FREEING-UP EDUCATION ANALYSIS

Using government data to help improve educational outcomes for young people

James Noble, George Hoare and Farooq Sabri
May 2016

CONTENTS

EXECUTIVE SUMMARY

Why an Education Impact Data Lab is needed

Many of the social programmes delivered in and around schools lack persuasive evidence about whether they are effective. The difficulty is that good impact evaluation requires time, money and expertise which is generally lacking amongst programme providers.

The Education Impact Data Lab (EIDL)[*] would help resolve this problem by enabling organisations to conduct impact evaluations using data which is already collected by schools and collated by the Department for Education (DfE). Essentially, it would be a short-cut to the data that organisations need to measure their impact and increase their effectiveness.

Importantly, the EIDL model would protect the privacy of individuals' data. Organisations that use the EIDL would only see aggregate results, they would not be given data about individual young people, and analysis would be conducted by DfE or its partners, so there would be no risk of data getting into the wrong hands.

The proposed model

We suggest that an EIDL could be run in a similar way to the successful Justice Data Lab, which is a free service launched by the Ministry of Justice in April 2013 for organisations working to reduce re-offending.

The EIDL could work through a simple model. Firstly, programme providers would tell the EIDL which young people they had worked with. The EIDL team would then identify these young people in the National Pupil Database (NPD) and provide aggregate outcome data for both this group and a statistically matched comparison group to assess whether the programme was associated with any difference. The results of the comparison would be sent back to the programme provider in an easy to understand report.

The Justice Data Lab has shown that NPC's data labs concept can work, and we have good reason to think there is a market for a similar service in education as there are many charities and other organisations running programmes in schools and local autonomy over spending decisions has increased demand for good quality evidence to inform decisions.

Benefits of the service

The Education Impact Data Lab service would have clear benefits:

- It would allow programme providers to understand their impact, improve their services and present a better case to schools.

- It would help schools make better decisions about which programmes to commission and get the best value for money.

---

[*] NPC runs a programme to develop Data Labs to help charities access reports on the outcomes of their interventions compared to a comparison group, see www.NPCdatalabs.org. To avoid confusion with the existing Fischer Family Trust 'Education Data Lab', our proposed education service is referred to as the Education Impact Data Lab (EIDL).

- It would prevent waste, by giving providers access to data that already exists rather than collecting their own.
- It would reduce spending on evaluation, freeing up resources. The cost of setting-up and running the EIDL would be far lower than commissioning independent studies to deliver the same results.
- It would help the education sector—and government— to develop an improved understanding of 'what works' based on reliable data and statistical methods.

Each of these benefits could link to better pupil outcomes; **hence we believe that an EIDL would make an important contribution to addressing the persistent attainment gap in schools.** As such, we recommend that the EIDL is piloted for one year to test and refine the concept and we suggest that this is funded by the government or by a research commissioner such as Education Endowment Foundation.

The aim of this report

This report makes the case for the EIDL based on consultation with organisations in the education sector. We outline the demand for an EIDL and what it could be used for, including case studies of organisations who have expressed an interest in using it.

By publishing the report we want providers of services for young people to appreciate the potential of the EIDL approach, and we want them to join us in lobbying the government to set it up. Actions that individual organisations can take include:

- Apply pressure to funders to fund a pilot of the service.
- Talk to colleagues, other organisations and membership bodies about the potential of the EIDL, to increase awareness and demand.
- Start to prepare your organisation to use the EIDL. In particular check that you are collecting data on service user's names, date of birth, school(s) attended and ideally Unique Pupil Number (UPN).
- Check that any information you give to service users states that you may share their data for the purposes of research, and gives them the opportunity to opt out or in. Keep in touch with developments of the EIDL, as we think consent issues should become clearer.
- Get in touch with NPC to discuss the EIDL further.

The report is based on:

- Qualitative interviews with seven charities in the education sector.
- A survey of 71 individuals—mostly from charities and social enterprises working in schools (n=62), with the remainder from private and public sector organisations[*].
- Discussions with stakeholders such as the Fischer Family Trust (FFT) and the Education Endowment Foundation (EEF).

---

[*] It should be noted that our survey—while providing a good indication of demand for an EIDL—is not necessarily representative of all education VCSEs. They are a diverse group, including many that operate outside the school environment and are not part of any networks that would help us identify them—so they may be underrepresented. We also know that access to data is a relatively specialist topic, so we suspect there is some response bias in the sample towards charities that are already engaged in impact measurement.

CHARITIES IN THE SCHOOL SYSTEM

Tackling low social mobility and educational disadvantage are important government priorities and a goal of many charities working in the education system. The statistics are compelling. In 2015, just 37% of disadvantaged[*] children achieved five good GCSEs, including in English and mathematics, compared to 65% of all other pupils.[1] Since poor academic performance is associated with lower wages and higher unemployment in adulthood[2], this attainment gap transmits poverty from one generation to the next.[3]

Successive governments have focused on schools' role in addressing educational disadvantage, but a wide range of voluntary, community and social enterprises (VCSEs) also provide services and support to young people.[4] This includes academic support (inside and outside the classroom); 'enrichment' activities like arts and sports, counselling, mentoring and other activities that contribute to a young person's personal and social development.[5] Often, these are services that schools cannot provide themselves, and although they may not be explicitly academically focused, they can make an important contribution to educational outcomes.

From *Every Child Matters* in the early 2000s[6], through to the focus on character in the recent education white paper[7], governments of all political persuasions have recognised the importance of the support, opportunities and experiences that children access outside the classroom. Indeed, the government's vision is '*for schools to increase their range of activities',* based on strong relationships with voluntary and sporting organisations, among others. Character traits of perseverance and resilience are seen as key, not only to academic success, but also to future job prospects and an individuals' place in wider society.

In recent years, the role of the VCSE sector in the school system has been influenced by Pupil Premium funding[8] and academy school status, both of which give schools more autonomy over commissioning. School leaders now have to make difficult choices about how to deploy the limited time and resources available to them for extra-curricular activities, and their decisions are closely scrutinised[†]. Unfortunately, there is not enough good evidence about the potential impact of different types of programmes, so schools often have to make decisions based on hunches, personal relationships and their perceptions of quality and impact. The demand for better evaluation is not just about the pressure of funding; education providers also want to better understand the difference they make.

'Evidence is a crucial tool which schools should use to inform their decision making and ensure that they identify the "best bets" for spending.'

Sutton Trust and EEF, July 2015[9]

The government has recognised this problem and has sought to address it by setting-up the What Works Centre for Educational Achievement, run by the Education Endowment Foundation (EEF) and Sutton Trust, which has made progress in trialling different educational programmes, synthesising evidence and communicating findings to schools. However, the EEF is not sufficiently resourced to evaluate all the programmes delivered in schools, many of which are too small scale for trial-based methodologies.

---

[*] Disadvantage is defined as pupils who were known to be eligible for free school meals in the past six years (years 6 to years 11), pupils who have been looked after, or were adopted from care.

[†] The National Audit Office (NAO) found that the impact of the Pupil Premium is being undermined by schools spending it on approaches that are not cost effective. To address this, NAO recommends that schools publish statements explaining the impact of their expenditure and that measures to *'expand the evidence base more quickly'* should be considered. The EIDL would help realise this aim.

WHY EVALUATING EDUCATIONAL OUTCOMES IS SO CHALLENGING

There are two core arguments for impact evaluation in and around schools:

- If providers have data about impact it will help them to improve the quality of their services and strengthen the overall evidence base; ultimately leading to better outcomes for young people.

- If commissioners have reliable and informative evidence about what is effective for their context they will make the best choices about which services to provide.

Our consultation with the sector has identified the following challenges to achieving these aims.

Six key challenges to evaluation

1. Charities and other providers have limited resources to commission or run impact evaluations.

In situations where evaluation would be helpful, the funds are generally not available to do it well. This has led to an abundance of weak, inconclusive evaluation studies which don't provide much useful information and waste time and resources. For example, our survey indicated that only 27% (n=16) of respondents had successfully conducted studies with control groups and just 10% (n=6) have done randomised controlled trials (which are seen as the highest standard of educational research)[*].

2. Charities and other providers tend not to have the expertise to conduct evaluations.

This is not surprising as evaluation is a specialist discipline and few providers are big enough to recruit or commission people with this knowledge or experience.

3. Impact on educational outcomes takes time.

A real challenge for charities is getting data on longer-term outcomes. For example, providers may work with pupils at the transition between primary and secondary school but have to wait at least five years to learn about the impact of this on academic outcomes at GCSE. Only 20% (n=12) of our survey respondents had been able to conduct long-term follow up with service users post-intervention.

'We clearly know we must keep on chasing this 'proof' but are increasingly frustrated by the lack of simple systems for tracking progress of individuals we work with. A couple of schools have been real success stories for us, but it's like two in three hundred or so!'

Senior manager, charity/voluntary organisation

---

[*] These figures are likely to be a significant overestimate because of the likely bias amongst our respondents towards those already interested in impact measurement.

4. **Getting data from schools is very difficult.**

All schools collect data on their pupils which can be used to look at the effectiveness of programmes, but they whether they are willing/able to share it varies. There is also disparity in which data systems schools use and the type of non-statutory data they collect. This means that charities working across a number of schools have to negotiate access with each school individually and then face the task of trying to synthesise inconsistent data. Hence we are not aware of any charitable providers that have been able to do this well.

'Getting information from schools varies from school to school. Some understand the importance of it and provide information quickly, whilst others are either less concerned about the data, or are worried about data protection issues.'

                    Policy and communications manager, charity/voluntary organisation

5. **Getting data from teachers, other school staff and pupils is also difficult.**

Without access to school level data, some providers attempt to get information directly from school staff or pupils through questionnaires or forms. Our survey indicated that half of charitable providers had attempted to get data from teachers within the last year and half from pupils. Not only is this data very unreliable, it is often duplicative. Providers expend effort and goodwill chasing over-stretched staff for data that often already exists in the school system. We suspect that across the school system many hours of teacher and charity staff time are being wasted in pursuit of inherently unreliable data.

'Getting hold of this data has been the most time consuming and frustrating part of our work. If we can get the data, we can process it and will use it, but we cannot get consistent, reliable data from schools without gargantuan efforts and endless nagging that spoils our relationship with the schools.'

                    Senior manager, charity/voluntary organisation

6. **Getting data from DfE is difficult.**

A lot of the data that organisations need already exists: schools have a statutory requirement to record the characteristics (age, ethnicity, Pupil Premium entitlement), attendance and academic achievement of *all* pupils and add it to the DfE's National Pupil Database (NPD), making it one of the richest education datasets in the world. Academic studies have been conducted using the National Pupil Database, but we think that it remains an underutilised resource. The challenge is that access to NPD is strictly limited because it contains personal data and very few VCSEs can meet the conditions required to access it[10] or have the analytical capability to use the data and interpret results.

Nevertheless, more and more charities are looking to independent evaluators to provide evidence for the impact of their programmes. In January the British Science Association commissioned Pro Bono Economics to evaluate the impact of its CREST award on GCSE science students' grades, and last month Action Tutoring undertook an evaluation of the impact of its small-group tutoring programmes carried out by the National Institute of Economic and Social Research. In both cases, NPD data and statistical comparison groups were used. This methodology would be used by the EIDL, and would open this important analysis to a wider group of charities who would otherwise struggle to access this data and interpret the results.

> In short, we want to see analysis of the National Pupil Database become routine for all providers working with young people in schools and believe the EIDL is the way to make this happen.

HOW AN EDUCATION IMPACT DATA LAB (EIDL) WOULD WORK

The precedent of the Justice Data Lab

Organisations working to improve outcomes for offenders used to face similar challenges to those discussed. This is why we made the case for[11], and supported the development of, the Justice Data Lab[12]—a Ministry of Justice (MoJ) service that enables not-for-profit and other organisations to get limited access to aggregate re-offending rates for their service users and to compare these results to a matched comparison group[*].

To use the Justice Data Lab organisations provide details of the offenders that they have worked with (eg, name and date of birth) to analysts in the MoJ, who then produce a report of the aggregate re-offending rate, frequency of offending, time to re-offending and severity of the re-offence. The MoJ also provides results from a matched control group of offenders with similar characteristics, created through a statistical technique called propensity score matching.[13]

Figure 1: PET Justice Data Lab results
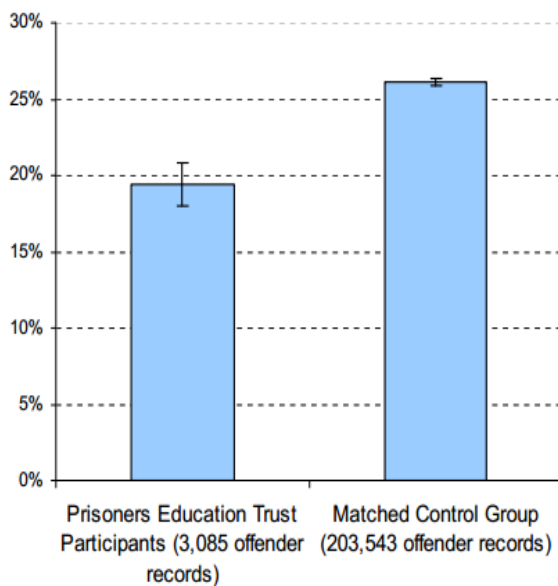
Source: Ministry of Justice (2014) Justice Data

Figure 1 shows an example of Justice Data Lab analysis for the Prisoners' Education Trust (PET) which provides grants to prisoners for distance learning courses, educational, arts and hobby materials.

PET sent data to the Justice Data Lab for prisoners who received a grant between 2002 and 2010. The one-year proven re-offending rate for these offenders was 19%, compared with 26% for a matched control group. This difference is statistically significant; meaning that PET grants are associated with a reduction in re-offending of between 5 and 8 percentage points.

PET has incorporated these results into a wider evaluation of their programme which has helped them to re-negotiate the delivery of programmes with a commissioner and to begin estimating the economic benefit of their grants.

We believe that a similar model could work for organisations that aim to deliver outcomes in education. Like the Justice Data Lab, to use the service organisations would provide details (such as name, date of birth, school) of a cohort of pupils they have worked with and in return they would receive a report on the educational outcomes for that group compared to a matched control group of young people with similar characteristics. The difference between the two results would be an indication of the impact the organisation's service has had on the beneficiaries' educational outcomes.

---

[*] The Justice Data Lab has won a number of awards; most notably in 2014 it was joint winner of the Statistical Excellence in Official Statistics award from the Royal Statistical Society.

Essential ingredients for a successful data lab

We think there are four criteria for a successful data lab. We describe these below along with a discussion of how the proposed EIDL would meet each of them, and some of the important issues to address.

1. There is an efficient and straightforward way to access analysis.

Limited access to the NPD is already possible, but access is tightly controlled and the onus is on the NPD user to analyse the data themselves. This system doesn't work for the majority of charities and other providers because:

- They cannot meet the access conditions;
- It would require too large a data extract to identify service users and a comparison group; and
- They do not have the analytical capability, resources or experience to use the data.

However, because the type of analysis that most providers want to conduct is broadly the same, there is scope to make analysis more routine. The Justice Data Lab provides us with a good template. It has established an effective process for secure data transfer and analysis that minimises the time and expertise needed.

To use the EIDL, organisations would need to collect or access data about their service users' names, date of birth, school(s) attended and ideally Unique Pupil Number (UPN)[14]. This data would need to be sent securely to DfE, who would then identify and link to individual pupils in the NPD.

2. It meets legal requirements and does not compromise individuals' right to privacy.

A crucial consideration for the EIDL is that the National Pupil Database contains sensitive personal information about individual children and young people and the DfE has legal and ethical obligations to protect this. This is why strict access controls are in place.

However, studies of programme effectiveness do not require the sharing of personal data with third parties: the proposed EIDL approach would provide organisations with the results of the analysis *but not the data itself*. No personal data would ever leave the EIDL and after the review period any data used in the analysis would be destroyed. A full Privacy Impact Assessment was conducted for the Justice Data Lab and this showed that the process complies with both the Data Protection Act (DPA) and the European Convention on Human Rights. We recommend the same approach for the EIDL, with clear procedures for information governance made available for public scrutiny.

The issue of service user **consent** is a potential barrier to establishing the EIDL. It relates to what, if anything, providers need to tell service users (or their parents) about plans to link their data to the NPD. It is a difficult issue, which depends upon what data you need from the NPD and what data you need to link to it. If a service is delivered in school the situation is clearer, because consent arrangements should be covered by the school's policies. If a service is delivered outside school hours, then providers may need to get permission for the linking to go ahead. However, this is probably only necessary on an 'opt-out' basis, ie, the provider needs to make clear that the linking will take place and can then assume people are happy with this unless they tell them otherwise. DfE do not currently make clear what process or wording they expect, in particular whether consent wording needs to refer to DfE / NPD specifically, or whether a more general reference to 'third parties' is acceptable. As part of the establishment of the EIDL it will be important to clarify these requirements.

More broadly, we encourage those developing the EIDL to follow the Justice Data Lab approach where it was decided that consent for data linking was not a requirement for submitting data to the service. Imposing stricter consent requirements on the EIDL would make little sense as there is no risk to individual's data or privacy, while it would diminish its use by curtailing retrospective analysis and increasing the administrative burden on service providers.

3.   It delivers useful information.

The consultation suggested that organisations would value measures of impact for the following variables:

*   Academic attainment (including Key Stage 2 results, Attainment 8 measures, for KS4, and pupil level value added scores).

*   Attendance at school.

*   Exclusions from school (which is the only behaviour variable available in the NPD).

To be genuinely useful, analysis of these variables would need to be provided alongside those of a statistically matched control group.

The organisations we consulted also emphasised the importance of understanding how the impact of their programme might vary for different groups of beneficiaries and demographic characteristics (in particular, Pupil Premium entitlement). Indeed, simply helping organisations to understand the characteristics of their service users would be a major benefit of the EIDL (see the Sea Cadets case study on page 18).

Some survey respondents suggested that the EIDL should be able to answer more bespoke questions. However, as this would likely increase the complexity and length of time for reporting, we recommend that the EIDL starts by running fairly limited analysis and moves into more complex areas over time (which is how the Justice Data Lab has developed since 2013).

4.   It delivers truthful information.

A successful data lab should deliver high quality information that helps people to draw the correct conclusions and is transparent about the methodology deployed.

The EIDL will need to consider which measures of attainment are appropriate. There are a range of possible measures to choose from, even within the same assessment stage. We suspect that choosing a single measure of attainment for all EIDL reports would be too limiting and that this would have to be negotiated between the provider and the team delivering the EIDL depending on the programme being analysed.

Another consideration is the fairly infrequent assessments of attainment produced by the school system. In England, pupils are tested at three Key Stages throughout their compulsory schooling (Key Stages 1, 2 and 4). This means there are essentially only three data points, so, for example, a charity working to help pupils transition from primary to secondary would only be able to look at Key Stage 4 attainment (a delay of 4-5 years between intervention and outcome). To address this problem it would be interesting to explore whether interim assessments conducted by schools themselves could be included in the variables available to the EIDL.

Another important aspect of this is the validity of the comparison group. The proposed approach is to derive comparison groups using a statistical approach called propensity score matching[13] which depends on having good data about the contextual factors that influence outcomes. This may not always be available. For example a programme that is focused on particularly disengaged young people or issues related to parenting would struggle to identify a good control group because there is no data on these factors in the NPD. In these cases, it may only be appropriate to report limited analysis (for example 'before and after' analysis without a comparison group).

How should results be reported?

Ease of use applies to reporting as well as access. We suggest that the EIDL follows the Justice Data Lab by providing a straightforward and consistent report that explains the findings in a way that is accessible to potential users. The main audience for the report will be service providers, schools and other commissioners—who would potentially use the reports to make decisions. As many organisations lack expertise to interpret statistical analysis, the report should use plain English and provide guidance on how the results can be interpreted, including limitations.[13]

A consistent reporting process should also help with the synthesis of findings across analysis.

Should results be published?

All analysis conducted by the Justice Data Lab is published. We support this in principle because it ensures transparency and is a fair exchange for receiving a high standard of evidence for little or no cost. Publishing the findings is also necessary to meet a key aim of the EIDL, which is to improve our shared understanding of what works for different young people in the education system

However, feedback from our interviews indicates mixed feelings towards publication because of a perceived risk to the organisation. Indeed, we suspect that uptake of the Justice Data Lab has been lower than expected because potential applications know that results will be published and they are worried about what it will show and how it could affect them.

The best option would be to maintain the publication principle as a default position, but that organisations are given a period of time to respond to the analysis. Potentially the report could incorporate providers' acknowledgement of the findings and a summary of how the organisation intends to respond to them.

'If results were to be published openly it would be helpful to have some explanatory material as to why a significant result would not have shown up and guidance on how to rectify this. Whilst obviously noting that not all programmes would ever get a significant result because they don't have the impact on education as hoped.'

Senior manager, charity/voluntary organisation

Who should run the EIDL?

The EIDL would need to be operated by the National Pupil Database data owner—the Department for Education—or by an established NPD processor such as the Fischer Family Trust or the National Foundation for Education Research.

THE POTENTIAL BENEFITS OF THE EIDL

The EIDL would be a significant step towards helping organisations overcome the challenges faced by charities working in schools. Nine in ten of the organisations in our survey (n=62) said that this service would be useful; only six organisations responded that it would not be useful.

'It sounds like a great idea and would be incredibly useful.'

Charitable/voluntary organisation

This section goes into detail on some of the main arguments for establishing an EIDL.

It would allow programme providers to better understand their impact and improve their services.

The purpose of evaluation is to learn and improve, so that we can do more with the resources available to us. The methodologies currently available to organisations help to an extent, but there is a significant challenge around the hard, statistical evidence of effect, which many would see as the bottom line. Data provided by the EIDL would help complete this picture so that qualitative and other evidence can be combined with statistical analysis to provide a clear narrative around impact. The organisations we consulted recognise this, of the 62 respondents to our survey who thought this EIDL would be useful, 97% (n=60) thought it would *'improve their charity's understanding of their impact on beneficiaries'* and 89% (n=55) said it would *'enable them to improve their services'*.

'Systematic access to anonymous data about pupils' educational attainment would allow us to evaluate our outcomes better, improve our offer and make the case for the effectiveness of our intervention better. This is a really great initiative that would be transformational for us'.

Senior manager, charity/voluntary organisation

Users of the service will need to be prepared for the fact that they may not get the results they want. Justice Data Lab users have received both inconclusive and negative results and overall the Justice Data Lab has been a dose of realism for the sector, highlighting just how hard it is for programmes to demonstrably reduce re-offending. But in the long run this is positive. The EIDL could be an antidote to inflated claims of impact and may cause organisations to reflect on whether they are using their resources and young peoples' time in the most effective way possible.

It would help programme providers present a better case to schools and other commissioners.

Providers of educational programmes are eager for a way to 'prove' what they do, so that they can sustain and grow services that they passionately believe in. This is particularly pertinent to organisations in the charity sector, which often work on a smaller scale and have limited capacity for independent evaluation. Without evidence from the EIDL there is a risk of charities being side-lined in favour of larger providers.

We expect that this will be the main driver for use of the EIDL. Of the 62 organisations in our survey that thought the EIDL would be useful, 89% (n=55) believed it would *'help them demonstrate their impact to funders'* and 76% (n=47) believed it would help them *'attract alternative sources of funding, such as payment by results and social investment'*.

'From what I know about this proposed service I think it would extremely useful to my organisation and the sector as a whole.'

Performance, impact and evaluation manager, charity/voluntary organisation

It would help schools make better decisions about which programmes to commission and to get the best value for money.

A clear corollary of providers having better evidence of impact is that the buyers of services will also be able to make more informed decisions.

We do not think that the EIDL will change this situation overnight, but it will introduce new evidence into the school system and help schools make increasingly rational decisions. It should also help to influence conversations and procurement decisions; once people know that this source of evidence is available they will hopefully begin to challenge service providers to produce it, thereby raising the standard of evidence across the board.

'*This ability would be invaluable to the 3rd sector and is much needed so funders and the tax payer can have confidence that money is well used.'*

Trustee, charity/voluntary organisation

We also believe that the government has an ethical duty to introduce the EIDL. If you consider the amount of young peoples' and school staffs' time being spent on well-meaning but untested interventions and the opportunity cost this represents, then you could argue that the government should do its utmost to enable providers to study the effect of their work.

'This is critical work if we are to ensure our resources are deployed in the most effective way to help young people'.

Chief executive, charity/voluntary organisation

It would prevent waste, by giving providers access to data that already exists, rather than having to collecting their own.

It has been a longstanding aim in the school system to reduce the amount of bureaucracy and time wasted on form-filling[*]. But for teachers and schools, collecting data for service provider evaluations must be one of the worst uses of time.

Moreover, teachers recognise this themselves, so the quality of data providers get from them is usually very poor. Similarly, the process of asking young people to fill in forms can be time-consuming and affect the relationships with the service provider.

---

[*] For example, through the National Agreement under Labour
http://webarchive.nationalarchives.gov.uk/20130401151715/http://www.education.gov.uk/publications/standard/Performanceachievementandstandards/SCH/Page6/DfES%200172%202003 and the Coalition government's Workload Challenge
https://www.gov.uk/government/news/tens-of-thousands-of-teachers-join-the-workload-challenge

The frustration is that much of this data collection is duplicative. For example, providers should not be trying to measure school attendance of their young people as this is recorded elsewhere. And yet this happens because the providers are not given access to this data. The EIDL would stop this from happening. Providers will still need to collect some data, in particular qualitative feedback, but they will be able to focus on more meaningful questions such as feedback on the services provided.

'It's a fantastic proposal and one we would fully support to not only help us but put less pressure on our partner schools to provide us with duplicate information.'

Project Manager, charity/voluntary organisation

It would reduce the amount spent on evaluation.

The cost of setting-up and running the EIDL would be far lower than the cost of commissioning independent studies to deliver the same results.

This is particularly pertinent if you consider the standard of evidence that the EIDL offers. Randomised control trials (RCTs) are often viewed as the gold standard for evaluation[15] and while the EIDL would not match these for methodological validity, it would be the next best thing (and would represent a level 3 analysis on the widely used NESTA standards of evidence[16]).

Indeed the proposed EIDL would have a number of benefits when compared to the RCT approach:

- It would be less demanding for schools and staff involved in the delivery of programmes.
- It will allow programmes that have already taken place to be evaluated, opening up a large volume of untapped evidence.
- It would be much less costly because it is based on data that already exists.

The EIDL would be particularly useful to organisations delivering 'promising' programmes that wanted to strengthen their evidence base before accessing funding for robust evaluation studies. In this way, the EIDL could be used as a 'pre-RCT' service. There is also the potential to combine the EIDL with an RCT methodology to track outcomes for 'treatment' and 'control groups' to deliver a cost-effective evaluation of long-term impact.[17]

It would help the education sector—and government—to develop an improved understanding of 'what works' based on reliable data and statistical methods.

The ambition across all areas of policy is to have a better understanding of what types of services work for different people in different circumstances. Government, academia and others have been pursuing this aim for many years and within education, organisations such as the EEF are already synthesising evidence about the effectiveness of different educational interventions[*]. We believe that an EIDL would complement and add to this work by widening access to robust evaluations in a very cost-effective way. Moreover, we think it is important to emphasise this particular benefit of the EIDL so that it is not just seen as a device for individual organisations to improve their market position, but rather that it delivers a social good, from which all providers will ultimately benefit.

---

[*] So far EEF has supported 44 projects to evaluate their impact, largely through Randomised Controlled Trials (RCTs).

RCTs are a study design that randomly assigns people to different groups to eliminate the possibility of external factors affecting the results so that better comparisons can be made.

Summary of benefits

As summarised in the table below, there are a number of benefits from developing an EIDL for service users, schools, policymakers, researchers, commissioners and ultimately children and young people themselves.

For EIDL users…	For government data holders…
• Provides a clear route to accessing government data. • Improves understanding of beneficiary outcomes. • Informs service development. • Responds to commissioner requirements.	• Provides more accessible and tailored access to data through a secure mechanism, thereby reducing the number of specific requests. • Allows requests for and uses of education data to be tracked, and feedback to be sought from users. • Route to engage with charity sector and learn from their work and expertise.
For policy makers, researchers and commissioners…	**For beneficiaries…**
• Supports evidence-based commissioning. • Routinely produces high-quality, low-cost evaluations of social policies and programmes. • As more organisations use the EIDL, it would be possible to conduct meta-analyses to understand what works. • Reorientation of focus and efforts, away from collecting duplicative data towards more useful data analysis, synthesis, interpretation and qualitative research.	• Alongside other forms of research, ultimately results in better outcomes for service users as the most effective programmes are provided. • Their time and efforts are spent with services about which there is greater confidence of positive outcomes.

POTENTIAL DEMAND FOR THE EIDL

Estimating the number of potential users

Providers of education services vary a lot by activities provided, size and maturity, and there are no infrastructure organisations coordinating or overseeing all this activity. This makes it difficult to estimate the number of organisations that would be suitable for an EIDL evaluation.

One estimate, calculated by the National Council for Voluntary Organisations (NCVO), and based on the International Classification of Non Profit Organisations, suggests there are 7,847 education charities in England.[18] However this figure is only indicative because the Charity Commission registration process does not exactly capture what we are looking for and does not include social enterprises. Moreover, while NPC is mainly interested in the charity and social enterprise sector, the proposed service could support public or private sector organisations—including schools. So the figure could be an underestimate of the total number of potential users.

Another way to look at it is to compare with the criminal justice sector, which already has an established data lab service. Around 234,000 people were supervised by probation at the end of September 2015[19], which compares to around one million pupils in just year 6 and year 11 of the school system as of January 2015[20]. Over 140 interventions have been evaluated through the Justice Data Lab, so a rather crude equation based on these figures would suggest that we should expect an even higher level of demand for the EIDL.

We expect that a key challenge will be ensuring that the service is advertised and promoted within the sector. The Justice Data Lab would not have been successful without the support of Clinks (the membership organisation for charities in the criminal justice sector). No similar organisation represents the *whole* sector in education, so in planning the EIDL serious thought will need to be given to how to market it. An important lesson from the Justice Data Lab is that, given the complexity of some of the ideas involved, it is not enough simply to tell people about the service, they also need encouragement to see the potential and have an opportunity to talk through the process and implications of using it.

Estimating demand amongst potential users

As well as considering the potential number of users, we also need to consider whether they would want to use the EIDL. Our survey gives us some indications of this. For example, we found that:

- **There is demand for this kind of data**. Most respondents (89%) had previously tried to access education data from at least one source. This included from schools (61%), teachers (48%) and young people directly (48%).
- **There are challenges in accessing data**. Around half (54%) told us that accessing data from these sources was 'difficult', with consent (54%), expertise (37%) and time required (70%) cited as key barriers.
- **Difficulties accessing data is holding organisations back**. Over six in ten respondents (62%) agreed that 'Our ability to attract funding and scale our impact is held back by a lack of robust data on the impact of our work'.

A further indicator of the potential demand for the service was that **thirty-eight organisations (over half of all respondents) agreed to be a pilot for the service.**

Examples of organisations that could use the EIDL

As part of our scoping we discussed the EIDL in more detail with seven organisations. The case studies below summarise these discussions and explain the type of services they provide and their need for an EIDL service.

Impetus-PEF

Impetus-PEF is a funder which supports a portfolio of c. 25 education/employment readiness organisations that focus on raising aspirations and educational attainment in secondary schools. They also work closely with organisations to improve their evidence and impact measurement.

Need: They have found it difficult to source credible counterfactual data and have tended to rely on local benchmarking and within-school control groups which are prone to selection-bias. They see considerable potential in the EIDL for three main reasons:

- As a tool to support experimentation and testing new ideas.
- To test whether an intervention has potential before investing in a more robust study like a Randomised Control Trial (RCT).
- As a way to understand their beneficiaries better—to check if organisations are reaching the right target group.

Analysis required: Attainment (including Attainment 8 measures), attendance, behaviour (as measured by exclusions), value added, prior attainment and destinations data. Individual level results and comparison against local LA benchmarks would be helpful.

London Youth

London Youth is a network of community youth organisations serving young people of all backgrounds across the capital.

Their vision is that all young Londoners grow up healthy, able to navigate a fulfilling career and make a positive contribution in their communities. They support people to develop confidence, resilience and relationship skills, delivering with and through a network of community youth organisations and at two residential centres.

Need: London Youth is interested in how the EIDL could provide them with a more credible evidence base. In particular, they are interested in looking at the relationship between their data on short-term outcomes and data on longer-term outcomes that the EIDL offers.

Analysis required: Educational attainment and attendance are priorities, but will vary depending on the needs of specific programmes.

School-Home Support (SHS)

SHS aims to increase attendance and participation at disadvantaged schools. They focus on engaging parents to address any issues at home and improving parenting skills generally. SHS practitioners work in around 100 schools. They are also involved in the Troubled Families programme[21] and provide remote support. Potential sample size is thousands.

Need: SHS relies on teacher reported assessments of pupil progress and some school management data on attendance. They have no reliable data on whether their services make a difference and have never been able to access the National Pupil Database, despite efforts to do so. They have limited resources to commission help and only limited internal capacity for analysis.

Analysis required: Attendance, Key Stage 2 attainment, Key Stage 1-Key Stage 2 value added, exclusions.

The Future Leaders Trust

The Future Leaders Trust (TFLT) is a charity that aims to raise children's educational achievement regardless of background by developing a network of exceptional school leaders.

The Trust provides intensive leadership development for senior leaders, aspiring headteachers and academy trust CEOs, and is delivering the government's Talented Leaders programme for schools struggling to recruit a great headteacher.

TFLT currently operates in over 600 challenging schools and has supported over 160 headteachers to headship.

Need: While TFLT's school partners are able to provide some data to support impact measurement, and TFLT have access to the National Pupil Database, they have limited time or resources to process this data. This means they cannot do pupil level impact measurement and are instead restricted to publicly available, school level performance data. Arrangements for data collection with schools are ad hoc, resource intensive and prone to inconsistency. They are also interested in the potential of the EIDL to analyse data by demographic and social characteristics.

Analysis required: Attainment (including Progress 8 measures), destination and exclusions (as this represents an early outcome for any improving schools).

Sea Cadets

Sea Cadets is a national youth charity working with 14,000 10 to 18 year-olds. The charity operates in 400 communities across the UK and offers a safe and inspiring environment for young people to learn new skills through progressive training and nautical adventure, based on the customs and traditions of the Royal Navy. The skills and qualifications achieved and values attained help to boost confidence and achievement.

Need: Sea Cadets believe that they are effective at reaching and engaging disadvantaged young people, but have only limited options for testing this. Amongst other things, they aim to encourage participation in education and improve academic results—but they have no way of testing this either.

Analysis required: The demographic profile of cadets, attendance, behaviour (exclusions), attainment, Key Stage 2-Key Stage 4 value added. Data on the destinations of young people post-16 would also be helpful.

Prince's Trust

The Prince's Trust (PT) runs two programmes that aim to help young people into employment, education or volunteering opportunities:

- Excel: Focuses on 'soft skills' as a way to raise school attainment.
- Fairbridge programme: For under 16s who are out of mainstream education.

Need: Prince's Trust do not currently collect any routine data on attainment.
They have previously attempted to conduct a one-off matched control using the National Pupil Database data, but lacked the skills and quality of data to do this successfully. They see particular value in mapping 'hard' educational outcomes (such as attainment) with the 'soft' outcomes that they currently collect such as self-confidence, and the potential for analysis that looks at programme effectiveness for specific subgroups of young people. The also see the potential for the EIDL in exploring new funding opportunities such as social investment and payment-by-results.

Outcomes required: Ultimately interested in attainment, attendance and behaviour.

Teach First

Teach First (TF) is a charity that runs a number of different programmes in education, including:

- Leadership development: which aims to develop the leadership and quality of teachers.
- Innovation Unit: which incubates start-up programmes that aim to tackle educational inequality.
- Teach First Futures programme: which aims to support and inspire disadvantaged students to make ambitious and informed choices for higher education.
- The Fair Education Alliance: a collaboration with 26 other programme providers who are aligned with TF's mission of tackling educational inequality.

Need: Few of the organisations supported by the Innovation Unit or the Fair Education Alliance have access to the National Pupil Database, or the resources and expertise to construct an adequate control group. The EIDL could be potentially valuable for all these organisations.

Outcomes required: Attainment data is the main priority, but dimensions of analysis is key (ie, how the data is matched: class, school, LA and characteristics). Data on the destinations of young people post-16 would also be helpful.

HOW THE EIDL SERVICE COULD BE FUNDED

Funding options

We think there are three options for funding the EIDL.

1. Grant-funded service

Given the many benefits of the EIDL for the education system as a whole, we think there is a strong argument for a free-to-use service like the Justice Data Lab. Indeed, eight in ten of survey respondents (n=55) agreed with the statement that *'The government has a responsibility to ensure that not-for-profit providers have access to data so that they can measure their impact'.*

Funding could be sought from a range of external organisations such as the Department for Education, the EEF, the Cabinet Office and the wider research community, including charitable research funders.

2. Chargeable service

A chargeable service is an option, but it will affect demand and therefore the supply of evidence from the service. The findings from our consultation suggest that a minority of organisations would be willing to pay to use the EIDL. Four in ten (n=26) respondents said that they would be prepared to pay something if they could afford it, with a similar proportion unsure (n=24). One in five (n=12) said that they would not be willing to pay. While we suspect that this is an underestimate of willingness to pay because it is difficult for people to value something that has not been developed and is still quite vague, it does indicate that a service which is free (or low cost) would be preferable to a full chargeable model (at least in the early stages).

Another argument against a chargeable services is that it would weaken the case for publishing reports from the EIDL—the Ministry of Justice were able to justify transparency on the basis that the service is free.

3. A hybrid funding model

It might be possible to charge charities on a sliding scale based on size and income. This would be preferred to model two, but it could be difficult to administer as size of organisation and income are usually not clear cut.

Another option is to have free core service with charges for extra or more complex analysis. Many interviewees thought that answering bespoke questions would be a valuable part of the EIDL, but adding this facility to the core service would increase the time and costs of reports. Hence, a service that is part chargeable might be the best way to maximise use of the service while generating some revenue.

Summary

We feel there is a strong argument that a free or low-cost option because:

- Beneficiaries of the service would not just be providers, but also schools, government departments and other commissioners/funders of education/youth services.
- Findings would serve the public good by improving shared knowledge and the commissioning of services.
- A chargeable service would discourage smaller charities from using the EIDL.

CONCLUSION

The need

Schools are commissioning more services directly and are rightly focused on outcomes rather than activities. But the system cannot function well without reliable evidence about what is effective: commissioners want this evidence and provider organisations are keen to provide it. The data to do this exists, but organisations struggle to access it or conduct robust data analysis.

The solution

An Education Impact Data Lab (EIDL) has the potential to transform the way evidence informs policy and practice in the education sector. The ultimate benefit of developing an EIDL will be a more efficient and effective system of educational programmes for young people. Each analytical report would provide feedback on impact that can be used by the provider of the service and/or by commissioners to make decisions about what needs improving, and what services should be commissioned.

Over time, the collection of reports will create a powerful repository of evidence around what types of programmes work best. This will complement the work of organisations like EEF to expand the evidence base, ultimately leading to better decisions across the charity, education and youth sectors about how to use resources more effectively and improve educational outcomes of young people.

Many of the charities we surveyed share this vision. They are enthusiast about the contributing to the evidence base and are interested in learning from evidence of what works. As part of the service, organisations should be supported with their data collection systems and processes to enable even those organisations at early stages of impact evaluation to participate.

The risks

The main risk of the EIDL is around privacy and data protection. However, the EIDL is based on the Justice Data Lab, which in its three years of operation, has had no data protection objections from organisations that have used its services or complaints from privacy rights groups or others.

The next steps

Our initial research has shown strong support for the development of an EIDL from the charity sector. We recommend that:

- The service should be piloted for a year to develop the processes and metrics.
- The pilot service should be funded by commissioners/research funders, such as the DfE or a research organisation such as EEF.
- We should establish an advisory panel of users and researchers to feed into the design of the service.
- We should agree that service user consent is not needed for the EIDL (as per the Justice Data Lab).

- The service should be free-to-use during the pilot period, with the option of a charging structure introduced at a later date.

- Potential users of the service should start to consider whether they are ready to use it, in particular whether the data they are collecting is good enough to enable matches to the NPD.

You can help

Thank you to all the organisations who have already contributed time to support the development of the case for the EIDL. We need your continued support to:

- Apply pressure to funders to fund a pilot of the service.

- Talk to colleagues, other organisations and membership bodies about the potential of the EIDL to increase awareness and demand.

- Start to prepare your organisation to use the EIDL. In particular check that you are collecting data on service user's names, date of birth, school(s) attended and ideally Unique Pupil Number (UPN).

- Check that any information you give to service users states that you may share their data for the purposes of research, and gives them the opportunity to opt out. Keep in touch with developments of the EIDL where consent issues should become clearer.

- Get in touch with NPC to discuss the EIDL further.

To discuss any issues raised within this report, please get in touch with either George Hoare or James Noble via info@thinknpc.org or 020 7620 4850.

Visit www.NPCdatalabs.org for more about our Data Labs project.

ACKNOWLEDGEMENTS

REFERENCES

[1] Department for Education (2016) *Revised GCSE and equivalent results in England, 2014-2015.* www.gov.uk/government/uploads/system/uploads/attachment_data/file/494073/SFR01_2016.pdf

[2] Department for Work and Pensions (2014) *An evidence review of the drivers of child poverty for families in poverty now and for poor children growing up to be poor adults.* www.gov.uk/government/uploads/system/uploads/attachment_data/file/285389/Cm_8781_Child_Poverty_Evidence_Review_Print.pdf

[3] National Audit Office (2015*) Funding for disadvantaged pupils.* www.nao.org.uk/wp-content/uploads/2015/06/Funding-for-disadvantaged-pupils.pdf

[4] Curvers, S., Kail, A. and Wharton, R. (2016) *School report: How can charities work best in the school system?* New Philanthropy Capital.

[5] McNeil, B., Reeder, N. and Rich, J. (2012) *A framework of outcomes for young people.* The Young Foundation. http://youngfoundation.org/wp-content/uploads/2012/10/Framework-of-outcomes-for-young-people-July-2012.pdf

[6] Department of Education (2003) *Every Child Matters.* www.education.gov.uk/consultations/downloadableDocs/EveryChildMatters.pdf

[7] Department of Education (2016) *Educational Excellence Everywhere.* www.gov.uk/government/uploads/system/uploads/attachment_data/file/508447/Educational_Excellence_Everywhere.pdf

[8] www.gov.uk/guidance/pupil-premium-information-for-schools-and-alternative-provision-setting

[9] Sutton Trust and Education Endowment Foundation (2015) *The pupil premium: Next steps.* www.educationendowmentfoundation.org.uk/uploads/pdf/Pupil_Premium_Summit_Report_-_FINAL_EDIT.pdf

[10] Department of Education (2015) *The national pupil database: User guide.*

[11] Rickey, B. and Pritchard, D. (2012) *Unlocking offending data: Charities' views on how access to offending data could help them improve outcomes for offenders.* New Philanthropy Capital. www.thinknpc.org/publications/unlocking-offending-data/

[12] https://www.gov.uk/government/publications/justice-data-lab

[13] Sabri, F. and Gyateng, T. (2015) *Understanding statistical significance: A short guide.* New Philanthropy Capital. www.thinknpc.org/publications/understanding-statistical-significance/

[14] https://www.gov.uk/government/publications/unique-pupil-numbers

[15] See for example www.gov.uk/government/publications/the-magenta-book https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/62529/TLA-1906126.pdf

[16] Puttick, R. and Ludlow, J. (2012) *Standards of evidence for impact investing.* Nesta. /www.nesta.org.uk/publications/standards-evidence-impact-investing

[17] This type of analysis is already being conducted by EEF and FFT. See /www.educationdatalab.org.uk/Research-Section/Research.aspx

[18] NCVO (2015) *UK Civil Society Almanac.*

[19] Ministry of Justice (2016) *Offender Management Statistics Quarterly: July to September 2015. Tables Prison Releases: Q3 2015 and Probation Q3 2015.* www.gov.uk/government/statistics/offender-management-statistics-quarterly-july-to-september-2015

[20] Department for Education (2015) Schools, pupils and their characteristics: January 2015. National Tables: SFR16/2015. /www.gov.uk/government/statistics/schools-pupils-and-their-characteristics-january-2015

[21] https://www.gov.uk/government/policies/support-for-families

TRANSFORMING THE CHARITY SECTOR

NPC is a charity think tank and consultancy which occupies a unique position at the nexus between charities and funders, helping them achieve the greatest impact. We are driven by the values and mission of the charity sector, to which we bring the rigour, clarity and analysis needed to better achieve the outcomes we all seek. We also share the motivations and passion of funders, to which we bring our expertise, experience and track record of success.

Increasing the impact of charities: NPC exists to make charities and social enterprises more successful in achieving their missions. Through rigorous analysis, practical advice and innovative thinking, we make charities' money and energy go further, and help them to achieve the greatest impact.

Increasing the impact of funders: NPC's role is to make funders more successful too. We share the passion funders have for helping charities and changing people's lives. We understand their motivations and their objectives, and we know that giving is more rewarding if it achieves the greatest impact it can.

Strengthening the partnership between charities and funders: NPC's mission is also to bring the two sides of the funding equation together, improving understanding and enhancing their combined impact. We can help funders and those they fund to connect and transform the way they work together to achieve their vision.

www.thinkNPC.org