



P R O V O C A T I O N S

BURDEN OF PROOF: IS EVIDENCE REALLY THE KEY TO GOOD POLICY DESIGN?

A FETL provocation by Stephen Exley



PROVOCATIONS

The meaning of the term 'provocation' is usually understood as 'incitement', something that provokes, arouses or stimulates – a message in search of a response. This series of short FETL papers has this in mind. Their aim is to stimulate interest and debate, to air a new or original idea with a view to eliciting thoughtful, open-minded responses. As with all FETL publications, we do not seek to offer the final word. We are all about what happens beyond the page, in the wider life of an idea. We hope readers will take them in this spirit and share with us their own ideas and responses and, indeed, their own provocations.



FOREWORD

Dame Ruth Silver

Evidence-based policy was one of the mantras of the Blair and Brown years and remains a cherished aspiration of policymaking in the UK. It is accepted, usually without much question, as an obviously good thing at which policy in education should aim. However, as Stephen Exley demonstrates in this excellent 'provocation', originally published as an article in *Tes*, things are not as straightforward as this suggests. Is it realistic, or even desirable, to make evidence the main driver of policy, he asks.

This is an essential but rarely asked question, which is why the Further Education Trust for Leadership (FETL) decided to republish Stephen's article. We are grateful to *Tes* for its permission to reproduce it here. I very much hope it will stimulate further reflection on this theme, and a more critical approach, in general, to dealing with evidence.

One of the reasons FETL was set up was to strengthen the research and evidence base on which policy in further education could draw. This is undoubtedly important. It is critical that leaders ground their decisions in an understanding of what works and, just as important, what doesn't. FETL's mission implies a belief that systems can change, that we can learn both to do things better and to avoid making the same mistakes two or three times – a recurrent problem in a sector where policy memory tends to be short.

But decisions are rarely a matter of applying directly the findings of research or the outcomes of monitoring and evaluation. As Stephen shows, there are numerous other factors in play, including the sector's mission and values, the local sub-systems in which providers operate and the priorities set by the needs of learners and the resources available to support them.

The job of education leaders, whether at national, local or institutional level, is not to accept the evidence presented uncritically, nor is it to seek out evidence that supports their prejudices, disregarding the rest. Instead, leaders must take a thoughtful, sensitive approach to reading and implementing evidence, using their own judgement and looking everywhere and elsewhere for new models of understanding, while also being prepared to be challenged and to change their thinking, if appropriate.

Stephen's article is timely and useful, particularly given the turbulent times we are currently attempting to navigate, with their attendant opportunities and threats, and the chronic short-termism that continues to afflict our politics. As leaders, we need to apply the same skills of criticism, reflection and judgement we increasingly demand of our learners.

Dame Ruth Silver is President of the Further Education Trust for Leadership

IS EVIDENCE REALLY THE KEY TO GOOD POLICY DESIGN?

While evidence-based policymaking is very much in vogue, critics say the skills sector is lagging behind. As a result, those who govern FE are destined to continue designing flawed policies and failing to learn from their mistakes. But is a greater use of evidence really the solution? Stephen Exley reports.

"This government expects more of policymakers. More new ideas, more willingness to question inherited ways of doing things, better use of evidence and research in policymaking and better focus on policies that will deliver long-term goals."

For a further education sector battered by decades of supposedly radical policy shifts, bold promises and false dawns, this section of the Modernising Government White Paper – a rallying cry for a new era of evidence-based, long-term policymaking – sounds like a breath of fresh air.

It really shouldn't. The document, drawn up in the early days of the New Labour government, is 20 years old. Looking back, it hasn't aged well. Just four years on from this bold statement of intent, Tony Blair, who was prime minister at the time, rowed back on the Tomlinson report's recommendation to remove GCSEs and A levels, concerned that the proposals could be a little too radical for the electorate to stomach.

Another four years on, Labour did press ahead with a major qualification reform in the shape of the 14-19 diploma, launched in a blaze of publicity. Yet two years – and more than £300 million in investment – later, the new coalition government announced that the qualification would be scrapped before it had even taken root.

Where is the robust evidence?

In theory, at least, evidence in education, as in policymaking, is very much in vogue. The Education Endowment Foundation was created in 2011 to support schools and colleges by providing evidence-based resources designed to improve practice. It was just one of seven What Works Centres opened across major policy areas, designed to “enable policymakers, commissioners and practitioners to make decisions based upon strong evidence of what works and to provide cost-efficient, useful services”.

Yet, while ministers have been at pains to ensure practice in the classroom and lecture hall is based on robust evidence and monitored for its impact, they haven’t held their own policymaking processes to the same high standards. In his foreword to the City & Guilds Group’s Sense & Instability 2019 report, which explores policymaking in the skills sector and was published last month, Matthew Taylor, chief executive of the Royal Society for the Encouragement of Arts, Manufactures and Commerce, notes that “policy continues to be developed with little or no supporting evidence”.

This is only the start of the problem. In the raft of recent reforms in the further education and skills sector, insufficient efforts have been made to incorporate any substantial means of evaluating their subsequent impact.

Unless the philosophy behind Modernising Government is finally and genuinely put into practice, some two decades late, policymakers are destined to continue making mistakes – and failing to learn from them. But how achievable – and, indeed, desirable – would it be for evidence to be the main driver behind policy?

In its third and latest Sense & Instability report, the City & Guilds Group focused on seven high-profile policies, which were collectively funded to the tune of £15 billion over a 14-year period.

The Department for Education’s (DfE) own Skills Index offers a stark insight into the cumulative effect of skills policies. Compared with the baseline in 2012, the overall impact of skills on the UK’s productivity has since dropped by 27 percentage points.

While some of this decline can be attributed to government funding cuts in FE, Chris Jones, chief executive of the City & Guilds Group, believes this also reflects major flaws in how policies are set and monitored.

“There are rather continuous examples of lack of measurement being built into the design of policy, linked to explicit outcomes and value for money,” he says.

“Policy is just not being designed well enough. The infrastructure, the necessary funding, the appropriate incentives [are not] being structured in a way that will allow the right outcomes to be derived. The direction of policy? Broadly correct. But the machinery of government and the process by which it implements and manages [policymaking] would appear to be woeful.”

Take the seven policies that Sense & Instability 2019 focuses on: Train to Gain; the Skills Pledge; the Work Programme; Advanced Learner Loans; the introduction of apprenticeship standards; T levels; and the National Retraining Scheme. The report assesses the extent to which evidence has been built into the design, implementation and assessment of each programme, using the following principles:

- Are clear success measures established in the design and consultation phase?
- Have these success measures been validated through the use of pilots?
- Were comparison groups, baseline-end assessments and longitudinal studies used to track outcome and impact?
- Was the extent to which any success measures were met evaluated?

The results reveal a somewhat limited use of evidence.

For Jones, the DfE is lagging behind other government departments. He cites the National Institute for Health and Care Excellence as an example of good practice. It focuses on producing evidence-based guidance for health and social care practitioners, as well as developing quality standards and performance metrics for providers.

“There are examples elsewhere around government of where I think they are thinking harder about what they do and how they measure impact over time, and [ensuring] the appropriate independence of views that are brought to bear,” Jones explains. “We don’t have this across skills.”

Even if there were a concerted focus on data evaluation and assessment in this context, the nature and rate of policy change in recent years would make this difficult, he adds. In the

past 30 years, there have been 70 secretaries of state responsible for skills policy compared with just 20 in charge of schools policy.

The same is true of skills ministers. Take the Sainsbury review, commissioned in 2015, which resulted in the creation of T levels. In a relatively short period of time, this policy has ended up on the desks of three different skills ministers: the review was commissioned by Nick Boles, subsequently picked up by Robert Halfon and is currently being implemented by Anne Milton.

“Even if they would like evidence-based policy, the nature and rate of change and responsible ministers for [skills] over a long period of time has actually often created these vacuums of policy,” Jones says.

And one of his main bugbears is how ministers have repeatedly hailed T levels as “gold-standard” qualifications before the first ones have even been taught. “As of today, we have no student that has studied a T level. No student has been examined in a T level. No student has progressed into a job with a T level. No student has progressed to university with a T level. How on earth can we say it is a gold standard?”

But while Jones is by no means alone in objecting to the phrase being bandied about without the evidence to back it up, does this really matter? It is a metaphor, after all.

For David Russell, chief executive of the Education and Training Foundation, the T-level example goes to the heart of the complex relationship between policy and evidence: it is an assertion that’s impossible to prove or disprove.

“There is truth as to whether a qualification is any good or not, but you have to unpick what is meant by ‘good’. There are things like reliability, validity, complexity, cost, accessibility.

“One of the things is reputation. So, the reputation of qualification is a real feature you can measure by going around and asking people. But reputation is only weakly connected to actual quality.

“One of the main things you can do to change the reputation of a qualification is change the way you talk about it.”

Secret garden

Russell, who has 16 years’ experience in the civil service, has seen at first hand how this plays out. Public understanding of education is very different from other public services. While people will base their views of, say, the NHS or train services on their own experience, the intricacies of qualifications are something of a secret garden.

“Most people have no real direct access to the qualification,” Russell says. “They don’t know what’s in it or what’s involved in it, and they’ll never find out. Because it’s complicated and we don’t have time, they probably don’t care that much. So, in that situation, where there’s a big gap between me making up my mind about the thing and me having direct access to the thing, what influencers are saying about it becomes really important.”

While concern is most commonly targeted at policies introduced on the basis of limited or partial evidence, it can also be used to question popular policies that don’t have the expected – or desired – impact.

Russell cites the Young Apprenticeships programme, established in 2004 to give 14- to 16-year-olds the opportunity to gain knowledge and skills in a specific vocational area and achieve a relevant qualification.

“People really believed in it,” he recalls. “There was a strong sense this was a great thing; these are disenchanted kids not doing well in maths and English. You need to contextualise their maths and English in a different way and get them engaged with employers. It sounded like a really good idea. But the evidence was that it wasn’t working at all.” The policy was soon scrapped.

Andy Westwood, professor of government practice at the University of Manchester, is a firm believer that evidence is just one of a number of factors behind policymaking. The others – including ideology, the values of a particular minister and media pressure – are less scientific.

“Lots of things come together to make policy at any given time,” says Westwood, a one-time special adviser to ministers at the former Department for Innovation, Universities and Skills.

“Academic evidence or user data is part of that mix. When you’re putting together Green Papers, White Papers, legislation, it’s going to be there – it’s just that it might not always be at the beginning of the queue.

“We expect policymaking to be more scientific than it is – or even more scientific than it should be. We actually want things to be based on values. In healthcare, say, you want things to be based on values of running a decent NHS with access for people equally in all parts of the country. The evidence might say to close provision in small places and stick it in bigger places, and it would be more efficient.”

‘Finding the sweet spot’

Russell recalls a policymaking model towards the end of the New Labour government that was based on three intersecting circles: “political”, “delivery” and “evidence”.

“Effective policymaking was about finding the sweet spot in the middle,” he explains.

“Something that took the political trick, was deliverable in the real world through the tools and levers you have available as a department of state, and had the evidence to back it up sufficiently to suggest it would be beneficial. That overlap between these three considerations is really, really small – and, sometimes, it’s non-existent.”

While Russell believes there are plenty of examples of research being twisted and cherry-picked to suit the agenda of a minister or political party, he nonetheless believes evidence is becoming more important to policymaking: “It’s staggering forward like a drunken man, lurching to the left and right, but he is making progress forwards.”

Evidence, he says, has become a new battleground: “Which evidence should we pay attention to? Which evidence is reliable? Which evidence should we place more weight on?”

“It’s been said that [former education secretary] Michael Gove didn’t do ‘evidence-based policymaking’, he did ‘policy-based evidence making’, says Russell. “The way I would put it is that he used evidence like a lawyer, not like a scientist.

“Gove deployed evidence to back up his argumentation like a lawyer would do, and ignored evidence he didn’t like. It’s not that he twisted evidence or pretended it said something that it didn’t. He just deployed it to make a case. Even then, that’s an acknowledgement that you have to play the evidence game.”

But, even if it were desirable, could policy ever be completely evidence based? Russell believes not – and the reason is that the voting public “generally don’t care that much about evidence”. It is important but is by no means the be-all and end-all.

“If you set out a programme for government that was entirely evidence based and rational, you’d never get to enact it because people wouldn’t like it,” says Russell.

On this, Westwood is in full agreement. “There are lots of times where evidence isn’t straightforward; it says different things,” he says. “It’s very rarely that clear-cut. You need to use judgement and bring in other factors.”

Stephen Exley is FE editor for Tes.

He tweets @stephenexley

This article originally appeared in the 7 June 2019 issue of Tes, under the headline “Burden of proof”. It is available online at <https://www.tes.com/magazine/article/evidence-really-key-good-policy-design>

To cite this paper:

FETL (Further Education Trust for Leadership). 2019.
Burden of Proof: Is evidence really the key to good policy design?. FETL.

This work is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International (CC BY-NC-SA 4.0) License. To view a copy of the license, visit <http://creativecommons.org/licenses/by-nc-sa/4.0/>



Published July 2019

The Further Education Trust for Leadership

Website: www.fetl.org.uk

Email: enquiries@fetl.org.uk

 [@FETforL](https://twitter.com/FETforL)