

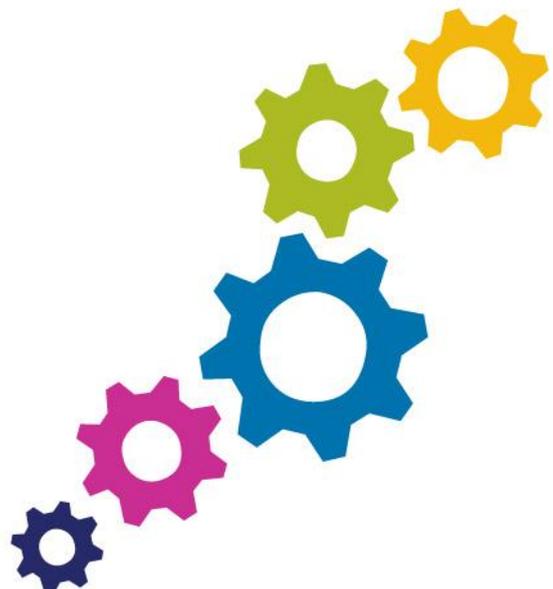


**Evidence for
Excellence in
Education**

Policy document

Evaluation Policy

National Foundation for Educational
Research (NFER)



● independent ● insights ● breadth ● connections ● outcomes

Contents

1	Introduction	1
2	Evaluation Fundamentals	1
3	Purposes and Methods of Evaluation	3
4	Contexts for Evaluation	5
5	Evaluation in Practice	6
6	Our Commitment to Achieving an Impact	13
	Background reading	15

1 Introduction

This document has been written by the National Foundation for Educational Research (NFER) to provide policy makers, practitioners, administrators, researchers and teacher educators with an overview of the principles NFER applies towards evaluation. These principles underpin the evaluation activities we undertake and our statements on national policy, whether responding to consultations or producing our own position papers and blog posts.

For over half a century, NFER's widely respected evaluations have informed educational policy and provided sound evidence in education debates and national reviews. The Evaluation Policy describes NFER's wealth of experience and our approach to methodology and design. It is illustrated by examples from the wide range of projects, large and small, that we have undertaken.

The national evaluation of the Technical and Vocational Educational Initiative (TVEI) ran from 1985 for over ten years. It built up a fund of evidence on the impacts of an enhanced curriculum with a specific technical and vocational focus for young people in the 14-19 age group. This paved the way for many subsequent evaluations in the school-to-work field, which remains a particular area of NFER evaluation specialism.

NFER's purpose is to provide evidence that improves education and learning. Evaluation plays a crucial role in national policy in supporting a system where decisions and recommendations are based on the best possible evidence.

2 Evaluation Fundamentals

NFER's evaluation studies investigate the implementation and impact of initiatives with the aim of informing decision making. They are conducted independently, systematically and ethically.

Evaluation is defined in the Government's Magenta Book as 'an objective process of understanding how a policy or other intervention is implemented, what effects it had, for whom, how and why' (HM Treasury, 2011, p. 11). In line with this, NFER evaluation can be defined as research into an identifiable initiative, intervention, programme or policy with an aim of appraising its effectiveness and informing decision-making. Thus this Evaluation Policy does not cover all of NFER's research activities: exploratory research, action research and self-evaluation consultancy, for example, are outside its remit.

Key characteristics of high quality evaluation are that it is independent and objective, systematic and ethical.

As an **independent** research organisation the Foundation is able to use its position to provide advice and shape projects without undue influence from political agendas or other initiatives. As an evaluation agency, NFER is independent of the development and

delivery of the initiative itself. The initiative is devised by the client or the client's delivery agency. It could be, for example, a national government policy, a new local government structure, an educational intervention, a training programme or a commercially published curriculum scheme. The NFER evaluation team stands apart from the delivery of the initiative, giving an outside perspective and a more **objective** view, although involvement early on in the design of an intervention allows for a more effective evaluation to be planned and delivered.

In designing an evaluation, the intended aims and impacts, nature, structure and stage of development of the initiative itself guide the work of evidence-gathering and reporting. These considerations define the research questions or hypotheses, which must be clarified and agreed at the outset. Based on this, research methods are selected from a wide repertoire and **systematically**

deployed to provide defensible and replicable evidence to address the questions identified. We collect data from a range of sources and stakeholders (including children, young people, parents and the wider community as well as professionals) to get a full picture of the processes and outcomes. Quality assurance procedures are in place, decisions are made on a reasoned basis and the progress of the research is recorded.

NFER's evaluation of Playing for Success, which aimed to raise standards for underachieving young people, showed significant improvements in attitude and motivation and some significant gains in mathematics and ICT. There was no clear evidence, however, of long-term gains in literacy attainment. Our independent, systematic investigations led to this objective assessment of its partial success.

The report on the evaluation of tailored consultancy support for local authorities by the National Youth Agency includes 11 case studies of named local authorities. These case study reports were agreed with research participants and published with their permission.

As part of NFER's research activities, evaluation projects adhere to the Foundation's Code of Practice, ensuring that data collection is conducted **ethically** and that high standards of consent, confidentiality and security are applied. Our reports are careful to set out the evidential basis for their conclusions and individual research participants remain anonymous unless they have given specific permission.

3 Purposes and Methods of Evaluation

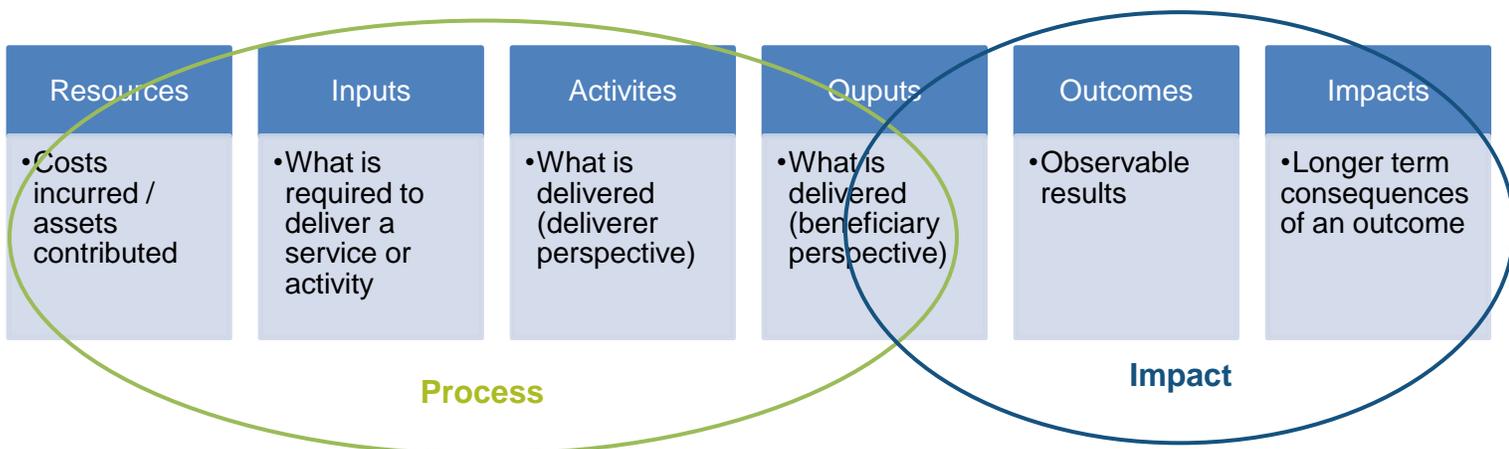
The purposes of evaluation fall broadly into the categories of *process* and *impact*. Methods can be categorised as *quantitative* and *qualitative*. The relationship between these two categorisations is complex, so that methods and purposes can be viewed as a cross cutting matrix. Many evaluations have more than one purpose and employ mixed methods. Value for money evaluation relates process and impact to the resources invested.

Purpose

The purpose of the evaluation is the bedrock upon which the research design is built. In general the purpose is to answer a number of specific research questions about the initiative.

In conceptualising the evaluation purpose, it is helpful to draw upon a logic model. This diagram (Figure 1) presents in schematic form the resources, inputs, activities, outputs, outcomes and impact by means of which an initiative is intended to achieve its goals.

Figure 1: Logic model



Correspondingly, *process* evaluation can be seen as an enquiry into the resources, inputs, activities and some of the outputs of an initiative, while *impact* evaluation has a focus on outputs, outcomes and impacts. **Process evaluation** addresses research questions about whether, how and how well the initiative was implemented, including the quality of involvement and the immediate reactions of the recipients. **Impact evaluation** aims to answer questions about the effects, both short-term and long-term, of participation in the initiative. It seeks to establish *additionality*, the effect of the intervention over and above what would have happened without it, and *attribution*, the extent to which any changes can be attributed to the intervention, rather than other factors.

Value for money evaluation sets process and impact evaluations in a wider context by enabling comparisons to be made with other initiatives.

Value for money evaluation considers how well the resources committed to an initiative have been utilised. It can incorporate cost-benefit analysis which translates outcomes and impacts into monetary values and relates them to the resources invested, measuring the return on investment an initiative generates.

In practice, most, if not all, evaluations have both process and impact elements and many also include an economic strand. However, it is possible for an evaluation to take a narrower focus, for example examining in depth just the activities and outputs of an intervention in order to recommend improvements. An evaluation of this kind is described as **formative** in purpose and is best suited to initiatives at an early stage of development.

Research methods

Cross-cutting with the purpose of the evaluation, a distinction can be made between *qualitative* and *quantitative* methods of data collection, each of which results in a different form of evidence. Broadly, **quantitative evidence** consists of measured variables based on large representative samples, whilst **qualitative evidence** captures detail and interactions drawn from smaller scale, but more in-depth, data collection. In NFER evaluations, we adopt qualitative and quantitative methods as appropriate, often in combination, to address the individual research questions.

The relationship between process and impact evaluation, on the one hand, and qualitative and quantitative methods, on the other, is not straightforward, and may be envisaged as a two-way matrix as shown in Figure 2.

Figure 2: Purposes of evaluation and types of evidence

Purpose	Process	Impact
Evidence		
Qualitative	Formative evaluation Experiences of deliverers and beneficiaries Observed facilitators and barriers to effectiveness	Reported and observed changes, information about the likely causes and reasons for the impact
Quantitative	Number of activities delivered and participants in initiative Survey of experiences of deliverers and beneficiaries	Measured attainment or progress Measured difference in attitude or engagement Measured change over time

4 Contexts for Evaluation

The features, intended beneficiaries and stage of development of the intervention shape the context for evaluation. Many evaluations have practical and budgetary constraints on their samples and methods.

The context for an evaluation is defined by the intervention itself and the policy and practice environment within which it is introduced.

Nature, aim and beneficiaries of initiative.

Ultimately, the intended beneficiaries in educational evaluations are usually children, young people or learners of all ages. The initiative may affect them directly, in the classroom or setting, or indirectly, for example through teacher development, structural changes or new policies. It may aim to bring about improvements in attainment or in attitudes, engagement or life choices.

NFER's evaluation of the professional development provided by the National Science Learning Centre gathered reports of the impact of its courses from teachers, technicians and pupils.

Our evaluation of City Challenge leadership strategies provided early evidence on new structures of system leadership in education.

Some initiatives are major national policies affecting a large proportion of the school population, while others are small-scale interventions targeted to bring about improvements in a specific area of focus.

Stage of development of initiative. At the earliest development stage of an intervention, the evidence collected may be small-scale and focus on the recipients' immediate responses, often to refine and develop the intervention further. Later, a larger-scale pilot may provide more systematic and objective evidence. Finally, a fully developed initiative may be subject to full-scale measurement of its impacts,

in comparison to non-participants, alongside a collection of evidence of the ways in which it is implemented. An individual initiative may be evaluated in a series of stages, with both process and impact findings feeding formatively into further development and evaluation. NFER's experience includes evaluations of initiatives at all stages of development.

The Phonics Screening Check for young children was fully rolled out nationally at the time of its three-year evaluation. Evidence included reported changes in response to its introduction, the views of stakeholders and an analysis of the literacy progress of pupils.

Practical constraints. The kind of evaluation we do is real-world research: it does not take place in a laboratory with carefully controlled conditions, but in functioning

classrooms and workplaces. The initiatives that we evaluate have been developed in specific practical and policy contexts, and the features of these contexts often define or constrain the nature of the study. There are two particular aspects of this that tend to be found in almost all of our evaluation projects.

Firstly, there may be constraints on the sample, resulting from pre-existing participation in the initiative. This places limits on the possibility of experimental¹ designs and the self-selected nature of the sample may weaken applicability of the evaluation evidence to people who are not 'early adopters' or enthusiasts.

Secondly, there are usually pre-existing financial assumptions about the cost of the evaluation, which may or may not be realistic in relation to the type and strength of evidence required of it.

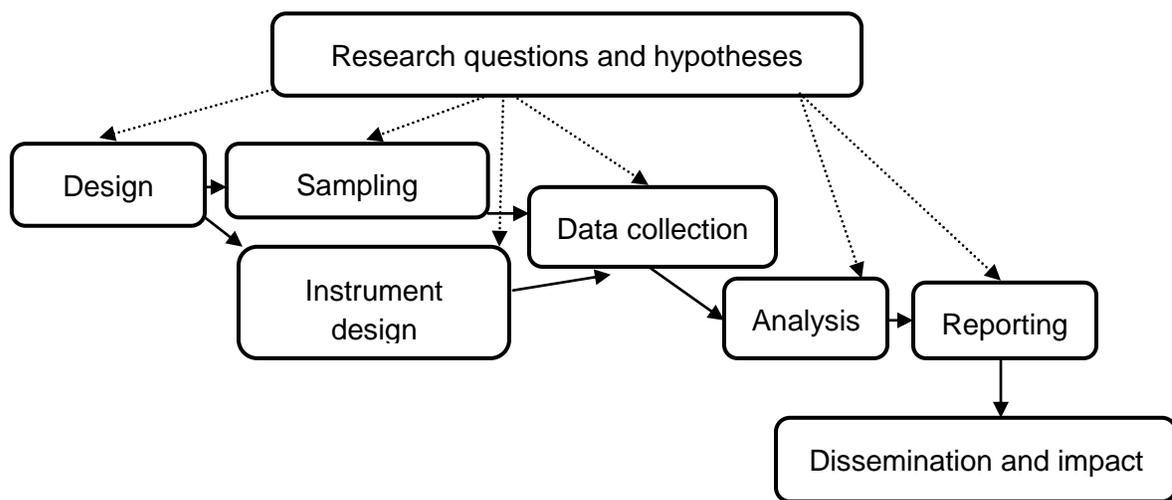
For these reasons, each evaluation project is individually designed. Any constraints or limitations are pointed out as part of a detailed research proposal. The aim is always to provide evidence that is fit for its purpose and to ensure that the claims made are firmly based on that evidence. We work with those commissioning evaluations to help ensure that the resources are proportionate to the strength of evidence that is required.

5 Evaluation in Practice

Evaluation in practice involves the processes of clarifying the research questions, designing the study, sampling, instrument development, data collection, analysis and reporting. The research questions shape and guide the evaluation throughout.

While every evaluation is individual and unique, it is possible to trace a number of processes or stages, most of which apply in most cases. These are summarised in Figure 3 and described in more detail below.

Figure 3: Elements of a typical evaluation



¹ Please see Section 5, *Design*, for a fuller discussion of this.

Scoping and clarification of research questions and hypotheses

The purpose of an evaluation is to provide answers to specified research questions, so the clear definition of those questions is an essential first step.

The research questions grow out of the context for the initiative, which includes what is already known about similar initiatives and the reason for the introduction of this one. Thus the **research and policy background** is an important element in the definition. Discussion with the client and the developer of the initiative at an early stage of the project also makes a key contribution.

Research questions for evaluation are often characterised by the formulation: What works?

However, this is too broad and loosely defined to underpin any meaningful study design, and must be broken down further. Generally, research questions for evaluation are variations of the more specific formulations: **In what sense(s) does it ‘work’? What works for whom, under what circumstances? How does it work? To what extent have the aims been achieved? Should we invest resource in this intervention?**

Sometimes, the research questions take the form of hypotheses, conjectures about causes and effects that can be investigated. Like research questions, clear and well-formulated hypotheses can form the basis for a well-designed evaluation study.

The evaluation of the In Harmony music programme developed a theory of change diagram in collaboration with the advisory group, to map out the intended impacts and define the research questions. This process involved clarifying the goals of the programme and working out key criteria for measuring impacts.

Study design

The design of the evaluation is underpinned by its purpose, in terms of the research questions to be investigated. Sometimes NFER works with partner organisations to provide a wider variety of complementary expertise. Study designs take account of real-world constraints and offer pragmatic proposals that will provide the best possible evidence within the context.

Many large-scale evaluations of high-profile national initiatives require the **measurement of impacts**. These impacts may be expressed, for example, in terms of pupil attainment, pupil/practitioner attitudes, wider outcomes such as wellbeing or the destination of school leavers, practitioner behaviour or changes in perception. A successful evaluation of this kind requires valid measures – either a well-matched assessment instrument or a well-designed survey.

In most cases, the optimum design for evidence of measured impacts is a **randomised controlled trial (RCT)**, where participants are randomly assigned to receive, or not to receive, the intervention, in order to eliminate selection bias.

The evaluation of a reading comprehension intervention for the Education Endowment Foundation used a randomised controlled trial with pre- and post-tests of pupil attainment in literacy.

This is described as an **experimental** design. The conduct of RCTs requires evaluators who are fully independent of the intervention. NFER's Trials Unit provides this, together with knowledge of the detailed procedures for co-ordination, randomisation, study design, recruitment, data management and statistical analysis necessary for a successful RCT. Randomisation of participants controls for both observable and non-observable background characteristics. This feature enables a causal connection to be drawn with respect to any impact of the intervention, contributing to conclusions about additionality and attribution.

The evaluation of the Room to Read literacy programme in Nepal used a bespoke test designed by NFER, adapted and administered by local specialists to young children in programme and comparison schools at two time points. Through multilevel modelling it was possible to control for background factors and identify improvements in literacy skills associated with participation in the programme.

Randomisation is not suitable for all initiatives, for a variety of practical/technical reasons. For example, an initiative may require its participants to be selected on relevant criteria, rather than randomly. Where an RCT is not possible, it is desirable to adopt a design that measures and controls for observable background characteristics, using a comparison group with similar characteristics that does not receive the intervention (but is not selected through a randomisation process) – a **quasi-experimental** design. Statistical modelling makes it possible to isolate whether or not participation in the intervention is associated with

the desired outcomes. In both RCTs and non-random designs, measurement of the variable of interest at more than one time point gives greater reliability by providing a baseline measure against which change can be assessed. Measurement of impacts is an essential part of **value for money analysis**, when impacts are considered in relation to costs.

Although measured impacts are often a requirement of an evaluation, these measures do not elucidate the actions, mechanisms, dynamics, interactions, reasoning and detailed causes at work as the intervention brings about its effects. For this, the research design must include investigation of the processes at work – the inputs, activities and outputs in the logic model. **Process investigations** can take many forms, for example, surveys, observations, one-to-one interviews or focus groups, often in combination. They can provide evidence of how, where and how much of the intervention was delivered and of the reactions and views of the participants and other stakeholders.

There are also many examples of evaluations where measured impacts are not required at all. Instead, the focus is on the **views and perceptions of the recipients** and the changes in attitudes and actions they report, in qualitative terms. Such evaluation gives evidence of whether an intervention or change in policy is working as expected, in the view of those most closely

NFER's evaluation of Tomorrow's Engineers used interviews with pupils, teachers and those delivering the programme to investigate changes in young people's perceptions of engineering as a career.

involved, together with identifying any barriers to its effectiveness. Participants may give convincing accounts of the impacts of an initiative upon their own and others' practice, views and attitudes. Careful thematic analysis can illuminate reasons, causes and effects within the wider context, while individual case studies can exemplify good practice. These methods are often useful as part of a wider research design that provides quantitative evidence.

Instrument design

The evaluation of Project Maths in Ireland used a test developed by NFER drawing on questions in international surveys.

The development or selection of validated research instruments is a crucial part of any evaluation.

NFER's assessment specialists provide expert advice whenever a **measure of pupil attainment** is needed, and in some cases a customised test or

assessment is developed for the evaluation. It is essential to give careful consideration to the relationship between the test and the aim of the intervention: the test should be neither so close to the intervention that any findings lack wider application; nor so removed from the focus of the intervention that any improvement is unlikely to be captured.

Questionnaires are developed in light of the research agenda and the characteristics of the recipients: questionnaires for pupils or parents, for example, require different wording from those for teachers or other practitioners. NFER teams have a depth of expertise in presenting and adapting questions for different users, developed through extensive experience of administering surveys. New questions, where needed, are developed, reviewed and piloted before use.

Questionnaire development for the Talk About Alcohol evaluation drew upon many years' experience of administering surveys exploring smoking, drinking and drug use among school-aged children.

For interviews and focus groups, evaluation teams develop **structured or semi-structured interview schedules**, which may be developed from existing schedules and/or piloted before use. We also use **other formats** such as ranking diagrams, life grids or verbal protocols where they can add structure or depth to the data.

Sampling

To match the evaluation design, sampling can take a wide variety of forms. For a large-scale high-stakes RCT or quasi-experiment, NFER statisticians calculate the **sample size** necessary to have sufficient power to detect changes of the type predicted by the initiative. The NFER Schools Database makes it possible to draw **nationally**

The 2009 Tellus4 survey represented the views of 253,755 children and young people in school years 6, 8 and 10. The sample of 3,699 primary, secondary and special schools and Pupil Referral Units was drawn to be representative of schools in each of the 152 local authorities in England.

representative samples that are stratified by a number of relevant variables. A smaller scale intervention at an early stage of development, by contrast, may be evaluated in a **small number** of volunteer schools recruited by the developer.

Designing a complex evaluation requires specifying **several separate or overlapping samples**. There is often an intervention sample and a comparison sample. Beyond this, a sample of schools may be drawn as the basis of a survey sample of teachers, a survey sample of headteachers and a test and attitude survey sample of pupils, all of which are analysed separately to answer different research questions. From the original large samples of schools, a smaller number may be selected for in-depth study, extending outwards to parents or local service providers. Longitudinal evaluations, where the data is collected repeatedly across a number of years, often include a number of cohorts of pupils who experience the intervention in successive academic years, and are then followed over time.

Data collection

Quantitative data collection requires detailed management, which at NFER is undertaken by our Research and Product Operations department. Research and Product Operations specialise in working with schools and have comprehensive survey planning and record-keeping systems for online and paper surveys, test trialling and visiting administrators. Research and Product Operations keep a database of school involvement in research projects and seek agreement from participants and take responsibility for the despatch and return of all data collection instruments, and for data entry and preparation. These procedures help ensure high response rates and **datasets that are complete, comprehensive and error-free**, an essential precursor of analysis.

Qualitative data collection requires expert judgement and is undertaken by our experienced research teams, sometimes augmented by research associates with specialist knowledge. Records of visits or telephone or online interviews are made in electronic or paper form to support accurate analysis, which is often structured by the agreed themes of the research agenda.

For both qualitative and quantitative data collection, the Foundation is committed to exploring and using **new technology** as it becomes available, and this forms an intrinsic part of the repertoire of methods. Online surveys are regularly used instead of or alongside paper versions, to suit the needs of the recipients. Audience response system (ARS) keypads are used to support focus groups. Webinars, virtual research environments and online focus groups have been used where appropriate in projects. A

ARS was used with three-year-old children as part of a study of the impact of Integrated Children's Centres in Wales. Using the handsets was a novel experience for the children and the fact they could see the results projected on the screen in front of them helped to enliven the discussion.

Methodology Working Group has ongoing responsibility for reviewing, piloting and introducing new methods.

Throughout all our data collection activities, we apply a series of processes to ensure that data is kept confidential and managed in a way consistent with our legal and ethical framework.

Analysis and emerging findings

Expert and meaningful analysis is at the heart of effective evaluation, transforming raw data into meaningful answers to the original research questions. This stage is shaped by the nature of the data in combination with the aims of the enquiry, and often proceeds through exploratory and confirmatory stages as patterns gradually emerge.

A wide range of **statistical techniques** is employed for the analysis of quantitative data, for example: descriptive analysis; multivariate analysis, including multi-level modelling and factor analysis; and psychometrics for test data, including Item Response Theory and equating.

Qualitative data progresses from detailed notes and illuminating quotations to insights about the **themes, relationships and patterns** revealed by the responses. Analytic conceptual frameworks are developed to reflect the research questions and emerging empirical findings. For larger datasets we use a qualitative data analysis package to assist in identifying common themes.

The emerging findings from the evaluation of Post-16 Work Experience Placement Trials were discussed in meetings with DfE staff and college practitioners. This process helped to draw out implications for policy and practice.

As patterns and findings begin to emerge, the research team meets to exchange impressions and insights. Gradually, the **key findings** are articulated in the course of these discussions and the evidence for them is set out and tested.

Reporting

At the report stage, the evaluation comes full circle, providing the developer or policy maker with answers to the research questions about their intervention or policy, together with the supporting evidence on which the conclusions are based. There may be a series of interim reports, each of which feeds its findings back formatively into the development of the initiative, in addition to a final summative report of conclusions.

In any report, the research team has to present the data in a way that is suited to its **audience and purpose** and which is **accurate, brief and clear**. Whilst summary data is essential to understand the overall messages, the evidential basis for the findings must remain available in detail for further interrogation if required.

These considerations lead to a very wide range of reporting styles and formats, from a brief set of slides or a glossy summary brochure to a lengthy and conventionally

structured research report or academic article. As the emerging findings become clearer, the format and approach for the report are also discussed with the client so that the messages from the evaluation emerge clearly, supported by key evidence.

6 Our Commitment to Achieving an Impact

NFER works to ensure that insights from evaluation studies are influential in improving policy and practice.

NFER's purpose is to provide independent evidence to inform improvements in education and the lives of learners. Evaluation projects are ideally placed to offer **meaningful insights on education that can influence policy and practice.**

Each initiative or policy that we evaluate can be expected to aim at some improvement, somewhere across the world of education or children's services. The evaluation evidence we collect and report provides a rich fund of information that can be turned to the use of practitioners and policy makers, and to the benefit of learners. The conclusions of our reports identify how effective initiatives have been, in what way, for whom and under what circumstances. They include recommendations for policy and practice with a firm research basis.

To support this impact, we work with clients throughout evaluation projects to ensure that insights and recommendations reach those who can make use of them. We are committed to keeping informed about the latest developments in education to ensure that our messages are as timely and as relevant as possible. There may be messages for those developing or implementing the initiative that will allow them to make changes to enhance its effectiveness. For this purpose, a brief summary of key messages at an interim point in the evaluation may be most appropriate. Alternatively, the evaluation may provide an opportunity for creating wider impact beyond the intervention itself. We may, for example, produce additional outputs beyond the main report that are targeted to particular groups of practitioners, such as guidelines for headteachers, or host events to enable a range of stakeholders to discuss the findings. Agencies responsible for professional and leadership development also play a role. Through engagement in this process, those directly involved can think through the implications for policy and practice and plan for implementing change.

We are also conscious that impact from our research is not created solely through effective communication of findings. An infrastructure is required that enables practitioners to systematically access the best available evidence on topics that are important to them (including for example the role NICE plays for the medical profession). Equally importantly, practitioners need to be equipped as critical consumers of research, and in the first place to believe in the value of evaluation findings to support their practice. NFER has an important role to play in this process and is supporting it in a variety of ways. For example, one approach to enabling teachers to engage with evaluation findings is by undertaking their own process of

NFER Thinks on STEM initiatives brought together evidence from a wide range of separate evaluations to draw general lessons about enhancing pupil engagement in this important area.

enquiry, focussed on identifying, implementing and testing the applicability of solutions to their own classroom challenges and aspirations. We have a range of products and resources for schools to use, including our Enquiring Schools programme, our Research Mark and our web guidance site for schools wishing to undertake their own enquiry.

Many of the messages from our research are reinforced across multiple projects, which places us in a unique position to synthesise findings for policymakers and practitioners, and to provide commentary on current issues for debate. The internet provides instant ways of sharing these valuable insights, through online position papers, blog posts, podcasts and social media. We continue to seek out and develop ways in which evaluation findings can be shared most effectively with stakeholders, maximising their impact for the benefit of learners and providing evidence for excellence in education.

Background reading

Reference works on evaluation and research

Cohen, L., Manion, L. and Morrison, K. (2007). *Research Methods in Education*. Sixth edn. Abingdon: Routledge.

HM Treasury (2011). *The Magenta Book: Guidance for Evaluation*. London: HM Treasury [online]. Available: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/220542/magenta_book_combined.pdf [4 December, 2013].

Palfrey, C., Thomas, P. and Phillips, C. (2012). *Evaluation for the Real World: the Impact of Evidence in Policy Making*. Bristol: Policy Press.

Robson, C. (2002). *Real World Research*. Second edn. Oxford: Blackwell.

Sharp, C and Nelson, J. (2012). *Process Evaluation*. Unpublished report.

St. Clair, R. (Ed) (2009). *Education Science: Critical Perspectives*. Rotterdam: Sense Publishers.

Project reports used as illustrative examples

Chamberlain, T., George, N., Golden, S., Walker, F. and Benton, T. (2010). *Tellus4 National Report* (DCSF Research Report 218). London: DCSF [online]. Available: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/221932/DCSF-RR218.pdf [4 December, 2013].

Education Endowment Foundation (2013). *Literacy Catch-Up Projects* [online]. Available: <http://educationendowmentfoundation.org.uk/projects/literacy-catch-up-projects> [4 December, 2013].

Griffiths, J. (2003). *NFER: The First Fifty Years 1946–1996*. Slough: NFER.

Jeffes, J., Jones, E., Wilson, M., Lamont, E., Straw, S., Wheeler, R. and Dawson, A. (2012). *Research into the Impact of Project Maths on Student Achievement, Learning and Motivation*. Slough: NFER [online]. Available: <http://www.ncca.ie/en/Conference/Impact-of-Project-Maths-Final.pdf> [4 December, 2013].

Lynch, S., Styles, B., Dawson, A., Worth, J., Kerr, D. and Lloyd, J. (2013). *Talk About Alcohol: an Evaluation of the Alcohol Education Trust's intervention in Secondary Schools*. Slough: NFER [online]. Available: <http://www.alcoholeducationtrust.org/pdfs/NFERfinal.pdf> [4 December, 2013].

Macleod, S., Straw, S. and Dawson, A. (2012). *Executive Summary of Evaluation of Tomorrow's Engineers 2011/12*. Slough: NFER [online]. Available:

http://www.engineeringuk.com/resources/data/Executive_summary_of_evaluation_of_Tomorrows_Engineers_FINAL1.pdf [4 December, 2013].

Rudd, P., Poet, H., Featherstone, G., Lamont, E., Durbin, B., Bergeron, C., Bramley, G., Kettlewell, K. and Hart, R. (2011). *Evaluation of City Challenge Leadership Strategies: Overview Report*. Slough: NFER [online]. Available: http://eprints.whiterose.ac.uk/73938/1/city_challenge_leadership_strategy_overview_evaluation_report_29072011_1_.pdf [4 December, 2013].

Sharp, C., Chamberlain, T., Morrison, J. and Filmer-Sankey, C. (2007). *Playing for Success: an Evaluation of its Long Term Impact* (DfES Research Brief 844). London: DfES [online]. Available: <http://webarchive.nationalarchives.gov.uk/20130401151715/https://www.education.gov.uk/publications/eOrderingDownload/RB844.pdf> [4 December, 2013].

Sims, D., Southcott, C., Lynch, S. and Wade, P. (2013). *Evaluation of the Work Experience Placement Trials*. London: DfE [online]. Available: <http://www.nfer.ac.uk/nfer/publications/WEXX01/WEXX01.pdf> [4 December, 2013].

Southcott, C., Stevens, E., Kettlewell, K. and Sims, D. (2012). *Evaluation of the NYA Tailored Support Offer to Local Authorities: Case-studies Report* (LGA Research Report). Slough: NFER [online]. Available: <https://www.nfer.ac.uk/nfer/publications/NYAS01/NYAS01.pdf> [4 December, 2013].

Straw, S. and MacLeod, S. (2013). *Improving Young People's Engagement with Science, Technology, Engineering and Mathematics (STEM)* (NFER Thinks: What the Evidence Tells Us). Slough: NFER [online]. Available: <http://www.nfer.ac.uk/nfer/publications/99927/99927.pdf> [4 December, 2013].

Walker, M., Bartlett, S., Betts, H., Sainsbury, M. and Mehta, P. (2013). *Evaluation of the Phonics Screening Check: First Interim Report* (DfE Research Report 286A). London: DfE [online]. Available: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/198994/DFE-RR286A.pdf [4 December, 2013].

Walker, M., Straw, S., and George, N. (2012). *Qualitative Evaluation of the National Science Learning Centre: Interim Report*. Slough: NFER.

NFER provides evidence for excellence through its independence and insights, the breadth of its work, its connections, and a focus on outcomes.

- independent
- insights
- breadth
- connections
- outcomes

© 2014 National Foundation for Educational Research

**National Foundation for
Educational Research**
The Mere, Upton Park
Slough, Berks SL1 2DQ

T: 01753 574123
F: 01753 691632
E: enquiries@nfer.ac.uk
www.nfer.ac.uk