British What Works Centres: what lessons for evidence-based policy in France?

March 2017

Translation funded by

Alliance for Useful Evidence

In partnership with
British What Works Centres: what lessons for evidence-based policy in France?

March 2017

This report was produced by Caroline Allard and Ben Rickey, project managers at the Agence nouvelle des solidarités actives in partnership with France Stratégie, the Secrétariat général pour la modernisation de l’action publique (SGMAP), the Commissariat général à l’égalité des territoires (CGET), Santé publique France, the Caisse nationale des allocations familiales (Cnaf) and the Caisse nationale de solidarité pour l’autonomie (CNSA).

In partnership with
# British What Works Centres: what lessons for evidence-based policy in France?

## Study partners

## Foreword

## Executive summary

## The report

### Introduction

### An international issue: making public policy more “evidence-based”

### “Evidence-based” public policy

### Producing a good standard of evidence

#### Experimental and quasi-experimental studies: towards higher standards of evidence

#### Economic evaluations: measuring the efficiency of interventions

#### Systematic literature reviews

### When the State promotes RCTs and quasi-experimental studies

### For a successful use of evidence by professionals and policymakers at all levels

#### Issues around “knowledge mobilisation”

#### Sharing the results of trials and systematic literature reviews in the form of guidelines and practices

#### Promoting a culture of evidence among policymakers and professionals

#### Creating organisations to strengthen the link between evidence and public policy

## British What Works Centres: a response?

### Presentation of What Works Centres

#### The general principles

#### Synthetic presentation

#### Background

#### Governance

### What the centres do

#### Synthesising existing research

#### Dissemination and promoting uptake

#### Evaluating the impact of innovations

### Key findings on What Works Centres

### What culture of evaluation and evidence for public policy in France?

### An administration focussed on needs analysis

### Evaluating public policy

### The development of RCTs and quasi-experimental studies

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Study partners</td>
<td>4</td>
</tr>
<tr>
<td>Foreword</td>
<td>10</td>
</tr>
<tr>
<td>Executive summary</td>
<td>11</td>
</tr>
<tr>
<td>The report</td>
<td>16</td>
</tr>
<tr>
<td>Introduction</td>
<td>17</td>
</tr>
<tr>
<td>An international issue: making public policy more “evidence-based”</td>
<td>23</td>
</tr>
<tr>
<td>“Evidence-based” public policy</td>
<td>23</td>
</tr>
<tr>
<td>Producing a good standard of evidence</td>
<td>24</td>
</tr>
<tr>
<td>Experimental and quasi-experimental studies: towards higher standards of evidence</td>
<td>26</td>
</tr>
<tr>
<td>Economic evaluations: measuring the efficiency of interventions</td>
<td>30</td>
</tr>
<tr>
<td>Systematic literature reviews</td>
<td>31</td>
</tr>
<tr>
<td>When the State promotes RCTs and quasi-experimental studies</td>
<td>33</td>
</tr>
<tr>
<td>For a successful use of evidence by professionals and policymakers at all levels</td>
<td>35</td>
</tr>
<tr>
<td>Issues around “knowledge mobilisation”</td>
<td>35</td>
</tr>
<tr>
<td>Sharing the results of trials and systematic literature reviews in the form of guidelines and practices</td>
<td>37</td>
</tr>
<tr>
<td>Promoting a culture of evidence among policymakers and professionals</td>
<td>39</td>
</tr>
<tr>
<td>Creating organisations to strengthen the link between evidence and public policy</td>
<td>40</td>
</tr>
<tr>
<td>British What Works Centres: a response?</td>
<td>43</td>
</tr>
<tr>
<td>Presentation of What Works Centres</td>
<td>43</td>
</tr>
<tr>
<td>The general principles</td>
<td>43</td>
</tr>
<tr>
<td>Synthetic presentation</td>
<td>45</td>
</tr>
<tr>
<td>Background</td>
<td>47</td>
</tr>
<tr>
<td>Governance</td>
<td>49</td>
</tr>
<tr>
<td>What the centres do</td>
<td>50</td>
</tr>
<tr>
<td>Synthesising existing research</td>
<td>50</td>
</tr>
<tr>
<td>Dissemination and promoting uptake</td>
<td>58</td>
</tr>
<tr>
<td>Evaluating the impact of innovations</td>
<td>66</td>
</tr>
<tr>
<td>Key findings on What Works Centres</td>
<td>69</td>
</tr>
<tr>
<td>What culture of evaluation and evidence for public policy in France?</td>
<td>72</td>
</tr>
<tr>
<td>An administration focussed on needs analysis</td>
<td>72</td>
</tr>
<tr>
<td>Evaluating public policy</td>
<td>73</td>
</tr>
<tr>
<td>The development of RCTs and quasi-experimental studies</td>
<td>78</td>
</tr>
<tr>
<td>Rapid development of these new methods</td>
<td>78</td>
</tr>
<tr>
<td>An ongoing movement but with relatively low output of few RCTs and quasi-experimental studies</td>
<td>81</td>
</tr>
</tbody>
</table>
RCTs and quasi-experimental studies as decision-making tools
Trials and quasi-experimental studies sometimes inform policy decisions
Systematic reviews are not well established in social policy in France

Public services on the whole little informed by robust evidence
Policymakers little informed by RCTs and quasi-experimental studies
“Effective” practice is difficult for practitioners to access

Scale-up support is under-developed

The key findings on evaluation and evidence in France

What Works Centres: what are the lessons for France?

What Works Centres Methods
Benefits
Challenges

The What Works Centres as organisations
Benefits
Challenges

Does the French context lend itself to the “What Works’ model?

Proposals
Proposal 1: Explore the opportunity of labelling or creating “What Works’ inspired initiatives or structures in France
Proposal 2: Create an alliance around evidence-based policies

Case studies of 5 What Works Centres
Early Intervention Foundation
Education Endowment Foundation
What Works Centre for Local Economic Growth
Centre for Ageing Better
Public health at the heart of the National Institute for Health and Care Excellence (NICE)

Perspectives
France Stratégie
Secrétariat général pour la modernisation de l’action publique (SGMAP)
Commissariat général à l’égalité des territoires (CGET)
Santé publique France
Caisse nationale des allocations familiales (Cnaf)
Caisse nationale de solidarité pour l’autonomie (CNSA)
Study partners

- **France Stratégie**, a government think tank that reports to the Prime Minister, and provides analysis, expertise and consultation to inform the future of public services...
- **The Secrétariat général pour la modernisation de l’action publique** (SGMAP), which helps government departments to deliver transformational policies through the implementation of innovative initiatives, a user focus, the evaluation of public policies, rolling out digital technology and promoting more transparent and collaborative public action.
- **The Commissariat général à l’égalité des territoires** (CGET), responsible for designing and implementing the national regional equality policy and for overseeing and coordinating it across government.
- **Santé publique France** is the national public health agency. With expertise in epidemiology, prevention and targeted intervention, the agency has a broad remit ranging from fact-finding to action.
- **The Caisse nationale des allocations familiales** (Cnaf) manages the family branch of French social security and has a country-wide presence through its departmental network of offices (caisses d’allocations familiales – Caf).
- **The Caisse nationale de solidarité pour l’autonomie** (CNSA) funds and supports the implementation of policies to promote independence among dependent older people and those with disabilities.

Members of the steering committee

*France Stratégie*: Marine Boisson-Cohen

*SGMAP*: Jean-Baptiste Bodin, Clément Lacouette-Fougère et Benoît Landau

*CGET*: Emmanuel Dupont

*Santé publique France*: Pierre Arwidson et Isabelle Vazeilles

*Cnaf*: Arthur Heim

*CNSA*: Julie Micheau et Amaëlle Penon
Acknowledgements and list of people interviewed

Ansia would like to thank all of those who contributed to this study. However, only the authors of the report are responsible for its content.

- People interviewed in the United Kingdom

**Alliance for Useful Evidence**
Pippa Coutts
Manager, Scotland

**Behavioural Insights Team**
David Halpern
Chief Executive

**Cabinet Office**
Alice Dowswell
Head of the What Works Team

**Centre for Ageing Better**
Dan Jones
Director of Innovation and Change

**Centre for Policy Development**
Martin Stewart-Weeks
Board Director

**Crisis**
Ligia Teixeira
Head of Research and Evaluation

**Early Intervention Foundation**
Kirsten Asmussen
Evidence Analyst

**Economic and Social Research Council**
Karen Powell
Senior research portfolio manager

**Education Endowment Foundation**
Danielle Mason
Head of Research and Publications

**Education Endowment Foundation**
Margaret Macadam
Deputy Director of evaluation, strategy, and analysis

**Education Endowment Foundation**
Jonathan Sharples
Senior Researcher

**Education Endowment Foundation**
James Turner
Deputy Chief Executive

**Education Endowment Foundation**
Camilla Neill
Evaluation Manager

**Education Endowment Foundation**
Emily Yeomans
Grants Manager

**Economic and Social Research Council**
Karen Powell
Senior research portfolio manager

**Economic and Social Research Council**
Margaret Macadam
Deputy Director of evaluation, strategy, and analysis

**EPPI-Centre**
David Gough
Director

**National Institute for Health and Care Excellence**
Beth Shaw

**National Institute for Health and Care Excellence**
Alexia Tonnel
Director of Evidence Resources

**Nesta**
Geoff Mulgan
Chief Executive

**Public Health England**
Gregor Henderson
National Lead for Mental Health and Wellbeing

**Public Health England**
Felix Greaves
Deputy Director, Science and Strategic Information
British What Works Centres: what lessons for evidence-based policy in France?

Public Health England
Alexia Clifford
Deputy Director, Marketing Activation

Public Health England
Elaine Rashbrook
Interim Deputy Director
Healthy People/National Lead for Older People

Public Health England
Justin Varney
National Lead for Adult Health & Wellbeing (Healthy People)

Public Health England
Tim Chadborn
Behavioural insight team

Public Policy Institute for Wales
Dan Bristow
Deputy Director

The Alliance for Useful Evidence
Jonathan Breckon
Director, Alliance for Useful Evidence

University of Cambridge
Mike Kelly
Former Head of Public Health Centre at NICE

What Works Centre for Local Economic Growth
Max Nathan
Research Fellow – Urban Programme

What Works Centre for Crime Reduction
Rachel Tuffin
College of Policing

What Works Centre for Local Economic Growth
Meg Kaufman
Project Manager

What Works Centre for Local Economic Growth
Rita Beden
External Affairs Manager

What Works Centre for Wellbeing
Nancy Hey
Director

What Works Scotland
Sarah Morton
Co-Director, Knowledge Exchange, Centre for Research on Families and Relationships

Institute of Global Health Innovation
Francoise Cluzeau
Associate Director
Global Health & Development Group ex NICE

People interviewed in France

Assistance publique -Hôpitaux de Paris
Martin Hirsch
Director

Caisse nationale d’allocations familiales
Bernard Tapie
Director of statistics, studies and research

Conseil national d’évaluation du système scolaire
Nathalie Mons
Chair

Cour des comptes
Jérôme Filippini
Secretary General

Cour des comptes
Barbara Falk
Deputy Secretary General

Cour des comptes
Laetitia Buffet
Project Manager reporting to the First Chair

Délégation interministérielle pour l’hébergement et l’accès au logement
Pascale Estecahandy

National Coordinator
Un chez-soi d’abord

Direction de la recherche, des études, de l’évaluation et des statistiques
Franck von Lennep

Direction de l’animation de la recherche, des études et des statistiques
Philippe Zamora
Deputy Director for the Tracking and Evaluation of Employment and Training Policies
Agence nouvelle des solidarités actives

Direction de l’évaluation, de la prospective et de la performance du ministère de l’Éducation nationale
Olivier Cosnefroy
Head of the Office for the evaluation of Educational initiatives and trials

Direction de l’évaluation, de la prospective et de la performance du ministère de l’Éducation nationale
Daniel Auverlot
Deputy director of Evaluations and attainment

Direction de l’évaluation, de la prospective et de la performance du ministère de l’Éducation nationale
Fabienne Rosenwald
Director of evaluation, Forecasts and performance

Direction générale de l’emploi de la formation professionnelle
Tristan Klein
Deputy head of unit for the delivery of ESF Programmes

Direction générale du Trésor
Michel Houdebine
Chief economist

École d’Économie de Paris
Marc Gurgand
Scientific Director
J-PAL Europe

Direction de la jeunesse, de l’éducation populaire et de la vie associative
Jean-Benoît Dujol
Director

Institut national de la jeunesse et de l’éducation populaire
Thibaut de Saint Pol
Director

France Stratégie (Commissariat général à la stratégie et à la prospective)
Fabrice Lenglart
Deputy Commissioner General

France Stratégie (Commissariat général à la stratégie et à la prospective)
Rosenn Desplatzh
Head of Public Policy evaluation

Inspection générale des affaires sociales
Stéphane Paul
Inspector General

Institut des politiques publiques
Antoine Bozio
Director

Laboratoire interdisciplinaire d’évaluation des politiques publiques – LIEPP
Bruno Palier
Director of Research
CNRS, co-director of LIEPP

Laboratoire interdisciplinaire d’évaluation des politiques publiques – LIEPP
Juliette Seban
Secretary General

Ministère de l’Éducation nationale
Olivier Noblecourt
Deputy Director of the cabinet

Plein Sens
Caroline Moriceau
Director of Studies

Secrétariat général pour la modernisation de l’action publique
Laure de la Bretèche
Secretary General for public Sector modernisation

Secrétariat général pour la modernisation de l’action publique
Clement Lacouette-Fougère
Evaluation officer

Société française de l’évaluation
Virginie Besrest
Board member

Université Marne-la-Vallée
Yannick L’Horty
Director of the research Federation “Work, Employment and Public Policies”
Participants at the workshop on 4 October 2016

- Caisse nationale d’allocations familiales (Cnaf)
  Arthur Heim
  Social investment officer

- Caisse nationale de solidarité pour l’autonomie (CNSA)
  Julie Micheau
  Scientific Director

- Commissariat général à l’égalité des territoires (CGET)
  Emmanuel Dupont
  Head of Scientific coordination and innovation

- Direction de l’évaluation, de la prospective et de la performance du Ministère de l’Éducation nationale
  Pascal Bessonneau
  Statistician

- Fonds d’expérimentations pour la jeunesse (Fej)
  Malika Kacimi
  Head of Unit

- Institut français de l’éducation
  Catherine Pérotin
  Deputy Director

- Institut français de l’éducation
  Marie-Claire Thomas

- Laboratoire interdisciplinaire d’évaluation des politiques publiques (LIEPP)
  Juliette Seban
  Secretary General

- Organisation for Economic Co-operation and Development (OCDE)
  Laura Skoratko
  Policy analyst

- Secrétariat d’État en charge des personnes handicapées et de la lutte contre l’exclusion
  Léopold Gilles
  Tackling Exclusion Advisor

- Direction générale de la cohésion sociale (DGCS)
  Michèle Arnaudie
  Social innovation advisor
### Participants at the thematic workshops

**Cnaf (on the topic of early childhood)**

- Sandrine Dauphin  
  Head of the Research Coordination and researcher network department
- Arthur Heim
- Danielle Boyer  
  Head of the Observatoire national de la petite enfance
- Laurent Ortalda  
  Head of early childhood at the Dep

**CNSA**

- Julie Micheau  
  Scientific Director

**Santé Publique France**

- Pierre Arwidson  
  Deputy director for preventative health and campaigns
- Isabelle Vazeilles

**CGET**

- Emmanuel Dupont  
  Head of scientific coordination and innovation
- Catherine Prévost  
  Officer at the Resources and regeneration department
- Brigitte Baccaini  
  Deputy Director of observation and statistical analysis
- Damien Kacza  
  Officer for evaluation at the Office for evaluation and trials
- Sylviane Le Guyader  
  Research officer at the Office for Forecasts and Studies
- Charlotte Le Nué  
  Head of local monitoring
Foreword

Many governments – and people – around the world are wrestling with similar problems. How to improve education; how to boost growth; how to cut crime; how to make people healthier and happier. As this report shows, they are also facing the same challenges in understanding “What Works” and ensuring this informs public policy decisions, and practice.

This report seeks to draw out lessons for France from the experience of the UK What Works Centres. It provides an interesting “outsider” perspective on the What Works Centres, most of which were less than three years old at the time the report was written. What emerges is a network of centres with a common philosophy and aim: improving the effectiveness of public services. These centres have already done a lot to strengthen evidence-informed decision-making, whether that be through evidence synthesis, evaluation or knowledge mobilisation. Thanks to their methodological rigour, most have become a respected authority on what works in their sector. However, their approach threw up a number of challenges, from sourcing high quality evidence to convincing decision-makers to draw on this evidence, and supporting the transfer or scale-up of proven interventions.

This report shows that the UK also has a lot to learn from French innovation in evidence-based policy making. From the youth experimentation fund to the implementation of proven approaches like Housing First, France has developed considerable expertise in recent years. The use of consensus conferences is also a unique feature of French practice. In fields as diverse as reoffending and education, these conferences have often had a direct impact on key policy decisions.

The report is a timely reminder that the what works movement is international. No one country or organisation own the idea. With this in mind, more direct Anglo-French collaboration and co-funding of research would be a step in the right direction. Indeed, whilst communication of what works has to suit the local policy context of the country, much of the underlying evidence base is international. That’s why we are keen to explore
opportunities for French and British funders to pool resources for instance to fund systematic reviews. We hope this report can provide the spark for this sort of cross-channel collaboration.

Jonathan Breckon, Director, Alliance for Useful Evidence
Michèle Pasteur, Chief Executive, Agence nouvelle des solidarités actives

Executive Summary

The search for effective solutions to ingrained social, health and economic challenges has driven a growing interest in evidence-based policy making in recent years. These challenges range from early years inequalities to school dropout, from unemployment to economic disparities between regions, from drug taking to the loss of autonomy among the elderly. Over the past decade, this interest has contributed to the rising popularity of social policy experimentation, and the use of robust evaluation methods like RCTs, in France. These methods have added a new dimension to the French tradition of public policy evaluation: the ability to establish a causal link between policy measures and changes in service users’ lives. Many in the policy community were hopeful that these approaches would enable them to allocate public spending on the most effective evidence-based interventions.

Indeed, since 2007 a lot has been achieved in France. Several large-scale trials have been implemented, including a quasi-experimental trial of the French minimum income scheme, the Revenu de solidarité active or “RSA”. A 200 million euro social experimentation youth fund was also set up – le Fonds d’expérimentation pour la jeunesse – that has funded hundreds of pilot projects and around 30 Randomized Controlled Trials (RCTs).

However, this optimism gave way to criticisms and questions regarding the feasibility of such trials, and their use in policy decision-making. In spite of the apparent popularity of social policy experimentation, only a few dozen have actually been conducted in the French context. When evaluations do exist, few decision-makers and professionals use them to inform their practice.
Those who would like to apply the lessons from these evaluations struggle to do so because their findings are rarely disseminated in an accessible format. As a result, the evidence produced rarely informs public policy decisions and frontline services.

These difficulties led Ansa and our partners to look for promising practices in the field of knowledge transfer. In 2016, we decided to conduct a study of the What Works Centres aimed at drawing lessons for the French context. Six partners co-funded and supported this study: the disability benefits agency (CNSA), the family benefits agency (CNAF), three agencies attached to the prime minister’s office – a foresight and analysis agency (France Stratégie), a modernisation agency (SGMAP), and a regional inequalities agency (CGET), and Public Health France. All of these agencies hope to draw on the experience of the What Works Centres to launch initiatives in their policy area.

Starting from the assumption that it takes time and resources to place “evidence” at the heart of policy and practice, the British government has supported the emergence of independent knowledge transfer bodies in a range of fields: the What Works Centres.

The primary objective of the nine What Works Centres is to support frontline professionals and decision-makers to apply the evidence of “what works” in fields ranging from educational inequality to local economic development. They focus particularly on the evidence of the effects of specific interventions at a local level, rather than large-scale policy levers. Thanks to their evidence syntheses, these centres are now identified by a growing number of stakeholders as the “one stop shop” to the evidence of what works in their sector. Most sought initially to focus primarily on the findings of experimental and quasi-experimental evaluations in their systematic evidence reviews. However, centres faced both a limited supply of robust RCTs and quasi-experimental research and an inability to assess the replicability in a British context of interventions developed abroad. These difficulties have led several centres to broaden the definition of evidence to include other types of evaluation or research. Conscience of
the need to strengthen the British evidence based, several have also sought to stimulate the production of new experimental research to test promising innovations in the UK.

**The What Works Centres disseminate the findings of their evidence synthesis and support professionals to adopt effective approaches.** They have made evidence about the relative impact and costs of different interventions available through widely accessible outputs, including usually an online interventions library or clearinghouse (sometimes called a toolkit). By providing a rapid response to the question “what works in my sector?” the centres have quickly established themselves as a useful resource for decision-makers and professionals. For instance, certain head teachers have compared the Education Endowment Foundation’s Teaching and Learning Toolkit to the “Which?” consumer magazine, because it is independent, authoritative and clear.

**The centres are conscious that making evidence of what works available is necessary but not sufficient to change practice.** They are therefore committed to providing support to put evidence into action. The majority of the centres target local commissioners and frontline practitioners, with tailored guidance and toolkits, and – in some instances – support to replicate effective approaches. The majority of the centres have only recently begun this sort of intensive work on adoption, so we know little about their impact on practice. Aware of the need to innovate in this field, the What Works Centres continue to test and improve their dissemination and adoption methods. Several also recognise that to deliver on their mission would require a cultural shift that requires long-term change management in their sector.

**All of the centres promote the use of trials, but only two actually support or fund such evaluations.** The What Works Centre for Crime Reduction has limited resources for this work, and has opted to directly support the set up and evaluation of trials in collaboration with police forces. The Education Endowment Foundation, on the other hand, has launched 127 evaluations, including 105 using experimental methods, and involving 7 000 schools
British What Works Centres: what lessons for evidence-based policy in France?

– the equivalent of one in four schools in England. By conducting such a large-scale programme of trials, EEF has been able to evaluate the effectiveness of promising innovations, in so doing strengthening the evidence base for what works to reduce the attainment gap between students from rich and poor families in the UK. These trials have also helped build strong links between EEF and local schools and educational charities tackling these issues on the frontline. In our view, much can be learnt from EEF’s experience by those seeking to implement similar trials in France.

Ansa and our partners think that the what works model is a promising approach for strengthening evidence based policy making. Some have suggested that differences between the UK and France mean the model could not be replicated in France. The countries do differ, both in the way the public sector is structured (France is, broadly speaking, more decentralised), and in the role “evidence” plays in the policy process. However, our initial analysis suggests these differences should not prevent the emergence of what works style centres in France. Indeed, the What Works Centres have faced many of the same questions and criticisms in the UK than they would almost certainly be subject to if introduced in France. The centres have also established themselves in fields where the public sector is structured very differently, both highly centralised and relatively decentralised.

We therefore intend to explore the feasibility of launching what works-inspired initiatives or organisations in France, either by creating them from scratch, or by labelling existing organisations. These centres could take the form of new organisations, teams within government agencies, or a hub-and-spoke network model. These centres can only succeed if they are set up in well-defined fields, have a clear mission, have sufficient resources, and are committed to changing practice over the long-term. Their methodological rigor, independence and transparency are also key success factors. Lastly, they must place the needs and expectations of decision-makers and professionals at the heart of their model if they want to have an impact on policy and practice.
Whatever the policy field, decision-makers and practitioners face the same challenges in developing and implementing policy that is based on evidence. There is currently no forum dedicated to discussing these specific challenges, and to promoting effective practice. **Ansac and our partners therefore propose to create a cross-cutting network bringing together policymakers, researchers and frontline professionals to discuss how to reinforce evidence-based approaches in France**, inspired by the Alliance for Useful Evidence.
British What Works Centres: what lessons for evidence-based policy in France?

The report
Introduction

The issues: responding to a demand for effectiveness
Public authorities are faced with multiple social issues whether it is high unemployment rates (which has hovered around 9 and 10% since the 2008 crisis), the education attainment gap, social exclusion, an ageing population or addictions. Each year public spending is used to offer responses to these issues or improve existing ones. Stakeholders on the frontline also innovate to develop new interventions.

However, very often these social, medical and economic challenges persist, indeed spread or get worse. Why? Do we know enough about the needs of target groups? Do current policy measures have positive, neutral or negative outcomes? Are they effective but implemented on too small a scale? If so, how can they be expanded or rolled out? A growing number of stakeholders worldwide are interested in the impact of their policies, often in a context of budgetary restrictions.

In response to these questions, the use of randomised control trials (RCTs)1 and quasi-experimental studies has taken off in the last ten years in France. This trend has raised hopes that these robust methods would ensure public resources were allocated to programmes with proven impact. This hope gave way to questions about the reasons behind the inadequate dissemination of the findings and their use in designing public policy programmes.

As highlighted in particular by the recent work of the Agence nouvelle des solidarités actives (Ansa) and France Stratégie1, this is linked to the difficulties encountered by stakeholders in:

- producing evaluations that can be genuinely used to inform decision making by public stakeholders;
- producing a synthesis of this evidence in a given field to draw out

---

1 “British What Works Centres: a response”, page 43.
key policy lessons, and identify measures that can be integrated into mainstream public services;
- circulating this evidence on projects or practices that “work” among frontline stakeholders as well as policy-makers in order to encourage their adoption.

More broadly, a growing number of policy makers and researchers are showing interest in the concept of “evidence-based policy” (“politique fondée sur la preuve” in French). The promising approaches and challenges in implementing evidence-based policies have been shared within networks at both the national and international level (J-Pal, Alliance for Useful Evidence, Evidence Works, etc.) for a number of years.

Robust, well conducted evaluations are needed to answer questions such as “What is the impact of a programme, intervention or treatment on beneficiaries?”, “on non-beneficiaries?”, “Can these findings be replicated in a larger population?”

France already has a strong culture of public policy evaluation. It is based mainly on a tradition of the evaluation of public action by state inspectorates and audit agencies: these use both quantitative and qualitative methods to analyse the mechanisms and outputs of public policies. In addition to a degree of independence, their strength lies in their excellent knowledge of the institutional context and of the stakeholders implementing these policies. The external evaluation of public policies has also developed outside these agencies, led in particular by local government, specialist consultancies (such as members of the Société française d’évaluation), sometimes in partnership with researchers.

One key challenge is establishing a causal relationship between a policy and improvements in the lives of beneficiaries. Such internal validity can only be achieved with certain methods and data. In the academic literature, only certain evaluation methods (RCTs for example) and certain implementation conditions (trials in particular) can establish the causal link between a policy and a measurable impact (return to work following
employment support for example). In addition, the results achieved for a particular group cannot necessarily be reproduced with another group, made up of different individuals (the external validity of results). Importantly, establishing the impact of an intervention does not necessarily offer insight into the mechanisms driving that intervention. In order to translate the results of an evaluation into operational policy findings, it is often beneficial to supplement these quantitative methods with qualitative ones analysing the intervention’s implementation.

The evidence to be taken into account for evidence-based policies is wide-ranging and includes the results of experimental and quasi-experimental studies. It extends to other types of research – ranging from needs analysis, the evaluation of the theoretical basis of programmes and policies, qualitative evaluations (based on observations, interviews and case studies), process evaluations, and quantitative evaluations. These other resources, of a more theoretical, descriptive or qualitative nature, are essential particularly in terms of producing diagnostics (analysis and understanding of needs) and recommendations.

This report focuses on the means of strengthening the relationship between evidence and public policy. It asks what lessons and inspiration can be drawn from the experience of the British What Works Centres, the main objective of which is to contribute to the adoption by practitioners and policy-makers of the findings of high-quality evidence.

What Works Centres (WWC): a brief overview
The centres were inspired by the National Institute for Health and Clinical Excellence (NICE), established in 1999. Thanks to its robust methods and its impact on medical practice, NICE has become a pioneer in evidence-based health policy. From 2011, eight new centres inspired by the NICE model were set up in the United Kingdom. Supported by David Cameron’s coalition government, these new centres were set up as long-term ventures.

2 See page 25 ‘evaluation methods’.
The WWCs cover a range of areas:
- reducing educational inequalities; (The Education Endowment Foundation)
- early intervention (The Early Intervention Foundation);
- local economic growth (The What Works Centre for Local Economic Growth);
- crime prevention (The What Works Centre for Crime Reduction);
- promoting wellbeing (The What Works Centre for Wellbeing);
- improving the quality of life of older people (The Centre for Ageing Better).

Two of the centres are also focussed on delivering this approach in two devolved nations of the United Kingdom: Scotland (What Works Scotland) and Wales (Public Policy Institute for Wales).

The approach of the study

The approach was based on the following questions: how do these centres operate? What are their strengths and limitations? What practices might inspire French stakeholders in strengthening the link between evidence and public policy? In which areas of intervention and under which conditions?

In order to respond to these questions, Ansa used a two-part methodology:
1– A study to understand how What Works Centres operate, their key activities, the impact they have had and any challenges identified.
2– A comparison of the lessons from WWCs and the key practices in terms of evaluation and evidence-based policy making in France.

Through monographies and other analysis, the aim was to shed light on WWCs, to produce a series of findings of relevance to the French context. We decided to conduct a short study (May to December 2016) in order to quickly share the initial findings with partners and stakeholders. It is based mainly on interviews with the British centres, some of their stakeholders and French stakeholders. These interviews were conducted with a limited number of people who were generally committed to evidence-based policy-making. It would be worth carrying out more interviews with a wider
range of stakeholders to provide more in-depth analysis of the What Works Centres and evidence-based policy in France. This qualitative research was supplemented by a literature review.

Given the chosen method, this study offers avenues for enquiry but on several points does not allow definitive conclusions to be drawn. It points to issues that would be worthy of further exploration – for example by asking frontline professionals and policymakers about their use of the tools provided.

In-depth interviews were conducted with the heads of five out of the nine centres and in some cases with members of their team. A literature review setting out their approach and work was carried out. The case studies for these five centres are appended to this report.

They are:
- the Early Intervention Foundation (see case study, page 109);
- the Education Endowment Foundation (see case study page 120);
- the What Works Centre for Local Economic Growth (see case study page 131);
- the Centre for Ageing Better (see case study page 143);
- National Institute for Health and Care Excellence (see case study page 156). This case study also refers to another public body, Public Health England. These two bodies were studied in view of their work evaluating, synthesising and disseminating evidence in the field of public health.

Interviews were conducted with the heads of the other WWCs to fill in the gaps in terms of understanding how they operate and the issues faced.

In order to understand the context and levers for setting up the WWCs, meetings were held with representatives of the Cabinet Office which oversees the network of What Works Centres, Nesta (an innovation foundation), and the Alliance for Useful Evidence, which promotes evidence-based policies in Great Britain and beyond. The team also
British What Works Centres: what lessons for evidence-based policy in France?


Over thirty French stakeholders involved in evaluation, synthesising evidence and promoting evidence-based public policy were interviewed. A workshop of experts and four thematic workshops were held with study partners on early childhood (Cnaf), ageing (CNSA), regional equality (CGET) and public health (Santé publique France). Interviews and workshops compared work by the WWCs and the key French stakeholders, as well as the main issues in the French context (the list of interviewees can be found on pages 5 to 7). In addition, a brief review of the British and French literature was conducted.

The findings are presented in this report in five main parts:

- Part 1 defines the various components of evidence-based policies and presents the initiatives launched in various countries in order to put this concept into practice (page 23);
- Part 2 describes how British What Works Centres operate and analyses their relevance and results (page 43);
- Part 3 offers an overview of French practice in terms of evaluation and evidence-based policies (page 72);
- Part 4 sets out how lessons from the experience of the What Works Centres can be applied in France (page 95);
- Part 5 sets out proposals for the French context (page 102).

---

An international issue: making public policy more “evidence-based”

“Evidenced-based” public policy

Having been established in the 1990s in the field of medicine, the “evidence-based policy” approach, spread to other areas of policy-making (tackling poverty, education, justice, environment etc.)\(^1\). The aim of this approach is to allow the sound use of available evidence by policymakers at all levels and by practitioners in their decision-making processes and practices.

Different types of evidence can be taken into account in this approach, such as scientific evidence, evidence relating to the management of the intervention (arising from the implementation) and detailed evidence relating to the situations and needs of individuals. This approach also presents issues around the accessibility of validated evidence, in particular in appropriate synthetic formats. Finally, it raises the question of integrating and ranking different types of evidence (scientific, empirical etc.).

The conference Evidence Works 2016 – a Global Forum for Government, organised by Nesta, the Alliance for Useful Evidence and Results for America in September 2016 was attended by 140 participants (ministers, senior officials, evaluators) from over forty countries. These representatives came together to discuss the best ways of implementing this evidence-based approach to policy, and its success provides a testament to the widespread interest in these approaches in public policy circles.

Discussions nevertheless emphasised the difficulties encountered in promoting and genuinely implementing these evidence-based policies: What is the best way of identifying projects with the biggest impact? Which policies offer a cost-effective response to such an issue? In most
areas of public policy, decision-makers and practitioners do not have access to clear, useful answers to these questions. Even when solutions are known, how can decision-makers and practitioners be encouraged to adopt them?

The answers to these questions have multiplied in the last twenty years, both nationally and internationally. Experimental and quasi-experimental studies and systematic reviews are part of the answer and offer a response to the need to identify effective interventions and support decision-makers and practitioners to adopt them.

**Producing a good standard of evidence**

Research into effective public policies has led to an increasing reliance on evaluation methods of various forms, ranging from qualitative or process evaluations to policy evaluation and RCTs. These methods are used for anything from a public policy area to individual policies or projects, or indeed specific practices and interventions.

As such, the term “evaluation” covers a very wide range of approaches and objectives. These various methods provide a range of insights, and are applicable to different contexts.

Below is a brief recap of the most common practices, which are also the subject of numerous articles.
EVALUATION METHODS

Needs analysis and theory-based evaluation can help, on the one hand, to define the needs to be addressed, and, on the other, to analyse promising interventions, setting out their scientific basis and expected outcomes.

Process evaluation is intended to analyse the implementation of a programme and often involves ensuring that the activities were indeed carried out, to a high standard and within the expected time-frame.

Experimental and quasi-experimental studies are intended to identify whether a programme or public policy has had measurable outcomes and to quantify these. This process is set out in more detail in the following paragraph.

Qualitative evaluations are in-depth investigations into a small sample of individuals, based on observations, interviews and case studies in particular. This type of evaluation can contribute, for example, to understanding what does and does not work in an intervention.

Quantitative evaluations are based on representative samples of beneficiaries of an intervention.

Economic evaluations fall into one of two categories. On the one hand these are economic models produced using empirical data. This type of analysis has been used in the past to assess major public investment such as in transport. On the other hand, the socio-economic evaluation of programmes involves comparing the cost of a programme with the results obtained (cost-effectiveness analysis) in certain cases assigning a monetary value to the outcome (cost-benefit analysis). These methods are also a means of comparing programmes and choosing the most efficient.
Experimental and quasi-experimental studies: towards higher standards of evidence

An evidence-based approach to policy means drawing on the best available evidence. When trying to establish the impact of a policy on beneficiaries, this means identifying the most robust way to demonstrate the effects of an action.

The purpose of experimental and quasi-experimental studies is to establish a causal link between an intervention (for example, financial or personal support) and impacts on individuals (for example, access or return to employment) in order to demonstrate whether a programme has produced the desired results directly. Such approaches were developed and refined in medical research before being transferred and adapted to the social sector.

The academic literature defines specific implementation conditions and/or econometric methods required to confirm the causal link between a programme and its impact. In particular, it is about setting objectives to work towards and developing indicators for those objectives (which can thus reduce the measured effects to a few indicators). The benefit of an intervention is thus calculated by the size of the impact measured and the degree of confidence in that result will depend on the quality of the evidence (particularly the presence of a valid counterfactual scenario, the selection and size of samples, controlling selection bias etc.). Ansa will refer to these as “experimental and quasi-experimental studies”.

The results of these studies are one form of ‘evidence’ to be taken into account as part of a broader body of knowledge which, taken together, can contribute to a “culture of evidence” and “evidence-based policies”. **This method has the benefit of offering a robust response** from an academic perspective to questions around the effectiveness of innovations or interventions. However, it presents biases and limitations to be taken into account at the evaluation design stage and when using the results. Moreover, it should be emphasised that it is not possible to apply it to all
types of intervention or policy. In particular, where a programme seeks to intervene in phenomena with very complex causes, or that have results that are too difficult to measure, such approaches may be ineffective and other methods must therefore be adopted.

In order to use evaluation in a transparent and robust way, scientific bodies have established classifications for the various evaluation methods based on their scientific standard. These scales or standards of evidence can range from the quality of the design of an intervention in terms of the prior scientific evidence, to before/after studies, to experimental and quasi-experimental methods. Two examples of such classification are the Maryland Scientific Scale and the Nesta scale. Depending on the theoretical approach, discipline or subject, the ease with which a high standard of evidence can be established will vary, as will these scales. It is, therefore, important to balance a push for high standard of evidence whilst ensuring the methods are appropriate to the intervention.

EXPERIMENTAL AND QUASI-EXPERIMENTAL METHODS

These methods seek to establish the average difference between a situation with the intervention and that which may have occurred in its absence, by comparing a sample of individuals benefiting from the intervention with individuals who don’t.

This control group aims to create a “counterfactual”, demonstrating what would have happened to the test group had they not received the intervention. As such, the “control” group that makes up the sample must be as similar as possible to the “treated” group. If the groups differ, it is easy to confuse the effects of the intervention with the differences (both observable and unobservable) between the two groups. Creating this control group is a key challenge in such methods, and the ability to identify a causal effect depends to a large extent on how this is done.

The research shows two sets of approaches to for identifying a valid counterfactual scenario.
1– Randomised control trials (RCTs). This method involves selecting people at random to benefit from the intervention (test group) and monitoring the people who are not selected (control group). Through random assignment, test and control groups can be created that are on average identical and as such the effects of the intervention can be isolated more easily. From this perspective, it is often seen as the highest standard for measuring impact.

2– Quasi-experimental design: these are methods that either mimic experimental situations by using statistical methods (such as statistical matching, double or triple difference) or identify situations that produce variations creating unequal access to treatment, in order to identify its impact (instrumental variable, regression discontinuity or other natural experiments). Within this category, each of the various methods includes a bias which must be taken into account when analysing results. Other evaluation and research methods are often to complement an impact evaluation. On the one hand, this includes research that enables an in-depth understanding of the needs or issue (ex. participatory observation) and, on the other, evaluation of processes so as to understand the conditions in which the intervention under evaluation was implemented.

In a French context, “social experimentation” refers to an innovation in social policy pilot tested on a small scale, in conditions that allow its effects to be evaluated with a view to rolling it out. It is a means of promoting innovation and is based on the principle of counterfactual impact evaluation. Many public bodies in France and in Organisation for Economic Cooperation and Development (OECD) countries, now trial certain programmes in order to confirm their feasibility, relevance and above all their impact. However, increasingly in France, the term social experimentation refers to pilot projects or innovations that do not benefit from a robust evaluation.

---

3 Here, the term ‘pilot project’ or ‘pilot’ refers to policies that are implemented on a small scale, in order to test their feasibility.
Internal validity: is the study conclusive?
RCTs and quasi-experimental studies contain inherent bias. Such bias must be addressed or the evaluation results risk being inconclusive. Denis Fougère\textsuperscript{xvi} describes the bias inherent in controlled experiments and describes the responses that may be provided in terms of experiment protocols and designs. He highlights the fact that practitioners of these methods applied to social sciences are all too often unaware of experiment protocols and designs and that “the observation was often made that findings deduced from social experiments often had little statistical significance and/or were not very robust”. As such, the response must involve addressing these aspects through the use of qualified evaluators.

External validity and issues around roll-out
Once the effects of an intervention on beneficiaries have been shown through an RCT or quasi-experimental study (internal validity), other questions arise to determine whether it can be disseminated or even rolled out:
– is the intervention that is under evaluation transferable to other contexts, groups of people or circumstances? For example, the experiment might have been conducted in a period of economic stability, whilst its roll out was prepared in a period of economic crisis (external validity)?
– could scaling up interfere with the overall cost and have indirect effects that were not measured in the trial (“equilibrium effects”)? For example, will it have unintended effects on the overall situation, such as an increase in the cost of consumer goods in the case of financial assistance being rolled out? Will rolling out the programme over the long period change the behaviour of beneficiaries?
– what can be done to ensure the conditions of the intervention under evaluation are respected (“fidelity”)?

The process of policy-making must therefore incorporate these questions if it is to take the roll-out decision on the basis of these results.

In the field of social employment policies, one can cite a number of trials that have led to successful dissemination on an international scale, including Individual Placement and Support (IPS).
INDIVIDUAL PLACEMENT AND SUPPORT (IPS)

Individual Placement and Support offers people with mental health problems unlimited intensive individual support in gaining and keeping employment in the mainstream job market. This method has been the subject of numerous controlled trials in the United States and Europe. These studies reported a positive impact with 42% to 61% of people accessing mainstream competitive jobs\textsuperscript{xviii}. The European trials showed that the beneficiaries of IPS were twice as likely to access employment than when using the classic “trained, then placed” approach\textsuperscript{xx}. IPS has been trialled successfully in the United Kingdom, Norway, Denmark, China (Hong Kong region), Canada, New Zealand and Australia. Since 2012, an independent think tank – the Centre for Mental Health – has helped to set up 14 IPS teams in England, ensuring fidelity with the initial model. In parallel, IPS was rolled out across England. In 2012, half of all secondary mental health services said that they had at least one IPS advisor\textsuperscript{xx}. The IPS method has also been disseminated in France, on a more modest scale. Several French charitable organisations adopted this method in projects under various names (emploi accompagné, Working First and Job Coach).

Economic evaluations: measuring the efficiency of interventions

Economic evaluations, in this context referring to cost-effectiveness and cost-benefit analysis, are a means of studying the cost of an intervention in relation to its effects. They are often based on the results of RCTs or quasi-experimental studies, which are used to compare the situations and trajectories of test and control groups. The costs and effects to be taken into account depend on the objectives of the evaluation and are based on hypotheses (whether or not to take into account fixed costs, cost opportunities for the beneficiary etc.)\textsuperscript{xxi}.

By adopting the same conventions in terms of value, forecast or effects included and excluded, they allow programmes with the same objectives to be compared. This type of analysis aids decision making in allocating resources between different programmes.


METHODS

Cost-effectiveness analysis: ratio between the cost of the programme and its effectiveness (for example the number of people accessing employment). Interventions will have effects in a range of areas and it is thus possible to conduct analysis of cost-effectiveness of several desired outcomes. For example, NICE in the United Kingdom uses the QALY (Quality Adjusted Life Year). One QALY is equivalent to one year of life in good health. As such, NICE calculates and compares the cost of programmes in pounds sterling per expected QALY.

Cost-benefit analysis: this analysis involves assigning a monetary value to the various benefits obtained and comparing them with the cost of the programme. In the case of social outcomes in particular (for example improving well-being, the avoided costs of disengagement from education or crimes not committed), giving them a financial value is particularly tricky and can lead to substantial variations depending on the selected results. Indeed, this means using hypotheses and placing a value on effects for which there is not currently a consensus in terms of social policy.

Systematic literature reviews

When attempting to resolve an issue, many interventions may exist with different effects and costs. Standard practice in synthesising research involves analysing the most striking developments in relation to a specific subject or controversy, based on known references. This type of synthesis, often called a literature review, is suited to many situations and can reveal interesting interventions.

In the context of evidence-based public policies – and the interventions covered by this report – one specific approach is predominant: systematic reviews. This method can be used to compare the impacts of interventions and thus distinguish between those that work or show promise and those that it would be better to rule out, in response to a specific question raised by practitioners. These reviews can also include an analysis of
the cost and benefit of the interventions. They can also be used to identify the interventions and sectors where there has been little in the way of evaluation and which require additional research. However, they are more complex and costly in terms of resources and time than a typical literature review.

Systematic reviews are based on specific methodological standards – the Cochrane definition for the medical sector, is given here:

**Systematic reviews** (according to Cochrane, see reminder below) are a robust scientific process in several stages:

1– Researching existing studies, clinical trials or other relevant studies, whether published or not, that respond to a given clinical question (for example, all the clinical trials carried out in order to evaluate the efficacy of a treatment).

2– Assessing the quality of each study and selecting the ones that meet a high predefined quality standard.

3– Synthesising the findings from the selected studies.

4– Finally, should the data allow, carrying out statistical analysis, known as meta-analysis. This means that quantified findings from various studies can be combined in one summary finding.

When they are based on the best studies, these reviews present a synthetic overview of evidence on a given issue and its reliability and are a means of analysing the relative effectiveness of all solutions.

**COCHRANE AND CAMPBELL COLLABORATION: MAJOR PLAYERS IN SYSTEMATIC REVIEWS**

Cochrane (previously known as Cochrane Collaboration) was established in 1993 to ensure that randomised studies were developed and used more widely in the medical world, so as to ensure quality decision-making on health and medical care. This international network of researchers and practitioners supplies the medical world with systematic reviews of all random controlled trials and the best standards of evidence available in the healthcare sector. Many
members of Cochrane identified the need for systematic reviews to evaluate interventions in the social sector in the same way: Campbell Collaboration was thus established in 2000. Campbell Collaboration publishes systematic literature reviews in education, crime and justice, international development and social welfare. The network produced 112 systematic literature reviews between 2004 and 2015xxv.

When the State promotes RCTs and quasi-experimental studies

Barriers to implementing these methods
The implementation of trials, whether they be RCTs or quasi-experimental studies, throw up real challenges for governments and political leaders. The discrepancy between the long-term nature of trials and the short-term nature of political mandates as well as the potentially high costs of these methods are key barriers. Trials are also based on the assumption that the effects of the measures proposed are uncertain, which runs contrary to the classic posture of policymakers.

“Trials are based on the idea that we don’t know, in advance, whether a measure will work or not. That idea is difficult to accept for politicians used to selling guaranteed results to the electorate.”

David Halpern, Director of the Behavioural Insights Team.

In order to accept trials – and to take part – frontline professionals must also accept the idea that their interventions can be evaluated in order to be improved or enriched.

Evaluation is thus sometimes seen as a sanction, a score, as calling into question their project or even their skills.

Political will
The development of trials therefore depends on effective incentives to overcome these barriers and is dependent on the political will, amongst public and private actors, to fund and mainstream these evaluations.
Incentives to carry out trials and evaluate what we have observed can take several forms:

- **funds for evaluation and experimentation.** Set up in 2009, the Fond d’expérimentation pour la jeunesse in France received an initial endowment of €230 million from the State and private sponsors for the period 2009–2012. The Education Endowment Foundation, one of the British What Works Centres, received £125 million to be spent over 15 years (see case study appended on page 120). Although they have different approaches, these organisations support and fund trials, with programmes led by frontline stakeholders and evaluations carried out by independent evaluators;

- **mainstreaming trials within government.** For example, the Behavioural Insights Team, which started life within Cabinet Office, has implemented many trials to test new ways to deliver public services, based on behavioural science or “nudges”;

- **support for trials or enshrining them in law.** The United Kingdom set up the Cross-Government Trial Advice Panel to support civil servants in implementing trials.

**BEHAVIOURAL INSIGHTS TEAM, OR “NUDGE UNIT”**

The Nudge Unit was set up in 2010 within the British Prime Minister’s Office to conduct randomised control trials of small incentives or changes based on research in behavioural psychology for the purposes of improving public services for users. The centre tested numerous innovations to make public services easier to use and enabling citizens to make “better” choices for themselves. In its first two years of existence, it has led dozens of trials in policy areas ranging from health to taxation, saving energy, crime reduction, employment and even economic growth. These trials have shown that a number of small actions (in particular in the interactions between public services and users) had significant effects and in some cases led to substantial savings with an equivalent (or even better) outcomes.

For example, it emerged that adding a comment about the fact that most people pay their taxes on time to a letter sent to taxpayers would
appear to further increase the proportion of taxes paid on time. It was also shown that asking jobseekers what they are going to do to find a job in the next two weeks, rather than what they had done over the last two weeks, at their regular interview, significantly improved their chances of returning to work.

For a successful use of evidence by professionals and policymakers at all levels

In order to take the appropriate decisions and improve their practice, policymakers and professionals need to have an understanding of the measures that stand a good chance of working in their policy area as well as the conditions for success.

The challenge goes beyond evaluating the impact of programmes and relates to evidence as a whole—particularly scientific evidence. Evidence-based policy approaches seek to meet this challenge, by drawing on the best available evidence in policy-making processes.

Issues around “knowledge mobilisation”

The problem of the adoption of scientific evidence by practitioners is well known and has been the subject of much study in the medical field. The adoption of the most effective medical interventions is often a very long process (an estimated 17 years according to Balas and Boren\textsuperscript{xxviii}). This issue of knowledge mobilisation, meaning the “dissemination, adoption and appropriation of evidence for use in practice and decision-making”\textsuperscript{xxix}, is particularly relevant in the social sector.

Faced with evolving growing and constantly evolving research base, public health practitioners face certain barriers that can be found in other areas:

- \textbf{lack of material in an appropriate format:} the results of studies and evidence is rarely presented in a way that is synthetic and accessible...
British What Works Centres: what lessons for evidence-based policy in France?

to laymen - professionals and decision-makers rarely find answers to their questions in a convenient format;
- **lack of continuous learning:** individuals within organisations operate with their areas of expertise that are both explicit but also implicit, acquired from experience, and with convictions. They are also part of an organisational culture. These factors do not immediately incentivise them to call into question the routines and hypotheses they take for granted;
- **lack of confidence of stakeholders,** professionals and policymakers in the findings of a given piece of research, particularly because the organisations to which they belong rarely have a culture of research;
- **resistance to change** among people and within organisations, difficulty in establishing buy-in among stakeholders (policymakers, professionals, beneficiaries, citizens) to a new solution or around the removal of an initiative that has not had a positive impact.

Policymakers, middle managers and frontline professionals are all stakeholders in complex decision-making, embedded in institutional and organisations practices, dependant on governance mechanisms and the capacity of organisations to change. These stakeholders must also take the broader context and political considerations into account. The aim of the processes of knowledge mobilisation is to help to incorporate robust evidence into the policy process.

As such the tools for tackling these challenges are multiple and complementary:
- sharing research findings (RCTs, quasi-experimental studies, research synthesis) in forms that professionals and decision-makers can use;
- cultural and organisational changes among decision-makers and professionals in order to create an environment that is conducive to evidence-based policies.

In line with the values of evidence-based effectiveness, British researchers have conducted a recent review of approaches to promoting the use of existing evidence and evaluations: *The Science of Using Science*. This review of the literature sheds light on the findings of current research.
and once again points to a lack of evidence in this area. As such the initiatives presented in the following sections are interesting leads rather than interventions that have already been proven to be effective.

**Sharing the results of trials and systematic literature reviews in the form of guidelines and practices**

Policymakers and professionals are in need of the information that will enable them to resolve general problems (how to reduce the attainment gap) or specific ones (is reducing class sizes “effective”?). For this, the results shared must give them:

- an overview of the solutions that might be applied, with recommended solutions which appear the most effective or promising;
- available interventions, with a focus on “guidelines for use” (for which group, in which context) and detailed implementation processes.

**Collaboration between researchers, research producers**

and professionals responsible for implementing services, emerges as a factor of success in this stage, in terms of ensuring that the evidence is presented in an accessible and user-friendly way and that it is useful. This collaboration can also continue thereafter in order to ensure that the research produced is more in line with operational needs.

Here we present two existing formats for sharing the results of evaluation and research synthesis in the form of presentations and recommendations: consensus conferences and clearing houses.
CONSENSUS CONFERENCE

The aim of a consensus conference is to identify key questions and determine areas of controversy to be overcome for a defined problem. By bringing together many different perspectives, drawing on international scientific knowledge and allowing structured debate, such a method can give rise to a core consensus which can form the basis of a shift in public policy or the assessment that there is not enough evidence to reach a decision. Results take the form of recommendations made by an independent panel following a presentation of the available evidence by experts.

The method of consensus conferences started life in the United States in the 1970s to determine the best action to take in a specific medical case. Initially made up only of experts, these conferences gradually opened up to include non-professionals and were used in social welfare and legal cases etc.

In France, the use of this method has grown over the last few years, with:
- the PIREF consensus conference on literacy (www.cndp.fr);
- the consensus conferences organised by the Haute Autorité de Santé in the medical field on 16 - 17 December 2005 on the treatment of psychopathy and on 25 - 26 January 2007 on criminal psychiatric evaluation;
- in the social field, the consensus conference on homeless people “Sortir de la rue” in 2007;
- in the justice field, the consensus conference on preventing reoffending on 14 -15 February 2013;
- in education, the consensus conferences organised by the Conseil national de l’évaluation du système scolaire (Cnesco): on repeating a school year in January 2015 (www.cnesco.fr section conférence de redoublement), on numeracy in November 2015, literacy in March 2016 and on streaming (the public sessions of which will take place in March 2017).
CLEARING HOUSES

These resources – often in the form of websites – list interventions whose effects have been proven through trials. They emerged in an American context characterised by a surge in the use of RCTs for social exclusion prevention programmes and a desire to make better use of restricted public budgets.

One of these is the *Blueprints for Healthy Youth Development*. The organisation Blueprints provides a database of programmes that promote the health and wellbeing of children and adolescents. Blueprints was one of the pioneers in researching and identifying evaluated programmes. Hosted by the University of Colorado in Boulder, it began in 1996 with work to prevent youth violence and went on to extend its remit. To date, over 1,400 programmes have been reviewed, with just 5% ultimately selected as model and promising programmes.

Other examples can be viewed below:
- the *What Works Clearinghouse – Institute of Education Sciences in the United States*;
- the *Washington State Institute for Public Policy*.

*Promoting a culture of evidence among policymakers and professionals*

In many countries, initiatives have been developed to encourage the adoption of evidence-based approaches by policymakers and professionals.

The initiatives presented here often target policymakers in different forms.
Creating organisations to strengthen the link between evidence and public policy

The paradigm shift in public policy-making and professional practice towards constantly incorporating findings of evaluative research is a long-term one: it is about transforming practices, skills, professional cultures, and the thinking of policymakers, frontline professionals and the general public. In order to mainstream evidence-based decision-making in the behaviour of policymakers, different approaches must be combined over many years, or even decades. As such, many international experts agree that there needs to be dedicated infrastructure to support actors from the design to implementation of evidence-based policies xxxv.
This infrastructure often takes the form of independent institutes:
– the Bureau for Economic Policy Analysis in the Netherlands;
– the Productivity Commission in Australia;
– What Works Centres in the United Kingdom.

They can also draw on dedicated teams within government such as the Behavioural Insights Team in the United Kingdom (now independent) and the Evidence Team in the United States.

These organisations play an essential role as brokers of knowledge, producing deliverables that must be relevant, robust and available ‘at the right time’. These documents take various forms, such as a “policy brief” (for example the Promising policies series by the Bureau for Economic Policy Analysis in the Netherlands) or the intervention database created by the What Works Centres in the United Kingdom.

The British WWCs are a specific example of this new infrastructure that could be a source of inspiration for French actors. The next section analyses their practices and the fourth draws lessons from their experience that could be applied in France.

ix Heim Arthur, Comment estimer le rendement de l’investissement social, pending publication.
x “[the principles] involve gauging the consequences of each project across the country, evaluating the productivity gains or the increase in production capacity of businesses as well as additional goods consumption and improvements in citizens’ living standards and by comparing these effects to the costs of investment.” in: Emile Quinet, L’évaluation socioéconomique des investissements publics, report published on 18 September 2013. Available online: www.strategie.gouv.fr/publications/levaluation-socioeconomique-investissements-publics-tome1.


A recent overview of these quasi experimental methods can be found in: Rozenn Desplat, Marc Ferracci, Comment évaluer l’impact des politiques publiques? Un guide à l’usage des décideurs et praticiens, France stratégie, September 2016.


Heim Arthur, “Comment estimer le rendement de l’investissement social”, pending publication.


Heim Arthur, “Comment estimer le rendement de l’investissement social”, pending publication.


Halpern David, Inside the Nudge Unit, WH Allen, 2015.


xxx Retakes, addressed at the consensus conference organised by CNESCO, represent a good example of the buy-in difficulties encountered.

Langer Laurent, Trippney Janice, Gough David, “The science of using science, Researching the use of evidence in decision-making”, April 2016, EPPI-Centre, University College London.


Puttick Ruth, “Why we need to create a ‘NICE for social policy’”, Nesta, May 2012.
British What Works Centres: a response?

British What Works Centres (WWC) were a response to a simple observation: a significant proportion of British public services are not, or only to a limited extent, “evidence-based”. In particular, these services are rarely designed and implemented using practices proven to be effective through RCTs or quasi-experimental studies. In education, analysis of 70 programmes funded by the Department for Education in England identified just two to three programmes that were subject to robust evaluation. Even in the medical field, the adoption of interventions is often a very long process (an estimated 17 years according to Balas and Boren). This situation is in part due to the lack of links between evaluators and policymakers, the difficulty in turning evaluation reports into operational lessons and the small part played by evaluations in political decisions.

“Despite the idea that policy and practice should be underpinned by rigorous evidence being widely accepted, billions of pounds are spent every year on social policies and programmes with little rigorous evidence on the impacts that these have.”

Why we need to create a “NICE for Social Policy”, NESTA, 2012.

Presentation of What Works Centres

The general principles

What Works Centres were designed as a response to this issue and to promote the adoption of interventions that “work”. They have three missions:

– producing a synthesis of existing research in their policy area through systematic literature reviews with a particular focus on RCTs and quasi-experimental studies. The centres turn this research into findings, or even practical guidelines, for professionals and decision-makers;
– disseminating and supporting adoption. These centres disseminate the
research, particularly through “toolkits” designed to help professionals identify the most effective interventions. They also support its adoption through events, training and dissemination programmes;

- **evaluating the impact of innovations.** Some centres also have dedicated funding to strengthen this evidence base through new trials.

Each centre focuses on a specific area of public policy, an approach with two benefits. On the one hand this allows them to exhaustively catalogue RCTs and quasi-experimental studies and as such act as a “one-stop shop” for them. On the other hand, centres can organise targeted dissemination that is tailored to professionals in that policy area. The areas are wide-ranging, from health to education, from economic development to well-being.

Each policy area presented specific challenges in terms of evaluation, research synthesis and dissemination. Each centre has had to adapt to the specificities of its policy area. Two of the centres are also focussed on delivering this approach in two of the devolved nations of the United Kingdom: Scotland (What Works Scotland) and Wales (Public Policy Institute for Wales).

**WHAT WORKS CENTRES: MAIN CHARACTERISTICS**

In terms of their remit they:

- **focus on research into effective solutions rather than needs analysis.** To do so the centres rank the programmes according to their effectiveness, taking into account their results, the robustness of evaluations carried out and their cost. Centres provide an indication of the costs of the various programmes in order to inform the choices of decision-makers in terms of their efficiency;

- **prioritise changing frontline practices** (bottom up approach), rather than national public policy agenda (top down approach). Consequently, they focus mainly on middle managers and professionals, although national decision-makers can sometimes be targeted;

- **supporting local professionals and managers** in adopting the
research findings. Taking inspiration from implementation science, the centres have tested a range of dissemination methods.

In terms of organisation and funding, the centres are:
– generally independent, and generally sit within a charitable organisation, foundation or university (with the exception of two public centres);
– very close to political authorities, as they benefit from public funding;
– sectoral, bringing together decision-makers, researchers and professionals in a given policy area;
– the recipients of long-term funding, over a period of 3 to 15 years;
– very unequal in terms of their overall budget (The National Institute for Clinical Excellence has 100 times the annual budget of the What Works Centre for Wellbeing).

Synthetic presentation

The What Works Centres emerged in three phases.

1– In 1999, the British Secretary of State for Health created the National Institute for Clinical Excellence (NICE). Its role was to evaluate drugs and medical devices, particularly in terms of cost-effectiveness. Its remit has gradually expanded to include clinical and public health guidance. Its aim is to feed these into the National Health Services (NHS) in order to quickly disseminate good practice. Indeed, thanks to its robust methods and influence on medical practice, NICE became a pioneer in evidence-based health policy.

2– The Education Endowment Foundation was established in 2011 with an endowment of £125 million inspired particularly by Race to the Top, an American programme.

3– Finally, from 2013, seven new What Works Centres opened their doors. Five are thematic and two are regional, covering the devolved nations of Scotland and Wales.
<table>
<thead>
<tr>
<th>Structure</th>
<th>Area</th>
<th>Launch date</th>
<th>Status</th>
<th>Annual budget (2015)</th>
<th>Length of funding</th>
<th>Workforce</th>
</tr>
</thead>
<tbody>
<tr>
<td>National Institute for Health and Care Excellence (NICE)</td>
<td>Health, public health</td>
<td>1999</td>
<td>Public</td>
<td>£62.5 million</td>
<td>Permanent</td>
<td>600 staff (approximately)</td>
</tr>
<tr>
<td>Education Endowment Foundation</td>
<td>The attainment gap</td>
<td>2011</td>
<td>Charity and Foundation</td>
<td>£14 million</td>
<td>2011-2026 [endowment £125m]</td>
<td>19 staff</td>
</tr>
<tr>
<td>Early Intervention Foundation</td>
<td>Children from conception to 18</td>
<td>2013</td>
<td>Charity and Foundation</td>
<td>Not provided</td>
<td>2013-2016</td>
<td>16 staff</td>
</tr>
<tr>
<td>What Works Centre for Local Economic Growth</td>
<td>Local economic growth</td>
<td>2013</td>
<td>Consortium</td>
<td>£1 million</td>
<td>2013-2016</td>
<td>10 FTE</td>
</tr>
<tr>
<td>What Works Centre for Crime Reduction</td>
<td>Crime prevention</td>
<td>2013</td>
<td>Public</td>
<td>Not provided</td>
<td>2013-2016</td>
<td>Not provided</td>
</tr>
<tr>
<td>Public Policy Institute for Wales</td>
<td>Wales</td>
<td>2014</td>
<td>Consortium of universities</td>
<td>£450 000</td>
<td>2014-2017</td>
<td>8 staff</td>
</tr>
<tr>
<td>What Works Centre for Scotland</td>
<td>Scotland</td>
<td>2014</td>
<td>Consortium of universities</td>
<td>£1 million</td>
<td>2014-2017</td>
<td>7 FTE</td>
</tr>
<tr>
<td>What Works Centre for Wellbeing</td>
<td>Wellbeing</td>
<td>2014</td>
<td>Social enterprise</td>
<td>£600 000</td>
<td>2014-2017</td>
<td>5 staff</td>
</tr>
<tr>
<td>What Works Centre for Ageing Better</td>
<td>Ageing well</td>
<td>2015</td>
<td>Foundation</td>
<td>On average £5 m/year on 10 years</td>
<td>2015-2025 [endowment: £50m]</td>
<td>20 staff</td>
</tr>
</tbody>
</table>
Background

The creation of these eight new centres from 2011 represents an unprecedented page in the history of public policy in the United Kingdom. This phenomenon can be explained by four factors in particular.

Firstly, evidence-based policy-making had gained in popularity in the United Kingdom, particularly under the Labour government elected in 1997. Indeed, the governments under Tony Blair were from a “reformist” left for which the effectiveness of public services was a key issue. From 1999, this government published a white paper emphasising that ‘evidence’ must be a central component of the most effective public policies. This vision was quickly put into practice through a range of initiatives.

On the one hand, numerous controlled trials were carried out, for example to test active labour market policies. On the other hand, new measures were launched to mainstream “evidence” in political decision-making and professional practice. The National Institute for Health and Clinical Excellence, established in 1999, was the flagship example. Between 2000 and 2004, another – less known – experiment was carried out in support of evidence-based policies: the UK Centre for Evidence Based Policy and Practice. A real forerunner of the What Works Centres, its funding was not renewed, seemingly in part due to being too academic in focus and as such not operational enough to aid decision-making.

Secondly, the idea for What Works Centres was the subject of sustained lobbying by leading figures, led by David Halpern (director of the Behavioural Insights Team) and Geoff Mulgan (director of the innovation agency, Nesta). Between 2011 and 2012 many references were made to a “NICE for Social Policy”, in blogs, reports and press articles. These stakeholders noted the absence of organisations responsible for connecting supply of ‘evidence’ with demand. To fill this gap, they turned to NICE and suggested replicating the NICE model for social policies, in areas such as education, justice and child protection. From 2012, the Alliance for Useful Evidence – a network of professionals with an interest in “evidence-based policies” – also made a contribution with publications and
British What Works Centres: what lessons for evidence-based policy in France?

Events on this topic. A long-standing player in the area of evidence-based public policy – the Dartington Social Research Unit – also made a tangible impact on the promotion of What Works Centres. Indeed, its contribution to a parliamentary report by Graham Allen played a part in the creation of the ‘Early Intervention Foundation’.

Thirdly, this lobbying was well received by certain senior officials, advisors and ministers in the new coalition government. This idea was incorporated in a public service reform strategy with a focus on openness and transparency. Thus, the idea of mirroring NICE in other fields was quickly raised in an Open public services white paper (2011). This concept of a NICE for social policy might have remained just an idea. However, it benefited from the support of a major figure, head of the civil service Jeremy Hayward, who, from 2012 raised the idea of a NICE for social policy. To offer the impetus for the creation of new centres, a team dedicated to the What Works Centres was set up within the Cabinet Office.

Finally, the Cabinet Office, with the support of the Economic and Social Research Council (ESRC), enlisted the support of funders for the creation of new centres. They carried out a consultation to identify the areas where there was a clear need for a What Works Centre. This consultation resulted in start-up funding, particularly from government departments. In the case of four centres, the ESRC provided significant co-funding. The Centre for Ageing Better is an exception as it received a grant from the Big Lottery Fund, a public body with greater independence than government departments.

**THE ROLE OF THE CABINET OFFICE**

A team of three people within the Cabinet Office coordinates and leads the What Works Centres, a role with five strands: support in establishing a framework for the new centres; support in establishing a framework for the new centres;
– support in securing renewed funding for the centres;
– developing links between centres and government departments (civil servants, political advisors etc.);
– leading the network of centres, with a quarterly meeting of the What Works Council, where centre directors share their practice and potential areas for collaboration;
– promoting the centres worldwide.

Thanks in part to the Cabinet Office, the centres have three levels of collaboration:
– sharing practice at quarterly What Works Councils, as well as in other contexts;
– the sharing of synthesis reports between the centres;
– collaborative research projects or trials.

Governance

The legal statuses of the centres vary: some are public, others charities and yet others are more akin to a programme of activity led by a consortium. As such, their formal governing bodies vary. Despite these differences, the centres share a lot in their conception of governance and the role given to stakeholders. Indeed, stakeholders from the worlds of research and public policy sit on the permanent bodies of most of them. The research projects they run can also include stakeholder participation. NICE is one example of this, as its Guidance is produced directly by a “Committee” of health professionals and other operational stakeholders.

To find out more: NICE Case Study (page 156)
What the centres do

Synthesising existing research

The What Works Centres (WWC) aim to bring research on “What Works” in their respective policy areas to the attention of stakeholders in the public and charity sectors. These reviews are primarily aimed at responding to the needs of frontline professionals and middle managers. To have access to up to date research in their policy area, all of the WWCs conducted an extensive programme of systematic reviews of intervention evaluations. By carrying out a series of reviews over approximately a three-year period, most of the centres covered a significant portion of the research in their policy area. Most of them make this scientific content accessible by presenting it in the form of a “toolkit” (see page 50).

In order to avoid the inherent bias in traditional literature reviews, these centres use the “systematic review” method to synthesise the research. This method is “systematic” in that it follows a robust research protocol, from a scoping exercise with the stakeholders via surveying and identifying studies, to selecting, evaluating and synthesising these. The What Works Centre for Economic Growth’s protocol is a good illustration of this general approach.

▶ To find out more: What Works Centre for Economic Growth Case Study (page 131)
All of the centres aim to draw on the “best existing studies” in carrying out these reviews. All of the centres place a strong emphasis on RCTs and quasi-experimental studies, with varying degrees of flexibility. A number of them have developed a hierarchy to distinguish between the relative robustness of the various evaluation methods, often inspired by the *Maryland Scientific Methods Scale (MSMS)*.

**MARYLAND SCIENTIFIC METHODS SCALE**

**Level 1.** Either (a) a cross-sectional comparison of treated groups with untreated groups, or (b) a before-and-after comparison of treated group, without an untreated comparison group.

**Level 2.** Comparison of outcomes in treated group after an intervention, with outcomes in the treated group before the intervention (“before and after” study).

**Level 3.** Comparison of outcomes in treated group after an intervention, with outcomes in the treated group before the intervention, and a comparison group used to provide a counterfactual
British What Works Centres: what lessons for evidence-based policy in France?

Level 4. Comparison of outcomes in treated group after an intervention, with outcomes in the treated group before the intervention, and a comparison group used to provide a counterfactual (e.g. difference in difference).

Level 5. Reserved for research designs that involve explicit randomisation into treatment and control groups, with Randomised Control Trials (RCTs) providing the definitive example.

Source: What Works Centre for Local Economic Growth.

Some automatically exclude evaluations with no counterfactual - the What Works Centre for Local Economic Growth only analyses studies that correspond to levels 3 to 5 of the MSMS. In being selective, this centre has been confronted with the limited supply of RCTs and quasi-experimental studies in the area of local economic growth.

“Some of the programmes aren’t amenable to the highest quality methods. Active labour market programmes, or for start-ups and business and advice – you can use randomisation. With housing and infrastructure, you can’t randomise where that goes...”

Max Nathan, Deputy Director, What Works Centre for Local Economic Growth (LSE).

Others include a larger sample of evaluations, whilst specifying the conditions for inclusion in the studies. The Early Intervention Foundation (EIF), for example, scores the studies, from the presence of a simple logical model, supported by arguments (level 1) to at least two evaluations using a control or comparison group (level 4). The EIF has adopted this approach in light of the difficulty in determining beyond dispute whether an intervention subject to evaluation has actually had positive and measurable outcomes. A handful of interventions are clearly effective or ineffective based on the evaluations carried out. However, many interventions are in a “grey area”, either because the evaluations show that they have no effect or a very limited effect, or because they do not manage to identify an effect.
(lack of statistical power, problems of attribution etc.). By including a wide range of evaluations, the EIF wants to give programmes the opportunity to be fine-tuned, improved and better designed over time, and to benefit from more robust evaluation methods. The EIF takes inspiration particularly from the “Standards of Evidence” produced by NESTA, where evaluation is part and parcel of the development from its design to its dissemination on a wider scale.

To find out more: Early Intervention Foundation Case Study (page 109)

**NESTA Standards of Evidence**

<table>
<thead>
<tr>
<th>Niveau 1</th>
<th>Vous pouvez décrire votre action, les besoins auxquels elle répond, de façon cohérente et convaincante.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Niveau 2</td>
<td>Une évaluation démontre des résultats positifs, mais vous n’avez pas établi qu’ils sont liés à votre action.</td>
</tr>
<tr>
<td>Niveau 3</td>
<td>Vous avez démontré qu’il existe un lien causal entre votre action et les résultats positifs grâce à un groupe témoin.</td>
</tr>
<tr>
<td>Niveau 4</td>
<td>Vous avez au moins une évaluation de réplication qui confirme le lien causal entre l’action et les résultats positifs.</td>
</tr>
<tr>
<td>Niveau 5</td>
<td>Vous disposez des guides, systèmes et processus permettant de réplicer l’action, et de reproduire un impact positif.</td>
</tr>
</tbody>
</table>

*Source: Standards of Evidence, Nesta, 2013.*
Cost analysis is a central component of most systematic reviews carried out by the WWCs. All of these include a calculation of the cost of interventions evaluated, allowing comparisons to be made. Depending on the purpose of the studies and the information available, some centres go as far as to carry out cost-effectiveness analysis. Such analysis is useful in comparing programmes with the same objective (for example academic attainment). In order to carry out such analysis, in its toolkit the Education Endowment Foundation has compared academic attainment through an evaluation on the basis of “school progress in months", with an estimate of the cost per pupil of each intervention.

Cost-effectiveness of educational interventions


Notes: Cost estimates are based on the likely costs of adopting or implementing an approach for a class of twenty five pupils. The interventions in bold are those for which the standard of evidence is high to very high based on the definition in the toolkit. School progress in months is estimated by comparing the change in scores of pupils in standardised tests between the start and end of the school year. As such, twelve month’s progress is the average of all pupils. Where a programme results in a class making more progress than average, that progress is quantified in additional months’ progress.
Some also include the benefits of interventions, expressed as a financial sum. The advantage of a cost-benefit analysis is that a wide range of effects can be included in the benefits. Indeed, this method is often used for preventative measures that have multiple effects. For example, the Early Intervention Foundation uses a number of cost-benefit analyses of interventions aimed at young children. The High Scope Perry Preschool Program, for example, had a cost-benefit ratio of 1:7 or 1:12, mainly due to higher employment and educational attainment levels, as well as lower levels of criminality and less reliance on social welfare. This type of analysis is often used to calculate the monetary value of preventative measures, which reduce costs.

The WWCs also examine the potential for learning wider lessons from specific studies. On the one hand, reviews generally include studies carried out in OECD countries, raising the question of whether findings in other political, institutional and cultural contexts can be transposed to the British context. On the other hand, the WWCs have come up against the “black box” aspect of scientific articles. Indeed, certain evaluations do not make available detailed presentations nor the method of the intervention, nor the conditions in which the intervention seems to produce results. The WWC for Crime Reduction has produced a framework for analysing interventions in order to fully incorporate this issue of transferability in their reviews (the “EMMIE” framework).
EMMIE FRAMEWORK

Effect size: the overall effect direction and size of an intervention and the confidence that should be placed on that estimate;
Mechanisms: the mechanisms activated by the policy, practice or programme in question;
Moderators or contexts relevant to the production/non-production of intended and major unintended effects of different sizes;
Implementation refers to the key sources of success and failure in implementing the policy, practice or programme;
Economic costs: The economic costs (and benefits) associated with the policy, practice or programme.

Systematic reviews are not a new practice. However, WWCs differ in particular in terms of their accessibility to a lay audience. In particular, all of the WWCs have made a greater effort to translate the analysis into findings for professionals.

For example, the eleven reviews of the WWC for Local Economic Growth all come with an Executive Summary which synthesises and explains the findings of the study in laymen’s terms for a professional readership.

“So there was a lot of work with academics to be more explicit with their recommendations, and to stop sitting on the fence. There is a creative tension between being rigorous and having an impactful message.”

Meg Kaufman, What Works Centre for Local Economic Growth at the Centre for Cities.

Although systematic reviews are the main type of research used by the centres, a number of them employ other methods. Most of the centres have an in-depth understanding of their subject area, in terms of the profile of the users in question, important socio-economic and health trends, needs and existing public service provision. All of this information enables the centres to prioritise subjects for research, evaluation and innovation. Two of the centres – the Centre for Ageing Better and the WWC for Wellbeing – wanted
to develop understanding of the poorly understood phenomena of ageing better and wellbeing. To do so, they set up and carry out sociological studies (surveys, interviews, participant observation) and consultations.

To find out more: Case Study of Centre for Ageing Better (page 143)

Finally, the regional centres are an exception in terms of their methodological approaches. Indeed, their research programme is more ‘responsive’ and developed jointly with the devolved authorities in Scotland and Wales. They draw on a wider range of research types and are a lot less attached to RCTs and quasi-experimental studies than the thematic centres. With this degree of heterogeneity in terms of method, when discussing the approaches behind ‘What Works’ it is probably legitimate to separate out the thematic and regional centres.

THE REGIONAL WHAT WORKS CENTRES: DIFFERENT APPROACHES

The Public Policy Institute for Wales was the result of a Labour Party manifesto commitment in the run up to the elections to the Welsh National Assembly in 2011. A consortium of universities responded to the call for proposals and the centre was officially launched in January 2014. Every six months, the team meets the government departments of Wales to draw up their work programme. It then enlists teams of academics who deliver the projects – often in the form of research synthesis reports – and arranges reporting to the minister, his or her advisors and senior officials. In two years, the centre has completed 60 research projects. Its close relationship with ministers, its responsiveness and its ability to coordinate numerous research projects has made it particularly useful for Welsh national decision-makers.

What Works Centre for Scotland was set in 2014 to help local actors to use evidence when making decisions about public service development. One of the centre’s objectives was to promote preventative and participatory approaches in the context of austerity. Their approach is highly devolved and their 2014-2016 work programme
was based mainly on partnerships with stakeholders in four council areas: Aberdeenshire, Fife, Glasgow and West Dunbartonshire.
In each council area, the centre co-produces an assessment of the local context, facilitates dialogue between stakeholders, offers a tangible definition for “effective services” and promotes the use of research in the planning and implementation of services. Unlike the thematic centres, its approach does not focus mainly on RCTs and quasi-experimental studies; in fact, any form of evidence that can inform public services is put to use (for example observational data).

The purist approach of the What Works Centres has attracted some criticism. In particular, some stakeholders call into question the idea of ranking research with randomised evaluations systematically coming first.

“There is no simple answer to the question of what counts as good evidence. It depends on what we want to know, for what purposes, and in what contexts we envisage that evidence being used... in practice, decision makers need to consider evidence in the round. Policymakers and practitioners need to weigh evidence relating to what works alongside evidence about cost, acceptability and distributional effects”

What counts as good evidence?, Alliance for Useful Evidence*

Dissemination and promoting uptake

Research synthesis is not an end in itself. The WWCs primarily aim to support frontline stakeholders in adopting effective practices for better dissemination of research.

All WWCs make their reviews accessible. For most centres, this work begins with the systematic reviews being translated into findings for professionals. It continues with the dissemination of these publications in digital formats and through events. In spite of their value, these individual reviews give just a partial snapshot of research.

As such, in order to become true “one-stop shops” to research most WWCs
have developed libraries of interventions aimed at offering a synthetic overview of the research. These libraries present a scientific assessment of the interventions, based on at least three criteria: the impact, the robustness of the research and the cost. Designed as genuine decision-making tools for civil servants, they are interactive and accessible. The Education Endowment Foundation’s Teaching & Learning Toolkit is a particularly good example. The Toolkit is an interactive tool available via the centre’s website. It is also dynamic. Updated every six months by the research team at Durham University, the number of programmes has increased from 11 to 34. Analysis of the programmes has also been updated to take into account new studies.

THE TEACHING & LEARNING TOOLKIT EXPLAINED

The Teaching & Learning Toolkit is a review of reviews of research into the effectiveness of the various educational interventions. As such, this synthesis includes mainly the results of experimental or quasi-experimental evaluations. The Toolkit presents analysis relating to various educational interventions (or 'strands'), based on three aspects:

- **Impact.** In order to compare the impact of interventions, the research team has translated the size of the effects observed during the evaluations into a single measure: additional months of educational progress for the beneficiaries (in relation to a counterfactual).

- **Evidence strength.** This measures both the quality and number of studies carried out, with a central focus placed on evaluations with a control or comparison group;

- **Cost.** This is an estimate of the annual per-pupil cost if the interventions were to be implemented in classes of 25 pupils, typically including the costs of teacher training or the production of teaching materials.

More informations at https://educationendowmentfoundation.org.uk/resources/about-the-toolkits/
Extract from the Teaching & Learning Toolkit

Source: educationendowmentfoundation.org.uk/evidence/teaching-learning-toolkit.

To find out more: Education Endowment Foundation case study (page 120)

THE OTHER WHAT WORKS CENTRES INTERVENTION LIBRARIES

What Works Centre for Local Economic Growth
– Toolkits: www.whatworksgrowth.org, Resources.
Programmes Library du Early Intervention Foundation: guidebook.eif.org.uk, Programmes library.
These libraries are accessible to a lay audience and are a means of improving access to research by practitioners and middle managers. They are interactive and include comparative analysis of the costs of various interventions, making them real decision-making tools. These tools represent a break with previous ways of presenting research results. Indeed, before they were created, a large proportion of scientific research was circulated mainly between researchers in paying academic journals. Three of the centres with such a library emphasised their importance in opening up a dialogue with professionals around research (local economic growth, crime reduction and education). Few of these tools have evaluated their impact. The exception is the Education Endowment Foundation. According to a survey of some 543 head teachers carried out by the National Audit Office in 2015, the Teaching & Learning Toolkit is used by some 64% of them to inform their decisions around pedagogical practice.\(^{\text{vi}}\)

“The toolkit responds to some of the constraints facing frontline professionals: lack of time, too much information and unable to sift through knowledge. The toolkit was a game changer in making education research useful to practitioners.”

Jonathan Sharples, chercheur expérimenté, Education Endowment Foundation.

Thus, making research accessible to professionals has been a necessary and useful step

However, none of the research centres see mere research dissemination – however clear it is – as sufficient to influence frontline practice. The WWCs therefore all support the uptake of this scientific material by stakeholders.

“Creating online resources isn’t enough. A big theme in knowledge mobilisation is that it is a social process, you have to create motivations and opportunities to bring that to life.”

Jonathan Sharples, Senior Researcher, Education Endowment Foundation.
Some centres support national and regional decision-makers in embedding research in their policy-making. The regional centres and, in particular that of Wales, are more focussed on national policymakers. As a sort of “What Works” service for the Welsh National Assembly, the Public Policy Institute for Wales is in constant dialogue with ministers, their advisors and senior civil servants. Most of their reports are thus produced quickly to inform policy-making and are fed back to policymakers orally. Thethematic centres generally have a more ad-hoc approach to helping national policymakers to integrate the research. The Education Endowment Foundation, for instance, produced a briefing paper for the Prime Minister to notify him of the positive findings of trials using teaching assistants.

Decentralisation has also opened up opportunities to influence political decisions at the regional or local level. In 2012, the first elections of the Police and Crime Commissioners took place in England and Wales with the aim of making the police service more accountable to the public. These elections coincided with the official launch of the WWC for Crime Reduction. The College of Policing – which emerged a short time after the WWC – created an academic panel to synthesise research findings to feed into the Commissioners’ programmes. This vast body of scientific research was then presented in the form of four key questions asked by decision-makers in the police service.

The thematic centres aim to promote the use of research mainly by middle managers and frontline professionals. The centres produce three types of tools to support adoption:

– libraries and systematic reviews to raise awareness of the types of intervention that appear to be more effective on the whole. For example, The Early Intervention Foundation’s report on parent-child interaction identified 17 programmes with positive and measurable benefits and 18 promising programmes;

– self-evaluation tools which enable stakeholders to compare their current practice with so-called “effective” interventions. As such, the EIF produced an Early Intervention Maturity matrix which enables local authorities to identify how well they are doing in terms of knowing their users, developing
their teams and services and policy levers etc. (see case study on page 109);
– intervention descriptions or guides which offer detailed definitions of the
  intervention methodology as well as background that may affect outcomes.
Thus, for every public policy, the WWC for Local Economic Growth has
produced a toolkit made up of several guides to specific interventions.
For example, the employment training review has been turned into a
toolkit made up of four guides: careers counselling; financial incentives;
pre qualifications and reminders (to improve attendance on programmes).

The centres circulate these toolkits online but also at policy discussion days,
conferences and via ad hoc workshops or training. The centres seek to
cover the whole chain, from research to the adoption of these practices
by professionals. The aim is that professionals are able to easily implement
practice which is ‘state of the art’ in terms of effectiveness.

TEACHING ASSISTANTS: FROM REVIEW TO ADOPTION BY TEACHING PROFESSIONALS

EEF’s work on teaching assistants are a real-life example of this.
Educated at least to GCSE level, these teaching assistants support
teachers, particularly in offering academic support to low-attaining
pupils. The systematic review showed that, on average, teaching
assistants had either no impact or a negative impact on pupils. These
findings were of particular concern to the team as teaching assistants
make up a significant proportion of those working in education in
England. At that time, 240 000 teaching assistants were employed at an
annual cost of £5 billion. As such, the EEF team wanted to trial different
ways of making use of these practitioners. “We funded 6 trials looking
at how to use teaching assistants to do one-to-one interventions, for
instance a literacy scheme. Every one of the studies showed positive
effects on the students with consistent effect sizes. EEF thus decided to
turn these research findings on the use of teaching assistants (including
six trials funded by the EEF), into an operational guide. This guide was
accompanied by resources to promote the adoption of good practice
such as a self-assessment guide. These resources were circulated to
every school in England as part of a national campaign.”
Some centres directly support the scaling up of effective programmes, such as the Centre for Ageing Better which publishes “scale-up” opportunities. Other centres have opted for initiatives aimed at building a professional culture around evaluation and evidence-based policy-making. For example, the WWC for Crime Reduction, hosted by the College of Policing, has introduced modules on evaluation and “evidence” in its initial training and continuing development courses for police officers. Several WWCs have emphasised that such a professional culture is a precondition for their success and will only change through long-term support.
“We need at least 10 years to do what our funders want us to do. We’ve needed 3 years to go through a critical mass of the evidence. A lot of what we do is about communicating about the evidence at the local level, and to support local governments to apply the findings locally. A large part of our role is cultural and institutional influence and change. That takes a lot of time to do – we’re dealing with large institutions and traditions of not doing evaluation, or not doing evaluation well.”

Max Nathan, Deputy Director, What Works Centre for Local Economic Growth (LSE).

Aware of the difficulty in changing professional practice country-wide, several centres are conducting more intensive dissemination activities in specific local areas:
– NICE’s eight implementation consultants are knowledge brokers at the regional level, raising awareness of NICE resources among health stakeholders;
– hospitals, GPs and local authorities;
– through the programme Pioneering Places, the EIF has worked with some 20 local authorities in identifying activities to better embed the principle of early intervention in their services. The EIF has recommended initiatives to improve the identification of the target population and then determined the best way of supporting them, drawing on effective interventions identified by systematic reviews;
– the Centre for Ageing Better has also partnered up with Greater Manchester to conduct pilot projects and apply ‘evidence’ to local decision-making;
– the EEF has developed a menu of interventions identified as effective and already offered by third-sector providers in England. As part of an initial pilot scheme in an area near London, participating schools receive co-funding and support from the EEF in exchange for adopting these proven practices.

WWCs test a range of models for disseminating “What Works”. These are real laboratories of evidence-based policies, the findings of which are not yet conclusive. However, these centres are usually well known to professionals in their fields. The centres have also had a number of notable successes in terms of disseminating ‘effective’ practice.
DISSEMINATION OF BODY CAMERAS IN LONDON

Body cameras are devices worn by police officers that record all of their interactions and operations. In order to test the benefits of these devices, the WWC for Crime Reduction conducted a randomised trial in London\textsuperscript{lxiv}. The positive results of this evaluation were widely disseminated and in October 2016 the Metropolitan Police announced that 22,000 police officers in London would be equipped with such devices\textsuperscript{lxv}.

However, the WWCs have little in the way of objective evidence of their impact on professional practice. The teams openly acknowledge that they do not yet know “What Works” in terms of rolling out effective practice. The centres are all seeking to improve their dissemination activities, with some going as far as to evaluate them. For example, the WWC for Crime Reduction’s work is continuously evaluated\textsuperscript{xvi}. The Education Endowment Foundation has even opted to carry out randomised trials of various dissemination methods.

**Evaluating the impact of innovations**

Funding innovations and conducting trials are not central to the remit of all the centres. Indeed, only two WWCs fund or carry out RCTs or quasi-experimental studies.

**The Education Endowment Foundation is by far the most prolific WWC in terms of launching new trials.** With an endowment of £125 million, the EEF funds around 20 new trials every year. The EEF funds both strands of trials: the innovative project and the evaluation. Projects are designed and implemented by frontline stakeholders – schools, charities or universities – and selected through a call for projects. By funding the project delivery, EEF encourages structures to submit to a scientific evaluation, despite the risk of these showing a negative or neutral impact. Depending on the stage of development, each project is assigned a level of support as a pilot project, an efficacy trial or an effectiveness trial. Some promising projects are the subject of pilots to test their feasibility before subjecting them to an
RCT. With the exception of these pilot projects, for each trial, the EEF funds a robust evaluation carried out by an external evaluator. There is a focus on experimental methods in order to analyse the project impact. The EEF has encountered technical difficulties and, in some cases, resistance or misunderstandings on the part of project leads. However, on the whole, the EEF has proved those critics who feared a strong resistance on the part of teachers to experimental evaluation methods, wrong. Indeed, it has carried out 127 evaluations including 105 as randomised trials in 7,000 primary schools, i.e. one in four schools in England.

“A few years ago, one RCT had been done in the education sector in England. Suddenly we’ve got over 100. It’s been like flicking a switch!”

Jonathan Sharples, Senior Researcher, Education Endowment Foundation.

EEF reviews feed into social experiments and vice versa. Indeed, the Toolkit has also helped the EEF to identify the types of intervention which would benefit from further research and which might be the subject of new trials funded by the EEF. Ultimately, these trials produce findings which feed into the Toolkit. The example of teaching assistants (see pages 63-64) is a good illustration of this virtuous circle.

“The evaluations of TAs reviewed for our toolkit suggested it was having a neutral, even negative impact. The toolkit brought the issue out – that has driven a lot of action at our end. So we have funded trials looking at how to use TAs to do one-to-one interventions…. [which] showed consistently positive effects on the students. These results then, in turn, changed the conclusions of our Toolkit.”

Jonathan Sharples, Senior Researcher, Education Endowment Foundation.

Analysts at the WWC for Crime Reduction also evaluate the impact of innovations. Owing to its limited resources, this centre evaluated the impact of just a few innovative projects implemented by the police service as part of pilot schemes. Using the findings from pilot projects as a starting point, the centre focusses on high-potential innovations for
trials. The roll-out trial of body cameras in London is a good illustration of this approach\textsuperscript{lvii}.

Finally, the Centre for Ageing Better does not fund trials at this stage. However, it has resources dedicated to the implementation of pilot projects and is planning to raise funds from academia in order to fund evaluations.

Observing the lack of scientific evaluations in their area, some centres are attempting to encourage or launch new trials. For example, NICE Guidance often recommends subjects for clinical trials. The WWC for Local Economic Growth offers personalised support in the design and launch of trials in its policy area, particularly for local authorities. This has often proved to be an arduous task. Thus, in Summer 2016, although the team was involved in the design of 20 pilot projects, just 4 were moving towards implementation. In order to “enhance” the quality of evaluations, several centres carry out initiatives to raise awareness of RCTs and quasi-experimental methods. In some cases, these initiatives are beginning to bear fruit.

“So we’re working with government departments to improve the quality of the evaluations they commission. Today, the invitation to tenders for evaluation projects sent out by government departments now often include Maryland Scale standards as minimum, and something about process evaluation.”

Max Nathan, Deputy Director, What Works Centre for Local Economic Growth (LSE).

A number of centres have expressed their frustration at their inability to evaluate the impact of innovations in Britain. Conversely, the capacity to support trials has been a real asset for the EEF and WWC for Crime Reduction: on the one hand in involving frontline stakeholders as innovators and, on the other, in pushing forward research on “What Works”.
Key findings on What Works Centres

In terms of synthesising existing research

- most What Works Centres have adopted a “purist” approach, with a heavy emphasis on RCTs and quasi-experimental studies, systematic reviews, and cost-benefit analysis;
- this approach has had benefits, enabling centres to identify the relative effectiveness of a range of interventions and thus to provide decision-makers and professionals with useful findings;
- however, the production of systematic reviews has not been without difficulty, in particular due to the lack of RCTs and quasi-experimental studies, the complexity in determining the effectiveness of an intervention and in clarifying the contexts in which an intervention works;
- the difficulties encountered in producing these reviews led a number of centres to adopt a more nuanced approach, for example in better taking into account the issue of transferability or scalability.

In terms of dissemination and support for adopting the findings

- the WWCs have turned certain evaluations into easy-to-read findings, particularly in their reviews and libraries of interventions.
  As such the centres have quickly made themselves valuable to policymakers and practitioners;
- making this scientific subject matter accessible has been necessary but not sufficient on its own for the stakeholders to take full advantage of it;
- as such, some centres support national and regional decision-makers in incorporating research into policy-making;
- however the centres mainly target middle managers and frontline professionals, developing a range of tools to support them in adopting effective interventions. A number of centres implement regional dissemination activities in order to promote adoption of effective interventions by professionals;
British What Works Centres: what lessons for evidence-based policy in France?

- the WWCs do not claim to have identified the most effective dissemination methods. They are continuing to test various approaches, including through experimental trials;
- some WWCs consider that their success depends on changing professional cultures in their sector, which requires long-term work on awareness-raising and training.

In terms of evaluation

- for the Education Endowment Foundation and the WWC for Crime Reduction, the ability to support trials has been a real asset, in improving knowledge of “What Works” in the British context and bringing them closer to frontline actors;
- the other centres’ activities are, on the whole, restricted to training, support and awareness-raising, with generally mixed results at this stage in terms of new trials launched.

xxxvi See the site www.rssenews.org.uk/articles/20100604_1.
xli For example, the trial “Employment Retention and Advancement (ERA) Demonstration” was carried out between 2003 and 2007.
xlii Led by Kings College London in partnership with other universities, the main aim of this centre was to promote the sharing of relevant research by researchers and practitioners in the area of public policy, in particular by synthesising research findings. Source: www.researchcatalogue.esrc.ac.uk/grants/H141251005/read.
xlv Graham Allen MP, Early Intervention: The Next Steps An Independent Report to Her Majesty’s Government,
xlviii What Works Centre for Crime Reduction, What Works Centre for Local Economic Growth, What Works Centre for Wellbeing, and What Works Scotland.
xlix Available online: www.nice.org.uk/process/pmg20/chapter/1-introduction-and-overview#whois-involved.
lxii See the site: www.eif.org.uk/publication/early-intervention-maturity-matrix.
lxiii See the site: www.whatworksgrowth.org/resources/employment-training-toolkit.
lxiv See the site: www.eif.org.uk/pioneering-place.
lxv See the site: whatworks.college.police.uk/About/Pages/Evaluation.aspx.
lxvii See the site: whatworks.college.police.uk/About/Pages/Evaluation.aspx.
What culture of evaluation and evidence for public policy in France?

An administration focussed on needs analysis

Observational data, such as needs analysis, is a key component of ‘evidence’. Indeed, it allows public policies to be developed in full knowledge of the challenges and problems, but also the affected groups and their characteristics and needs.

France has a dense network of organisations the main function of which is to observe issues in French society, in particular statistical services, national observatories and research centres. At a national level, collecting and analysing observational data is carried out by national statistical services (Institut national de la statistique et des études économiques, Institut national d’études démographiques) and government departments (for example Direction de l’évaluation, de la prospective et de la performance [Depp] at the Education Ministry (Ministère de l’Éducation nationale), the Direction de l’animation de la recherche, des études et des statistiques [Dares] at the Employment Ministry (Ministère du Travail et de l’Emploi), the Direction de la recherche, des études, de l’évaluation et des statistiques [Drees] at the Social Affairs and Health Ministry (Ministère des Affaires sociales et de la Santé), by public bodies or divisions and teams within such bodies (for example France Stratégie, CGET, Santé publique France and SGMAP), also within social security bodies (for example Cnaf, CNSA and Cnav). There are also teams that have this monitoring function at the local level, within local authorities or local branches of state departments, as well as independent bodies (for example in the case of health the Observatoires régionaux de santé). Research centres also play a central role in implementing research projects, in some cases for or with the organisations listed above. Thanks to these actors, France has a rich pool of data and studies relating to the nature and scale of these social policy issues. Their research can be both quantitative and qualitative and is, in principle, geared towards public decision-making.
Financial and human investment and technological advances have improved the quality and responsiveness of monitoring bodies in France in recent years. Thanks to these efforts, France now has a growing number of surveys, cross-sectional and longitudinal, that can be used to analyse the characteristics of individuals, their needs, and changes in these over time (for example, the employment surveys carried out by the DARES). There have been significant improvements in terms of opening up and sharing public data ("open data"), a policy piloted by Etalab and the Head of Data (Administrateur général des données). For example, the site data.gouv.fr offers access to over 355,000 free and reusable public data.

Despite these advances, difficulties relating to observational data in France persist. Indeed, there are many policy areas for which there is no national standardised data production, nor information systems allowing exploitation and dissemination. For example, in the area of disability, at the level of the Département the Maisons départementales des personnes handicapées (MDPH) have each developed their own information system so the data so that the data entered there cannot be exploited. The design by the CNSA of a national standardised framework of their information system and their support for rolling it out has helped to remedy this. Indeed, some of these data will be incorporated into the new national health data system (Système national des données de santé - SNDS). The SNDS, which is in the process of being set up, combines two approaches: free access to anonymous health data and secure access to personal health data for the purposes of research, studies and evaluation in the public interest. Indeed, this secure access opens up the possibility of carrying out evaluations (quasi-experimental methods) at a low cost using large samples.

Evaluating public policy

Public policy evaluation has grown significantly since the 1960s without becoming completely institutionalised. This method, which arose out of the coming together of American social sciences and national decision-makers in several areas of public policy, was first developed in the United States.
In France the use of public policy evaluation goes back to the 1960s with the streamlining of budgets (Rationalisation des choix budgétaires – RCB). Led by senior civil servants from the budget directorate, the RCB (1965-1977) was a set of programmes intended to ensure public spending decisions were made on the basis of forecasting their socio-economic impacts. At the end of the 1980s, initiatives of the Michel Rocard government mainstreamed the evaluation of public policies in a new set of administrative reform policies. A set of measures jointly led by the Parliament, the Commissariat général au plan and an independent body (Conseil scientifique de l’évaluation then the Conseil national de l’évaluation) thus established evaluation as an essential vehicle for the modernisation of public policies and strengthening the democratic legitimacy of public action. In parallel, a new obligation to evaluate programmes put out to tender, in particular in European structural funds, reinforced evaluation practice within local authorities.

From 2000, two successive government reform agendas combined the promotion and evaluation of public policies with the search for more effective public intervention: the Loi organique relative aux lois de finance (LOLF) and the Révision générale des politiques publiques (RGPP). Although in reality these two initiatives were more audit and inspection type initiatives, they nevertheless contributed to increasing the visibility of the vocabulary and issues of evaluation within the State and government agencies.

2012 saw a turning point in public policy evaluation with the public action modernisation agenda (Modernisation de l’action publique – MAP) the objective of which was to evaluate “all public policies” by the end of the five-year term. This cross-government evaluation programme was delivered mainly by the state inspectorates.
Between 2012 and 2016, the Secrétariat général de la modernisation de l’action publique (SGMAP) implemented 80 public policy evaluations – known as “EPPs” – as part of the public services modernisation agenda. EPPs involved government departments and state providers, social security bodies and local authorities. These evaluations, most frequently delivered by general inspectorates but also by experts, use a combination of methods (literature reviews, quantitative analysis, surveys and consultations) with project management support provided by the SGMAP. Evaluations produced recommendations for policymakers. A review of the implementation of these recommendations shows that two out of three completed EPPs resulted in follow-up (ministerial decisions, action plans implemented, laws passed etc.).

For example, the EPP Éducation prioritaire, completed in 2013, contributed to changes such as the overhaul of the mapping of education action zones and the objective of reducing the attainment gap being inscribed in the school reform law of 18 July 2013 (Loi pour la refondation de l’école).

Moreover, local authorities are becoming increasingly committed to evaluation and have been for the last twenty years, as they have been affected by laws on decentralisation, the increased importance of European funding and other reforms. Indeed, advances resulting from the first local trials, in particular in the 1990s, helped to disseminate evaluation practices to the various levels of government. “Evaluation structures were set up first by regional authorities, mostly in the interests of empowering and lending legitimacy to this level of government, followed by the conseils départementaux (county council equivalent) for social policy, and municipalities particularly for the purposes of evaluating urban policy.” states the LIEPP report, Les scènes multiples de l’évaluation.
In terms of stakeholders, promotion of public policy evaluation is carried out by a combination of government departments and state and independent bodies:

- **general inspectorates** including the Inspection générale des affaires sociales (IGAS), and the Inspection générale des finances (IGF);
- **the national auditor (Cour des comptes)** which was given the remit of evaluating public policies under the constitutional reform of 2008;
- **government department directorates**: in particular the Direction de l’animation de la recherche, des études et des statistiques (The directorate for overseeing research, studies and statistics - Dares), the Direction de la recherche, des études, de l’évaluation et des statistiques (The Directorate for research, studies, evaluation and statistics - Drees), and the Direction de l’évaluation, de la prospective et de la performance (The Directorate for evaluation, forecasts and preformance - Depp);
- **several cross-government stakeholders** including:
  - the Secrétariat général pour la modernisation de l’action publique (The Secretariat General for the Modernisation of Public Action – SGMAP) through its EPP-MAP agenda;
  - France Stratégie, with its remit to contribute to the evaluation of public policy, has been commissioned to carry out a number of evaluations including through the Comité de suivi des aides publiques aux entreprises et des engagements (the Committee for monitoring public grants for businesses and commitments)\(^6\). It has also published a guide to evaluation\(^7\);
  - The Commissariat général à l’égalité des territoires (The General Commission for Regional Equality – CGET), the evaluation remit of which is delivered by the “Direction des stratégies territoriales” (Directorate for regional strategies);

\(^6\) Also the Commission nationale d’évaluation des politiques d’innovation (National Commission for the Evaluation of Innovative Policies - CNEPI) and the Commission d’étude des effets de la loi pour la croissance et l’activité (the Commission for the Study of the Effects of the Law for Growth and Activity).

\(^7\) Rozenn Desplatz, Marc Ferracci, *Comment évaluer l’impact des politiques publiques? Un guide à l’usage des Décideurs et praticiens*, France Stratégie, September 2016. This briefing paper offers an overview of the various methods of public policy evaluation, their benefits and limitations: controlled, or randomised experimentation, natural experiments, cost/benefit analysis, modelling (structural approach), the qualitative method (observation of a sample, interviews etc.), etc. The guide is illustrated with examples of how these various approaches to evaluation are used.
- **Parliament**: with the Comité d’évaluation et de contrôle des politiques publiques (Public Policy Evaluation and Oversight Committee - CEC) of the National Assembly;
- **local authorities** are particularly committed to issues around evaluation, at the regional, departmental and metropolitan levels;
- **the Conseil économique, social et environnemental (Economic, Social and Environmental Council – Cese)** whose advisors are from civil society, has had a Délégation à la prospective et à l’évaluation des politiques publiques (Public Policy Forecasting and Evaluation Delegation) since 2010. See also its regional equivalents (Cesr);
- **stakeholders from the non-profit and private sectors** several hundred of which are members of the Société française d’évaluation (French Society for Evaluation - SFE) established in 1998. Its members implement public policy evaluations and commit to the principles of pluralism and independence set out in the SFE charter;
- **researchers**, for example at the Laboratoire interdisciplinaire d’évaluation des politiques publiques (Interdisciplinary Laboratory for Public Policy Evaluation – Liepp) at Sciences Po who develop economic and quantitative approaches to evaluation in its policy areas, or the École d’économie de Paris, which hosts J-Pal Europe and the Institut des politiques publiques (IPP).

This overview of public policy evaluation shows a paradox in terms of French evaluation culture. On the one hand, it is evidence of an undeniable buzz around an approach and a field of research that is now an established part of the landscape. On the other, it highlights the weaknesses of an approach that is not yet fully embedded both in terms of its tools and methods and from an institutional perspective. The sphere of evaluation brings together competing, or at the very least divergent, methodological approaches, that commentators have termed different “schools of thought” of evaluation

In addition, there are challenges around publicising, disseminating and mobilising evaluations. The vast majority remain under used by decision-makers and practitioners with the wider public remaining unfamiliar with them. Evaluation reports are not systematically made public: the scientific journals that publish evaluations tend to be fee-paying. Another explanation
put forward is the poor readability of evaluation reports. Still another that the approach to policy evaluation does not always offer responses to some of the questions raised by decision-makers and frontline professionals.

The development of RCTs and quasi-experimental studies

**Rapid development of these new methods**

Unlike other developed countries, particularly in the English-speaking world, in France RCTs and quasi-experimental studies have become an established part of social policy relatively recently. Several factors have led to a surge in their use in the last ten years. The Loi organique relative à l’expérimentation (the Organic Law on Experimentation) of 1 August 2003, opened the way to those wishing to conduct social experiments in France. However, this practice has mainly developed since 2007, instigated in particular by Martin Hirsch. His leadership resulted in substantial investment in such approaches, particularly in trialling a new welfare benefit, the revenu de solidarité active (RSA) and the creation of an experimental youth fund, the Fond d’expérimentation pour la jeunesse. This movement has had a number of successes, for example in identifying effective programmes that have been disseminated (see page 91). Some experiments also produced useful findings which were brought to the attention of public decision-makers (see page 85). This trend has raised hopes that these robust methods would ensure public resources were allocated to programmes with proven impact. However, questions have arisen in its wake on the feasibility and use of RCTs and quasi-experimental studies, particularly for social and economic policy.

A dynamic that is continuing and being built upon: THE FOND D’EXPÉRIMENTATION POUR LA JEUNESSE

The Fond d’expérimentation pour la jeunesse (Youth experimentation fund – FEJ) was set up to fund innovative activities aimed at excluded young people, implemented on a small scale and subject to
robust evaluation. It is overseen by the Mission d’animation du FEJ (FEJ Programme Team – MAFEJ) initially reporting to the Direction de la jeunesse, de l’éducation populaire et de la vie associative (Directorate for young people, popular education and voluntary associations – DJEPVA).

Between 2009 and 2012, the FEJ benefited from an endowment of €230 million from the state and private donors, including €31 million from the Total Foundation and the €2 million from the Union des industries et des métiers de la métallurgie (Union of Metallurgy Industry and Trades – UIMM). Over this period, each experiment was the result of a partnership between a structure acting as project lead and an independent evaluator. As such, the aim was to find out the effects of the projects and steer public policies towards the most effective interventions. Between 2009 and 2014, over 700 projects with an average duration of two years, were initiated, reaching over 500 000 young people.

The FEJ encouraged project leads and evaluators to conduct RCTs and quasi-experimental studies. Is Guide méthodologique pour l’évaluation des expérimentations sociales (Methodological Guide for the Evaluation of Social Experiments) states that in order to “measure the effects of a new programme, ideally the situation of the programme beneficiaries should be observed and compared with what would have happened for the same beneficiaries in the absence of the programme.” As such, the FEJ placed an emphasis on counterfactual evaluations based on a random assignment.

Since its creation, around 23 evaluations have been carried out using these experimental and quasi-experimental methods – an unprecedented output for social policies in France. However, these 23 projects were just a small minority of some 700 trials funded, a tiny proportion given the initial ambition. Jean Bérard, Alexia Prétari and Mathieu Valdenaire offer two factors to explain the difficulties met: “these methods were not suitable for all of the projects and
not all teams were in a position to implement them”[8]. The Centre de recherche pour l’étude et l’observation des conditions de vie (Research Centre for Studying and Observing Living Conditions – Credoc) offers a more in-depth analysis of the difficulties encountered in implementing randomised evaluations for the FEJ[8]. This assessment cites three key barriers: the refusal of project leads on ethical grounds, insufficient numbers in some cases and the risk of contagion between the test and control groups[8].

The remit, funding and management of the FEJ have evolved since 2012. Firstly, in 2012 there was a temporary pause in “programme 163” grants aimed at funding FEJ calls for projects. Secondly, the FEJ Programme Team (MAFEJ) was commissioned to manage the programme “La France S’engage” (France gets involved) an initiative launched by the French President on 24 June 2014. Finally, since 1 January 2016, the MAFEJ has reported to the Institut national de la jeunesse et de l’éducation populaire (National Institute for Young People and Popular Education – INJEP), itself reporting to the Directeur de la jeunesse, de l’éducation populaire et de la vie associative (Director for young people, popular education and voluntary associations – DJEPVA). The INJEP has made evidence synthesis and dissemination a priority with the aim of better sharing FEJ findings among stakeholders in the youth sector. The recent publication De l’éducation à l’insertion: dix résultats du Fonds d’expérimentation pour la jeunesse (From Education to Employability: ten findings from the Fonds d’expérimentation pour la jeunesse) is an example of this dissemination work.

The experience of the FEJ and how it has evolved tells us a great deal about the conditions required to successfully implement robust evaluations and integrate them into the policy process in France.

---

[8] The risk that the test and control groups influence each other’s behaviour for example.
An ongoing movement but with relatively low output of few RCTs and quasi-experimental studies

Currently, although certain initiatives have petered out, the movement around trials and social experimentation continues.

The emergence of new stakeholders has led to these approaches being shared once more. These stakeholders include in particular those involved in conducting RCTs and quasi-experimental studies, such as the CNRS’s research federation CNRS “Travail, Emploi, Politiques publiques” (Work, Employment and Public Policies – TEPP) and J-Pal Europe. Such methods are at the heart of two research centres created in recent years: the Institut des politiques publiques (Institute of Public Policies – IPP) at the École d’économie de Paris and the Laboratoire interdisciplinaire de l’évaluation des politiques publiques (Liepp) at Sciences Po (see above).

Finally, Ansa has emerged as a “trial coordinator” responsible for supporting and implementing trials as well as acting as an intermediary between project leads, evaluators and funders. The large-scale trial of Parler Bambin (Speak Toddler), led by Ansa and funded by La France s’engage, is a good illustration of the role of this coordinator.

**ANSA, OPERATIONAL LEAD OF THE NATIONAL TRIAL, PARLER BAMBIN**

*Parler Bambin* is a trial intended to implement Parler Bambin on a large scale and evaluate its effects through a randomised control evaluation. *Parler Bambin* is based on enriching linguistic interactions between the adult and young child (3 to 36 months) through habits that promote the development of language and the pleasure of communicating in nurseries (“crèches”). The operational strand led by Ansa, involves coordinating the programme with all stakeholders, providing a new framework for disseminating *Parler Bambin*, structuring training provision and managing a national network of practitioners from 2015 to 2020. Ansa works with an alliance of pioneering towns (Courcouronnes, Le Havre and Lille) and the initiators of the project in Grenoble.
British What Works Centres: what lessons for evidence-based policy in France?

The research-action programme is led by an interdisciplinary research team: economists from the J-Pal École d’Économie de Paris (Marc Gurgand) and language learning specialists the Laboratoire dynamique du langage (Sophie Kern). This Randomised control trial (RCT) is based on comparing a test group and control group of crèches, selected according to the date on which they entered the programme. Ultimately it will be used to measure the effects of Parler Bambin on the overall development of children, parental involvement and professional practice. Thanks to the support of La France s’engage, the programme covers 96 crèches nationally.

Some of these stakeholders implement RCTs and quasi-experimental studies and also promote these methods, like the SPARK programme.

**SOCIAL POLICY ANALYSIS FOR ROBUST KNOWLEDGE (SPARK)**

Launched in 2014, SPARK is a programme aimed at raising awareness among public decision-makers of RCTs and quasi-experimental studies. Funded by the European Commission, it was led by the London School of Economics in partnership with J-Pal Europe and the Alliance for Useful Evidence. In particular, SPARK led to the production of a social experimentation guide aimed at civil servants\textsuperscript{\textlxxvi}.

The desire to evaluate the impact of programmes or interventions led to an increased role for the state, particularly in the statistical services within government departments. On the one hand, ad hoc teams can be set up to carry out such evaluations. As such, the Dares set up a project team to carry out a quantitative analysis of the Garantie jeune (the Youth Guarantee). On the other hand, in some cases permanent teams are set up to conduct evaluations, such as the Bureau de l’évaluation des actions éducatives et des expérimentations (the Bureau of educational action and experimentation) at the Direction...
The popularisation of social experimentation has favoured the emergence of the “experimental reflex” within government and a section of civil society. Thus, a number of new trials have been launched in the last five years, such as *Un chez soi d’abord* (A home first), the *Garantie jeune* (the Youth Guarantee), *Territoires zéro chômage de longue durée* (Zero long term employment areas) and *Lits d’accueil médicalisés* (Beds for homeless people with severe health problems – LAMS). Yet, these ‘trials’ have rarely been subject to RCTs with a test and control group (with the exception of *Un chez soi d’abord*). Indeed, these projects have mostly been subject to a “before/after” or quasi-experimental evaluation: methods that are useful for certain types of programme but are less conductive to establishing a causal link between the programme and the effects observed.

**EXPERIMENTING TO TACKLE POVERTY: UN CHEZ SOI-D’ABORD**

*Un chez-soi d’abord* is a trial conducted on four sites – Lille, Marseille, Paris, Toulouse – aimed at providing access to housing for some 400 homeless people with mental health problems. This project takes its inspiration from the Housing First trials conducted in North America and in other European countries, which showed that around 90% of people remained in housing after 12 months. Four years after taking up housing, 85% of *Un chez-soi d’abord* beneficiaries are still there according to the last review of the research, which compares the experiences of over 350 beneficiaries with a control group. The *Délégation interministérielle à l’hébergement et à l’accès au logement* (Cross-government Delegation for Accommodation and access to Housing – Dihal), which coordinates trials at the national level, has embedded the programme in the law (Social Action Code) and plans on rolling it out in 16 additional areas by 2022.
Finally, in addition to these small-scale trials, some public stakeholders are now using quasi-experimental methods to evaluate the impact of certain reforms, such as the evaluation of the crédit d’impôt pour la compétitivité et l’emploi (tax credit for competitiveness and employment – CICE) under the auspices of France Stratégie. However, this practice remains under-developed in France today.

**EVALUATING REFORMS**

The example of evaluating the impact of the crédit d’impôt pour la compétitivité et l’emploi (tax credit for competitiveness and employment – CICE) under the auspices of France Stratégie.

The monitoring committee for the crédit d’impôt pour la compétitivité et l’emploi (CICE), the remit of which has been extended to include the monitoring of business grants and assistance, sits within France Stratégie. It publishes annual reports. The committee’s 2016 report assessed the effect of CICE on the basis of a retrospective analysis of the individual behaviour of beneficiary companies in 2013 and 2014. Evaluations of the CICE were conducted by three independent academic research teams working in parallel using different methodologies. The teams attempted to neutralise effects relating to the economic climate and those relating to the political economy so as to isolate the pure effects of the CICE.

The report concluded that this programme, at a cost of around €20 billion, had the following effects:

- a probable effect on the survival of businesses, which requires more detailed evaluation;
- after the second year, no observable effect on investment, research and development and exports;
- in terms of jobs and wages, differences in evaluation. One team (TEPP) concluded that between 45 000 and 115 000 jobs had been created or safeguarded, where as the other team (Liepp) does not suggest a positive effect on employment or per capita wages.
RCTs and quasi-experimental studies as decision-making tools

Trials and quasi-experimental studies sometimes inform policy decisions

Following a number of trials conducted in France, policymakers have taken the findings on-board, showing that it is possible to incorporate evidence in policy-making. Thus, the limitations of a benefit for young people, the Revenu Contractualisé d’Autonomie (RCA), which became clear in the light of its evaluation, were taken into account when modelling the Garantie Jeune. Another trial compared the impact on jobseekers of support by private providers with the programme Cap Vers l’Entreprise. This trial informed public debates on whether or not there is added value in using private providers to deliver the public employment service.

WHEN EXPERIMENTATION INFORMS DECISION-MAKING: FROM THE RCA TO THE GARANTIE JEUNE

Young people with few qualifications are finding it increasingly difficult to join the labour market, despite the large number of programmes aimed at them. Following the lack of results of the Contrat d’insertion dans la vie sociale (Social integration contract – CIVIS) in 2005, the Ministère de la Jeunesse (Youth Ministry) trialled the Revenu Contractualisé d’Autonomie (Contractual independence allowance – RCA) from 2011 to 2013. The innovative aspect of the RCA lies in the possibility of receiving the benefit alongside any income from work as a continuation of the fixed term contract, whatever the young person’s employment status.

A controlled trial was carried out, comparing the experiences of 2500 young people having benefited from the RCA and those of 2500 CIVIS beneficiaries. The employment circumstances of RCA and CIVIS beneficiaries 1 and 2 years after entering the programme were comparable. Two explanations were put forward to explain these
negative results in the case of the RCA: on the one hand, the benefit was too low to have an impact on their outcomes (€250 maximum); on the other hand, the advice and support was inadequate in terms of its content and intensity.

The findings from the RCA were taken into account when designing the Garantie jeune. The latter offered a higher allowance of €450 per month with structured support in accessing employment and training, including regular work placements\textsuperscript{\textsuperscript{\textdegree}}. These changes bore fruit in light of the initial evaluation findings for the first cohort – 14 months after entering the programme, some 40.4% of beneficiaries were in employment, compared with 34.1% of the control group\textsuperscript{\textsuperscript{\textdegree}}.

In some cases, French stakeholders have been inspired by trials carried out in other OECD countries. Programmes that have been shown to work overseas, have been implemented in France in the form of pilot projects or trials, to evaluate whether the intervention can be replicated in the French context.

**TRIALLING FOREIGN PROGRAMMES: FOCUS ON 3 INTERVENTIONS IMPORTED INTO FRANCE**

**Anti-smoking measures**

The campaign “Moi(s) Sans Tabac” (My tobacco-free month), first implemented in November 2016 and led by Santé publique France, was inspired by the British campaign “Stoptober”. Studies have shown that after 28 days, the chances of stopping smoking for good increase five-fold. “Stoptober” has been shown to be effective with a clear increase in the number of attempts at stopping associated with the campaign (50% more than in other months of the year).
Reducing inequality from early childhood

“Jeux d’enfants” (Child’s play) is an early intervention programme aimed at under-fives and intended to improve children’s life chances from a young age. This project arose directly out of the American project Carolina Abecedarian, which was the subject of a controlled evaluation showing it to be effective in supporting the cognitive, psychological, emotional and social development of the young child over a 30 year period. The programme has been the subject of a pilot project conducted by the city of Lille with the support of Ansa.

Access to employment for people suffering from mental health problems

Working First 13 is an innovative programme based on the Individual Placement and Support methodology developed in the United States. It is aimed at those who are farthest from the labour market in Marseille.

Systematic reviews are not well established in social policy in France

Systematic reviews are a method of comparing the relative impacts of interventions in a specific policy area (for example, support for jobseekers). It aims to distinguish those that work from those that show promise and those with a neutral or negative effect. In France, systematic reviews are already used in the field of health, particularly within the Haute autorité de santé (National Health Authority – HAS) and the Institut national de la santé et de la recherche médicale (National Institute of Health and Research – Inserm). In social policy, this method is also used in certain consensus conferences (see page 38), which are based on a synthesis of existing evidence. In the field of justice, for example, the consensus conference on preventing re-offending drew on systematic reviews. In education, the Conseil national d’évaluation du système scolaire (National Conference for the Evaluation of the School System – Cnesco) has also organised four consensus conferences since its creation, including one on retakes. These examples and the findings they showcased are evidence of the extent to which such initiatives can improve the quality of political decisions.
SYSTEMATIC REVIEWS: PUPIL RETAKES

In early 2015, the Cnesco and the Institut français de l’éducation (French Education Institute – Ifé) organised the consensus conference “Lutter contre les difficultés scolaires: le redoublement et ses alternatives” (Tackling difficulties at school: repeating a school year and alternatives). Whilst, since the 1970s, a number of countries have been united in promoting a policy automatically moving up to the next academic year, retaking a year remains a common practice in France. A systematic review of literature showed that retakes, at best have no effects, but above all may prove detrimental both in terms of pupil attainment and developing their self-esteem. Drawing, in part, on the findings of this systematic review, the panel thus concluded that there should be a “reduction in retakes and the implementation of more effective solutions”.

Public services on the whole little informed by robust evidence

Policymakers little informed by RCTs and quasi-experimental studies

Policy-making draws on different sources of analysis when drawing up new measures and laws. It is important to distinguish between two very different ways in which ‘evidence’ can thus be embedded in policy-making: the will of policymakers and a legal duty.

On the one hand, in developing new measures or laws, policymakers can call on advisors, public bodies and research centres in order to gain insight. Indeed, such bodies share the research they carry out as part of their programme of work (usually set on an annual basis). These bodies generally base their recommendations on in-depth analysis of the needs of individuals and current public service provision.
However, these stakeholders do not regularly draw on systematic reviews of RCTs and quasi-experimental studies. Finally, even where ex ante evaluations are carried out prior to the implementation of a measure, their results are not systematically compared with an ex post evaluation as a means of providing more robust evidence of the actual impact.

On the other hand, with exceptions under the Loi organique n°2009-403 of 15 April 2009, all bills must be accompanied by an impact assessment carried out by the government. These assessments, although different from evaluations, set out the objectives of, and reasons behind, the new law, the current state of the law in that area, how the bill fits into European law, the evaluation of the economic, financial, social and environmental provisions of the bill and the procedure for implementation and consequences thereof. These impact assessments are usually conducted to very tight deadlines by central government without the support of researchers specialising in evaluation and without taking into account the results of available RCTs and quasi-experimental studies.

With this in mind, the Commission d’étude des effets de la loi pour la croissance et l’activité within France Stratégie put together an original initiative aimed at reporting the state of the art in terms of evidence and informing the public and legislative debate.

AN EX ANTE EVALUATION BASED ON A REVIEW OF EVALUATIONS
THE COMMISSION D’ÉTUDE DES EFFETS DE LA LOI POUR LA CROISSANCE ET L’ACTIVITÉ WHICH SITS WITHIN FRANCE STRATÉGIE

The aim of the Loi pour la croissance et l’activité ("Law for growth and work", the so-called "Macron law") was to establish the conditions for a recovery in the economy and labour market in France. An independent evaluation committee was set up to carry out an ex ante evaluation of the impact of different strands of this law. This committee was made up of French researchers who were experts in the areas covered by the law. It analysed the effects on the economy and employment of laws of similar measures implemented in other countries. To this end,
it analysed available evidence on the five themes of the bill: Sunday working hours, mobility, reform of industrial tribunals, regulated legal professions and commercial real estate. In January 2015 the committee published five briefing papers reporting on the findings of their reviews aimed at informing public debate.

“Effective” practice is difficult for practitioners to access

Access to current research is a key issue for evidence-based policies. In France, initiatives aimed at promoting the dissemination of best practice to practitioners through guides or libraries, such as the “experience bank” of the Union nationale des centres communaux d’action sociale (National Union of Municipal Social Centres – Unccas)\(^{lxxxix}\). These practices are chosen when they seem to offer appropriate responses either to a legislative change or to user needs, or demand on the part of practitioners. However, these practices have rarely been subject to an RCT or quasi-experimental study. More generally, there are very few outputs aimed at sharing the findings of what “works” with practitioners in the social field. As such, practitioners lack the tools to enable them to base their practice on ‘evidence’.

Scale-up support is under-developed

In France, the idea that innovations that have been trialled on a small scale can be disseminated is far from new. Indeed, public services in France have often been inspired by local innovations. The system of Centres de Protection maternelle et infantile (Mother and Infant Protection Centres – PMI) established in 1945, for example, was largely inspired by the Association alsacienne et lorraine de puériculture (the Alsace and Lorraine Childcare Association) established in 1920 by the paediatrician Paul Rohmer. However, the idea that the scaling up of innovations requires support is more recent. This idea is at the heart of the French model of social experimentation. Indeed, social experimentation is generally defined as “a method used to evaluate social innovations on a small scale before
considering scaling them up”… Some pilot programmes, such as the Parents’ toolkit have, indeed, had convincing results and have been scaled up.

**FROM EXPERIMENTATION TO ROLL-OUT: THE PARENTS’ TOOLKIT**

The Parents’ Toolkit is aimed at establishing fruitful discussion between schools and families in order to do more to involve parents in their children’s education. It contains tools to be used by teaching staff to facilitate discussion with families at workshops. A randomised control trial showed that it has a number of positive effects against certain criteria: “pupils more motivated, a reduction in school violence and truancy”. The programme was then disseminated in a structured way and since 2010 has been gradually rolled out, reaching about one quarter of French lower secondary schools to date.

In the last few years, the practice of supporting scaling up has grown in France. Several programmes have emerged, such as Ansa’s Accélérateur d’innovation sociale (Social Innovation Accelerator), the support for scaling up proposed by Avise and the Antropia programme launched by the Essec. Finally, since 2014 the label La France s’engage has funded and supported social innovations in scaling up.

**LA FRANCE S’ENGAGE**

La France s’engage is an initiative led by the French President since 2014. It is intended to promote and facilitate the roll-out of socially innovative initiatives in the general interest, led by individuals, charitable organisations, foundations and businesses. The aim is to promote the emergence of new types of intervention that meet the needs of French society, through a continuous call for projects. Since 2014, 62 projects have been accredited and benefited from funding and support.
The key findings on evaluation and evidence in France

The culture of evaluation in France is marked to a large degree by the evaluation of public policies. Despite the undeniable enthusiasm for this approach, it has yet to become established, both in terms of its tools and methodology and from an institutional perspective. The sphere of evaluation brings together competing, or at the very least divergent, methodological approaches that commentators have termed different “schools of thought”. Most evaluation remains under used by decision-makers and practitioners with the wider public remaining unfamiliar with them.

From 2007, there has been a sharp increase in the use of RCTs and quasi-experimental studies, in particular due to the movement around social experimentation. This has resulted in unprecedented production of such studies, particularly around issues of employability and young people. Since 2012, this trend has continued at a more modest scale.

Despite the stated interest in these robust methods, few RCTs and quasi-experimental studies have been conducted in the social sector in France in recent years. This method is often seen as complex, slow and costly, both by public authorities and frontline professionals. Evaluators have come up against barriers to implementation as illustrated by the experience of FEJ. This has led some stakeholders to question its technical feasibility and as such to favour other evaluation methods. The difficulties encountered demonstrate the inherent complexity of these methods, but also a lack of expertise and experience of these methods on the part of both evaluators and project leads.

The use of other methods associated with evidence-based policies – the synthesis and dissemination of evidence – remains extremely limited in the social, economic and educational spheres in in France. The research centres and teams that we interviewed generally do not conduct them. As such, this method which is practically non-existent, requires further development.
Interventions with positive impacts are rarely scaled up in a structured way. As an example, only a small number of the many social experiments funded by the FEJ were properly rolled out. Indeed, as highlighted in a recent report of the SGMPA and the CGET, “the public sphere is not currently organised in such a way as to facilitate the scaling up of these initiatives”\textsuperscript{xxiii}. Finally, these scale up programmes generally struggle to identify those innovations that have been the subject of an RCT or quasi-experimental study. They tend to roll out “innovative” or “promising” initiatives but rarely those that have shown themselves to be “effective”.

\textsuperscript{ixix} SPENLEHAUER Vincent, “La (f)utilité gouvernementale de l’évaluation des politiques publiques, quelques leçons américaines et françaises”.
\textsuperscript{lx} JACOB Steeve, VARONE Frédéric et GENARD, Jean-Louis, (Dir.) [2007], L’évaluation des politiques au niveau régional, Bruxelles, Peter Lang, Coll. Action publique.
\textsuperscript{lxxi} For more informations on the EPPs: www.modernisation.gouv.fr/laction-publique-se-transforme/en-evaluant-ses-politiques-publiques.
\textsuperscript{lxv} La Vigie de l’Évaluation, February 2014: www.evaluationvigie.fr/les-quatre-chapelles-de-levaluation.
\textsuperscript{lxvi} On this point we can also refer to the findings in the survey of public policy evaluation carried out every two years by the SFE with the support of the SGMAP which produces quantitative and qualitative analysis of references of evaluations catalogued since 2007 in the Observatoire de l’évaluation des politiques publiques (OEPP). In 2016, almost 1 800 references were thus catalogued in the OEPP. See online: www.sfe-asso.fr/ressources-sfe/barometre-evaluation.
\textsuperscript{[ii]} ALDEGHIS Isa, JAUNEAU Pauline, MULLER Lara, Bilan méthodologique des évaluations conduites par le CRÉDOC dans le cadre du Fonds d’expérimentation pour la Jeunesse, CRÉDOC, December 2012.
\textsuperscript{lxvii} European Commission, Guide to Testing Social Policy Innovation, June 2014.
\textsuperscript{lxviii} Décret n°2016-1940 dated 28 December 2016 relatif aux dispositifs d’appartements de coordination thérapeutique “Un chez soi d’abord”.
\textsuperscript{lxx} BRUN-SCHAMME Amandine, DESPLATZ Rozenn, NABOULET Antoine, Rapport 2016 Comité de suivi du Crédit d’impôt pour la compétitivité et l’emploi, France Stratégie, September 2016.
\textsuperscript{lxxi} Décret n°2011-128 dated 31 January 2011 relatif à l’expérimentation d’un revenu contractualisé d’autonomie.
British What Works Centres: what lessons for evidence-based policy in France?

lxxxi www.gouvernement.fr/action/la-garantie-jeunes.
lxxv Prime Minister circular dated 15 April 2009.
lxxviii The briefing notes can be found here: www.strategie.gouv.fr/presse/communiques-de-presse/installation-de-commission-evaluation-de-loi-croissance-lactivite.
lxxix www.unccas.org/-banque-d-experience-
xc LE BARS Sylvie, PRIGENT Marion, RICKEY Ben et AITOUT Claire-Selma, L’expérimentation sociale à l’épreuve du terrain, March 2014.
What Works Centres: what are the lessons for France?

In this last chapter, we set out lessons that can be learned from the What Works Centres. Based on an overview of the practices of these centres, to what extent can they be a source of inspiration for France? The study highlights the benefits and some limitations of the approach embodied by the What Works Centres.

What Works Centres: Methods

Benefits

The What Works Centres use traditional methods within an evidence-based policy approach: systematic literature review, knowledge transfer through toolkits, and trials.

What sets their approach apart is the ambition to apply these methods across a whole public policy area. Thus, these centres have tested these methods on a large scale, finding their limitations and innovating to adapt them to social policies in the British context. We see the following practices developed by the WWC as worthwhile and a potential source of inspiration in France:

- **Major systematic literature review programmes.** By conducting ambitious review programmes around the effectiveness of a range of interventions, the centres have sought to offer decision-makers and practitioners useful answers to the questions they have. The difficulties encountered in producing these reviews led a number of centres to adopt a more nuanced approach, for example in better taking into account the issue of transferability of the interventions identified to the British context, or in changing the criteria for selecting the evaluations to be taken into account;

- **Innovations in the methods of dissemination.** The WWCs have turned hundreds of evaluations into easy-to-read findings for practitioners and
public decision-makers, particularly through the intervention libraries. Here, the innovation lies in combining systematic reviews conducted to a high scientific standard, with findings presented in a synthetic, easy-to-understand and easy-to-use way. The WWCs are nevertheless aware that it is not enough to publish these materials in order for stakeholders to make use of them. As such, the centres have increased the number of dissemination activities through events, training and dissemination programmes.

- **The development of new capabilities in conducting experiments.**
  Another advantage of their approach lies in the fact that they test new approaches, including through random trials. In the cases of the Education Endowment Foundation (EEF) and the WWC for Crime Reduction, the ability to support experiments has been a real asset, bringing them closer to frontline stakeholders. Indeed, the former funded a series of educational trials, unprecedented in England in terms of their scale. 127 evaluations, of which 105 using randomised methods, between 2011 and 2016. In doing so they have developed an expertise in the implementation of randomised trials, particularly in terms of correcting the bias which, as we have seen, can result in their failure - and have contributed to the development of a broader panel of researchers with the expertise to conduct such trials – a condition for successfully developing high quality randomised trials. Evaluating the impact of one hundred innovations has been a real asset for the Education Endowment Foundation, in terms of enriching knowledge but also bringing the centre closer to stakeholders in education. As such, the expertise developed by the EEF represents, in itself, one benefit of the approach for French researchers seeking to conduct trials.

**Challenges**

The What Works Centres transferred to the social sphere, evidence-based approaches that had initially been developed in the medical field.
This raised questions and difficulties that include:

- **The production of systematic reviews has in some cases been difficult due to the lack of evidence of a high enough standard, the complexity in determining the effectiveness of an intervention and in clarifying the contexts in which an intervention works.** Some centres have adopted a more nuanced approach, including “promising” interventions or interesting leads. Other centres have been more “purist” in their focus on experimental and quasi-experimental evaluations, in spite of the challenges they met (such as local economic growth). Some stakeholders continue to criticise the idea that RCTs and quasi-experimental studies is at the top of the evidence hierarchy, and should be the main type of “evidence” to be used by decision-makers and practitioners. Most centres tend to promote the use of RCTs through training, awareness-raising and support, with mixed results;

- **Most of the centres have yet to demonstrate a measurable impact on professional practice.** The lack of long term data in terms of the impact of these relatively new organisations is one explanation. Yet, there remain uncertainties about the effectiveness of the approach which involves identifying “What Works” in general in OECD countries then disseminating it in the United Kingdom. In particular, there are questions around the transferability of the practices shown to be effective (“external validity”), and the consequences of scaling them up (“equilibrium effects”). Conducting a trial in the United Kingdom is the preferred way of ensuring an intervention is transferable.
The What Works Centres as organisations

**Benefits**

The other benefit of the What Works Centres lies in their existence as a programme that is:

- **part of a long process** of transforming professional practice. Some centres have long-term endowments that are consistent with this remit, but most only have three-year funding cycles which is a source of fragility;
- **independent** as not part of central government, thus mitigating the impact of changes in ministerial approach in their area. However, most are close to government departments as they are in large part dependent on public funding. Two of them are even public bodies;
- **sectoral**, covering a field that is not restricted by the silos that characterise public policy. This approach seems to be particularly effective when the sector corresponds to a clear body of professionals (for example teachers in the case of the EEF).

Assigning specific organisations with a “What Works” remit presents a number of benefits:

- **a “one-stop shop” for evidence for practitioners.**
  As a programme, the centres have the remit, resources and independence to become a one-stop shop providing frontline professionals with access to “What Works”. In an information-heavy world, this meets a key requirement for practitioners to quickly access advice that is both clear and evidence-based. By having a long-term presence, the centres can become the reference in their area (for example NICE for healthcare practitioners). In addition, by updating their libraries, the centres aim to provide up to date research;
- **a common thread in supporting changes in practice over the long-term.**
  The centres aim to influence the practice of thousands of practitioners and decision-makers. This involves supporting stakeholders in adopting promising or effective practices, which often requires intensive dissemination work. Moreover, in order for practitioners and decision-makers to adopt such practices, they must themselves be convinced of the benefit of basing their activities on evidence, which means a real cultural shift. Setting up centres
is a way of ensuring that these changes are supported over the long term;
- a “missing link” between decision-makers, academics and practitioners. The centres aim to connect academics, decision-makers and practitioners within a single “evidence ecosystem”. Thus, they can promote a new dialogue between these stakeholders on the role of evidence in public policy. This dialogue is enhanced when the centres have the means of trialling the innovations put forward by frontline stakeholders or when they set up participatory bodies such as NICE with practitioners, frontline professionals, academics and service users.

Challenges

The “What Works” model is inspired by evidence-based medicine and particularly NICE, which is based on the dissemination of “evidence” in a specific area (health) to a specific body of health professionals (doctors, nurses etc.). The centres have reproduced this model in areas where these conditions are not often met:
- some centres have been set up for areas that were not clearly identified or with an unclear scope such as early intervention, well-being and ageing well. As such, these centres have had to not only specify the scope of their policy area, in consultation with stakeholders, but also to convince stakeholders of the relevance of that area;
- in most cases, centres do not have a clearly defined professional body. The centres on crime reduction (police) and education (teachers, teaching assistants) are exceptions, although the potential target of their dissemination extends far beyond this type of professional. The other centres have had to use empirical methods to establish a network of professionals that may be interested in their work. This exercise is made easier where there is a core group of professionals who are particularly affected by the subjects, such as economic development officers in the case of the WWC for Local Economic Growth. The task becomes infinitely more difficult where the scope of the policy area is vague, such as well-being and, to a lesser degree, ageing well.
These centres have taken a pragmatic approach to overcoming these problems of scope and target. However, the need to define their scope is evidence of the fragility of what they are trying to do. Indeed, it is difficult to establish a scope and a group of professionals from scratch, although grappling with cross-cutting issues can be an objective in to increase multi-disciplinarity and provide a better response to it. This issue would require vigilance if the “What Works” centre were to be developed in certain policy areas in France.

**Does the French context lend itself to the “What Works” model?**

Analysis of the French context in view of the “What Works” approach has confirmed the initial hypothesis of a limited reliance on evidence-based approaches: RCTs, systematic reviews, dissemination of findings to practitioners and decision-makers. In this respect, the French context shares many commonalities with the British context (particularly prior to the launch of the WWCs).

Yet, from a cultural perspective, the idea that it is desirable and possible to base policies on evidence is more contested in France, particularly in the social end education fields. Advocates for this approach (particularly the many stakeholders involved in trials) appear to be in the minority in the scientific community as well as among decision-makers and frontline professionals.

From the point of view of organising public action, there are differences between the French and British contexts. On a number of occasions, the French stakeholders interviewed cited a greater degree of decentralisation in the United Kingdom and greater autonomy of stakeholders at the local level (for example in the educational system) as key factors in the success of the What Works Centres. One fear is that the degree of centralisation in France means that the centres are of little or no relevance. Without purporting to produce a comparative analysis of the organisation
of public policy in the two countries, it is important to address this subject. Analysis that contrasts a “centralised” France with a “decentralised” United Kingdom is based on a simplified view of the situation. This analysis does not take into account the decentralisation of a number of policy areas in France since the 1980s or the high degree of centralisation of certain policy areas in the United Kingdom (the health system for example).

The differences in the structures of public action in the two countries cannot alone represent a barrier to transferring the ‘What Works’ approach to France. However, it is clear that the structure of public action in a specific policy area has consequences for the positioning and procedures of any What Works Centre. The study tends to show that:

- where a WWC operates in a policy area that is relatively decentralised, this can represent a justification for its existence. Indeed, where policies are decentralised, it is all the more important to inform the decisions of local elected representatives, officials and practitioners. Where these stakeholders have more room for manoeuvre, the question of their ability to make decisions informed by evidence arises, given that they generally have limited time and resources. For example, in England, school heads have been granted a devolved budget for educational measures to help lower the attainment gap for disadvantaged pupils (the “pupil premium”). With this in mind, the Education Endowment Foundation’s Toolkit includes key information on the most effective measures for such pupils;

- however, the WWCs can also operate in more “centralised” policy areas. NICE, the original WWC, operates in the health sector with a public health service managed in a highly centralised way. Far from being a barrier to the success of NICE, this centralisation has made it possible to mainstream NICE’s recommendations in the health system in a more systematic way through, for example, auditing processes. It is important to remember that, even where there is a strong national policy framework, frontline officials and practitioners almost always have room for manoeuvre.
Proposals

Taking inspiration from the What Works Centres and drawing on the survey of French stakeholders, Ansa and its partners make two proposals. They are intended to support the emergence of evidence-based policies and practices in France by strengthening the link between evidence and public policy.

Proposal 1: Explore the opportunity of labelling or creating “What Works” inspired initiatives or structures in France

Like the British What Works Centres, these structures would be a means of strengthening the role of “evidence” in decision-making and public policy-making, in a specific policy area. Such teams or structures would act as one-stop shops to evidence for policymakers and frontline professionals.

This study contains recommendations around the possible remit that such organisations may have in France as well as a number of conditions for success. Three organisational models are proposed in order to embed these ‘centres’ in the French institutional landscape.

Main missions

- Produce systematic reviews. Systematic reviews must be carried out robustly, using recognised and transparent methods and classification criteria, and to a high standard (see for example the PRISMA method\(^9\)). They contribute both to advancing research and making such research accessible by systematically publishing their work. Each “centre” is intended to become a “one-stop shop” through which to access the RCTs and quasi-experimental studies produced in a given policy area. This remit involves a substantial amount of work, drawing on non-negligible human

\(^9\) PRISMA for preferred reporting items for systematic reviews and meta-analysis (www.prisma-statement.org).
and financial resources: concentrating resources on one stakeholder or pooling the resources of several stakeholders are possible approaches to ensuring that work is to a high-standard.

- because, in addition to the systematic reviews it may be necessary to include more descriptive or qualitative types of evidence to inform practices and decision-making, these centres should act more broadly as “resource centres” for the evaluation of programmes and the identification of interventions that “work” for a given issue;

- disseminate syntheses, promising or proven interventions and recommendations in a suitable format that meets the expectations and needs of national and local policymakers as well as frontline professionals. The ‘centres’ should support stakeholders in implementing practices proven to be effective through dissemination activities as well as play a role in disseminating the methods and results.

Additional mission

- Experimentation capability. Over the last ten years, there has been a marked increase in the number of RCTs and quasi-experimental studies in France, particularly as part of trials. However, the number of such studies conducted in France remains low, making evidence-based policy-making difficult. Moreover, in every area of public policy, there are often dozens of promising interventions that are little known beyond the specific geographical area and which may lend themselves to such evaluation. We therefore recommend that these structures are given the means to fund both intervention and robust evaluations, drawing inspiration from the practices of the Education Endowment Foundation. On the one hand, this remit would be a means of building the evidence base of “What Works” in the French context and, on the other, of bringing the centre closer to the frontline, in turn increasing the relevance of the centre’s outputs.
**British What Works Centres: what lessons for evidence-based policy in France?**

**THE KEY FACTORS OF SUCCESS, FROM AN ANALYSIS OF THE WHAT WORKS CENTRES:**

- a long-term presence meaning access to funding over at least three years\(^{10}\);
- substantial resources to deliver the core and additional activities proposed\(^{11}\);
- independence in setting their programme of work and publishing their work; this independence can be achieved through the organisation’s charter, funding arrangements and governance.
- transparency to be achieved by the systematic publication of its reviews and evaluations, specifying the methods used in its work;
- requirements of decision-makers and practitioners to be placed at the heart of its activities by involving them in planning and implementing systematic reviews, dissemination activities, scale-up and trials.
- high standards through the use of evaluation classification to distinguish between “types of evidence” and “standards of evidence”;
- particular attention to be paid to the transferability of interventions having been shown to be effective in other countries to the French context.

**Organisational structure of a centre**
The most appropriate organisational structure for a body inspired by the What Works Centres will depend on several factors specific to the public policy area in which it operates:

- the organisation of public action, its degree of centralisation, the institutional landscape, and the professionals in question;
- existing practice in terms of evidence-based policies in France, whether in relation to statistics, evaluation, more qualitative work or the dissemination of evidence to decision-makers and practitioners;

---

\(^{10}\) The financial security of the British What Works Centres: gives them the visibility to begin studies and they have been able to develop a substantial base of systematic literature reviews in the first three years of their existence.

\(^{11}\) The British What Works Centres: have a budget of around £1 million per year for core activities (literature reviews and dissemination) and the Education Endowment Foundation has a budget of £14 billion a year for all of its activities including trials with controlled evaluation.
– the maturity and quantity of evaluations and systematic reviews. The What Works inspired centres could adopt the three organisational structures set out below.

➤ Model 1

A team “hosted” by central government with the central remit set out above. This team would have dedicated human and financial resources, its own visual identity and a web site, as well as a programme of work resulting in publications accessible by the public.

<table>
<thead>
<tr>
<th>Strengths</th>
<th>Weaknesses</th>
</tr>
</thead>
<tbody>
<tr>
<td>By mobilising existing resources and expertise, this team could be set up relatively quickly. With its close links to national decision-makers, this team would be well placed to identify the “hot topics” for these decision-makers and influence policy-making.</td>
<td>Embedded in central government, these teams will not enjoy the same independence as the What Works Centres in terms of setting their programme of work and publishing their work. These teams would also run the risk of being abolished when government restructuring takes place. With their position close to the centre, they may potentially have greater difficulty in disseminating findings to frontline professionals.</td>
</tr>
</tbody>
</table>
Model 2
An independent structure with a dedicated team pursuing a programme of work developed with its board. This structure could be new or the result of merging organisations that already have “What Works” type remits in the policy area in question.

<table>
<thead>
<tr>
<th>Strengths</th>
<th>Weaknesses</th>
</tr>
</thead>
<tbody>
<tr>
<td>The independence of the structure will allow it to set a programme of work and systematically publish its studies as well as to carry out work in a transparent way. It will have a greater ability to respond to the expectations of frontline professionals. As a third party stakeholder it will potentially be better placed to establish a dialogue between national decision-makers, practitioners and academics.</td>
<td>The structure will potentially have greater difficulty in establishing links with national decision-makers. It may take a relatively long time to set up such a structure and will require substantial funding which will need to be sustained.</td>
</tr>
</tbody>
</table>

Model 3
A “hub-and-spoke network” led by a stakeholder whose main role will be to plan and coordinate the activities of other stakeholders in a given policy area. The centre would be assigned modest human and financial resources and will mobilise the resources of partners around a joint work programme.

<table>
<thead>
<tr>
<th>Strengths</th>
<th>Weaknesses</th>
</tr>
</thead>
<tbody>
<tr>
<td>By mobilising existing stakeholders, the centre will become established in the institutional landscape, without risk of “duplication” with existing stakeholders. It would also be able to benefit from both the expertise and networks of partner stakeholders.</td>
<td>In the absence of substantial resources, its success would depend on the partners’ level of commitment. The centre thus risks becoming a coordination group rather than a centre with its own programme of work.</td>
</tr>
</tbody>
</table>
Proposal 2: Create an alliance around evidence-based policies

Whatever the policy area, stakeholders are faced with the same challenges in developing and implementing evidence-based policy: limited resources, lack of familiarity with evaluation methods, difficulties integrating the long timescales of evaluations into short policy cycles. These challenges are experienced both by policy-makers, but also by frontline professionals and academics. There is currently no resource shared by all stakeholders dedicated to promoting these approaches and facilitating discussions between stakeholders.

Ansa and its partners thus recommend establishing a cross-cutting network that brings together policy-makers, academics and frontline professionals around the issue of evidence-based policy and practice, inspired by the Alliance for Useful Evidence in the United Kingdom. This alliance could, in particular, help to:

- facilitate discussion on these issues by organising workshops, conferences and “webinars”;
- raise awareness among public decision-makers of evidence-based policy methods by incorporating them in their initial training (Ena, Inet), or continuous development (e.g. CNFPT), or in ad-hoc courses;
- advance thinking around evidence-based policies by publishing practical guides and thinkpieces, for example on “standards of evidence”, with the different types of knowledge and evaluation also to be taken into account.

The knowledge to be taken into account for evidence-based policies is broad and includes, but is not limited to, the findings of impact evaluations. It extends to other resources - ranging from needs analysis, the evaluation of the theoretical basis of the various programmes and policies, qualitative evaluations (based on observations, interviews and case studies), process evaluations, quantitative evaluations (descriptive statistics, correlation) as well as the expertise of citizens or users in terms of their needs and required responses. These other resources, of a more theoretical, descriptive or qualitative nature, are essential particularly in terms of producing assessments (analysis and understanding of needs) and recommendations.
Case studies of 5 What Works Centres
Early Intervention Foundation

At a glance

<table>
<thead>
<tr>
<th>Sector</th>
<th>Intervention for children from conception to 18</th>
</tr>
</thead>
<tbody>
<tr>
<td>Launch date</td>
<td>February 2013</td>
</tr>
<tr>
<td>Organisational type</td>
<td>Charity belonging to the WWC network</td>
</tr>
<tr>
<td>Size of team</td>
<td>16 people approximately</td>
</tr>
<tr>
<td>Annual budget</td>
<td>Not public</td>
</tr>
<tr>
<td>Website</td>
<td><a href="http://www.eif.org.uk">www.eif.org.uk</a></td>
</tr>
</tbody>
</table>

Context

The Early Intervention Foundation (EIF) was set up in February 2013 with an endowment from the Department for Education. It was initially supported by a consortium of organisations including 4Children, The Local Government Association and Achievement for All 3As with 3 years of initial funding from the Department for Education. The centre was set up at Graham Allen, MP’s initiative, following a report he published on early intervention, demonstrating the impact and savings it generated for the public purse\textsuperscript{13}, drawing inspiration from the Washington State Institute for Public Policy in particular. The Dartington Social Research and Michael Little\textsuperscript{14} were consulted on this work. The aim of EIF is to:
- turn this into evidence-based recommendations to share with local public decision-makers, providers and in some cases funders;
- shifting investment to effective early interventions, to not only support children and their families but also create a positive impact on wider society and generate savings.

\textsuperscript{13} grahamallenmp.co.uk/early_intervention.
\textsuperscript{14} dartington.org.uk/about/staff/dr-michael-little
Organisational structure

The centre is a charitable organisation that belongs to the national network of What Works Centres. It is governed by a board of trustees. The centre has a team of 16 people responsible for analysing evaluations (evidence analysts), the economic analysis of interventions, avoided costs as well as the dissemination of recommendations. It draws on an evidence panel made up of academics and experts who offer support. Their role is to provide expert advice on how to fulfil the centre’s mission, which can include evidence syntheses and recommendations. It is currently funded by a group of public stakeholders including government departments and national agencies, research institutes and private donors. Funding is increasingly targeted at projects/themes chosen with the donors.

Activities

Reviews

Objectives

The EIF produces reviews of evaluations of early interventions covering a broad age group - from early years to adolescence. This is a means of determining, based on a set of RCTs and quasi-experimental studies, which interventions have a positive impact, which show promise and finally which could benefit from improvements or further research.

Actions

As part of its work, in July 2016 the EIF published a review of evaluations of interventions on the parental-child relationship\(^{15}\) with a focus on existing programmes in the United Kingdom. This review relates to interventions aimed at improving the parent-child relationship from before birth to the age of 5, in order to contribute to childhood development in terms of attachment, behaviour and in cognitive terms, as well as preventing any future difficulties. Of the 100 interventions identified in a systematic review of

literature published in 2015, 75 were selected for more in-depth study in light of the existing evaluations and the cost of interventions\[16\].

The report presents the aggregate findings by theme to provide an overview then goes on to break down the findings for each programme. The review identified 17 programmes as having had positive and measurable impacts (level 3 or 4) whereas 18 programmes are rated as promising (level 2). Finally, some interventions have been evaluated scientifically but do not show any impact, or have not been evaluated in a robust enough way.

The EIF is working on a number of other themes. For example, the Home Office has commissioned studies on interventions to prevent youth violence and integrate young gang members. As such, the EIF has published a number of reports on the subject, particularly on risk factors, how to identify young people at risk and interventions that may or may not work, as well as what the various evaluations can teach us.

**Methods**

In order to carry out such reviews, the EIF conducts an in-depth review of the existing evaluations for each intervention\[17\]. For each intervention, a systematic review is conducted within the EIF as well as an analysis of the costs of the intervention. These systematic reviews are then studied by external experts, first individually, then brought together in a panel to agree with the EIF team on the first rating, based on the quality of the underlying studies.

In order to ensure that the synthetic reviews reflect current research as closely as possible, these are finally examined by the evidence panel. The academics on the panel draw on their expertise on research design, the quality of measures used, the confidence it is possible to have in the results, to inform the team having conducted the review.

---

\[16\] The study does not include all existing parenting programmes for early years in the United Kingdom as no such database exists.

\[17\] This differs from the meta-analytical approaches according to which indicators for various interventions with the same purpose are aggregated to produce a set of scores for various types of practices. The EIF approach facilitates the comparison between the various interventions on the basis of specific studies.
This work is carried out in conjunction with practitioners of the intervention who may offer additional information or correct misconceptions.

**Evaluating the evaluations**

In the report “Foundations for Life: What Works to support parent child interaction in the early years”, the EIF rates the interventions identified. Each intervention is rated both on the basis of the quality of its evaluation and on whether or not it had a positive impact. Quality criteria are based on the Maryland Scientific Scale, Blueprints, and the NESTA Standards of Evidence. They also compare the estimated cost per child supported based on a method developed by the EIF, so that the costs are comparable. The scale developed for a systematic review will be adapted to the purpose of the study and the experiences of previous reviews will be fed into it.

The EIF evaluate the impact on the basis of the most robust studies: a quasi-experimental study takes precedence over a before-after study with no control group. In order to achieve level 3, for example, the most robust studies must be consistent with each other and improve at least one key outcome (attachment, behaviour, cognitive development) for children. The level can be higher (3+) if the level 2+ studies provide corroborating evidence. Interventions are also rated on whether they can be replicated - level 4 in particular corresponds to evaluations conducted in at least two different contexts. This does not, however, include evaluations designed specifically to test the intervention at scale - as is the case with EEF, for example. Finally, the EIF takes into account the presence of theories of change in the scoring of interventions. Indeed, before carrying out a more robust evaluation, this method is a means of checking the theoretical basis of the intervention, by showing the links between the intervention and the desired outcomes. The EIF is working with other WWCs to develop a quality standard for theories of change and so qualify them.

---

We have developed a way to recognise theories of change appropriately informed by scientific research. We focus on those that are more than a logic model, than program description. The theory of change also has to have a clear basis in the scientific research literature. There should also be clarity about how scientific research evidence has informed the programme’s target population, content and dose.

Dr Kirsten Asmussen, Evidence Analyst, EIF.

Figure 1. A dynamic model of the development of evidence

Effectiveness
Evidence from at least two high-quality evaluations* demonstrating positive impacts across populations and environments lasting a year or longer. This evidence may include adaptations to meet the needs of different target populations.

Efficacy
Evidence from at least one rigorously conducted evaluation* demonstrating a statistically significant positive impact on at least one child outcome.

Preliminary Evidence
Evidence of improving a child outcomes from a study involving at least 20 participants, representing a wide of the sample using validated instruments.

Logic Model
Key elements of the logic model are being confirmed and verified in relation to practice and the underlying scientific evidence. Testing of impact is underway but evidence of impact at Level 2 not yet achieved.

No Effect
A finding of no effect on measured child outcomes in a high-quality impact evaluation. The next step is to return to the verification and confirmation of the logic model.

*High quality evaluations do not need to be randomised controlled trials if a coherent and robust counterfactual can be provided in other ways.

Source: www.eif.org.uk.
Challenges and lessons
The EIF is faced with the complexity of determining beyond dispute whether an intervention subject to evaluation has actually had positive and measurable outcomes.

“Criteria is used to inform decisions about the strength of evidence, but these are rarely clear-cut decisions. I call this “500 shades of grey’, since we need to weigh multiple criteria to determine the extent to which programmes have met the threshold for each of the levels.”

Dr Kirsten Asmussen, Evidence Analyst, EIF.

As such, the EIF considers that it is important to pay attention to interventions in this ‘grey area’ about which studies have not given conclusive results, are not robust enough or have not given the expected outcomes. The EIF thus wants to give programmes a chance to be refined, improved and better designed, particularly by drawing on the findings of the evaluations and then to be subject to a more robust evaluation. Indeed, the EIF hopes that highlighting the projects in need of improvement will encourage the stakeholders in the sector to move these interventions up the EIF scale.

The EIF has also observed that the literature is continuously changing. The production of new more rigorous studies can change how the interventions were rated, sometimes leading to the promising or proven status being revisited. These changes sometimes result in confusion in the minds of the EIF’s target audience.

The Family Nurse Partnership programme, for example, was subject to randomised evaluation a number of times in different contexts in the United States and once in the Netherlands. These are highly qualified nurses who regularly visit young mums in early pregnancy until the child is 2 years old – the young women are generally around 19 years old. It is a programme with a clear target population and the outcomes measured in these studies
are very positive at the time of the child’s birth and in early childhood as well as in the years that follow. As such, the intervention has been extensively rolled out in United Kingdom. A randomised evaluation is under way and initial conclusions relating to the findings of the antenatal visits, are rather disappointing compared with what other studies initially suggested (although they are sufficiently positive as to achieve a 4+ rating).

One explanation is the fact that the monitoring of the health of young, disadvantaged pregnant women in the United Kingdom is of much better quality than in the United States. In the studies conducted, the counterfactual (and as such the point of comparison for evaluating the effect) is therefore different (young women having very little access to healthcare in the United States and young women having more regular access to healthcare in the United Kingdom). These findings have led the programme leads in the United Kingdom to improve the antenatal programme - as such the intervention has been adapted to the UK context rather than simply replicated.

**Trials**

The EIF does not carry out trials itself nor does it have the resources to commission evaluations. In some cases the centre has shared its expertise with local authorities to secure research funding through calls for projects. The EIF’s aims is also that the results of the synthetic reviews inform future programmes of research by research bodies. Of the lessons learned in terms of trials, EIF points out that some commissioners have financed sizeable randomised control trials, but for a variety of reasons, it failed. In particular, excessive attrition leading to insufficiently large sample sizes is a frequent difficulty.

“It is better to start with a good pilot project.”

*Dr Kirsten Asmussen, Evidence Analyst, EIF.*
Dissemination and promoting uptake

Objectives
The EIF’s main activity is to carry out very comprehensive evidence reviews of early interventions, from early childhood to young adulthood. The EIF has also taken initiatives to disseminate the findings. The EIF now puts more of a focus on encouraging stakeholders to embed the scientific results synthesised by the EIF in their decision-making process.

Methods
The uptake of new interventions depends to a great extent on the way public services are structured and delivered in the country in question.

For programmes targeting children, local authorities are the main actors in the United Kingdom as they are responsible for children’s services (child protection and foster care, social work). Although their activity is regulated by law, local authorities have flexibility in the choice of programmes they fund. As such, one of the EIF’s main objectives is to convince them to take up or follow its recommendations. The EIF is continuing to think about the best way of promoting the uptake of effective interventions in decentralised policy areas. One suggestion is for local authorities to be able to choose from a sample group of interventions with proven benefits.

Until now, the EIF has targeted mainly local authorities both in terms of its publications and dissemination activities on the ground. In addition to local authorities they believe it would be worthwhile reinforcing dissemination of the EIF’s findings and recommendations to the following target audiences, through tailored publications and events:
- child protection and parenting services;
- the Department for Education and other government departments with links to this subject including the Department of Health;
- the Department for Communities and Local Government, responsible for implementation;
- frontline professionals and local structures: they are the best placed to
identify the difficulties they face and, where possible, recognise what could help them:
– the practitioners who develop the programmes;
– and parents.

Online publications
The EIF’s publications on the results of the reviews produced as well as the methods used can be accessed via the EIF website. It is the main vehicle for disseminating the work and results of the EIF. For example, the report on parent-child interaction (cited above) sets out all of the programmes studied, with the evaluation methods, and proposes a set of recommendations on the types of interventions to be implemented, the initiatives to prioritise etc. The website thus presents a certain number of practical tools to enable stakeholders to get to grips with the subject:
– a “Maturity matrix”, which enables local authorities to identify how well they are doing in terms of knowing their users, developing their staff, their services, levers for action and how well prepared they are to take preventative action;
– tools to help stakeholders to justify an “investment” in preventative initiatives and studies into the costs of late intervention as well as the benefits of preventative intervention;
– tools to enable the police to take preventative action in their dealings with young people, in cooperation with other relevant local stakeholders.

Stakeholder support
Over time, the EIF has trialled different methods to promote the uptake of effective actions by stakeholders, in particular local authorities.

As part of its “pioneering places” programme, the EIF worked with 20 areas to identify initiatives to better embed the principle of early intervention in their services. The EIF recommended initiatives to improve the identification of the target population and then the best way of supporting them, drawing on effective interventions.

The EIF also runs workshops for sharing practice and advice on the ways of implementing such interventions, ensuring outcomes and effectively targeting the relevant population etc.

**Challenges and lessons**

Disseminating information, however well presented, is not enough. As such, the EIF is exploring the best ways of ensuring that decision-makers and frontline professionals use the results and tools made available to them.

In addition to decision-making, the quality of implementation of these programmes is a key issue, particularly in terms of targeting the right population. This subject is addressed in the book by Dr Kirsten Asmussen, "The evidence-based parenting practitioner’s handbook" in relation to parenting interventions.

One of the strands that can be emphasised when investing in programmes targeting early childhood and adolescence relates to the long-term avoided costs. However, the results of interventions in terms of avoided costs are not always conclusive. For interventions aimed at young children, the reduction of some costs is only visible many years after the intervention. These benefits may be difficult to evaluate or may not result directly in cost reductions. The difficulty in tracing the financial benefits may lead to such programmes to be called into question by policy-makers, even if they have many other benefits.

**Looking ahead**

EIF has a number of aspects of its work that it aims to strengthen over the coming months and years:

- Develop dissemination actions and promoting adoption: this is about drawing on findings from scientific studies on the subject (see work led by Nesta) in order to develop new practices for dissemination. The EIF also wants to reach all stakeholders, not just local authorities, particularly by developing tailored tools and approaches.
- Influence programmes of researchers and stakeholders who can lead trials
to build on existing studies where these are inadequate.
- Continue to convince decision-makers of the importance of early intervention.

**Lessons for France**

**Lessons relating to the What Works Centre approach**
- Determining definitively whether an intervention has a positive measurable effect and under what conditions is a complex task. However, some interventions were able to be scientifically validated. Promising interventions whose impact has yet to be tested represent an opportunity. These interventions can be the subject of scientific trials to generate findings and ideas for improvement.
- The EIF is currently seeking to determine the best methods for disseminating effective and promising interventions. In particular, the EIF is considering new approaches that would take into account what has already been tested and ongoing trials in the United Kingdom, particularly those by the EEF. The best dissemination method will depend in part on the administrative and political context of the country and policy area.

**Specific lessons related to early childhood**
- There are many studies on interventions, particularly parent-child interaction and parenthood, and many of them have been reviewed and synthesised by the EIF: as such, this represents a rich evidence base for discussion in this area, to be compared with existing studies or those already completed in France.
- In interventions targeting the most at-risk families and children, the quality of the identification of target populations is essential in successfully implementing interventions and achieving the desired effects (whilst keeping the cost of the intervention under control).
Education Endowment Foundation

<table>
<thead>
<tr>
<th>At a glance</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Sector</strong></td>
</tr>
<tr>
<td><strong>Launch date</strong></td>
</tr>
<tr>
<td><strong>Organisational type</strong></td>
</tr>
<tr>
<td><strong>Size of team</strong></td>
</tr>
<tr>
<td><strong>Annual budget</strong></td>
</tr>
<tr>
<td><strong>Website</strong></td>
</tr>
</tbody>
</table>

**Context**

The Education Endowment Foundation (EEF) was set up in 2011 by The Sutton Trust, in partnership with the Impetus Trust with an endowment of £125 million from the Department for Education. The Secretary of State for Education, Michael Gove launched this initiative in November 2010, with the aim of raising attainment in under-performing schools. It was inspired by Race to the Top, an ambitious programme introduced by Obama in the United States. The EEF committed to spending the initial endowment in full by 2026. In March 2013, the EEF was designated by the government as the What Works Centre for education.

The EEF’s aim is to tackle the attainment gap by raising educational attainment among the poorest pupils. Through experimentation, research synthesis and the dissemination of knowledge, the EEF aims to

---

help decision-makers invest in programmes that are the most effective (in terms of impact) and the most efficient (in terms of the cost of the service). Thus, the EEF targets head teachers, teachers, parents, academics and policymakers in education. This role is all the more important against the backdrop of decentralised school budgets with schools receiving specific funding for pupils from the most disadvantaged families.

Organisational structure

The centre is an independent charity with a board of seven trustees. The management and trustees receive strategic advice from an Advisory Board made up of 11 members, comprising leading figures from education, public policy and business. The EEF has 19 staff led by the chief executive, Sir Kevan Collins. In addition to the administrative and communication departments, the EEF has three teams: Grants, Evaluation and Dissemination and Impact. The members of these teams come from different backgrounds (government departments, local authorities, think tanks and charities). A number of them have worked as teachers, giving them direct experience of the day-to-day lives of their main target audience.

The EEF’s independence from the government enables it to develop and pursue a long-term vision. Moreover, the centre’s approach has kept it away from controversies around education policy. For example, since 2011, owing to the roll-out of Free Schools and Academies the government has been accused in some quarters of “liberalising” education in England. The EEF’s focus – on the teaching methods used in the classroom – has largely prevented it from getting embroiled in these discussions.
Activities

Evidence synthesis

From its creation, the EEF has carried out a great deal of synthesis of research in education. Drawing on research by Durham University for the Sutton Trust, in 2011 the EEF released its Teaching & Learning Toolkit. It is a “review of reviews” of evaluations of educational interventions. This synthesis relates mainly to the results of experimental or quasi-experimental evaluations. This review has provided insight into the latest developments in the educational sciences. Unlike previous syntheses, the Teaching & Learning Toolkit was designed as a decision-making tool for practitioners. The Toolkit can be used to compare different educational interventions based on three criteria: impact, evidence strength (quality, quantity and consistency of evaluations) and cost. In order to compare the impact of interventions, the research team translated the results of evaluations into a single measurement: additional months’ progress for the test group (compared with the control group). Arriving at such a succinct and accessible presentation of the research is a difficult task. Indeed, there is a tension between certain academics (who want to remain very exact and argue, for example, that using the months’ progress measurement only is flawed) and the needs of decision-makers who would be “paralysed” if they were to receive too much information.

“The toolkit is a meta-analysis of metaanalyses. Researchers have done that before. The big difference was we did that in a way that was accessible for teachers and head teachers.”

James Turner, Deputy Chief Executive.

The Toolkit is now an interactive tool available via the EEF website. It is also dynamic. Updated every six months by the research team at Durham University, the number of education interventions it covers has increased from 11 to 34 since 2011. In addition to adding new interventions, the analysis of interventions has also been updated to take into account new studies.
on interventions already in the Toolkit. EEF also funded specific literature reviews to fill the gaps in the Toolkit.

The Toolkit has helped the EEF to identify the types of intervention which would benefit from further research and which might be the subject of new trials funded by the EEF. Ultimately, these trials produce findings which feed into the Toolkit. The example of teaching assistants is a good illustration of this virtuous circle.

“The evaluations of TAs reviewed for our toolkit suggested it was having a neutral, even negative impact. The toolkit brought the issue out. That has driven a lot of action at our end. So we have funded trials looking at how to use TAs to do one-to-one interventions…. [which] showed consistently positive effects on the students. These results then, in turn, changed the conclusions of our Toolkit.”

Jonathan Sharples, chercheur expérimenté, Education Endowment Foundation.

Extrait du Teaching & Learning Toolkit

Source: educationendowmentfoundation.org.uk/evidence/teaching-learning-toolkit.
Trials

At the time the EEF was set up in 2011, few experimental evaluations had been carried out in education in England. As such, one of the EEF’s objectives was to produce numerous experimental evaluations to test the effectiveness of educational interventions in the English context. Thanks to its substantial endowment, the EEF had the necessary resources to fund both innovative projects and their evaluations. This funding of the organisations that conduct trials is one of the factors behind the programme’s success: it encourages organisations to subject their project to a robust evaluation despite the organisational risks that that might represent for them, if the impacts of the programme tested prove neutral or negative.

Since 2011, every year the EEF has held two “open” funding rounds each year and a themed round for one aspect of education – such as digital technologies or financial education. The EEF receives an average of 200 responses for each round out of which around ten projects are selected. Applications are mainly from charitable organisations working in schools, but also from schools and academics. The quality of proposals from schools tend to be lower but where they have a practice that they are seeking to disseminate, it is often easier for them to convince other schools to take part.

The “Grants” team first evaluates the project proposals and scores them according to their development stage, which corresponds to three levels of EEF support:

- for projects at an early stage in their development, the EEF proposes a **pilot project** to test the validity of the concept, without impact evaluation;
- for projects that have already been tested but where there is little evidence of their impact, EEF propose an **efficacy trials** test to evaluate the project in ideal conditions;
- for projects with more robust evidence as to their impact, the EEF offers an **effectiveness trial** to evaluate the project in “real world” conditions with a view to scaling it across the country. The EEF can thus support projects throughout their development, from design to scaling.
Once the project has been identified, the EEF evaluation team prepares the evaluation in collaboration with the project lead. The EEF uses external evaluators in order to ensure the independence of evaluations. The EEF generally favours experimental evaluations. On the one hand, these methods are generally used to analyse the impact of the project (105 of 127 evaluations to date have been so called “randomised” evaluations). On the other hand, the EEF requires an evaluation of the process that enables the conditions for success of the intervention to be analysed. Finally, the EEF focusses on the longitudinal effects of the projects and as such has created a tool for using data from the Department of Education to monitor the school results of project beneficiaries until the end of their school career.

Despite many successes, the EEF has encountered difficulties with certain evaluations. A number of trials selected by the EEF did not go ahead as the project leader was unable to incorporate evaluation in the project. These difficulties often came down to the fact that the evaluation was not sufficiently taken into account in the design of the proposed trial. As such, the EEF changed its approach.

“Now, before the project goes through the grants committee, we have a meeting bringing together the provider and the evaluator. The idea is that they take a co-created evaluation and project to the grants committee for sign-off. This means we can ensure, for example, that the intervention will have a sufficient sample to have the right effect size.”

Camilla Nevill, Evaluation Manager.

Initially, EEF came up against a lack of familiarity with these evaluation methods on the part of project leads and a lack of expertise in experimental methods on the part of researchers. As such, the EEF has had to raise awareness among stakeholders and support skills development among evaluators. To this end, the evaluation team has produced many resources available on the EEF site \[21\]. The EEF also points to the fact that over time it has

\[21\] educationendowmentfoundation.org.uk/evaluation/resources-centre.
got better at scoping and designing evaluations, incorporating the process evaluation and with a better design to guard against the risk of attrition.

The results of evaluations have also generated some debate. Indeed, of around 50 evaluations that have been finalised, only 12 were judged sufficiently promising to receive further financial support from the EEF – in most cases for large-scale trials. Unsurprisingly, it is difficult for some project leads to accept that their intervention has not had measurable effects on pupils. However, some projects have had very interesting results, like the six trials on teaching assistants.

“So we have funded six trials looking at how to use TAs to do one-to-one interventions, in the form of a literacy scheme. Every one of the studies showed consistently positive effects with consistent effect sizes, which is exceptional.”

Jonathan Sharples, Senior Researcher, Education Endowment Foundation.

On the whole, the EEF has proved those critics who feared a strong resistance from teachers to experimental evaluation methods, wrong. Indeed, it has carried out 127 evaluations including 105 as randomised trials in 7,000 schools, i.e. one for every four schools in England.

“A few years ago, one RCT had been done in the education sector in England. Suddenly we’ve got over 100. It’s been like flicking a switch!”

Jonathan Sharples, Senior Researcher, Education Endowment Foundation.

Dissemination and promoting uptake

Between 2011 and 2015, the Teaching and Learning Toolkit was at the heart of the EEF’s dissemination activities. This Toolkit now synthesises the results of some 10,000 evaluations on a page of the website. As such, it represents a sort of “one-stop shop” to access education research. It is clear, accessible and above all acts as a decision-making tool for education practitioners.
“The toolkit responds to some of the constraints facing frontline professionals: lack of time, too much information and unable to sift through knowledge. The toolkit was a game changer in making education research useful to practitioners.”

Jonathan Sharples, Senior Researcher, Education Endowment Foundation.

As such, since 2011, the EEF has worked hard to raise awareness among education practitioners of the existence and use of the Toolkit, with very interesting results in terms of access to the tool.

“Now 2/3 of head teachers report using the toolkit. Some say it’s like a “Which” consumer magazine, because it’s independent, authoritative, and clear.”

Jonathan Sharples, Senior Researcher, Education Endowment Foundation.

This success is of course down to the accessibility of the tool but also to the position the Toolkit holds as an additional resource to supplement the expertise of practitioners and local knowledge. However, the EEF’s Dissemination and Impact team is conscious that the dissemination of the tool is probably not enough to transform practice.

“Creating online resources isn’t enough. A big theme in knowledge mobilisation is that it is a social process, you have to create motivations and opportunities to bring that to life.”

Jonathan Sharples, Senior Researcher, Education Endowment Foundation.

Consequently, the EEF has decided to carry out more targeted dissemination activities. In particular, since 2015, the EEF has run two activities based on two methods:

- **Campaigns.** These are campaigns to raise awareness and promote up-take of effective interventions, such as the campaign on teaching assistants. Indeed, the EEF decided to turn these research findings on

the use of teaching assistants (including six trials funded by EEF), into an operational guide\textsuperscript{23}. This guide was accompanied by resources to promote the adoption of good practice such as a self-assessment guide.

- **Menus.** The approach is simple: create a list of interventions from the Toolkit and use financial incentives and support to promote their uptake by schools. An initial pilot project is under way in a local authority in East London.

### Teaching assistants campaign: extract from the self-assessment guide.

#### A self-assessment guide

**RECOMMENDATIONS I AND II**

**Deployment of TAs in the classroom**

<table>
<thead>
<tr>
<th>I</th>
<th>INEFFECTIVE</th>
</tr>
</thead>
<tbody>
<tr>
<td>TAs routinely have a direct, informal teaching role with lower-attaining pupils and/or those with SEN. They regularly adopt the status of ‘primary educator’ to these pupils.</td>
<td></td>
</tr>
<tr>
<td>This arrangement causes a ‘specialisation’ effect: high-attained pupils have fewer opportunities to engage with whole-class teaching and have fewer interactions with teachers and peers.</td>
<td></td>
</tr>
<tr>
<td>Teachers assume that a high amount of TA support creates the needs of lower-attaining and/or SEN pupils, and as a result spend less time working with these groups.</td>
<td></td>
</tr>
<tr>
<td>Personal relationships may be good, but there is a lack of coherence and effective teamwork between teachers and TAs in the classroom.</td>
<td></td>
</tr>
<tr>
<td>The TAs are either mobile in the classroom and support whole-class teaching (such as delivering equipment, or working at the whiteboard).</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>I</th>
<th>IMPROVING</th>
</tr>
</thead>
<tbody>
<tr>
<td>There is variability in the quality and appropriateness of teachers’ monitoring-by-moment decision-making regarding TA deployment.</td>
<td></td>
</tr>
<tr>
<td>In some lessons, TAs are effectively supplementing the whole-class teaching, although this is not consistent.</td>
<td></td>
</tr>
<tr>
<td>There is a tendency to drift back towards TAs working in an isolated teaching role with lower-attaining and/or SEN pupils, particularly as the lesson progresses.</td>
<td></td>
</tr>
<tr>
<td>There is little or no attempt to refocus the TA during lessons in response to the real-time needs of pupils.</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>I</th>
<th>EXEMPLARY</th>
</tr>
</thead>
<tbody>
<tr>
<td>The learning needs of all pupils are met, first and foremost, through high-quality teaching by the teacher, teachers spend at least as much time working with lower-attainers as others.</td>
<td></td>
</tr>
<tr>
<td>TAs work with a range of pupils within the class and supplement and extend teachers’ work rather than replace them. Support for specific individuals or groups is structured so it helps them access general classroom teaching.</td>
<td></td>
</tr>
<tr>
<td>Teachers and TAs work effectively as a team, with a shared understanding of their respective roles in achieving lesson objectives. TAs’ work is monitored by frequent documentation regarding TA deployment and driven by pupil needs.</td>
<td></td>
</tr>
<tr>
<td>TAs support whole-class delivery at the front of the class, now the room or undertake administrative tasks to ease teachers’ workload.</td>
<td></td>
</tr>
</tbody>
</table>

**Source:** educationendowmentfoundation.org.uk.


128
Although the EEF has developed its dissemination methods, the team acknowledges that they do not yet know “What Works” in this area. Indeed, very few knowledge-mobilisation projects have been evaluated to verify their effects on professional practice. As such, the EEF decided to test these dissemination methods using actual trials. The EEF is currently trialling four approaches to disseminating effective literacy practices. Each project has a passive strand – where resources are sent to schools – and an active strand – where these resources are accompanied by events (conferences, ateliers etc.). The idea is to check whether the events have added value when compared with simply circulating resources.

Looking ahead

EEF has a number of aspects of its work that it aims to strengthen over the coming months and years:

- Having been highly focused on the production and synthesis of knowledge since 2011, the EEF is to dedicate itself more to the dissemination of knowledge and the uptake of effective practice over the next five years. The centre now produces more explicit guidelines, in addition to highlighting the findings of studies.
- The EEF will move from the widespread dissemination of the Toolkit towards more targeted dissemination activities to promote the uptake of effective practice by practitioners.
- Until now, the EEF mainly trialled interventions aimed at improving pupil attainment. It will now invest more in trialling interventions aimed at developing social and emotional (non-cognitive) skills.

Lessons for France

- The EEF had a substantial endowment to be spent over 15 years, allowing it to engage in large scale activities over the long-term.
- The EEF has been able to take a systematic approach to using the
evidence base. Indeed, the centre acts across the whole knowledge chain, from synthesis to evaluation to dissemination.

- The Teaching & Learning Toolkit has been at the heart of the EEF’s success. A real “one-stop shop” for education research and a decision-making tool, it provides a response tailored to the constraints of frontline practitioners. From the early years it has also helped to show how the EEF could be useful and to raise its profile.

- Making the research available in an accessible way in this Toolkit has been necessary but not sufficient to promote the uptake of effective practice by teachers. Practitioners must be offered opportunities and motivated to take up effective interventions.

- The evaluator and the project lead must communicate with each other from the trial design stage in order for the project to incorporate the demands of a scientific evaluation.

- More generally, project leads and researchers required support in acquiring the skills related to experimental evaluation methods.

- The EEF works with the EIPPEE\textsuperscript{24}, a European network that promotes the use of research in education.
What Works Centre for Local Economic Growth

<table>
<thead>
<tr>
<th>At a glance</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Sector</strong></td>
</tr>
<tr>
<td><strong>Launch date</strong></td>
</tr>
<tr>
<td><strong>Organisational type</strong></td>
</tr>
<tr>
<td><strong>Size of team</strong></td>
</tr>
<tr>
<td><strong>Annual budget</strong></td>
</tr>
<tr>
<td><strong>Website</strong></td>
</tr>
</tbody>
</table>

Context

The What Works Centre for Economic Growth was set up in October 2013 by a consortium led by the London School of Economics. Between 2013 and 2016, the centre received £3.1 million of initial funding from three funders – the Economic and Social Research Council, the Department for Communities and Local Government (DCLG) and the Department for Business (BIS).

The aim of the centre is to provide relevant solutions to local and national policymakers through:
- systematically reviewing the evidence base on the impact of policies for local economic growth;
- helping local authorities, Local Enterprise Partnerships, central government and business understand and make better use of evidence in designing and delivering policy;
- improving the quality of the UK evidence base by helping to develop “demonstrator” projects – local policy experiments that develop and test innovative approaches.
The term “local economic growth” is relatively vague, as it does not fall neatly within sectors set out by research or public policy. The team therefore decided to use a relatively narrow definition: programmes that directly or indirectly improve local productivity and employment (job creation, return to work).

**Organisational structure**

The centre is run by a partnership of three organisations:
- **The London School of Economics** (LSE). The centre is led by Professor Henry Overman and his deputy Doctor Max Nathan who represent it at an institutional level. The LSE is also responsible for designing and proofreading all of the systematic reviews and evaluation work. Where the LSE is heavily involved.
- **Arup.** This consultancy, whose expertise includes infrastructure, transport, planning and economic growth, is responsible for systematic reviews. Two senior consultants are what works centre leads at Arup and involve other analysts in the production of reviews.
- **Centre for Cities.** This think tank specialises in analysing the economic growth of major British cities. Under the leadership of Meg Kaufman, six people in total are involved in the What Works Centre at Centre for Cities, primarily on communications (website, documentation, events).

As such, the centre is atypical in that it is not made up of a single organisation but is an initiative led by a partnership between three organisations.

It has a “virtual” team made up of representatives of the three partners which in turn bring in their colleagues to carry out specific assignments (such as a review). Despite the risks associated with this model of management, the centre has worked well on the whole, thanks to a clear allocation of roles and shared management processes (monthly meetings, shared calendar).
The centre also draws on two panels:

- The *User Panel* is made up mainly of policymakers working in local economic development and is led by Tom Bridges, Chief Economic Development Officer at Leeds City Council. The members of the Panel meet four times a year. This panel helps to set the centre’s research priorities and determine the best way of communicating with decision-makers. Discussion may focus on, for example, the best formats for disseminating findings or the layout of the website. Some panel members are also asked to give ad-hoc presentations on specific subjects. Their feedback, judged “very useful” by the team, helps the centre to make their work accessible and relevant for their target audience;

- The *Academic Panel* reviews the centre’s work to ensure it is as robust as possible. This panel is made up of around 10 academics from the London School of Economics and other universities. The panel is “virtual” in that its members do not meet face to face. Their role involves supporting the Centre’s research activities. Until now, this has mainly meant support in developing systematic reviews and toolkits. Thus, one to two members of the Panel are assigned to each review to offer advice on various aspects of research, helping them to find their bearings in the literature, to model their research strategy and draw conclusions. Academics are paid for their time (around 1 to 5 days per review) either directly in the case of the LSE academics, or by their institution in the case of other universities. According to the LEG team, the members of the panel have played a “crucial” role in the reviews. The centre is funded in large part by government departments, who comment on its deliverables prior to publication. This close relationship with the state potentially makes the centre less independent, but it is considered an advantage by the centre, particularly in terms of its profile and impact.

“Because government funds us they have to listen to us. Plus, having a government logo on your resources adds credibility for your audience.”

Meg Kaufman, What Works Centres Coordinator at the Centre for Cities.
Activities

Evidence synthesis

In the last three years, the centre has conducted 11 systematic reviews (“evidence reviews”). Each review covered an aspect of local economic development, such as employment training, estate renewal and sport and culture. The themes were initially defined in the call for tenders; themes were then agreed upon by the management team and funders. These reviews are carried out using the same method:
- the scope of the review is determined in conjunction with the User Panel in order to establish the inclusion criteria for evaluations;
- a search is carried out for all of the RCTs and quasi-experimental studies in OECD countries that fall within that scope and are published in English;
- those that meet the centre’s methodological standards are selected;
- the quality of evaluation methods and the clarity of project implementation are evaluated;
- findings from these evaluations and the evidence base more broadly are synthesised in order to draw conclusions.

The centre only analyses evaluations carried out using experimental or quasi-experimental methods. This means only evaluations with a credible counterfactual are analysed in their reviews. The centre uses the Maryland Scientific Scale to select evaluations, which must correspond to levels three to five on this scale (see next page). Finally, this analysis is set out in a technical report of around thirty pages, with appendices and a brief summary. These reports are all available on the centre’s website.
Methodology for systematic reviews

**Maryland Scientific Methods Scale**

**Level 1.** Either (a) a cross-sectional comparison of treated groups with untreated groups, or (b) a before-and-after comparison of treated group, without an untreated comparison group.

**Level 2.** Use of adequate control variables and either (a) a cross-sectional comparison of treated groups with untreated groups, or (b) a before-and-after comparison of treated group, without an untreated comparison group.

**Level 3.** Comparison of outcomes in treated group after an intervention, with outcomes in the treated group before the intervention, and a comparison group used to provide a counterfactual (e.g., difference in difference).

**Level 4.** Comparison of outcomes in treated group after an intervention, with outcomes in the treated group before the intervention, and a comparison group used to provide a counterfactual (e.g., difference in difference).
Level 5. Reserved for research designs that involve explicit randomisation into treatment and control groups, with Randomised Control Trials (RCTs) providing the definitive example.

The 11 reviews all produced interesting findings. The LSE reach team also noted a trend in the studies on sport and culture, the public realm and urban renewal: a significant portion of the programmes evaluated do not show measurable positive economic impacts but can have other worthwhile effects (quality of life, amenity benefits).

“We have found that a lot of the economic arguments made don’t cash out – but there are lots of other reasons you would want to do a programme.”

Max Nathan, Deputy Director, What Works Centre for Local Economic Growth (LSE).

Major sporting and cultural events: Sporting and cultural events rarely have major effects on the local economy and in most cases have a neutral effect. The wage effects of projects are generally low and limited either to the immediate surroundings of the venue, or to a specific category of workers. Given the high cost of sporting and cultural events, they are rarely an effective investment in terms of impact on economic growth. However, they have a societal value and can have effects on the well-being of inhabitants, for example.

whatworksgrowth.org/policy-area/sport-and-culture.

In carrying out these reviews, this centre has been confronted with the limited supply of RCTs and quasi-experimental studies in the area of local economic development. This is down to two factors. Area-based initiatives, such as enterprise zones, urban renewal or the construction of infrastructure, do not lend themselves to the most scientific evaluation methods. For example, just 21 of the 1050 evaluations of urban renewal initiatives
reviewed meet the minimum standards. As such, the centre is considering alternatives to ‘randomisation’, such as longitudinal studies. On the other hand, even in cases where the initiatives do lend themselves better to such methods (personal support for jobseekers or businesses), very few evaluations of this type have been carried out.

“Some of the programmes aren’t amenable to the highest quality methods. Active labour market programmes, or for start-ups and business and advice – you can use randomisation. With housing and infrastructure, you can’t randomise where that goes.”

Max Nathan, Deputy Director, What Works Centre for Local Economic Growth (LSE).

The centre’s research team also noted that the available evaluations are generally not sufficiently attentive to the description of the initiative itself. As such, this failing makes the dissemination of “effective” initiatives difficult as the centre often has little in the way of information on their processes and conditions for success.

“Academics are obsessed with research design, sometimes to the exclusion of everything else. As a policy maker or practitioner, you need to understand the nitty gritty of the how the programme works!”

Max Nathan, Deputy Director, What Works Centre for Local Economic Growth (LSE).

In light of these findings, the research team has been forced to make changes to its review methods. Initially, the centre had the intention of carrying out traditional meta-analysis by calculating the average size of effects in the studies. This approach is appropriate where there are numerous evaluations with a counterfactual that have identical outcomes, but it was impossible in the case of the centre’s reviews which were based on a small number of evaluations. As such, the centre decided to draw out the overall “trends” of evaluations: overall did they show effects that were positive, neutral, negative or mixed? Moreover, the centre analysed whether the effects varied according to the type of initiative, for instance for their review on employment training.
Employment training: For basic skills or interventions aimed at boosting the overall employability of individuals, shorter programmes (under six months or even under four months) have a greater and more marked impact on the employment situation of beneficiaries than longer programmes. These short programmes are also less costly, which suggests that they are more effective than the long programmes.

Trials

The centre aims to launch pilot projects and local trials so as to address the lack of evidence highlighted in their systematic reviews. However, it does not have the necessary resources to fund these trials directly. The centre has pursued two strands of activity aimed at encouraging the launch of new trials:

Firstly, the LSE team has offered personalised support for designing and launching trials in its policy area, particularly for local authorities. Indeed, the centre has a continuous call for projects in place. In theory, their advice should be mainly around the ways of modelling an initiative so that it is “evaluatable” using scientific methods. In practice, the team often begins by raising awareness of experimental and quasi-experimental methods among local officials and councillors. In addition, time is spent on funding applications and designing the trial.

This preparation for launching a trial is often long-winded and in some cases does not bear fruit. Thus, although the LSE team has been involved in the design of 20 pilot projects, just 4 are on the verge of being launched.

Secondly, the centre carries out awareness-raising initiatives on experimental and quasi-experimental methods. With its guide, How to evaluate, the centre aims at helping public officials to adopt these methods. The content of this guide is also fully embedded in their programme of local workshops. Finally, the centre is attempting to ‘raise’ the standard of evaluations funded by central government.
“So we’re working with government departments to improve the quality of the evaluations they commission. Today, the invitation to tenders for evaluation projects sent out by government departments now often include Maryland Scale standards as minimum, and something about process evaluation.”

Max Nathan, Deputy Director, What Works Centre for Local Economic Growth (LSE)

Dissemination and promoting uptake

The centre aims to raise awareness among stakeholders in local economic development of the evidence base so that they use it more in the design and implementation of policies. Although local initiatives mainly target businesses, the main target audience for their dissemination are officials in national and local government. The centre raised awareness among this group in three stages.

The team began by making their evidence reviews as accessible as possible to a professional audience, particularly by producing an Executive Summary. Indeed, for each of the first three reviews, the three partners held a six-hour meeting to make the study findings accessible to policymakers.

This work enabled them to determine the appropriate format and language which the research teams attempted to reproduce in subsequent reviews. Thus, the centre has produced 11 reviews and the WWC portal acts as a “one-stop shop” to the evidence of what works in the field of local economic development.

“So there was a lot of work with academics to be more explicit with their recommendations, and to stop sitting on the fence. There is a creative tension between being rigorous and having an impactful message.”

Meg Kaufman, What Works Centres Coordinator at the Centre for Cities.

The team went on to produce several toolkits whose aim was to encourage the uptake of practices that have proven results. Each toolkit covers a review topic, such as employment training, and brings together several
guides covering interventions in more detail. Using the evaluations catalogued during the systematic review process as a starting point, these guides offer more detailed advice on the design of interventions and their conditions for success. For example, the employment training review has been turned into a toolkit made up of four guides: careers counselling; financial incentives; pre-qualifications and reminders (to improve attendance on courses).

“The toolkits are not a step-by-step guide to doing a local economic growth plan. We have suggestions for how to do things better, but we don’t want people thinking that it is the final word on effectiveness.”

Meg Kaufman, What Works Centres Coordinator at the Centre for Cities.

### Employment training: the guides

<table>
<thead>
<tr>
<th>Policy Design Tool</th>
<th>What does it aim to do?</th>
<th>How secure is the evidence?</th>
<th>How much does it cost?</th>
<th>How effective is it?</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Careers Counselling</strong></td>
<td>Improve take-up and completion of training</td>
<td></td>
<td>£££</td>
<td>★★★★★</td>
</tr>
<tr>
<td><strong>Financial incentives</strong></td>
<td>Improve take-up and completion of training</td>
<td></td>
<td>£££</td>
<td>★★★★★</td>
</tr>
<tr>
<td><strong>Pre-qualification</strong></td>
<td>Improve performance and completion of training</td>
<td></td>
<td>£££</td>
<td>★★★★★</td>
</tr>
<tr>
<td><strong>Reminders</strong></td>
<td>Improve attendance to training programmes</td>
<td></td>
<td>£££</td>
<td>★★★★★</td>
</tr>
</tbody>
</table>

Source: www.whatworksgrowth.org/resources/employment-training-toolkit.

Moreover, the centre has, since its creation, used various methods to encourage the dissemination and uptake of its outputs. These include sending out reviews through its newsletter and Twitter account, thus enabling it to reach several thousand people. The centre has also organised national and local events to promote the uptake of its publications by practitioners.
These events have been an opportunity to share work with practitioners who remain wary of randomisation. Some of these events have been an opportunity for lively debate on these issues and have generally secured practitioners’ buy-in to the process. Aware that this up-take is often a social and localised process, the centre has organised a number of bespoke local workshops over the last year.

“It’s not enough to just create a mailing list and send out your report. Their format is unfamiliar, and their methodological rigour is stark. We need more time to explain to people that RCTs are not unethical. There needs to be a fair amount of introduction that was better done in person. Our most impact and uptake comes from meeting and talking to people.”

Meg Kaufman, What Works Centres Coordinator at the Centre for Cities.

Looking ahead

The centre has two main focuses over the coming months and years:

- This centres’ future will be clearer after its core funding has been renewed in early 2017.
- At this stage, the team considers that the centre will spend less time synthesising research and will focus instead on dissemination and experimentation activities. Indeed, the centre will probably focus on the uptake of effective practices through the dissemination of toolkits and the promotion of evaluation with a view to launching trials.

Lessons for France

- The centre operates in an area which is lacking in RCTs and quasi-experimental studies. Nevertheless, the centre has opted to maintain high standards when selecting evaluations for inclusion in its reviews. Consequently, these are based on a relatively small number of studies, which limits the centre’s ability to learn general lessons.
- In preparing the evidence reviews, the centre observed creative tension
between scientific standards and the clarity of the message. In this respect, discussions between researchers and communication professionals produced reviews that were both accessible to practitioners and scientifically robust.

- **The centre’s remit was to launch pilot projects and local trials, but it did not have the resources to fund them.** The centre supports the design of potential trials, a long process which in many cases does not bear fruit. Indeed, at this stage no trials have been launched with its support, although four are currently in the pipeline.

- **In order to deliver on its remit, the centre must support a cultural shift among public stakeholders, a long and complex process.** Indeed, it has taken three years to carry out its programme of systematic evidence reviews. The centre is now focussing on the immense task of promoting scientific evaluations and evidence-based public policies. This requires substantial work on awareness-raising, training and, in some cases, personal support for civil servants.

- **The centre has a relatively small target audience but one which is not united in a single professional body.** In a sense, the centre performs this role in bringing together a network of local economic growth practitioners.
Agence nouvelle des solidarités actives

Centre for Ageing Better

At a glance

<table>
<thead>
<tr>
<th>Sector</th>
<th>Ageing well</th>
</tr>
</thead>
<tbody>
<tr>
<td>Launch date</td>
<td>2015</td>
</tr>
<tr>
<td>Organisational type</td>
<td>Charitable foundation</td>
</tr>
<tr>
<td>Size of team</td>
<td>20 staff</td>
</tr>
<tr>
<td>Annual budget</td>
<td>£5 million</td>
</tr>
<tr>
<td>Website</td>
<td><a href="http://www.ageing-better.org.uk">www.ageing-better.org.uk</a></td>
</tr>
</tbody>
</table>

Context

The ageing population has brought the quality of life of older people to the forefront of public debate in western countries.

This issue has been identified by the Big Lottery Fund (BLF) as one of its priorities. The BLF is a public body responsible for distributing the funds raised by the National Lottery for good causes and funds projects, charitable organisations and institutions. It has a long-standing commitment to the issue of ageing through various programmes and research. Recent programmes include a call for projects to disseminate or scale up home care and dementia initiatives which have had a proven impact and a call for innovative projects “Fulfilling Lives: Ageing Better” launched in 2014 and 2015 respectively.

In addition, in 2013, the House of Lords Committee on Public Services and Demographic Change recommended the creation of a What Works Centre in its report Ready for Ageing? to keep momentum alive in a context where initiatives with a high degree of goodwill fizzle out due to lack of attention.
In the discussions that followed, the Big Lottery Fund recognised the need to invest on the production and dissemination of evidence (in particular RCTs / quasi-experimental studies) in this area and decided to fund the Centre for Ageing Better over a ten-year period through the creation of an endowment in 2015. £50 million of funding for the Centre for Ageing Better was announced at the same time as funding for 14 local programmes under Fulfilling Lives: Ageing Better which received a total of £80 million.

Organisational structure and objectives

The Centre for Ageing Better (CfAB) is an independent structure with an endowment from the Big Lottery Fund to be spent over a ten-year period. The centre is administered by a board of trustees who work or have worked in research, public policy, the charity sector, and civil society, on issues around ageing and the loss of autonomy. The chair, Lord Filkin CBE, is a member of the House of Lords and chaired the report Ready for Ageing. The Centre for Ageing Better aims to:

- develop, share and implement findings from evidence to help people to age better;
- support innovation to address the challenges and opportunities of an ageing population;
- contribute to the societal changes (at the national and local level) needed to improve quality of life for older people.

The centre currently has a team of 20 which is set to increase to around 25-30 permanent staff members. Its team is made up in particular of:

- evidence managers and officers: - six academics specialising in social sciences and qualitative research, mostly with knowledge of issues around poverty, well-being and behavioural science, health and health research (for example, chronic pain management and use of art therapy in a clinical context)

25 The expertise sought was in applied research, particularly social sciences, reviews, the ‘translation’ of knowledge and evaluation, rather than primary research or public health.
– communications managers and network officers: five people with a background in media, communications and public affairs;
– managers responsible for supporting innovation and the implementation of new practices: four people with experience of civil society, housing associations, of resident participation or indeed with public sector providers;
– a back-office team (finance and administration, logistics etc.).

The centre sets up advisory panels made up of experts from academia, professional life and people affected for each of its subjects of the work.

The key partners of this centre are local authorities (primarily Greater Manchester) to implement projects and advise them on their strategy, government departments and secretaries of state with which the centre shares results of research, networks of stakeholders whose aim is to promote one or more aspects of ageing well, charitable organisations and businesses. The CfAB works closely with universities (for example, Manchester Institute for Collaborative Research on Ageing). The centre plays an essential role in advising the Big Lottery Fund on ageing as well as the funding streams allocated to it. Finally, the centre works with Public Health England on targeted campaigns.

The CfAB has attempted to ensure its activities complement the remit of existing stakeholders in the ageing sector, particularly NICE and Public Health England.

Who does what in the United Kingdom?

NICE produces systematic reviews and guidelines on a range of ageing and disability related clinical interventions, including assistive devices and disability aids (such as hearing aids, for example). These guidelines are updated regularly. NICE also develops reviews and guidelines on mental health and disability, both in terms of prevention and support. For example, NICE has produced a set of guidelines containing “mid-life” approaches to delay or prevent frailty, disability and dementia. 
Public Health England (PHE) runs programmes relating to the ageing population: the aim of these programmes is to promote initiatives by stakeholders in the health sector as well as by other relevant stakeholders, to help a certain number of institutions to put research findings into practice and implement relevant initiatives and finally to contribute to campaigns by the National Health Service (NHS). These programmes draw on existing scientific evidence, on research commissioned from academics and on collaboration with NICE. Topics under investigation include dementia, ageing well, fall prevention and support for carers. PHE also helps to run pilot projects. In some of these areas, PHE works towards producing consensus statements. These involve bringing together a set of key stakeholders in the area under investigation, be they frontline professionals or academics. For example, to study the scientific evidence available and agree on initiatives to be implemented, in particular for the institutions present. In 2014 the Blackfriars Consensus on dementia prevention was held\textsuperscript{xcvi}. It was supported by 60 organisations and experts in the prevention of dementia and non-communicable diseases.

Activities

Analyse need

There is no single recognised definition of “ageing well”. As such the CfAB decided to carry out analysis in order to give this concept substance. As part of this, it devised the Later Life in 2015 project with Ipsos MORI. This project has several qualitative and qualitative strands, detailed below. Through this project, work strands for the centre have been defined with participation of older people.

English longitudinal study of ageing (ELSA)

The CfAB used a longitudinal study\textsuperscript{26} of a cohort in order to study ageing\textsuperscript{xcvii}:

\textsuperscript{26} The longitudinal study is also a way of identifying correlations and pinpointing causal links to be tested – for example, the study was able to identify that older people who did volunteer work were more likely to be in good health, have an active social life and experience well-being, health without having determined the direction of the causal link. The CfAB also commissions primary research on return to work initiatives, contributing to communities and digital technology.
this study was set up 14 years ago and in the beginning included over 10,000 men and women aged between 50 and 100. Every year for 12 years the cohort was interviewed (the latest data was collected in 2015). The study is very broad in scope and includes biometric, socio-economic and well-being information etc. It has mainly been used by the CfAB to define the segments of the elderly population for the Later Life study.

Later Life in 2015

The CfAB partnered with IPSOS MORI to undertake a major social research project exploring people’s well-being in later life based on their own responses. The longitudinal study cited above was used in the more recent phase, to split the population into 6 segments according to their experiences, circumstances and levels of well-being. In-depth case studies were carried out with around twenty people to gain an understanding of what quality of life means to the older population, their situation based on criteria ranging from health to their living environment (housing, local neighbourhood etc.) and in order to set priorities for improving the situation according to their point of view. Finally, an additional survey of over 1,000 representative people was carried out to confirm the aspects identified in the qualitative survey.

Source: laterlife.ageing-better.org.uk.
Work strands
The CfAB’s work strands have been informed by the Later life in 2015 study, consultations of all stakeholders, and taking into account existing evidence and evaluations of interventions. The CfAB is currently working on seven themes depicted as personal objectives for ageing people, with an initial focus on certain aspects of these issues:

– being in fulfilling work – with programmes to ensure that work and the working environment are suitable and programmes that support the over-50s in returning to work;
– living in a suitable home – with a programme under way on adapting housing for old age;
– contributing to communities – supporting people who are less likely to get involved in community activity;
– managing major life changes – support with the transition into retirement;
– keeping physically active – helping people to maintain and improve their muscle strength and balance;
– living in age-friendly communities – promoting a place-based approach to supporting all aspects of healthy ageing;
– getting the most out of digital – helping people to stay or get involved in digital technology.

Evidence synthesis
In 2016, the CfAB began a programme of reviews of evaluations of interventions aimed at improving the quality of life of older people. The methodology is in the process of being agreed and shall be adapted according to the subject and the available literature.

Examples of systematic reviews and syntheses of existing research
– the role of home adaptations in improving later life: Evidence review;
– evidence Briefing - The benefits of making a contribution to your community in later life;
– fulfilling work: What do older workers value about work and why?
The CfAB is supporting a review led by PHE of analysis into the cost-effectiveness of interventions on improving the muscle strength and balance of older people. The CfAB should also be launching a call for good practice which would be unlikely to be found in the existing scientific literature. This approach is evidence of the centre’s openness to projects that show promise but have not yet been the subject of an RCT or quasi-experimental study.

The findings
This programme of reviews is intended to identify the interventions proven to be effective or promising in at least one of the desired strands for improvement (for example, being infulfilling work). This impact should be set against any potential negative effects on the other strands (for example, an intervention which leads to greater financial security but has a strong negative effect on mental health would not be seen as having much value).

The quality of evaluations will be an issue for the centre in this area: the expected outcomes often fairly open, it can be less appropriate or more complicated to carry out randomised studies in this context. Moreover, according to the quality criteria used by the centre in conducting evaluations, it is not clear in the case of some interventions trialed on a small scale whether they can be rolled out. Finally, evaluations often have too short a time scale, for example, the beneficiaries of interventions on preparing for retirement are generally monitored for just two years.

Incentives
There is already a substantial body of evidence on interventions that have had a positive impact on people (the positive effect of regular physical activity on wellbeing has already been proven). As such, one of the issues for the centre is to identify and gather scientific findings from the evaluation of interventions that incentivise older people to change their behaviour in a sustained manner before proposing better interventions for local authority service providers.
British What Works Centres: what lessons for evidence-based policy in France?

“Our main aim is to understand what works so that more people can benefit from these interventions.”

Dan Jones – Director of Innovation and Change, CfAB.

Some incentives reviewed relate to action to be taken before (or even well before) the age of fifty, to prevent the difficulties identified by the cohorts of the current study in the United Kingdom. For example:

– start saving from your first job in order to have enough savings by retirement age;
– take exercise, don’t smoke etc. to have a better quality of life later.

**Avenues for investigation based on correlations**

Finally, there are a certain number of correlations which open up interesting avenues for investigation but the effect of which is still unproven. For example, there is a strong correlation between volunteering and well-being for people in later life and these effects can be found when people stop and/or take up volunteering. But it is yet to be proven whether programmes promoting volunteering will have the same impact on well-being.

**Housing**

A systematic review of all of the literature available in English on the role of home adaptations in improving later life has been commissioned; it will be published in early 2017. The questions raised by the systematic review have been defined with a group of stakeholders, including practitioners and the people with lived experience, in order for the responses identified to be relevant and usable. At this stage, the CfAB is not focussing on arrangements for alternative homes, or intergenerational homes etc. in particular because 90% of people over the age of 60 are home-owners in England.
Trials

As part of CfAB’s approach and in a context of a shortage of robust evidence, trials are a means of identifying interventions with a positive impact and that can be scaled up. The CfAB’s role is to:
- identify the areas where evaluations are required;
- provide methodological support in implementing and monitoring pilot projects;
- respond to academic calls for projects to fund RCTs or quasi-experimental studies for promising interventions, offering its methodological support throughout the trial.

At this stage, the CfAB is not planning to run trials but is already helping to identify areas where there are gaps in evaluated interventions and supporting the implementation of pilot programmes.

In most cases, the centre is seeking to offer match funding in order to ensure the commitment of stakeholders who take them on. If the project lead were in a position to fund the whole pilot, the CfAB would offer technical assistance with implementation and/or evaluation and might fund the “evaluation” part of the project. The evaluations will probably be carried...
out by universities or qualified research bodies. The CfAB is also considering other approaches to promoting innovation, for example, a competition for the design of user-friendly home adaptations. As such, the CfAB is testing a number of approaches to develop innovation and increase scientific knowledge in its area.

Examples of pilot projects
There are not enough trials for return to work programmes. As such, the CfAB is setting up pilot programmes and evaluations for innovative programmes that show promise but have yet to be tested. A partnership with Greater Manchester is allowing pilot programmes to be set up in order to test such approaches, starting with a process of co-design with older jobseekers in 2016. In these pilots, the CfAB is funding the design of interventions and their evaluation. Subsequently there would be pilots on a larger scale with more robust external evaluation. The CfAB also hopes to identify pilot projects for transport adapted for older passengers, preparing for retirement and getting the most out of digital in early 2017.

Source: www.ageing-better.org.uk.
Involvement in intervention design
In order to design new interventions for evaluation, the CfAB plans on securing the participation of the people with lived experience. Randomised control trials (RCT) in the field of international development have shown that the participation of users increased the quality of the design of services, in some cases significantly.

“At the innovation end where we’re interested in identifying solutions that could work, our working assumption is that is the most effective will be user-led.”

Dan Jones – CfAB.

Dissemination and promoting uptake

CFAB’s target stakeholders are just as much the older people themselves as practitioners and policymakers. Interventions can be just be about anything from changing the behaviour of people and professional practice, the availability of housing and facilities to planning issues. The CfAB is not planning a top-down dissemination of the findings, particularly in the form of a catalogue of interventions as it believes that uptake by such a wide-ranging group of stakeholders risks being quite low.

“Our working assumption is that pushing out, broadcast, isn’t an effective method of knowledge dissemination.”

Dan Jones – CfAB.

There is a desire to target those who make decisions and implement on the ground. The partnership with Greater Manchester Combined Authority (GMCA) is emblematic, with a local approach to ageing well: it’s about experimenting with them but also supporting them to apply the evidence-base at a local level. GMCA and the centre are working together in particular on spatial planning, public spaces, new housing, local economic growth, health, employment training, and informal local facilities (for example, ensuring that shops allow older people to use their toilets).
“This local-level approach is based on the assumption that creating an environment that is adapted to older people is one of the key ways of improving their quality of life.”

Dan Jones – CfAB.

Looking ahead

The Centre for Ageing Better was set up in 2015: as such it is now ready to launch its activities, whether systematic reviews or pilot projects. In-depth research has been done into what older people need to age well, laying the groundwork and setting the direction of travel for the centre for the years to come. The CfAB approach differs from most What Works Centres. Indeed, the CfAB places a greater emphasis on promoting innovation and will probably support a range of projects, not just those with shown to be “effective” through an RCT or quasi-experimental study. The local partnership approach also differs from those of the other What Works Centres – it will be interesting to see how successful this approach is.

Lessons for France

- Understanding the needs of older people, as expressed by themselves and observed in ethnographic and longitudinal studies, is a participatory way of clearly identifying the priorities for a subject area with such a broad scope as ageing well.
- Partnering up with one or more local authorities is a promising approach both for launching pilot projects and for supporting local decision-makers to base their services on evidence.
- Finally, looking at creating the right environment and using “nudges” are interesting approaches to preventing a loss of independence, allowing people to remain in their homes etc.
This report presents a set of recommendations for the government, local authorities, the social security system etc., in tackling an ageing population. House of Lords. Ready for Ageing?. 14 March 2013. Available online: www.publications.parliament.uk/pa/ld201213/ldselect/ldpublic/140/140.pdf.

Available online: www.nice.org.uk/guidance/ng16.


Description available online: www.elsa-project.ac.uk.

Resources available online: www.ageing-better.org.uk/our-work/later-life-2015.


Resource available online: www.ageing-better.org.uk/news/people-likely-benefit-volunteeringleast-likely-get-involved.
Public health at the heart of the National Institute for Health and Care Excellence (NICE)

<table>
<thead>
<tr>
<th>At a glance</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Sector</strong></td>
<td>Evaluating medical equipment, medicines and clinical guidelines, public health care</td>
</tr>
<tr>
<td><strong>Launch date</strong></td>
<td>1999</td>
</tr>
<tr>
<td><strong>Organisational type</strong></td>
<td>An Independent public body</td>
</tr>
<tr>
<td><strong>Size of team</strong></td>
<td>600 staff</td>
</tr>
<tr>
<td><strong>Annual budget</strong></td>
<td>£62.5 million</td>
</tr>
<tr>
<td><strong>Website</strong></td>
<td><a href="http://www.nice.org.uk">www.nice.org.uk</a></td>
</tr>
</tbody>
</table>

**Context**

Issues around public health such as food, sexual behaviour etc., are not just a result of individual lifestyles but also to a large degree a result of context (family and social circumstances etc.). They can be influenced by a range of stakeholders. There is therefore a large and diverse target audience for promoting evidence-based approaches.

“Our public health guidance does much more than just list a set of recommended public health interventions. They focus much of their scope on organisations needing to work together at local level to build strategies, assess local needs, plan, select and coordinate the delivery of well targeted public health interventions.”

Alexia Tonnel, NICE.
The focus on having public health interventions that have proven to be effective goes back to the late 1990s in the United Kingdom. At that time the Department of Health set up the Health Development Agency (HDA), to conduct systematic analysis of research on the subject.

Around the same time, NICE was established by the Department of Health, with the objective of evaluating medical products and equipment, particularly in terms of cost-effectiveness. Its remit has gradually increased to include guidance for health and social care. The aim was to provide the National Health Services (NHS) and the social care sector with this information in order to quickly disseminate good practice.

HDA and NICE merged in 2005. Evidence reviews and public health guidelines are overseen by the Public Health Guidance directorate of NICE27.

Separately, in 2013, 140 organisations were merged to form Public Health England (PHE), an executive agency of the Department of Health, to coordinate the delivery of public health across the UK. Beyond health protection PHE deals with public health issues and, in particular conducts systematic literature reviews and advises on public health policies.

PHE works closely with NICE on subjects of joint interest and helped to set up the What Works Centre for Well-Being.

**Organisational structure**

NICE is a non-departmental public body funded by the Department of Health.

NICE has around 600 staff, roughly 30 of whom work on public health 2016. In the field of public health, NICE’s role involves developing guidelines for the prevention of health problems for the NHS, local authorities, the general public and the private and voluntary sectors. It also produces briefings for

27 In 2017, the public health guidance programme has now been merged with the clinical and medicine management guideline producing teams.
British What Works Centres: what lessons for evidence-based policy in France?

local authorities to advise them on the public health measures that have a positive impact and good value for money.

In order to produce its guidance, NICE draws on systematic reviews of scientific literature on the topic area, either conducted in-house or commissioned from academics. Public health studies are, for the most part, conducted in-house.

In addition, and in an increasingly coordinated way, PHE produces, in-house, scientific analysis of interventions in order to issue guidance aimed at the general public and the Department of Health, particularly in terms of public policy.

**Activities**

**Evidence synthesis**

NICE develops guidance based on systematic reviews of public health research, including impact evaluations. NICE has tackled over 60 public health topics: one of the first studies was on the obesity epidemic and its context. Topics now range from issues around alcohol consumption, to tuberculosis and AIDS prevention. The process of evaluation, decision-making and drafting guidelines are described in the manual: “Developing NICE guidelines”\(^\text{28}\).

28 www.nice.org.uk/process/pmg20/chapter/1-introduction-and-overview.
Topic referred to NICE

Scoping
- Developer drafts scope, including key issues and questions
- Stakeholders comment on draft scope
- Final scope published

Guideline development
- Review questions agreed
- Literature search
- Call for evidence from stakeholders if needed
- Evidence reviews and economic analyses prepared
- Committee discusses evidence reviews and expert testimony and develops draft recommendations

Consultation on draft guideline
- Stakeholders comment on draft guideline

Guideline revised
- Committee discusses and revises guideline in response to stakeholders' comments
- Developer writes responses to stakeholders' comments

Sign off at NICE
- Guidance Executive signs off guideline

Publication
- Confidential advance copy released to stakeholders that commented on draft guideline
- Guideline, NICE Pathway and information for the public published
- Resources to support implementation published

Updating
- Regular checks to determine if update needed
- Part or all of guideline updated according to usual process and methods

Source: www.nice.org.uk/process/pmg20/chapter/1-introduction-and-overview.
Evaluation methods – issues and limitations of the “traditional” methods

In order to evaluate the effects of an intervention, NICE considers the “best available evidence”. Ideally, this should mean well run randomised controlled trials, as for drugs or medical devices. But in the field of public health, NICE takes a wider range of types of research into account, for two main reasons:

– randomised control trials are only conducted in some arenas of public health: they are well suited as a response to specific questions, but are not suited to all types of public health intervention;
– public health issues are multidimensional depending both on the individual in question (and any relationship they have with a practitioner) and their environment. In order to tackle these questions, it is necessary to draw on sociology, anthropology, geography and political science.

“You can use RCTs for well-defined interventions, targeting a specific public, aiming for a pre-identified result on people’s health, but these cases are relatively rare.”

Mike Kelly, University of Cambridge.

As such, this has led NICE to also take into account studies based on observation, other qualitative research, interviews with experts and economic modelling. For example, in the case of research into preventing and reducing obesity, it is also about working with practitioners in sport, nutrition, education, transport, local authorities, nurses and architects.

In the field of public health, NICE works with PHE to produce guidelines. The subjects of studies are generally chosen by the Department of Health or the NHS, from a library of subjects. They are then tackled by order of priority as defined jointly in particular with PHE, the NHS and the Department of Health, mainly based on the priorities put forward by local authorities and professional bodies. Priorities are reviewed every year and can evolve according to changing needs.
The process takes around two years, from the launch to the publication of NICE guidelines, taking into account all of the time spent consulting practitioners from a range of sectors and stakeholders from civil society. The systematic review itself takes 3 to 6 weeks.

Public health subjects are tackled in partnership between NICE and PHE. PHE expertise will be engaged at different stages, from defining the review question, reviewing studies and being involved in the final validation process.

The **Public Health Advisory Committee (independent advisory committee)** – composed of health professional experts, researchers, health economists, social workers, non-profit organisation employees, civil society members – independently debate and agree all the recommendations that underpin a guideline. Their decisions are based on systematic reviews, other forms of evidence synthesis, economic evaluation and expert input.

Guidelines include research recommendations – highlighting areas of uncertainties in the evidence base that may benefit from further research.

The **Citizens Council** provides insight into the broader views of the general public. This insight is then taken into account by the committee when developing the guidelines. Moreover, all guidelines are subject to a consultation period with stakeholders. This is an opportunity for the work of the committee to be further validated by public scrutiny. This also supports the transparency of NICE’s processes.

Finally, the studies are reviewed regularly to ensure they are up to date – the frequency depends on the study.
Example: PH 53 – Weight management: lifestyle services for overweight or obese adults.
Guidelines published on the NICE site include recommendations for different local and national stakeholders – particularly to promote a locally integrated approach to the issue in partnership with local charitable organisations (such as walking groups), health practitioners, Public Health England etc. For example, one recommendation relates to the fact that those responsible for commissioning lifestyle weight management services are aware of:
– the number of adults who are overweight or obese locally;
– the range of lifestyle weight management programmes available;
– continuing professional development or training opportunities on weight management.
Guidelines are based on randomised trials on the one hand and expert reports on the other, thus combining the two approaches. Finally, this study, published in May 2014, is an updated version of the guidelines issued in 2006 and will be updated again in 2017.

PHE also produces systematic or rapid evidence reviews for specific interventions, at the request of the Department of Health. The institution can provide answers within a few months and as such does its best to adapt to the political timetable. As such, its work complements that of NICE. In particular, it has conducted studies and disseminated guidelines on the e-cigarette, the consumption of sugar, and alcohol misuse.

Trials
NICE itself does not conduct “primary research”, nor does it commission it from third parties as this does not fall within its remit and is the role of other bodies. But, based on the secondary studies conducted, it can make research recommendations. Indeed, all of the guidelines published by NICE are based on such research.
contain a chapter on “Research Recommendations”. NICE communicates these recommendations to the National Institute for Health Research, which decides on research subjects and budgets and can take into account this work.

NB. when conducting studies in certain areas of public health, it is important to bear in mind that some habits (in terms of dietary patterns, physical exercise, smoking for example) take decades to change not days or months.

**Dissemination and promoting uptake**

**NICE and PHE both play an essential role in disseminating and supporting the uptake of practices and interventions with a proven impact.**

**Reaching the general public**

PHE’s role involves disseminating prevention and treatment practices in a number of areas: e-cigarette, weight control and reduction, alcohol misuse etc. This can be through a range of means of communication such as public campaigns (such as “5 fruit and vegetables a day”) smartphone applications or web tools.\(^{(30)}\)

The means of publicising these campaigns is devised by the PHE Health Social Marketing team which relies heavily on behavioural studies, including “nudges” working alongside the PHE Behavioural Insight Team. PHE records quantitative data on the results of campaigns (number of clicks, tools downloaded etc.). The impact of the campaigns themselves have not yet been evaluated. As such, the general public can be reached via practitioners and local authorities who are targeted by PHE and NICE recommendations.

\(^{(30)}\) For example the One You programme and related apps for monitoring alcohol consumption or stopping smoking, www.nhs.uk/oneyou/apps#MkLwS5RxYyI4PH1Z.97.
Reaching practitioners and local authorities
The implementation of numerous public health initiatives is the responsibility of local authorities. NICE carries out initiatives and develops tools to reach these stakeholders.

Online tools
NICE has a very comprehensive website aimed at practitioners seeking to implement guidelines (the majority of which are clinical content), which includes:

- **all guidelines and their recommendations** by “theme”. The recommendations are presented as well as the stakeholders that may/must be involved in that area. Research methods and the studies reviewed are also accessible. Recommendations on public health are mainly around initiatives that local authorities can deliver alongside local stakeholders;
- **an interactive flowchart** to identify relevant guidance (pathways);
- **a self-assessment tool** (quality standards) which enables organisers to identify their standard of knowledge and implementation of guidelines and to work towards the next steps. For example, in public health one indicator is based on the fact that local authorities and health bodies give access to healthy food and drink in public vending machines and put them in prominent positions across all venues. This is one of the obesity recommendations.
- **access to research** sites enabling users to access research articles.

This site is accessible to the general public.
“NICE plays a role not only in the production of guidelines, but also in facilitating access to primary and secondary research as well as grey literature.”

Alexia Tonnel, NICE.

For example, NICE Evidence Search offers access to recommendations from bodies other than NICE, to bibliographical resources as well as to other forms of evaluation review, case studies and ‘grey literature’ produced by trusted organisations. These scientific studies can be used by public health professionals in conjunction with local data on the population, local studies and information supplied by PHE, for example.

PHE runs a blog dedicated to Public Health, also targeted at a professional audience that shares all public health recommendations and directs to links aimed at the general public\(^{31}\).

For example, there are detailed and user-friendly explanations of the review of evidence on e-cigarettes as a quitting aid. These provide responses to questions raised during the public debate that took place following the PHE recommendation to use e-cigarettes\(^{32}\).

\(^{31}\) publichealthmatters.blog.gov.uk.

\(^{32}\) publichealthmatters.blog.gov.uk/2015/09/18/e-cigarettes-a-public-health-response.
NICE and PHE recognise that the publication of guidelines in and of itself is not enough for stakeholders to adopt the findings.

For some time NICE invested in local support in particular with dedicated consultants who supported local stakeholders and could feed back so as to improve guidelines on implementation. NICE also set up an implementation strategy group which was influential. But it now has more of a standardised approach, with, in particular, implementation notes and the possibility of contacting NICE for help.

PHE can coordinate local dialogue through its local agencies, in order to jointly develop approaches with local stakeholders. Many recommendations have been developed by NICE and PHE on how local authorities can put guidelines into practice. These two stakeholders identify new areas for improvement for decision-makers.
Local authorities like to do things a certain way. We are all looking for the best methods for implementation. In the case of well-being, well-being measurements have already generated adjustment on local political decisions.”

Gregor Henderson, PHE.

In terms of medical stakeholders, whilst the implementation of NHS guidelines (reflecting NICE’s research) is not compulsory, practitioners must be aware of them, and there are many auditing tools and financial impact documents made available to practitioners to support the uptake of guidelines.

Adoption in public policy and law
NICE guidelines are not intended for national public policies but aimed at local stakeholders.

PHE, however, is commissioned to make recommendations bearing on national public policies. This requires a more confidential approach, to avoid spreading contradictory messages and ensure that the government is ready to share information. Some subjects, such as guidelines on e-cigarettes or the sugar tax resulted in lively public debate, both in the tabloids and scientific journals.

Looking ahead

NICE and PHE have a number priorities for their work on collating and disseminating public health evidence over the coming months and years:

- NICE has an evidence synthesis process spread over a number of years, whereas PHE is more responsive to political directives and changes in context.
- NICE and PHE work in an increasingly collaborative and integrated way on issues around public health; NICE guidance is now co-branded by PHE to promote uptake.
- NICE and PHE seek to reach decision-makers and above all local authorities, in a more effective way; they run projects jointly so as to go further in this area, particularly by developing the most suitable tools.
Lessons for France

- The public health interventions identified can only be effectively implemented through a local partnership of several actors, not just health specialists. It is worth taking this into account when evaluating programmes.
- Transparency in the methods used to review evidence, consulting several types of stakeholder and publishing all research, all help to boost the credibility of NICE and its recommendations.

"Everything we do is in the public domain, including the role of committees and their members."

Françoise Cluzeau, NICE.

It can be difficult to have randomised trials for all public health topics: it is important to be open to other approaches to have at least some indication of the policy levers.

What Works Centre for Well-Being

<table>
<thead>
<tr>
<th>At a glance</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Sector</strong></td>
</tr>
<tr>
<td><strong>Launch date</strong></td>
</tr>
<tr>
<td><strong>Organisational type</strong></td>
</tr>
<tr>
<td><strong>Size of team</strong></td>
</tr>
<tr>
<td><strong>Annual budget</strong></td>
</tr>
<tr>
<td><strong>Website</strong></td>
</tr>
</tbody>
</table>
This What Works Centre was set up very recently and is funded by the ESRC, the Department of Health and seven other government departments. It published its first systematic literature reviews in September/October 2016.

This work is based on defining well-being and approaches to measuring it. The centre uses the official definition for well-being in the United Kingdom: “Well-being, put simply, is about ‘how we are doing’ as individuals, communities and as a nation and how sustainable this is for the future”. Personal well-being is a particularly important dimension for the centre and is defined as how satisfied we are with our lives, our sense that what we do in life is worthwhile, our day to day emotional experiences (happiness and anxiety) and our wider mental well-being.

The centre acts as a hub. It relies on a central team of 5 staff members to ‘translate’ the findings of scientific studies for target stakeholders. This team is made up of a Director, a Head of Evidence, a Programme Manager and a Head of Communications. The centre also has 53 researchers which it works with, in particular in conducting evidence reviews.

The centre has been commissioned to conduct systematic literature reviews on these topics:

- work and learning;
- community well-being;
- culture and sport;
- well-Being across the life course.

The centre adapts its review methods and inclusion criteria according to the nature and quality of available research. In its initial phase of work, the team is focussing as much as possible on systematic literature reviews where these are available and on ‘translating’ them into results that can be easily understood by a wide audience. It relies on an evidence panel to ensure the quality of the reviews produced and scores the evaluations using the GRADE model (also adopted by NICE and Cochrane collaboration).

33 whatworkswellbeing.files.wordpress.com/2016/02/what-works-wellbeing-methods-guide-july-2016.pdf.
In order to produce relevant studies, the centre also draws on qualitative research, particularly in order to understand why there has or has not been an impact and why the impacts are experienced differently by different people or in different contexts.

Finally, it should be emphasised that well-being is a very broad, multi-disciplinary area which is closely linked to, or even an integral part of, public health issues. As such it is worth going beyond the research already catalogued in peer reviewed journals, for example by speaking to experts, following “Ariadne’s thread” of references in the texts cited, carrying out research by hand in the various publications etc.

The next stage will be to consolidate communication resources. After two years of preparing the target audiences for evidence-based results and drawing inspiration from existing practices, the centre is considering a communication campaign to target specifically each stakeholder: government, local government, private sector, etc.

Finally, the centre draws on current studies to identify the most effective approaches for the adoption of scientific findings. Indeed, the centre led on the production of a review of what works in research use: ‘the science of using scientific knowledge’.

34 whatworkswellbeing.org/evidence-program/the-science-of-using-science-knowledge.
Perspectives
France Stratégie

This report, produced by Ansa with the support of its partners, offers a useful contribution to thinking on the mainstreaming of evaluation in France.

The creation of nine What Works Centres in the United Kingdom with the support of the Cabinet Office, was a response to three major priorities: promoting impact evaluations for interventions and programmes in the education, employment, social and health fields; synthesising and disseminating findings; supporting practitioners and local policymakers in putting them into practice more effectively.

These three priorities mirror several of the France Stratégie’s own analyses and recommendations. They reflect recent changes in public policy and services.

The effectiveness and efficiency of public programmes are a major challenge in most mature democracies where a high level of public spending goes alongside a certain disillusionment with public institutions. The state of public opinion supports the demand for evaluation. In France, there is a gap between the interest in impact evaluation and its actual application. The dominant approach to evaluation remains the evaluation of public policies, which relates mainly to the relevance, coherence and implementation of policies. The utility of this approach is not contested and its benefits are irreplaceable. Yet, an increase in the number and quality of impact evaluations is needed. Their specificity lies in their ability to determine whether a practice has a measurable effect for its beneficiaries. France Stratégie sought to better understand the aims and methods in


Interest in approaches and structures geared towards identifying programmes that “work” for a given issue is driven by other factors behind the transformation of public services. The areas covered by the What Works Centres are not down to chance. The social state, historically reliant on social insurance, transfer mechanisms and benefits, is today shifting the focus of its interventions towards services, programmes to support investment in education and training, and also maintaining health and personal independence. This shift, sometimes described as a social investment strategy, operates in the presence of new social challenges related to tackling inequality (early school attainment gap, gender inequality, health inequality etc.), combating poverty, unemployment and exclusion, as well as the ageing population. This shift brings with it innovation, inventiveness and complexity as well as the emergence of an unprecedented demand to identify “relevant” programmes and therefore for evaluation. Where previously the response to a risk was a corresponding benefit and the effect was evaluated by the rate of coverage and living standards, now there are tailored services, impact evaluation and even the calculation of savings³⁷.

Finally, approaches and structures that take their inspiration from the What Works Centres should be considered in the context of public policy that covers a broader scope than the state and its agencies, includes social security bodies and local authorities, incorporates the European level and contributed to by stakeholders from the third sector. Joined up, decentralised and devolved, public action has developed the room for manoeuvre of local managers and practitioners, established

the right of local authorities to experiment, increased the share of project funding. These changes justify tools to pool knowledge, share good practice or proven interventions, interfaces that make it easier to promote and share them. The state has an essential role in setting up such initiatives or structures that are a public good. The practical processes for implementing or raising the profile of such approaches or structures, the detail of their remit and their policy levers, the nature of the support the State may offer them and the involvement of government departments still require investigation. These different aspects will be investigated in more detail as part of collaborative work between France Stratégie and Ansa in the coming months.

Marine Boisson-Cohen, lead expert and deputy director, Département Société et Politiques sociales (Society and Social Policy Department).

38 Dominique Bureau and Marie-Cécile Naves (coord.), Quelle action publique pour demain?, France Stratégie, 2015.
Secrétariat général pour la modernisation de l’action publique (SGMAP)

The report produced by Ansa responds to the initial objective which was to document the growing interest in “evidence-based public policy”, an approach that is still little-known in France.

The case studies of the British What Works Centres, at the heart of the report (not to be confused with a research or evaluation exercise), offers interesting insight, in particular into the methods and good practices developed by these structures.

The use of systematic reviews synthesising the available evidence of what works presents undeniable benefits, although it is clear on reading the report that they can be awkward to implement owing to:

– the lack of knowledge and experimentation in certain areas of public policy,
– the complexity around the transfer of ‘effective’ interventions from one context to another.

The intensification of impact evaluations without a doubt represents an interesting avenue for promoting public decisions informed by evidence. Yet, the difficulty, highlighted in the report, of using this type of quantitative method, is a reason to remain open to other types of evidence that are more descriptive and qualitative in nature, and ultimately is an argument for using “mixed” methods.

Finally, the SGMAP has noted the particular care taken by What Works Centres to disseminate their findings to practitioners who they support through the uptake and implementation phase. Targeting frontline professionals and promoting the use of innovative methods in local areas, are worth pursuing in more detail in France in the wake of thinking by the SGMAP and the CGET on “scaling up.”
The SGMAP shares the wariness mentioned at the end of the study on the effectiveness of the “What Works” models as a whole. Without the benefit of hindsight and evaluation, such effectiveness is not yet proven. Were it to be so, this effectiveness should be viewed in relation to the financial resources made available for the development of these organisations when assessing the model’s cost-effectiveness. As things stand, we suggest experimenting the methods the What Works Centres have developed in the French context rather than adopting the model wholesale.

The brief overview of evaluation in France offers an interpretation that is mostly in line with SGMAP thinking. In France, questions remain as to the feasibility and use of impact evaluations. The number of impact evaluations conducted in France remains small and research from other countries is rarely transferable as is. These studies are rarely disseminated beyond academic circles and are met with little interest on the part of decision-makers.

The study’s findings, far from ending the debate on the role and future of evidence-based policies in France, are the starting point for a movement towards policies based, if not on evidence, at least on knowledge. The SGMAP is in favour of further our thinking by testing approaches or structures aimed at supporting evidence-based public policy making. This is not the place to settle questions around the appropriateness and feasibility of creating dedicated structures or networks but to move forward by capitalising on the achievements of a number of existing initiatives in order to translate this approach into action. The SGMAP would also like this work to be promoted in order to feed into the debate around the role of evaluations and the dissemination of their findings to stakeholders responsible for public policy making and delivery. Synergies should also be identified between the study and the findings of the “meta-evaluation” conducted by the SGMAP of its overall evaluation of public policies. This evaluation recommends adapting approaches more closely to policy-makers’ needs. It highlights that defining the expected use of an evaluation in advance, at the time recommendations are
drawn up (law, action plans, experimentation etc.) is a necessary condition if its findings are to be taken into account. On this point, there is a clear link with then What Works model.

Local initiatives, embodied in particular by local public innovation labs, are driving interest in disseminating proven interventions. This dissemination of solutions must be organised, in particular by the State, in order to be of maximum benefit to our fellow citizens.

The SGMAP intends to take part in this collective exercise, by working in the coming months towards building the operational basis of such an approach to public policy.
British What Works Centres: what lessons for evidence-based policy in France?

Commissariat général à l’égalité des territoires (CGET)

The CGET’s interest in British What Works Centres covers two separate issues: an abundance of scientific research which is now becoming difficult to rank and synthesise; a recurrent difficulty in evaluating public policy programmes and ensuring these evaluations result in action. As Ansa’s report shows very clearly, the creation of the What Works Centres offers very interesting responses to these two issues and represents a genuine source of inspiration for French public policy actors.

To make the most of this, it is important to consider the What Works Centres as a device for evidence synthesis (the main tool of which is the systematic review), dissemination of meaningful solutions (to practitioners) and methodological support (social experiments mainly).

The interest of focusing on the centre’s as organisations is that it does not reduce the legitimacy of the What Works Centres to a single method of evaluation (impact evaluation). It is on this condition that their inclusion in the French landscape of knowledge production and evaluation seems to us to be possible and desirable.

We think that in France there is a lack of structures with the ability to organise (by major thematic areas) systematic monitoring of scientific outputs with the aim of comparing independent and robust syntheses (that is taking responsibility for ranking evidence and the quality of arguments) designed for the purposes of evaluation and/or as a decision-making tool (i.e. in a format that shows options and alternatives).

In this respect, impact evaluations (as defined and used by the British What Works Centres) are just one source of information among others. Nevertheless, we place significant importance on them results because we think that this type of evaluation is useful and must be rolled out more systematically. They demand a level of scientific ambition that is
For all of that, let us not overstate their ability to mechanically state the facts and to be naturally conclusive. Given the many simplifications resulting from their methodology (and their focus on effectiveness for the purposes of evaluation), we cannot consider them the only way of evaluating a public policy programme and, particularly, of instigating a transformation of public action (which must remain, and let us not lose sight of this, the ultimate aim of the What Works Centres).39

Put simply, French-style What Works Centres from systematic literature reviews to public debate, should help to showcase and make accessible a range of research and opinions. We can therefore ask ourselves whether methods such as consensus conferences, should ultimately be one of the key approaches of these What Works Centres?

---

39 The aim here is not to highlight the limitation or flaws of impact evaluations (and above all of the political arguments behind them). It simply a case of reiterating that for some they are part of a rationale that replaces judgement with measurement and deliberation with evidence: in many respects they tend to reduce the issues around understanding to explanation and, in so doing, to overstate the role of explanation in transitioning to action (the, to say the least, erratic relationship between the use of impact evaluations and roll-out or scaling up is just one of many examples of this). This confusion is often maintained by reducing the “overall impact of a programme” (an expression which, let us be clear, is a means of justifying many things and/or to not say much about it) to its direct and immediate impact on an unequivocal aspect of the beneficiary. This arises from an epistemological ambition - that is never satisfied - to treat a hypothesis (or causality) in strict isolation. This often results in impact evaluations showing little interest in the organisational, financial and political characteristics of public policies, not to mention that this focus on the beneficiary is ultimately not without consequences in terms of ways of thinking about social justice, defining social rights and conditions for accessing them (theme of social investment). Added to this are the frequent difficulties in implementing them, interpreting the findings and the obvious black box effect when looking at understanding the conditions in which the findings were obtained. So many complex political and scientific issues that are only crystallised by impact evaluations and that would, of course, benefit from being explored in a more systematic way (even if some of them may evolve quickly with the emergence of “big data”),
Ultimately, for the CGET, this exercise has the benefit of breathing life back into the debate, which has fizzled out to a degree, about the contribution of researchers to evaluation and about the means of using evaluation for operational purposes. As such, we hope that it is communicated to the relevant stakeholders very soon (in the form of comparative perspectives and why not a feasibility study).

Emmanuel Dupont.
Santé publique France

Santé publique France (Public Health France) supported the project of Ansa and its partners as it closely reflects the agency’s priorities, in terms of contributing to the development of local health prevention and promotion initiatives that have proven effectiveness. There are two possible options. The first is to import, adopt and evaluate in the French context interventions that are effective in other countries and other contexts. The second is to innovate, experiment and/or evaluate interventions created in France.

It is in this way, for example, that Santé publique France adapted the Public Health England campaign “Stoptober” the effectiveness of which had been proven with an increase in attempts at quitting across the Channel (Brown, Kotz et al. 2014). The French version was put in place last November under the name “Mois sans Tabac” and the results of the evaluation of the French campaign will be known at a later date. Santé publique France is also involved in an ambitious experimental project led by CNAMTS with the support of the DGS (Department of Health), to prevent type 2 diabetes. This trial has been developed based on experiences in the US (Albright and Gregg 2013), Finland and China in this area and has cut new diabetes cases by over 50% through an intensive programme of assistance in improving diet and an increase in physical activity.

Santé publique France is also involved, alongside the Mission interministérielle de lutte contre les conduites addictives (Cross-government unit for combating addictive behaviour), in evaluating grassroots projects as part of the Commission interministérielle de prévention des conduites addictives (Cross-government commission for combating addictive behaviour – CIPCA). A call for applications has been launched as a result of which 80 grassroots projects have been identified. Five projects have
been selected by a cross-government panel and financial and human resources have been invested by the MILDECA to describe and evaluate these five projects. This work is currently under way.

Beyond Santé publique France’s involvement in preventative healthcare, which is its specific policy area, the importance of this partnership lies in the fact that there are extremely close links between economic, social and health matters.

As such, effective initiatives in health may be derived from the economic or social sectors and vice versa. For example, access to housing can improve health. Kellam has shown that improving the classroom environment in the first years of primary school has effects on long-term health, in particular reducing smoking 15 years later (Kellam, Mackenzie et al. 2011). Parenting support programmes have shown reductions in the consumption of psychoactive substances (Spoth, Redmond et al. 2001). Moffitt and his colleagues have shown, through the Dunedin birth cohort in New Zealand, links between the phenomena observed in childhood and a set of situations that fall within both the social and health sectors (Moffitt, Arseneault et al. 2011).

As such, beyond the benefit of sharing methods for synthesising literature, experimentation and dissemination which should not be neglected, this partnership is above all important because of the entanglement and interactions between the areas of intervention of the project partners.
Caisse nationale des allocations familiales (Cnaf)

The Caisse nationale des allocations familiales (French family benefit agency) wanted to support the study conducted by the Ansa for two main reasons:
– firstly, to feed into the institution’s strategic vision, one of the main pillars of which is social investment (or prevention and early intervention). This should prove a valuable and tangible way of setting the public policy agenda and ranking programmes and interventions according to their economic return and in terms of human capital against others. The study of What Works Centres can help to make this concept a reality for the Family branch in particular as part of thinking on creating a social investment fund which would conduct evaluations;
– the Cnaf’s remit also includes the evaluation of public policies led by its statistical and research division (Direction des statistiques, des études et de la recherche). As such, the second reason relates, on the one hand, to a better understanding of impact evaluations and, on the other, to the way in which - in general terms - the findings of evaluation can be better shared with public stakeholders.

The first benefit of the report is methodological. The substantial literature review exercise carried out by the centres is a noteworthy aspect which we feel is relevant to France. Each centre conducts this exercise in their area of expertise, resulting in a shared library that can contribute both to scientific and political thinking for frontline stakeholders. The organisation of evaluation by thematic area, particularly in terms of early childhood, young people or support for parenthood is another contribution of the British centres.

The impact evaluation methodologies that can identify a causal relationship and as such determine the effectiveness of an intervention are not developed enough in France. Yet, they would appear to be one of the useful decision-making tools to justify and intervention, spending
incurred but also in terms of accountability for public policy to citizens and beneficiaries. The study does emphasise that these methodologies lend themselves better to local trials and interventions than to national policies. This is why, our interest is all the more keen as the social investment fund, we wish to create would have precisely the powers to fund trials and evaluate local programmes. A particularly striking aspect of the evidence put forward in the study is the way that this method complements other more traditional methods. The benefit of this approach and its scientific robustness is that it offers a means of evaluating the effectiveness of an intervention and, combined where necessary with other types of evaluations, it also becomes possible to explain why and how the intervention can be improved. As such, it is also the improvement of existing services that is at stake.

In France we lack a hub to organise dialogue between researchers and decision-makers. The work that might be organised as part of our social investment project might draw inspiration from the British centres, as structures dedicated to knowledge transfer and stakeholder support. These centres produce three types of tools to support the appropriation of research findings by stakeholders: libraries and systematic reviews; self-assessment tools; intervention descriptions or guides. Even if the effectiveness of the What Works Centres in terms of adoption has yet to be proven, they offer interesting avenues to be explored in making research findings accessible in a form that can be easily appropriated by the various stakeholders. The benefit lies in targeting in particular frontline professionals who are familiar with the implementation of programmes. The issue is around being able to generate a demand for evaluation or even experimentation in order to spur the creation of centres in France. The report offers valuable food for thought on conditions for roll-out. The independence of structures leading on evaluations is essential in maintaining research standards. Financial and human resources to carry out literature reviews, quality evaluations and the dissemination of findings to stakeholders are also needed. A partnership with different institutions to pool resources could prove to be one avenue to make such centres a reality.
British What Works Centres: what lessons for evidence-based policy in France?

A robust evaluation, on the one hand, and reports on frontline experiences on the other, seem to us to be the main benefits of the What Works Centres. On these two points, the study conducted by Ansa will feed into Cnaf’s thinking on the creation of a social investment fund. We also suggest widespread circulation of this report to encourage the setting up of an institutional support network for similar evaluation measures.

Sandrine Dauphin.
Caisse nationale de solidarité pour l’autonomie (CNSA)

By conducting a study into the British What Works Centres, Ansa is feeding into an important debate on issues and methods that contribute to more evidence-informed public policy.

The CNSA – the French benefits agency for older and disabled people – is very directly affected by this compelling issue of evaluation. It is helping to drive public policies for which social demand is undergoing rapid transformation, as much in the field of disability as in the loss of independence of the elderly. Policy responses are changing and there is a real need for evaluation both of existing policies and emerging responses. By way of illustration, two areas in which Ansa’s thinking can be implemented jump out from the point of view of the CNSA.

The first is that of support for innovation. In its role as an agency, the CNSA is regularly asked about or comes to support programmes or interventions described as “innovative”. As is the case of many other stakeholders, it lacks a method and benchmark to establish where these initiatives fit in in terms of existing practices, to evaluate their relevance in principle. When it wishes to promote evaluation initiatives among grantees, it notes that they share this lack of methodological frameworks.

Finally, in this dialogue with the grantees the issue that emerges very quickly is: who must evaluate? In other words, with whom lies the “burden of proof”? The second one relates to the so-called “non-medical” interventions. Social care is based on a wide range of interventions aimed at the elderly who have lost their independence or people with disabilities, which are only
partially evaluated and there is no framework for these practices as exists in the medical field. Without condemning the legitimate need of frontline stakeholders to invent new practices to offer psychological support, we need to come together to find the evidence that they are good practices and support the dissemination of those that are proven to be effective. And the question arises regularly: is it just about importing programmes proven in other countries? We need to get better at assessing the likely impact of such interventions in the French context, both to redefine support practices and rethink funding frameworks, as today some interventions are funded on a discretionary basis due to a lack of evidence, or as part of daily living assistance due to a lack of certainty as to the nature of the impact.

By proposing to explore a tailored transfer of British What Works Centres to France, Ansa provides important insight and subjects for examination for similar programmes for validating practices, supporting scale up or transposition of practices proven overseas, support for trials with the aim of measuring impact and dissemination of best practices.

The thematic approach of the British centres suggests that public policy areas more or less able to organise this evidence ranking exercise. Cross-cutting analysis of the various centres has already enabled Ansa to highlight the success factors of this approach. The analysis of French resources that can be drawn on in research and evaluation suggests that we are not starting from scratch, far from it. However, everything remains to be done in terms of the systematic setting of evidence standards, the ranking of evidence and, above all, of dissemination.

Building on this first, highly stimulating, report, phase 2 of Ansa’s project should be a way of considering in a very pragmatic way what these thematic evidence centres could be.
In terms of the CNSA, it is necessary to choose between a public remit relating to the support for the elderly or people with disabilities. And, given the British experience, we have not failed to observe that the theme should be narrow.
Graphic Design: Camille Esayan camilleesayan.fr.
This report was produced by Caroline Allard and Ben Rickey, project managers at the Agence nouvelle des solidarités actives in partnership with France Stratégie, the Secrétariat général pour la modernisation de l’action publique (SGMAP), the Commissariat général à l’égalité des territoires (CGET), Santé publique France, the Caisse nationale des allocations familiales (Cnaf) and the Caisse nationale de solidarité pour l’autonomie (CNSA).

Ansa_Aso
www.solidarites-actives.com
Agence nouvelle des solidarités actives
28, rue du Sentier 75002 Paris
Tél.: 00 33 1 43 48 65 24

British What Works Centres: what lessons for evidence-based policy in France?