonds d'expérimentation pour la Jeunesse[27] ran large scale trials, with control groups, on issues such as parental involvement in schools.

As well as increased supply of research and evidence, there is growing demand for it. With shrinking budgets and limited resources, governments are recognising the value of evidence when making procurement decisions. Programmes like the US' Investing in Innovation Fund (i3)[28] offers a glimpse of what the future might hold, when for the first time, government funding of education programmes was allocated on the basis of the strength of evidence behind them.

Students themselves are increasingly demanding more evidence. Universities, schools and colleges are increasingly being held accountable for their performance, through both formal evaluation and more ad hoc user feedback. For instance, Key Information Sets in the UK[29] were introduced in 2012 to help students make better informed decisions on the higher education courses they select, drawing together data collected on student employment outcomes, course satisfaction, teaching hours and entry requirements.[30] In addition, there are online open platforms like 'My Edu'[31] and 'Professor Performance'[32] that are making it easier for pupils to feedback on their teaching experiences.

There are then many other free resources available that are dedicated to education which help teachers conduct and implement their own research on products and services, such as: 'DIY Evaluation Guide'[33] and 'The What Works ProCESs'.[34] Equally, the potential to test which programmes work is becoming increasingly feasible with the explosion of 'big data'.[35,36]

Numerous product and service evaluation sites have been created with the goal of empowering education professionals to make evidence–informed decisions. By assessing the strength of evidence against cost and usability, sites such as Best Evidence Encyclopaedia,[37] Graphite[38] and Blueprints for Healthy Youth Development,[39] edSurge[40] and ClassroomWindow[41] make it much easier to understand different initiatives. In addition, the EEF[42] is having a huge impact on increasing the prominence of evidence in education spending decisions across the United Kingdom. A survey conducted in 2013 found that 36 per cent of school education leaders were using their Teaching Learning Toolkit to inform their purchasing decisions,[43] which has been stimulated by a need to demonstrate the rationale for pupil premium funding (see text box for further details).

The Teaching and Learning Toolkit[44]

The EEF is an independent grant–making charity established to help break the link between family income and educational achievement.

The Teaching and Learning Toolkit was developed by the EEF, Sutton Trust and Durham University to provide objective guidance of what works across 33 topics in education practices and programmes, each summarised in terms of their average impact on attainment, the strength of the evidence supporting them and their cost, *"helping schools to get 'educational bang for their buck".* The quality of evidence is tiered into five categories of 'rigor' ranging from very limited to very extensive. A survey conducted by the National Foundation for Educational Research (NFER) in 2013 found that 36 per cent of school education leaders were using the tool to inform their purchasing decisions.[45]

The Toolkit is a live resource which will be updated on a regular basis as findings on education programmes emerge.

1.1 Opportunity for businesses

It's clear that there is a lot of activity underway, but as the initiatives listed above show, much of this effort is primarily focused on education at a government policy level or at practitioners, with little attention paid to the role that education businesses could and should play in this landscape.[46]

Some sectors already take evidence very seriously as they recognise the role evidence plays in helping bring goods to market, or they are in fact obliged to invest in evidence production. In the pharmaceutical sector, clinical trials are a regulatory necessity before a drug can be sold, and in the environmental field, ever more sophisticated reporting requirements, and internal tools, are bringing clarity to such issues as carbon reduction.[47] In agriculture there is a long history and extensive business sponsorship of research to test the effects of new pesticides or crops, often involving collaborations with Higher Education Institutes.

Education businesses are not yet required to collect evidence of impact before selling goods. Nonetheless many educational publishing companies, such as Pearson or the NFER, use research to develop and test materials and assessments. However, changes underway in government, investor and consumer practices, are prompting more education developers to take evidence of impact seriously as the likely condition for continued success in educational markets. There is then an opportunity for businesses to get ahead of the curve.

Demonstrating impact could increase demand. As well as evaluation to help identify unforeseen harm, demonstrating a positive difference would boost the credibility of a product or service, as well as enhancing the reputation of the company overall.

The signalling effect that evidence has is increasingly being recognised as a point of differentiation. Impact investment is one such field that is becoming more 'evidence aware', recognising that bringing effective goods to market could help sustain profits. For example, Nesta has launched an Impact Investment Fund.[48,49] As well as seeking financial return, it also explicitly seeks to understand and strengthen social impact by making rigorous evaluation an integral part of the investment package. This means that as well as growing output of the product or service, resource also needs to be allocated to evidence the impacts. The hypothesis is that evidence enables the most effective products and services to be selected, backed and scaled, increasing chances of profitable success in the longer term.

Evidence of impact could therefore be an increasingly important source of competitive advantage. Yet at the present time, most education businesses have barely begun to understand that this is a lens through which they will increasingly be judged, nor as a valuable component of strategic and future thinking.[50]

1.2 The need to rethink evidence of impact

This new evidence–aware world offers businesses important opportunities to get ahead if they can demonstrate the effectiveness of their goods and services in promoting various clearly identified desirable outcomes. Yet too often we know that evidence of impact is missing.

Frequently 'evidence of impact' is measured in terms of financial return, sales, and profitability, with positive figures being presented as success, with few education businesses having actual evidence of their impact upon clearly specified educational outcomes in specific sectors of specific user groups, nor have they developed plans for how they will seek to evaluate impact in the future. Most claims of proven learning practices do not extend beyond what Perkins (2010) calls *"good practice – We've done it, we like it and it feels like we make an impact".*[51]

The push for corporate social responsibility, where businesses are aware of and attempt to minimise their negative and maximise their positive social and environment impacts, has grown rapidly over the past decade. Yet this can often simply be a marketing exercise, or allocation of funds as grants to charitable causes, rather than focussed attention on the actual impacts of the businesses own goods and services.

So what is to be done? If learning outcomes are to improve then we need evidence of what is and isn't working to promote certain desired outcomes, whether this is in terms of individual students or professional learning, or more broadly conceived as organisation or sector learning . To achieve this we need practical tools and frameworks that help business leaders, entrepreneurs and investors deepen their understanding of what counts as 'good evidence', and to make it easier for information to be generated and analysed appropriately to generate evidence, and develop effective systems to ensure it is widely shared and applied. It can be helped by digital technologies that make feedback and aggregation easier than before, but it also needs to start with some simple principles: that any new approach is tested and improved in the light of experience.

Overall, the drive for evidence in education businesses is not fundamentally just a research activity, although it will undoubtedly need to be informed by a research mind–set, instead it is better conceived of as an ongoing learning activity, a process generating useful guidance to inform continual improvement.

In the next section of this paper we outline an approach and a framework that could be used in response to this challenge, and discuss the wider strategic changes needed if education businesses are truly going to 'do well by doing good'.

SECTION 2: HOW COMPANIES CAN APPLY AND GENERATE EVIDENCE

In Section 1 we outlined why businesses should be committed to defining themselves as learning organisations first and foremost by applying evidence to evaluate and improve impact on educational outcomes. In this section we discuss how evidence generation can be both achievable for education businesses.

1. The roles evidence should play in education businesses

We have highlighted that businesses need to take evidence more seriously. A key facet of this is getting over a common misperception that gathering evidence of impact is a simple pass or fail test. Instead, what's important is to embed a commitment to gathering evidence over time, to iterate and help improve performance. In effect, this means embedding the principles of good learning into the practices of education businesses.

We now outline key ideas central to rethinking the role of evidence.

1.1. Asking the right questions

The Pearson Efficacy Framework[52] provides a useful set of questions in order to prompt product teams to think through whether a product is likely to achieve efficacy.

- What are the efficacy goals that are to be achieved?
- What is the evidence behind these claims?
- How will these goals be met?
- Does the team have the capacity, in terms of knowledge, skills and relationships, in order to effectively meet these goals?

The questions should not be asked once, nor will there be a single set of right answers, instead these will need revisiting and revising as the product develops to ensure it stays on track. In particular attention needs to be paid to clearly articulating the efficacy goals, demonstrating the quality and trustworthiness of the evidence behind these claims, and setting out a well thought through action plan to show how goals will be met and evaluated. Finally it will be important to identify the team's individual and collective capacities, and any professional development needs and supporting resources required.

1.2. Evidence across the innovation lifecycle

There are three overlapping roles that evidence plays in the development stages of products and services.

Firstly, before commencing development, companies should be aware of the synthesised knowledge base relevant to their field of activity and apply it to product or service design. Innovators should conduct an evidence audit to ensure they engage with and apply the best of what is already known about effective learning and teaching most relevant to the product or output concerned. So that, in short, *"innovation can stand on the shoulders of previous progress."*[53]

This evidence can come from traditional research sources, such as academia, but it can also be combined with corporate metrics, customer feedback and awareness of other competitors in the field. All of these types of evidence should be thought of as prompts. Sometimes very creative ideas will go well beyond any existing evidence – but still warrant being developed and later tested. Innovation involves novelty – and by definition new things won't have evidence of impact to support them. Overly rigid requirements to follow evidence can stifle innovation at the early stage – though it's always valuable to be acutely aware of what the existing evidence says. A balance then needs to be struck between being considerate and thoughtful, but open to experimentation and new discoveries.[54]

Secondly, companies should continue to invest in appropriate R&D, that is to say measure, analyse, learn and improve the design and development of products and services as they move from being an idea to a working prototype. During these phases ideas often evolve rapidly, with decisions to either pivot or preserve based on this cycle of testing and iteration.[55] Some very promising ideas may turn out not to work at all; others require tweaking or adaptation. Good judgement is needed as well as good measurement, since some ideas that later turn out to be very successful may falter on their first outings.

Thirdly, there is a need to continue to collate evidence of impact over time as products and services grow and scale, with an aim of providing robust evidence to demonstrate the efficacy of impact over time and across differing contexts. The more mature the idea, and the greater the scale on which it is being implemented, the greater the need for a rigorous analysis of evidence.

What this evidence looks like, and the methods used to generate it, will vary at different stages of the innovation journey, between companies and across different settings. These are key points we will return to.

1.3 Moving beyond a test of pass and fail

Assessments of educational impact or efficacy have traditionally focused on making pass/fail judgments as to whether a product or service leads to a significant impact on the intended outcomes trial. This is often summarised as the 'what works' question.

The rhetoric of finding 'what works' is maturing, with an increasing recognition of needing to go further, to understand what is working now, for whom, when and why.[56] It is therefore important to understand where the product or service has been developed, tested and shown to be effective, both in terms of geographies, user group, and also the specific learning context. For instance, research undertaken in a school may tell us a lot about what worked in that context, but less about whether the same impacts will be seen elsewhere, such as in the home or other out of school contexts, an issue known as external validity.[57]

Related to this is the need to understand how much a product or service can be changed and altered by users and still have a positive impact. This is often referred to as adhering with 'fidelity' to the original model.[58]

1.4. What counts as good evidence

The question of what counts as good evidence of impact in social sciences is highly contentious. Some argue that there are hierarchies of evidence, with some research methods being inherently better than others.[59] For interventions, RCT and quasi experiments are regarded as the gold standard for good reasons in intervention research, but they are not always possible, desirable or ethical (parents increasingly accustomed to choice in education are unlikely to agree to random allocation of children to particular pre–schools or schools and consent may be problematic in interventions). Many questions of interest may not be about a specific intervention. Our view is that what counts as high quality evidence, *"depends on what we want to know, for what purposes, and in what contexts we envisage that evidence being used."*[60]

This requires less concern with the prioritisation of certain research methods, and instead collecting appropriate information using well tried and respected methods of analysis appropriate to the task, helping to bring greater focus on the usefulness of evidence in understanding learning and making decisions for further improvement. This is where standards of evidence can provide a useful framework.

2. Standards of Evidence

As we outlined in the earlier sections, there is a lack of clarity as to what good evidence looks like. Many organisations use terms like 'top tier' or 'proven' to classify programmes and help decision makers to select those interventions that are deemed to be working. These commonly draw upon studies where the intervention has been evaluated using random assignment. This can be useful in fields where there is a strong evidence base of well conducted RCTs, however, there are many fields where these types of evaluations may be lacking. Equally, there are instances where this method isn't appropriate, such as when the intervention is at an early stage of development, is localised or involves a small sample size. This can make it hard to judge and compare innovations which have alternative types of evidence appropriate to their differing stages of development.

Standards and scales of evidence that enable different types of evidence to be considered can therefore be exceptionally useful in helping to judge the strength of different products and services.

Nesta has adopted a Standards of Evidence framework to guide the Nesta Impact Investment Fund,[61] and to build the evidence behind a range of their innovation programmes. The advantages of the framework include enabling a diverse range of products and services to be considered and evaluated in an appropriate way, and through providing a common language for communicating impacts.

In Figure 2 we outline an amended version of the Nesta Standards of Evidence, modified for use in the education field. These standards of evidence distil the insights from many other frameworks that have been developed over the years (see the textbox). It doesn't pretend to be definitive, but it hopefully balances simplicity and sophistication in a way that can provide a common language for discussing the evidence of impact for a particular product, service, intervention or programme.

Project Oracle and the development of the Nesta Standards of Evidence

The Standards of Evidence were originally developed for use by Project Oracle, a Greater London Authority programme to support the evaluation of youth services across London. Nesta then adapted these further to help guide the decisions in Nesta Impact Investments, as well as across a number of Nesta's other innovation programmes.

When developing Nesta Impact Investments we needed a robust impact measurement framework, yet we could not find anything appropriate in the venture world. Nesta's first starting point was to adapt and adopt the Project Oracle Standards of Evidence, selected in recognition that the bar for evidence needs to balance against ensuring that innovation can flourish, and that the bar is not set so high that it becomes an unwieldy millstone of hindrance to growth and development. The Standards of Evidence are therefore tiered and incremental, helping grow the evidence behind products and services at a rate appropriate to them.

For further details about Project Oracle see: **www.project-oracle.com**

The following table shows the **Nesta Standards of Evidence**, adapted to capture education impact. What is expected at each stage is clear; however how the evidence is generated to meet these criteria is deliberately flexible and open. This means that the research methods at each stage are illustrative, not exhaustive. Well justified, alternatives that produce the required information are also considered.

Figure 2

Level	What we'd like to see	Suggested Method	Criteria to be met (Quality – how certain are we that it works?)
1	A justification for why the product/ service could have an impact on specified educational outcomes.	A clearly articulated account of impact for why the service/product could have impact on educational outcomes, and why that would be an improvement on the current situation.	There is a clear rationale to show why the product/service could have a positive impact on an educational outcome The key elements required: • A description of the product/service. • An explanation for how it could positively impact on one (or more) specified educational outcomes. • An explanation of how the educational outcome could be measured. The description will include the context in which the product/service operates, specific target populations, and recruitment and referral processes of these target populations, and clear documentation about what participants receive (at Level 4 this becomes an understanding of how it is delivered). At this stage there wouldn't be an expectation for impact data about the product/service; however, there would be an expectation that the product/service is situated in any available benchmark information and data, for instance data about the problem to be tackled, information about similar initiatives being developed etc.
2	Data has been gathered to show some change amongst those using the service/product.	Pre and post-survey evaluation; regular interval surveying, panel study, as well as other well-conducted methods.	At Level 2 data will show that there is a change in the measure of the educational outcome among the recipients of the product or service. At this stage, data can begin to show effect but it will not demonstrate direct causality.
3	Data shows an impact is happening because of the product/ service, whilst also demonstrating less impact amongst those who don't receive the product/ service.	Robust methods using a control group (or another well justified method) that isolates the impact of the product/ service. Random selection of participants strengthens evidence at this level, however, there is a need to have a sufficiently large sample at hand (scale is important in this case).	At Level 3 the evaluation design would involve using a comparison group to isolate the educational impacts of the product/service. The objective of Level 3 is to have confidence that the product/service is contributing to the impact that has been measured. Ideally at Level 3 a randomised control trial (RCT) would be used with at least one long-term follow up of the outcomes (however, in some instances an RCT is not appropriate so well justified alternatives would be considered). All products/services at Level 3 will be well documented, with necessary skills, training – and other delivery requirements – outlined clearly, to enable effective replication in alternative places, situations, and contexts.

4	Clear explanation as to why and how the service/product is having the impact observed and evidenced so far. An independent evaluation validates the impact observed. In addition, at Level 4 the product/service delivers impact at a reasonable cost, suggesting that it could be replicated and purchased in multiple locations.	At this stage, a robust independent evaluation that investigates and validates the nature of the impact is required. This might include endorsement via commercial standards, industry kitemarks etc. Documentation is needed to standardise delivery and processes. In addition, data on costs of production and acceptable price point for customers is also expected.	There will be standardised product/service documentation and processes, to show what is delivered, how it is delivered/produced, and the costs, so that if needed, the product/service could be produced by a third party successfully and get the same impacts. A high quality, independent evaluation/validation exercise that clearly shows the impact the product/service is having. There will be strong understanding of the market, and evidence that the cost of delivery matches what potential purchasers would be willing to pay for the product/service. This could include cost–benefit or cost–effectiveness analysis.
5	Demonstrable evidence that the product/service could be operated up by someone else, somewhere else and scaled up, whilst continuing to have positive and direct impact on the outcomes, and remaining a financially viable proposition.	Methods could include multiple replication evaluations; future scenario analysis; fidelity evaluation.	There will be multiple evaluations of the product/service in different settings (at least two evaluations; one of which has been independently undertaken) to demonstrate that the product/service can be used in different settings (which could be in different settings geographically and/or with different types of product/service users). For a service, it will also be proven that it can be delivered by different staff. There will be detailed findings about 'dosage', for instance, does giving more of the product or service create a greater impact; and targeting, for instance, are the same results found when the product/service is used in different areas or communities?

Source: Puttick, R. and Ludlow, J. (2012) 'Standards of Evidence for Impact Investing.' London: Nesta. Note: These are an amended and edited version of the 'Standards for London', which underpin Project Oracle, a Greater London Authority programme.[62]

An evidence continuum

The Standards of Evidence are on a 1 to 5 scale, with Level 1 being the starting point, moving up to Level 5 where a product or service has very strong evidence of effectiveness. The standards set out a continuum of gathering evidence that is appropriate for different stages of development, and the likely resources for evaluation that are available. As illustrated in Figure 2, the Standards effectively map out an evidence journey from innovative new ideas, to proven models operating at large scale.

Level 1 represents a low threshold, appropriate to very early–stage innovations, which may still be only at the idea stage, involving little more than a clear articulation of why the intervention is needed, what educational outcomes it will aim to improve, and why this could be better than what currently happens. These arguments need to be made with reference to the existing knowledge base.

As the levels are progressed, it will be expected that data is collected to isolate the impact of the product or service, that findings are independently validated, and then at Level 5, we expect to see demonstrable evidence that the product or the service can be delivered at multiple locations and still deliver a strong, positive impact. In other words, the positive impact is scalable. As products and services move up the Standards of Evidence and continue to have a positive impact, our certainty and confidence in them will also increase.

Of course, not all innovations will be successful, for instance some may produce weak results at an early stage, which means development will cease before they reach Level 5. Equally, some innovations may reach a scale associated with Level 5, but without the accompanying evidence of impact. Standards of evidence are a framework to assess how confident we are in the impact of a product or service, helping inform its development; yet evidence is only one facet of the decision making process. People may choose to buy or adopt without knowing fully whether it will work or not.

Generating evidence throughout the innovation lifecycle

As we outlined above, the Standards of Evidence framework can help to ensure that evidence is generated in a staged way, that is both appropriate to the lifecycle of products and services, and that it helps developers better understand what and when to measure, and how to approach this. At each stage the most appropriate methods can be selected.

At Level 1 innovators start with crafting a logic model (see Appendix for further details), which very early–stage innovations may be just an idea, with the logic model setting what the target outcomes are, and how the product or service could achieve these. This is the theoretical foundation for any product or service; ensuring that the principles of effective learning – identified as those appropriate for systems, organisations, professionals or students – are systematically 'baked' into the design. At this stage innovators should engage with the existing knowledge base in their area of learning or teaching. For example, for those innovators developing products to support effective teaching John Hattie's meta–analyses of effective pedagogy would be a useful place to start.[63] The Pupil Premium Toolkit is another source of innovation. Or for those developing digital learning technologies, the Education Endowment Foundation's recent meta–analyses of digital technologies[64] or Luckin et al.'s, review and proposed eight effective learning themes would be directly relevant.[65] These eight themes are:[66]

- Learning from experts.
- Learning with others.
- Learning through making.
- Learning through exploring.
- Learning through inquiry.
- Learning through practising.
- Learning from assessment.
- Learning in and across settings.

As the product or service turns from an idea to a prototype, it has reached Level 2 and can start to be tested, with the logic model or theory of change work at Level 1 forming the foundations for how this will be structured. At Level 2, correlation is sought. By this we mean that a positive impact is being seen when the product or service is being used, but we cannot isolate this impact to it. Methods that could be selected here include pre and post surveys, whereby a baseline is taken at the start and then follow-up data is collected after the intervention to assess any change. It may be possible to compare change to external benchmarks, for example in terms of standardised reading or maths tests and students gains in months.

If positive findings are found at Level 2, then the product would move up to be tested with more in-depth methods. It is at this point that we are seeking plausible evidence of potential causality between the product and the associated outcomes, whereby we isolate the impacts seen with the intervention being deployed. Methods at this stage could involve randomised control trials, or another well justified method.

As the higher levels are reached, the evaluations used will seek to ensure that these impacts are replicable in different contexts and settings. This may require businesses to adapt products to fit certain contexts, such as tailoring to suit the language or cultural conditions, but whilst ensuring positive impacts are maintained.

Even if a product or service does reach Level 5 we do not see an end to the evidence journey. Evidence at all stages may only ever be partial and time bound, and needs to therefore be used as part of an ongoing process of learning.[67]

At each stage of the Standards of Evidence it is imperative that the evaluation design is tailored to fit the product or service, and the method is appropriate, meaning that product and service developers take a bespoke approach each time. In essence: not selecting the method before defining the question; ensuring methods are appropriate for the product or service and its stage of development; to not impede or hinder its development.

3. How innovators can use Standards of Evidence

In this section we outline some considerations to aid innovators in the adaptation and tailoring of the framework. They are intended to help set innovators off in the right direction of travel, steering them through the design and delivery of evidence collation, rather than being overly prescriptive. We also provide two examples of how the standards could support specific products and services.

Clearly define intended learning outcomes

Each product or service will be different in terms of the target educational outcomes they seek to impact, for instance, enrolment and attendance rates, standardised test scores, progression into further education, work or as part of professional learning. In addition, some products and services focus upon impacts on individual pupils, teachers or parents whilst others might focus at the school or system level. The selection of these will influence the types of data points and the timeline for seeing an impact. Innovators need to clearly define the outcome(s) the product or services is seeking to improve and specify how each will be measured and when. This is a requirement at Level 1. The next four levels (Levels 2–5) of the Standards involve gathering increasingly robust evidence to help evaluate whether the product or service is having a positive impact on these defined outcomes within a specified timeframe.

Draw on a range of sources of evidence

It's worth noting the decades of debate and contention surrounding different research methodologies, and their perceived appropriateness and strengths, is indicative of the fact that there is no single 'correct' way to evaluate and capture impact. Yet reference should be made to existing guides, such as the *Handbook of Evaluation* by Shaw, Green and Mark, which outlines the criteria for high quality evaluation, and helps aid the selection of appropriate methodologies.[68]

We do not start with the assumption that some methods are intrinsically superior to other methods of evaluation. Rather the most appropriate method(s) for evaluating a specific product or service in a particular context will vary depending on how to best maximise valid and reliable information about learning outcomes in a way that is practical and cost–effective. Validity is important, it refers to whether an evaluation can suitability answer the questions intended, which helps determine the types of methods to be used to adequately answer the question set. Independent of validity, but still of importance is reliability, which is another way of saying consistent, in which we mean if a test is repeated multiple times that the same results will be seen.

Primary research of educational impact generated through the standards of evidence could be combined with other sources of existing data analytics, management information, and other existing data collection processes, such as consumer feedback and complaints.[69] In all fields the strongest insights come from the use of multiple types of evidence. The key point is to triangulate findings, to orchestrate and access the most useful evidence in order to build up a rich understanding of whether effective learning is occurring, for who and under what conditions.

The aim should always be to draw on a range of different sources of evidence and assess whether they converge. For example an innovator seeking to improve mathematics competencies for primary schools through a digital platform may gather information from usage patterns, standardised tests, and surveys of student engagement. If convergence of data doesn't occur, a diagnostic approach to understanding the causes of the perceived discrepancies should result in a deeper awareness of what is working, for who and why.

Task responsibility for evaluation

Responsibility has to be tasked for setting the evaluation design and strategy for products and services, and assessing them against the Standards of Evidence framework. If the evaluations are undertaken by staff internally it could help ensure continuity, and preserve commercial sensitivities; but external verification brings extra credibility, and possibly reduces bias from creeping in. This is particularly important at the higher levels of the standards.

As the Standards of Evidence will often be applied to products already being used in the field, how close should those responsible for evaluation be to the product development team and sales team? Too far and they could lose the potential to inform and optimise development in real time, yet if too close their findings and analysis could lose independent objectivity.

Related to this is who in the company listens to any concerns and critiques raised by researchers? Who do they report to? Who should be responsible for sense checking and responding to their findings? Business leaders will need to determine what endorsement or verification is needed in order to document where a product or service is on its evidence journey.

Encourage progression through the levels

How quickly should products or services move through the Standards of Evidence? We have already outlined that the evaluation should be designed to be relevant to the individual product or service, but there is a need to consider how continual progression through the levels is encouraged. For instance, if all companies can move through Levels 1 and 2 relatively quickly should a time limit be set to encourage rapid progression to the higher levels?

Define failure

In the quest for finding what is working, there is a need to agree upon a definition of failure. When should you persevere with developing, adapting, and seeking more definitive conclusions about a product, and when should you pull the plug entirely? And who decides on this?[70] This is complicated further by research rarely providing conclusive results.

Yet, the Standards of Evidence framework should help to ensure that these decisions are more likely to be made at the earlier stages – Levels 1, 2 and 3 – and thus minimises the risk that an ineffective product or service is taking to scale without sufficient evidence of impact.

The products or services that do not rise through the standards, despite attempts at improvement should be viewed as opportunities for developing a deeper understanding of what doesn't work well enough, for who and under what circumstances.

The design principles outlined here will help guide how the framework could be adopted within an education business. But it is worth highlighting that the Standards of Evidence for a business context are a new application of the framework, and therefore will need iteration and change to ensure they embed and are appropriate. In the same way that generating evidence should be an ongoing process in the product or service's lifecycle, so too should a company's approaches to generating this in the first place, helping ensure it utilises the most effective research methods and tools.

In the final section of this paper we will outline how the Standards of Evidence could be applied by others in the wider education ecosystems, such as venture capitalists and startup developers.

SECTION 3: SHIFTING THE FIELD – THE USE OF THE STANDARDS OF EVIDENCE ACROSS EDUCATION

We are witnessing an influx of new ideas, talent and capital into the field of education business. In order to capitalise on this potential, effective approaches to evidence and evaluation are needed in order to select, iteratively improve and scale the most effective innovations. To become a truly evidence–led organisation, with evidence of impact incorporated into product and services development, any startup or established company – within education or elsewhere – will need to undergo a strategic shift in ways of working.

In this section we outline a number of ways in which education businesses, entrepreneurs and investors could adopt and operationalise the proposed approach and the Standards of Evidence framework.

How a large established organisation with an extensive product range is embedding 'efficacy' to improve learners' lives

Education companies have a responsibility to make a positive impact on its learners, yet their contribution is not often well understood. Pearson, a large education company which publishes textbooks, produces online learning programmes and operates entire institutions, has recently set out their commitment to measure, report and improve on learning outcomes. We reflect on their journey, as told in *An Incomplete Guide to Delivering Learning Outcomes,* and how the Standards of Evidence might help to improve outcomes across a major education business.

Pearson is an established organisation managing a large portfolio of education goods, from e-learning tools to direct delivery offerings. In this context, each product and service differs in its intended education outcomes. As a result, there was historically no common language across the company to understand and discuss what counts as quality evidence of impact at different stages of the product lifecycle.

To meet this challenge, Pearson developed the efficacy framework (see Foreword). The framework was designed to embed standards of evidence that fit the diversity of its products and intended outcomes in different international contexts, whilst also realising the need to ensure consistent approaches to evidence across their portfolio. Whilst Pearson will not be able to demonstrate the evidence for impact overnight, a number of practical steps are being taken towards promoting a more explicit evidence-based approach – the 'path to efficacy'.

These standards cannot just be applied to the outcomes, e.g. students' performance. In order to fully embrace efficacy, the standards are built into the R&D process to enable employees to better understand how to move products and services from ideation to scaling through the use of evidence. During the idea stage, developers should be able to demonstrate that they have met the criteria for Level 1 before receiving investment capital. During the prototyping stage developers should be required to generate evidence of impact that meet Level 2 and potentially Level 3 before being able to move into market implementation. At stages of early

and late implementation, comprehensive plans should be in place to move up the ladder of standards whilst applying lessons generated from evaluations to continuous adaptation and improvement. Systemically applying the standards in this way to the development of new products and services will ensure the next generation of company outputs are evidence–based.

All current products and services across the portfolio are being rated against the standards by reviewing the existing evidence base for products and services in a systematic way. The evidence base is an ever-expanding network of information helped by the increasing sophistication of technology to capture data and knowledge about pedagogy. Across the portfolio it is likely that some products will already meet the standards of Level 3 or even 4. It is also possible that a large number or products and services in the market may not yet meet the desired criteria, but wherever the current state, teams should then be required to set out plans to move up the standards over a certain period. Products or services will go under continuous and extensive review and as a result, the potential for sun-setting may be considered. To mobilise Pearson's staff during this process, incentives and KPIs for leaders across the company should be linked to implementing plans to move up the Standards of Evidence across the portfolio. Incentives should be linked to the quality of the evidence generated and its effective use for improvement, rather than whether a certain level has been reached. This is important as it focuses the work on effective organisational processes and avoids the tendency to 'game' set targets.

These standards are going to be applied throughout all of Pearson's activities, including assessing acquisitions. When companies grow through acquisition, decisions are typically made based on a criterion that assesses the health of the target company and the revenue–generating potential of the venture. This 'evidence for impact' category needs to be included in all acquisition decisions, which determines which Level of the Standards of Evidence the potential target has reached and how they can be improved.

Embarking on this journey will involve a significant resource commitment. As the central goal of Pearson is to make a demonstrable impact on learning outcomes, the company leaders will need to invest in new processes and data collection studies to provide clear justification for the choices that the company makes in the future about its investments, partnerships, and product development. Furthermore, there needs to be an investment in developing 'evidence literacy' across the business, so that leaders at all levels of the business can understand evidence and data. This may include the recruitment of new staff with different skill sets as well as investing in professional development to support a culture shift that puts efficacy at the heart of the company's core business.

There needs to be commitment to systematically apply research evidence to inform design, development and implementation decisions. Yet, simply sharing academic research reports with product or services leaders is unlikely to result in effective evidence–informed decision making. Knowledge mobilisation strategies need to be employed to support the effective engagement and use of the evidence of impact that is being generated.

How Standards of Evidence can be used by education startups and education accelerators

Educational accelerators have a powerful role to play in supporting educational entrepreneurs develop an evidence–based approach. One such accelerator is MaRS, a Canadian innovation agency established to capture the commercial potential of Toronto's $1 billion annual science and technology research spending. Here Krista Jones, MaRS Practice Lead for Education Innovation, discusses how MaRS and the educational entrepreneurs could adopt the Standards of Evidence in their work.

In our experience entrepreneurs in educational startups rarely consider generating evidence for impact. They can typically clearly articulate the problem they are solving for students or teachers, the potential addressable market and the engagement of their current user base. Yet, these same entrepreneurs often struggle to state the educational outcomes they aim to improve, and cannot outline plans for how to generate evidence that they are making a positive impact. The assumption seems to be that if you have good intentions and a happy user–base then learning is bound to follow.

The strength of the Standards of Evidence approach is that it allows entrepreneurs to begin and use the 'evidence journey' in a way that is appropriate to their stage of development and proportionate to the funds available. Entrepreneurs should be able to meet Levels 1 and 2 without the need for outside research expertise. It is at these stages they should begin to look at other similar innovations and establish the parameters for true impact evaluation once they reach Level 3. They need to think of user adoption and engagement as two of the key elements required to create a successful education product or service and remember that they are not sufficient to determine impact. Rather than seeing these foundational evidence stages as additional requirements of evaluation that sit outside of the design and prototyping process, they need to reframe their thinking comparing this impact evidence to the ROI studies that business process innovations have been doing for years. The evidence of impact should be gathered, analysed and used to iterate on the design as part of the normal 'validated learning cycles' of a lean approach.

During scale–up, entrepreneurs should work in partnership with a research institution to help embed evaluation into the ongoing development and improvement cycles. **Stronger evidence of impact at Levels 3 and 4 will not only give a unique competitive advantage in the market, but will also make them a far superior acquisition target for larger companies or investment firms committed to efficacy.** Note that there will always be the natural tension between learning efficacy and product adoption, the best innovations will be those which master the art of adoption and provide a high level of efficacy. By the end of Level 2 a high potential educational venture should have created a measurement model that allows them to determine the optimal trade–off.

The rapid growth of educational accelerators such as Imagine K12 in Silicon Valley, Socratic labs in New York City, Learn Launch in Boston and MaRS in Toronto have a powerful role to play in supporting educational entrepreneurs develop an evidence–based approach. Achieving Level 1 – identifying target outcomes and outlining a logic model based on the literature – should be a prerequisite for gaining acceptance to any programme. In addition, understanding the standards of evidence, and how to generate evidence to achieve different levels should be a core component of the curriculum of any accelerator course.

Education entrepreneurs are seeking to solve two problems. Firstly, can they develop revenue (or at least committed use) through meeting a user's need? And secondly, can they show their product or service results in positive educational outcomes for that user? Which means users not only like using the product or service, but it works! Entrepreneurs are already taught how to map out a logical business model and how to iterate, based on rapid customer feedback. Educational entrepreneurs should also be equipped to outline a logical learning model and iterate based on evidence of impact on education outcomes.

The biggest gap that needs to be evaluated is the timeframe necessary to gather both types of information. Startups' most precious commodity is time; time to market, time to profit, time it takes to find product market fit. Often times the learning outcomes are not measureable

in the timeframe that the business of being a startup requires. In today's startup world there is an expectation that entrepreneurs will iterate themselves to a viable product with minimal cash, in an expedited manner. It will be important once again to manage the trade–offs that will inevitably be necessary.

When education entrepreneurs embark on pilots designed to test their products with real customers, they are often consumed with the logistics of the pilot and don't collect crucial evidence to test the products efficacy regarding learning outcomes. It is crucial that startups begin their pilot with a strong hypothesis (Level 1) and a specific course of action on how to test that hypothesis to approach Level 2. This will require identifying a limited number of measurable metrics to be measured over the limited course of the pilot so the entrepreneurs can identify patterns in learning and can measure the strength of their learning hypothesis. Comparison techniques with corporate training and self–directed learning will be key components of a successful measurement programme.

To move to Level 3 and beyond, accelerators could help broker partnerships with local universities and research institutions. Researchers, and particularly graduate students, are often comfortable setting up trials that control for one variable at a time and are used to the statistical analysis that makes efficacy data more relevant. These researchers could be embedded with startup teams to help develop and implement an evaluation plan to move up the Standards of Evidence. Embedding researchers into the process of product development could also play a role in developing research capacities across the startup team and this improves the ability of the company to effectively generate and apply evidence as it grows. As their product or service evolves through contact with a mass market, so will their impact on learning.

The Education Innovation team at MaRS (Krista Jones, Aron Solomon and Joseph Wilson) believe that *"an iterative approach to collecting data and measuring impact is necessary for startups to capture the impact they are having on students over short bursts of time and is vital for the long–term sustainability of education systems. The process of impact measurement will need to be accelerated and carefully integrated into the rapid iterative product development models of today's lean startups. By embracing a Standards of Evidence approach in the world of education startups, everyone from the makers of things to those who invest in them will have a more clearly–defined sense of what is worthy of time, attention, focus. For EDUpreneurs, this will become the necessary corollary to the Business Model Canvas and the logical answer to future inquiries as to whether an innovation not only fits but serves the market."*

How do Standards of Evidence work in the real world?

Nesta Impact Investments is a £25 million early–stage investment fund, launched in 2012. It is for ventures that achieve a positive impact in the fields of education, ageing, and community sustainability, with Standards of Evidence the core strategy for impact measurement. Here the fund's director, Joe Ludlow, comments on their experiences of using such an evidence framework.

There were two main observations that lead us to develop the Standards of Evidence For Impact Investing. Firstly, that most impact measurement in investing involves simply counting units of output by an organisation, which is often a weak proxy for effect on a social outcomes. And secondly, that in early–stage investing, where innovations are often still prototypes in development, growth in volume may tell you little or nothing about the progress a venture has made as a result of the investment.

What we learned from the Project Oracle approach was that the Standard of Evidence reached by a venture could be seen as a proxy for the 'impact risk', by this we mean the certainty that it will have a positive effect. Within our impact model is the idea that a successful impact investment is one that reduces its impact risk (or increases its Standard of Evidence), as well as being one which increase its volumes of output.

So far, so good, in concept. Now we have had nearly a year of applying this thinking in

practice, working with real companies seeking to have a positive impact, and make a profit. Whilst we remain convinced that increasing rigour of evaluation is an important dimension of improved impact performance, a number of practical issues are clear:

- Rarely outside of healthcare have we found companies familiar with evaluation approaches. Most companies are more familiar with talking about their reach or number of customers and volumes and perhaps perceived benefits and customer satisfaction.
- Companies design products for the widest possible reach, often with multiple features designed to deliver different outcomes for different users, which can complicate evaluation design.
- We learned early on that in some fields, like state education and health, administrative data is increasingly collected and made open, and that this in theory helps reduce the cost of causality studies. However, in practice there hasn't been an occasion where administrative data has provided appropriate data points for the interventions we are funding, requiring creative evaluation design.
- Setting up control groups in trials – where certain people don't receive the product or service – can be complicated; it's counter–intuitive for early–stage companies not to want to sell a product to everyone!
- The question of who should pay for the evaluation often arises – should it be the company, or should it be the impact investor? As a consequence, we are gaining a sense, intuitively, of what feels like cost–effective evaluation too – trying to balance what good quality means with what can be justified within the inevitable pressure on budgets.

This way of thinking about impact is not yet in the mainstream. Part of our challenge going forward is to communicate what we are doing, what we are finding and what that means for the investors in our fund, the customers of our portfolio companies and for the field of impact investing.

CONCLUDING COMMENTS

It's a statement of the obvious that education should draw on the world's stores of accumulated knowledge and acquired critical faculties. Yet as we've shown, a remarkable amount of education practice doesn't apply basic principles of good education to itself. It ignores the lessons of experience and evidence, and doesn't build in reflection and evaluation. The result is that ideas in education often spread more because they're appealing or convenient, than because they work.

We've shown here how evidence could become much more integrated into the day to day work of businesses. We've also shown that this work needs to start from clear principles: that any new method should be tested and improved in the light of experience; that what's already known should be orchestrated and made easily available to the busy people who have to make decisions; that a clearer understanding should be developed of what is meant by impact for different products and services and a more structured approach to establishing what works, so that what doesn't work can be ditched and what does be spread.

Governments and public bodies will continue to have a vital role in promoting and in part supporting the research and development to help establish what works. But business needs to be at the forefront too. The work Pearson is doing to embed efficacy represents a major breakthrough. It's part of a bigger shift in business towards greater accountability and a tougher commitment to excellence – and there are signs that the best firms recognise that they'll gain competitive advantage by being ahead of the curve.

Hopefully, too, the public will become more engaged as more intelligent purchasers and users. Ultimately we need evidence to be reinforced by consumer demand.

The drive to apply the principles of evidence and learning to education will take time to become embedded and to deliver results. But it goes with the grain of change – with a better informed public, more demanding requirements by governments for the investment of public monies in education, with new tools coming from data, and with the desire of professionals to do the best they can for learners. Our hope is that the ideas set out in this report will be adapted and adopted widely by as many as possible of the thousands of businesses involved in education – and that the ultimate result will be better education for many hundreds of millions of young people who deserve the best.

APPENDICES

1. Further reading

John Hattie – evidence studies in education	
Summary	John Hattie has written three seminal texts which highlight existing studies in education and what this evidence means. The 2009 meta study collates and compares evidence from more than 50,000 studies in education, covering more than 80 million students. In his follow up book *Visible learning for teachers* he reflects on the data in the context of the classroom and he primarily advocates that evidence is most useful when teachers also engage in evaluation and learning.
Link/reference	• Hattie, J. (2011) 'Visible learning for teachers: Maximizing impact on learning.' Abingdon: Routledge. • Hattie, J. and Anderman, E.M. (eds.) (2013) 'International guide to student achievement.' Abingdon: Routledge. • Hattie, J. (2009) 'Visible learning: A synthesis of over 800 meta-analyses relating to achievement.' Abingdon: Routledge.
An Introductory Guide to Implementation	
Summary	Though there is a growing understanding of 'what works' it is important to recognise that the challenge can be implementation of an evidence-based programme. CES aim to bridge this gap through tools and consultancy to services. This guide provides an overview of their implementation methodologies
Link/reference	Burke, K., Morris, K. and McGarrigle, L. (2012) 'An Introductory Guide to Implementation: Keep Terms, Concepts, Frameworks.' Dublin: Centre for Effective Services. www.effectiveservices.org/images/uploads/file/publications/Guide%20to%20implementation%20concepts%20and%20frameworks%20Final%20for%20web%20v2.pdf
Hack education – The Audrey Test – questions for ed-tech start ups	
Summary	Guiding questions for practitioners in their buying decisions.
Link/reference	Watters, A. (2012) 'The Audrey Test: Or, What Should Every Techie Know About Education?' hackeducation.com/2012/03/17/what-every-techie-should-know-about-education
How to Decide What Technology Fits in the Classroom	
Summary	A useful summary of questions to guide buyers when considering suitable education (teched) products and servcies.
Link/reference	Baresghian, T. (2012) 'How to Decide What Technology Fits in the Classroom.' www.pbs.org/mediashift/2012/10/how-to-decide-what-technology-fits-in-the-classroom300
	Evidence-based business?
Summary	Blog outlining why evidence is important and how businesses should be held just as accountable as government services if intended for public good and tied in with public services.
Link/reference	www.nesta.org.uk/blogs/geoffs_blog/evidence_based_business

Rise of the Datavores	
Summary	The vast availability of data makes it possible to experiment and measures in near real time. Digitisation of educational products and services makes it possible to quickly and accurately capture information on the learning outcomes. The internet can be *"a 'lab' for business experimentation, and in particular through the use of randomised controlled trials that can help businesses measure the impact of their activities more accurately, quickly and more cheaply than before."*
Link/reference	Bakhshi, H. and Mateos-Garcia J. (2012) 'Rise of the Datavores: How UK Businesses Analyse and Use Online Data.' London: Nesta. www.nesta.org.uk/library/documents/Datavores.pdf
Decoding Learning. The proof, promise and potential of digital education	
Summary	A lot of money has been invested in technology in education but little has been done to understand the impact upon education outcomes. The report looks at how technology has been used and aims to develop a better understanding of the impact on learning.
Link/reference	Luckin, R., Bligh, B., Manches, A., Ainsworth, S., Crook, C. and Noss, R. (2012) 'Decoding Learning. The proof, promise and potential of digital education.' London: Nesta. www.nesta.org.uk/library/documents/DecodingLearningReport_v12.pdf
Alive in the Swamp	
Summary	Report that looks at digital innovations and systems and how best to asses and implement. It found over the last five years digital information across the world has increased nine times, providing more opportunities to gather feedback, evaluated response and monitor learner outcomes.
Link/reference	Fullan, M. and Donnelly, K. (2013) 'Alive in the Swamp: Assessing digital innovations in education.' London: Nesta.
Managing change	
Summary	Highlights the multiple partners involved in delivering effective education and recognises the challenges of the debate between politicians and education practitioners.
Link/reference	Morris, E. (2012) Managing change – The relationship between education and politics. Article in 'Better Evidence.' online magazine: www.betterevidence.org/uk-edition/issue-9/managing-change-the-relationship-between-education-and-politics
Don't waste funds on 'neuro-myths', academic warns	
Summary	Sergio Della Sala, professor of human cognitive neuroscience at the University of Edinburgh stated that investing time and money on initiatives that have not been proven to work is *"a wasteful use of limited resources"*
Link/reference	Belgutay, J. (2013) Call for schools to beware 'neuro-myths'. 'TES' 24 May 2013 www.tes.co.uk/article.aspx?storycode=6336312
Ben Goldacre – Building Evidence into Education	
Summary	A paper examining how evidence can empower teachers and to help them identify the best strategies.
Link/reference	Goldacre, B. (2013) 'Building evidence into education.' Bad science series. London: Department for Education. dera.ioe.ac.uk/id/eprint/17530

2. Further information on methods

Logic model

- Logic models graphically illustrate programme components and clarify inputs, activities and outcomes, and have been defined as *"a systematic and visual way to present and share your understanding of the relationships among the resources you have to operate your programme, the activities you plan, and the changes you hope to achieve".*[71]
- Detailed guide to developing logic models from W.K. Kellogg Foundation.
- www.wkkf.org/knowledge-center/resources/2006/02/WK-Kellogg-Foundation-Logic-Model-Development-Guide.aspx

Randomised Controlled Trials (RCT)

- RCTs randomly assign participants to two groups, the 'treatment' group where individuals receive the intervention, or to the 'control' group where individuals do not receive the intervention. Comparing the impacts seen in the two groups enables the effectiveness of the intervention to be identified.
- For further details and to create your own RCT go to www.randomiseme.org

Standardised measurement instruments

The instruments are tools, such as questionnaires or other surveying methods that derive empirical results to demonstrate the impact on certain outcomes. If evidence is available, standardised instruments can offer greater relative confidence to the evaluator, and furthermore, may allow for comparisons between programmes.

As examples:

- Rosenberg's Self-Esteem Scale
  www.wwnorton.com/college/psych/psychsci/media/rosenberg.htm
- Goodman's Strengths and Difficulties Questionnaire
  www.sdqinfo.com
- Birchwood's Social Functioning Scale
  www.thorn-initiative.org.uk/ FileStore/Filetoupload,188455.en.docx
- Frick's Alabama Parenting Questionnaire
  cyfernetsearch.org/sites/default/files/InstrumentFiles/Alabama%20Parenting%20Questionnaire%20%28Parents%20of%20Children%206%20-%2018%29.pdf

3. Details of organisations and other sources of information

Center for Data Driven Reform in Education	
Theme	Framework and guidance to make it easier to evaluate education initiative.
Summary	CDDRE highlight the importance of research and desire for reform to generate more evidence in education. Provide a step-by-step research guide for teachers to implement an evaluation.
Link/reference	www.cddre.org/solutions/reform.html Slavin, R. E. (2012) 'Educational psychology: Theory into practice (10th Ed.).' Boston: Allyn & Bacon. Slavin, R.E., Madden, N.A., Chambers, B., and Haxby, B. (2009) 'Two million children: Success for All.' Thousand Oaks, CA: Corwin.
Top Hat Monocle	
Summary	Application for professors making it easier to interact with students – also allows the opportunity to use the data from the interactions to assess outcomes (results) including the ability to track how your class is doing over time.
Link/reference	www.tophat.com/tour
Education Endowment Foundation – toolkit	
Theme	A decision-making guide for buyers.
Summary	Aims to open up evaluation to a wider audience and make it easier for practitioners/policymakers to know what to invest in and how to evaluate. An overview of 30 topics in education practices and with a summary of the evidence to support it now used by a third of school leaders. EEF have also designed a DIY evaluation tool kit which takes teachers through the research process step-by-step and gives them some tips on what to look for.
Link/reference	(1) educationendowmentfoundation.org.uk/toolkit/about-the-toolkit (2) educationendowmentfoundation.org.uk/news/a-third-of-school-leaders-using-the-toolkit (3) educationendowmentfoundation.org.uk/uploads/pdf/EEF_DIY_Evaluation_Guide_(2013).pdf
Centre for Effective Services	
Theme	To enhance evidence informed service by promoting cross collaboration, capacity building in the 'real world'.
Summary	Aims to connect design and delivery of services with scientific/technical knowledge of what works.
Link/reference	www.effectiveservices.org
Blueprints for Healthy Youth Development	
Theme	Increasing availability of what works to practitioners/policymakers. Grading of programmes based on evidence.
Summary	Provides research and guidance on what has been proven to work in youth development. The review includes project information, developmental stage – ages 0–2 etc., target population, funding strategies, rating for evidence, cost-benefit analysis and the intended impact e.g. early cognitive development, positive relationships with peers. Applies 'standards' requires a minimum of one RCT and two quasi-experimental evaluations to be considers as 'promising' and two RCTs and 12 months of positive intervention after programme ends.
Link/reference	www.blueprintsprograms.com

Rate My Teacher	
Theme	Increasing consumer demand for evidence of impact in education.
Summary	A rating site for teachers across a range of education providers, primarily based in the US – rating based on – location, easiness, helpfulness, clarity, with the option to add comments.
Link/reference	www.ratemyteachers.com www.ratemyteachers.com/cindy-bond/28839-t

Evidence Informed Policy and Practice in Education in Europe (EIPPEE)	
Theme	Evidence in education increasingly prominent in policymaking.
Summary	The EIPPEE Network is an international network of individuals and organisations interested in evidence-informed policy and practice in education. EIPPEE is a two-year project, from March 2011 to August 2013. The project aims to increase the use of evidence to inform decision making in education policy and practice across Europe.
Link/reference	www.eippee.eu/cms/Default.aspx?tabid=3179

Institute for Effective Education – University of York	
Theme	Increasing evidence studies and opening this up to a wider audience.
Summary	Evaluate programmes and practices across a range of education settings. IEE recognise that research may be of differing qualities. • **Better: Evidence-based Education magazine** They have created an education magazine to provide examples of what works in education in a simple straight forward manner to open up the evidence to policymakers and educational leaders. • **Best Evidence Encyclopaedia:** an education 'encyclopaedia' to provide examples of what works in education programmes/practices/products in a simple straightforward manner to open up the evidence to teachers, policymakers and children.
Link/reference	www.bestevidence.org.uk www.betterevidence.org www.york.ac.uk/iee

Key Information Sets	
Theme	Increasing transparency in education, students are increasingly savvy about the educational choices they make and there is an increasing demand for evidence to make these decisions.
Summary	Key Information Sets (KIS) are comparable sets of information about full or part-time undergraduate courses and are designed to meet the information needs of prospective students. (This includes reliably gathered information on employment and accreditation; student satisfaction, cost and accommodation; study information; entry information).
Link/reference	www.hefce.ac.uk/whatwedo/lt/publicinfo/kis
Related links	Example of an evaluation: unistats.direct.gov.uk/Subjects/Overview/10007783-LL73/ReturnTo/Subjects

Graphite – digital tool evaluator app created by the Gates Foundation	
Theme	Making it easier for teachers to navigate the market place of digital tools.
Summary	Graphite is a free web service that rates and reviews digital learning resources, making it easier for teachers to navigate the marketplace of digital tools. Graphite recognises the role of the teacher and context in understating whether the tool will work in their classroom by building a community of teachers to use the products and share their feedback.
Link/reference	(1) Bill Gates, 'A New Way to Connect Teachers With Technology.' blog, June 24th 2013 siliconsurvival.wordpress.com/2013/06/24/bill-gates-on-connecting-teachers-with-technology (2) Main website: www.graphite.org

World Bank – Strategic Impact Evaluation Fund	
Theme	Importance of evidence at policy level.
Summary	The fund was set up to generate evidence of what works in key development areas to inform future policymaking. The Strategic Impact Evaluation Fund will run from 2012–2018 and aims to generate cross-country evidence of innovative programmes to inform international policy addressing health, education and water supply/sanitation. The World Bank has also supported two impact evaluation funds to grow the evidence base of programmes that have been invested in. The Spanish Impact Evaluation Fund a $14 million fund which was closed in in 2012 evaluated innovative programmes in health, education and social assistance – providing funding, training (17 field-based evaluation workshops) and sharing of knowledge. The fund focused on priority programmes designed to improve human outcomes with three types of evaluations (Cluster Fund, Innovation Fund and Quick Wins Fund). The World Bank celebrate case studies of what works in their monthly note series – *From Evidence to Policy* – outlining case studies of what has been proven to work in areas of health, education and other related human development measures.
Link/reference	(1) web.worldbank.org/WBSITE/EXTERNAL/EXTABOUTUS/ORGANIZATION/EXTHDNETWORK/EXTHDOFFICE/0,,contentMDK:22383030~menuPK:6508083~pagePK:64168445~piPK:64168309~theSitePK:5485727,00.html (2) web.worldbank.org/WBSITE/EXTERNAL/EXTABOUTUS/ORGANIZATION/EXTHDNETWORK/EXTHDOFFICE/0,,contentMDK:23177912~pagePK:64168445~piPK:64168309~theSitePK:5485727,00.html

Evidence for Policy and Practice Information and Coordinating (EPPI) Centre	
Theme	Recognise the importance of evidence in education – making it easier to understand and access for policymakers and practitioners.
Summary	The EPPI-Centre aims to promote evidence informed policy and practice and conducts systematic reviews across a range of subjects (including education) providing information on the research and evidence of a topic e.g. impact of teaching assistants in the classroom. Part of the Social Sciences research department at the Institute of Education.
Link/reference	eppi.ioe.ac.uk/cms/Default.aspx?tabid=473

Iterative Best Evidence Synthesis	
Theme	Improve transparency of what works in teaching.
Summary	Aims to provide trustworthy evidence on what works in education to improve outcomes for diverse learners, to stimulate continuous improvement in education. Prioritisation of research is based on approaches which can best accelerate outcomes for all diverse learners. Created in response to ***demand from teachers***
Link/reference	www.educationcounts.govt.nz/__data/assets/pdf_file/0020/102539/BES-Introductory-Flyer.pdf www.educationcounts.govt.nz/__data/assets/pdf_file/0014/108023/Evidence-for-Improvement-and-Public-Good.pdf
The Centre for the Use of Research and Evidence in Education (CUREE)	
Theme	CUREE aim to *"help teachers make informed decisions about the most effective and efficient approaches to use in their own context."*
Summary	CUREE is an internationally acknowledged centre of expertise in evidence-based practice in all sectors of education. The staff use their knowledge and skills in teaching, research, communications and knowledge management to produce high quality research, CPD and tools and resources. They provide advice on cost-effectiveness and a range of user-friendly research. They also review and appraise evidence to ensure it is 'useful'.
Link	www.curee.co.uk

ENDNOTES

1. Ries, E. (2011) 'The Lean Startup.' London: Penguin.
2. Roberts, K., Radjou, N. and Prabhu, J. (2012) 'Jugaad Innovation: Think Frugal, Be Flexible, Generate Breakthrough Growth.' Jossey-Bass.
3. www.oecdeducationtoday.blogspot.ca/2012/11/private-vs-public-expenditure.html
4. www.tes.co.uk/article.aspx?storycode=6358755
5. The Standards of Evidence framework is based on those developed for Project Oracle. For further details see: project-oracle.com
6. Puttick, R. and Ludlow, J. (201) 'Standards of Evidence for Impact Investing.' London: Nesta.
7. Pearson defines a 'product' as all that they or an external party "produces to sell to customers – content, technology, service and commercial elements which are combined to meet learner and customer needs". As cited Barber, M. and Rizvi, S. (2013) 'The incomplete guide to delivering learning outcomes.' London: Pearson. Available at: www.efficacy.pearson.com
8. As cited in Barber, M. and Rizvi, S. (2013) 'The incomplete guide to delivering learning outcomes.' London: Pearson. Available at: www.efficacy.pearson.com
9. Puttick, R. (2011) 'Ten steps to transform the use of evidence.' Blog, available online: www.nesta.org.uk/library/documents/TenStepsBlog.pdf
10. Nutley, S., Powell, A. and Davies, H. (2013) 'What Counts as Good Evidence?' London: Alliance for Useful Evidence. Available at: www.alliance4usefulevidence.org/assets/What-Counts-as-Good-Evidence-WEB.pdf
11. Andrews, T. and Knaan, I. (2011) 'Scared Straight: Don't Believe the Hype (Facts from Coalition for Juvenile Justice).' Blog, available at: www.reclaimingfutures.org/blog/juvenile-justice-reform-Scared-Straight-Facts-vs-Hype
12. Washington State Institute for Public Policy (2007) 'Evidence-based Juvenile Offender Programs: Program Description, Quality Assurance, and Cost.' Available at: www.wsipp.wa.gov/rptfiles/07-06-1201.pdf
13. Wilson, V. (2007) 'Reducing class size: Does the evidence support the Scottish Executive's policy?' Available at: www.scotedreview.org.uk/pdf/187.pdf
14. Mo, D., Swinnen, J., Zhang, L., Yi, H., Qu,Q., Boswell, M. and Rozelle, S.(March 2013) 'Can one laptop per child reduce the digital divide and educational gap? Evidence from a randomized experiment in migrant schools in Beijing.' Working Paper 233. Available at: www.iis-db.stanford.edu/pubs/23675/olpc_paper_March_31_2012_Web.pdf
15. The Alliance for Useful Evidence, details available at: www.alliance4usefulevidence.org
16. Mulgan, G. and Puttick, R. (2013) 'Making evidence useful: The case for new institutions.' London: Nesta. Available at: www.nesta.org.uk/library/documents/MakingEvidenceUseful.pdf
17. Evidence for Policy and Practice Information and Co-ordinating Centre, for further details see: www.eppi.ioe.ac.uk/cms
18. www.york.ac.uk/iee
19. www.qub.ac.uk/research-centres/CentreforEffectiveEducation/Research/ImprovingChildrensLives
20. www.educationnorthwest.org
21. www.ies.ed.gov/ncee/wwc
22. thelearningcurve.pearson.com
23. Development Impact Evaluation Initiative, for further details see: web.worldbank.org/WBSITE/EXTERNAL/EXTDEC/EXTDEVIMPEVAINI/0,,menuPK:3998281~pagePK:64168427~piPK:64168435~theSitePK:3998212,00.html
24. Systems Approach for Better Education Results, for further details see: web.worldbank.org/WBSITE/EXTERNAL/TOPICS/EXTEDUCATION/0,,contentMDK:22930156~pagePK:148956~piPK:216618~theSitePK:282386,00.html
25. Centre for Educational Research and Innovation (CERI), for further details: www.oecd.org/edu/ceri/evidenceineducationlinkingresearchandpolicy.htm
26. Randomise Me, for further details see: randomizeme.org/?page=3
27. Fonds d'expérimentation pour la Jeunesse, for further details see: www.experimentation.jeunes.gouv.fr/
28. Investing in Innovation Fund, for further details see: www2.ed.gov/programs/innovation/index.html
29. For more information on Key Information Sets see: www.hefce.ac.uk/whatwedo/lt/publicinfo/kis/
30. Key Information Sets, for further details see: www.hefce.ac.uk/whatwedo/lt/publicinfo/kis/
31. MyEdu, for further details see: www.myedu.com/

32. Professor Performance, for further details see: www.professorperformance.com/
33. Coe, R., Kime, S.,Nevill, C. and Coleman, R. (2013) ' The DIY Evaluation Guide.' London: The Education Endowment Foundation. Available at: educationendowmentfoundation.org.uk/uploads/pdf/EEF_DIY_Evaluation_Guide_2013.pdf
34. Centre for Effective Services (2011) 'The What Works Process: Evidence informed improvement in child and family services.' Dublin: Centre for Effective Services.
35. Bakhshi, H. and Mateos-Garcia, J. (2012) 'Rise of the Datavores.' London: Nesta. Available at: www.nesta.org.uk/home1/assets/features/rise_of_the_datavores_report
36. Fullan, M. and Donnelly, K. (2013) 'Alive in the swamp, assessing digital innovations in education.' London: Nesta. Available online: www.nesta.org.uk/library/documents/Alive_in_the_Swamp.pdf
37. Best Evidence Encyclopaedia, for further details see: www.bestevidence.org.uk/
38. Graphite, for further details see: www.graphite.org/
39. Blueprints for Healthy Youth Development, for further details see: www.blueprintsprograms.com/
40. EdSurge, for further details see: www.edsurge.com/
41. Classroom Window, for further details see: classroomwindow.com/press/
42. Education Endowment Foundation, for further details see: educationendowmentfoundation.org.uk/
43. Education Endowment Foundation (2013) 'A third of school leaders using the Toolkit.' 16 May 2013. See: educationendowmentfoundation.org.uk/news/a-third-of-school-leaders-using-the-toolkit
44. Education Endowment Foundation toolkit, for further details see: educationendowmentfoundation.org.uk/toolkit/about-the-toolkit/
45. Education Endowment Foundation (2013) 'A third of school leaders using the Toolkit.' 16 May 2013. See: educationendowmentfoundation.org.uk/news/a-third-of-school-leaders-using-the-toolkit
46. It should be noted that some resources, such as the Best Evidence Encyclopaedia do include products from businesses, and positively, there are plans for the Toolkit to do so in the future too. Yet at the moment the role education businesses play is under explored.
47. http://www.carbonstatement.com/company/what-we-do.aspx
48. Puttick, R. and Ludlow, J. (2012) 'Standards of evidence for impact investing.' London: Nesta. Available at: www.alliance4usefulevidence.org/assets/StandardsofEvidenceforImpactInvesting.pdf
49. Nesta Impact Investment Fund, for further details see: www.nestainvestments.org.uk/
50. Mulgan, G. in Barber, M. and Rizvi, S. (eds) (2013) 'Asking more: the path to efficacy.' London: Pearson.
51. Perkins, D. (2010) 'Fidelity-Adaptation and Sustainability.' Presentation series on developing evidence information practice for children and young people: the 'why and what'. Organised by the Centre for Effective Services (www.effectiveservices.org) in Dublin, Cork and Galway, October 2010.
52. Barber, M. and Rizvi, S. (2013) 'The incomplete guide to delivering learning outcomes.' London: Pearson. Available at: efficacy.pearson.com/
53. Sharples, J. (2013) 'Evidence for the Frontline.' London: The Alliance for Useful Evidence. Available at: www.alliance4usefulevidence.org/assets/EVIDENCE-FOR-THE-FRONTLINE-FINAL-5-June-2013.pdf
54. Kuhn, T. S. (1962) 'The Structure of Scientific Revolutions.' Chicago Il: University of Chicago Press.
55. Ries, E. (2011) 'The lean Start-up: How Today's Entrepreneurs Use Continuous Innovation to Create Radically Successful Businesses.' London: Portfolio Penguin.
56. Puttick, R. (2012) 'We don't need what works.' London: Nesta. Available at: www.nesta.org.uk/assets/blog_entries/wedontneedwhatworksweneedtoknowwhatisworking
57. Allen, R. (2013) 'Evidence-based practice: why number-crunching tells only part of the story.' Available at: ioelondonblog.wordpress.com/2013/03/14/evidence-based-practice-why-number-crunching-tells-only-part-of-the-story/
58. For example see Social Research Unit (2013) 'Implementing Evidence Based Programmes.' See: dartington.org.uk/how-to-implement-evidence-based-programmes-webinar-for-yjb-effective-practice-series/
59. Nutley et al., (2007) ibid.
60. Nutley, S.M., Walter, I. and Davies, H.T.O (2007) 'Using Evidence: how research can inform public services.' Bristol: The Policy Press.
61. Puttick, R. and Ludlow, J. (2012) 'Standards of evidence for impact investing.' London: Nesta. Available at: www.nesta.org.uk/publications/assets/features/standards_of_evidence_for_impact_investing
62. For further details on the development of Project Oracle and the Oracle Standards of Evidence, see Ilic, M. and Bediako, S. (2011) 'Project Oracle: Understanding and sharing what really works.' in Puttick, R. (ed.) 'Using Evidence to improve Social Policy and practice: perspectives on how research and evidence can influence decision making.' London: Nesta. Available at: www.nesta.org.uk/library/documents/EssayBookWeb.pdf

63. Hattie, J. (2008) 'Visible Learning: A Synthesis of Over 800 Meta-Analyses Relating to Achievement.' Abingdon: Routledge.

64. Higgins, S., Xiao, Z. and Katsipataki, M. (2012) 'The Impact of Digital Technology on Learning: A Summary for the Education Endowment Foundation.' Available at: educationendowmentfoundation.org.uk/uploads/pdf/The_Impact_of_Digital_Technologies_on_Learning_FULL_REPORT_(2012).pdf

65. Luckin R., Bligh, B. Manches, A., Ainsworth, S., Crook, C. and Noss, R. (2012) 'Decoding Learning: the potential and promise of digital education.' London: Nesta.

66. Taken directly from p. 9 of Luckin, R., Bligh, B., Manches, A., Ainsworth, S., Crook, C. and Noss, R. (2012) 'Decoding Learning: the potential and promise of digital education.' London: Nesta.

67. Nutley et al., ibid p. 19.

68. www.sagepub.com/refbooks/Book217583

69. Bakhshi, H. and Mateos-Garcia, J. (2012) 'Rise of the Datavores.' London: Nesta. Available at: www.nesta.org.uk/home1/assets/features/rise_of_the_datavores_report

70. In 2012 Nesta held a 'Failure Fest' to discuss less than successful education projects. See: www.nesta.org.uk/assets/events/nesta_failure_fest_education_innovation_and_enterprise

71. W.K. Kellogg Foundation (2004) 'W.K. Kellogg Foundation Logic Model Development Guide.' Available at: www.wkkf.org/knowledge-center/resources/2006/02/WK-Kellogg-Foundation-Logic-Model-Development-Guide.aspx

Nesta...

PEARSON

Nesta

1 Plough Place
London EC4A 1DE

research@nesta.org.uk
www.twitter.com/nesta_uk
www.facebook.com/nesta.uk

www.nesta.org.uk

February 2014