

What Works Centres britanniques : quels enseignements pour des politiques fondées sur la preuve en France ?

Mars 2017

En partenariat avec



What Works Centres britanniques : quels enseignements pour des politiques fondées sur la preuve en France ?

Mars 2017

Ce rapport a été rédigé par Caroline Allard et Ben Rickey, responsables de projet à l'Agence nouvelle des solidarités actives en partenariat avec France Stratégie, le Secrétariat général pour la modernisation de l'action publique (SGMAP), le Commissariat général à l'égalité des territoires (CGET), Santé publique France, la Caisse nationale des allocations familiales (Cnaf) et la Caisse nationale de solidarité pour l'autonomie (CNSA).

En partenariat avec



cget



What Works Centres britanniques : quels enseignements pour des politiques fondées sur la preuve en France ?

Les partenaires de l'étude	4
Synthèse	10
Le rapport	14
Introduction	15
Un enjeu partagé au niveau international : des politiques publiques davantage « fondées sur la preuve »	22
Des politiques publiques « fondées sur la preuve »	22
Produire un niveau de preuve de qualité : les évaluations d'impact et les synthèses de la recherche	23
<i>L'évaluation d'impact : vers des niveaux de preuve plus élevés</i>	25
<i>Évaluations socio-économiques : une mesure de l'efficacité de dispositifs</i>	30
<i>Les revues systématiques de littérature</i>	31
<i>Quand l'État favorise l'expérimentation avec évaluation d'impact</i>	33
Réussir l'utilisation des connaissances par les professionnels et les décideurs politiques à tous les niveaux	35
<i>Les enjeux du « transfert de la connaissance »</i>	36
<i>Mettre les résultats d'évaluation d'impact et de revues systématiques de littérature à disposition sous forme de recommandations et de pratiques</i>	38
<i>Favoriser une culture de la preuve au niveau des décideurs et professionnels</i>	40
<i>Créer des organisations pour renforcer le lien entre évaluation d'impact et politiques publiques</i>	41
Les What Works Centres britanniques : une réponse ?	45
Présentation des What Works Centres	45
<i>Les principes généraux</i>	45
<i>Présentation synthétique</i>	47
<i>Contexte de leur émergence</i>	49
<i>Gouvernance des centres</i>	52
L'action des centres	52
<i>Synthétiser la recherche existante</i>	52
<i>Diffuser et accompagner l'adoption</i>	61
<i>Évaluer l'impact d'innovations</i>	70
Les enseignements clés sur les What Works Centres	72
Quelle culture de l'évaluation et de la preuve pour les politiques publiques en France ?	76
Un dispositif administratif orienté vers une compréhension fine des problématiques	76
L'évaluation des politiques publiques	78
L'émergence de l'évaluation d'impact et de l'expérimentation sociale	82
<i>Un développement rapide de ces nouvelles méthodes</i>	82
<i>Un mouvement qui se poursuit mais une relativement faible production d'évaluations d'impact</i>	82

Les évaluations d'impact comme outil d'aide à la décision	89
<i>Les expérimentations et évaluations d'impact éclairent parfois la décision publique</i>	89
<i>Les revues systématiques sont peu développées dans les politiques sociales en France</i>	91
Une action publique globalement peu informée par les évaluations d'impact	92
<i>Des décideurs politiques peu éclairés par les évaluations d'impact</i>	92
<i>Les pratiques « efficaces » sont difficilement accessibles aux professionnels</i>	94
Un accompagnement au changement d'échelle peu développé	94
Les enseignements clés sur l'évaluation et la preuve en France	96
Les What Works Centres : quels enseignements pour la France ?	99
Méthodes des What Works Centres	99
<i>Intérêts</i>	101
<i>Limites</i>	102
Les What Works Centres en tant que dispositif	102
<i>Intérêts</i>	102
<i>Limites</i>	103
Le contexte français se prête-t-il au modèle « What Works » ?	104
Propositions	104
Proposition 1 : Étudier l'opportunité de reconnaître ou de créer des démarches ou des structures d'inspiration « What Works » en France	106
Proposition 2 : Créer une alliance autour des politiques fondées sur la preuve	111
Études de cas de 5 What Works Centres	113
Early Intervention Foundation	114
Education Endowment Foundation	126
What Works Centre for Local Economic Growth	138
Centre for Ageing Better	151
National Institute for Health and Care Excellence (NICE)	165
Points de vue	181
France Stratégie	182
Secrétariat général pour la modernisation de l'action publique (SGMAP)	185
Commissariat général à l'égalité des territoires (CGET)	188
Santé publique France	191
Caisse nationale des allocations familiales (Cnaf)	194
Caisse nationale de solidarité pour l'autonomie (CNSA)	197

Les partenaires de l'étude

- **France Stratégie**, organisme de réflexion, d'expertise et de concertation, rattaché au Premier ministre.
- **Le Secrétariat général pour la modernisation de l'action publique** (SGMAP), qui accompagne les administrations dans la conduite de politiques de transformation à travers la mise en place de démarches innovantes, la prise en compte des usagers, l'évaluation des politiques publiques, la diffusion du numérique et la promotion d'une action publique plus transparente et collaborative.
- **Le Commissariat général à l'égalité des territoires** (CGET), chargé de concevoir et de mettre en œuvre la politique nationale d'égalité des territoires et d'en assurer le suivi et la coordination interministérielle.
- **Santé publique France** est l'agence nationale de santé publique. Forte de ses compétences en épidémiologie, en prévention et dans l'intervention auprès des publics, l'agence couvre, par ses missions, un champ large d'activités, de la connaissance à l'action.
- **La Caisse nationale des allocations familiales** (Cnaf) pilote la branche famille de la Sécurité sociale et est présente sur tout le territoire grâce à son réseau départemental de caisses d'allocations familiales (Caf).
- **La Caisse nationale de solidarité pour l'autonomie** (CNSA) finance et accompagne la mise en œuvre des politiques d'autonomie des personnes âgées dépendantes et des personnes handicapées.

Membres du comité de pilotage

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

Personnes rencontrées et remerciements

L'Ansa tient à remercier toutes les personnes ayant contribué à cette étude. Cependant, les auteurs du rapport sont les seuls responsables des propos avancés.

■ Personnes interviewées au Royaume-Uni

Alliance for Useful Evidence

Pippa Coutts
*Alliance for Useful Evidence
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*

Economic and Social Research Council

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 Nevill
Evaluation Manager

Education Endowment Foundation

Emily Yeomans
Grants Manager

Education Endowment Foundation

Danielle Mason
*Head of Research
and Publications*

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*

What Works Centres britanniques : **quels enseignements pour des politiques fondées sur la preuve en 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
Head of the Alliance
for Useful Evidence

University of Cambridge

Mike Kelly
Former Head of Public
Health Centre at NICE

**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 Local
Economic Growth**

Max Nathan
Research Fellow
– Urban Programme

**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**

Imperial College London
Françoise Cluzeau
Associate Director
Global Health & Development
Group ex NICE

■ Personnes interviewées en France

**Assistance publique
-Hôpitaux de Paris**

Martin Hirsch
Directeur

**Caisse nationale
d'allocations familiales**

Bernard Tapie
Directeur des
statistiques, des études
et de la recherche

**Conseil national d'évaluation
du système scolaire**

Nathalie Mons
Présidente du Cnesco

Cour des comptes

Jérôme Filippini
Secrétaire général

Cour des comptes

Barbara Falk
Secrétaire générale adjointe

Cour des comptes

Laetitia Buffet
Chargée de mission auprès
de Premier Président

**Délégation interministérielle
pour l'hébergement
et l'accès au logement**

Pascale Estecahandy

Coordinatrice nationale
Un chez-soi d'abord

**Direction de la recherche,
des études, de l'évaluation
et des statistiques**

Franck von Lennepe

**Direction de l'animation
de la recherche, des études
et des statistiques**

Philippe Zamora
Sous-Directeur Suivi
et Évaluation des Politiques
d'Emploi et de Formation

**Direction de l'évaluation,
de la prospective
et de la performance
du ministère
de l'Éducation nationale**
Olivier Cosnefroy
Chef de bureau
de l'évaluation des
actions éducatives et des
expérimentations

**Direction de l'évaluation,
de la prospective
et de la performance
du ministère
de l'Éducation nationale**
Daniel Auverlot
Sous-directeur
des évaluations
et de la performance
scolaire

**Direction de l'évaluation,
de la prospective
et de la performance
du ministère
de l'Éducation nationale**
Fabienne Rosenwald
Directrice de l'évaluation,
de la prospective
et de la performance

**Direction générale
de l'emploi de la formation
professionnelle**
Tristan Klein
Chef-adjoint Mission appui
au déploiement des
programmes du Fonds
social européen

Direction générale du Trésor
Michel Houdebine
Chef économiste

École d'Économie de Paris
Marc Gurgand
Directeur scientifique
J-PAL Europe

**Direction de la jeunesse,
de l'éducation populaire
et de la vie associative**
Jean-Benoît Dujol
Directeur

**Institut national
de la jeunesse
et de l'éducation populaire**
Thibaut de Saint Pol
Directeur

**France Stratégie
(Commissariat
général à la stratégie
et à la prospective)**
Fabrice Lenglard
Commissaire général adjoint

**France Stratégie
(Commissariat
général à la stratégie
et à la prospective)**
Rosenn Desplatz
Responsable en évaluation
des politiques publiques

**Inspection générale
des affaires sociales**
Stéphane Paul
Inspecteur général

**Institut des politiques
publiques**
Antoine Bozio
Directeur

**Laboratoire interdisciplinaire
d'évaluation des politiques
publiques – LIEPP**
Bruno Palier

Directeur de recherche
CNRS, co-directeur du LIEPP

**Laboratoire interdisciplinaire
d'évaluation des politiques
publiques – LIEPP**
Juliette Seban
Secrétaire générale

**Ministère de l'Éducation
nationale**
Olivier Noblecourt
Directeur adjoint du cabinet

Plein Sens
Caroline Moriceau
Directrice d'études

**Secrétariat général
pour la modernisation
de l'action publique**
Laure de la Bretèche
Secrétaire générale pour
la modernisation de l'action
publique

**Secrétariat général
pour la modernisation
de l'action publique**
Clement Lacouette-Fougère
Chargé de mission
évaluation

**Société française
de l'évaluation**
Virginie Besrest
Membre du CA

Université Marne-la-Vallée
Yannick L'Horty
Directeur de la
fédération de recherche
« Travail, Emploi
et Politiques publiques »

What Works Centres britanniques : **quels enseignements pour des politiques fondées sur la preuve en France ?**

■ Participants à l'atelier du 4 octobre 2016

Caisse nationale d'allocations familiales (Cnaf)

Arthur Heim
Chargé de mission sur l'investissement social

Caisse nationale de solidarité pour l'autonomie (CNSA)

Julie Micheau
Directrice scientifique

Commissariat général à l'égalité des territoires (CGET)

Emmanuel Dupont
Responsable de l'animation scientifique et de l'innovation

Direction de l'évaluation, de la prospective et de la performance

du Ministère de l'Éducation nationale

Pascal Bessonneau
Statisticien

Fonds d'expérimentations pour la jeunesse (Fej)

Malika Kacimi
Chef de Mission

Institut français de l'éducation

Catherine Pérotin
Directrice adjointe

Institut français de l'éducation

Marie-Claire Thomas

Laboratoire interdisciplinaire d'évaluation des politiques publiques (LIEPP)

Juliette Seban
Secrétaire générale

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
Conseiller Lutte contre les exclusions

Direction générale de la cohésion sociale (DGCS)

Michèle Arnaudies
DGCS/CT Innovation sociale

■ Participants aux ateliers thématiques

Cnaf (sur le sujet de la petite enfance)

Sandrine Dauphin
Responsable du département de l'animation de la recherche et du réseau des chargés d'études

Arthur Heim

Danielle Boyer
Responsable de l'Observatoire national de la petite enfance

Laurent Ortalda
Responsable petite enfance au Dep

CNSA

Julie Micheau
Directrice scientifique

Santé Publique France

Pierre Arwidson
Directeur adjoint de la prévention et de la promotion de la santé

Isabelle Vazeilles

CGET

Emmanuel Dupont
Responsable
de l'animation scientifique
et de l'innovation

Catherine Prévost
Chargée de mission au Pôle
ressources et valorisations

Brigitte Baccàini
Sous-directrice
de l'observation et des
analyses statistiques

Damien Kacza
Chargé de mission
évaluation au Bureau
de l'évaluation
et de l'expérimentation

Sylviane Le Guyader
Chargée de mission études
au Bureau de la prospective
et des études

Charlotte Le Nué
Responsable
de la veille territoriale

Synthèse

La recherche de solutions efficaces à des difficultés sociales, sanitaires et économiques persistantes alimente un intérêt croissant pour une approche de « politique publique fondée sur la preuve ». Ces difficultés prennent diverses formes, des inégalités sociales dès la petite enfance au décrochage scolaire, du chômage aux inégalités économiques entre territoires, de comportements à risque à la perte d'autonomie des personnes âgées.

En France, l'expérimentation sociale avec évaluation d'impact¹, en particulier, a connu un certain essor depuis une dizaine d'années.

Cette approche est venue compléter la riche tradition française d'évaluation des politiques publiques en apportant une promesse nouvelle : mesurer l'effet causal des politiques publiques sur leurs bénéficiaires. Cet élan a suscité l'espoir de disposer de méthodes rigoureuses qui permettraient d'allouer les ressources aux dispositifs dont les effets auraient été démontrés. Toutefois, lui ont succédé des interrogations relatives à la faisabilité et à l'utilisation des évaluations d'impact, notamment pour les politiques sociales et économiques. Le nombre de ces expérimentations reste relativement faible. S'il existe des éléments de preuve d'impact sur certaines interventions à l'étranger, on manque de mesure d'impact probante dans le contexte français, les évaluations réalisées à l'étranger n'étant pas toujours transposables. Quand ces évaluations d'impact existent, peu de décideurs et de professionnels ont le réflexe de les étudier. Ceux qui souhaiteraient s'emparer des enseignements des évaluations peinent à le faire, puisque leurs résultats sont rarement diffusés sous une forme largement appropriable. Par conséquent, ces évaluations d'impact sont peu connues des praticiens et des décideurs publics et peu mobilisées pour éclairer les décisions politiques ou pour informer les pratiques sur le terrain.

¹ L'évaluation d'impact est entendue ici au sens d'une évaluation qui cherche à mesurer un impact causal, en identifiant le contrefactuel. Voir définition page 25.

Pour identifier des démarches à même de renforcer le lien entre l'évaluation d'impact et les politiques publiques, l'Agence nouvelle des solidarités actives (Ansa) et ses partenaires, France Stratégie, le Secrétariat général pour la modernisation de l'action publique (SGMAP), le Commissariat général à l'égalité des territoires (CGET), Santé publique France, la Caisse nationale des allocations familiales (Cnaf) et la Caisse nationale de solidarité pour l'autonomie (CNSA), ont mené en 2016 une étude du fonctionnement des What Works Centres britanniques.

Partant du principe qu'il faut du temps et des moyens pour amener « la preuve » au cœur de l'action publique, le gouvernement britannique a soutenu ces dernières années le développement de structures dédiées au transfert de la connaissance – notamment des évaluations d'impact – dans différents champs de l'action publique : les What Works Centres.

Les neuf What Works Centres britanniques ont pour objectif principal de contribuer à l'appropriation des enseignements des évaluations d'impact par les professionnels et les décideurs politiques au Royaume-Uni. Dans des domaines aussi divers que l'éducation des enfants défavorisés, la prévention précoce ou le développement économique local, les interventions étudiées portent d'abord sur des pratiques de terrain. Grâce à leurs revues systématiques, ces centres sont désormais reconnus par de nombreux acteurs de leur domaine comme des « guichets uniques » vers les évaluations d'impact. La réalisation de ces revues systématiques n'a pas été sans difficulté, se heurtant parfois au manque d'évaluations d'impact existantes et à la complexité liée au transfert des interventions « efficaces » d'un contexte à un autre. Certains centres ont ainsi été amenés à diversifier leur approche (notamment à compléter les évaluations d'impact par d'autres types de « preuves » de nature plus empiriques ou qualitatives), d'autres à susciter et à financer de nouvelles évaluations d'impact pour identifier les interventions les plus efficaces ou les innovations les plus prometteuses dans le contexte britannique.

Les What Works Centres diffusent les enseignements de leurs revues systématiques et aident les professionnels à s'en emparer. Ils mettent

What Works Centres britanniques : quels enseignements pour des politiques fondées sur la preuve en France ?

à disposition sous des formes didactiques les résultats de différentes interventions, permettant de comparer leurs effets et leurs coûts. Ainsi, ils se sont rendus rapidement utiles pour les décideurs politiques et les professionnels. En complément de ces actions, certains centres accompagnent les acteurs pour une meilleure intégration des connaissances dans les politiques qu'ils mettent en œuvre. La plupart des centres ciblent les cadres intermédiaires et les professionnels de terrain, avec des outils d'aide aux professionnels et des actions territoriales d'essaimage. Ces actions en sont à leur début : leur efficacité n'est pas encore démontrée, même si certaines semblent prometteuses. Les What Works Centres continuent d'améliorer et de tester leurs approches de diffusion. Ils reconnaissent également que la réussite de leur mission passera par une **évolution des cultures professionnelles**, ce qui exige une action dans la durée.

Tous les centres promeuvent l'évaluation d'impact. Deux accompagnent et financent des évaluations et des expérimentations. En particulier, *Education Endowment Foundation* a engagé 127 évaluations (dont 105 avec randomisation) impliquant 7 000 écoles, soit une école sur quatre en Angleterre. Un tel déploiement a permis de faire émerger des innovations et d'en mesurer l'efficacité, de nouer des relations fortes entre ce centre et les acteurs de terrain et d'enrichir la base de connaissance sur ce qui « fonctionne » pour réduire des inégalités scolaires. L'expérience de ce centre est riche d'enseignements pour ceux qui souhaitent conduire de telles évaluations d'impact en France.

Bien que le Royaume-Uni et la France diffèrent, notamment sur la place de la « preuve » pour justifier les choix d'action publique, ces différences ne semblent pas insurmontables : l'étude montre que les What Works Centres ont trouvé leur place dans des domaines où les décisions sont centralisées, comme décentralisées, en ciblant notamment les professionnels de terrain, qu'ils soient *a priori* sensibles ou non à cette démarche.

L'Ansa et ses partenaires souhaitent promouvoir l'étude de la création ou de la reconnaissance de démarches ou de structures d'inspiration « What Works » en France comme un moyen prometteur de faire progresser les politiques publiques fondées sur la preuve. Ces centres pourraient prendre la forme de nouvelles structures, d'équipes au sein des administrations ou de centres « réseau en étoile » coordonnant un partenariat d'acteurs. De tels centres ne réussiront que s'ils se consacrent à des missions claires, dans la durée, dans des domaines bien définis et disposent des moyens suffisants. Leur rigueur méthodologique, leur indépendance et leur transparence seront également des facteurs clés de réussite. Enfin, pour avoir un impact sur les politiques et pratiques, il s'agira de mettre les attentes des décideurs et des praticiens au cœur de leur action.

Quel que soit le domaine d'intervention, les acteurs font face aux mêmes défis pour élaborer et mettre en œuvre des politiques fondées sur la preuve. Aujourd'hui, il n'existe pas de support dédié à la promotion de ces approches qui soit commun à l'ensemble des parties. L'Ansa et ses partenaires proposent ainsi de **créer un réseau transversal réunissant décideurs politiques, chercheurs et professionnels de terrain autour de la question des politiques publiques fondées sur la preuve**, à l'image d'*Alliance for Useful Evidence* au Royaume-Uni.

What Works Centres britanniques : **quels enseignements
pour des politiques fondées sur la preuve en France ?**

Le rapport

Introduction

Les enjeux : répondre à une exigence d'efficacité

Les pouvoirs publics font face à des problématiques sociales multiples, que ce soit le taux de chômage (qui oscille entre 9 et 10% depuis la crise de 2008), les inégalités scolaires, l'exclusion sociale, le vieillissement de la population ou encore les comportements à risque (dont le tabagisme). Des dépenses publiques sont mobilisées chaque année pour proposer des réponses à ces problématiques ou améliorer celles existantes. Les acteurs de terrain innovent également pour développer de nouvelles interventions.

Cependant, très souvent ces problématiques persistent, voire s'étendent ou s'intensifient. Pourquoi ? A-t-on une connaissance suffisante des besoins des publics ciblés ? Ces actions ont-elles des effets positifs, neutres ou négatifs ? Sont-elles efficaces mais déployées à trop petite échelle ? Si c'est le cas, comment faudrait-il les essayer ou les généraliser ? De plus en plus d'acteurs publics à travers le monde se posent la question de l'impact de leurs politiques, dans un contexte de contraintes budgétaires.

Tenant de répondre précisément à ces questions, l'expérimentation sociale avec évaluation d'impact contrôlée² a connu en France un essor au cours des dix dernières années. Cet élan a suscité l'espoir de disposer de méthodes rigoureuses permettant de concentrer les ressources sur des projets dont les effets auraient été démontrés. À cet espoir ont succédé des interrogations sur les raisons d'une insuffisante diffusion des résultats produits et de leur utilisation dans la définition de dispositifs de politique publique.

Comme le soulignent notamment des travaux récents de l'Agence nouvelle des solidarités actives (Ansa) et de France Stratégie¹, cela est lié aux difficultés que les acteurs ont rencontrées pour :

² L'évaluation d'impact contrôlée est entendue ici dans un sens précis, défini dans la partie « Les What Works Centres britanniques : une réponse », page 45.

What Works Centres britanniques : **quels enseignements pour des politiques fondées sur la preuve en France ?**

- réaliser des évaluations susceptibles de constituer une réelle aide à la décision pour les acteurs publics ;
- faire la synthèse de ces connaissances sur un champ pour en tirer des enseignements, méthodes et outils opérationnels, appropriables par les services publics et leurs partenaires opérateurs, permettant d'orienter la décision publique ;
- organiser la circulation de ces connaissances sur les projets ou pratiques qui « marchent » parmi les acteurs de terrain et parmi les décideurs afin d'en garantir l'essaimage.

D'une manière plus large, des acteurs publics et des chercheurs se mobilisent actuellement autour de la notion d'« *evidence-based policy* », traduite ici par « politique fondée sur la preuve »ⁱⁱ. Les difficultés à surmonter et les perspectives pour mettre en œuvre des politiques fondées sur la preuve sont partagées au sein de réseaux et de communautés, au niveau national ou international (J-Pal, *Alliance for Useful Evidence*, *Evidence Works*, etc.) depuis plusieurs années.

Pour répondre à des questions telles que « Quel est l'impact d'un dispositif, d'une intervention ou d'un traitement sur les bénéficiaires ? », « sur les non-bénéficiaires ? », « Ces résultats sont-ils transposables ou non à une plus large population ? » ... des évaluations solides, bien menées, pouvant apporter une réponse dans des contextes définis, sont nécessaires.

La France dispose déjà d'une importante culture de l'évaluation des politiques publiques. Elle repose principalement sur une forte tradition d'évaluation administrative de l'action publique, menée par des corps d'inspection et de contrôle internes à l'État : ceux-ci conjuguent méthodes quantitatives et méthodes qualitatives pour analyser les mécanismes et les résultats des politiques publiques. Outre une certaine indépendance, leur force est leur excellente connaissance du contexte institutionnel et des acteurs mettant en œuvre ces politiques. Par ailleurs, les pratiques d'évaluation externe des politiques publiques se développent : elles sont menées notamment par les collectivités, des cabinets de conseils spécialisés (tels les membres de la Société française d'évaluation), parfois en lien avec des chercheurs.

Une évaluation d'impact doit être capable d'identifier correctement l'effet causal d'un dispositif sur les bénéficiaires. Cette validité interne des résultats dépend de façon cruciale de la méthode employée au regard des données disponibles. Dans la littérature académique, seulement certaines méthodes d'évaluation (l'évaluation aléatoire par exemple) et certaines conditions de mises en œuvre (expérimentations notamment) permettent d'établir le lien causal entre un dispositif et un impact mesurable (retour à l'emploi suite à un accompagnement par exemple). Également, les effets obtenus sur une population donnée ne sont pas forcément directement obtenus pour un même dispositif appliqué à une autre population, composée d'individus différents (la validité externe des résultats). Enfin, établir l'impact d'une intervention ne renseigne pas nécessairement sur les mécanismes par lesquels ce dispositif agit. Afin de traduire les résultats d'une évaluation en mesures opérationnelles, il est souvent utile de compléter l'approche quantitative par une approche qualitativeⁱⁱⁱ.

La connaissance à prendre en compte pour les politiques fondées sur la preuve est large et inclut les résultats des évaluations d'impact. Elle s'étend à d'autres ressources – allant de l'analyse des besoins, l'évaluation des fondements théoriques des différents programmes et politiques, des évaluations qualitatives (s'appuyant sur des observations, des entretiens, des études de cas), des évaluations de processus, jusqu'à des évaluations quantitatives (statistique descriptive, corrélation)³. Ces autres ressources, de nature plus théorique, descriptive ou qualitative, sont essentielles notamment pour produire des diagnostics (analyse et compréhension des besoins) ainsi que l'élaboration de recommandations.

Ce rapport s'intéresse aux moyens de renforcer le lien entre l'évaluation d'impact et les politiques publiques. Il interroge les enseignements et les inspirations que l'on peut tirer de l'expérience des neuf What Works Centres britanniques qui ont pour objectif principal de contribuer à l'appropriation

3 Voir page 24 « Méthodes d'évaluation ».

What Works Centres britanniques : quels enseignements pour des politiques fondées sur la preuve en France ?

des enseignements des évaluations d'impact par les professionnels et les décideurs politiques.

Les What Works Centres (WWC) : présentation synthétique

Ces centres s'inspirent tous du *National Institute for Health and Clinical Excellence* (NICE), fondé en 1999. Grâce à ses méthodes rigoureuses et à son influence sur les pratiques médicales, NICE est devenu un pionnier de la politique de santé fondée sur la preuve. À partir de 2011, huit nouveaux centres inspirés du modèle de NICE ont été créés au Royaume-Uni. Soutenus par le gouvernement de coalition de David Cameron, ces nouveaux centres ont pu inscrire leur action dans la durée.

Les WWC couvrent divers domaines :

- l'éducation des enfants défavorisés ; (*The Education Endowment Foundation*)
- la prévention précoce (*The Early Intervention Foundation*)^{iv} ;
- le développement économique local (*The What Works Centre for local economic growth*) ;
- la prévention de la criminalité (*The What Works Centre for crime reduction*) ;
- la promotion du bien-être (*The What Works Centre for Wellbeing*) ;
- l'amélioration de la qualité de vie des personnes vieillissantes (*The Centre for Ageing Better*).

Deux centres visent aussi la territorialisation de cette approche, au niveau de deux des nations du Royaume-Uni : l'Écosse (*What Works Scotland*) et le Pays de Galles (*Public Policy Institute for Wales*).

La démarche de l'étude

La démarche s'est appuyée sur les questionnements suivants : comment fonctionnent ces centres ? Quelles sont leurs forces et leurs limites ? Quelles pratiques pourraient inspirer les acteurs français pour renforcer le lien entre évaluation d'impact et politique publique ? Pour quels domaines d'intervention et dans quelles conditions ?

Pour répondre à ces questions, l'Ansa a déployé une méthodologie en deux temps :

- 1- la conduite d'une étude pour comprendre le fonctionnement des What Works Centres, les actions menées, les résultats obtenus, les éventuelles limites identifiées ;
- 2- la comparaison entre le retour d'expérience des WWC et les principales pratiques d'évaluation, de synthèse et d'utilisation des résultats des évaluations en France.

Ce travail a eu pour objectif de fournir un éclairage à la fois monographique et analytique des WWC afin d'en tirer une série d'enseignements pour le cas français. Le choix a été fait d'une étude menée sur un temps court (mai à décembre 2016), pour mettre rapidement à disposition des partenaires et des acteurs intéressés ces premiers enseignements. Elle repose principalement sur des entretiens avec les centres britanniques, certaines de leurs parties prenantes et des acteurs français. Ces entretiens ont été conduits avec un nombre limité de personnes qui étaient, par ailleurs, généralement des parties prenantes des démarches de politiques fondées sur la preuve. Il conviendrait de réaliser davantage d'entretiens avec une plus grande diversité d'acteurs pour approfondir les analyses des What Works Centres et des politiques fondées sur la preuve en France. Ce travail a été complété par une analyse documentaire.

Compte-tenu de la méthode choisie, cette étude ouvre des pistes de travail mais ne permet pas de tirer, sur plusieurs points, de conclusions définitives. Elle signale des enjeux qui mériteraient d'être enrichis d'explorations complémentaires – par exemple auprès des professionnels de terrain et des décideurs politiques sur leur utilisation des outils mis à disposition.

Pour cinq des neuf WWC, des entretiens approfondis ont été conduits avec les responsables de chaque centre, éventuellement des membres de leur équipe. Une étude documentaire présentant leur démarche et les travaux produits a été menée. Les études de cas de ces cinq centres se trouvent en annexe du présent rapport.

What Works Centres britanniques : quels enseignements pour des politiques fondées sur la preuve en France ?

Il s'agit de:

- l'*Early Intervention Foundation* (voir l'étude de cas page 114);
- l'*Education Endowment Foundation* (voir l'étude de cas page 126);
- le *What Works Centre for Local Economic Growth* (voir l'étude de cas page 138);
- le *Centre for Ageing Better* (voir l'étude de cas page 151);
- *National Institute for Health and Care Excellence* (voir l'étude de cas page 165). Cette étude de cas aborde aussi un autre organisme public, *Public Health England*. Ces deux organismes ont été étudiés au titre de leurs travaux d'évaluation, de synthèse et de diffusion de la preuve dans le domaine de la santé publique.

Des entretiens ont été réalisés avec les responsables des autres WWC pour compléter la compréhension des fonctionnements et des enjeux rencontrés.

Afin de comprendre le contexte et les leviers de la mise en œuvre des WWC, des rencontres ont aussi été menées avec des représentants du *Cabinet Office* (services du Premier ministre britannique) qui animent le réseau des What Works Centres, de *Nesta* (une agence d'innovation) ainsi que de l'*Alliance for Useful Evidence*, qui promeut en Grande-Bretagne et au-delà les politiques basées sur la preuve. L'équipe a aussi assisté à un comité de pilotage du réseau des WWC en juin 2016 et à la conférence internationale *Evidence Works 2016* en septembre 2016⁴.

Dans le cadre d'entretiens individuels, plus de trente acteurs français investis dans l'évaluation, la synthèse de la recherche et dans la promotion des politiques publiques fondées sur la preuve ont été interrogés. Un atelier d'experts et quatre ateliers thématiques avec les partenaires de l'étude, sur les sujets de la petite enfance (Cnaf), le vieillissement (CNSA), l'égalité des territoires (CGET), et la santé publique (Santé publique France) ont été menés. Les entretiens et les ateliers ont porté sur une comparaison entre les actions menées par les WWC et par les principaux acteurs français, ainsi que sur les principaux enjeux du contexte français (la liste des personnes interrogées se trouve pages 8 et 9). Cette approche a été complétée par une revue de littérature rapide britannique et française.

Les résultats de ces travaux sont présentés dans ce rapport selon cinq grandes parties :

- la 1^{re} partie définit les différentes composantes des politiques fondées sur la preuve, et présente des initiatives lancées dans divers pays pour rendre ce concept opérationnel (page 22) ;
- la 2^e partie décrit le fonctionnement des What Works Centres britanniques et analyse leur pertinence et résultats (page 45) ;
- la 3^e partie donne un aperçu des pratiques françaises en matière d'évaluation et de politiques fondées sur la preuve (page 76) ;
- la 4^e partie tire les enseignements de l'expérience des What Works Centres pour la France (page 99) ;
- la 5^e partie émet des propositions pour le contexte français (page 106).

i Agence nouvelle des solidarités actives, L'expérimentation sociale à l'épreuve du terrain www.solidarites-actives.com/pdf/Ansa_Contribution16_ExpérimentationSociale_VF.pdf, 2014 ; France stratégie, *Quelle action publique pour demain ?*, 2015.

ii Le présent rapport s'appuie notamment sur l'article de Laurent et al. « Pourquoi s'intéresser à la notion d' « evidence-based policy » ? », *Revue Tiers Monde*, n°200, Octobre-décembre 2009 – p.853-873.

iii Rozenn Desplat, Marc Ferracci, *Comment évaluer l'impact des politiques publiques ? Un guide à l'usage des décideurs et praticiens*, France Stratégie, septembre 2016.

iv Définition de la prévention précoce : prévenir, dès la grossesse et pendant l'enfance, des troubles du développement chez l'enfant et le futur adolescent grâce au ciblage des populations vulnérables.

v Conférence Evidence Works 2016 – Global Forum Gouvernement. Disponible en ligne : www.nesta.org.uk/event/evidence-works-2016-global-forum-government.

Un enjeu partagé au niveau international : des politiques publiques davantage « fondées sur la preuve »

Des politiques publiques « fondées sur la preuve »

L'approche « *evidence-based policy* » ou « politique fondée sur la preuve », après avoir été formalisée dans les années 1990 pour la médecine, s'est étendue à d'autres sphères de la décision publique (lutte contre la pauvreté, éducation, justice, environnement, etc.)^{vi}. Cette approche a pour objectif **de permettre l'utilisation judicieuse des connaissances disponibles par des décideurs publics à tous les niveaux et des professionnels de terrain dans leurs processus décisionnels et leurs pratiques.**

Différentes connaissances peuvent être prises en compte dans cette approche, comme les connaissances scientifiques, les connaissances dans la gestion d'un dispositif (issues de la mise en pratique), la connaissance fine des situations et des besoins des personnes. Cette approche présente aussi des enjeux d'accessibilité des connaissances validées, notamment sous des formats synthétiques adaptés. Enfin elle pose la question de l'intégration de preuves de différentes natures (scientifiques, empiriques, etc.) et de leur hiérarchisation.

La conférence *Evidence Works 2016 – a Global Forum for Government*, organisée par Nesta, *Alliance for Useful Evidence* et *Results for America* en septembre 2016 a réuni 140 participants (ministres, hauts fonctionnaires, évaluateurs) de plus de quarante pays. Ces représentants se sont réunis pour échanger autour des meilleures manières de mettre en œuvre cette approche des politiques fondées sur la preuve, témoignant de l'intérêt généralisé porté à l'efficacité des politiques publiques.

Les échanges ont toutefois souligné les difficultés rencontrées pour favoriser et véritablement appliquer ces politiques fondées sur la preuve : Comment identifier les projets ayant le plus d'effets ? Quels sont les dispositifs qui répondront à telle problématique au meilleur coût ? Dans la plupart des

domaines de la politique publique, les décideurs et professionnels n'ont pas accès à des réponses claires et utiles à ces questions. Même lorsque ces solutions sont connues, comment inciter décideurs et professionnels à s'en emparer ?

Les réponses à ces questions se sont multipliées au cours des vingt dernières années, à la fois au niveau national et international. Si l'on se concentre sur la problématique de l'évaluation d'impact, qui fait partie du champ des connaissances pour les politiques publiques fondées sur la preuve, ces réponses passent par la production d'évaluations d'impact, les revues systématiques de littérature répondant à des questions d'efficacité d'intervention, et l'accompagnement du changement dans les pratiques des décideurs et des professionnels pour qu'ils s'en saisissent.

Produire un niveau de preuve de qualité : les évaluations d'impact et les synthèses de la recherche

La recherche d'efficacité des politiques publiques a conduit au recours de plus en plus important aux méthodes d'évaluation sous diverses formes, allant des évaluations qualitatives ou de processus de projets aux évaluations de politiques publiques dans leur ensemble, en passant par l'évaluation d'impact. Ces méthodes portent sur des objets allant de politiques publiques globales à des dispositifs ou projets, voire des pratiques et interventions spécifiques.

Le terme d'évaluation recouvre donc un ensemble très large d'approches et d'objectifs. Ces différentes méthodes permettent d'apporter des connaissances de nature variée. Elles s'appliquent chacune dans des contextes et circonstances spécifiques.

Le focus qui suit présente un rappel succinct de pratiques répandues, qui sont traitées dans de nombreux articles par ailleurs.

MÉTHODES D'ÉVALUATION^{vii}

L'analyse du besoin et l'évaluation des fondements théoriques d'un programme **vont permettre d'une part de bien définir le besoin auquel on cherche à répondre, les différentes manières d'aborder le problème, les actions et le public à cibler et, d'autre part, d'identifier et de vérifier les fondements théoriques (scientifiques) des actions choisies et les effets attendus.**

L'évaluation de processus **est destinée à suivre la mise en œuvre d'un programme : s'assurer que les actions ont bien lieu, se déroulent avec la qualité et dans les délais attendus.**

L'évaluation d'impact **a pour objet d'identifier si un programme ou une politique publique a eu des effets mesurables et de les quantifier. Cette démarche est approfondie dans le paragraphe suivant.**

Les évaluations qualitatives^{viii} **consistent en une investigation approfondie d'un petit échantillon d'individus, s'appuyant notamment sur des observations, des entretiens, des études de cas. Ce type d'évaluation peut contribuer, par exemple, à la compréhension de ce qui fonctionne ou non dans une intervention.**

Les évaluations quantitatives **reposent sur des échantillons représentatifs, à partir de plusieurs centaines d'individus.**

Les évaluations socio-économiques^x **désignent deux types de travaux. D'une part, il s'agit de modèles économiques calibrés à partir de données empiriques. Ce type de calcul a été historiquement employé pour l'évaluation de grands investissements publics comme dans les transports^x. D'autre part, l'évaluation socio-économique de programmes consiste à comparer le coût d'un programme par rapport aux résultats obtenus (analyse coût/efficacité), dans certains cas en associant une valeur monétaire au résultat (analyse coût/bénéfice). Elle permet aussi de comparer les programmes entre eux et d'aider à choisir les programmes les plus efficaces.**

L'évaluation d'impact : vers des niveaux de preuve plus élevés

Les approches de politiques fondées sur la preuve reposent sur le choix de s'appuyer sur les connaissances proposant le meilleur niveau de preuve. Dans le cadre de l'évaluation de l'impact d'un dispositif sur des bénéficiaires, il s'agira de rechercher ainsi le niveau de preuve le plus élevé disponible pour démontrer les effets produits par une action. Le niveau de preuve correspond au degré de confiance que l'on a dans la capacité de la conception et du protocole de recherche à identifier le lien de cause à effet.

Les évaluations d'impact ont pour objet d'établir un lien de causalité entre une intervention (par exemple un accompagnement financier ou humain) et des impacts pour des personnes (par exemple l'accès ou le retour à l'emploi) pour démontrer si un programme a produit directement les effets recherchés. La recherche médicale a développé et approfondi ces approches, avant qu'elles ne soient transposées et adaptées dans le champ social.

La littérature académique définit des **conditions de mise en œuvre spécifiques et/ou des méthodes économétriques pour parvenir à établir l'impact d'un dispositif au sens causal**^{ki}. Il s'agit notamment de définir des objectifs que l'on cherche à atteindre et d'élaborer des indicateurs représentant ces objectifs (pouvant ainsi réduire les effets mesurés à quelques dimensions). L'intérêt d'une intervention se mesure alors à la taille de l'impact mesuré et le degré de confiance en ce résultat dépendra de la qualité de la preuve (notamment la présence d'un contrefactuel valide, la sélection et la taille des échantillons, le contrôle de biais de sélection, etc.). Dans un souci de simplicité, l'Ansa désignera les évaluations d'impact avec contrefactuel par « évaluation d'impact » (voir encadré suivant « L'évaluation d'impact et l'évaluation sociale »).

Les évaluations d'impact constituent une des techniques d'évaluation. Leurs résultats constituent une forme de « preuve », à considérer dans un champ plus large de connaissances qui, ensemble, sont susceptibles

de contribuer à une « culture de la preuve » et à des « politiques fondées sur la preuve ». **Cette méthode a l'avantage d'apporter une réponse robuste du point de vue académique aux interrogations sur l'efficacité d'innovations ou de dispositifs en place.** Elle présente toutefois des biais et limites à prendre en compte à la conception de l'évaluation et lors de l'utilisation des résultats. Par ailleurs, il est souligné **qu'il n'est pas possible de la mettre en œuvre pour tout type d'intervention ou de dispositif.** En particulier, lorsqu'un programme cherche à intervenir sur des phénomènes aux causalités trop complexes, où les résultats sont trop difficiles à mesurer, les évaluations d'impact peuvent être infructueuses et il faut donc se saisir d'autres méthodes.

Pour s'emparer de manière transparente et rigoureuse des différentes évaluations réalisées, **des organismes scientifiques ont établi des classifications des différentes méthodes d'évaluations sur la base de leur qualité scientifique et de leur valeur probante.** Ces échelles de niveau de preuve peuvent aller de la qualité de la conception d'une intervention sur la base de connaissance scientifique préalable, aux études avant/après, aux évaluations d'impact avec contrefactuel. On peut citer la *Maryland Scientific Scale* et l'échelle de Nesta comme exemples de classification^{xii}. Selon l'approche théorique, la discipline ou l'objet traité, il sera plus ou moins facile de construire des preuves de niveau élevé, et ces échelles de classification pourront donc varier. Il s'agit alors de mettre en balance un niveau de preuve élevé par rapport à des méthodes d'évaluation pertinentes pour l'action envisagée.

L'ÉVALUATION D'IMPACT ET L'EXPÉRIMENTATION SOCIALE

L'évaluation d'impact estime la différence, en moyenne, entre une situation avec intervention et celle qui aurait prévalu en son absence, en comparant un échantillon d'individus affectés par l'intervention à des individus non affectés. En effet, comme un individu ne peut être à la fois affecté et non affecté, la situation sans intervention est par nature inobservable et doit être estimée. C'est le contrefactuel.

Le contrefactuel fait appel à un échantillon non affecté par l'intervention dont on veut mesurer l'effet, pour représenter ce qui se serait passé pour l'échantillon affecté s'il ne l'avait pas été. Le groupe « témoin » qui le constitue doit donc être le plus proche possible de l'échantillon « test » qu'il doit représenter. Sans quoi, la comparaison risque de confondre l'effet de l'intervention avec les différences (observables et inobservables) qui existent entre les deux groupes. L'identification du contrefactuel est la principale difficulté de l'évaluation d'impact et la capacité à identifier un effet causal dépend de la stratégie d'identification employée.

La recherche met en évidence deux ensembles de méthodes permettant d'identifier un contrefactuel valide.

Deux types d'évaluation d'impact^{xiii xiv}

1- évaluation expérimentale, aussi appelée expérimentation aléatoire contrôlée, randomisée. Cette méthode consiste à tirer au sort les personnes qui vont rentrer dans le dispositif (groupe test) et à suivre les personnes qui ne sont pas retenues lors du tirage (groupe témoin). Par son assignation aléatoire, elle permet de créer des groupes test et témoin en moyenne identiques et, ainsi, de mieux isoler les effets dus à l'intervention. De ce point de vue, elle est souvent considérée comme la méthode de référence et le niveau le plus élevé de mesure des impacts.

2- évaluation quasi-expérimentale : il s'agit de méthodes qui imitent les situations expérimentales à partir de méthodes statistiques (comme l'appariement statistique, les doubles ou triples différences) ou en identifiant ce qui dans la « nature » ou les règles institutionnelles produit des variations créant un hasard d'accès au traitement, pour identifier son effet (variable instrumentale, discontinuité de régression, ou d'autres expériences naturelles)^{xv}. Au sein de cette famille, les différentes méthodes comportent chacune des biais et des limites, qu'il faut prendre en compte dans l'analyse des résultats.

D'autres méthodes d'évaluation et de recherche viendront compléter la réalisation d'une évaluation d'impact. Il s'agit, d'une part, de la recherche permettant de comprendre en profondeur le besoin

ou la problématique (observation, recherche fondamentale) et, d'autre part, d'évaluations de processus permettant de comprendre les conditions de réalisation de l'intervention évaluée.

L'expérimentation sociale^{xvi} désigne une innovation de politique sociale initiée dans un premier temps à petite échelle, dans des conditions qui permettent d'en évaluer les effets, en vue d'une généralisation. Elle permet de valoriser l'innovation et repose en principe sur une évaluation d'impact (avec contrefactuel). De nombreux acteurs publics, en France et dans les pays de l'Organisation de coopération et de développement Économiques (OCDE), ont le souci de tester ainsi à petite échelle de nouveaux programmes, afin d'en valider la faisabilité, la pertinence et surtout l'impact. Le terme d'expérimentation sociale désigne toutefois de plus en plus souvent aujourd'hui en France des projets pilotes⁴ ou d'innovations qui ne se dotent pas nécessairement d'une évaluation d'impact.

Validité interne : est-ce que l'évaluation d'impact est concluante ?

Les évaluations d'impact comportent des biais inhérents à leur démarche. Ces biais doivent être traités, au risque sinon d'aboutir à des résultats d'évaluation non-concluants. Denis Fougère^{xvii} décrit les biais inhérents aux expériences contrôlées et décrit des réponses qui peuvent y être apportées par les protocoles et plans d'expérimentation. Il attire l'attention sur le fait que la question des protocoles et des plans d'expérience sont trop souvent ignorés par les praticiens de ces méthodes appliquées aux sciences sociales, et que l'« *on a souvent pu constater que les estimations déduites des expérimentations sociales étaient statistiquement peu significatives et/ou peu robustes* ». La réponse passe donc par la prise en compte de ces aspects, en s'appuyant sur des évaluateurs compétents.

Validité externe et enjeux de généralisation

Une fois que les effets d'une intervention sur les bénéficiaires sont

⁴ Le terme de projet pilote, ou de pilote, désigne ici des dispositifs qui sont mis en œuvre à petite échelle, notamment pour en tester leur faisabilité.

démontrés dans le cadre d'une évaluation d'impact (validité interne), d'autres questions se posent pour décider de sa capacité à être essayée voire généralisée :

- l'intervention évaluée est-elle transposable à d'autres contextes, d'autres groupes de personnes, d'autres circonstances? Par exemple, une crise économique est-elle survenue alors que la situation économique était stable lors de l'expérimentation (validité externe) ?
- le passage à l'échelle peut-il perturber la situation globale et avoir des effets indirects, non mesurés dans l'expérimentation? Par exemple, aura-t-il des effets non-anticipés sur la situation globale, tels qu'une montée des prix de biens de consommation lorsqu'une aide financière est généralisée? Généraliser le programme sur un temps long changera-t-il le comportement des bénéficiaires ?
- comment garantir le respect des modalités de l'intervention évaluée (ou « fidélité ») ?

Le processus de décision politique se doit donc d'intégrer ces questionnements s'il veut s'appuyer sur les résultats d'une expérimentation pour décider de sa généralisation.

Dans le domaine des politiques sociales et de l'emploi, on peut citer un certain nombre d'expérimentations ayant fait l'objet d'un essai réussi à l'échelle internationale, dont *Individual Placement and Support* (IPS).

INDIVIDUAL PLACEMENT AND SUPPORT (IPS)

***Individual Placement and Support* offre un suivi intensif individualisé, illimité dans le temps, à l'accès et au maintien à l'emploi sur le marché ordinaire du travail des personnes souffrant de troubles psychiques. Cette méthode a fait l'objet de nombreuses expérimentations contrôlées aux États-Unis et en Europe. Ces études ont souligné un impact positif, avec un taux d'accès à l'emploi en milieu ordinaire de 42% à 61%^{xviii}. Les expérimentations européennes ont montré que les bénéficiaires d'IPS avaient deux fois plus de chances d'accéder à un emploi que**

dans l'approche classique linéaire « formé puis placé »^{xxix}. IPS a fait l'objet d'expérimentations réussies au Royaume-Uni, Norvège, Danemark, Chine (région de Hong Kong), Canada, Nouvelle Zélande et Australie. Depuis 2012, un think tank indépendant – le *Centre for Mental Health* – a accompagné la création de 14 équipes d'IPS en Angleterre, en veillant à la fidélité au modèle initial. En parallèle, IPS a fait l'objet d'une adoption plus large en Angleterre. En 2012, la moitié des services de santé mentale secondaires disaient disposer d'au moins un conseiller IPS^{xx}. La méthode IPS se diffuse également en France, à une échelle plus modeste. Plusieurs associations françaises ont adopté cette méthode dans des projets sous des appellations diverses (emploi accompagné, *Working First* ou *Job Coach*).

Évaluations socio-économiques : une mesure de l'efficacité de dispositifs

Entendues ici au sens des analyses coût/efficacité et coût/bénéfice, les évaluations socio-économiques permettent de mettre en regard le coût d'un dispositif par rapport aux effets obtenus. Elles reposent sur les résultats d'évaluations d'impact, qui permettent de comparer les situations et trajectoires entre les groupes test et témoin. Les coûts et effets à prendre en compte dépendent des objectifs de l'évaluation et reposent sur des hypothèses fortes (prendre en compte ou non les coûts fixes, les coût/opportunités pour le bénéficiaire, etc.)^{xxi}.

En adoptant les mêmes conventions de valeur, de projection ou d'effets pris en compte ou non, elles permettent alors de comparer des programmes poursuivant de mêmes objectifs. Ce type d'analyses constitue une aide à la décision dans l'allocation de ressources entre différents programmes.

MÉTHODES

Analyse coût/efficacité^{xxii} : **ratio entre le coût du programme et son efficacité (par exemple le nombre de personnes accédant à un emploi). Des interventions auront des effets sur plusieurs axes et il est alors possible de mener des analyses du rapport coût/efficacité sur plusieurs**

résultats recherchés. Par exemple, l'institut NICE au Royaume-Uni utilise le QALY (Quality Adjusted Life Year) : année de vie corrigée en fonction de la qualité de vie de la personne. Un QALY équivaut à une année de vie en bonne santé. NICE mesure et compare alors le coût des programmes en livres sterling par QALY espéré.

Analyse coût/bénéfice : **cette analyse reposera sur l'attribution d'une valeur monétaire aux différents bénéfices obtenus et la comparaison avec le coût du programme. En ce qui concerne en particulier les effets sociaux (par exemple l'amélioration du bien-être, coûts évités d'un décrochage scolaire ou de crimes non commis), leur donner une valeur financière est particulièrement délicat et peut induire de fortes variations selon les résultats retenus. Cela implique en effet de faire des hypothèses très fortes et de valoriser des effets pour lesquels il n'y a pas aujourd'hui de consensus pour les politiques sociales^{xxiii}.**

Les revues systématiques de littérature

Face à une problématique que l'on souhaite résoudre, de nombreuses interventions peuvent exister, avec des effets et des coûts différents. Les pratiques courantes de synthèse de la recherche vont viser à analyser les développements les plus marquants sur un sujet particulier ou une controverse, à partir de références connues et peuvent répondre à plusieurs préoccupations simultanées. Ce type de synthèse est adapté dans de nombreuses situations et peuvent donner à voir des interventions intéressantes.

Dans le contexte de politiques publiques fondées sur la preuve – et des dispositifs traités dans ce rapport – une approche spécifique prédomine : les revues systématiques de littérature. Cette méthode permet de **comparer les impacts des interventions et donc de distinguer celles qui fonctionnent ou sont prometteuses** de celles qu'il vaut mieux écarter, pour une question précise, posée par des praticiens. Ces revues seront d'autant plus intéressantes si elles y associent une analyse des coûts, voire des bénéfices, des interventions. Elles permettent aussi d'identifier les interventions et champs d'action où il y a eu peu d'évaluations et qui nécessitent des

What Works Centres britanniques : **quels enseignements pour des politiques fondées sur la preuve en France ?**

recherches complémentaires. Elles sont toutefois complexes et coûteuses en ressources et en temps.

Les revues systématiques de littérature reposent sur des normes méthodologiques précises – la définition de Cochrane pour le milieu médical est rappelée ici :

La revue systématique (selon Cochrane, voir focus ci-dessous) est une démarche scientifique rigoureuse constituée de plusieurs étapes :

- 1– chercher toutes les études existantes, essais cliniques ou autres études appropriées, publiées ou non, qui ont répondu à une même question clinique précise (par exemple, tous les essais cliniques effectués afin d'évaluer l'efficacité d'un traitement ;
- 2– évaluer la qualité de chaque étude et sélectionner celles qui correspondent à un standard de qualité élevé et défini au préalable ;
- 3– effectuer la synthèse des résultats obtenus dans les études sélectionnées ;
- 4– finalement, si les données le permettent, effectuer une analyse statistique, appelée méta-analyse. Elle permet de regrouper les résultats quantifiés provenant de plusieurs études en un résultat résumé^{xxiv}.

Lorsqu'elles s'appuient sur les meilleures études, ces revues présentent une vision synthétique et globale des connaissances sur une question donnée, de leur fiabilité, et permettent d'analyser la relative efficacité de l'ensemble des solutions.

**COCHRANE ET CAMPBELL COLLABORATION :
DEUX ACTEURS MAJEURS DE LA REVUE SYSTÉMATIQUE**

Cochrane (anciennement connu sous le nom de Cochrane Collaboration) a été fondé en 1993 pour que les études randomisées soient développées et plus utilisées dans le monde médical, afin de permettre une prise de décision de qualité relative à la santé et aux soins médicaux. Ce réseau international de chercheurs et de professionnels fournit au monde médical des revues systématiques de toutes les évaluations aléatoires contrôlées et du meilleur niveau

d'évaluation disponible dans la santé. Il fait aujourd'hui figure de référence dans ce domaine.

De nombreux membres de Cochrane identifiaient le besoin de revues systématiques pour évaluer de la même manière les interventions dans le champ du social : *Campbell Collaboration* est né de ce constat en 2000. *Campbell Collaboration* publie des revues systématiques de littérature dans les champs de l'éducation, de la criminalité et de la justice, du développement international et de la protection sociale. Le réseau a produit 112 revues systématiques de littérature entre 2004 et 2015^{xxv}.

Quand l'État favorise l'expérimentation avec évaluation d'impact

Des freins face à l'évaluation d'impact

La mise en œuvre d'expérimentations avec évaluation d'impact ou d'évaluations d'impact de dispositifs n'est, *a priori*, pas une évidence pour les gouvernements et les responsables politiques. Le décalage entre le temps long de l'expérimentation et le temps court des mandats politiques, le coût potentiellement élevé des évaluations d'impact constituent des freins à la mise en œuvre. Fondamentalement, expérimenter revient à assumer, à l'encontre des postures habituelles de décideurs politiques, que les effets des mesures proposées sont incertains.

“La démarche expérimentale repose sur l'idée qu'on ne sait pas, d'emblée, si une mesure va fonctionner ou pas. Cette idée est difficile à accepter pour des décideurs politiques habitués à vendre des résultats certains à l'électorat.”

David Halpern, directeur de Behavioural Insights Team.

Pour accepter les expérimentations avec évaluation d'impact ou les évaluations d'impact de dispositifs – et y participer – les professionnels de terrain doivent également accepter l'idée que leurs dispositifs puissent être évalués pour être ensuite améliorés ou enrichis par des apports extérieurs.

What Works Centres britanniques : quels enseignements pour des politiques fondées sur la preuve en France ?

L'évaluation est ainsi parfois vécue comme une sanction, une notation, une remise en question de leurs actions, voire de leurs compétences.

Une volonté politique

Le développement d'évaluations d'impact et d'expérimentations avec évaluation d'impact dépend donc d'incitations efficaces à dépasser ces freins et repose sur la volonté politique – des acteurs publics comme de certains acteurs privés – de financer et d'institutionnaliser ces évaluations.

Les incitations à expérimenter et évaluer que nous avons observées peuvent prendre plusieurs formes :

- **les fonds pour l'évaluation et l'expérimentation.** Créé en 2009, le Fond d'expérimentation pour la jeunesse en France bénéficie d'une dotation initiale de 230 millions d'euros de la part de l'État et de financeurs privés pour la période 2009–2012. L'*Education Endowment Foundation*, qui fait partie des neuf What Works Centres britanniques, a reçu 125 millions de livres sterling à dépenser sur 15 ans (voir étude de cas en annexe, page 126). Bien qu'ayant des approches différentes, ces fonds accompagnent et financent des expérimentations ou des évaluations d'impact, les programmes étant portés par des acteurs de terrain et les évaluations réalisées par des évaluateurs indépendants.
- **l'intégration de démarches d'expérimentation au sein des administrations.** Par exemple, l'équipe *Behavioural Insights Team*, intégrée pendant plusieurs années aux services du Premier ministre britannique, a mis en œuvre de nombreuses expérimentations avec évaluation d'impact relatives aux interactions entre les services et les citoyens, en s'appuyant sur les sciences du comportement.
- **l'accompagnement à l'expérimentation ou l'inscription dans la réglementation.** Le Royaume-Uni a créé au sein de son administration le *Cross-Government Trial Advice Panel*^{xxvi} (conseil interministériel d'experts pour l'expérimentation), afin d'accompagner les agents de l'État vers la mise en œuvre d'expérimentations.

BEHAVIOURAL INSIGHTS TEAM, OU « NUDGE UNIT »^{xxvii}

Le *Nudge Unit* a été créé en 2010 au sein des services du premier ministre britannique pour mener des expérimentations aléatoires de petites incitations ou changements fondés sur la recherche en psychologie comportementale à des fins d'amélioration des services publics pour les usagers. Le centre a testé de nombreuses innovations pour faciliter le recours aux services publics et pour inciter les citoyens à faire de « meilleurs » choix. Au cours de ses deux premières années d'existence, il a conduit des dizaines d'expérimentations dans des domaines aussi variés que la santé, les impôts, les économies d'énergie, la réduction de la criminalité, l'emploi et même la croissance économique. Ces expérimentations ont permis de démontrer que nombres de petites actions (en particulier dans les interactions entre les services publics et les usagers) avaient des effets notables et permettaient parfois de réaliser des économies substantielles pour un résultat équivalent (voire meilleur).

Par exemple, on s'est rendu compte qu'ajouter dans un courrier destiné aux contribuables un commentaire sur le fait que la plupart d'entre eux paient leurs impôts à temps augmenterait encore le taux de paiement dans les délais. On a aussi pu démontrer que demander aux personnes en recherche d'emploi ce qu'elles vont faire pour trouver un travail dans les deux prochaines semaines, plutôt que les deux dernières, lors de leur entretien régulier, améliorerait significativement leur chance de retour à l'emploi.

Réussir l'utilisation des connaissances par les professionnels et les décideurs politiques à tous les niveaux

Pour prendre des décisions pertinentes et améliorer leurs pratiques, les décideurs politiques et les professionnels ont besoin d'avoir une compréhension des dispositifs qui ont de bonnes chances de fonctionner

dans leur domaine et leurs conditions de réussite.

Cet enjeu est partagé plus largement pour la connaissance dans son ensemble, au-delà des seules évaluations d'impact de dispositifs – notamment la connaissance scientifique. Les approches de politique fondée sur la preuve cherchent à répondre à cet enjeu, démontrant une volonté de renouveler les démarches d'aide à la décision publique, en s'appuyant sur la meilleure connaissance disponible.

Les enjeux du « transfert de la connaissance »

L'appropriation des connaissances scientifiques par un public de professionnels est une problématique bien connue et étudiée dans le domaine médical. L'adoption des interventions médicales les plus efficaces est souvent un processus très long (17 ans selon une estimation de Balas et Boren^{xxviii}). Cette question du transfert de la connaissance, au sens de « diffusion, adoption et appropriation des connaissances en vue de leur utilisation dans la pratique et les prises de décisions »^{xxix}, se pose tout autant dans le domaine social.

Face à une recherche qui évolue en continu et à la multiplication des sources de connaissance, un certain nombre de freins à l'utilisation de ces connaissances sont cités au sujet de la recherche interventionnelle en santé publique et se retrouvent dans d'autres domaines :

- **manque de matière sous une forme appropriée** : les résultats des études et les connaissances sont rarement présentés de manière synthétique et accessible pour des non-initiés – les professionnels et décideurs y trouvent rarement des réponses à leurs problématiques dans un format qui leur soit utile ;
- **manque d'apprentissage en continu** : les individus au sein des organisations fonctionnent avec leurs compétences explicites mais aussi implicites, acquises au fil de l'expérience, avec des convictions. Ils dépendent également d'une culture organisationnelle. Ces éléments ne les incitent pas d'emblée à remettre en question des routines et des hypothèses tenues pour acquises ;

- **manque de confiance des acteurs**, des professionnels et des décideurs **dans les résultats d'une recherche donnée**, notamment parce que les organisations auxquelles ils appartiennent reposent rarement sur une culture de recherche ;
- **résistance au changement** des personnes et des organisations, difficulté à créer l'adhésion des parties prenantes (décideurs politiques, professionnels, bénéficiaires, citoyens) à une nouvelle solution ou à la suppression d'une action qui n'a en fait pas d'effet positif^{xxx}.

Les décideurs, cadres intermédiaires, professionnels de terrain sont tous parties prenantes de processus de décision complexes, imbriqués dans des fonctionnements institutionnels et organisationnels, dépendant des dispositifs de gouvernance et de la capacité des organisations à changer. Ces acteurs doivent aussi tenir compte de la compréhension globale du contexte, de l'évolution de la situation, des enjeux politiques dans ces processus. Les démarches de transfert de la connaissance ont pour objet de les aider à y incorporer les connaissances et, en particulier, les résultats d'évaluations d'impact et de leurs synthèses.

Les leviers d'action pour répondre à ces enjeux se situent donc à plusieurs niveaux qui entreront en jeu de manière complémentaire :

- la mise à disposition de résultats de la recherche (évaluation d'impact, synthèse de la recherche) sous des formes que les professionnels et décideurs peuvent utiliser ;
- l'évolution culturelle et organisationnelle chez les décideurs et professionnels, pour créer un environnement favorable aux politiques fondées sur la preuve.

En cohérence avec les valeurs d'efficacité fondée sur la preuve, des chercheurs britanniques se sont penchés sur les démarches favorisant l'utilisation des connaissances et les évaluations existantes : *The Science of Using Science*. Un premier travail de revue systématique de littérature avec un niveau d'évaluation robuste et de littérature relevant des sciences sociales (i.e. de ressources de nature plus qualitative) met en avant les résultats de recherche existants et pointe encore un manque de preuve

What Works Centres britanniques : **quels enseignements pour des politiques fondées sur la preuve en France ?**

dans ce domaine^{xxx}. Les actions présentées dans les parties suivantes représentent donc des pistes intéressantes, plutôt que des interventions qui seraient déjà formellement démontrées comme efficaces.

Mettre les résultats d'évaluation d'impact et de revues systématiques de littérature à disposition sous forme de recommandations et de pratiques

Les décideurs et professionnels ont besoin d'informations qui leur permettent de résoudre des problématiques générales (comment réduire les inégalités scolaires ?) ou spécifiques (est-ce que la diminution de l'effectif des classes est « efficace » ?). Pour cela, les résultats mis à disposition doivent leur donner :

- une vue d'ensemble des solutions qui pourraient y répondre, avec des recommandations des solutions qui semblent les plus efficaces ou prometteuses,
- des interventions disponibles mettant en avant les « indications d'utilisation » (pour quel public, dans quel contexte) et les modalités détaillées de mise en œuvre.

La collaboration entre chercheurs, producteurs de la recherche et professionnels, chargés de la mise en œuvre des services, apparaît comme un facteur de succès dans cette étape, pour garantir que la preuve est présentée de manière accessible et utilisable et qu'elle est utile. Cette collaboration peut aussi se poursuivre au-delà pour favoriser une meilleure adéquation de la recherche produite aux besoins opérationnels.

Nous présentons ici deux formes existantes de mise à disposition des résultats d'évaluation et de synthèse de la recherche, sous forme d'intervention et de recommandations : les conférences de consensus et les Clearing Houses.

CONFÉRENCE DE CONSENSUS

L'objectif d'une conférence de consensus est d'identifier les questions incontournables et de cerner les points de controverse à dépasser pour un problème défini. En mobilisant de nombreux regards croisés,

en s'appuyant sur les connaissances scientifiques internationales, en permettant un débat structuré, une telle méthode peut aboutir à la construction d'un socle de consensus susceptible de poser les bases d'une évolution de la politique publique, d'un dissensus ou d'un constat que les preuves sont insuffisantes pour statuer. Les résultats prennent la forme de recommandations émises par un jury indépendant, au terme de la présentation publique de rapports d'experts faisant la synthèse des connaissances.

La méthode des conférences de consensus a été initiée aux États-Unis dans les années 70, pour déterminer la meilleure manière d'agir dans un cas médical spécifique. Initialement composées uniquement d'experts, ces conférences se sont progressivement ouvertes aux non-professionnels et sont mises en œuvre dans les champs de la protection sociale, juridiques, etc.

En France, le recours à cette méthode a augmenté au cours des dernières années, avec :

- la conférence de consensus en décembre 2003 sur l'apprentissage de la lecture par le PIREF (www.cndp.fr);
- les conférences de consensus pratiquées par la Haute Autorité de Santé dans le domaine médical, ainsi les 16 et 17 décembre 2005 sur la prise en charge de la psychopathie et les 25 et 26 janvier 2007 sur l'expertise psychiatrique pénale;
- dans le champ social, la conférence de consensus sur les personnes sans-abri « Sortir de la rue » en 2007;
- dans le champ de la justice, la conférence de consensus sur la prévention de la récidive, les 14 et 15 février 2013;
- dans le domaine de l'éducation, les conférences de consensus organisées par le Conseil national de l'évaluation du système scolaire (Cnesco) : sur le redoublement en janvier 2015 (www.cnesco.fr rubrique conférence de redoublement), sur la numération en novembre 2015, sur la lecture en mars 2016 et sur la différenciation pédagogique (dont les séances publiques auront lieu en mars 2017).

CLEARING HOUSES

Ces ressources – souvent sous forme de sites internet – listent les interventions dont les effets ont été démontrés par des évaluations d'impact robustes et en précisent les modalités. Elles ont émergé dans un contexte américain caractérisé par un fort développement de l'évaluation d'impact de programmes de prévention des exclusions sociales et par une volonté de mieux employer des budgets publics restreints.

On y trouve *Blueprints for Healthy Youth Development*^{xxxii} (« des modèles pour un développement sain des jeunes »). L'organisme *Blueprints* met à disposition un registre de programmes qui promeuvent la santé et le bien-être des enfants et des adolescents. *Blueprints* a été un des pionniers dans la recherche et l'identification de programmes ayant eu des évaluations d'impact. Accueilli au sein de l'université du Colorado à Boulder, il a démarré en 1996 par des travaux pour prévenir la violence des jeunes et a ensuite étendu son champ d'action. À ce jour, plus de 1 400 programmes ont été revus, pour seulement 5% finalement retenus comme des programmes modèles ou prometteurs.

D'autres exemples à consulter pour aller plus loin :

- le *What Works Clearinghouse – Institute of Education Sciences aux États-Unis*^{xxxiii} ;
- le *Washington State Institute for Public Policy*^{xxxiv}.

Favoriser une culture de la preuve au niveau des décideurs et professionnels

Dans de nombreux pays, des actions ont pour objet spécifique de conduire les décideurs et professionnels à utiliser les méthodes et apports des politiques fondées sur la preuve.

Les quelques initiatives présentées ici ciblent souvent les décideurs politiques, sous différentes formes.

Type d'initiative	Initiative	Pays
Réseaux de promotion des politiques publiques fondées sur la preuve	Alliance for Useful Evidence	Royaume-Uni
	Results for America	États-Unis
	African Evidence Network	Continent Africain ⁵
Réseaux de citoyens	Evidence for Democracy	Canada
Campagnes d'information tout public	Ask for Evidence	Royaume-Uni
Services de conseil intégrés à l'administration	Cross-Government Trial Advice Panel	Royaume-Uni
	Commission on Evidence-Based Policy making	États-Unis
Formation continue des décideurs publics et des fonctionnaires	Les évaluations d'impact ont été intégrées à la formation des fonctionnaires dans le cadre de leur redéfinition du Policy Profession	Royaume-Uni
	Formation par le bureau d'évaluation	Afrique du Sud
Obligations réglementaires	L'Office of Management and Budget lie désormais le financement des programmes publics par l'État davantage à des exigences de preuves, en distinguant : <i>concept grants</i> (pour développer et expérimenter des idées prometteuses), <i>robust evaluations</i> (pour mettre en œuvre des approches ayant fait leurs preuves)	États-Unis

Créer des organisations pour renforcer le lien entre évaluation d'impact et politiques publiques

Le changement de paradigme dans la prise de décision des politiques publiques et les pratiques des professionnels pour intégrer au fil de l'eau les résultats de la recherche évaluative se joue sur un temps long : il s'agit

⁵ L'Afrique du Sud et le Malawi en sont les membres fondateurs.

What Works Centres britanniques : **quels enseignements pour des politiques fondées sur la preuve en France ?**

de transformer des pratiques, des compétences, des cultures de travail, le regard des décideurs, des professionnels de terrain et du grand public. Pour intégrer de manière durable, dans les comportements des décideurs, une prise de décision informée par la preuve, il faut combiner des approches complémentaires, cohérentes entre elles, successives ou simultanées, sur plusieurs années, voire des décennies. **De nombreux experts internationaux s'accordent donc pour dire qu'il faut une (ou des) infrastructure dédiée(s) pour accompagner les acteurs de la conception à la mise en œuvre des politiques basées sur la preuve^{xxv}.**

Ces infrastructures prennent souvent la forme d'instituts indépendants :

- le Bureau for Economic Policy Analysis aux Pays-Bas,
- le Productivity Commission en Australie,
- les What Works Centres au Royaume-Uni.

Elles peuvent également reposer sur des équipes dédiées au sein des administrations, telles que le *Behavioural Insights Team* au Royaume-Uni (désormais indépendant), et l'*Evidence Team* aux États-Unis.

Ces organisations jouent en général un rôle de courtage de la connaissance, produisant des livrables qui doivent être pertinents, robustes et disponibles « au bon moment ». Ces écrits prennent des formes diverses, tels qu'un « policy brief » (par exemple *Promising policies series* du Bureau for Economic Policy Analysis du Pays-Bas) ou des listings d'intervention réalisés par les What Works Centres au Royaume-Uni.

Les WWC britanniques sont un cas concret de cette nouvelle infrastructure, potentiellement source d'inspiration pour les acteurs français. La prochaine partie analyse leurs pratiques et la quatrième tire les enseignements de leur expérience pour la France.

vi Laurent et al. « Pourquoi s'intéresser à la notion d' « evidence-based policy » ? », Revue *Tiers Monde*, n°200, Octobre-décembre 2009 – p.853-873.

vii Poverty Action Lab, *Introduction to Evaluations* (consulté le 24 novembre 2016). Disponible en ligne : www.povertyactionlab.org/research-resources/introduction-evaluations.

- viii** Guide méthodologique pour l'évaluation des expérimentations sociale. 6 avril 2009. Disponible en ligne : www.experimentation.jeunes.gouv.fr/IMG/pdf/guide-pour-l-evaluation-des-experimentations.pdf.
- ix** Heim Arthur, *Comment estimer le rendement de l'investissement social*, à paraître.
- x** « [les principes] consistent à jauger les conséquences de chaque projet sur l'ensemble de la collectivité nationale, en évaluant les gains de productivité ou de capacité de production apportés aux entreprises ainsi que les suppléments de consommation marchande et les améliorations de qualité de vie procurés aux citoyens, et en permettant de comparer ces effets aux coûts de l'investissement. » dans : Emile Quinet, *L'évaluation socioéconomique des investissements publics*, rapport publié le 18 septembre 2013. Disponible en ligne : www.strategie.gouv.fr/publications/evaluation-socioeconomique-investissements-publics-tome1.
- xi** Desplat Rozenn, Ferracci Marc, *Comment évaluer l'impact des politiques publiques ? Un guide à l'usage des décideurs et praticiens*, France Stratégie, septembre 2016.
- xii** What Works Centre for Local Economic Growth. Maryland Scientific Scale. Disponible en ligne : www.whatworksgrowth.org/resources/the-scientific-maryland-scale. Putick Ruth & Ludlow Joe. *Standards of Evidence*. Nesta, Octobre 2013, p.2. Disponible en ligne : www.nesta.org.uk/sites/default/files/standards_of_evidence.pdf.
- xiii** Des définitions simples sont données pages 34 et 35 du rapport *L'expérimentation sociale à l'épreuve du terrain*, Contributions n°16, Agence nouvelle des solidarités actives, mars 2014. Disponible sur www.solidaritesactives.com, rubrique Publications.
- xiv** Présentation approfondie de méthodes d'évaluation d'impact, Pauline Givord, *Méthodes économétriques pour l'évaluation de politiques publiques*. WPD3E n°G2010-08, 2010.
- xv** Un panorama récent de ces méthodes quasi-expérimentales se trouve dans : Rozenn Desplat, Marc Ferracci, *Comment évaluer l'impact des politiques publiques ? Un guide à l'usage des décideurs et praticiens*, France stratégie, septembre 2016.
- xvi** Claire Alspektor, « L'expérimentation sociale : innover par la preuve ? », *Les notes d'analyse et de synthèse de la MASSP*, n°24, novembre 2015.
- xvii** Fougère Denis, « Les méthodes d'expérimentation en question », *Éducation & formations*, n°81, mars 2012, p. 41-47.
- xviii** Becker, D. R. et al., *A Mental Health Learning Collaborative on Supported Employment*. 2011, Psychiatric Services.
- xix** Campbell K., Bond G. R., and Drake R. E., *Who Benefits From Supported Employment: A Meta-analytic Study*. 2009.
- xx** Briefing 47 : *Barriers to employment*, 2012, Centre for Mental Health, www.cnwl.nhs.uk/wp-content/uploads/2012/11/Briefing47_Barriers_to_employment.pdf.
- xxi** Heim Arthur, « Comment estimer le rendement de l'investissement social », à paraître.
- xxii** Cellini Stephanie, Kee James, « Cost-effectiveness and cost-benefit analysis », dans *Handbook of practical programme evaluation*, 2010, chapitre 21.
- xxiii** Weimer David L. & Vining Aidan R., (editor) (2009) *Investing in the disadvantaged. Assessing the benefits and costs of social policies*. Georgetown university press, Washington D.C.
- xxiv** Cochrane Suisse. *Revue systématique*. Disponible en ligne : swiss.cochrane.org/fr/les-revues-syst%C3%A9matiques-systematic-reviews.
- xxv** Campbell Collaboration. Disponible en ligne : www.campbellcollaboration.org.
- xxvi** The Cross-Government Trial Advice Panel. Disponible en ligne : www.gov.uk/government/uploads/system/uploads/attachment_data/file/532459/Cross-Government_Trial_Advice_Panel.pdf.
- xxvii** Halpern David, *Inside the Nudge Unit*, WH Allen, 2015.
- xxviii** Balas, EA, Boren SA. *Managing clinical knowledge for health care improvement*. In: Bommel J. MCrCray AT,

What Works Centres britanniques : quels enseignements pour des politiques fondées sur la preuve en France ?

editors. *Yearbook of Medical Informatics 2000: Patient-Centred Systems*. Stuttgart, Germany: Schattauer Verlagsgesellschaft mbH; 2000: 65-70.

xxix Cambon Linda, Alla François, « Recherche interventionnelle en santé publique, transfert de connaissance et collaboration entre acteurs, décideurs et chercheurs, le défi français de l'innovation », *GIS-IRESP, Questions de santé publique*, n°27, décembre 2014.

xxx Le redoublement, traité lors de la conférence de consensus de la CNESCO, constitue un bon exemple des difficultés d'adhésion rencontrées.

xxxi Langer Laurenz, Tripney Janice, Gough David, « The science of using science. Researching the use of evidence in decision-making », avril 2016, EPPI-Centre, University College London.

xxxii Blueprints for Healthy Youth Development.

Disponible en ligne : www.colorado.edu/cspv/blueprints/28.

xxxiii What Works Clearinghouse – Institute of Education Sciences aux États-Unis. Disponible en ligne : ies.ed.gov/ncee/wwc.

xxxiv Washington State Institute for Public Policy. Disponible en ligne : www.wsipp.wa.gov.

xxxv Puttick Ruth, « Why we need to create a "NICE for social policy" », Nesta, mai 2012.

Les What Works Centres britanniques : une réponse ?

Les What Works Centres (WWC) britanniques sont nés d'un constat simple : une partie importante des services publics britanniques sont peu ou pas « fondés sur la preuve ». En particulier, ces services sont rarement conçus et mis en œuvre en appliquant les pratiques démontrées comme étant efficaces par des évaluations d'impact. Dans le champ éducatif, une analyse de 70 programmes financés par le ministère de l'Éducation anglais a identifié seulement deux à trois programmes qui avaient fait l'objet d'une évaluation rigoureuse^{xxxvi}. Même dans le domaine de la médecine, l'adoption des interventions les plus efficaces est souvent un processus très long (17 ans selon une estimation de Balas et Boren^{xxxvii}). Cette situation tient, entre autres, au manque de liens entre évaluateurs et décideurs politiques, à la difficulté à traduire des rapports d'évaluation en enseignements opérationnels et au faible poids des évaluations dans les décisions politiques.

“Bien que l'idée d'avoir des politiques et des pratiques fondées sur des preuves rigoureuses soit largement admise, des milliards de livres sterling sont dépensées chaque année sur des politiques et programmes ayant très peu de preuve rigoureuse quant à leur impact.”

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

Présentation des What Works Centres

Les principes généraux

Les What Works Centres ont été conçus pour répondre à cet enjeu et favoriser l'adoption des dispositifs qui « fonctionnent »^{xxxviii} en portant ces trois missions :

– **synthétiser la recherche existante** dans leur domaine, par des revues systématiques de littérature en donnant une place privilégiée aux évaluations d'impact scientifiques. Les centres transposent cette recherche

What Works Centres britanniques : quels enseignements pour des politiques fondées sur la preuve en France ?

en enseignements, voire en recommandations concrètes, pour les professionnels et décideurs,

- **diffuser et accompagner l'adoption.** Ces centres diffusent la recherche, notamment par le biais de bibliothèques d'intervention (« *toolkit* »), conçues pour aider les professionnels à identifier les interventions les plus efficaces. Ils accompagnent également son adoption par le biais d'événements, de formations et de programmes d'essaimage,
- **évaluer l'impact d'innovations.** Certains centres disposent également des ressources dédiées pour alimenter ce corpus de recherche par l'expérimentation et l'évaluation d'impact.

Chaque centre investit un domaine spécifique de la politique publique, une approche qui comporte deux avantages. D'une part, ceci leur permet de faire un recensement exhaustif des évaluations d'impact et donc se positionner comme « guichet unique » d'accès à cette matière. D'autre part, les centres peuvent organiser une diffusion ciblée et adaptée aux professionnels du domaine. Les domaines sont divers, allant de l'éducation à la santé, du développement économique au bien-être. Chaque domaine a présenté des défis spécifiques en termes d'évaluation, de synthèse de la recherche et de diffusion. Chaque centre a dû s'adapter aux spécificités de son domaine. Deux centres visent aussi la territorialisation de cette approche au niveau de deux des nations du Royaume-Uni : l'Écosse (*What Works Scotland*) et le Pays de Galles (*Public Policy Institute for Wales*).

WHAT WORKS CENTRES : PRINCIPALES CARACTÉRISTIQUES

En termes de missions, ils :

- se concentrent sur la recherche relative aux solutions efficaces, plus qu'à l'analyse des besoins. **Pour ce faire, les centres procèdent à un classement des dispositifs selon leur efficacité, en tenant compte à la fois de leurs résultats, de la rigueur des évaluations réalisées et de leurs coûts. Les centres fournissent une indication synthétique des coûts des différents dispositifs afin d'éclairer les choix des décideurs sur leur efficience ;**

- visent en priorité à changer les pratiques sur le terrain (approche par le bas, « *bottom up* »), plutôt que les orientations de politiques publiques nationales (approche par le haut, « *top down* »). Par conséquent, ils ciblent principalement les cadres intermédiaires et professionnels, même si les décideurs nationaux peuvent parfois être dans leur cible ;
- accompagnent des professionnels et des cadres locaux à s'emparer des enseignements de la recherche. S'inspirant de la science de l'application (« *implementation science* »), les centres testent diverses méthodes de diffusion et d'essaimage.

En termes d'organisation et de financement, les centres sont :

- statutairement indépendants, portés généralement par une association, fondation ou université (à l'exception de deux centres publics) ;
- très proches des pouvoirs publics, puisque bénéficiant de subventions publiques ;
- sectoriels, reliant les décideurs, chercheurs et professionnels sur un domaine ;
- investis dans la durée, avec des financements de 3 à 15 ans ;
- très inégalement dotés en moyens (rapport de 1 à 100 entre le budget annuel du *What Works Centre for Wellbeing* et celui du *National Institute for Clinical Excellence*).

Présentation synthétique

On peut distinguer trois phases dans l'émergence des neuf What Works Centres.

1– En 1999, le ministre de la santé britannique crée le **National Institute for Clinical Excellence (NICE)**. Sa mission est d'évaluer les médicaments et les matériels médicaux, notamment en termes d'efficacité par rapport à leur coût. Son champ d'action s'est élargi progressivement aux directives cliniques et à la santé publique. Son but est d'apporter ces éléments au *National Health Services* (NHS) pour diffuser les bonnes pratiques rapidement. En effet, grâce à ses méthodes rigoureuses et à son influence

What Works Centres britanniques : **quels enseignements pour des politiques fondées sur la preuve en France ?**

sur les pratiques médicales, NICE devient un pionnier de la politique de santé fondée sur la preuve.

2- En 2011, l'**Education Endowment Foundation** est fondé avec une dotation de 125 millions de livres sterling, s'inspirant notamment de *Race to the Top*, un programme américain.

3- Enfin, à partir de 2013, **sept nouveaux What Works Centres** ouvrent leur porte. Cinq sont thématiques et deux sont territoriaux, s'adressant aux administrations décentralisées de l'Écosse et du Pays de Galles.

Structure	Domaine	Date de création	Statut	Budget annuel (2015)	Durée de financements	Taille de l'équipe
National Institute for Health and Care Excellence (NICE)	Santé, santé publique	1999	Public	£63,3 million	Pérenne	600 salariés (environ)
Education Endowment Foundation	Inégalités scolaires	2011	Association et Fondation	£14 million	2011-2026 (dotation £125m)	15 salariés
Early Intervention Foundation	Enfants de la conception à 18 ans	2013	Association et Fondation	Non communiqué	2013-2016	16 salariés
What Works Centre for Local Economic Growth	Développement économique local	2013	Groupe-ment	£1 million	2013-2016	10 ETP
What Works Centre for Crime Reduction	Prévention de la criminalité	2013	Public	Non communiqué	2013-2016	Non communiqué

Public Policy Institute for Wales	Pays de Galles	2014	Groupe-ment d'univer-sités	£450 000	2014-2017	8 salariés
What Works Centre for Scotland	Écosse	2014	Groupe-ment d'univer-sités	£1 million	2014-2017	7 ETP
What Works Centre for Wellbeing	Bien-être	2014	Entreprise sociale	£600 000	2014-2017	6 salariés
What Works Centre for Ageing Better	Bien-vieillir	2015	Fondation	En moyenne £5 m/an sur 10 ans	2015-2025 (dota-tion : £50m)	15 salariés (environ)

Contexte de leur émergence

La création de ces huit nouveaux centres à partir de 2011 est une page inédite dans l'histoire des politiques publiques au Royaume-Uni.

Ce phénomène peut s'expliquer par quatre facteurs en particulier.

Premièrement, le Royaume-Uni **a bénéficié d'un engouement pour l'évaluation d'impact et des politiques « fondées sur la preuve »**, sous l'impulsion notamment du gouvernement travailliste élu en 1997. En effet, les gouvernements de Tony Blair représentaient un courant de la gauche, dite souvent réformiste, pour laquelle l'efficacité des services publics était un enjeu clé. Dès 1999, ce gouvernement a publié un livre blanc soulignant que « la preuve » doit être un des composants centraux de politiques publiques plus efficaces^{xxix}. Cette vision a été rapidement déclinée en différentes initiatives^{xl}. D'une part, de nombreuses expérimentations contrôlées ont été menées, par exemple pour tester les politiques actives en faveur de la participation au marché du travail^{xi}. D'autre part, de nouveaux dispositifs se sont lancés pour inscrire la « preuve » dans la décision politique et les pratiques des professionnels. *Le National Institute*

What Works Centres britanniques : **quels enseignements pour des politiques fondées sur la preuve en France ?**

for Health and Clinical Excellence, créé en 1999, en est l'exemple phare. Entre 2000 et 2004, une autre expérience – moins connue – a été conduite pour soutenir des politiques basées sur la preuve : le *UK Centre for Evidence Based Policy and Practice*^{xlii}. Véritable précurseur des What Works Centres, ses financements n'ont pas été reconduits, vraisemblablement en partie à cause de son orientation trop académique et donc insuffisamment opérationnelle pour une aide à la décision.

Deuxièmement, **l'idée des What Works Centres a fait l'objet d'un plaidoyer appuyé par des personnalités influentes**, en première ligne par David Halpern (directeur du *Behavioural Insights Team*) et Geoff Mulgan (directeur de l'agence d'innovation Nesta). Entre 2011 et 2012 de nombreuses références sont faites à un « *NICE for Social Policy* », dans des blogs, rapports^{xliii} et articles de presse^{xliv}. Ces acteurs ont constaté l'absence d'organisations chargées de faire le lien entre l'offre et la demande pour « la preuve ». Pour combler ce manque, ils se sont tournés vers NICE et ont donc proposé de transposer le modèle de NICE aux politiques sociales, dans des domaines tels que l'éducation, la justice, et la protection de l'enfance. À partir de 2012, l'*Alliance for Useful Evidence* – réseau pour des professionnels intéressés par les « politiques basées sur la preuve » – a également apporté sa pierre à l'édifice avec des publications et des événements à ce sujet. Un acteur historique de l'expérimentation – *Dartington Social Research Unit* – a aussi contribué de manière tangible à la promotion des What Works Centres. En effet, sa participation à un rapport parlementaire de Graham Allen a contribué à la création de l'*Early Intervention Foundation*^{xlv}.

Troisièmement, **ce plaidoyer a été bien reçu par certains des hauts fonctionnaires, conseillers et ministres du nouveau gouvernement de coalition**. Cette idée a été intégrée dans une stratégie de réforme des services publics orientés vers l'ouverture et la transparence. Ainsi, l'idée de dupliquer NICE sur d'autres champs est rapidement évoquée dans le *Livre blanc des services publics ouverts* (2011)^{xlvi}. Ce concept de NICE pour les politiques sociales aurait pu en rester à l'état d'idée. Cependant, il a bénéficié d'un soutien d'une figure importante, celle

de Jeremy Hayward, chef de la fonction publique, qui a évoqué dès 2012 l'idée d'un NICE pour les politiques sociales^{xvii}. Pour impulser la création de nouveaux centres, le *Cabinet Office* (services du Premier ministre) a été doté d'une équipe dédiée aux What Works Centres.

Enfin, le **Cabinet Office**, avec le soutien de l'*Economic and Social Research Council*⁶ (ESRC), **a mobilisé des financeurs autour de la création de nouveaux centres**. Ils ont mené un travail de concertation pour identifier les champs où il y avait un besoin clair pour un What Works Centre. Cette concertation a permis de mobiliser des financements de base pour lancer de nouveaux centres, notamment de la part des ministères. Dans le cas de quatre centres, l'ESRC a apporté des co-financements importants^{xviii}. Le centre sur le « bien vieillir » fait figure d'exception puisqu'il a bénéficié d'une dotation du *Big Lottery Fund* (« Grand fond de la loterie »), un organisme public disposant d'une plus grande indépendance que les ministères.

LE RÔLE DU CABINET OFFICE

Une équipe de trois personnes au sein du Cabinet Office (service du Premier ministre) joue un rôle de coordination et d'animation des What Works Centres, une action qui se décline en cinq axes :

- appuyer au cadrage des nouveaux centres ;
- appuyer à l'obtention et au renouvellement des financements des centres ;
- tisser des liens entre les centres et les ministères (fonctionnaires, conseillers politiques etc.) ;
- animer le réseau des centres, avec une rencontre trimestrielle du What Works Council, où les directeurs des centres échangent autour de leurs pratiques et de potentielles collaborations ;
- promouvoir les centres à l'international.

Grâce en partie au Cabinet Office, les centres ont trois niveaux

⁶ Principal financeur public de la recherche en sciences économiques et sociales.

What Works Centres britanniques : quels enseignements pour des politiques fondées sur la preuve en France ?

de collaboration :

- des échanges de pratiques lors des What Works Councils trimestriels, voire dans d'autres contextes ;
- l'utilisation des travaux de synthèse entre centres ;
- des projets collaboratifs de recherche ou d'expérimentation.

Gouvernance des centres

Les centres ont divers statuts : certains publics, d'autres associatifs, d'autres encore sont plutôt un programme d'activité porté par un consortium. De ce fait, leurs instances formelles de gouvernance sont diverses. Malgré ces différences, les centres partagent beaucoup dans leur conception de la gouvernance et dans la place accordée aux parties prenantes. En effet, la plupart intègre dans leurs instances permanentes des acteurs des mondes de la recherche et de l'action publique. Les projets de recherche peuvent également intégrer des instances participatives. NICE représente un exemple à cet égard, puisque les recommandations de leur *Guidance* sont élaborées directement par un « *Committee* », composé de professionnels de la santé et d'autres acteurs opérationnels^{xlix}.

- Pour en savoir plus : étude de cas sur NICE (page 165)

L'action des centres

Synthétiser la recherche existante

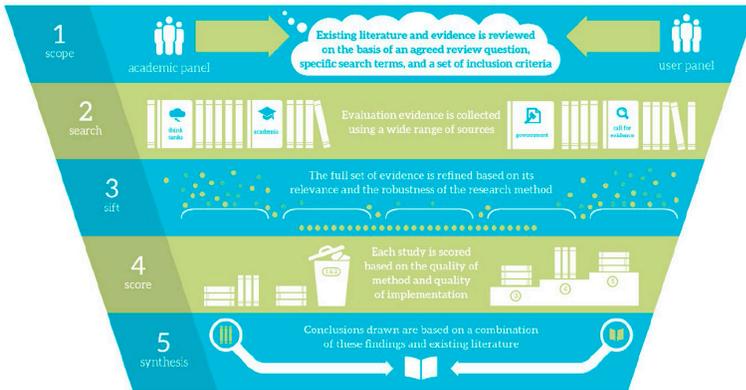
Les What Works Centres (WWC) visent à porter les travaux de recherche sur « ce qui fonctionne » dans leur domaine à la connaissance des acteurs publics et associatifs. Ces revues visent avant tout à répondre aux besoins des professionnels de terrain et des cadres intermédiaires. Afin de disposer de l'état de l'art de la recherche sur leur domaine, les WWC ont tous engagé un programme conséquent de revues systématiques des évaluations d'interventions. En réalisant une série de revues sur une période d'environ trois ans, la plupart des centres ont traité une partie

importante de la recherche dans leur domaine. La plupart présente cette matière scientifique de manière accessible sous forme de bibliothèque d'interventions ou « *toolkit* » (voir page 62).

Pour éviter les biais inhérents dans les revues de littérature classiques, **ces centres effectuent des synthèses de la recherche selon la méthode de « revue systématique »**. Cette méthode est « systématique » puisqu'elle respecte un protocole de recherche rigoureux : exhaustivité des études portant sur un sujet spécifique, critères de sélection et choix des bases documentaires explicites. Ces revues sont élaborées selon un protocole de recherche défini, allant du cadrage avec les parties prenantes, en passant par le recensement et l'identification des études, à leur sélection, puis leur évaluation et leur synthèse. Le protocole du *What Works Centre for Economic Growth* illustre bien cette approche globale.

► Pour en savoir plus : étude de cas *What Works Centre for Economic Growth* (page 138)

Méthodologie des revues systématique du *What Works Centre for Economic Growth*



Source : www.whatworksgrowth.org.

What Works Centres britanniques : quels enseignements pour des politiques fondées sur la preuve en France ?

Les centres ont tous l'ambition de s'appuyer sur les « meilleures études existantes » dans le cadre de ces revues. Les centres donnent tous une place privilégiée aux évaluations d'impact, avec différents niveaux de flexibilité. Plusieurs ont élaboré une hiérarchie permettant de distinguer la relative rigueur des différentes méthodes d'évaluations, s'inspirant souvent de *Maryland Scientific Methods Scale* (MSMS)¹.

MARYLAND SCIENTIFIC METHODS SCALE

Niveau 1. **Corrélation des résultats avec la présence ou l'intensité de l'intervention, sans volonté d'établir une situation contrefactuelle.**

Niveau 2. **Comparaison entre les résultats pour le groupe test après une intervention avec les résultats pour ce groupe avant l'intervention (évaluation avant/après).**

Niveau 3. **Comparaison entre les résultats pour un groupe de traitement après une intervention et ceux d'un groupe de comparaison pour créer une situation contrefactuelle, avec un risque d'effets de variables non-observable importants (par exemple une méthode d'appariement statistique).**

Niveau 4. **Comparaison entre les résultats du groupe test après une intervention, et pour un groupe témoin pour créer une situation contrefactuelle plus robuste que le niveau 3, comme des expériences naturelles.**

Niveau 5. **Les évaluations intégrant l'assignation aléatoire aux groupes tests et témoins (« randomisation »).**

Source : What Works Centre for Local Economic Growth¹. Traduction de l'Ansa.

Certains excluent d'emblée les évaluations ne disposant pas d'un contrefactuel – le *What Works Centre for Local Economic Growth* n'analyse que les études correspondant aux niveaux 3 à 5 du MSMS. En étant aussi sélectif, ce centre a fait face au faible nombre d'évaluations d'impact scientifiques dans le domaine du développement économique local.

“Certains types de dispositifs ne se prêtent pas aux méthodes d'évaluations les plus rigoureuses. Pour les actions d'activation des chômeurs, de conseils aux entreprises, d'incubation des start-up, il est possible de randomiser. Mais on ne peut pas randomiser l'emplacement des logements et de l'infrastructure... Il y a des domaines du développement économique local qu'on ne peut pas évaluer par ces méthodes, même si d'autres méthodes peuvent convenir en ces cas, comme l'appariement statistique.”

Max Nathan, directeur adjoint, What Works Centre for Local Economic Growth (LSE).

D'autres incluent un panel plus large d'évaluations, tout en précisant les conditions d'inclusion des études. Le *Early Intervention Foundation* (EIF), par exemple, réalise une cotation des études, allant de la présence d'un simple modèle logique et argumenté (niveau 1) à la réalisation d'au moins deux évaluations d'impact rigoureuses (niveau 4)ⁱⁱⁱ. L'EIF a adopté cette approche face à la difficulté à définir de manière incontestée si une intervention évaluée a effectivement des effets positifs et mesurables. Une poignée d'interventions sont clairement efficaces ou inefficaces selon les évaluations réalisées. Cependant, beaucoup d'interventions se situent dans une « zone grise », soit parce que les évaluations estiment précisément un effet nulle ou très faible, soit puisqu'elles ne parviennent pas à identifier un effet (par manque de puissance statistique, problèmes d'attribution etc.). En incluant un panel large d'évaluations, l'EIF souhaite ainsi donner une chance aux programmes d'être affinés, améliorés et mieux conçus sur une certaine durée de mise en œuvre, et de faire l'objet d'une évaluation plus robuste dans un deuxième temps. L'approche d'EIF s'inspire notamment des « *Standards of Evidence* » élaboré par NESTA, où l'évaluation accompagne le développement de projets de sa conception à son essaimage à grande échelleⁱⁱⁱⁱ.

► Pour en savoir plus : étude de cas Early Intervention Foundation (page 114)

What Works Centres britanniques : quels enseignements pour des politiques fondées sur la preuve en France ?

Standards of Evidence de NESTA



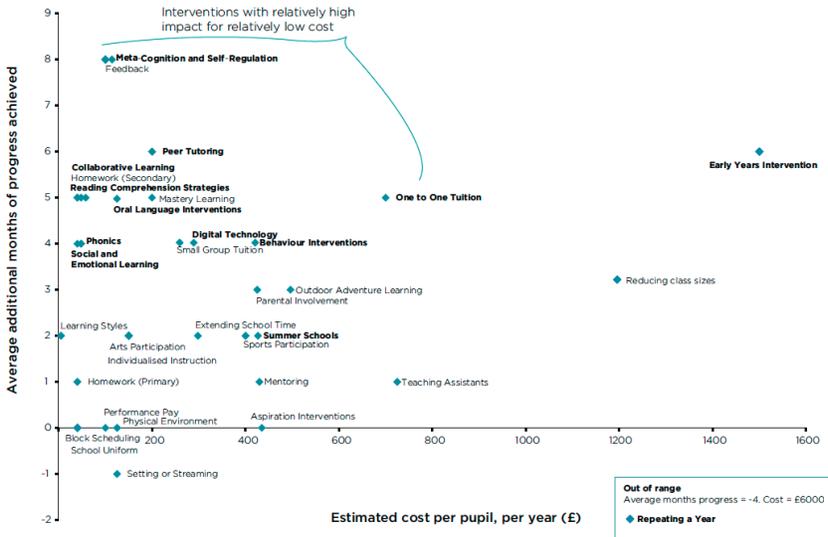
Source : Standards of Evidence, Nesta, 2013.

Traduction : Ben Rickey, Ansa.

L'analyse des coûts est un composant central de la plupart des revues systématiques menées par les WWC. Tous intègrent un calcul du coût des interventions évaluées, permettant de les comparer entre elles. Certains centres, selon les objets d'études et les informations disponibles, vont jusqu'à faire des analyses coût/efficacité. Cette analyse est utile quand on doit comparer des dispositifs qui visent un seul objectif (par exemple de meilleurs

résultats scolaires). Pour effectuer de telles analyses, l'*Education Endowment Foundation* a mis en regard dans sa boîte à outils le progrès scolaire mesuré par des évaluations selon un critère de « nombre de mois de progrès éducatif », avec un calcul du coût de chaque intervention par élève.

Analyse coût/efficacité des interventions pédagogiques



Source : The Sutton Trust – EEF Teaching and Learning Toolkit and Technical Appendices, 2012^{iv}.

Notes : Le coût par élève par an se base sur une classe composée de 25 élèves. Les interventions soulignées en gras sont celles pour lesquelles le niveau de preuve est important à très important selon la définition de la boîte à outils. Le mois de progrès éducatif est estimé en comparant l'évolution des scores des élèves lors des tests standardisés entre le début et la fin de l'année scolaire. Douze mois de progrès est donc la moyenne de tous les élèves. Lorsqu'un dispositif fait progresser une classe plus que la moyenne, ce progrès est quantifié en mois additionnels de progrès.

Certains intègrent également les bénéfices des interventions, exprimés en montant financier. L'intérêt de l'analyse coût/bénéfice est de considérer une large gamme d'effets dans les bénéfices. Cette méthode est d'ailleurs souvent utilisée par des actions de prévention ayant de multiples effets. Par exemple, l'*Early Intervention Foundation* s'appuie sur de nombreuses analyses coût/bénéfice réalisées à l'égard des interventions auprès des enfants en bas âge. Le High Scope Perry Preschool Program par exemple avait un ratio coût/bénéfice entre 1 à 7 ou de 1 à 12, principalement grâce à des niveaux plus élevés d'emploi et de salaire, et des taux plus faibles de criminalité et de recours aux minima sociaux^{iv}. Ce type d'analyse est souvent utilisé pour calculer la valeur monétaire des actions préventives, qui réduisent des coûts de prise en charge, on parle alors de coûts évités.

Les WWC s'interrogent également sur la capacité à tirer des enseignements généraux des études spécifiques. D'une part, les revues incluent généralement des études réalisées dans les pays de l'OCDE, soulevant la question de la transposabilité au contexte britannique des résultats obtenus dans d'autres contextes politiques, institutionnels et culturels. D'autre part, les WWC se sont heurtés au côté « boîte noire » des articles scientifiques. En effet, certaines évaluations d'impact ne mettent à disposition ni les présentations détaillées ni les modalités de l'intervention ni les conditions dans lesquelles l'intervention semble produire des effets. Le *WWC for Crime Reduction* a élaboré un cadre d'analyse des interventions qui cherche à mieux intégrer pleinement cette question de la transposabilité dans leurs revues (grille « EMMIE »).

EMMIE FRAMEWORK^{vi}

Effect size : **le type et taille des effets d'une intervention, ainsi que la confiance que nous avons en ces analyses ;**

Mechanisms : **les mécanismes activés par la politique, pratique ou dispositif en question ;**

Moderators **ou contextes qui impactent la présence/absence des effets attendus ou inattendus ;**

Implementation **désigne les éléments contribuant à la réussite**

ou l'échec de la politique, pratique ou dispositif;

Economic costs : coûts (et bénéfices) associés à la politique, pratique ou dispositif.

La revue systématique n'est pas une nouvelle pratique. Cependant, les revues des WWC se distinguent notamment par leur accessibilité à un public non-expert. En particulier, les WWC ont tous fait un effort accru de traduction des analyses en enseignements pour les professionnels. Par exemple, les onze revues du WWC for *Local Economic Growth* sont toutes accompagnées d'une fiche (« *Executive Summary* ») qui synthétise et vulgarise les enseignements de l'étude pour un public professionnel.

“Il y avait beaucoup de travail avec les chercheurs pour que leurs recommandations soient plus explicites, et d'arrêter de ménager la chèvre et le chou. Il existe une tension créatrice entre être scientifique et avoir un message clair pour les professionnels.”

Meg Kaufman, coordinatrice du What Works Centre for Local Economic Growth au Centre for Cities.

Bien que la revue systématique soit le principal type de recherche utilisée par les centres, plusieurs recourent à d'autres méthodes.

La plupart des centres ont une compréhension approfondie de la problématique qui les concerne, à la fois des profils des publics concernés, des facteurs socio-économiques ou de santé importants, des comportements et des dispositifs. Cet ensemble d'éléments permet aux centres de prioriser des sujets de recherche, d'évaluation et d'innovation. Deux centres – le *Centre for Ageing Better* et le *WWC for Wellbeing* – ont voulu approfondir la compréhension des phénomènes mal compris que sont le bien-vieillir et le bien-être. Pour ce faire, ils ont mobilisé et conduit des travaux sociologiques (enquêtes, entretiens, observation participante) et de concertation.

► Pour en savoir plus : étude de cas Centre for Ageing Better (page 151)

What Works Centres britanniques : **quels enseignements pour des politiques fondées sur la preuve en France ?**

Enfin, les centres territoriaux font figure d'exception en termes d'approche méthodologique. En effet, leur programme de recherche est plus « réactif » et co-construit avec les pouvoirs publics en Écosse et au Pays de Galles. Ils s'appuient sur un panel plus large de types de recherche et sont beaucoup moins attachés aux évaluations d'impact que les centres thématiques. Avec une aussi grande hétérogénéité de méthode, il sera probablement légitime de parler d'approches « What Works », en distinguant les centres thématiques et territoriaux.

LES WHAT WORKS CENTRES TERRITORIAUX : D'AUTRES APPROCHES

Public Policy Institute for Wales a été l'une des mesures du programme du parti Travailleiste lors des élections de l'Assemblée du Pays de Galles en 2011. Un consortium d'universités a répondu à l'appel d'offre et le centre a été formellement lancé en janvier 2014. Tous les six mois, l'équipe rencontre les ministères du Pays de Galles afin d'élaborer leur programme de recherche. Elle mobilise ensuite des équipes universitaires qui réalisent les projets – souvent des synthèses de la recherche – et organise leur restitution auprès du ministre, ses conseillers et des hauts fonctionnaires. En deux ans, le centre a ainsi fait réaliser 60 projets de recherche. Sa proximité avec les ministres, sa réactivité, et sa capacité à coordonner la réalisation de nombreux projets de recherche l'a rendu particulièrement utile pour les décideurs nationaux de cette administration.

What Works Centre for Scotland a été fondé en 2014 pour aider les acteurs locaux à s'appuyer sur la preuve dans leurs prises de décision relatives à l'amélioration et la réforme des services publics. Un des objectifs du centre était de promouvoir des approches préventives et participatives dans un contexte d'austérité. Leur approche est très décentralisée et leur programme de travail entre 2014 et 2016 reposait principalement sur des collaborations avec des acteurs dans quatre territoires : Aberdeenshire, Fife, Glasgow et West Dunbartonshire. Dans chaque territoire, le centre co-construit un diagnostic du contexte local, facilite l'acculturation entre acteurs, définit concrètement

ce qu'on comprend par « services efficaces », et promeut l'utilisation de la recherche dans la planification et la mise en œuvre des services. À la différence des centres thématiques, son approche privilégie moins l'évaluation d'impact, toute forme de recherche pouvant informer les services publics étant mobilisées (par exemple les données d'observation).

L'approche rigoriste des What Works Centres a attiré des critiques.

En particulier, certains acteurs interrogent l'idée d'appliquer une hiérarchie de la recherche qui place les évaluations randomisées systématiquement en tête.

“Il n'y a pas de réponse simple à la question « qu'est-ce qu'un bon niveau de preuve ». Cela dépend de ce que nous voulons savoir, pour quelles fins et dans quel contexte nous envisageons d'utiliser la recherche... il faut reconnaître que, en pratique, les décideurs doivent traiter la recherche dans son ensemble. En effet, des décideurs politiques et professionnels doivent mettre la recherche relative à ce qui fonctionne dans la balance avec des analyses relatives aux coûts, à l'acceptabilité, et aux effets distributionnels.”

What counts as good evidence?, Alliance for Useful Evidence¹⁶.

Diffuser et accompagner l'adoption

Les synthèses de la recherche ne sont pas une fin en soi. Les WWC visent surtout à accompagner les acteurs de terrain à adopter des pratiques efficaces, par une meilleure diffusion de la recherche.

Tous les WWC rendent accessibles leurs revues. Pour la plupart des centres, ce travail commence par la traduction des revues systématiques en enseignements pour les professionnels. Il se poursuit par la diffusion de ces publications, par des voies numériques et dans le cadre d'événements. Malgré leurs intérêts, ces revues individuelles donnent un aperçu partiel de la recherche.

What Works Centres britanniques : quels enseignements pour des politiques fondées sur la preuve en France ?

Afin de devenir de véritables « guichets uniques » vers la recherche, la plupart des WWC ont donc élaboré des **bibliothèques d'interventions qui visent à donner une vision synthétique de la recherche.**

Ces bibliothèques présentent une analyse scientifique de la relative pertinence de chaque intervention, selon au moins trois critères : l'impact, la rigueur de la recherche et le coût. Conçues comme de véritables aides à la décision pour les agents publics, elles sont interactives et faciles d'accès. **Le Teaching & Learning Toolkit (« Malette de pédagogie et d'apprentissage ») de l'Education Endowment Foundation présente un exemple particulièrement abouti.** Le Toolkit proposé est un outil interactif intégré au site internet du centre. Il est également évolutif. Mis à jour tous les six mois par l'équipe de recherche de l'université de Durham, il est passé de 11 à 34 dispositifs éducatifs. L'analyse des dispositifs a également été mise à jour pour prendre en compte de nouvelles études.

LE TEACHING & LEARNING TOOLKIT EXPLIQUÉ

Le Teaching & Learning Toolkit est une revue des revues de la recherche sur l'efficacité de différentes interventions pédagogiques. Cette synthèse porte donc principalement sur les résultats d'évaluations expérimentales ou quasi-expérimentales. Le Toolkit présente des analyses relatives à différentes « interventions pédagogiques » (ou « strands »), selon trois dimensions :

- l'impact. **Pour comparer l'impact des interventions, l'équipe de recherche a traduit la taille des effets observés dans le cadre d'évaluations en une seule mesure : mois supplémentaires de progrès éducatif pour les bénéficiaires (relatif à un contrefactuel).**
- la rigueur de la preuve. **Il s'agit d'une mesure à la fois de la qualité et du nombre d'études réalisées, avec une place centrale donnée aux évaluations d'impact avec contrefactuel ;**
- le coût. **Il s'agit d'une estimation du coût annuel par élève si des interventions étaient mises en œuvre dans des classes de 25 élèves, intégrant typiquement les coûts de la formation de l'enseignement ou la production de ressources pédagogiques.**

Plus d'informations sur educationendowmentfoundation.org.uk, rubrique Our work/About the toolkit .

Extrait du Teaching & Learning Toolkit



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

► Pour en savoir plus : étude de cas *Education Endowment Foundation* (page 126)

LES AUTRES BIBLIOTHÈQUES D'INTERVENTIONS DES WHAT WORKS CENTRES

What Works Centre for Local Economic Growth

- Policy Reviews : www.whatworksgrowth.org, rubrique Policy reviews.
- Toolkits : www.whatworksgrowth.org, rubrique Ressources.

Crime Reduction Toolkit : whatworks.college.police.uk, rubrique Toolkit.

Programmes Library du Early Intervention Foundation :

guidebook.eif.org.uk, rubrique Programmes library.

NICE Guidance : www.nice.org.uk www.nice.org.uk/guidance.

Accessibles pour un public non-expert, **ces bibliothèques permettent de démocratiser l'accès à la recherche** pour les professionnels et cadres intermédiaires. Interactives et présentant des analyses comparatives

What Works Centres britanniques : **quels enseignements pour des politiques fondées sur la preuve en France ?**

des coûts de différentes interventions, elles sont également de véritables **outils d'aide à la décision**. Ces outils représentent une rupture avec la présentation habituelle des résultats de recherche. En effet, avant leur création, une grande partie de la recherche scientifique circulait principalement entre chercheurs dans des revues académiques payantes. Trois des centres disposant d'une telle bibliothèque ont particulièrement insisté sur leur importance pour ouvrir un dialogue avec les professionnels autour de la recherche (*local economic growth, crime reduction et education*). Nous disposons de peu d'éléments d'évaluation sur l'utilisation de ces outils, à l'exception de celui de l'*Education Endowment Foundation*. D'après une enquête menée par le *National Audit Office* en 2015 auprès d'un échantillon de 543 chefs d'établissements, le *Teaching & Learning Toolkit* est utilisé par 64% d'entre eux pour informer leurs décisions sur les pratiques pédagogiques^{lviii}.

“Le Toolkit répond aux contraintes de temps auxquels les professionnels de terrain font face : manque de temps, trop d'informations, incapacité à identifier celles qui sont utiles. Le Toolkit a changé la donne en rendant la recherche sur l'éducation utile pour des professionnels.”

Jonathan Sharples, chercheur expérimenté, Education Endowment Foundation.

Rendre ainsi la recherche accessible pour les professionnels a donc été une étape nécessaire et reconnue comme utile.

Pendant, aucun des centres ne considère qu'une simple diffusion de la recherche – aussi claire soit-elle – soit suffisante pour faire évoluer à grande échelle les pratiques sur le terrain. Par conséquent, les **WWC accompagnent tous l'adoption de cette matière scientifique par les parties prenantes.**

“Créer des ressources ne suffit pas. Le transfert de la connaissance est un processus social, il faut créer des motivations et des opportunités de faire vivre la connaissance.”

Jonathan Sharples, chercheur expérimenté, Education Endowment Foundation.

Certains centres accompagnent des décideurs nationaux et régionaux à intégrer la recherche dans les politiques qu'ils conduisent. Les centres territoriaux et, en particulier, celui du Pays de Galles, sont davantage orientés vers les décideurs nationaux. Véritable « service What Works » pour l'Assemblée du Pays de Galles, le *Public Policy Institute for Wales* est en dialogue continu avec les ministres, leurs conseillers et les hauts fonctionnaires. La majorité de leurs livrables est ainsi produite rapidement pour éclairer une décision politique et fait l'objet d'une restitution orale auprès de ces décideurs. La plupart des centres thématiques accompagne de manière plus ponctuelle les décideurs nationaux à s'approprier la recherche. L'*Education Endowment Foundation* a ainsi produit une note d'information à l'intention du Premier ministre pour l'informer des résultats positifs des expérimentations conduites à l'égard des aides-enseignants.

La décentralisation a également ouvert des opportunités d'influencer des décisions politiques au niveau régional ou local. En 2012, les premières élections des Police and Crime Commissioners ont eu lieu en Angleterre et au Pays de Galles, avec l'objectif de rendre les services policiers plus responsables vis-à-vis du public. Ces élections ont eu lieu avec le lancement officiel du WWC for Crime Reduction. Le College of Policing – qui a émergé peu après le WWC – a créé un panel académique afin de synthétiser les enseignements de la recherche pour alimenter les programmes des Commissioners. Cette vaste matière scientifique a ensuite été présentée sous quatre questions clés que posent les décideurs des services policiers^{lix}.

Les centres thématiques visent à favoriser l'utilisation de la recherche principalement par les cadres intermédiaires et professionnels de terrain. Les centres produisent trois types d'outils pour accompagner cette appropriation :

- des bibliothèques et revues systématiques pour les sensibiliser aux types d'intervention qui semblent globalement plus efficaces. Par exemple, le rapport sur les relations parents-enfants de l'*Early Intervention Foundation* a identifié 17 programmes comme ayant des impacts positifs et mesurables, et 18 programmes comme prometteurs^x;
- des outils d'auto-évaluation qui permettent aux acteurs de comparer

What Works Centres britanniques : **quels enseignements pour des politiques fondées sur la preuve en France ?**

leurs pratiques actuelles aux interventions dites « efficaces ». L'EIF a ainsi produit une « *Maturity matrix* » pour la petite enfance, qui permet aux collectivités locales d'identifier où elles en sont en termes de connaissance de leurs publics, du développement de leurs équipes ou de leurs services, et des leviers d'action, etc. (voir étude de cas page 114)^{lxii} ;

– des descriptifs d'intervention ou guides, qui définissent plus précisément les modalités de l'intervention et les aspects contextuels pouvant influencer sur les effets. Le *WWC for Local Economic Growth* a ainsi produit, pour chaque politique publique, une boîte à outils composée de plusieurs guides portant sur des interventions spécifiques. Par exemple, la synthèse sur la formation professionnelle a été transformée en boîte à outils composée de quatre guides : l'orientation professionnelle ; les incitations financières ; les pré-qualifications ; et les rappels (pour améliorer la participation aux formations)^{lxiii}.

Les centres diffusent ces outils par voies numériques, mais aussi dans le cadre de journées de restitution, lors de conférences et dans des formations *ad hoc*.

Les centres cherchent ainsi à couvrir toute la chaîne, de la recherche à l'appropriation de ces pratiques par les professionnels. Dans l'idéal, les professionnels peuvent ainsi facilement mettre en œuvre des pratiques qui sont « l'état de l'art » en termes d'efficacité.

LES AIDES-ENSEIGNANTS : DE LA REVUE À L'APPROPRIATION PAR LES PROFESSIONNELS DE L'ÉDUCATION

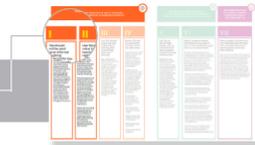
Les travaux de l'EEF autour des aides-enseignants (*teaching assistants* en anglais) traduisent ce principe en réalité. Disposant *a minima* d'un certificat général d'études secondaires du premier cycle (GCSE), ces aides-enseignants appuient les enseignants, notamment dans le soutien scolaire des élèves en difficulté. La revue systématique a montré que les aides-enseignants observaient, en moyenne, soit pas d'effets soit des effets négatifs pour les élèves ciblés. Ce constat a particulièrement interpellé l'équipe car les aides-enseignants représentent une part importante des professionnels de l'éducation en Angleterre : 240 000 en poste pour un coût annuel

de 5 milliards de livres sterling. L'équipe EEF voulait donc expérimenter différentes manières de déployer ces professionnels. « On a mené des expérimentations qui ont montré que, lorsqu'ils sont déployés de façon ciblée, les aides-enseignants ont un impact positif pour les élèves » a expliqué Jonathan Sharples d'EEF. EEF a ensuite décidé de traduire les enseignements de la recherche sur le déploiement des aides-enseignants (dont les six expérimentations financées par EEF) en un guide opérationnel. Ce guide a été accompagné d'outils pour favoriser l'appropriation des bonnes pratiques, tels qu'une auto-évaluation des pratiques des aides-enseignants. Ces outils ont été diffusés à chaque établissement scolaire en Angleterre dans le cadre d'une campagne d'essai « grandeur nature ».

Aides-enseignants : extrait de l'outil d'auto-évaluation

A self-assessment guide

RECOMMENDATIONS I AND II Deployment of TAs in the classroom



1 INEFFECTIVE

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.

This arrangement causes a 'separation' effect: high-supported pupils have fewer opportunities to engage with whole-class teaching, and have fewer interactions with teachers and peers.

Teachers assume that a high amount of TA support meets the needs of lower-attaining and/or SEN pupils, and as a result spend less time working with these groups.

Personal relationship may be good, but there is a lack of cohesive and effective teamwork between teachers and TAs in the classroom.

TAs rarely rove around the classroom and support whole-class teaching (such as demonstrating equipment, or working at the whiteboard).

2 IMPROVING

There is variability in the quality and appropriateness of teachers' moment-by-moment decision-making regarding TA deployment.

In some lessons, TAs are effectively supplementing the whole-class teaching, although this is not consistent.

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.

There is little or no attempt to redeploy TAs during lessons in response to the 'real time' needs of pupils.

3 EXEMPLARY

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.

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.

Teachers and TAs work effectively as a team, with a shared understanding of their respective roles in achieving lesson objectives. Teachers' moment-by-moment decisions regarding TA deployment are driven by pupils' needs.

TAs support whole-class delivery at the front of the class, rove the room or undertake administrative tasks to ease teachers' workload.

Source : educationendowmentfoundation.org.uk

What Works Centres britanniques : quels enseignements pour des politiques fondées sur la preuve en France ?

Certains accompagnent directement le changement d'échelle des projets « efficaces », comme le *Centre for Ageing Better* qui lance des appels à projets « scale up ». D'autres centres ont décidé de mener des actions d'acculturation globale des professionnels à l'évaluation d'impact et des politiques basées sur la preuve. Par exemple, le *WWC for Crime Reduction*, hébergé par l'organisme chargé de la formation des policiers (*College of Policing*), a intégré des modules sur l'évaluation et la « preuve » dans la formation initiale et continue des officiers. Plusieurs WWC ont souligné que cette acculturation des professionnels est une pré-condition pour leur succès et doit s'inscrire dans la durée.

“On a besoin d'au moins 10 ans pour faire ce que les financeurs attendent. On avait besoin de 3 ans pour traiter une masse critique de la recherche. Maintenant, on doit communiquer au sujet de la preuve au niveau local... une grande partie de notre rôle est d'impulser un changement culturel et institutionnel. Cela prend beaucoup de temps.”

Max Nathan, directeur adjoint, What Works Centre for Local Economic Growth (LSE).

Conscients qu'il est difficile de changer les pratiques professionnelles et les dispositifs privilégiés par des organismes gestionnaires sans une mobilisation collective sur un territoire, **plusieurs centres mettent en œuvre des actions d'essaimage plus intensives sur des territoires spécifiques :**

- les huit « *implementation consultants* » de NICE sont des courtiers de la connaissance au niveau régional, sensibilisant les acteurs de la santé – services hospitaliers, généralistes, collectivités territoriales – aux ressources proposées par NICE.
- dans le cadre du programme *Pioneering Places*, EIF a travaillé avec 20 territoires pour identifier des actions permettant de mieux intégrer le principe de la prévention précoce dans leurs services^{lxiii}. EIF a préconisé des actions permettant d'améliorer l'identification de la population cible et ensuite a déterminé la meilleure manière de les accompagner, s'appuyant sur les interventions efficaces identifiées par ses revues systématiques. Le *Centre for Ageing Better* a également noué

un partenariat avec le *Grand Manchester*, il s'agit d'expérimenter avec eux, mais aussi d'éclairer les décisions locales avec « la preuve ».

– l'EEF a élaboré un menu d'interventions identifiées comme efficaces et déjà proposées par des opérateurs associatifs en Angleterre.

Dans le cadre d'un premier projet pilote dans une collectivité près de Londres, les écoles volontaires peuvent bénéficier d'un cofinancement et d'un soutien de l'EEF pour l'adoption de ces pratiques.

Les WWC testent divers modèles de diffusion et d'essaiage de « ce qui fonctionne ». Ce sont de véritables laboratoires des politiques basées sur la preuve, dont les résultats ne sont pas encore conclusifs. Certes, ces centres sont généralement bien identifiés par les professionnels de leurs domaines. Les centres ont également eu quelques réussites notables en termes d'essaiage des pratiques « efficaces ».

L'ESSAIMAGE DES CAMÉRAS VIDÉO CORPORELLES (CVC) À LONDRES

Les caméras vidéo corporelles sont des appareils portés par des policiers, permettant d'enregistrer toutes leurs interactions et leurs opérations. Afin de tester les apports de ces appareils, le WWC for Crime Reduction a réalisé une évaluation randomisée à Londres^{lxiv}. Les résultats positifs de cette évaluation ont été largement diffusés et le Metropolitan Police a annoncé, en octobre 2016, que 22 000 officiers de la police à Londres seront équipés de tels appareils^{lxv}.

Cependant, **les WWC disposent de peu d'éléments objectifs quant à leur impact sur les pratiques professionnelles.** Les équipes reconnaissent ouvertement qu'ils ne savent pas encore ce qui « fonctionne » en matière d'essaiage des pratiques efficaces. Les centres cherchent donc tous à améliorer leurs actions d'essaiage, certains allant même jusqu'à les évaluer. Ainsi, le *WWC for Crime Reduction* fait l'objet d'une évaluation globale et continue de son action^{lxvi}. L'*Education Endowment Foundation* a même décidé de soumettre différentes méthodes de diffusion à une véritable expérimentation aléatoire.

Évaluer l'impact d'innovations

Le financement d'innovations et la réalisation d'évaluations d'impact ne constituent pas une mission centrale de tous les centres. En effet, seulement deux WWC financent ou réalisent directement des expérimentations et des évaluations d'impact.

L'Education Endowment Foundation est de loin le WWC le plus prolifique en matière de lancement de nouvelles expérimentations. Dotée de 125 millions de livres sterling, l'EEF finance environ 20 expérimentations chaque année. EEF finance les deux volets de l'expérimentation : le projet innovant et l'évaluation d'impact. Les projets sont conçus et portés par des acteurs de terrain – écoles, associations, voire universitaires – et sélectionnés par appel à projets. En finançant le fonctionnement du projet, EEF incite les structures à se soumettre à une évaluation scientifique, malgré les risques que cela peut représenter si les impacts s'avèrent nuls ou négatifs. En fonction du stade de développement, chaque projet est attribué à un niveau d'accompagnement : projet pilote, expérimentation d'efficacité et expérimentation à grande échelle. Plusieurs projets jugés prometteurs ont fait l'objet de pilotes permettant de tester leur faisabilité, sans les soumettre à une évaluation d'impact. À l'exception de ces projets pilotes, pour chaque expérimentation l'EEF finance une évaluation d'impact rigoureuse réalisée par un évaluateur externe. Les méthodes expérimentales sont privilégiées pour analyser les effets du projet. EEF a rencontré des difficultés techniques et, parfois, des résistances ou des méconnaissances de la part des porteurs de projet. Cependant, l'EEF a globalement donné tort à ses détracteurs qui craignaient une résistance forte de la part des enseignants aux méthodes d'évaluation expérimentales. En effet, il a engagé 127 évaluations dont 105 avec « randomisation », impliquant 7000 écoles, soit une école sur quatre en Angleterre.

“Il y a quelques années il y avait seulement eu une évaluation « randomisée » dans le secteur de l'éducation en Angleterre. Tout d'un coup nous en avons plus de 100. C'est comme si on avait appuyé sur un bouton!”

Jonathan Sharples, chercheur expérimenté, Education Endowment Foundation.

Les expérimentations sociales sont alimentées par les revues d'EEF, et inversement. En effet, le *Toolkit* a également aidé l'EEF à identifier les types d'interventions pour lesquels la recherche méritait d'être renforcée et qui pourraient faire l'objet de nouvelles expérimentations financées par l'EEF. À terme, ces expérimentations produisent des résultats qui alimentent le *Toolkit*. L'exemple des aides-enseignants (cf. pages 66-67) illustre bien ce cycle vertueux.

“Les évaluations montraient que les aides-enseignants n'avaient pas d'effets ou des effets négatifs sur les élèves. Le Toolkit a fait ressortir cette problématique. Ce constat nous a conduit à plusieurs. On a mené des expérimentations qui ont montré que lorsqu'ils sont déployés de façon ciblée, les aides-enseignants ont un impact positif pour les élèves. Ces résultats ont, à leurs tours, fait évoluer les constats dans notre Toolkit.”

Jonathan Sharples, chercheur expérimenté, Education Endowment Foundation.

Les analystes du WWC for Crime Reduction évaluent également l'impact des innovations. En raison de ses moyens limités, ce centre a évalué l'impact de seulement quelques projets innovants mis en œuvre par les services de la police dans le cadre de projets pilotes. Partant des enseignements des projets pilotes, le centre cible des innovations à fort potentiel pour leurs expérimentations. L'expérimentation de l'essai de caméras vidéo corporelles (CVC) à Londres illustre bien cette approche^{lxvii}.

Enfin, le **Centre for Ageing Better ne porte pas à ce stade d'expérimentation.** Cependant, il dispose des moyens dédiés pour la mise en œuvre de programmes pilotes et envisage de mobiliser des financements du monde académique pour en financer l'évaluation d'impact.

Constatant le manque d'évaluations scientifiques sur leur domaine, certains centres tentent d'inciter au lancement de nouvelles évaluations d'impact.

Par exemple, les *Guidances* de NICE préconisent souvent des sujets qui devraient faire l'objet d'essais cliniques ou d'expérimentations. Le *WWC for Local Economic Growth*, pour sa part, propose un accompagnement

What Works Centres britanniques : **quels enseignements pour des politiques fondées sur la preuve en France ?**

individualisé au cadrage et au lancement d'expérimentations dans son domaine, notamment auprès des collectivités territoriales. Ce travail s'est révélé souvent laborieux. Ainsi, à l'été 2016, bien que l'équipe soit impliquée dans la conception de 20 projets pilotes, seulement 4 sont sur le point d'être effectivement lancés. Afin de « relever » la qualité des évaluations d'impact, **plusieurs centres mènent des actions de sensibilisation aux méthodes d'évaluation d'impact**. Dans certains cas, ces actions commencent à porter leurs fruits.

“On travaille avec les ministères pour améliorer la qualité des évaluations qu'ils financent. Aujourd'hui, certains appels d'offre de l'État exigent a minima une évaluation correspondant au niveau 3 de l'échelle scientifique de Maryland et demandent aussi une évaluation de processus.”

Max Nathan, directeur adjoint, What Works Centre for Local Economic Growth (LSE).

Plusieurs centres ont exprimé leur frustration quant à leur incapacité à évaluer l'impact d'innovations dans le contexte britannique. À l'inverse, la capacité à soutenir des expérimentations a été un réel atout pour l'EEF et *WWC for Crime Reduction* : d'une part, pour impliquer les acteurs de terrain en tant que porteurs d'innovation et, d'autre part, pour pousser la recherche sur « ce qui fonctionne ».

Les enseignements clés sur les What Works Centres

En matière de synthèse de la recherche existante

- **la plupart des What Works Centres ont adopté une approche plutôt « rigoriste », avec un accent fort sur les évaluations d'impact avec contrefactuel, les revues systématiques d'évaluation et des analyses de coûts ;**
- **cette approche a eu des avantages, permettant aux centres d'identifier la relative efficacité d'un panel d'interventions et donc d'apporter des enseignements utiles aux décideurs et professionnels ;**

- cependant, la production des revues systématiques n'a pas été sans difficulté, en particulier en raison du manque d'évaluations d'impact scientifiques, de la complexité à déterminer l'efficacité d'une intervention et à clarifier les contextes dans lesquels une intervention fonctionne;
- ces difficultés ont conduit plusieurs des centres à nuancer leur approche, en intégrant des évaluations sans contrefactuel, ou en prenant en compte la question de la transposabilité.

En matière de diffusion et d'accompagnement à l'adoption des résultats

- les WWC ont transformé des centaines d'évaluations en enseignements intelligibles notamment dans leurs revues et bibliothèques d'intervention. Les centres se sont rendus ainsi utiles pour des décideurs et les professionnels;
- rendre cette matière scientifique accessible a été nécessaire mais pas suffisant pour que les acteurs s'en saisissent;
- certains centres accompagnent donc des décideurs nationaux et régionaux à intégrer la recherche dans les politiques qu'ils conduisent;
- les centres ciblent toutefois principalement les cadres intermédiaires et professionnels du terrain, élaborant divers outils pour les accompagner à adopter des interventions efficaces. Plusieurs centres mettent en œuvre des actions d'essai territorial afin de favoriser une mobilisation collective des professionnels;
- les WWC ne prétendent pas avoir la « science infuse » en matière des méthodes d'essai. Ils continuent de tester différentes approches, y compris dans le cadre d'une expérimentation;
- certains WWC jugent qu'ils ne pourront réussir qu'à la condition d'intégrer l'évolution des cultures professionnelles à leurs missions, qui nécessite un long travail de sensibilisation et de formation.

En matière d'évaluation d'impact

- pour l'*Education Endowment Foundation* et *WWC for Crime Reduction*, la capacité à soutenir des expérimentations a été un véritable atout, pour améliorer la connaissance de ce qui « fonctionne » dans le contexte britannique et pour les rapprocher des acteurs de terrain;

What Works Centres britanniques : quels enseignements pour des politiques fondées sur la preuve en France ?

– l'action des autres centres se limite globalement à la formation, l'accompagnement et la sensibilisation à l'évaluation d'impact, avec un bilan globalement mitigé à ce stade en termes de nouvelles expérimentations lancées.

xxxvi Voir sur le site www.rssnews.org.uk/articles/20100604_1.

xxxvii Balas, EA, Boren SA. *Managing clinical knowledge for health care improvement*. In : Bommel J, McCraty AT, editors. *Yearbook of Medical Informatics 2000 : Patient-Centred Systems*. Stuttgart, Germany : Schattauer Verlagsgesellschaft mbH; 2000, pages 65-70.

xxxviii Cabinet Office, *What works? : Evidence for decision makers*, 2014, Cabinet Office.

xxxix Cabinet Office (1999) *Modernising government* London, Stationery Office, 66pp. Voir le site www.cabinet-office.gov.uk/moderngov/whtpaper/index.htm.

xl Ces initiatives sont énumérées dans un livre publié en 2000, qui utilise déjà le terme « what works » : Davies, H.T.O., Nutley, S.M. and Smith, P.C. (eds) (2000) *What works? Evidence-based policy and practice in public services*, Bristol: Policy Press.

xli Par exemple, l'expérimentation « Employment Retention and Advancement (ERA) Demonstration » conduite entre 2003 et 2007.

xlii Porté par Kings College London en collaboration avec d'autres universités, l'objectif principal du centre était de favoriser l'échange de la recherche pertinente entre chercheurs et praticiens dans le domaine des politiques publiques, notamment par la synthèse des résultats de la recherche. Source : www.researchcatalogue.esrc.ac.uk/grants/H141251005/read.

xliii MULGAN Geoff, « Introduction », dans PUTTICK Ruth, *Using Evidence to Improve Social Policy and Practice : Perspectives on how research and evidence can influence decision making*, 2011, Alliance for Useful Evidence/NESTA. PUTTICK Ruth, *Why we need to create a "NICE for Social Policy"*, May 2012, NESTA.

xliv WALKER David, *Would a version of the health standards body Nice work for social policy?*, 11 June 2012, The Guardian. Disponible en ligne : www.theguardian.com/public-leaders-network/2012/jun/11/nice-standards-for-policy-making.

xlv Graham Allen MP, *Early Intervention: The Next Steps An Independent Report to Her Majesty's Government*, HM Government, 2011. Disponible en ligne : <http://grahamallenmp.co.uk/static/pdf/early-intervention-7th.pdf>.

xlvi HM Government, *Open Public Services White Paper*, HM Government, 2011, page 29.

xlvii WINTOUR Patrick, Cabinet secretary calls for social policy 'kite mark' to highlight quality initiatives, 10 January 2012, The Guardian. Disponible en ligne : www.theguardian.com/politics/2012/jan/10/cabinet-secretary-social-policy-kite-mark.

xlviii Il s'agit des centres sur la prévention de la criminalité, le développement économique local, le bien-être et l'Écosse.

xlix Disponible en ligne : www.nice.org.uk/process/pmg20/chapter/1-introduction-and-overview#whos-involved.

I MADALENO Margarida et WAIGHTS Sevrin, *Guide to scoring methods using the Maryland Scientific Methods Scale*, What Works Centre for Local Economic Growth. Disponible en ligne : www.whatworksgrowth.org/resources/scoring-guide.

II What Works Centre for Local Economic Growth, *About our reviews*, 2014. Disponible en ligne : www.whatworksgrowth.org/public/files/Methodology/14-03-20_About_our_Reviews.pdf.

III ASMUSSEN Kirsten, FEINSTEIN Leon, MARTIN Jack, CHOWDRY Haroon, *Foundations for Life : What works to support*

- parent child interaction in the early years*, July 2016, Early Intervention Foundation. Disponible en ligne : www.eif.org.uk/publication/foundations-for-life-what-works-to-support-parent-child-interaction-in-the-early-years.
- liii PUTTICK Ruth, LUDLOW Joe, *Standards of Evidence*, Nesta, 2013. Disponible en ligne : www.nesta.org.uk/sites/default/files/standards_of_evidence.pdf.
- liv <https://educationendowmentfoundation.org.uk/resources/teaching-learning-toolkit>.
- lv Heckman, J., Moon, S., Pinto, R., Savelyev, P., Yavitz, A. "The rate of return to the HighScope Perry Preschool Program", *Journal of Public Economics*, Vol. 94, Issues 1–2, 2010, pp. 114–128. URL : www.sciencedirect.com.
- lvi JOHNSON Shane, TILLEY Nick, & BOWERS Kate, "Introducing EMMIE: an evidence rating scale to encourage mixed-method crime prevention synthesis reviews", *Journal of Experimental Criminology*, Vol 11, 2015, pp.459–473.
- lvii NUTLEY Sandra, POWELL Alison, et DAVIES Huw, *What counts as good evidence?*, Alliance for Useful Evidence, 2013.
- lviii Il s'agit d'un sondage de 543 directeurs d'établissements scolaires, réalise début 2015. Source : National Audit Office, *Funding for disadvantaged pupils Survey evidence from pupils, parents and school leaders*, 30 June 2015, page 30. www.nao.org.uk/report/funding-for-disadvantaged-pupils.
- lix whatworks.college.police.uk/Research/overview/Pages/default.aspx.
- lx ASMUSSEN Kirsten, FEINSTEIN Leon, MARTIN Jack, CHOWDRY Haroon, *Foundations for Life : What works to support parent child interaction in the early years*, July 2016, Early Intervention Foundation.
- lxi Voir sur le site : www.eif.org.uk/publication/early-intervention-maturity-matrix.
- lxii Voir sur le site : www.whatworksgrowth.org/resources/employment-training-toolkit.
- lxiii Voir sur le site : www.eif.org.uk/pioneering-place.
- lxiv GROSSMITH Linda et al. *Police, Camera, Evidence: London's cluster randomised controlled trial of Body Worn Video*, College of Policing, November 2015. Voir sur le site : whatworks.college.police.uk/Research/Documents/Police_Camera_Evidence.pdf.
- lxv Voir sur le site : news.met.police.uk/news/rollout-of-body-worn-cameras-191380.
- lxvi Voir sur le site : whatworks.college.police.uk/About/Pages/Evaluation.aspx.
- lxvii GROSSMITH Linda et al. *Police, Camera, Evidence: London's cluster randomised controlled trial of Body Worn Video*, College of Policing, November 2015. Voir sur le site : whatworks.college.police.uk/Research/Documents/Police_Camera_Evidence.pdf.

Quelle culture de l'évaluation et de la preuve pour les politiques publiques en France ?

Un dispositif administratif orienté vers une compréhension fine des problématiques

L'observation des problématiques (enjeux, publics, besoins) est un élément central de la « preuve ». En effet, elle permet d'élaborer des politiques publiques en connaissant à la fois les enjeux et les problèmes posés mais aussi les publics concernés ainsi que leurs caractéristiques et leurs besoins.

La France dispose pour cela d'un réseau dense d'organismes dont la fonction principale est l'observation des problématiques de la société française, en particulier des services statistiques, des observatoires et des laboratoires de recherche. Au niveau national, cette fonction est assurée par les services statistiques nationaux (Institut national de la statistique et des études économiques, Institut national d'études démographiques) et ministériels (par exemple la direction de l'évaluation, de la prospective et de la performance [Depp] au ministère de l'Éducation nationale, la direction de l'animation de la recherche, des études et des statistiques [Dares] au ministère du Travail et de l'Emploi, la direction de la recherche, des études, de l'évaluation et des statistiques [Drees] au ministère des Affaires sociales et de la Santé), par des organismes publics ou des directions et des équipes au sein de ces organismes (par exemple France Stratégie, CGET, Santé publique France, SGMAP), également au sein des organismes de sécurité sociale (par exemple Cnaf, CNSA, Cnav). Il existe également des équipes dotées de cette fonction d'observation au niveau local, au sein des collectivités ou des services déconcentrés de l'État, ainsi que des structures indépendantes (par exemple les observatoires régionaux de santé). Les laboratoires de recherche jouent également un rôle central, en réalisant des projets de recherche, parfois pour ou avec les organismes cités ci-dessus. Grâce à ces acteurs, la France dispose d'un vivier riche de données et d'études traitant la nature et l'envergure de ces

problématiques. Les travaux et recherches sont quantitatifs et qualitatifs, en principe orientés vers la décision publique.

Des investissements financiers et humains et des avancées technologiques ont amélioré la qualité et la réactivité des structures d'observation en France au cours des dernières années. Grâce à ces efforts, la France dispose désormais d'un nombre croissant d'enquêtes et d'études périodiques ou longitudinales permettant d'analyser les caractéristiques et les évolutions des situations et des besoins des personnes (comme par exemple les enquêtes emploi de la Dares). Des améliorations notables ont eu lieu en matière d'ouverture et de partage des données publiques (« open data »), une politique pilotée par l'Etalab et l'Administrateur général des données. Le site data.gouv.fr, par exemple, permet l'accès à plus de 355 000 données publiques gratuites et réutilisables.

Malgré ces avancées, les difficultés relatives aux données d'observation en France persistent. En effet, il existe de nombreux domaines pour lesquels il n'y a pas de production homogène de données sur tout le territoire national, ni de systèmes d'information permettant leur exploitation et diffusion. Par exemple, dans le champ du handicap, les maisons départementales des personnes handicapées (MDPH) ayant développé chacune leur propre système d'information, les données qui y sont renseignées ne peuvent être exploitées. La conception par la CNSA d'un cadre national harmonisé de leur système d'information et l'appui à son déploiement permettra d'y remédier. Certaines de ces données seront par ailleurs intégrées au nouveau Système national des données de santé (SNDS). Le SNDS, en cours de création, allie deux approches : un accès libre aux données anonymes et un accès sécurisé aux données personnelles de santé à des fins de recherches, d'études ou d'évaluation d'intérêts publics. Cet accès sécurisé ouvre d'ailleurs la possibilité de réaliser des évaluations d'impact (quasi expérimentales) à moindre coût pour des échantillons importants.

L'évaluation des politiques publiques

Parmi les outils d'analyse de l'action publique, l'évaluation des politiques publiques a connu un essor important depuis les années 1960, sans atteindre une institutionnalisation totalement aboutie^{lxviii}. Cette méthode, issue de la rencontre des sciences sociales américaines et des décideurs nationaux dans plusieurs domaines d'action publique^{lxix}, s'est d'abord développée aux États-Unis.

Le recours à l'évaluation des politiques publiques remonte, en France, aux années 1960 avec la Rationalisation des choix budgétaires (RCB). Portée par des hauts fonctionnaires de la direction du budget, la RCB (1965-1977) consistait en un ensemble de programmes d'actions visant à fonder les décisions de dépenses publiques sur la base de la prévision de leurs impacts socio-économiques. À la fin des années 1980, les initiatives du gouvernement de Michel Rocard inscrivent l'évaluation des politiques publiques dans un nouveau cycle de politiques de réformes administratives. Un dispositif partagé entre le Parlement, le Commissariat général au plan et une instance indépendante (Conseil scientifique de l'évaluation puis Conseil national de l'évaluation) consacre alors l'évaluation comme un vecteur déterminant de modernisation des politiques publiques et de renforcement de la légitimité démocratique de l'action publique. Parallèlement, la mise en œuvre des politiques contractuelles, principalement les fonds structurels européens et les contrats de plan État/régions, consacre l'obligation d'évaluer les programmes contractualisés, de même que la conduite des premières expérimentations locales ont permis la diffusion de pratiques d'évaluation au sein des collectivités territoriales^{lxx}.

À partir des années 2000, deux chantiers successifs de réforme de l'administration associent la promotion de l'évaluation des politiques publiques à la recherche d'une plus grande performance de l'intervention publique : la loi organique relative aux lois de finance (LOLF) et la révision générale des politiques publiques (RGPP). Si ces deux démarches ont, dans les faits, plus relevé de démarches d'audit et d'inspections, elles ont

contribué cependant à une mise en visibilité plus grande du lexique et des enjeux de l'évaluation au sein de l'État et des administrations.

En 2012, l'évaluation des politiques publiques a connu un nouveau tournant sous l'impulsion de la démarche de modernisation de l'action publique (MAP) dont l'objectif était d'évaluer « l'ensemble des politiques publiques » d'ici la fin du quinquennat. Le dispositif, de portée interministérielle a notamment mobilisé les corps d'inspections générales.

L'ÉVALUATION DE POLITIQUES PUBLIQUES : LE PROGRAMME DU SGMAP

Entre 2012 et 2016, le Secrétariat général de la modernisation de l'action publique (SGMAP) a mis en œuvre 80 évaluations de politiques publiques – dites « EPP » – dans le cadre de la modernisation de l'action publique^{boxi}. Les EPP impliquent les administrations et les opérateurs de l'État, les organismes de sécurité sociale et les collectivités territoriales. Ces évaluations, confiées le plus souvent à des services d'inspections générales mais également à des personnalités qualifiées, sont réalisées avec une mixité de méthodes (études documentaires, analyses quantitatives, enquêtes et consultation^{boxii}) et un appui d'ingénierie apporté par les équipes du SGMAP. Les évaluations aboutissent à des scénarios de transformation, composés de recommandations pour les décideurs politiques. Le suivi de la mise en œuvre de ces recommandations montre que deux EPP achevées sur trois ont connu des suites (décisions ministérielles, mise en place de plan d'action, promulgation d'une loi, etc.).

Par exemple, l'EPP *Éducation prioritaire* achevée en 2013 a contribué à des évolutions telles que la refonte de la cartographie des zones d'enseignement prioritaire, ou encore l'inscription de l'objectif de réduction des écarts de réussite scolaire dans la loi pour la refondation de l'école du 18 juillet 2013.

Par ailleurs, les collectivités territoriales investissent l'évaluation de manière

What Works Centres britanniques : quels enseignements pour des politiques fondées sur la preuve en France ?

croissante, et ce depuis une vingtaine d'années, au gré des lois de décentralisation, du poids de l'Europe et des réformes qui les touchent. En effet, les avancées produites par les premières expérimentations locales, notamment dans les années 90, ont permis de diffuser les pratiques d'évaluation aux différents niveaux territoriaux. « Des structures d'évaluation ont été mises en place d'abord par des collectivités de niveau régional, le plus souvent dans un souci d'autonomisation et de légitimation de cet échelon, avant d'être suivies par des conseils départementaux pour les compétences sociales et par les collectivités infraterritoriales (communes et intercommunalités) dans le cadre notamment des évaluations de la politique de la ville. » précise le rapport du LIEPP *Les scènes multiples de l'évaluation*. D'autre part, les politiques contractuelles (fonds européens, plans État-région) obligent à l'évaluation des politiques contractualisées^{6xiii}.

Au niveau des acteurs, la promotion de l'évaluation des politiques publiques est portée par un ensemble composite d'administrations et d'organismes, au sein et en dehors de l'État :

- **les corps d'inspection générale**, dont l'Inspection générale des affaires sociales (IGAS), et l'Inspection générale des finances (IGF) ;
- **la Cour des comptes** qui s'est vu confier une prérogative d'évaluation des politiques publiques par la réforme constitutionnelle de 2008 ;
- des **directions de ministères** : en particulier, la Direction de l'animation de la recherche, des études et des statistiques (Dares), la Direction de la recherche, des études, de l'évaluation et des statistiques (Drees), et la Direction de l'évaluation, de la prospective et de la performance (Depp) ;
- Plusieurs **acteurs interministériels** dont :
 - le Secrétariat général pour la modernisation de l'action publique (**SGMAP**) via son programme des EPP-MAP ;
 - **France Stratégie**, dont le décret prévoit qu'il participe à l'évaluation des politiques publiques, s'est vu confier plusieurs missions d'évaluation, dont le Comité de suivi des aides publiques aux entreprises et des engagements⁷. Il a également publié un guide de l'évaluation⁸ ;

7 Également la Commission nationale d'évaluation des politiques d'innovation (CNEPI) et la Commission d'étude des effets de la loi pour la croissance et l'activité.

8 Rozenn Desplat, Marc Ferracci, *Comment évaluer l'impact des politiques publiques ? Un guide à l'usage des*

- le Commissariat général à l'égalité des territoires (**CGET**), dont la compétence évaluative est intégrée au sein de la « direction des stratégies territoriales » ;
- **le législateur** : avec le Comité d'évaluation et de contrôle des politiques publiques (CEC) de l'Assemblée nationale ;
- **des collectivités locales** sont particulièrement investies sur les questions d'évaluation, au niveau régional, départemental et métropolitain ;
- **le Conseil économique, social et environnemental (Cese)**, dont les conseillers sont issus de la société civile, s'est doté depuis 2010 d'une Délégation à la prospective et à l'évaluation des politiques publiques. Voir également ses déclinaisons en régions (Ceser) ;
- **des acteurs associatifs et privés** dont plusieurs centaines sont adhérents de la Société française d'évaluation (SFE) créée en 1998. Ses membres mettent en œuvre des évaluations de politiques publiques et s'engagent à respecter des principes de pluralisme et d'indépendance établis dans la charte de la SFE ;
- **les chercheurs**, par exemple ceux réunis au sein du Laboratoire interdisciplinaire d'évaluation des politiques publiques (Liepp) de Sciences Po qui développe des approches économiques et quantitatives dans ses axes de recherche portant sur l'évaluation ou l'École d'économie de Paris, qui héberge J-Pal Europe et l'Institut des politiques publiques (IPP).

Ce panorama de l'évaluation des politiques publiques témoigne d'un paradoxe en matière de partage d'une culture française de l'évaluation. D'un côté, il atteste d'un engouement indéniable pour une démarche et un domaine de recherche aujourd'hui bien inscrit dans le paysage. De l'autre, il met en évidence les faiblesses d'une démarche non encore totalement stabilisée, tant du point de vue de ses outils et méthodes que du point de vue institutionnel. La sphère de l'évaluation fait coexister des approches méthodologiques et stratégiques sinon

décideurs et praticiens, France Stratégie, Septembre 2016. Ce document de travail propose un panorama des différentes méthodes d'évaluation des politiques publiques, de leurs avantages et leurs limites : l'expérimentation contrôlée (ou aléatoire), l'expérience naturelle, l'analyse coût/bénéfice, la modélisation (approche structurelle), la méthode qualitative (observation d'un échantillon, entretiens, etc.), etc. Des exemples d'utilisation de ces différentes démarches d'évaluation viennent illustrer ce guide.

concurrentes du moins divergentes, que les commentateurs ont pu qualifier de « chapelles » de l'évaluation^{boxiv}.

À cela s'ajoute des défis en termes de publicité, de diffusion et de mobilisation des travaux d'évaluation. Ils restent dans leur grande majorité sous appropriés par les décideurs et les praticiens et méconnus d'un plus large public^{boxv}. Les rapports d'évaluation ne sont pas systématiquement rendus publics ; les revues scientifiques qui publient des travaux d'évaluation sont généralement payantes. À un autre niveau, une des explications avancées est la faible lisibilité des rapports d'évaluation. Une autre hypothèse serait que l'approche d'évaluation des politiques n'apporte pas toujours des réponses à certaines questions que posent les acteurs, décideurs ou professionnels de terrain.

L'émergence de l'évaluation d'impact et de l'expérimentation sociale

Un développement rapide de ces nouvelles méthodes

Contrairement à d'autres pays développés, notamment anglo-saxons, l'évaluation d'impact est une pratique qui s'est implantée relativement récemment dans les politiques sociales en France. Plusieurs facteurs ont conduit à un développement rapide au cours des dix dernières années. La loi organique du 1^{er} août 2003 relative à l'expérimentation a ouvert la voie à ceux qui souhaitent conduire des expérimentations sociales en France. Cependant, il a fallu attendre 2007 pour que cette pratique se développe notamment sous l'impulsion de Martin Hirsch. Ce portage, au sein de la société civile et de la sphère institutionnelle et politique, s'est traduit par un investissement conséquent dans des évaluations d'impact, notamment via l'expérimentation du revenu de solidarité active (RSA) et la création du Fond d'expérimentation pour la jeunesse. Cette dynamique a eu quelques réussites, par exemple en identifiant des dispositifs efficaces qui ont fait l'objet d'un essaimage (voir page 95). Certaines expérimentations menées ont également produit des

enseignements utiles qui étaient portés à connaissance des décideurs publics (voir page 89). Cet essor a ouvert la perspective de disposer de méthodes rigoureuses permettant de concentrer les dépenses publiques sur les projets dont les effets auraient été démontrés. À cette ouverture ont succédé des interrogations, en raison notamment des difficultés de mise en œuvre des évaluations d'impact et de l'insuffisante capitalisation sur les expérimentations déjà réalisées.

**UNE DYNAMIQUE QUI SE POURSUIT ET ÉVOLUE :
LE FOND D'EXPÉRIMENTATION POUR LA JEUNESSE**

Le Fond d'expérimentation pour la jeunesse (FEJ) a été créé pour financer des actions innovantes en faveur des jeunes, mises en œuvre à une échelle limitée et évaluées rigoureusement. Son animation est assurée par la mission d'animation du FEJ (MAFEJ) rattachée initialement à la Direction de la jeunesse, de l'éducation populaire et de la vie associative (DJEPVA).

Entre 2009 et 2012, le FEJ a bénéficié d'une dotation de 230 millions d'euros de la part de l'État et de contributeurs privés, notamment la Fondation Total, à hauteur de 31 millions d'euros et l'Union des industries et des métiers de la métallurgie (UIMM), à hauteur de deux millions d'euros. Pendant cette période, chaque expérimentation a associé une structure porteuse d'un projet à un évaluateur indépendant. L'objectif est ainsi de connaître les effets des projets et d'orienter les politiques publiques vers les dispositifs les plus efficaces. Entre 2009 et 2014, plus de 700 projets d'une durée moyenne d'environ deux ans, ont été initiés, touchant plus de 500 000 jeunes.

Le FEJ a incité les porteurs de projet et évaluateurs à recourir à une évaluation d'impact rigoureuse. Son Guide méthodologique pour l'évaluation des expérimentations sociales précise ainsi que pour *« mesurer les effets d'un nouveau dispositif, il faudrait idéalement pouvoir observer la situation des bénéficiaires du dispositif et la comparer avec ce qui se serait passé pour ces mêmes bénéficiaires en l'absence du dispositif »*. Le FEJ a ainsi mis en avant les évaluations avec

contrefactuel, reposant sur une distribution aléatoire des bénéficiaires.

Depuis sa création, environ 23 évaluations ont été réalisées selon ces méthodes expérimentales et quasi-expérimentales – une vague de production inédite pour les politiques sociales en France. Cependant, ces quelques 23 projets représentent une minorité parmi les 700 expérimentations financées, une proportion faible au vu de l'ambition initiale. Jean Bérard, Alexia Prétari et Mathieu Valdenaire soulignent deux facteurs d'explication : « ces méthodes n'étaient pas adaptées à l'ensemble des projets et toutes les équipes ne sont pas à même de les réaliser »^[9]. Le Centre de recherche pour l'étude et l'observation des conditions de vie (Credoc) fournit une analyse plus approfondie des difficultés rencontrées dans la mise en œuvre des évaluations randomisées pour le FEJ^[10]. Ce bilan cite trois obstacles clés : le refus des porteurs de projet pour raison éthique, les effectifs parfois trop faibles et le risque de contagion entre les groupes test et témoin⁹.

La mission, les financements et l'animation du FEJ ont évolué depuis 2012. Premièrement, on constate en 2012, une interruption temporaire des subventions du « programme 163 » destinées à financer les appels à projets du FEJ. Deuxièmement, la mission d'animation du FEJ (MAFEJ) s'est vue attribuer l'animation du programme « La France S'engage », démarche lancée par le président de la République le 24 juin 2014. Enfin, depuis le 1^{er} janvier 2016, la MAFEJ est rattachée à l'Institut national de la jeunesse et de l'éducation populaire (INJEP), lui-même rattaché au Directeur de la jeunesse, de l'éducation populaire et de la vie associative (DJEPVA). L'INJEP fait de la production de travaux de synthèse et de capitalisation une priorité, avec l'objectif de mieux diffuser les enseignements du FEJ auprès des acteurs de la jeunesse. La récente publication *De l'éducation à l'insertion : dix résultats du Fonds d'expérimentation pour la jeunesse*, constitue un exemple du renforcement de cette démarche de diffusion.

9 Le risque que les groupes test et témoin s'influencent dans leur comportement par exemple.

L'expérience et les évolutions du FEJ sont riches d'enseignements sur les conditions de réussite d'une bonne intégration de l'évaluation d'impact à l'élaboration et à la mise en œuvre des politiques sociales en France.

Un mouvement qui se poursuit mais une relativement faible production d'évaluations d'impact

Aujourd'hui, si certaines initiatives se sont essouffées, le mouvement autour de l'évaluation d'impact et de l'expérimentation sociale se poursuit.

L'émergence de nouveaux acteurs a renouvelé le portage de ces approches. Ces acteurs sont notamment ceux de la recherche spécialisés dans l'évaluation d'impact, tels que la Fédération de recherche du CNRS « Travail, Emploi, Politiques publiques » (TEPP) et J-Pal Europe. L'évaluation d'impact est également une des méthodes utilisées par deux laboratoires créés au cours des dernières années : l'Institut des politiques publiques (IPP) à l'École d'économie de Paris et le Laboratoire interdisciplinaire de l'évaluation des politiques publiques (Liepp) à Sciences Po (voir supra).

Enfin, l'Ansa a émergé en tant qu'« acteur tiers » chargé de l'accompagnement et de la mise en œuvre d'expérimentations, jouant un rôle de médiation entre porteurs de projet, évaluateurs scientifiques et financeurs.

L'expérimentation du programme national Parler Bambin, portée par l'Ansa et financée par La France s'engage illustre bien la nature de ce rôle d'acteur tiers.

L'ANSA, PILOTE OPÉRATIONNEL DE L'EXPÉRIMENTATION NATIONALE PARLER BAMBIN

Le programme national Parler Bambin est une expérimentation visant à mettre en œuvre Parler Bambin à grande échelle et à évaluer ses effets, par une évaluation aléatoire contrôlée. Parler Bambin repose sur l'enrichissement des interactions langagières entre l'adulte et le jeune enfant (3 à 36 mois) grâce à des postures favorisant le développement du langage et le plaisir de communiquer. Le volet opérationnel, porté

par l'Ansa, consiste à coordonner le programme avec toutes les parties prenantes, apporter un nouveau cadre de diffusion de Parler Bambin, structurer une offre de formation et animer un réseau national de professionnels de 2015 à 2020.

L'Ansa s'appuie sur une alliance de villes pionnières (Courcouronnes, Le Havre, Lille) et des acteurs historiques grenoblois. Le programme de recherche-action est porté par une équipe de recherche interdisciplinaire : le J-Pal Ecole d'Économie de Paris (Marc Gurgand) et le laboratoire dynamique du langage (Sophie Kern). Cette recherche repose sur la comparaison entre un groupe test et un groupe témoin de crèches, composés par tirage au sort de leur date d'entrée dans le programme. Elle permettra, à terme, de mesurer les effets de Parler Bambin sur le développement global des enfants, l'implication des parents et les pratiques professionnelles. Grâce au soutien de La France s'engage, le programme comptabilise aujourd'hui 96 crèches sur le territoire national.

Certains de ces acteurs mettent en œuvre des expérimentations et des évaluations d'impact, et promeuvent également ces méthodes, à l'image du programme SPARK.

SOCIAL POLICY ANALYSIS FOR ROBUST KNOWLEDGE (SPARK)

SPARK est un programme de sensibilisation des décideurs publics à l'évaluation d'impact et à l'expérimentation sociale lancé en 2014. Financé par la Commission européenne, il a été porté par le London School of Economics en partenariat avec J-Pal Europe et l'Alliance for Useful Evidence. SPARK a notamment conduit à la production d'un guide portant sur l'expérimentation sociale à destination des décideurs publics^{xxvi}.

La volonté d'évaluer l'impact de dispositifs ou interventions s'est traduite par une **montée en compétences de l'État**, notamment dans les services statistiques ministériels. D'une part, des équipes *ad hoc* peuvent être constituées pour réaliser des évaluations d'impact. La Dares a ainsi créé

une équipe projet pour réaliser l'évaluation quantitative de la Garantie jeune. D'autre part, des équipes pérennes sont parfois créées pour réaliser des évaluations, telles que le Bureau de l'évaluation des actions éducatives et des expérimentations à la Direction de l'évaluation, de la prospective et de la performance (Depp) du ministère de l'Éducation nationale.

La popularisation de l'expérimentation sociale a favorisé l'émergence du « réflexe expérimental » au sein des pouvoirs publics et d'une partie de la société civile. Ainsi, quelques nouvelles expérimentations ont été lancées au cours des cinq dernières années, telles qu'Un chez soi-d'abord, la Garantie jeune, Territoires zéro chômage de longue durée et les Lits d'accueil médicalisés (LAMS). Ces « expérimentations » ont toutefois rarement bénéficié d'une évaluation avec assignation aléatoire au groupe test et témoin (Un chez soi-d'abord fait figure d'exception à cet égard). En effet, ces projets ont les plus souvent été soumis à une évaluation « avant/après » ou quasi-expérimentale ; des méthodes utiles pour certains types de dispositif mais moins à même d'établir un lien causal entre le dispositif et les effets observés.

EXPÉRIMENTER POUR LUTTER CONTRE LA PAUVRETÉ : UN CHEZ SOI-D'ABORD

Un chez-soi d'abord est une expérimentation menée sur quatre sites – Lille, Marseille, Paris, Toulouse – visant l'accès au logement de 400 personnes sans domicile fixe, souffrant de troubles psychiques. Ce projet s'inspire des expérimentations *Housing First* conduites en Amérique du Nord et dans d'autres pays européens, qui ont montré un taux de maintien dans le logement à 12 mois avoisinant les 90%. Quatre ans après leur entrée dans le logement, plus de 85% des personnes y sont toujours selon le dernier bilan de la recherche, qui compare les parcours de plus de 350 bénéficiaires à un groupe témoin. La Délégation interministérielle à l'hébergement et à l'accès au logement (Dihal), qui coordonne l'expérimentation au niveau national, a inscrit le dispositif dans le code de l'action sociale^{lxvii} et envisage son déploiement à 16 territoires supplémentaires à l'horizon 2022.

Enfin, au-delà de ces expérimentations à petite échelle, certains acteurs publics ont désormais recours aux **méthodes quasi-expérimentales** pour évaluer l'impact de certaines réformes, telle que l'évaluation du crédit d'impôt pour la compétitivité et l'emploi (CICE) sous l'égide de France Stratégie^{lxviii}. Cette pratique reste toutefois peu développée en France aujourd'hui.

L'ÉVALUATION DE RÉFORMES

L'exemple de l'évaluation d'impact du crédit d'impôt pour la compétitivité et l'emploi (CICE) sous l'égide de France Stratégie. Le comité de suivi du crédit d'impôt pour la compétitivité et l'emploi (CICE), dont les missions ont été étendues au suivi des aides publiques aux entreprises et aux engagements, est installé à France Stratégie. Il publie des rapports annuels. Le rapport 2016 du comité fournit des éléments d'appréciation de l'effet du CICE fondés sur une analyse *a posteriori* des comportements individuels des entreprises bénéficiaires au cours des années 2013 et 2014. Des évaluations du CICE ont été réalisées par trois équipes de recherche universitaires indépendantes qui ont travaillé en parallèle et sur la base de méthodologies distinctes. Les équipes ont tenté de neutraliser les effets de la conjoncture et ceux des autres mesures de politique économique, afin d'isoler les effets purs du CICE. Le rapport conclut que ce dispositif, dont le coût s'élève à 20 milliards d'euros environ, a eu les effets suivants^{lxix} :

- un effet probable sur la survie des entreprises, qui doit être plus précisément évalué ;
- au terme de la deuxième année, pas d'effet observable sur l'investissement, la recherche et le développement et les exportations ;
- sur l'emploi et les salaires, des écarts d'évaluation, une équipe (TEPP) conclut à la création ou la sauvegarde de 45 000 à 115 000 emplois, alors que l'autre (Liepp) ne met pas en évidence d'effet positif sur l'emploi ni sur les salaires par tête.

Les évaluations d'impact comme outil d'aide à la décision

Les expérimentations et évaluations d'impact éclairent parfois la décision publique

Suite à certaines expérimentations menées en France, des décideurs politiques se sont emparés de leurs enseignements, montrant qu'il est possible d'intégrer des éléments probants à la décision politique. Ainsi les limites du Revenu Contractualisé d'Autonomie (RCA), soulevées dans son évaluation, ont été prises en compte lors de la modélisation de la Garantie Jeune. Une autre expérimentation a comparé l'impact sur les demandeurs d'emploi de l'accompagnement des opérateurs privés avec le programme Cap Vers l'Entreprise. Cette expérimentation a informé les débats publics sur la plus-value ou non du recours à des opérateurs privés au titre du service public de l'emploi.

QUAND L'EXPÉRIMENTATION ÉCLAIRE LA DÉCISION : DU RCA À LA GARANTIE JEUNE

Les jeunes peu qualifiés connaissent des difficultés croissantes d'insertion sur le marché du travail, en dépit des nombreux dispositifs qui leur sont destinés. Après le manque de résultats du CIVIS (contrat d'insertion dans la vie sociale) en 2005, le ministère de la Jeunesse a expérimenté le « Revenu Contractualisé d'Autonomie » (ou RCA) de 2011 à 2013. L'innovation du RCA réside dans le cumul de l'allocation et de ressources d'activité éventuelles, dans la continuité du contrat sur une durée déterminée, quelle que soit la situation professionnelle du jeune¹⁰⁰⁰.

Une expérimentation contrôlée a été menée, comparant les parcours de 2 500 jeunes bénéficiaires du RCA et ceux de 2 500 bénéficiaires de CIVIS. Les jeunes bénéficiaires du RCA et ceux en CIVIS connaissaient une situation comparable d'emploi 1 an et 2 ans après l'entrée dans le dispositif. Deux explications ont été

émises pour expliquer ces résultats négatifs pour le RCA : d'une part l'allocation était trop faible pour avoir un impact sur leurs parcours (250 euros maximum); d'autre part, l'accompagnement était inadapté en contenu et en intensité.

Les enseignements du RCA ont été pris en compte dans la conception de la Garantie jeunes. Cette dernière propose, sur un an, une allocation plus importante – 450 euros mensuels – avec un accompagnement structuré visant l'accès à l'emploi ou à la formation et intégrant des immersions régulières en entreprise^{boxi}. Ces évolutions ont porté leurs fruits au vu des premiers résultats d'évaluation pour une première cohorte – 14 mois après leur entrée, 40,4% des bénéficiaires sont en emploi, au lieu de 34,1% pour le contrefactuel^{boxii}.

Les acteurs français ont parfois été inspirés par une expérimentation conduite dans d'autres pays de l'OCDE. Des dispositifs ayant fait leur preuves à l'étranger ont été mis en œuvre en France sous forme de projet pilote ou d'expérimentation, pour évaluer la capacité à transposer l'intervention dans le contexte français.

LES EXPÉRIMENTATIONS DE DISPOSITIFS ÉTRANGERS : ZOOM SUR 3 INTERVENTIONS IMPORTÉES EN FRANCE

Lutte contre le tabagisme

La campagne « Moi(s) Sans Tabac », mise en œuvre pour la première fois en novembre 2016 et portée par Santé publique France, s'inspire de la campagne britannique *Stoptober*. Des études ont montré qu'au-delà de 28 jours d'abstinence, les chances de rester non-fumeur de manière permanente sont multipliées par 5. *Stoptober* a fait la preuve de son efficacité avec une augmentation nette du nombre de tentatives d'arrêt associées à l'opération (+ 50% par rapport aux autres mois de l'année).

Réduire des inégalités dès la petite enfance

« Jeux d'enfants » est un programme d'intervention précoce destiné aux enfants de 0 à 5 ans et visant à favoriser les chances de réussite des enfants dès le plus jeune âge. Ce projet est directement issu du projet américain Carolina Abecedarian, qui a fait l'objet d'une évaluation contrôlée démontrant son efficacité à soutenir le développement cognitif, psychoaffectif et social du jeune enfant sur 30 ans. Le dispositif fait l'objet d'un projet pilote conduit par la ville de Lille, avec l'appui de l'Ansa.

Accès à l'emploi des personnes souffrant de troubles psychiques

Working First 13 est un programme innovant basé sur la méthodologie *Individual Placement Support*, l'équivalent anglo-saxon de l'emploi accompagné. Il s'adresse aux personnes éloignées de l'emploi sur le territoire de Marseille^{boxiii}.

Les revues systématiques sont peu développées dans les politiques sociales en France

La revue systématique de littérature est une méthode qui permet de comparer des impacts des interventions relatives à un domaine spécifique ayant fait l'objet d'une évaluation d'impact (par exemple l'accompagnement des demandeurs d'emploi). Elle vise ainsi à distinguer celles qui fonctionnent ou sont prometteuses, de celles ayant des impacts nuls ou négatifs. En France, la pratique des revues systématiques est déjà développée dans le domaine de la santé, notamment au sein de la Haute autorité de santé (HAS) et de l'Institut national de la santé et de la recherche médicale (Inserm). Dans le champ social, cette méthode est également intégrée dans les conférences de consensus (voir page 38), qui reposent sur une synthèse des connaissances existantes. Dans le domaine de la justice, par exemple, la conférence de consensus sur la prévention de la récidive s'est appuyée sur une revue systématique. Dans le champ éducatif, le Conseil national d'évaluation du système

What Works Centres britanniques : quels enseignements pour des politiques fondées sur la preuve en France ?

scolaire (Cnesco) a également organisé quatre conférences de consensus depuis sa création, dont une sur le redoublement. Ces exemples et les résultats qu'ils ont mis en avant démontrent l'avantage de telles démarches pour la qualité des décisions politiques.

LA REVUE SYSTÉMATIQUE : LE REDOUBLEMENT DES ÉLÈVES

Début 2015, le Cnesco et l'Institut français de l'éducation (Ifé) ont organisé une conférence de consensus intitulée « Lutter contre les difficultés scolaires : le redoublement et ses alternatives ». Alors qu'un certain nombre de pays se sont, depuis les années 1970, ralliés à la promotion automatique des élèves, la pratique du redoublement est encore courante en France. Une revue systématique des études a souligné que le redoublement, au mieux n'a pas d'effet, mais peut surtout s'avérer nocif tant pour la réussite scolaire des élèves que pour le développement de leur estime de soi^{100xiv}. S'appuyant entre autres sur les enseignements de cette revue systématique, le jury a ainsi conclu qu'il fallait « faire baisser le redoublement et mettre en place d'autres solutions plus efficaces »^{100xv}.

Une action publique globalement peu informée par les évaluations d'impact

Des décideurs politiques peu éclairés par les évaluations d'impact

La décision publique est alimentée par différentes sources d'analyses lors de l'élaboration de nouvelles mesures ou lois. Il est important de distinguer deux manières très différentes dont « la preuve » s'intègre ainsi dans la décision publique : la volonté des décideurs et l'obligation légale.

D'une part, pour l'élaboration de nouvelles mesures ou lois, les décideurs peuvent saisir en vue d'un éclairage (rapport) les conseils, hauts conseils, observatoires, voire laboratoires de recherche. Ces derniers mettent par

ailleurs à disposition les travaux qu'ils conduisent au titre de leur programme de travail (généralement fixé annuellement). Ces organismes fondent généralement leurs préconisations sur des analyses approfondies des besoins des personnes et des réponses apportées par les pouvoirs publics.

Cependant, ces acteurs ne s'appuient pas de façon régulière sur des revues systématiques d'évaluations d'impact. Enfin, même lorsque des évaluations *ex ante* sont réalisées avant la mise en place d'une mesure, leurs résultats ne sont pas systématiquement confrontés à une évaluation *ex post* permettant d'établir de manière plus robuste l'impact réel.

D'autre part, sauf exceptions prévues par la loi organique n°2009-403 du 15 avril 2009, tous les projets de lois sont accompagnés d'une *étude* d'impact réalisée par le Gouvernement. Ces études, bien différentes des *évaluations* d'impact, définissent les objectifs poursuivis, exposent les motifs du recours à une nouvelle législation, l'état actuel du droit dans le domaine visé, l'articulation du projet avec le droit européen, l'évaluation des conséquences économiques, financières, sociales et environnementales des dispositions du projet et les modalités d'application envisagées ainsi que leurs conséquences^{boxvi}. Ces études d'impact sont généralement réalisées dans des délais très contraints par les administrations centrales, sans l'appui de chercheurs spécialisés de l'évaluation et sans prise en compte des résultats des évaluations d'impact disponibles.

À cet égard, la Commission d'étude des effets de la loi pour la croissance et l'activité au sein de France Stratégie a constitué une démarche originale, visant à rendre de compte de l'état de l'art en matière de preuves et à informer le débat public et législatif.

**UNE ÉVALUATION EX ANTE REPOSANT SUR DES REVUES D'ÉVALUATIONS D'IMPACT
LA COMMISSION D'ÉTUDE DES EFFETS DE LA LOI POUR LA CROISSANCE ET L'ACTIVITÉ
HÉBERGÉE À FRANCE STRATÉGIE**

**L'ambition de la loi pour la croissance et l'activité (dite « loi Macron »)
était de créer les conditions d'un rebond de l'économie et de l'emploi**

en France. Une Commission d'évaluation indépendante a été créée pour évaluer *ex ante* l'impact que pourraient avoir les différents volets de cette loi^[xxxvii]. Cette commission était composée de chercheurs français et étrangers, experts sur les domaines visés par la loi. Elle a analysé les effets sur l'économie et l'emploi des lois et dispositions analogues mises en œuvre dans des pays étrangers^[xxxviii]. À cette fin, elle a réalisé des revues d'évaluations d'impact sur les cinq thèmes du projet de loi : le travail du dimanche, la mobilité, la réforme des prud'hommes, les professions réglementées du droit et l'immobilier commercial. La commission a publié, en janvier 2015, cinq notes de synthèse restituant les enseignements de leurs revues ayant vocation à alimenter le débat public.

Les pratiques « efficaces » sont difficilement accessibles aux professionnels

L'accès à l'état de l'art de la recherche est un enjeu clé des politiques basées sur la preuve. En France, des initiatives visent à favoriser l'essaimage de « bonnes pratiques » auprès des professionnels par des guides ou des bibliothèques, tels que la banque d'expérience de l'Union nationale des centres communaux d'action sociale (Unccas)^[xxxix]. Ces pratiques sont choisies lorsqu'elles semblent apporter des réponses pertinentes soit à une évolution législative, soit à un besoin des usagers, soit à une demande des professionnels. Cependant, ces pratiques ont rarement fait l'objet d'une évaluation permettant de mesurer leur impact. Plus globalement, il existe très peu de productions visant à diffuser les enseignements de ce qui « fonctionne » auprès des professionnels du champ social. Ces derniers manquent donc d'outils pour leur permettre de baser leurs pratiques « sur la preuve ».

Un accompagnement au changement d'échelle peu développé

En France, l'idée que l'on puisse diffuser des innovations testées à petite échelle est loin d'être nouvelle. Les services publics en France ont d'ailleurs

souvent été inspirés par des innovations locales. Le système des centres de Protection maternelle et infantile (PMI) créés en 1945, par exemple, fut largement inspiré par l'Association alsacienne et lorraine de puériculture, créée en 1920 par le pédiatre Paul Rohmer.

Cependant, l'idée qu'il convient d'accompagner le changement d'échelle des innovations est plus récente. Cette idée est au cœur du modèle français de l'expérimentation sociale. En effet, l'expérimentation sociale est généralement définie comme « une méthode qui permet l'évaluation d'innovations sociales déployées à petite échelle avant d'en envisager leur généralisation »^{xc}. Certains des dispositifs expérimentés, tels que la Mallette des parents ont d'ailleurs eu des effets probants et ont fait l'objet d'un essaimage à plus grande échelle.

DE L'EXPÉRIMENTATION À L'ESSAIMAGE : LA MALLETTTE DES PARENTS

La Mallette des parents vise à instaurer « des échanges fructueux entre l'école et les familles, pour mieux impliquer les parents dans la réussite de leurs enfants »^{xcii}. Elle contient des outils que les équipes éducatives utilisent pour animer la discussion avec les familles lors d'ateliers-débats. Une expérimentation qui a montré des effets très positifs sur certains critères : « des élèves plus motivés, une réduction de la violence scolaire et de l'absentéisme ». Ensuite, le dispositif a fait l'objet d'un essaimage structuré et, depuis 2010, a été étendu progressivement pour atteindre aujourd'hui environ un quart des collèges français.

Au cours des dernières années, la pratique d'accompagnement au changement d'échelle s'est développée en France. Plusieurs programmes ont émergé, tels que l'Accélérateur d'innovation sociale de l'Ansa, l'accompagnement au changement d'échelle proposé par l'Avise et le programme Antropia lancé par l'Essec. Enfin, depuis 2014 le label La France s'engage finance et accompagne des innovations sociales à changer d'échelle.

SOUTIEN AU CHANGEMENT D'ÉCHELLE : LA FRANCE S'ENGAGE

La France s'engage est une démarche portée par le Président de la République, depuis 2014. Elle a vocation à mettre en valeur et faciliter l'extension d'initiatives socialement innovantes, d'intérêt général, portées par des individus, des associations, des fondations, des entreprises. L'enjeu est de faire émerger de nouveaux modes d'intervention répondant aux besoins de la société française, par le biais d'un appel à projets permanent. Depuis 2014, 62 projets ont ainsi été labellisés.

Les enseignements clés sur l'évaluation et la preuve en France

La culture de l'évaluation en France est très marquée par l'évaluation des politiques publiques. **Malgré l'engouement indéniable pour cette approche, elle n'est pas encore stabilisée, tant du point de vue de ses outils et méthode que du point de vue institutionnel. La sphère de l'évaluation fait coexister des approches méthodologiques et stratégiques sinon concurrentes du moins divergentes, que les commentateurs ont pu qualifier de « chapelles » de l'évaluation^{xcii}. Des travaux d'évaluation restent également sous-appropriés par les décideurs et les praticiens et méconnus d'un plus large public.**

À partir de 2007, le recours à l'évaluation d'impact a fortement augmenté, en raison notamment du mouvement autour de l'expérimentation sociale. **Il s'est traduit par une production inédite d'évaluations d'impact d'innovations, notamment autour des questions d'insertion et de la jeunesse. Depuis 2012, cette dynamique s'est poursuivie de manière plus modeste au travers des expérimentations et des évaluations d'impact.**

Aujourd'hui, malgré l'intérêt affiché pour ces méthodes rigoureuses, peu d'évaluations d'impact ont été réalisées dans le champ social

en France **au cours des dernières années. Cette méthode est souvent vue comme complexe, lente et coûteuse, à la fois par les pouvoirs publics et les professionnels de terrain. Les évaluateurs ont rencontré des obstacles à sa mise en œuvre comme en témoigne l'expérience du Fej. Ceci a conduit certains acteurs à interroger sa faisabilité technique et donc à privilégier le recours à d'autres méthodes d'évaluations. Les difficultés rencontrées témoignent de la complexité inhérente à ces méthodes, mais également d'un manque d'expertise et d'expérience sur ces méthodes, de la part des évaluateurs comme des porteurs de projet.**

Le recours aux autres méthodes associées aux politiques fondées sur la preuve – la synthèse et la diffusion des évaluations impacts – reste extrêmement limité **dans les domaines sociaux, économiques et éducatifs en France. Les laboratoires et équipes de recherche que nous avons interrogés n'en réalisent en général pas. Cette méthode représente donc une pratique quasi inexistante, à développer.**

Les interventions à impact positifs font rarement l'objet d'un changement d'échelle structuré. **À titre d'exemple, seules quelques-unes des dizaines d'expérimentations sociales financées par le Fej ont fait l'objet d'un véritable essaimage. En effet, comme souligne un récent rapport du SGMAP et du CGET « la sphère publique n'est pas organisée aujourd'hui de manière à faciliter le changement d'échelle de ces initiatives»^{lxviii}. Enfin, ces programmes de changement d'échelle peinent généralement à identifier des innovations ayant fait l'objet d'une évaluation d'impact. Ils tendent donc à essaimer des innovations « pertinentes » ou « prometteuses » mais rarement celles qui se sont révélées « efficaces ».**

lxviii LACQUETTE-FOUGERE Clément, LASCOURMES Pierre, « L'évaluation : un marronnier de l'action gouvernementale », *Revue française d'administration publique*, n°148, 2013, p. 859-875.

lxix SPENLEHAUER Vincent, « La (f)utilité gouvernementale de l'évaluation des politiques publiques, quelques leçons américaines et françaises ».

lxx JACOB Steeve, VARONE Frédéric et GENARD, Jean-Louis. (Dir.) (2007), *L'évaluation des politiques au niveau régional*, Bruxelles, Peter Lang, Coll. Action publique.

lxxi Pour plus d'information sur les EPP :

What Works Centres britanniques : quels enseignements pour des politiques fondées sur la preuve en France ?

- www.modernisation.gouv.fr/l'action-publique-se-transforme/en-evalu-ant-ses-politiques-publiques.
- lxxii** www.modernisation.gouv.fr/l'action-publique-se-transforme/en-evalu-ant-ses-politiques-publiques.
- lxxiii** LACQUETTE-FOUGERE Clément, LASCOURMES Pierre, « Les scènes multiples de l'évaluation », les problèmes récurrents de son institutionnalisation, LIEPP, Policy Paper n°1, mai 2013.
- lxxiv** La Vigie de l'Évaluation, février 2014 : www.evaluationvigie.fr/les-quatre-chapelles-de-levaluation.
- lxxv** On pourra renvoyer sur ce point aux résultats du baromètre de l'évaluation des politiques publiques réalisé tous les deux ans par la SFE avec le soutien du SGMAP et qui produit une analyse quantitative et qualitative des références d'évaluations recensées depuis 2007 dans l'Observatoire de l'évaluation des politiques publiques (OEPP). En 2016, près de 1 800 références étaient ainsi répertoriées dans l'OEPP. Voir en ligne : www.sfe-asso.fr/ressources-sfe/barometre-evaluation.
- [i]** BERARD Jean, PRETARI Alexia, et VALDENAIRES Mathieu, « Introduction générale, Expérimenter : oui, mais comment ? » dans *De l'éducation à l'insertion : dix résultats du Fonds d'expérimentation pour la jeunesse*, Paris, La documentation française, 2014, pages 13 à 30.
- [ii]** ALDEGHI Isa, JAUNEAU Pauline, MULLER Lara, Bilan méthodologique des évaluations conduites par le CRÉDOC dans le cadre du Fonds d'expérimentation pour la Jeunesse, CRÉDOC, décembre 2012.
- lxxvi** European Commission, *Guide to Testing Social Policy Innovation*, June 2014.
- lxxvii** Décret n°20161940 du 28 décembre 2016 relatif aux dispositifs d'appariements de coordination thérapeutique « Un chez soi-d'abord ».
- lxxviii** COSNEFROY, Olivier, Évaluation des rythmes scolaires, présentation, juin 2016 : www.ortej.org/IMG/pdf/presentation_ortej_15-06-16.pdf.
- lxxix** BRUN-SCHAMME Amandine, DESPLATZ Rozenn, NABOULET Antoine, *Rapport 2016 Comité de suivi du Crédit d'impôt pour la compétitivité et l'emploi*, France Stratégie, septembre 2016.
- lxxx** Décret n°2011-128 du 31 janvier 2011 relatif à l'expérimentation d'un revenu contractualisé d'autonomie.
- lxxxi** www.gouvernement.fr/action/la-garantie-jeunes.
- lxxxii** Comité scientifique en charge de l'évaluation de la Garantie Jeunes, *Rapport intermédiaire*, novembre 2016 : dares.travail-emploi.gouv.fr/IMG/pdf/rapport_garantie_jeunes.pdf.
- lxxxiii** www.workingfirst13.fr.
- lxxxiv** Cnesco et Institut français de l'éducation, conférence de consensus « Lutter contre les difficultés scolaires : le redoublement et ses alternatives ? Partie 2 : Le redoublement en France et dans le monde : de l'étude de ses impacts à la croyance en son utilité », décembre 2014.
- lxxxv** Cnesco et Institut français de l'éducation/ENS Lyon, conférence de consensus « Lutter contre les difficultés scolaires : le redoublement et ses alternatives ? Recommandations du jury », janvier 2015.
- lxxxvi** Circulaire du Premier ministre du 15 avril 2009.
- lxxxvii** www.strategie.gouv.fr/sites/strategie.gouv.fr/files/atoms/files/lettre_mission_macron_jpf.pdf.
- lxxxviii** Les notes de synthèse se trouvent ici : www.strategie.gouv.fr/presse/communiqués-de-presse/installation-de-commission-devaluation-de-loi-croissance-lactivite.
- lxxxix** www.unccas.org/-banque-d-experience-.
- xc** LE BARS Sylvie, PRIGENT Marion, RICKY Ben et AITOUT Claire-Selma, L'expérimentation sociale à l'épreuve du terrain, mars 2014.
- xc i** www.education.gouv.fr/cid53083/le-dispositif-mallette-des-parents.html.
- xcii** La Vigie de l'Évaluation, février 2014 : www.evaluationvigie.fr/les-quatre-chapelles-de-levaluation.
- xciii** Secrétariat général pour la modernisation de l'action publique et le Commissariat général à l'égalité des territoires, *Ensemble, accélérons !*, décembre 2015. www.modernisation.gouv.fr/les-services-publics-se-simplifient-et-innovent/par-la-co-construction/changement-d-echelle-des-innovations-ensemble-accelérons.

Les What Works Centres : quels enseignements pour la France ?

Dans ce dernier chapitre, sont restitués les enseignements qui peuvent être tirés des What Works Centres. À partir d'un aperçu des pratiques de ces centres, dans quelle mesure peuvent-ils constituer une source d'inspiration dans le cas français ? L'étude met en évidence des avantages et certaines limites à la démarche incarnée par les What Works Centres.

Méthodes des What Works Centres

Intérêts

Les What Works Centres ont recours à des méthodes classiques au sein des démarches de politiques fondées sur la preuve : revue systématique de littérature, transfert de connaissance par boîte à outils, expérimentation sociale avec évaluation d'impact.

La singularité de leur démarche tient à l'ambition d'appliquer ces méthodes sur un domaine entier de la politique publique. Ces centres ont ainsi mis ces méthodes à l'épreuve du terrain, testant leurs limites et innovant pour les adapter aux politiques sociales dans le contexte britannique. L'Ansa considère que les pratiques développées par les WWC sont intéressantes et constituent potentiellement des sources d'inspiration dans le cas français s'agissant :

- **des programmes conséquents de revues systématiques de littérature.**

En réalisant des programmes ambitieux de revues portant sur l'efficacité d'un panel d'interventions, les centres ont cherché à apporter des enseignements utiles aux décideurs et professionnels en répondant aux questions que ceux-ci posent. Les difficultés rencontrées dans la réalisation de ces revues ont conduit certains centres à nuancer leur approche, par exemple pour mieux prendre en compte la question de la transposabilité des interventions identifiées au contexte britannique, ou en faisant évoluer

les critères de sélection des évaluations à prendre en compte ;

■ **des innovations dans les méthodes de diffusion et d'essaimage.**

Les WWC ont transformé des centaines d'évaluations en enseignements intelligibles par des professionnels et décideurs publics, notamment au sein de bibliothèques d'intervention. L'innovation réside ici dans la combinaison des revues systématiques de haut niveau scientifique, avec la présentation des résultats de manière synthétique, faciles à appréhender et à s'en saisir, donnant aux centres un rôle de référence et une visibilité dans leur domaine. Les WWC sont toutefois conscients qu'il ne suffit pas de mettre cette matière à disposition pour que les acteurs s'en saisissent. Les centres ont donc multiplié des actions de diffusion par le biais d'événements, de formations et de programmes d'essaimage ;

■ **le développement d'une nouvelle capacité en conduite d'expérimentations.**

Un autre intérêt de leur démarche tient au fait qu'ils testent de nouvelles approches, y compris dans le cadre d'expérimentations aléatoires. Pour l'*Education Endowment Foundation* (EEF) et *WWC for Crime Reduction*, la capacité à soutenir des expérimentations a été un véritable atout, les rapprochant des acteurs de terrain. Ce premier a d'ailleurs financé la réalisation d'une vague d'expérimentations éducatives inédite par son ampleur en Angleterre : 127 évaluations dont 105 selon des méthodes expérimentales randomisées entre 2011 et 2016. Ce faisant, ils ont développé une expertise dans la mise en œuvre d'expérimentations aléatoires – notamment pour remédier aux biais d'expérimentation qui, on l'a vu, peuvent faire échouer celles-ci – et ont contribué au développement d'un plus large panel de chercheurs compétents pour conduire de telles expérimentations – condition de réussite du développement d'expérimentations aléatoires réussies. La capacité à soumettre une centaine d'innovations à une évaluation d'impact rigoureuse a été un véritable atout de l'*Education Endowment Foundation*, pour enrichir la connaissance mais aussi rapprocher le centre des acteurs de l'éducation. L'expertise ainsi développée par EEF représente, en soi, un intérêt de la démarche pour les acteurs français cherchant à conduire des expérimentations. La dynamique lancée montre le potentiel de développement d'une telle activité.

Limites

Les What Works Centres ont transposé aux champs sociaux des approches basées sur la preuve développées initialement dans le domaine médical.

Ceci a soulevé des questions et des difficultés rappelées ici :

- **la production des revues a été parfois difficile en raison du manque d'évaluations d'impact de niveau suffisant, de la complexité à déterminer l'efficacité d'une intervention et à clarifier les contextes dans lesquels une intervention fonctionne.** Certains centres ont nuancé leur approche, en incluant des interventions « prometteuses » ou des pistes intéressantes. D'autres centres auront sans doute à faire des choix plus net, en particulier dans les champs qui se prêtent le moins aux évaluations randomisées (tels que le développement économique local). De plus, certains acteurs continuent de critiquer l'idée que l'évaluation d'impact devrait être le type de « preuve » à porter en priorité à la connaissance des décideurs et professionnels. La plupart des centres tentent de stimuler la production de ces évaluations par des actions de formation, de sensibilisation et d'accompagnement, avec un bilan mitigé ;
- **les centres n'ont pas encore démontré un impact mesurable sur les pratiques des professionnels.** Le manque de recul sur l'impact de ces jeunes organisations constitue une explication. Il reste toutefois des incertitudes relatives à l'efficacité de la démarche qui consiste à identifier « ce qui fonctionne » en général dans les pays de l'OCDE, puis de les diffuser au Royaume-Uni. Se pose en particulier des questions de la transposabilité des pratiques démontrées comme efficaces (« validité externe »), et des conséquences de leur changement d'échelle (« effets d'équilibre »). Un recours important à l'expérimentation au Royaume-Uni est la manière privilégiée de s'assurer de la transposabilité d'une intervention.

Les What Works Centres en tant que dispositif

Intérêts

L'autre intérêt des What Works Centres tient à leur existence en tant que dispositif qui est :

- **inscrit dans un temps long**, celui de la transformation des pratiques professionnelles. Certains centres ont des dotations de longue durée qui sont cohérents avec cette mission, mais la plupart bénéficient de financements de trois ans seulement, source de fragilité ;
- **indépendant**, puisque non intégré au sein des administrations centrales, permettant d'atténuer l'impact des changements d'orientations ministérielles sur leur action. La plupart sont toutefois proches des pouvoirs publics, puisqu'ils dépendent pour grande partie de financements publics. Deux sont même des établissements publics ;
- **sectoriel**, traitant un champ qui n'est pas contraint par les silos qui caractérisent l'organisation de l'action publique. Cette approche semble particulièrement efficace lorsque le secteur correspond à un corps de professionnel bien identifié (par exemple les enseignants pour l'EEF).

Doter des organisations spécifiques des missions « *What Works* » présente plusieurs avantages :

- **un « guichet unique » vers les évaluations d'impact pour les professionnels.** En tant que dispositif, les centres ont à la fois le mandat, les moyens et l'indépendance pour devenir le guichet unique vers « ce qui fonctionne » pour les professionnels de terrain. Dans un monde riche en information, ceci répond à un enjeu clé pour les professionnels d'accès rapide à des conseils à la fois clairs et fondés sur la preuve. En s'inscrivant dans la durée, les centres peuvent devenir la référence dans leur domaine (par exemple NICE, EEF). De surcroît, en mettant à jour leurs bibliothèques, les centres visent à apporter un état de l'art toujours actualisé ;
- **un « fil rouge » dans l'accompagnement des changements de pratiques dans la durée.** Les centres visent à influencer sur les pratiques de milliers de professionnels et sur les décideurs. Ceci implique l'accompagnement des acteurs à s'emparer des pratiques prometteuses ou efficaces,

ce qui requiert souvent des démarches intensives d'accompagnement à l'essaiage. De plus, pour que ces professionnels et décideurs s'emparent de cette matière, ils doivent eux-mêmes être convaincus de l'intérêt de fonder leur action sur la preuve, un véritable changement culturel. La création de centres représente une manière d'assurer l'accompagnement de ces changements dans la durée ;

■ **un « maillon manquant » entre décideurs, chercheurs et professionnels.**

Les centres visent à relier, au sein d'un même « écosystème de la preuve », les chercheurs, décideurs et professionnels. Ils ont pu ainsi stimuler un nouveau dialogue entre ces acteurs sur la place des évaluations d'impact dans l'action publique. Ce dialogue se trouve renforcé lorsque les centres disposent des moyens pour expérimenter des innovations proposées par les acteurs de terrain ou lorsqu'ils mettent en place des instances participatives comme NICE avec les praticiens, professionnels de terrains, chercheurs et personnes concernées.

Limites

Le modèle « What Works » s'est inspiré de la médecine basée sur la preuve et notamment du NICE, qui repose sur la diffusion de « la preuve » dans un domaine bien identifié (santé) à un corps professionnel de la santé identifié (médecins, infirmiers, etc.). Les centres ont reproduit ce modèle dans des domaines où ces conditions ne sont pas souvent réunies :

■ **certains centres se sont construits dans des domaines qui n'étaient pas bien identifiés ou ayant un périmètre flou**, tels que la prévention précoce, le bien-être et le bien- vieillir. Ces centres ont ainsi été obligés non seulement de préciser le périmètre du domaine en concertation avec les acteurs, mais aussi de convaincre les acteurs de la pertinence de ce domaine ;

■ **les centres ne disposent pas, dans leur majorité, d'un corps professionnel bien délimité.** Les centres sur la criminalité (policiers) et l'éducation (enseignants, aides enseignants) font figures d'exception, bien que les cibles potentielles de leur diffusion dépassent largement un type de professionnels. Les autres centres ont dû constituer de manière empirique un réseau de professionnels pouvant s'intéresser à leurs travaux. Ce travail est facilité lorsqu'il existe un socle de professionnels particulièrement concernés par

What Works Centres britanniques : quels enseignements pour des politiques fondées sur la preuve en France ?

le sujet, tels que des agents de développement économique dans le cas de *WWC Local Economic Growth*. Cette tâche devient infiniment plus difficile lorsqu'il s'agit d'un domaine à périmètre flou tels que le bien-être et, dans une moindre mesure, le bien-vieillir.

Ces centres ont tenté de surmonter ces problématiques de périmètre et de cible de manière pragmatique. Cependant, la nécessité de bâtir ainsi leur champ témoigne de la fragilité de leur démarche. En effet, il est difficile de créer un domaine et un groupe de professionnels *ex nihilo*, même si s'emparer de problématiques transversales peut être un objectif pour aller plus loin dans l'interdisciplinarité et y apporter une meilleure réponse. Ceci représente un point de vigilance si l'approche « *What Works* » devait être développée sur certains champs en France.

Le contexte français se prête-t-il au modèle *What Works* ?

L'analyse du contexte français au regard de l'approche « *What Works* » a confirmé l'hypothèse initiale d'un faible recours aux pratiques associées aux « politiques fondées sur la preuve » : évaluations d'impact, revues systématiques, diffusion des enseignements auprès des professionnels et des décideurs. En ce sens, le contexte français (notamment avant le lancement des WWC) affiche de nombreux points communs avec le contexte britannique.

Sur un plan culturel, l'idée qu'il est souhaitable et possible de fonder des politiques sur la preuve semble toutefois plus contestée en France, notamment dans les champs sociaux et éducatifs. Les promoteurs de cette démarche (notamment les nombreux acteurs impliqués dans des expérimentations sociales ou éducatives avec évaluation d'impact) restent minoritaires dans la communauté scientifique comme parmi les décideurs et professionnels de terrain.

Sur le plan de l'organisation de l'action publique, il existe des divergences entre les contextes français et britanniques. À plusieurs reprises, les acteurs

français interrogés ont cité la plus grande décentralisation du Royaume-Uni et une plus forte autonomie des acteurs au niveau local (par exemple, dans le système éducatif) comme des facteurs clés de réussite des What Works Centres. Une crainte est que le degré de centralisation de la France ne rende de tels centres peu ou pas pertinents. Sans prétendre à une analyse comparée de l'organisation de l'action publique dans les deux pays, il est important d'aborder ce sujet. L'analyse qui consiste à contraster une France « centralisée » à un Royaume-Uni « décentralisé » repose sur une perception simplifiée de la situation. Cette analyse ne prend en compte ni la décentralisation de nombreuses politiques en France depuis les années 1980, ni le fort degré de centralisation de certaines politiques au Royaume-Uni (le système de santé par exemple).

Les différences dans les structures de l'action publique des deux pays ne peuvent pas représenter, à elles seules, un obstacle au transfert de l'approche « What Works » en France. Il est toutefois évident que la structure de l'action publique d'un domaine a des conséquences sur le positionnement et les modalités d'un éventuel What Works Centre. L'étude tend à montrer que :

- **lorsqu'un WWC agit sur un champ où l'action publique est relativement décentralisée ceci peut représenter une justification importante pour leur action.** En effet, lorsque les politiques sont décentralisées, il est d'autant plus important d'éclairer les décisions des élus, agents et professionnels locaux. En effet, lorsque ces acteurs disposent d'une plus large marge de manœuvre se pose la question de leur capacité à prendre des décisions éclairées par la preuve, alors qu'ils disposent généralement de peu de temps et de ressources. Par exemple, en Angleterre, les directeurs d'établissement scolaire bénéficient depuis peu d'un budget dévolu aux actions pédagogiques pour les élèves en difficulté (« *pupil premium* »). Dans ce contexte, le *Toolkit* de l'*Education Endowment Foundation* apporte des éléments cruciaux sur les actions les plus efficaces pour ces élèves ;
- **cependant, les WWC peuvent également agir dans des domaines plus « centralisés ».** NICE, WWC originel, agit sur le domaine de la santé avec un service de public de santé dont le pilotage est extrêmement centralisé. Loin d'être un obstacle à la réussite de NICE, cette centralisation a permis

What Works Centres britanniques : **quels enseignements pour des politiques fondées sur la preuve en France ?**

d'intégrer les recommandations de NICE de manière plus systématique dans le système de santé à travers, par exemple, des processus d'audit. Plus généralement, même lorsqu'une politique est fortement cadrée au niveau national, des agents et professionnels du terrain disposent presque toujours des marges de manœuvre.

Propositions

En s'inspirant de l'expérience des What Works Centres et s'appuyant sur l'enquête menée auprès des acteurs français, l'Ansa et ses partenaires avancent deux propositions. Elles visent à soutenir l'émergence de politiques et de pratiques fondées sur la preuve en France, en renforçant le lien entre l'évaluation d'impact et l'action publique.

Proposition 1 : Étudier l'opportunité de reconnaître ou de créer des démarches ou des structures d'inspiration « What Works » en France

À l'instar des What Works Centres britanniques, ces structures permettraient de renforcer la prise en compte de la « preuve » dans la décision et l'élaboration des politiques publiques, domaine par domaine. De telles démarches ou structures constitueraient l'accès privilégié aux enseignements des évaluations d'impact pour les décideurs et les professionnels de terrain.

La présente étude conduit à proposer des missions spécifiques que ce type de structures pourrait porter en France, et quelques conditions de réussite. Trois modèles organisationnels sont proposés pour inscrire ces « centres » dans le paysage institutionnel français.

Missions socles

► **produire des revues systématiques d'évaluations d'impact.** Les revues systématiques de littérature doivent être menées avec rigueur, selon des méthodes et des critères de classification reconnus et transparents, et répondant à un niveau d'exigence élevé (voir par exemple la méthode PRISMA¹⁰). Elles contribuent à la fois à l'avancement de la recherche et à rendre cette recherche accessible en publiant systématiquement leurs travaux. Chaque « centre » a ainsi vocation à devenir un point d'accès privilégié

10 PRISMA for preferred reporting items for systematic reviews and meta-analysis (www.prisma-statement.org).

vers les évaluations d'impact produites dans son domaine. Cette mission représente un travail conséquent, reposant sur des moyens humains et financiers à ne pas négliger : la concentration des moyens sur un seul acteur ou la mutualisation des moyens de plusieurs acteurs sont des modalités possibles pour aboutir à un travail de qualité ;

► **parce qu'il peut être nécessaire de compléter les revues systématiques d'évaluations d'impact par d'autres types de « preuves »**, de nature plus descriptives ou qualitatives, pour éclairer la pratique et la décision, ces centres auraient plus largement la vocation à constituer des « centre de ressources » pour l'évaluation des dispositifs et l'identification des dispositifs qui « marchent » sur une problématique donnée ;

► **diffuser de manière adaptée des synthèses, des interventions intéressantes, des recommandations**, dans un format répondant aux attentes et aux besoins des décideurs publics, au niveau national et local, et des professionnels de terrain. Les « centres » accompagneraient les acteurs dans l'appropriation et la mise en œuvre des pratiques démontrées comme efficaces par le biais d'actions d'essaiage et joueront un rôle de vulgarisation des méthodes et des résultats.

Mission complémentaire

► **se doter d'une capacité d'expérimentation**. Au cours des dix dernières années, le recours aux évaluations d'impact a sensiblement augmenté en France, notamment dans le cadre d'expérimentations. Cependant, le nombre d'évaluations d'impact réalisées en France reste faible, rendant difficile l'élaboration de politiques fondées sur la preuve. De plus, dans chaque champ de politiques publiques, il existe souvent des dizaines d'interventions prometteuses, méconnues en dehors d'un territoire spécifique et qui pourraient se prêter à une telle évaluation. Il est donc proposé de doter ces structures des moyens de financer conjointement l'intervention et son évaluation d'impact, en s'inspirant des pratiques de l'*Education Endowment Foundation*. Cette mission permettrait d'une part de développer la connaissance sur les interventions dans le contexte français, et d'autre part de rapprocher le centre du terrain, renforçant en retour la pertinence des productions du centre.

LES PRINCIPAUX FACTEURS CLÉS DE SUCCÈS, ISSUS DE L'ANALYSE
DES WHAT WORKS CENTRES :

- s'inscrire dans la durée, et donc disposer de financements pluriannuels sur au moins 3 ans¹¹ ;
- disposer de moyens conséquents pour assurer les missions socles et complémentaires proposées¹² ;
- être indépendants dans la définition de leur programme de travail et la publication de leurs travaux ; cette indépendance peut être assurée par les statuts juridiques de la structure, ses modalités de financement et sa gouvernance ;
- être transparents, en publiant systématiquement ses revues et évaluations, et en précisant les méthodes utilisées lors de ses travaux ;
- mettre les attentes des décideurs et des praticiens au cœur de leur action, en les associant à la programmation et à la mise en œuvre des revues systématiques, des actions de diffusion, d'essaimage, ou d'expérimentations ;
- avoir un niveau d'exigence élevé, par le recours à des classifications d'évaluation permettant de distinguer des « natures de preuve » et des « niveaux de preuve » ;
- prêter une attention particulière à la transférabilité au contexte français des dispositifs ayant été démontrés comme efficaces dans d'autres pays.

Organisation d'un centre

Le modèle d'organisation le plus pertinent pour une structure inspirée des What Works Centres dépendra de plusieurs facteurs relatifs au champ de politique publique dans lequel elle s'inscrit :

- **l'organisation de l'action publique**, son degré de centralisation,

¹¹ La sécurité financière des What Works Centres britanniques leur a donné une visibilité pour lancer des études, et ils ont pu construire un socle conséquent de revues systématiques de littérature lors de leurs trois premières années.

¹² Les What Works Centres britanniques disposent d'un montant de l'ordre d'un million de livres par an pour les missions socles (revues de littérature et diffusion) et l'*Education Endowment Foundation* dispose de 14 millions de livres par an pour l'ensemble de ses missions, dont les expérimentations avec évaluation contrôlée.

What Works Centres britanniques : **quels enseignements pour des politiques fondées sur la preuve en France ?**

le paysage institutionnel, et les professionnels concernés ;

- **les pratiques existantes en matière de politiques fondées sur la preuve** en France, qu'il s'agisse de statistiques, d'évaluation, de travaux de nature plus qualitative, et de diffusion de la preuve aux décideurs et praticiens ;
- **la maturité et la quantité de travaux** d'évaluations d'impact et de revues systématiques.

En particulier, des centres d'inspiration « *What Works* » pourraient prendre différentes formes d'organisation, présentées ici.

► **Modèle 1**

Une équipe « hébergée » au sein d'administrations centrales qui porte les missions socles énoncées précédemment. Cette équipe disposera de moyens humains et financiers dédiés, d'une identité visuelle propre et d'un site internet, ainsi que d'un programme de travail donnant lieu à des publications publiques.

Avantages	Limites
En mobilisant des ressources et expertises existantes, cette équipe pourrait se monter relativement rapidement. Par sa proximité avec les décideurs nationaux, cette équipe sera bien placée pour identifier des questions « à enjeu » pour ces décideurs et influencer sur la décision publique.	En s'inscrivant dans une administration centrale, ces équipes ne jouiraient pas de la même indépendance que les What Works Centres pour définir leur programme de travail et publier leurs travaux. Ces équipes courraient également le risque d'être supprimées lors de réorganisations de l'administration. En se positionnant proche du centre, elles pourraient potentiellement avoir plus de difficultés à diffuser les enseignements auprès des professionnels de terrain.

► **Modèle 2**

Une structure indépendante avec une équipe dédiée qui poursuit un programme de travail élaboré avec son instance de gouvernance. Cette structure pourrait être nouvelle, ou être créée par la fusion d'organisations exerçant déjà des missions de type « *What Works* » dans le domaine.

Avantages	Limites
<p>L'indépendance de la structure permettra de fixer un programme de travail et de publier de manière systématique ses études, et de conduire ses travaux de manière transparente. Elle aura une plus grande capacité à répondre aux attentes des professionnels de terrain. Acteur tiers, elle sera potentiellement mieux placée pour instaurer un dialogue entre décideurs nationaux, professionnels, et chercheurs.</p>	<p>La structure aura potentiellement plus de difficultés à établir des liens avec les décideurs nationaux. Le lancement d'une telle structure peut être un processus relativement long, et elle requerra des financements conséquents qu'il conviendra de pérenniser.</p>

► **Modèle 3**

Un centre « réseau en étoile », porté par un acteur dont la fonction principale sera de planifier et de coordonner les actions conduites par d'autres acteurs sur un domaine. Le centre sera doté de moyens humains et financiers modestes, et agira par capillarité en mobilisant les ressources des partenaires autour d'un programme de travail commun.

Avantages	Limites
<p>En mobilisant des acteurs existants, le centre s'inscrira dans le paysage institutionnel, sans risquer de « faire doublon » avec les acteurs existants. Il pourra également bénéficier à la fois de l'expertise et du réseau des acteurs partenaires.</p>	<p>En l'absence de moyens conséquents, sa réussite dépendra de la volonté des partenaires. Le centre risque de devenir ainsi un groupe de coordination plus qu'un centre avec son propre programme de travail.</p>

Proposition 2 : Créer une alliance autour des politiques fondées sur la preuve

Quel que soit le domaine d'intervention, les acteurs font face aux mêmes défis pour élaborer et mettre en œuvre des politiques fondées sur la preuve : ressources contraintes, méthodes d'évaluation mal connues, difficultés à articuler temps de l'évaluation et temps politiques. Ces défis sont ressentis non seulement du côté des décideurs publics, mais également des professionnels de terrain et des chercheurs. Aujourd'hui, il n'existe aucun support commun à l'ensemble des parties prenantes dédié à la promotion de ces approches, à la rencontre et à l'échange des acteurs intéressés par ces questions pour les accompagner à faire face à ces défis.

L'Ansa et ses partenaires proposent ainsi de créer un réseau transversal réunissant décideurs publics, chercheurs et professionnels de terrain autour de la question des politiques publiques fondées sur la preuve, à l'image d'*Alliance for Useful Evidence* au Royaume-Uni. Cette alliance pourrait notamment contribuer à :

- animer le débat autour de ces questions par l'organisation de temps d'ateliers, de colloques, et de « webinaire » ;
- sensibiliser les décideurs publics aux méthodes des politiques fondées sur la preuve en l'intégrant dans leur formation initiale (Ena, Inet), ou continue (ex. CNFPT), ou formations *ad hoc* ;
- avancer les réflexions autour des politiques fondées sur la preuve, par la publication de guides pratiques et de documents de réflexions, par exemple sur les « standards de preuve », les différents types de connaissances et d'évaluation devant également être pris en compte¹³.

13 La connaissance à prendre en compte pour les politiques fondées sur la preuve est large et inclut les résultats des évaluations d'impact sans toutefois s'y limiter. Elle s'étend à d'autres ressources – allant de l'analyse des besoins, l'évaluation des fondements théoriques des différents programmes et politiques, des évaluations qualitatives (s'appuyant sur des observations, des entretiens, des études de cas), des évaluations de processus, jusqu'à des évaluations quantitatives (statistique descriptive, corrélation), également l'expertise des citoyens ou usagers sur leurs besoins et les réponses à y apporter. Ces autres ressources, de nature plus théorique, descriptive ou qualitative, sont essentielles notamment pour produire des diagnostics (analyse et compréhension des besoins), ainsi que l'élaboration de recommandations.

Études de cas de 5 What Works Centres

Early Intervention Foundation

Fiche d'identité	
Périmètre	Intervention auprès d'enfants, de la conception à 18 ans
Date de création	Février 2013
Statut	Association à but non lucratif, appartient au réseau des WWC
Taille de l'équipe	16 personnes environ
Budget annuel	Non Public
Site web	www.eif.org.uk

Historique

L'*Early Intervention Foundation* (EIF) a été créé en février 2013 par une dotation du ministère de l'Éducation. Il a été initialement soutenu par un groupement d'organisations dont *4Children*, *The Local Government Association* et *Achievement for All 3As* avec un financement initial du « *Department for Education* » pour 3 ans.

Graham Allen, parlementaire, a été à l'initiative de ce centre, grâce à un rapport qu'il a produit sur les interventions pour la petite enfance jusqu'aux jeunes adultes (« *early interventions* »), montrant leur impact et les coûts évités par rapport à des interventions plus tardives¹⁴, notamment inspiré par le *Washington State Institute for Public Policy*. Le *Dartington Social Research Unit* et *Michael Little*¹⁵ ont été consultés pour ces travaux.

L'objectif d'EIF est de :

- déterminer les programmes d'intervention auprès de la petite enfance

¹⁴ grahamallenmp.co.uk/early_intervention.

¹⁵ dartington.org.uk/about/staff/dr-michael-little.

et d'adolescents qui fonctionnent : les interventions les plus efficaces, et leurs coûts en comparaison des bénéfices obtenus;

- traduire cela en recommandations s'appuyant sur les enseignements de la recherche, à destination des décideurs publics locaux, des opérateurs, et éventuellement de financeurs;
- promouvoir les interventions ciblant les enfants qui permettent de résoudre des problèmes de manière moins coûteuse et à un stade où ils sont moins compliqués que des interventions plus tardives ne pourraient le faire.

Organisation

Le centre est une association à but non lucratif et fait partie du réseau national des What Works Centres. Il est gouverné par un conseil d'administration.

Le centre est doté d'une équipe de 16 personnes chargées selon leurs compétences respectives de l'analyse des évaluations (« *evidence analysts* »), de l'analyse économique des interventions et des coûts évités qui s'y rapportent et enfin de la dissémination des recommandations.

Il s'appuie sur un « *evidence panel* », composé de chercheurs universitaires et experts, qui viennent en appui de la structure. Ils ont en particulier pour rôle d'examiner les synthèses et les recommandations produites par le centre.

La façon de nommer les membres du panel est en cours de modification.

Il est financé aujourd'hui par un ensemble d'acteurs publics, dont des ministères et des agences nationales, des instituts de recherche et des fonds privés volontaires. Les financements sont maintenant de plus en plus ciblés sur des projets/thèmes choisis avec les financeurs.

Action

Synthétiser

Objectifs

L'EIF produit des synthèses d'évaluations portant sur les interventions

What Works Centres britanniques : quels enseignements pour les politiques fondées sur la preuve en France ?

auprès d'enfants – principalement la petite enfance et les adolescents. Cela permet de déterminer, à partir de l'ensemble des évaluations d'impact scientifiques, quelles interventions ont un impact positif, lesquelles sont prometteuses et enfin lesquelles pourraient bénéficier d'amélioration ou de recherche complémentaire.

Actions menées

Parmi ses travaux, l'EIF a publié en juillet 2016 une synthèse des évaluations pour les interventions sur la relation parent-enfant¹⁶ se concentrant sur des programmes existants au Royaume-Uni. Cette synthèse concerne les interventions visant à améliorer la relation parent-enfant d'avant la naissance jusqu'à l'âge de 5 ans, afin de contribuer au développement des enfants en termes d'attachement, de comportement, et sur le plan cognitif, et de prévenir des difficultés futures. Parmi 100 interventions identifiées par une revue systématique de littérature publiée en 2015, 75 interventions ont été retenues pour être étudiées de manière plus approfondie au vu des évaluations existantes et le coût des interventions¹⁷.

Le rapport présente les résultats de manière agrégée par thématique pour avoir une vue d'ensemble, et fait ensuite la distinction des résultats pour chaque programme. La revue a identifié 17 programmes comme ayant des impacts positifs et mesurables (niveau 3 ou 4), tandis que 18 programmes sont identifiés comme étant prometteurs (niveau 2). Enfin certaines interventions ont été évaluées de manière scientifique mais ne montrent aucun impact, ou n'ont pas été évaluées de manière assez robuste.

L'EIF travaille sur plusieurs autres thèmes. Par exemple, le ministère de l'intérieur britannique a commandé des études sur les interventions pour prévenir la violence juvénile et l'intégration de jeunes dans des gangs. L'EIF a donc publié plusieurs rapports à ce sujet, notamment sur les facteurs de risque, comment identifier les jeunes personnes présentant ces risques

¹⁶ www.eif.org.uk/publication/foundations-for-life-what-works-to-support-parent-child-interaction-in-the-early-years.

¹⁷ L'étude ne regroupe pas l'ensemble de tous les programmes de parentalité pour la petite enfance existants au Royaume-Uni, car une telle base de données n'existe pas.

et les interventions qui peuvent fonctionner ou non, et ce que les différentes évaluations peuvent nous apprendre.

Méthodes

Afin de réaliser de telles synthèses, l'EIF étudie de manière approfondie les évaluations existantes pour chaque intervention de manière individuelle¹⁸. Pour chaque intervention, une revue systématique est menée au sein de l'EIF, ainsi qu'une analyse des coûts de l'intervention. Ces revues systématiques sont ensuite étudiées par des experts extérieurs, à titre individuel puis réunis en panel pour s'accorder avec l'équipe de l'EIF sur une première notation, à partir de la qualité des études sous-jacentes.

Afin de s'assurer que les synthèses reflètent au mieux l'état de l'art de la recherche, celles-ci sont enfin examinées par l'*evidence panel*. Les chercheurs du Panel apportent leurs expertises sur la qualité dans la construction des mesures des différentes études, la confiance que l'on peut avoir dans les résultats, etc. à travers des échanges avec l'équipe ayant réalisé la synthèse.

Ce travail est mené en lien avec les opérateurs des interventions, qui peuvent apporter des informations complémentaires ou corriger les incompréhensions.

Évaluer les évaluations

Dans le rapport « *Foundations for Life: What Works to support parent child interaction in the early years* », l'EIF effectue une classification¹⁹ des interventions identifiées. Chaque intervention est notée selon à la fois la qualité de son évaluation et l'existence d'un impact positif ou non. Les critères de qualité de l'évaluation s'inspirent du *Maryland Scientific Scale*, les travaux de *Blueprints*, et les « *Standards of Evidence* » de NESTA.

¹⁸ Cela diffère des approches de méta-analyses qui vont agréger des indicateurs pour différentes interventions ayant le même objectif, donnant lieu à un ensemble de scores pour différents types de pratiques. L'approche d'EIF facilite la comparaison entre différentes interventions sur la base des études spécifiques.

¹⁹ www.eif.org.uk/wp-content/uploads/2016/07/What-do-the-EIF-ratings-mean.pdf.

What Works Centres britanniques : quels enseignements pour les politiques fondées sur la preuve en France ?

Elles comparent aussi le coût estimé de l'intervention par enfant selon une méthode définie par l'EIF, pour que les coûts soient comparables entre évaluation. L'échelle définie pour une revue systématique sera adaptée à l'objet de l'étude et se nourrira de l'expérience acquise sur des revues précédentes.

L'EIF évaluera l'impact en se fondant sur les études les plus robustes : une étude quasi-expérimentale prendra le pas sur une étude avant-après sans échantillon de comparaison. Pour obtenir le niveau 3, les études les plus robustes doivent notamment être cohérentes entre elles et obtenir au moins un résultat principal (attachement, comportement, cognitif) pour les enfants. Le niveau peut être renforcé (3+) si des études de niveau 2+ apportent des éléments concordants. De nombreux autres critères sont pris en compte pour une approche fine des revues systématiques.

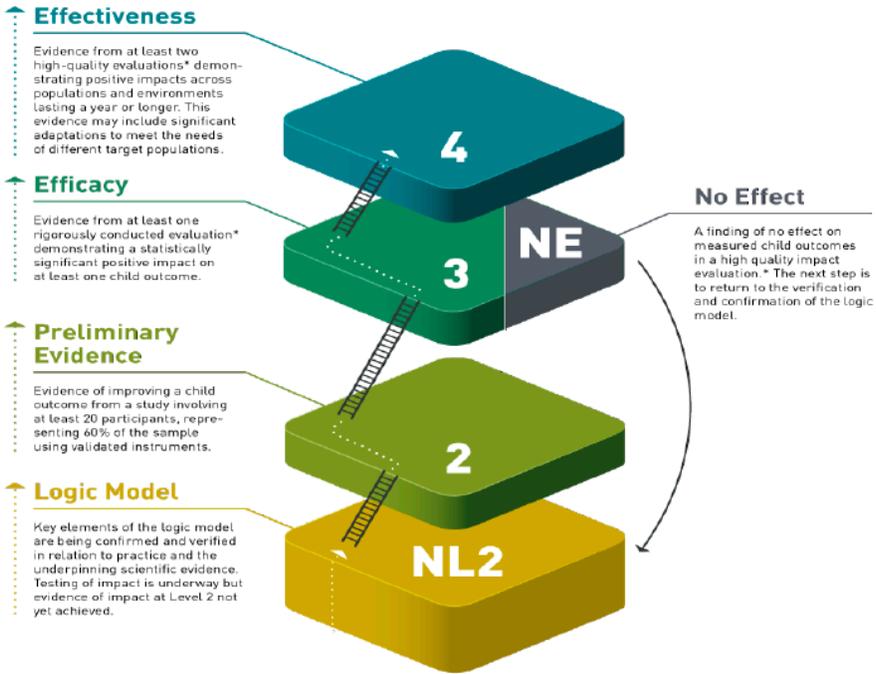
Les interventions sont également notées sur leur réplicabilité – le niveau 4 correspond notamment à des évaluations réalisées dans au moins deux contextes différents. Cela n'inclut pas des évaluations spécifiques en cas de changement d'échelle – comme peut le faire EEF par exemple.

Enfin, l'EIF prend en compte la présence des théories du changement dans la cotation des interventions. En effet, avant d'avoir des évaluations plus probantes, ces schémas permettent de vérifier la cohérence de l'intervention en précisant les liens entre l'intervention et les résultats visés. L'EIF travaille avec d'autres WWC pour développer un standard de qualité des théories du changement et ainsi les qualifier.

“Nous avons développé une façon de reconnaître les théories du changement correctement alimentées par des études scientifiques. Nous cherchons celles qui sont plus qu'un modèle logique, qui sont plus des schémas de programme. Il faut aussi que cette théorie du changement s'appuie de manière claire sur la littérature scientifique. Par exemple, elle doit montrer clairement les preuves scientifiques qui amènent à définir la population cible, avec quelle régularité il faut les voir, et avec quelle intensité.”

Dr Kirsten Asmussen, Evidence Analyst, EIF.

Figure 1. A dynamic model of the development of evidence



*High quality evaluations do not need to be randomised control trials if a relevant and robust counter-factual can be provided in other ways.

Source : www.eif.org.uk.

Obstacles rencontrés et enseignements de la démarche

L'EIF est confronté à la complexité de définir de manière incontestée si une intervention évaluée a effectivement eu des effets positifs et mesurables.

“Un ensemble de critères est utilisé pour définir la qualité des évaluations et donc le niveau de classement des interventions, sans que cela permette toujours de trancher de manière définitive. J'appelle cela les

What Works Centres britanniques : **quels enseignements pour les politiques fondées sur la preuve en France ?**

“500 nuances de 'grade' (critères d'évaluation des évaluations)” car il faut prendre en considération un ensemble de critères pour déterminer dans quelle mesure une intervention atteint les différents seuils.”

Dr Kirsten Asmussen, Evidence Analyst, EIF.

L'EIF considère donc qu'il est important de prêter attention aux interventions qui sont dans cette « zone grise », à propos desquelles les études ne sont pas concluantes, ne sont pas assez robustes, ou ne donnent pas les résultats espérés. L'EIF souhaite ainsi donner une chance aux programmes d'être affinés, améliorés et mieux conçus, notamment grâce aux enseignements des évaluations, et ensuite de faire l'objet d'une évaluation plus robuste. En effet, l'EIF espère que la mise en avant de projets à améliorer incitera les acteurs du secteur à faire 'monter en grade' ces interventions dans l'échelle de l'EIF.

L'EIF a également constaté que la littérature évolue en continue. En effet, la production de nouvelles études plus poussées peut changer la cotation des interventions, en remettant parfois en cause celles qui paraissaient prometteuses ou prouvées avant. Ces évolutions génèrent parfois de la confusion dans l'esprit du public cible de l'EIF.

Le programme « *Family Nurse Partnership* » a par exemple été évalué plusieurs fois dans différents contextes de manière randomisée aux États-Unis et une fois aux Pays-Bas. Il s'agit d'infirmières très qualifiées qui rendent visite régulièrement à des jeunes mamans du début de grossesse jusqu'aux 2 ans de l'enfant – les jeunes femmes ont en général autour de 19 ans. C'est un programme avec un public ciblé, et les effets mesurés dans ces études sont très positifs tant au moment de la naissance de l'enfant et les premières années que dans les années suivantes. L'intervention est aussi déployée très largement au Royaume-Uni. Une évaluation randomisée y est en cours et les premières conclusions, sur les résultats obtenus grâce aux visites prénatales, sont plutôt décevantes par rapport à ce que les autres études laissaient espérer (bien qu'elles soient suffisamment bonnes pour obtenir la note 4+).

Parmi les explications envisagées, il y a le fait que le suivi sanitaire des jeunes femmes enceintes et défavorisées est de meilleure qualité au Royaume-Uni qu'il ne l'est aux États-Unis. Dans les études menées, le contrefactuel (et donc le point de comparaison pour évaluer l'effet) est donc différent (jeunes femmes ayant très peu accès à des soins aux E-U, jeunes femmes ayant accès de manière plus régulière à des soins au R-U). Ces résultats poussent les responsables du programme au Royaume-Uni à améliorer le programme prénatal – en changeant de pays, il s'agit donc plutôt d'une adaptation que d'une réplique de l'intervention.

Expérimenter

L'EIF n'a pas pour rôle d'expérimenter directement ni n'a de moyens pour commander des évaluations de programme. Le centre a quelques fois apporté son expertise à des collectivités locales pour obtenir des financements de recherche dans des appels à projet. L'EIF a aussi pour objectif que les résultats des synthèses informent les futurs programmes de recherche des instituts de recherche.

Parmi les leçons retenues pour l'expérimentation, l'EIF souligne qu'il est arrivé que des commanditaires financent une évaluation randomisée conséquente, mais que suite à différentes raisons, le projet échoue. En particulier, en raison d'une attrition trop importante, une trop petite taille des échantillons est l'un des problèmes fréquemment rencontrés.

“Il vaut mieux faire un bon projet pilote d'abord.”

Dr Kirsten Asmussen, Evidence Analyst, EIF.

Diffuser et accompagner l'adoption

Objectifs

La principale activité d'EIF est la réalisation de synthèses très complètes au sujet des interventions auprès des enfants, de la petite enfance aux jeunes adultes. L'EIF a également amorcé des actions de diffusion des

résultats obtenus. L'EIF porte maintenant des efforts plus importants sur l'intégration, par les parties prenantes, des résultats scientifiques synthétisés par l'EIF dans leurs modes de décisions.

Méthodes

L'adoption de nouvelles interventions repose beaucoup sur l'organisation de l'action publique du pays concerné.

Pour ces programmes ciblant les enfants, les collectivités locales sont l'acteur principal au Royaume-Uni car elles sont responsables des services d'enfance et de la famille (protection de l'enfance et placement, travail social de secteur). Bien que leur action soit encadrée par la loi, les collectivités ont des marges de manœuvre dans le choix des programmes qu'elles financent. Un des principaux objectifs de l'EIF est donc de convaincre d'adopter ou tenir compte de ses préconisations. L'EIF continue de réfléchir à la meilleure manière de favoriser l'adoption d'interventions efficaces lorsqu'une compétence est ainsi décentralisée. Entre autres, il a suggéré que les collectivités aient le choix dans un panel d'interventions aux positifs prouvés.

Jusqu'à présent, l'EIF a ciblé principalement les collectivités locales tant dans ses publications que ses actions de diffusion sur le terrain. Au-delà des collectivités, il leur semble intéressant de disséminer l'approche et les résultats de l'EIF auprès des cibles suivantes, en développant des publications et des actions adaptées :

- les services de protection de l'enfance et ceux qui interviennent sur la parentalité ;
- le *Department for Education* (ministère de l'Éducation) et les autres directions générales de l'administration en lien avec ce sujet, dont le *Department of Health* ;
- le *Department for Communities and Local Government* qui doit réaliser la mise en œuvre ;
- les professionnels de terrain et les structures locales : ils sont au premier rang pour identifier les difficultés qu'ils rencontrent et éventuellement reconnaître ce qui pourrait les aider ;

- les opérateurs qui développent les programmes ;
- les parents.

Publications en ligne

Le site web permet d'accéder aux publications de l'EIF sur les résultats des synthèses produites, ainsi que sur les méthodes pour obtenir ces synthèses. Il est le véhicule principal de la diffusion des travaux et des résultats de l'EIF. Par exemple, le rapport sur les relations parents-enfants (cité plus haut) présente l'ensemble des dispositifs étudiés, avec les méthodes d'évaluation, et met en avant un ensemble de recommandations sur les types d'interventions à mettre en œuvre, les actions à privilégier, etc. Le site internet présente aussi un certain nombre d'outils pratiques pour permettre aux parties prenantes de s'emparer du sujet :

- une « *Maturity matrix* », qui permet aux collectivités locales d'identifier où elles en sont en terme de connaissance de leurs publics, du développement de leurs équipes et de leurs services, des leviers d'action et de leur niveau de préparation à mettre en œuvre des actions de prévention ;
- des outils pour aider les acteurs à justifier d'un « investissement » dans des actions de prévention, et des études menées sur les coûts d'une intervention tardive et les bénéfices d'intervention de prévention ;
- des outils pour permettre à la police d'agir en prévention auprès des jeunes, en coopération avec les autres acteurs locaux pertinents.

Accompagnement d'acteurs

L'EIF a testé au fil du temps différentes méthodes pour favoriser l'adoption des actions efficaces par les parties prenantes, en particulier les collectivités.

Dans le cadre de son programme « *pioneering places* », EIF a travaillé avec 20 territoires pour identifier des actions permettant de mieux intégrer le principe de la prévention précoce dans leurs services²⁰. EIF a préconisé des actions permettant d'améliorer l'identification de la population cible et ensuite la meilleure manière de les accompagner, s'appuyant sur les interventions efficaces.

20 www.eif.org.uk/pioneering-place.

L'EIF propose aussi des ateliers, des partages de pratique, du conseil sur la façon de mettre en œuvre de telles interventions, de s'assurer du résultat, de bien cibler le public concerné, etc.

Obstacles rencontrés et leçons tirées

Diffuser les informations, aussi bien présentées soient-elles, ne suffit pas. L'EIF effectue donc des recherches pour identifier les meilleurs moyens pour que les décideurs et les professionnels de terrain s'emparent des résultats et des outils mis à leur disposition.

Au-delà de la prise de décision, la qualité de la mise en œuvre de ces programmes représente un enjeu conséquent, notamment en ciblant le bon public. Ce sujet est traité dans le livre de Dr Kirsten Asmussen « *The evidence-based parenting practitioner's handbook* » en ce qui concerne les interventions sur la parentalité.

L'un des axes qui peut être mis en avant pour investir dans des programmes ciblant la petite enfance et les adolescents a trait aux coûts évités sur le long terme. Cependant, les résultats des interventions en termes de coûts évités ne sont pas toujours concluants. Pour les interventions auprès d'enfants en bas âge, la réduction de certains coûts n'est visible que de nombreuses années après l'intervention. Ces bénéfices peuvent aussi être difficiles à évaluer ou ne pas se traduire directement par des réductions de coûts. La difficulté à recueillir les bénéfices financiers peut conduire à remettre en question de tels programmes d'un point de vue politique, même s'ils ont de nombreux autres avantages par ailleurs.

Perspectives

- Développer les actions de dissémination et favoriser l'adoption : il s'agit notamment de développer de nouvelles pratiques pour cette diffusion, s'inspirant des enseignements des d'études scientifiques sur ce sujet (cf. travaux portés par Nesta). L'EIF a aussi la volonté de toucher l'ensemble des parties prenantes, au-delà des collectivités locales, notamment

en développant des outils et des approches spécifiques.

- Influencer les programmes de recherche et les acteurs qui pourront porter les expérimentations pour aller plus loin là où les études ne sont pas suffisantes.
- Continuer à convaincre les décideurs de l'importance des interventions de prévention auprès des enfants.

Enseignements

Enseignements communs pour les What Works Centres

- Déterminer de manière définitive si une intervention a un effet positif mesurable, et sous quelles conditions, est une tâche complexe. Toutefois, certaines interventions ont pu être validées de manière scientifique. De plus, la présence d'interventions prometteuses, mais dont l'impact n'a pas encore été testé, représente une opportunité. Ces interventions peuvent faire l'objet d'expérimentations scientifiques et ainsi être améliorées.
- L'EIF cherche aujourd'hui à déterminer les meilleures méthodes de diffusion des interventions efficaces et prometteuses. L'EIF envisage notamment de nouvelles actions qui tiendraient compte de ce qui a été déjà testé et des essais en cours au Royaume-Uni, notamment par l'EEF. La meilleure méthode d'essai dépendra avant tout du contexte administratif et politique du pays et du champ ciblé.

Enseignements plus spécifiques à la petite enfance

- De nombreuses études existent sur les interventions, notamment sur les relations parents-enfants et la parentalité, et un grand nombre d'entre elles ont été revues et synthétisées par l'EIF : cela constitue donc une base d'analyse très riche pour ce sujet, à confronter avec ce qui existe déjà ou a déjà été fait en France.
- Dans les interventions ciblant les familles et les enfants les plus à risque, la qualité de l'identification des publics-cible est essentielle pour mener à bien les interventions et obtenir les effets escomptés (tout en maîtrisant le coût de l'intervention).

Education Endowment Foundation

Fiche d'identité	
Périmètre	Réduction des inégalités scolaires
Date de création	Juillet 2011
Statut	Association
Taille de l'équipe	15 salariés
Budget annuel	14 millions de livres sterling en 2015, dont 12 millions de subventions pour des innovations et de la recherche
Site web	educationendowmentfoundation.org.uk

Historique

L'*Education Endowment Foundation* (EEF) a été créé en 2011 par l'association *The Sutton Trust*, en partenariat avec une fondation – *Impetus Trust* – et avec une dotation de 125 millions de livres sterling du ministère de l'Éducation anglais. Le ministre de l'Éducation, Michael Gove a lancé cette initiative en novembre 2010, dont l'objectif affiché était d'élever le niveau des études dans les écoles « en difficultés ». Elle s'inspirait de *Race to the Top*, un programme ambitieux de l'administration d'Obama aux États-Unis²¹. L'EEF s'est engagée à dépenser la totalité de sa dotation initiale à l'horizon 2026. En mars 2013, l'EEF a été désigné What Works Centre pour l'amélioration du niveau scolaire par le gouvernement.

L'objectif d'EEF est de s'attaquer aux inégalités scolaires en relevant le niveau de réussite scolaire chez les élèves les plus pauvres dans les écoles les plus en difficultés. Par le biais de l'expérimentation, de la synthèse

21 www.gov.uk/government/news/new-endowment-fund-to-turn-around-weakest-schools-and-raise-standards-for-disadvantaged-pupils.

de la recherche et de la diffusion de la connaissance, l'EEF vise à soutenir les décideurs à investir dans les dispositifs les plus efficaces (en termes d'impact) et les plus efficaces (en termes de coût du service). Ainsi, l'EEF cible à la fois les directeurs d'établissements scolaires, enseignants, parents d'élèves, chercheurs et décideurs politiques de l'éducation nationale. Ce rôle est d'autant plus important dans un contexte de décentralisation des budgets vers des établissements scolaires. En outre, les écoles reçoivent des financements spécifiques pour les élèves venant des familles les plus pauvres.

Organisation

Le centre est une association indépendante sous l'égide d'un conseil d'administration composé de sept administrateurs. La direction et le conseil d'administration de l'EEF bénéficient de conseils stratégiques de la part d'un « *Advisory Board* » constitué de 11 membres, personnalités des mondes de l'éducation, de la politique publique et de l'entreprise. EEF est dotée d'une équipe de 19 personnes sous la direction du directeur Sir Kevan Collins. En plus des pôles administratifs et de communication, EEF s'organise en trois équipes : *Grants, Evaluation et Dissemination and Impact*. Les membres de ces équipes viennent d'horizons différents (ministères, collectivités territoriales, think tanks et d'associations). Plusieurs ont travaillé en tant qu'enseignants, leur donnant une expérience directe du quotidien de leur principal public cible.

L'indépendance d'EEF vis-à-vis du gouvernement lui permet d'élaborer et poursuivre une vision à long terme. De plus, l'approche du centre l'a tenu à l'écart des polémiques sur la politique d'éducation. Par exemple, depuis 2011, du fait de l'essaimage des *Free Schools et Academies* le gouvernement a été accusé par certains de « libéraliser » l'éducation en Angleterre. Le focus d'EEF – sur les méthodes pédagogiques déployées au sein de la classe – l'a, en quelques sortes, protégé de ces débats.

Action

Synthétiser

Dès sa création, EEF a mené un travail important de synthèse de la recherche en éducation. S'appuyant sur des travaux réalisés par l'université de Durham pour le *Sutton Trust*, l'EEF a diffusé dès 2011 son *Teaching & Learning Toolkit*. Il s'agit d'une synthèse des synthèses de la recherche sur l'efficacité de différentes interventions éducatives. Cette synthèse porte donc principalement sur les résultats d'évaluations expérimentales ou quasi-expérimentales. Cette revue a permis de faire un état de l'art des sciences de l'éducation. Contrairement à d'autres travaux de synthèse préexistants, le *Teaching & Learning Toolkit* est conçu comme un outil d'aide à la décision pour les professionnels. Le *Toolkit* permet de comparer différentes « interventions pédagogiques » selon trois dimensions : l'impact, la rigueur de la preuve (la qualité et le nombre d'évaluations), et le coût. Pour comparer l'impact des interventions, l'équipe de recherche a traduit les résultats des évaluations en une seule mesure : mois supplémentaires de progrès éducatif pour le groupe test (comparé au groupe témoin). Arriver à un tel niveau de concision et avec des conclusions utiles est une tâche difficile. En effet, il existe une tension entre les chercheurs (qui souhaitent rester très précis et arguent par exemple que la mesure unique en mois de progrès est imparfaite) et les besoins des décideurs, qui seraient « paralysés » s'ils recevaient trop d'information.

“Le Toolkit est une méta-analyse des méta-analyses. Des chercheurs ont déjà fait ce genre de synthèse. La différence est que nous l'avons présenté d'une manière accessible aux enseignants et aux directeurs d'établissements scolaires.”

James Turner, Deputy Chief Executive.

Le *Toolkit* est désormais un outil interactif intégré au site web d'EEF. Il est également évolutif. Mis à jour tous les six mois par l'équipe de recherche de l'université de Durham, il est passé de 11 à 34 interventions éducatives.

Outre l'ajout de nouvelles interventions, l'analyse des interventions a également été mise à jour pour prendre en compte de nouvelles études. EEF a également financé des revues de littérature spécifiques permettant de combler des manques dans le Toolkit.

Le Toolkit a également aidé l'EEF à identifier les types d'interventions pour lesquels la recherche mériterait d'être renforcée, et qui pourraient faire l'objet de nouvelles expérimentations financées par l'EEF. À termes, ces expérimentations produisent des résultats qui alimentent le Toolkit. L'exemple des aides-enseignants (*teaching assistants* en anglais) illustre bien ce cycle vertueux.

“Les évaluations montraient que les aides-enseignants n'avaient pas d'effets ou des effets négatifs sur les élèves. Le Toolkit a fait ressortir cette problématique. Ce constat a conduit à plusieurs actions à l'EEF. On a mené des expérimentations qui ont montré que lorsqu'ils sont déployés de façon ciblée les aides-enseignants ont un impact positif pour les élèves. Ces résultats ont, à leurs tours, fait évoluer les constats dans notre Toolkit.”

Jonathan Sharples, chercheur expérimenté, Education Endowment Foundation.

Extrait du Teaching & Learning Toolkit



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

Expérimenter

Au moment de la création d'EEF en 2011, peu d'évaluations expérimentales avaient été réalisées dans le secteur de l'éducation en Angleterre.

Un des objectifs de l'EEF était donc de produire de nombreuses évaluations expérimentales permettant de tester l'efficacité des interventions éducatives dans le contexte anglais. Grâce à sa dotation importante, l'EEF disposait des moyens nécessaires pour financer à la fois les projets innovants et leurs évaluations. Ce financement des structures portant les expérimentations fait partie des facteurs de succès du dispositif : il incite les structures à se soumettre à une évaluation scientifique malgré les risques organisationnels que cela peut représenter pour elles, si les impacts du programme testé s'avèrent nul ou négatif.

Depuis 2011, l'EEF lance chaque année deux vagues d'appel à projets

« ouverts », et une vague « thématique » visant un aspect de l'éducation – tels que les technologies numériques ou l'éducation financière. EEF reçoit en moyenne 200 réponses pour chaque vague pour une dizaine de projets retenus. Les candidatures proviennent principalement d'associations qui interviennent dans les écoles, mais également des établissements scolaires et des chercheurs. Les écoles tendent à avoir des propositions de qualité moindre, mais lorsqu'elles ont une pratique qu'elles cherchent à diffuser, elles convainquent plus facilement d'autres écoles de participer.

L'équipe de « Grants » évalue d'abord la maturité des projets et leur attribue une note parmi les trois niveaux de développement, correspondant à trois niveaux d'accompagnement d'EEF :

- ▶ pour les projets à un stade précoce de leur développement, EEF propose un **projet pilote** pour tester la validité du concept, sans qu'il ait d'évaluation d'impact ;
- ▶ pour les projets ayant déjà été testés mais disposant de peu d'éléments sur leur impact, EEF propose une **expérimentation d'efficacité** (*efficacy*) pour évaluer le projet dans des conditions idéales ;
- ▶ pour les projets bénéficiant d'éléments plus probants quant à leur impact, EEF propose une **expérimentation à grande échelle** (*effectiveness*) pour

évaluer le projet en conditions réelles de déploiement afin de préparer une éventuelle généralisation. L'EEF peut ainsi accompagner les projets tout au long de leur développement, de la conception à la généralisation.

Une fois le projet identifié, l'équipe d'évaluation d'EEF prépare le lancement de l'évaluation en collaboration avec le porteur de projet. EEF a recours aux évaluateurs externes afin de s'assurer de l'indépendance des évaluations. L'EEF privilégie des évaluations de qualité. D'une part, les méthodes expérimentales sont généralement utilisées pour analyser les effets du projet (105 des 127 évaluations lancées jusqu'alors ont été des évaluations dites « randomisées »). D'autre part, l'EEF demande une évaluation de processus permettant d'analyser les conditions de réussite de l'intervention. Enfin, EEF porte une attention aux effets longitudinaux des projets et a ainsi créé un outil permettant de suivre, au sein de la base de données de l'Éducation nationale, les résultats scolaires des bénéficiaires de ses projets jusqu'à la fin de leur scolarité.

Malgré de nombreuses réussites, l'EEF a rencontré des difficultés dans le cadre de certaines évaluations. Quelques expérimentations qui ont été retenue par l'EEF n'ont pas été lancées car le porteur de projet n'a pas pu intégrer l'évaluation dans le projet. Ces difficultés ont souvent tenu au fait que l'évaluation n'a pas été suffisamment prise en compte dans la conception de l'expérimentation proposée. L'EEF a donc changé son approche.

“Aujourd'hui, avant que le projet ne passe au dernier comité de financement, on réunit le porteur de projet et l'évaluateur. L'idée est de co-construire l'évaluation et faire en sorte qu'elle s'intègre au projet avant qu'il soit validé. Ceci permet de s'assurer, par exemple, que l'échantillon soit suffisant pour avoir l'effet de taille nécessaire pour bien évaluer l'impact du projet.”

Camilla Nevill, Evaluation Manager.

Plus globalement, l'EEF a fait face à un manque de connaissance

What Works Centres britanniques : **quels enseignements pour les politiques fondées sur la preuve en France ?**

de l'évaluation par les porteurs de projet, et à un manque d'expertise dans les méthodes expérimentales parmi les chercheurs. L'EEF a donc du sensibiliser les acteurs et accompagner une montée en compétence des évaluateurs. L'équipe d'évaluation a ainsi produit de nombreuses ressources disponibles sur le site internet de l'EEF²². L'EEF souligne aussi s'être amélioré au fil du temps dans la construction des dispositifs d'évaluation, en intégrant l'évaluation de processus, en ayant une meilleure conception face au risque d'attrition.

Les résultats des évaluations ont également généré des débats. En effet, sur environ 50 évaluations restituées, seulement 12 des projets évalués ont été jugés suffisamment prometteurs pour bénéficier d'un second soutien financier de l'EEF – le plus souvent pour une expérimentation à grande échelle. En effet, il est difficile pour certains porteurs de projet d'accepter que leur intervention n'ait pas eu d'effets mesurables sur les élèves. Cependant, certains projets ont eu des résultats très intéressants à l'image des six expérimentations sur les aides-enseignants.

“Nous avons financé six expérimentations pour tester différents modèles d'intervention des aides-enseignants, axant leur travail sur des actions d'alphabétisation mises en place en individuel (« one to one ») avec l'élève. Toutes ces interventions ont eu des effets positifs, ce qui est exceptionnel.”

Jonathan Sharples, chercheur expérimenté, Education Endowment Foundation.

Globalement, l'EEF a donné tort à ses détracteurs qui craignaient une résistance forte de la part des enseignants quant aux méthodes d'évaluation expérimentales. En effet, il a engagé 127 évaluations dont 105 avec « randomisation », impliquant 7000 écoles, soit une école sur quatre en Angleterre.

“Il y a quelques années il y avait eu seulement une évaluation « randomisée » dans le secteur de l'éducation en Angleterre.”

22 educationendowmentfoundation.org.uk/evaluation/resources-centre.

Tout d'un coup nous en avons plus de 100. C'est comme si on avait appuyé sur un bouton!"

Jonathan Sharples, chercheur expérimenté, Education Endowment Foundation.

Diffuser et accompagner l'adoption

De 2011 à 2015, le Teaching and Learning Toolkit a été au cœur de l'action de diffusion de l'EEF. Ce Toolkit synthétise désormais les résultats de 10 000 évaluations d'impact sur une page du site internet. Il représente donc une sorte de « guichet unique » vers la recherche en éducation. Il est clair, accessible et, surtout, constitue une aide à la décision pour des professionnels de l'éducation.

"Le Toolkit répond aux contraintes de temps auxquels les professionnels de terrain font face : manque de temps, trop d'informations, incapacité à identifier celles qui sont utiles. Le Toolkit a changé la donne en rendant la recherche en éducation utile pour des professionnels."

Jonathan Sharples, chercheur expérimenté, Education Endowment Foundation.

Depuis 2011, l'EEF a donc fait un travail intensif de sensibilisation des professionnels de l'éducation à l'existence et à l'utilisation du Toolkit, avec des effets très intéressants en termes d'accès à l'outil.

"Aujourd'hui, deux-tiers des directeurs d'établissements scolaires disent qu'ils utilisent le Toolkit. Certains disent que c'est comme le magazine « Que choisir ? » car il est clair, indépendant et fait autorité dans son domaine."

Jonathan Sharples, chercheur expérimenté, Education Endowment Foundation.

Ce succès tient certes à l'accessibilité de l'outil mais également au positionnement du Toolkit en tant que ressource complémentaire au savoir-faire des professionnels et à la connaissance locale. Cependant, l'équipe de Dissemination and Impact de l'EEF est

What Works Centres britanniques : quels enseignements pour les politiques fondées sur la preuve en France ?

consciente que la diffusion de l'outil ne suffit probablement pas pour transformer les pratiques.

“Créer des ressources ne suffit pas. Le transfert de la connaissance est un processus social, il faut créer des motivations et des opportunités de faire vivre la connaissance.”

Jonathan Sharples, chercheur expérimenté, Education Endowment Foundation.

Par conséquent, l'EEF a décidé de mener des actions de diffusion plus ciblées. Depuis 2015, l'EEF a notamment lancé deux actions suivant deux méthodes :

- **les campagnes.** Il s'agit de campagnes de sensibilisation et d'accompagnement à l'appropriation d'interventions efficaces, à l'image de la campagne relative aux aides-enseignants²³. En effet, l'EEF a décidé de traduire les enseignements de la recherche sur le déploiement des aides-enseignants (dont les six expérimentations financées par EEF) en un guide opérationnel²⁴. Ce guide a été accompagné d'outils pour favoriser l'appropriation des bonnes pratiques, tels qu'une auto-évaluation des pratiques des aides-enseignants.
- **les menus.** L'approche est simple : constituer une liste d'interventions issues du *Toolkit*, et favoriser leur adoption par les écoles par le biais d'une incitation financière et d'un accompagnement. Un premier projet pilote est en cours dans une collectivité à l'est de Londres.

²³ educationendowmentfoundation.org.uk/campaigns/making-best-use-of-teaching-assistants.

²⁴ educationendowmentfoundation.org.uk/public/files/Publications/Campaigns/TA_Guidance_Report_MakingBestUseOfTeachingAssistants-Printable.pdf.

Campagne sur les aides-enseignants : extrait de l'outil d'auto-évaluation

A self-assessment guide

RECOMMENDATIONS I AND II Deployment of TAs in the classroom



1 INEFFECTIVE

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.

This arrangement causes a 'separation' effect: high-supported pupils have fewer opportunities to engage with whole-class teaching, and have fewer interactions with teachers and peers.

Teachers assume that a high amount of TA support meets the needs of lower-attaining and/or SEN pupils, and as a result spend less time working with these groups.

Personal relationship may be good, but there is a lack of cohesive and effective teamwork between teachers and TAs in the classroom.

TAs rarely rove around the classroom and support whole-class teaching (such as demonstrating equipment, or working at the whiteboard).

2 IMPROVING

There is variability in the quality and appropriateness of teachers' moment-by-moment decision-making regarding TA deployment.

In some lessons, TAs are effectively supplementing the whole-class teaching, although this is not consistent.

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.

There is little or no attempt to redeploy TAs during lessons in response to the real time needs of pupils.

3 EXEMPLARY

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 work with lower-attainers as others.

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.

Teachers and TAs work effectively as a team, with a shared understanding of their respective roles in achieving lesson objectives. Teachers' moment-by-moment decisions regarding TA deployment are driven by pupils' needs.

TAs support whole-class delivery at the front of the class, rove the room or undertake administrative tasks to ease teachers' workload.

Source : educationendowmentfoundation.org.uk.

Bien que l'EEF ait fait évoluer ses méthodes de diffusion, l'équipe reconnaît qu'ils ne savent pas encore ce qui « fonctionne » en la matière. En effet, très peu de projets de transfert de connaissance ont été évalués pour vérifier leurs effets sur les pratiques des professionnels. L'EEF a donc décidé de tester ces méthodes de diffusion dans le cadre de véritables expérimentations. L'EEF expérimente actuellement quatre approches de diffusion des pratiques efficaces en alphabétisation. Chaque projet a un volet passif – où les ressources sont envoyées aux écoles – et un volet actif – où ces ressources sont accompagnées d'actions de mobilisation (conférences, ateliers etc.). L'idée est de vérifier si les actions de mobilisations ont une plus-value par rapport à la simple diffusion de ressources.

Perspectives

- Ayant été très focalisé sur la production et la synthèse de la connaissance depuis 2011, EEF se consacrera davantage à la diffusion des connaissances et d'adoption des pratiques efficaces sur les cinq années à venir. Le centre produit désormais des recommandations plus explicites, au-delà de la mise en avant des résultats des études.
- L'EEF passera de la diffusion générique du *Toolkit* à des actions de diffusion plus ciblées afin de favoriser l'adoption de pratiques efficaces par les professionnels.
- Jusqu'alors l'EEF a expérimenté principalement des interventions visant à améliorer le niveau d'études des élèves, il investira désormais davantage dans l'expérimentation d'interventions visant le développement des compétences sociales et émotionnelles (non-cognitives).

Enseignements

- L'EEF avait une dotation financière importante à dépenser sur 15 ans, permettant d'avoir une action d'envergure qui s'inscrit dans la durée.
- L'EEF a pu aborder la mobilisation de la recherche de manière systémique. En effet, le centre agit sur toute la chaîne de la connaissance – de la synthèse, à l'évaluation, et à la diffusion.
- Le *Teaching & Learning Toolkit* a été au cœur de la réussite de l'EEF. Véritable « guichet unique » vers la recherche en éducation et aide à la décision, il apporte une réponse adaptée aux contraintes des professionnels de terrain. Il a permis aussi dès les premières années de montrer en quoi l'EEF pouvait être utile et de lui donner de la visibilité.
- Mettre à disposition la recherche de manière accessible dans ce *Toolkit* a été nécessaire mais pas suffisant pour favoriser l'adoption des pratiques efficaces par les enseignants. Il faut créer des opportunités et des motivations pour les professionnels d'adopter des interventions efficaces.
- L'évaluateur et le porteur de projet doivent échanger dès le cadrage des expérimentations, pour que le projet intègre les exigences d'une évaluation scientifique.

- Plus globalement, il a fallu accompagner une montée en compétences sur les méthodes d'évaluation expérimentales auprès des porteurs de projet et des chercheurs.
- EEF est en lien avec le EIPPEE²⁵, un réseau européen qui intervient pour favoriser l'utilisation de la recherche en éducation.

What Works Centre for Local Economic Growth

Fiche identité	
Périmètre	Développement économique local
Date de création	Octobre 2013
Statut	Partenariat
Taille de l'équipe	Environ 10 équivalents temps plein
Budget annuel	1 million de livres sterling par an 2013-2016
Site web	www.whatworksgrowth.org

Historique

Le What Works Centre for Economic Growth a été créé en octobre 2013 par un groupement porté par la London School of Economics. Entre 2013 et 2016, le centre bénéficie d'un financement initial de 3,1 millions de livres sterling de la part de trois financeurs - l'*Economic and Social Research Council*, le *Department for Communities and Local Government (DCLG)*, et le *Department for Business (BIS)*.

L'objectif du centre est d'apporter des solutions pertinentes aux décideurs politiques locaux et nationaux en :

- réalisant des revues systématiques de la recherche scientifique sur l'impact des politiques de développement économique local ;
- sensibilisant les collectivités et leurs partenaires locaux, les administrations centrales et les entreprises afin qu'ils comprennent cette recherche et qu'ils s'en emparent davantage dans la conception et la mise en œuvre des politiques ;
- améliorant la qualité de la recherche au Royaume-Uni par le lancement de projets pilotes et d'expérimentations locales.

Le terme développement économique local est relativement flou, n'étant ni un domaine de la recherche ni une politique publique à part entière.

L'équipe du centre a donc décidé de définir un périmètre relativement étroit : les programmes qui, directement ou indirectement, améliorent la productivité locale et la situation d'emploi (création d'emploi, retour à l'emploi).

Organisation

Le centre est porté par un groupement de trois structures :

- **London School of Economics (LSE).** Le centre est piloté par le Professor Henry Overman et son adjoint Doctor Max Nathan qui en assurent la représentation institutionnelle. La LSE est également responsable du cadrage et de la relecture de toutes les revues systématiques et des travaux autour de l'évaluation. Lorsque la LSE est très impliquée dans la réalisation de revues, Dr Nathan mobilise d'autres chercheurs de la LSE.
- **Arup.** Ce bureau d'étude spécialisé entre autres dans l'infrastructure, le transport, l'urbanisme et le développement économique a été chargé de la réalisation des revues systématiques. Deux consultants seniors assurent le pilotage du centre au sein d'Arup et mobilisent de nombreux analystes dans la réalisation des revues.
- **Centre for Cities.** Ce think tank est spécialisé dans l'analyse du développement économique des grandes villes britanniques. Sous l'égide de Meg Kaufman, six personnes au total sont mobilisées à mi-temps sur le What Works Centre, essentiellement sur les actions de communication (site web, documentation, événements).

Le centre est donc atypique car il n'est pas constitué d'une seule organisation mais il est une initiative portée par un groupement de trois organisations. Il est doté d'une équipe « virtuelle » composée de représentants des trois partenaires qui mobilisent à leur tour leurs collègues dans la réalisation des travaux spécifiques (tels qu'une revue). Malgré des risques associés à ce modèle de management, le centre a globalement bien fonctionné grâce à une répartition claire des rôles et des processus de management partagés (réunions mensuelles, calendrier partagé).

What Works Centres britanniques : quels enseignements pour les politiques fondées sur la preuve en France ?

Le centre s'appuie également sur deux panels :

- Le *User Panel* est composé essentiellement de décideurs publics du développement économique local et est présidé par Tom Bridges, directeur du développement économique à *Leeds City Council*. Les membres du Panel se réunissent quatre fois par an. Ce panel aide à définir les priorités du centre en matière de recherche et la meilleure manière de communiquer auprès des décideurs. Les échanges peuvent par exemple se centrer sur les meilleurs formats pour la diffusion des enseignements ou la structure du site web. Certains membres du panel sont également sollicités de manière *ad hoc* sur des sujets spécifiques. Leurs retours, jugés « très utiles » par l'équipe, aident le centre à rendre leurs travaux accessibles et utiles pour leur public cible ;
- L'*Academic Panel*, composé de chercheurs, revoit les travaux pilotés par le centre afin qu'ils soient le plus scientifique possible. Ce panel est composé d'environ 10 chercheurs de la *London School of Economics* et d'autres universités. Le panel est « virtuel » puisqu'aucune réunion collective n'est tenue entre ses membres. Leur fonctionnement consiste à soutenir les travaux de recherche du Centre. Jusqu'alors, il s'agit principalement d'un soutien dans l'élaboration des revues systématiques et les *toolkits*. Ainsi, un à deux membres du Panel sont assignés à chaque revue pour apporter des conseils sur plusieurs aspects de la recherche, les aidant à s'orienter dans la littérature, à modéliser leur stratégie de recherche et à élaborer des conclusions. Les chercheurs sont rémunérés pour le temps passé (1 à 5 jours par revue environ), soit directement pour les chercheurs du LSE, soit via leur institution pour les chercheurs d'autres universités. Selon l'équipe du LEG, les membres du panel ont joué un rôle « crucial » dans les revues. Le centre est financé en grande partie par des ministères, qui commentent ses livrables avant publication. Cette proximité de l'État rend le centre potentiellement moins indépendant, mais elle est considérée comme un atout par le centre, notamment pour sa visibilité et son impact.

“Comme l'État nous finance, ils doivent nous écouter. En plus, avoir un logo d'un ministère sur vos publications ajoute de la crédibilité pour les lecteurs.”

Meg Kaufman, coordinatrice du What Works Centres au Centre for Cities.

Action

Synthétiser

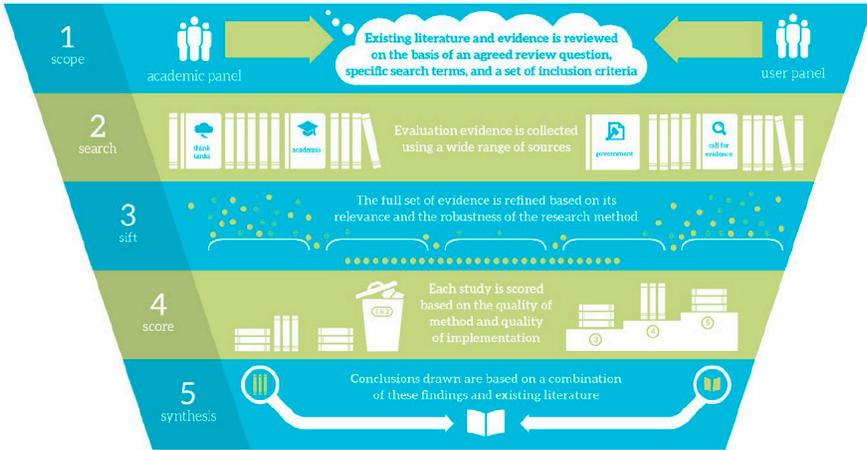
Lors de ces trois premières années, le centre a réalisé 11 revues systématiques (« *evidence reviews* »). Chaque revue portait sur un aspect des politiques de développement économique local, tel que la formation professionnelle, le renouvellement des quartiers sensibles et les actions de sport et de culture. Les thèmes étaient définis initialement dans l'appel d'offre de formation du centre ; les thèmes suivants ont ensuite été établis par l'équipe de direction et les financeurs. Ces revues sont réalisées selon la même méthode :

- cadrer le périmètre de la revue en lien avec le *User Panel* afin d'établir des critères pour la sélection des évaluations ;
- rechercher toutes les évaluations d'impact réalisées dans ce périmètre dans les pays de l'OCDE et publiées en anglais ;
- sélectionner celles qui remplissent les standards méthodologiques du centre ;
- évaluer la qualité des méthodes d'évaluation et la clarté des modalités de mise en œuvre du projet ;
- synthétiser les enseignements de ces évaluations ainsi que la recherche plus largement afin de tirer des conclusions.

Le centre ne traite que les évaluations d'impact réalisées selon des méthodes expérimentales. En effet, seules les évaluations ayant constitué un contrefactuel crédible sont analysées dans le cadre de leurs revues. Le centre utilise le *Maryland Scientific Scale* afin de sélectionner les évaluations, qui doivent correspondre au niveau trois à cinq de cette échelle (voir page suivante). Enfin cette analyse est restituée dans un rapport technique de trente pages environ, accompagné de ses annexes et d'une courte synthèse. Ces rapports sont tous disponibles sur le site internet du centre.

What Works Centres britanniques : **quels enseignements pour les politiques fondées sur la preuve en France ?**

Méthodologie des revues systématique



Source : www.whatworksgrowth.org.

What Works Centre for Local Economic Growth

ÉCHELLE SCIENTIFIQUE DE MARYLAND

Niveau 1. **Corrélation des résultats avec la présence ou l'intensité de l'intervention, sans volonté d'établir une situation contrefactuelle.**

Niveau 2. **Comparaison entre les résultats pour le groupe test après une intervention avec les résultats pour ce groupe avant l'intervention (« évaluation avant – après »).**

Niveau 3. **Comparaison entre les résultats pour un groupe de traitement après une intervention avec la situation du groupe avant l'intervention et un groupe de comparaison pour créer une situation contrefactuelle (« difference in différence »).**

Niveau 4 : **Comparaison entre les résultats du groupe test après une intervention avec la situation de ce groupe avant l'intervention et un groupe témoin pour créer une situation contrefactuelle plus robuste que le niveau 3 (« difference in différence »).**

Niveau 5 : **Les évaluations intégrant l'assignation aléatoire aux groupes**

tests et témoins (« randomisation »).

Traduction de l'Ansa

Les 11 synthèses ont toutes fait émerger des enseignements intéressants. L'équipe de recherche de la LSE a également remarqué une tendance pour les études concernant le sport et la culture, l'aménagement des espaces publics et le renouvellement urbain : une bonne partie des programmes évalués ne démontre pas d'impacts positifs mesurables sur le plan économique mais peuvent avoir d'autres effets intéressants (qualité de vie, qualité de l'environnement).

“On a trouvé que beaucoup d'arguments économiques en faveur de ces programmes ne tiennent pas en termes de retours. Cependant, il y a plein d'autres raisons de faire ces programmes.”

Max Nathan, Directeur adjoint, What Works Centre for Local Economic Growth (LSE).

Des événements sportifs et culturels majeurs : Les effets économiques mesurables des projets sportifs et culturels sur l'économie locale sont rarement importants et sont le plus souvent nuls. Les effets des projets en termes de salaires sont généralement faibles et limités soit aux environs du lieu de l'événement, soit à une catégorie spécifique de travailleurs. Étant donné le coût important des événements sportifs et culturels, ils sont rarement un investissement efficace en termes d'impact sur la croissance économique locale. Cependant, ils ont une valeur sociale et peuvent avoir des effets sur le bien-être des habitants par exemple.

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

Dans la réalisation de ces revues, le centre a fait face au faible nombre d'évaluations d'impact scientifiques dans le domaine du développement économique local. Ce constat tient à deux facteurs. D'une part, les actions visant une zone géographique, telles que les zones d'entreprises, le renouvellement urbain ou la construction d'infrastructures, ne se prêtent

What Works Centres britanniques : **quels enseignements pour les politiques fondées sur la preuve en France ?**

pas aux méthodes d'évaluation les plus scientifiques. Par exemple, seulement 21 des 1050 évaluations recensées portant sur les actions de renouvellement urbain répondent à leurs exigences minimales. Le centre réfléchit donc aux alternatives à la « randomisation », comme les études longitudinales. D'autre part, même lorsque les actions s'y prêtent mieux (accompagnement individuel de chômeurs ou d'entreprises), très peu d'évaluations de ce type ont été réalisées.

“Certains types de dispositifs ne se prêtent pas aux méthodes d'évaluations les plus rigoureuses. Pour les actions d'activation des chômeurs, de conseils aux entreprises, d'incubation des start-up, il est possible de randomiser. Mais on ne peut pas randomiser l'emplacement des logements et de l'infrastructure... Il y a des domaines du développement économique local qu'on ne peut pas évaluer par ces méthodes, même si d'autres méthodes peuvent convenir dans ces cas, comme l'appariement statistique.”

Max Nathan, directeur adjoint, What Works Centre for Local Economic Growth (LSE).

L'équipe de recherche du centre a également constaté que les évaluations disponibles ne sont généralement pas suffisamment attentives à la description de l'action même. Ce manque rend donc difficile l'essai d'actions « efficaces » puisque le centre dispose souvent de peu d'éléments sur leurs modalités et conditions de réussite.

“Des chercheurs sont obsédés par l'évaluation rigoureuse de l'impact, souvent à l'exclusion de tout le reste ! Quand on est décideurs politiques ou professionnel, on a besoin de savoir comment le programme marche dans les moindres détails.”

Max Nathan, directeur adjoint, What Works Centre for Local Economic Growth (LSE).

Face à ces constats, l'équipe de recherche a été contrainte de faire évoluer ses méthodes de revue. Initialement, le centre avait l'intention de réaliser des méta-analyses classiques en calculant la taille moyenne des effets dans les études. Cette approche est appropriée lorsqu'on dispose

de nombreuses évaluations avec un contrefactuel, mais était impossible pour les revues du centre qui se fondaient sur peu d'évaluations. Le centre a donc décidé de dégager des « tendances » globales des évaluations : ont-elles montré des effets globalement positifs, neutres, négatifs, ou mixtes. De plus, le centre analyse si ces effets varient en fonction du type d'action, à l'image de leur revue sur la formation professionnelle.

Formation professionnelle : Pour des compétences de bases ou des interventions visant l'employabilité globale des personnes, des programmes plus courts (moins de six mois, voire moins de quatre mois) ont des effets plus importants et plus prononcés sur l'emploi des bénéficiaires que des programmes plus longs. Ces programmes courts sont également moins coûteux, ce qui suggère une plus grande efficacité que les programmes longs.

► whatworksgrowth.org/policy-area/employment-training.

Expérimenter

Le centre vise à lancer des projets pilotes et d'expérimentations locales afin de combler le manque d'évaluations d'impact scientifiques souligné par leurs revues systématiques. Cependant, il ne dispose pas des ressources nécessaires pour financer ces démarches directement. Pour stimuler le lancement de ces évaluations, le centre a agi sur deux plans.

D'une part, l'équipe de la LSE propose un accompagnement individualisé au cadrage et au lancement d'expérimentations dans son domaine, notamment auprès des collectivités territoriales. En effet, le centre a un appel à projets permanent. En théorie, leurs conseils devraient porter essentiellement sur la manière de modéliser une action pour qu'elle soit « évaluable » selon les méthodes scientifiques. En pratique, l'équipe commence souvent par un travail de sensibilisation des agents et des élus locaux aux méthodes d'évaluation d'impact. À ce temps s'ajoute le temps de mobilisation des financements et de la conception de l'expérimentation.

What Works Centres britanniques : quels enseignements pour les politiques fondées sur la preuve en France ?

Cette préparation du lancement d'une expérimentation est souvent longue et parfois sans débouché. Ainsi, bien que l'équipe de la LSE soit impliquée dans la conception de 20 projets pilotes, seulement 4 sont sur le point d'être lancés.

D'autre part, le centre mène une action de sensibilisation aux méthodes d'évaluation d'impact. Avec son guide « *How to evaluate* », le centre vise à aider les agents publics à s'approprier ces méthodes. Le contenu de ce guide est également pleinement intégré dans leur programme d'ateliers locaux. Enfin, le centre tente de « relever » le niveau des évaluations financées par les administrations centrales.

“On travaille avec les ministères pour améliorer la qualité des évaluations qu'ils financent. Aujourd'hui, certains appels d'offre de l'État exigent a minima une évaluation correspondant au niveau trois de l'Échelle Scientifique de Maryland, et demande aussi une évaluation de processus.”

Max Nathan, directeur adjoint, What Works Centre for Local Economic Growth (LSE)

Diffuser et accompagner l'adoption

Le centre vise à sensibiliser les parties prenantes du développement économique local aux enseignements de la recherche afin qu'elles s'en emparent davantage dans la conception et la mise en œuvre des politiques. Bien que les actions de développement local ciblent au final les entreprises, le principal public pour leurs actions de diffusion est les agents des administrations centrales et des collectivités locales. Pour sensibiliser ce public, le centre a procédé en trois étapes.

Dans un premier temps, l'équipe a tenté de rendre les synthèses les plus accessibles possibles pour un public professionnel, notamment avec l'élaboration de recto-verso (« *Executive Summary* »). En effet, pour chacune des trois premières synthèses, les trois partenaires du centre ont fait une réunion d'échanges de six heures environ afin de vulgariser

les enseignements de l'étude pour les décideurs politiques. Ce travail a permis de clarifier le bon format et le niveau de langage que les équipes de recherche ont cherché à reproduire lors des synthèses suivantes. Ainsi, le centre a produit 11 synthèses et le portail du WWC forme un « guichet unique » vers les évaluations d'impact dans le domaine du développement économique local.

“Il y avait beaucoup de travail avec les chercheurs pour que leurs recommandations soient plus explicites, et pour arrêter de ménager la chèvre et le chou. Il existe une tension créatrice entre être scientifique et avoir un message clair pour les professionnels. »

Meg Kaufman, coordinatrice du What Works Centres au Centre for Cities.

Dans un deuxième temps, le centre a produit plusieurs boîtes à outils pour faciliter l'adoption des dispositifs ayant eu des effets positifs mesurables. Chaque boîte à outils porte sur le périmètre d'une synthèse, telle que la formation professionnelle, et réunit plusieurs guides qui traitent des interventions plus en détail. Partant des évaluations recueillies dans le cadre de la revue systématique, ces guides apportent des conseils plus précis sur les modalités des interventions et leurs conditions de réussite. Par exemple, la synthèse sur la formation professionnelle a été transformée en boîte à outils composée de quatre guides : l'orientation professionnelle, les incitations financières, les pré-qualifications et les rappels (pour améliorer la participation aux formations).

“Les toolkits ne sont pas de simples guides séquentiels qui expliquent comment élaborer un plan local de croissance économique. Certes nous faisons des recommandations pour apporter des améliorations, mais on ne veut pas que les gens croient que c'est le dernier mot sur ce qui est “efficace”.”

Meg Kaufman, coordinatrice du What Works Centres au Centre for Cities.

What Works Centres britanniques : **quels enseignements pour les politiques fondées sur la preuve en France ?**

Toolkit de formation professionnelle : présentation des guides

Policy Design Tool	What does it aim to do?	How secure is the evidence?	How much does it cost?	How effective is it?
Careers Counselling helps individuals choose appropriate training	Improve take-up and completion of training		£££	
Financial incentives are payments offered before, during or after training	Improve take-up and completion of training		£££	
Pre-qualification courses are a pathway to further education or training	Improve performance and completion of training		£££	
Reminders provide people with information about their training by text or email	Improve attendance to training programmes		£££	

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

What Works Centre for Local Economic Growth

Par ailleurs, depuis sa création, le centre a investi dans des actions de diffusion et d'appropriation de ses livrables. Il s'agit de l'envoi des synthèses à travers sa newsletter et son compte twitter – permettant ainsi de toucher plusieurs milliers de personnes. Le centre a également organisé des événements nationaux et locaux de restitution afin de favoriser l'appropriation de ses livrables par les professionnels.

Ces événements ont été l'occasion de confronter les travaux à des professionnels encore méfiants de la « randomisation ». Certains de ces événements ont été l'occasion de débats animés sur ces questions et ont globalement favorisé l'adhésion des professionnels à la démarche. Conscient que cette appropriation est souvent un processus social et local, le centre a, depuis plus d'un an, organisé de nombreux ateliers locaux sur mesure.

“Il ne suffit pas de créer une liste de diffusion et d'envoyer les synthèses. Leur format est inhabituel, leur rigueur méthodologique était un peu austère. On avait besoin de temps pour expliquer aux gens que les évaluations randomisées ne sont pas immorales. Pour ce faire, il faut

introduire le sujet, et c'est mieux de faire ça en personne. Ce sont les événements qui ont eu le plus de débouchés et ont mené le plus souvent à l'adoption de pratiques."

Meg Kaufman, coordinatrice du What Works Centres au Centre for Cities.

Perspectives

- Les perspectives pour le centre seront plus claires suite au renouvellement de ses financements de base prévu début 2017.
- À ce stade, l'équipe considère que le centre se consacrera moins à la synthèse de la recherche pour plus se concentrer sur des actions de diffusion et d'expérimentation. En effet, le centre se concentrera probablement sur l'adoption des pratiques efficaces par le biais de la diffusion des boîtes à outils et sur la promotion de l'évaluation en vue du lancement des expérimentations.

Enseignements

- **Le centre agit sur un domaine qui est relativement « pauvre » en évaluation d'impact scientifique.** Le centre a toutefois décidé de rester très exigeant sur la rigueur des évaluations qui seront prises en compte dans ses revues. Par conséquent, ces dernières se fondent sur un nombre relativement limité d'études, ce qui affaiblit la capacité du centre à tirer des enseignements globaux.
- **Dans la synthèse de la preuve, le centre a fait le constat d'une tension créatrice entre rigueur scientifique et clarté du propos.** À cet égard, les échanges entre chercheurs et professionnels de la communication ont été clés pour produire des synthèses à la fois accessibles pour les professionnels et rigoureuses sur le plan scientifique.
- **Le centre avait une mission de lancement de projets pilotes et d'expérimentations locales, sans être doté des ressources pour les financer.** Le centre appuie le cadrage de potentielles expérimentations, un processus long et souvent sans débouché. En effet, aucune

What Works Centres britanniques : **quels enseignements pour les politiques fondées sur la preuve en France ?**

expérimentation n'a été lancée à ce stade avec son appui, bien que quatre soient actuellement à l'étude.

- **Pour réussir ses missions, le centre doit accompagner un changement culturel au sein des acteurs publics, processus long et complexe.** En effet, il a fallu trois ans pour faire la synthèse de la recherche. Désormais, le centre se consacre à la tâche vaste de promouvoir la réalisation d'évaluations scientifiques et l'application de la recherche aux politiques publiques. Ceci nécessite un travail lourd de sensibilisation, de formation et parfois d'accompagnement individuel des agents publics.
- **Le centre bénéficie d'un public cible relativement petit en nombre mais qui n'est pas réuni dans un organisme professionnel.** Le centre joue dans un certain sens ce rôle de constitution d'un réseau de professionnels du développement économique local.

Centre for Ageing Better

Fiche identité	
Périmètre	Le bien-vieillir
Date de création	2015
Statut	Charitable foundation
Taille de l'équipe	20 salariés
Budget annuel	5 millions de livres sterling
Site web	www.ageing-better.org.uk

Historique

Le vieillissement de la population a mis la question de la qualité de vie des personnes âgées au cœur du débat public dans les pays occidentaux.

Cet enjeu a été identifié par le *Big Lottery Fund* (BLF) comme une priorité dans ses actions. Le BLF est un organisme public en charge de la distribution des fonds recueillis par la *National Lottery* pour les causes d'intérêt général, et finance des projets, associations, institutions. Il finance depuis longtemps un ensemble d'actions pour les seniors à travers ses différents programmes et commissionne des études pour identifier les besoins des personnes, les projets qui ont tendance à fonctionner et en déduire quels projets financer dans le futur. Parmi les programmes récents, un appel à projet pour l'essaimage ou le changement d'échelle d'initiatives portant sur les soins à domicile et la démence et dont l'impact avait été démontré et un appel à projet d'innovations « *Fulfilling Lives Ageing Better* » ont été lancés en 2014 et 2015.

D'autre part, en 2013, le comité de la Chambre des Lords travaillant

What Works Centres britanniques : quels enseignements pour les politiques fondées sur la preuve en France ?

sur le service public et les changements démographiques a préconisé la création d'un What Works Centre dans son rapport « *Ready for Ageing?* » pour maintenir la dynamique, dans un contexte où de bonnes initiatives émergent puis disparaissent, faute d'attention continue^{xciv}.

Dans les échanges qui ont suivis, le *Big Lottery Fund* a reconnu le besoin d'investir sur la production et la diffusion de résultats scientifiques (en particulier des évaluations d'impact) dans ce domaine et a choisi de financer le *Centre for Ageing Better* pour 10 ans via la création d'un fond, en 2015. Le financement du *Centre for Ageing Better*, de 50 millions de livres sterling, a été annoncé simultanément au financement de 14 programmes locaux de « *Fulfilling Lives Ageing Better* » qui ont reçu un total de 80 millions de livres sterling.

Organisation et objectifs

Le *Centre for Ageing Better* (CfAB) est une fondation indépendante, avec une dotation du *Big Lottery Fund* à dépenser en 10 ans. Le centre est gouverné par un conseil d'administration dont les membres travaillent ou ont travaillé dans des domaines de la recherche, des politiques publiques, associatifs, de la société civile, en lien avec le vieillissement et la perte d'autonomie.

Le président, Lord Filkin CBE, est membre de la Chambre des Lords et a présidé le rapport *Ready for Ageing*. *Centre for Ageing Better* a pour objectifs de :

- développer, partager et mettre en œuvre les résultats de la recherche et d'évaluations d'impact pour aider les personnes à mieux vieillir;
- soutenir l'innovation face aux défis et aux opportunités du vieillissement de la population;
- contribuer aux changements sociétaux (aux niveaux national et local) nécessaires à une meilleure qualité de vie pour les personnes âgées.

Le centre, aujourd'hui constitué d'une vingtaine de personnes, doit s'élargir jusqu'à environ 25-30 personnes en régime stable. Ses équipes se composent en particulier de :

- chargés de mission responsables des revues scientifiques et du suivi des

programmes : six personnes venant du monde universitaire : spécialistes des sciences sociales, de la recherche qualitative, pour la plupart connaissant les enjeux de la pauvreté, du bien-être, des sciences comportementales, de la santé et de la recherche de la santé (par exemple gestion individuelle des douleurs chroniques, utilisation de l'art-thérapie dans un contexte clinique)²⁶ ;

- responsables de la communication et de la relation avec les parties prenantes : cinq personnes ;
- chargés du soutien à l'innovation et de l'implémentation des nouvelles pratiques : quatre personnes, avec des expériences dans la société civile, chez des bailleurs sociaux, dans la participation des habitants ou encore avec des prestataires dans le domaine des services publics ;
- une équipe ressource (finance et gestion administrative, logistique, etc.).

Le centre met en place des panels consultatifs composés d'experts du monde universitaire, professionnel et de personnes concernées selon les sujets des travaux menés.

Les partenaires clés de ce centre se composent de collectivités (en premier lieu, le *Greater Manchester*) pour mettre en œuvre des projets et les conseiller sur leur stratégie, de ministères et secrétaires d'État auxquels le centre fournit des résultats de recherche, de réseaux d'acteurs ayant pour objet de promouvoir un ou plusieurs aspects du bien-vieillir, d'associations caritatives et d'entreprises. Le CfAB travaille de manière rapprochée avec des universités (par exemple, *Manchester Institute for Collaborative Research on Ageing*). Le centre joue un rôle essentiel de conseil auprès du Big Lottery Fund sur le vieillissement, en lien avec les programmes de financement qui y sont alloués. Le centre collabore enfin avec *Public Health England* sur des campagnes ciblées.

Le CfAB a tâché d'inscrire son action en complémentarité avec les missions des acteurs déjà actifs dans le domaine du bien-vieillir, et notamment NICE et *Public Health England*.

²⁶ Les compétences recherchées étaient dans la recherche appliquée, surtout en sciences sociales, les synthèses, la « traduction » de connaissances et l'évaluation, plutôt que la recherche primaire ou la santé publique.

Qui traite de quoi au Royaume-Uni

NICE produit des revues systématiques et des recommandations concernant les aides techniques correspondant à des situations de handicap (par exemple les aides auditives). Ces recommandations sont remises à jour régulièrement. NICE développe aussi des synthèses et recommandations concernant la santé mentale et les situations de handicap, tant au niveau de la prévention que de l'accompagnement. Par exemple, NICE a produit un jeu de recommandations pour des actions en « milieu de vie » afin de prévenir la fragilité, le handicap et la démence^{xcv}.

Public Health England (PHE) conduit des programmes en lien avec le vieillissement de la population : ces programmes ont pour objectifs de promouvoir des actions d'acteurs de la santé ou d'autres acteurs concernés, de contribuer à l'appropriation des résultats scientifiques et à la mise en œuvre d'actions pertinentes par un certain nombre d'institutions et enfin de contribuer aux campagnes menées par *National Health Services* (NHS). Ces programmes s'appuient sur les recherches scientifiques existantes, sur des commandes de travaux auprès d'universitaires et sur une collaboration avec NICE. Parmi les sujets étudiés, on trouve la démence, le vieillissement en bonne santé, la prévention des chutes, le soutien aux aidants. PHE soutient aussi la mise en œuvre de projets pilotes. PHE met en œuvre dans certains de ces domaines une démarche de « *consensus statement* ». Elle consiste à rassembler un ensemble d'acteurs pertinents en lien avec le domaine étudié, qu'il s'agisse de professionnels de terrain ou de chercheurs par exemple, pour étudier les connaissances scientifiques disponibles et se mettre d'accord sur les actions à mettre en œuvre, en particulier pour les institutions présentes. En 2014 s'est tenu le *Blackfriars Consensus* sur la prévention de la démence^{xcvi}, soutenue par 60 organisations et experts dans la prévention de la démence et des maladies non-transmissibles.

Action

Analyser le besoin

Il n'existe pas une seule définition du « bien-vieillir ». CfAB a donc décidé de conduire des analyses permettant de donner corps à cette notion. Dans ce cadre, il a conçu le projet *Later Life* in 2015 avec Ipsos MORI. Ce projet avait plusieurs volets qualitatifs et quantitatifs, détaillés ci-dessous. Ce projet a permis de définir de manière collaborative des axes de travail pour le centre.

English longitudinal study of ageing (ELSA)

Le CfAB s'est appuyé sur l'étude longitudinale²⁷ d'une cohorte pour étudier le vieillissement^{xcvii} : cette étude a été mise en place il y a 14 ans et comprenait à son début plus de 10 000 hommes et femmes âgée de 50 à 100 ans. La cohorte a été interrogée tous les deux ans pendant 12 ans (les dernières données ont été recueillies en 2015). L'étude est très large et inclut des éléments biométriques, sociaux-économiques, sur le bien-être, etc. Elle a surtout été utilisée par le CfAB pour définir les segments de population âgée de l'étude *Later Life*.

Later Life in 2015^{xcviii}

Le CfAB a élaboré avec IPSOS MORI une étude sociale de grande envergure, explorant ce que signifie le bien-être pour les personnes âgées à partir de leurs propres réponses.

L'étude longitudinale citée plus haut a été utilisée dans la vague la plus récente, pour segmenter la population âgée en 6 groupes de personnes ayant une vision et un niveau de bien-être homogènes. Des études de cas approfondies ont été menées avec une vingtaine de personnes pour comprendre leur vision de la qualité de la vie pour les seniors,

²⁷ L'étude longitudinale permet aussi d'identifier des corrélations et donc de mettre en avant des relations de causes à effets à tester – par exemple, on a pu identifier que les personnes âgées qui faisaient du bénévolat avaient une probabilité plus élevée d'être en bonne santé, d'avoir une vie sociale active et de ressentir du bien-être, sans avoir déterminé dans quel sens était la causalité. Le CfAB commande aussi des études de recherche primaire sur le retour à l'emploi, l'engagement local, le numérique.

What Works Centres britanniques : **quels enseignements pour les politiques fondées sur la preuve en France ?**

leur situation selon des critères allant de la santé à leur environnement de vie (logement, quartier, etc.) et pour qu'ils puissent déterminer les actions prioritaires pour améliorer la situation selon leur point de vue. Enfin, une enquête complémentaire a été conduite auprès de plus de 1 000 personnes représentatives pour confirmer des aspects identifiés dans l'enquête qualitative.

Later life in 2015: An analysis of the views and experiences of people aged 50 and over

What did we do?

Below is a brief summary of what the research entailed



Cluster analysis to identify segments of people aged 50 years and older with similar levels of wellbeing



In depth qualitative work with 24 people who matched the profiles of these segments



Diary work, to monitor wellbeing over a week



Return visits to 12 of the participants from stage 2, plus photography work



A quantitative survey of c.1400 people aged 50 years and older

Source : laterlife.ageing-better.org.uk.

Les axes de travail

Les axes de travail du CfAB ont été déterminés à partir de l'étude *Later life in 2015*, de consultations auprès d'un ensemble d'acteurs et en prenant en compte des résultats existants de recherche et des évaluations d'impact sur des interventions.

CfAB travaille aujourd'hui sur sept thématiques représentées sous la forme d'affirmations d'objectifs par les personnes vieillissantes, en se concentrant sur certains aspects de ces questions en premier lieu :

- *j'ai un travail épanouissant* – avec des programmes pour que l'emploi et l'environnement de travail soient adaptés et des programmes pour aider les personnes de plus de 50 ans à retrouver un emploi;
- *je vis dans un logement qui me convient* – avec un programme en cours concernant l'adaptation du logement au vieillissement;
- *je contribue à la société* – soutenir les personnes qui ont moins tendance à participer à la vie de la société;
- *je peux faire face à de grands changements dans ma vie* – apporter un soutien pour le passage à la retraite;
- *Je suis actif physiquement* – aider les personnes à maintenir ou à renforcer leur musculature et leur équilibre;
- *je vis dans un environnement favorable aux personnes âgées*;
- *je suis actif et connecté* – aider les personnes à rester ou devenir impliqué dans les technologies numériques.

Synthétiser

En 2016, le CfAB a commencé un programme de revues des évaluations d'interventions visant à améliorer la qualité de vie des personnes âgées. Il n'y a pas encore de synthèse de revues systématiques d'évaluations disponible. La méthodologie est en cours d'élaboration et sera adaptée en fonction du sujet et de la littérature disponible.

Exemples de revues systématiques commandées ou de synthèse de travaux existants

- revue systématique sur l'adaptation du logement pour les seniors^{xcix};
- synthèse sur ce que l'on sait de l'impact d'une contribution à la société par les seniors leur apportant satisfaction sur leur bien-être et leurs connexions sociales^c;
- revue systématique des évaluations sur les dispositifs permettant aux personnes au-delà de 50 ans de continuer à s'impliquer dans un travail épanouissant.

CfAB apporte son soutien à une revue portée par PHE sur les analyses de coût/efficacité des interventions relatives au renforcement musculaire

What Works Centres britanniques : quels enseignements pour les politiques fondées sur la preuve en France ?

et à l'équilibre des personnes âgées. CfAB devrait aussi lancer un appel à bonnes « pratiques » qui auraient peu de chance de se retrouver dans la littérature scientifique existante. Cette approche témoigne de l'ouverture du centre à des projets prometteurs mais n'ayant pas fait l'objet d'une évaluation d'impact.

Quels résultats

Ce programme de revue systématique a vocation à identifier des dispositifs démontrés comme efficaces ou prometteurs sur au moins un des axes d'amélioration souhaités (par exemple permettre d'être actif plus longtemps dans un travail épanouissant). Cet effet devra être arbitrée avec de potentiels effets négatifs sur les autres axes d'amélioration (par exemple : une intervention qui donne lieu à une amélioration de la sécurité financière mais a un effet négatif fort sur la santé mentale ne paraît pas très pertinente).

La qualité des évaluations sera un enjeu dans ce secteur pour le centre : les résultats espérés étant assez ouverts, il peut être moins adapté ou plus compliqué de mettre en œuvre des études randomisées dans ce contexte. De plus, selon les critères de qualité des évaluations du centre, certaines interventions testées à petite échelle ne permettent pas non plus de savoir si elles sont généralisables. Les évaluations portent enfin souvent sur une échelle de temps trop court, par exemple les bénéficiaires d'interventions de préparation à la retraite sont généralement suivis seulement pendant un à deux ans.

Les incitations

Il existe déjà un corpus important de recherches sur des interventions qui ont un impact positif sur les personnes (avoir une activité physique régulière a un effet positif déjà bien démontré sur le bien-être). L'un des enjeux pour le centre est alors d'identifier et de réunir des résultats scientifiques **d'évaluation de dispositifs qui vont inciter les personnes âgées à adapter leur comportement** pour ensuite proposer aux fournisseurs de service des collectivités les meilleures interventions.

“Notre question principale est de comprendre ce qui marche pour que plus de personnes puissent bénéficier des interventions.”

Dan Jones – Director of Innovation and Change, CfAB.

Certaines incitations étudiées portent sur les actions à mener avant (voire bien avant) l'âge de la cinquantaine, pour prévenir les difficultés identifiées par les cohortes actuellement étudiées au Royaume-Uni. Par exemple :

- commencer à épargner dès son premier emploi pour avoir assez d'épargne à l'âge de la retraite ;
- faire du sport, ne pas fumer etc. pour avoir une meilleure qualité de vie plus tard.

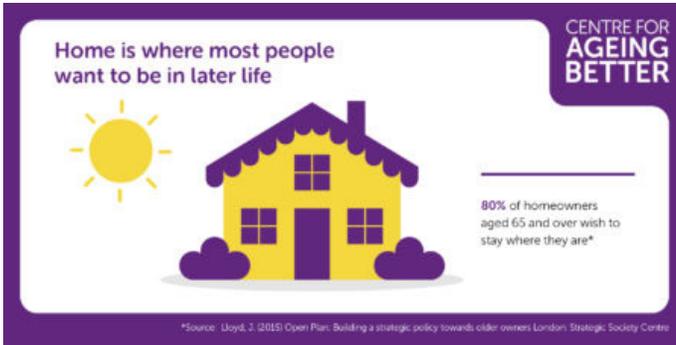
Des pistes de travail grâce aux corrélations

Enfin, il y a un certain nombre de corrélations qui donnent des pistes intéressantes d'interventions, mais dont l'effet reste à démontrer. Par exemple il y a une forte corrélation entre le bénévolat et le bien-être pour les personnes à la retraite, et l'on retrouve ces variations lorsque les personnes arrêtent et/ou reprennent une activité de bénévolat. Mais il reste à démontrer si des programmes favorisant le bénévolat auront effectivement un effet sur le bien-être des personnes.

Le logement

Une revue systématique de toute la littérature disponible en anglais a été commandée sur l'adaptation du logement pour les personnes âgées vivant à domicile ; elle sera publiée début 2017. Les questions posées par la revue systématique ont été définies avec un ensemble de parties prenantes, y compris des professionnels, des personnes concernées, pour que les réponses identifiées soient pertinentes et utilisables. Le CfAB ne se concentre pas à ce stade sur les modalités de logements alternatifs, comme les logements intergénérationnels, etc. notamment car 90% des personnes de plus de 60 ans sont propriétaires occupants en Angleterre.

What Works Centres britanniques : **quels enseignements pour les politiques fondées sur la preuve en France ?**



Source : www.ageing-better.org.uk.

Expérimenter

Dans l'approche du CfAB et dans le contexte d'un manque d'évaluations d'impact, l'expérimentation permet d'identifier des dispositifs ayant un impact positif et pouvant être essaimés à plus grande échelle. Le CfAB a vocation à :

- identifier les champs où des évaluations sont requises ;
- accompagner méthodologiquement la mise en œuvre et le suivi de programmes pilotes ;
- répondre à des appels à projet académiques pour financer l'évaluation d'impact des interventions prometteuses, en apportant son appui méthodologique tout au long de l'expérimentation.

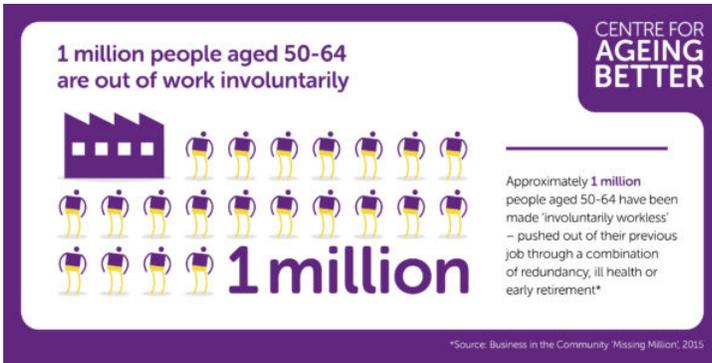
À ce stade, le CfAB ne porte pas d'expérimentation avec évaluation d'impact mais accompagne déjà l'identification de champs qui manquent d'interventions évaluées et la mise en œuvre de programmes pilotes.

Dans la plupart des cas, le centre cherche des cofinancements pour les projets afin de s'assurer de l'engagement d'acteurs qui pourront ensuite les reprendre. Si l'acteur portant le projet était en capacité de financer l'ensemble du pilote, CfAB apporterait une assistance technique

dans la mise en œuvre et/ou l'évaluation et pourrait financer la partie « évaluation » du projet. Les évaluations seront probablement menées par des universités ou des bureaux d'étude qualifiés. CfAB pense aussi mettre en œuvre d'autres approches pour promouvoir l'innovation, par exemple un concours dans la conception d'adaptations au logement faciles pour les utilisateurs. CfAB teste donc un certain nombre d'approches permettant de développer l'innovation et d'accroître les connaissances scientifiques dans son domaine.

Exemple de projets pilotes

Il n'y a pas suffisamment d'évaluations d'impact sur les dispositifs de retour au travail. Le CfAB met donc en œuvre des programmes pilotes et des évaluations de dispositifs innovants, prometteurs et non encore testés. Un partenariat avec le grand Manchester permet la mise en œuvre de programmes pilotes pour tester de telles approches, via un appel à projet mené en 2016. Dans ces pilotes, le CfAB finance la conception des interventions et de l'évaluation. Des pilotes à plus grande échelle avec des évaluations externes plus poussées interviendront dans un second temps. Le CfAB espère aussi identifier des projets pilotes sur les transports adaptés aux personnes âgées, la préparation à la retraite et la connexion numérique dans les premiers mois de 2017.



Source : www.ageing-better.org.uk.

La participation dans la définition des interventions

Pour concevoir de nouvelles interventions qui seront évaluées, le CfAB prévoit la collaboration avec des personnes concernées. Des évaluations d'impact (RCT) du champ du développement international ont démontré que l'implication de personnes concernées augmentait la qualité du design de services, dans certains cas de manière significative.

“Pour l'innovation, là où nous cherchons des solutions qui pourraient marcher, nous parions que la qualité des solutions est meilleure lorsqu'elles sont co-construites avec les utilisateurs eux-mêmes.”

Dan Jones – CfAB.

Diffuser et accompagner l'adoption

Les acteurs cibles pour le CfAB sont autant les personnes seniors et âgées elles-mêmes que les professionnels et les décideurs politiques. Les interventions peuvent en effet relever tant du changement de comportement des personnes et de pratiques professionnelles, de la disponibilité de logements et d'équipements adaptés que de l'aménagement du territoire.

Le CfAB ne prévoit pas une dissémination descendante des résultats, notamment sous la forme de catalogues d'interventions, pensant que l'adoption par des acteurs si variés risque d'être assez faible.

“Notre hypothèse de travail est que l'émission d'informations sur un mode descendant n'est pas une méthode efficace pour la diffusion de connaissances.”

Dan Jones – CfAB.

Il y a une volonté de cibler ceux qui décident et mettent en œuvre de manière opérationnelle. Le partenariat avec le grand Manchester est emblématique, avec une approche locale au bien-vieillir : il s'agit d'expérimenter avec eux mais aussi d'intégrer dans les conseils prodigués par le CfAB les connaissances basées sur la preuve. L'agglomération

et le centre collaborent notamment sur l'urbanisme, les espaces publics, les nouveaux logements, le développement économique local, la santé, les formations professionnelles, les infrastructures informelles d'un quartier (par exemple, s'assurer que les magasins acceptent que les personnes âgées utilisent leurs toilettes).

“Cette approche territoriale est sous-tendue par la logique qu'avoir un environnement adapté aux personnes âgées fait partie des changements importants pour augmenter leur qualité de vie.”

Dan Jones – CfAB.

Perspectives

Le *Centre for Ageing Better* a été créé en 2015 : il en est donc au lancement de ses activités, qu'il s'agisse de revues systématiques ou de projets pilotes. Un travail approfondi a été mené sur les besoins des personnes âgées pour une vie de qualité, jetant une base et donnant une direction pour l'action du centre dans les années à venir. L'approche du CfAB diffère de la plupart des *What Works Centres*. En effet, le CfAB met un accent plus important sur la promotion de l'innovation et envisage de considérer un éventail de projets, pas uniquement ceux ayant fait leurs preuves dans le cadre d'une évaluation d'impact. L'approche de partenariat local est également différente de celles des autres *What Works Centres* - le retour d'expérience sur cette approche promet d'être intéressant.

Enseignements pour la France

- La compréhension des besoins des personnes âgées, exprimés par elles-mêmes, observés par des études ethnographiques et par des enquêtes longitudinales permet d'identifier de manière claire et partagée les axes prioritaires pour un sujet au périmètre variable qu'est le bien-vieillir.
- La collaboration avec une ou plusieurs collectivités semble être un bon levier tant pour lancer des projets pilotes que pour soutenir des décideurs

What Works Centres britanniques : quels enseignements pour les politiques fondées sur la preuve en France ?

locaux à fonder leurs dispositifs sur la preuve.

- Enfin, l'approche par l'environnement et l'incitation au changement de comportement sont des axes intéressants pour la prévention de la perte d'autonomie, le maintien au domicile, etc.

xciv Ce rapport présente un ensemble de recommandations pour le gouvernement, les collectivités locales, le système de sécurité sociale, etc. pour faire face au vieillissement de la population. House of Lords. Ready for Ageing?. 14 mars 2013. Disponible sur le web : www.publications.parliament.uk/pa/ld201213/ldselect/ldpublic/140/140.pdf.

xcv Disponible sur le web : www.nice.org.uk/guidance/ng16.

xcvi Disponible sur le web : www.gov.uk/government/news/call-for-new-policy-focus-on-brain-health-to-reduce-the-risk-of-dementia.

xcvii Description disponible sur le web : www.elsa-project.ac.uk.

xcviii Ressources disponibles sur le web : www.ageing-better.org.uk/our-work/late-life-2015.

xcix Cahier des charges de la revue. Ressource disponible sur le web : 16881-presscdn-0-15-pagely.netdna-ssl.com/wp-content/uploads/2016/04/Evidence-review-scope-The-role-of-home-adaptations-in-improving-late-life.pdf.

c Ressource disponible sur le web : www.ageing-better.org.uk/news/people-likely-benefit-volunteering-least-likely-get-involved.

La santé publique au sein du National Institute for Health and Care Excellence (NICE)

Fiche identité	
Périmètre	Évaluation de matériels médicaux, médicaments et directives cliniques, santé publique, protection sociale
Date de création	1999
Statut	Organisme public indépendant
Taille de l'équipe	600 salariés
Budget annuel	/
Site web	www.nice.org.uk

Historique

Les questions de santé publique comme l'alimentation, le comportement sexuel, etc. sont non seulement individuelles mais aussi fortement contextuelles (milieu familial, social, etc.) et peuvent être influencées par de nombreux acteurs, qui peuvent donc jouer un rôle dans les recommandations d'un What Works Centre sur le sujet.

“Les recommandations en santé publique de NICE n'ont pas seulement pour objet d'identifier des interventions « directes » auprès du public mais aussi d'expliquer comment les décideurs locaux et les fournisseurs de services devront travailler ensemble pour mettre en œuvre les stratégies locales recommandées par NICE – comment collaborer, comment s'impliquer, etc.”

Alexia Tonnel, NICE.

What Works Centres britanniques : quels enseignements pour les politiques fondées sur la preuve en France ?

L'enjeu d'avoir des interventions de santé publique dont l'efficacité a été démontrée remonte à la fin des années 90 au Royaume-Uni. Le *Department of Health* avait alors mis en œuvre une organisation, le « *Health Development Agency* » (HDA), pour mener à bien des analyses systématiques de recherches sur le sujet.

À peu près au même moment, NICE est créé par le ministre de la santé britannique avec l'objectif d'évaluer les médicaments et les matériels médicaux, notamment en termes d'efficacité par rapport à leur coût. Son champ d'action s'est élargi progressivement aux directives cliniques et à la protection sociale. Le but était d'apporter ces éléments au National Health Services (NHS) pour diffuser les bonnes pratiques rapidement.

HDA et NICE ont fusionné en 2005. Les synthèses d'études scientifiques et les recommandations concernant la santé publique sont portées par la direction *Public Health Guidance* au sein de NICE.

De manière séparée, en 2013, la fusion de 140 organisations a donné naissance à **Public Health England (PHE)**, une agence du ministère de la Santé, pour aborder le sujet de la santé publique de manière plus rationalisée. Au-delà de ses missions de sécurité sanitaire, PHE porte des sujets de santé publique, mène notamment des revues systématiques de littérature et émet des recommandations pour la mise en œuvre de politiques de santé publique.

PHE travaille en liens étroits avec NICE sur les sujets communs et a contribué à la création du What Works Centre for Well-Being (pour le bien-être).

Organisation

NICE est un organisme public indépendant (*Non Departmental Public Body*) financé par le *Department of Health*.

Le centre est doté d'une équipe de 600 personnes dans son ensemble et d'une trentaine de personnes travaillant sur les sujets de santé publique. Dans le champ de la santé publique, NICE a pour rôle de développer des recommandations pour la prévention de problèmes de santé auprès de NHS, des collectivités locales, du grand public et du secteur privé et associatif. Il produit aussi des « Briefings » pour les collectivités locales, pour conseiller sur les actions de santé publique qui ont un impact positif et un bon rapport effet/prix.

Pour produire ses recommandations, NICE s'appuie sur des revues systématiques de littérature scientifique portant sur les sujets étudiés, portées en interne ou commandées auprès d'universitaires. Concernant la santé publique, ces études sont menées pour la plupart en interne.

De manière complémentaire, et de plus en plus coordonnée, PHE produit aussi, en interne, des analyses scientifiques des interventions pour émettre des recommandations adressées au grand public et au *Department of Health*, notamment en termes de politique publique.

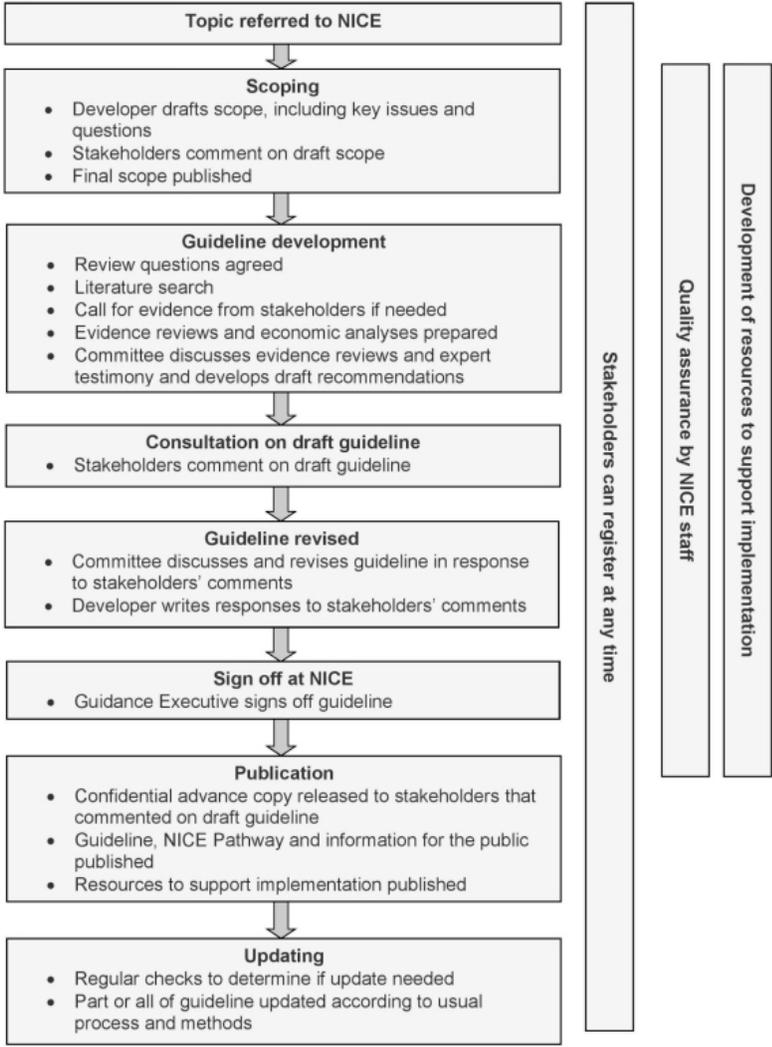
Action

Synthétiser

NICE développe des recommandations à partir des études systématiques des évaluations d'impact menées sur des interventions de santé publique. NICE a traité plus de 60 sujets de santé publique : une des premières études portait sur l'obésité, son épidémie, son contexte. Les sujets vont maintenant des problématiques liées à la consommation d'alcool, à la tuberculose jusqu'à la prévention contre le SIDA. Les processus d'évaluation, de prises de décision, de rédaction des recommandations sont décrits dans le manuel : « *Developing NICE guidelines* »²⁸.

28 www.nice.org.uk/process/pmg20/chapter/1-introduction-and-overview.

What Works Centres britanniques : **quels enseignements pour les politiques fondées sur la preuve en France ?**



Source : www.nice.org.uk/process/pmg20/chapter/1-introduction-and-overview.

Les méthodes d'évaluation – enjeux et limites des méthodes « classiques »

Pour évaluer les effets d'une intervention, NICE prend en compte les « meilleures études scientifiques disponibles ». Idéalement, il s'agit d'évaluations randomisées bien menées, comme dans le champ des médicaments ou appareillages médicaux. Mais pour la santé publique, il s'agit d'aborder la question avec une pluralité de méthodes, pour deux raisons :

- **les évaluations randomisées sont conduites seulement pour certains aspects de la santé publique** : elles conviennent bien pour répondre à des questions précises, mais cela ne convient pas toujours aux modalités d'intervention en santé publique ;
- **les évaluations en lien avec la médecine clinique peuvent être peu nombreuses**. En effet, les problèmes de santé publique sont transversaux et dépendent autant de la personne concernée (et de son éventuelle relation avec un praticien) que de son environnement. Pour aborder ces questions, il est nécessaire d'intégrer la sociologie, l'anthropologie, la géographie, les sciences politiques.

“Vous pouvez utiliser des évaluations randomisées pour des interventions bien définies, auprès d'un public ciblé, visant un résultat préalablement identifié sur la santé des personnes, mais ces cas sont assez rares.”

Mike Kelly, University of Cambridge.

Cela conduit donc NICE à prendre en compte aussi des études se fondant sur l'observation et des approches qualitatives, les entretiens d'experts, la modélisation économique.

Par exemple, pour les travaux sur la prévention et la réduction de l'obésité, il s'agissait aussi de travailler avec des professionnels du sport, de la nutrition, de l'éducation, des transports, des autorités locales, des infirmiers, des architectes.

What Works Centres britanniques : quels enseignements pour les politiques fondées sur la preuve en France ?

Dans le domaine de la santé publique, NICE se coordonne avec PHE pour le développement de recommandations. Les sujets d'études sont en général choisis par le Ministère de la Santé ou par NHS, dans une bibliothèque de sujets. Ils sont ensuite traités dans un ordre de priorité défini de manière commune notamment avec PHE, NHS et le Ministère de la Santé, principalement à partir des priorités mises en avant par les collectivités locales et les corps professionnels. Les priorités sont revues chaque année et peuvent évoluer en fonction des besoins et de l'évolution du contexte.

Une synthèse systématique par NICE prend environ deux ans, de son lancement à la publication des recommandations, en tenant compte de tous les temps de concertation mobilisant des professionnels de secteurs divers et des acteurs de la société civile. La revue systématique en elle-même prend 3 à 6 semaines.

Les sujets de santé publique sont abordés en collaboration avec un expert de PHE pour cadrer l'étude, identifier le type d'évaluations à prendre en compte, comprendre les enjeux, et en fin de course pour revoir les travaux d'évaluation de NICE et les accompagner dans la validation finale.

Le Public Health Advisory Committee (comité consultatif indépendant)

– composé de professionnels experts de la santé, de chercheurs, d'économistes de la santé, de travailleurs sociaux, de salariés d'associations et de membres de la société civile – élabore les recommandations de manière indépendante à partir des revues systématiques, des évaluations économiques, d'autres formes de synthèse d'études et les apports d'experts.

Les recommandations peuvent aussi concerner la mise en œuvre de recherches futures, là où les études ne sont pas suffisantes.

Les *Citizens Councils* (Conseils de citoyens) apportent un regard pouvant représenter l'opinion du grand public. Ce regard est ensuite pris en compte par le comité qui développe les recommandations. Par ailleurs, les

comités consultatifs sont ouverts au public pour garantir la transparence et l'indépendance des travaux menés.

Enfin les études sont elles-mêmes revues régulièrement afin d'assurer leur mise à jour – la fréquence dépend des études.

Exemple : PH 53 – la gestion du poids : services dans la vie courante pour les adultes en surpoids ou obèses²⁹.

Les « recommandations » publiées sur le site de NICE recouvrent des recommandations pour les différents acteurs locaux ainsi que nationaux – notamment pour favoriser une approche intégrée localement de la problématique, en lien avec des associations locales (comme des associations de marche), des professionnels de santé, *Public Health England*, etc. Par exemple, une recommandation porte sur le fait que le responsable au sein d'une collectivité de l'achat ou de la mise en œuvre des services de gestion du poids dans la vie courante :

- soit au courant des caractéristiques de la population en surpoids ou obèses dans sa collectivité ;
- connaisse l'ensemble des services possibles de ce type ;
- et continue à se former sur ce sujet.

Les recommandations reposent sur des évaluations randomisées d'une part et sur des rapports d'experts d'autre part, mêlant donc les deux approches. Enfin, cette étude, publiée en mai 2014, est une mise à jour de recommandations émises en 2006, et sera de nouveau mise à jour en 2017.

PHE produit aussi des études systématiques de recherches scientifiques

pour des interventions précises, à la demande du *Department of Health*. L'institution peut apporter des réponses en quelques mois et s'adapte donc mieux au temps politique. Elle intervient donc de façon complémentaire à NICE. Elle a notamment mené des études et promulgué

What Works Centres britanniques : quels enseignements pour les politiques fondées sur la preuve en France ?

des recommandations sur la e-cigarette, la consommation de sucre, l'abus d'alcool.

Expérimenter

NICE n'effectue pas elle-même de recherches « primaires » et n'en commande pas non plus auprès de tiers, car elle n'est pas missionnée pour cela et c'est le rôle d'autres organismes. Mais, à partir des études secondaires menées, elle peut recommander des évaluations et études spécifiques. En particulier, toutes les recommandations (« guidelines ») publiées par NICE contiennent un chapitre sur des évaluations complémentaires recommandées. NICE communique ces recommandations au *National Institute for Health Research*, qui décide des sujets de recherche et des budgets et peut prendre en compte ces travaux.

À noter : en menant des études sur certains sujets de santé publique, il faut prendre en compte le fait que certaines habitudes (en termes d'alimentation, d'activité physique, de consommation de tabac par exemple) prennent des décennies plutôt que quelques jours ou mois pour évoluer.

Diffuser et accompagner l'adoption

NICE et PHE jouent tous les deux un rôle essentiel dans la diffusion et l'accompagnement à l'adoption de pratiques et d'interventions dont l'impact est démontré.

Atteindre le grand public

PHE en particulier a le rôle de diffuser auprès du grand public des pratiques de prévention ou de traitement d'un certain nombre de sujets : e-cigarette, contrôle ou réduction du poids, abus d'alcool, etc. Cela peut passer par un ensemble d'actions de communication comme la diffusion de messages (type « 5 fruits et légumes par jour »), le développement d'applications pour

smartphone ou d'outils sur internet³⁰.

La façon de diffuser ces messages est conçue par l'équipe *Health Social Marketing* (marketing social de la santé) de PHE, qui s'appuie fortement sur les études comportementales, dont les « *nudges* » (incitations douces), en collaboration avec l'équipe « comportementaliste » au sein de PHE (*Behavioural Insight Team*). PHE relève des données quantitatives sur les résultats des campagnes (nombres de clics, d'outils téléchargés,...). Il n'y a pas encore d'évaluation d'impact de ces campagnes elles-mêmes. Le grand public peut aussi être atteint via les praticiens et les collectivités locales, auprès de qui PHE et NICE mènent des actions.

Atteindre les professionnels et les collectivités locales

La mise en œuvre de nombreuses actions en lien avec la santé publique dépend des collectivités locales. C'est d'autant plus vrai que ce champ touche à des acteurs et secteurs variés. NICE mène donc des actions et développe des outils pour atteindre ces parties prenantes.

Outils en ligne

NICE propose un site web très fourni à destination de professionnels cherchant à mettre en œuvre les recommandations (en particulier dans le monde médical), où l'on peut trouver :

- **l'ensemble des recherches** menées par thèmes de recommandations (*guidance*). Les recommandations y sont présentées, ainsi que les acteurs qui peuvent/doivent agir sur ce sujet. Les méthodes de recherche et les études analysées sont aussi accessibles. Les recommandations présentées pour la santé publique portent principalement sur des actions que les collectivités locales peuvent mener et sur les acteurs locaux à solliciter en fonction des sujets;
- **des chemins logiques d'approches** pour identifier les recommandations pertinentes (*pathways*);
- **un outil d'auto-évaluation** (*quality standards*) qui permet aux structures

³⁰ Par exemple le programme One You et les applis qui y sont reliées pour contrôler sa consommation d'alcool ou arrêter de fumer. www.nhs.uk/oneyou/apps#MkLw5RVxYTI4PH1Z.97.

What Works Centres britanniques : quels enseignements pour les politiques fondées sur la preuve en France ?

d'identifier leur niveau de connaissance et de mise en œuvre des recommandations et de travailler vers les étapes suivantes. Par exemple en santé publique, un indicateur repose sur le fait que les autorités locales et les organismes de santé donnent l'accès à des aliments et boissons sains dans les distributeurs à disposition du public et les mettent en évidence dans tous ces circuits de distribution. Cela fait partie des recommandations de lutte contre l'obésité.

– **des accès à des sites de recherche**, permettant aux utilisateurs d'accéder à des articles de recherche scientifique.

Ce site est accessible à tout public.

The screenshot shows the NICE Evidence Services interface. At the top, there is a search bar and navigation tabs for 'NICE Pathways', 'NICE Guidance', 'Standards and indicators', and 'Evidence services'. Below the search bar, the page title is 'Maternal and child nutrition overview'. The main content area displays a flowchart titled 'Maternal and child nutrition' with the question 'What type of recommendation are you interested in?'. The flowchart branches into three categories: 'Strategy, policy and commissioning', 'Training', and 'Primary and secondary care and community health services'. Each category has a sub-link: 'Advice for teachers and carers about nutrition for children in pre-school day care', 'Advice for manufacturers of goods for women who may become pregnant, and for pregnant and breastfeeding women', and 'Interventions and advice about breastfeeding for women who may become pregnant, or who are pregnant or breastfeeding' respectively. On the right side, there is a sidebar with tabs for 'About', 'Resources', 'Information for the public', and 'Quality standards'. Below these tabs, there is a list of resources under the heading 'Maternal and child nutrition overview', including 'Strategy, policy and commissioning for maternal and child nutrition', 'Training in maternal and child nutrition', 'Maternal and child nutrition in primary and secondary care and community health services', and 'Interventions and advice about breastfeeding for women who may become pregnant, or who are pregnant or breastfeeding'.

“Au-delà de la production de recommandations, NICE joue un rôle dans la collecte et la diffusion d'une base plus large d'évaluations scientifiques auprès de son public, à travers la plateforme NICE Evidence Services.”

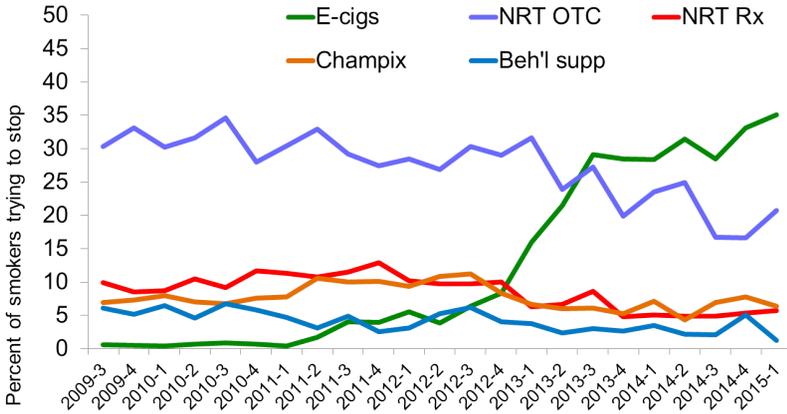
Alexia Tonnel, NICE.

Par exemple, le système de recherche NICE Evidence Search fournit un accès à des recommandations d'autres organismes que NICE, à des ressources bibliographiques, ainsi qu'à d'autres formes de synthèses d'évaluations, études de cas et « littératures grises » produites par des

organismes de confiance. Ces études scientifiques peuvent être utilisées par des professionnels de la santé publique en combinaison avec des données locales sur la population, des études locales et des informations fournies par PHE, par exemple.

PHE tient un blog dédié à la santé publique, également à destination des professionnels, qui communique sur l'ensemble des recommandations de santé publique – et renvoie le cas échéant vers des liens grand public³¹.

Par exemple, on y trouve les explications très détaillées et didactiques sur la synthèse réalisée concernant les e-cigarettes comme outil d'aide à l'arrêt du tabagisme. Cela permet notamment de répondre au début public qui avait eu lieu suite à la recommandation faite par le PHE d'utiliser les e-cigarettes³².



Source : publichealthmatters.blog.gov.uk.

Graphique montrant que les e-cigarettes ont maintenant la préférence des utilisateurs comme aide pour arrêter de fumer.

³¹ publichealthmatters.blog.gov.uk.

³² publichealthmatters.blog.gov.uk/2015/09/18/e-cigarettes-a-public-health-response.

What Works Centres britanniques : quels enseignements pour les politiques fondées sur la preuve en France ?

NICE et PHE reconnaissent que la publication de recommandations en soi est loin d'être suffisante pour que les acteurs s'approprient les méthodes et les résultats.

NICE a pendant un temps investi sur l'accompagnement local, en particulier avec des consultants dédiés qui accompagnaient les acteurs locaux et pouvaient faire un retour pour améliorer les recommandations de mise en œuvre. NICE avait aussi créé un groupe stratégique pour la mise en œuvre des recommandations, qui était influent. Mais elle a une approche plus standardisée aujourd'hui, avec en particulier des notes sur la mise en œuvre et la possibilité de se saisir de NICE.

PHE peut animer le dialogue local, grâce à ses antennes locales, pour co-construire les approches avec les acteurs locaux. De nombreuses recommandations ont été développées par NICE et PHE pour le déploiement des recommandations par les autorités locales. Ces deux acteurs identifient de nouvelles pistes de développement pour atteindre les décideurs.

“Les collectivités locales aiment faire les choses d'une certaine manière. Nous sommes tous en recherche des bonnes méthodes pour la mise en œuvre. [... Dans le cas de l'étude sur le bien-être], un aspect intéressant a été que les mesures du bien-être ont permis d'ajuster des décisions sur les politiques locales. ”

Gregor Henderson, PHE.

En ce qui concerne les acteurs médicaux, même si l'application des recommandations de NHS (reprenant les travaux de NICE) n'est pas obligatoire, elle fait souvent partie de la responsabilité des professionnels et il y a de nombreux mécanismes d'audit et financiers pour mettre en œuvre ces recommandations.

L'adoption par les politiques publiques et la loi

Les recommandations de NICE ne visent pas les politiques publiques nationales mais les actions portées par des acteurs locaux.

PHE, à l'opposé, est mandaté notamment pour faire des recommandations portant sur les politiques publiques nationales. Cela demande une approche plus confidentielle pour éviter d'envoyer des messages contradictoires et s'assurer que le gouvernement est prêt à enclencher les démarches. Certains sujets comme les recommandations sur la e-cigarette ou la taxe sur le sucre ont déclenché des débats publics agités, tant dans les tabloïds que les revues scientifiques.

Perspectives

- NICE dispose d'un programme de synthèse d'évaluations scientifiques qui s'étend sur plusieurs années, tandis que PHE agit de manière plus réactive face aux commandes politiques et aux évolutions du contexte.
- NICE et PHE travaillent de manière de plus en plus collaborative et intégrée sur les questions de santé publique ; les recommandations (guidance) produites par NICE vont être aussi labellisées par PHE pour en favoriser l'appropriation.
- NICE et PHE cherchent à atteindre les décideurs, et surtout les collectivités locales, de manière plus efficace : ils portent de manière coordonnée des projets pour aller plus loin dans ce domaine, notamment en développant des outils plus adaptés.

Enseignements

- Les solutions identifiées et les interventions de santé publique peuvent être portées de manière efficace par des acteurs très variés, et pas uniquement les spécialistes de la santé. Il est donc intéressant d'en tenir compte dans les dispositifs évalués.
- La transparence dans les méthodes de synthèse de la recherche, la consultation de plusieurs ensembles de parties prenantes, la publication de l'ensemble des travaux contribuent à la reconnaissance de NICE et à la valorisation de ses recommandations.

What Works Centres britanniques : **quels enseignements pour les politiques fondées sur la preuve en France ?**

“Tout ce que nous faisons est du domaine public, y compris le rôle des comités et leurs membres.”

Françoise Cluzeau, NICE.

Il peut être difficile d'avoir des évaluations randomisées pour l'ensemble des questions de santé publique : il est important d'être ouvert à d'autres approches pour avoir a minima des indications sur les leviers d'action.

Zoom sur le What Works Well-Being – le bien-être

Fiche identité	
Périmètre	Comprendre ce que le gouvernement, les collectivités locales, les associations, les individus et les entreprises peuvent faire pour améliorer le bien-être des personnes.
Date de création	Octobre 2014
Statut	Organisation indépendante
Taille de l'équipe	6 salariés
Budget annuel	600 000 livres sterling
Site web	whatworkswellbeing.org

Ce What Works Centre a été créé très récemment, financé notamment par ESRC, le *Department of Health* et 7 autres ministères. Il a publié ses premières revues systématiques en septembre/octobre 2016.

La définition du bien-être et les approches pour le mesurer sont la base de ce travail. Le centre utilise la définition officielle du bien-être au Royaume-Uni : « *Le bien-être veut dire “comment on va” en tant qu'individus, en tant que communautés et en tant que nation, et dans quelle mesure notre mode de vie est durable.* » Le bien-être individuel est

particulièrement important pour le centre, défini comme notre satisfaction avec nos vies, l'impression que ce que nous faisons a une valeur, nos émotions quotidiennes (joie, anxiété) et notre bien-être mental au sens large.

Le centre joue un rôle de *hub*. Il repose sur une équipe centrale de 6 salariés pour « traduire » les résultats d'études scientifiques à destination des acteurs ciblés. Cette équipe est composée de la directrice, des responsables de recherche, des responsables des programmes et des responsables de la communication. Le centre inclut aussi 53 chercheurs auxquels le centre fait appel pour mener des travaux de revues et de recherche.

Le centre a été chargé de réaliser des revues systématiques autour des sujets suivants :

- la formation professionnelle et le travail ;
- le bien-être pour les personnes non-actives ;
- le bien-être de la collectivité ;
- la culture et le sport.

Le centre adapte sa méthode de revue et ses critères d'inclusion en fonction de la nature et de la qualité de la recherche disponible³³.

Sur sa première phase d'activité, l'équipe du centre se concentre autant que possible sur des revues systématiques de littératures quand elles sont disponibles et sur leur « traduction » en résultat compréhensible par un large public. Elle s'appuie sur un « *evidence panel* » pour assurer la qualité des synthèses produites et note la qualité des évaluations selon le modèle « *GRADE* » (aussi adopté par NICE et le Cochrane collaboration).

Pour produire des études pertinentes, le centre se nourrit aussi de recherches qualitatives, notamment pour comprendre pourquoi il y a eu (ou non) un impact, et pourquoi les impacts sont vécus différemment par différentes personnes ou dans différents contextes...

What Works Centres britanniques : **quels enseignements pour les politiques fondées sur la preuve en France ?**

Enfin, il faut souligner que le bien-être est un champ très large, multidisciplinaire et très lié, voire inclus dans les questions de santé publique. Il est donc intéressant d'aller au-delà de la recherche déjà répertoriée et revue par des pairs, par exemple en parlant à des experts, en suivant le « fil d'Ariane » des références des textes cités, en effectuant des recherches manuelles dans différentes publications...

La prochaine étape sera d'étoffer les ressources de communication. Après deux ans à préparer les publics cibles à accueillir des résultats scientifiques et à s'inspirer de pratiques existantes, le WWC envisage un travail de communication ciblant spécifiquement chacun des acteurs essentiels : le gouvernement, le secteur privé, les régions, etc.

Enfin, le centre prévoit de s'appuyer sur les études en cours pour identifier les approches les plus efficaces pour l'adoption de résultats scientifiques. En effet, le centre s'intéresse de près à un sujet émergent appelé « la science de l'utilisation de la connaissance scientifique³⁴. »

France Stratégie

Le présent rapport, réalisé par l'Ansa avec le soutien de ses partenaires, contribue utilement à la réflexion sur l'institutionnalisation de l'évaluation en France et ses modalités.

L'examen des objectifs qui ont présidé à la création de neuf *What Works Centres* au Royaume-Uni avec le soutien du *Cabinet Office*, et les missions qu'ils remplissent actuellement, mettent à jour trois grandes priorités : la promotion des évaluations d'impact des interventions et des dispositifs dans les domaines éducatif, de l'emploi, social et sanitaire ; la capitalisation et la diffusion de leurs enseignements ; l'accompagnement des professionnels et des décideurs publics locaux pour mieux s'en emparer.

Ces trois priorités rencontrent plusieurs des propositions et considérations des travaux de France Stratégie. Elles font écho à des évolutions de l'action publique.

L'efficacité et l'efficience des interventions publiques sont un enjeu majeur dans la plupart des grandes démocraties où un niveau élevé des dépenses publiques se conjugue à un certain désenchantement à l'égard des institutions et de l'action publique³⁵. Cet état de l'opinion soutient une exigence d'évaluation. En France, un écart existe entre la situation de l'évaluation d'impact et l'intérêt dont elle fait l'objet. La démarche évaluative dominante demeure l'évaluation des politiques publiques, qui porte principalement sur la pertinence, la cohérence et la mise en œuvre des politiques. L'utilité de cette démarche n'est pas contestée et ses apports ne sont pas substituables. Le nombre et la qualité des évaluations d'impact appellent toutefois à être

³⁵ Sur la crise de confiance dans les institutions, Jean Pisani-Ferry, Fabrice Lenglard, Daniel Agacinski, Gilles Bon-Maury (dir.), *Lignes de faille. Une société à réunifier*, France Stratégie, 2016.

renforcés. Leur spécificité est de tester la relation de causalité entre un dispositif et les objectifs poursuivis, et de distinguer si une pratique a un effet mesurable pour ses bénéficiaires. France Stratégie a souhaité en faire mieux connaître les finalités et les méthodes dans *Comment évaluer l'impact des politiques publiques ? Un guide à l'usage des décideurs et praticiens* (Rozenn Desplat et Marc Ferracci, 2016)³⁶.

L'intérêt pour des démarches ou des structures orientées vers l'identification des dispositifs « qui fonctionnent » sur une problématique donnée est alimenté par d'autres facteurs de transformation de l'action publique. Les domaines investigués par les *What Works Centres* ne sont pas le fait du hasard. L'État social, historiquement assis sur des techniques d'assurance sociale, des mécanismes de transfert et des prestations monétaires, réoriente aujourd'hui ses interventions vers des services, des dispositifs d'accompagnement, des investissements dans l'éducation, la formation, également dans la préservation de la santé et l'autonomie des personnes. Cette nouvelle direction, parfois qualifiée de *stratégie d'investissement social*, s'opère en présence de nouveaux défis sociaux qui ont trait à la lutte contre les inégalités (inégalités précoces de développement, scolaires, entre les femmes et les hommes, de santé, etc.), à la lutte contre la pauvreté, le chômage et l'exclusion, ainsi qu'au vieillissement démographique. Cette évolution emporte de l'innovation, du foisonnement et de la complexité et elle fait apparaître une demande inédite d'identification des dispositifs « pertinents » et d'évaluation. Là où à un risque, l'on faisait correspondre une prestation et l'on évaluait l'effet par un taux de couverture et le niveau de vie, se substitue des enjeux de personnalisation des services, d'évaluation d'impact, voire d'évaluation des rendements³⁷.

³⁶ www.strategie.gouv.fr/publications/evaluer-limpact-politiques-publiques.

³⁷ Marine Boisson-Cohen, Bruno Palier et al., *L'investissement social : quelle stratégie pour la France ?*, enseignements du séminaire partenarial CNAF-DGCS- France Stratégie-LIEPP, à paraître courant 2017. En savoir plus : investissementsocial.org. Voir également Arthur Heim, *Peut-on estimer le rendement de l'investissement social ?*, Note d'analyse, France Stratégie, 2017.

What Works Centres britanniques : **quels enseignements pour des politiques fondées sur la preuve en France ?**

Des démarches et des structures d'inspiration *What Works Centres* sont enfin à considérer dans un contexte d'action publique qui embrasse un ensemble plus vaste que l'État et ses démembrements, comprend les organismes sociaux et les collectivités territoriales, intègre l'échelon européen, et à laquelle contribuent les acteurs de l'économie sociale et solidaire (ESS). Partenariale, décentralisée et déconcentrée, l'action publique a développé les marges de manœuvre de l'encadrement de proximité et des professionnels, établi un droit à l'expérimentation des collectivités territoriales, accru la part des financements sur projets. Ces transformations justifient des outils de mutualisation des connaissances, d'échanges des bonnes pratiques ou des interventions validées, d'interface qui faciliteront leur valorisation et leur capitalisation. L'État a vocation à jouer un rôle essentiel vis-à-vis de telles démarches ou structures relevant d'un intérêt commun³⁸. Les modalités concrètes de mise en place ou en visibilité de telles démarches ou structures, le détail de leurs missions et de leurs leviers d'intervention, la nature du soutien que l'État pourrait leur apporter et l'implication des services ministériels demeurent toutefois à investiguer. Ces différents aspects feront l'objet d'un approfondissement dans le cadre d'un travail collaboratif entre France Stratégie et l'Ansa dans les prochains mois.

Marine Boisson-Cohen, experte référente et adjointe au directeur, département Société et Politiques sociales.

38 Dominique Bureau et Marie-Cécile Naves (coord.), *Quelle action publique pour demain ?*, France Stratégie, 2015.

Secrétariat général pour la modernisation de l'action publique (SGMAP)

Le rapport d'étude établi par l'Ansa répond à l'objectif initial qui consistait à documenter l'intérêt croissant pour des « politiques publiques fondées sur la preuve », intérêt qui se heurte à une faible acculturation de ces méthodes en France.

L'étude monographique du fonctionnement des What Works Centres britanniques, au cœur du rapport (qui ne saurait se confondre avec un travail de recherche ou une démarche d'évaluation), apporte des éléments d'éclairage intéressants, en particulier sur les méthodes et bonnes pratiques développées par ces structures.

Le recours à des revues systématiques synthétisant les connaissances disponibles de ce qui marche présente des apports indéniables, bien que l'on comprenne à la lecture du rapport qu'elles peuvent être malaisées à mettre en œuvre du fait :

- du manque de connaissances et d'expérimentations dans certains champs de politique publique,
- de la complexité liée au transfert des interventions « efficaces » d'un contexte à un autre.

L'intensification des évaluations d'impact constitue sans nul doute une voie intéressante pour promouvoir des décisions publiques éclairées par la « preuve ». Pour autant, la difficulté, pointée dans le rapport, de recourir à ce type de méthodes quantitatives invite à rester ouvert à d'autres types de « preuves », de nature plus descriptives ou qualitatives, et plaide in fine pour l'usage autant que faire se peut de méthodes « mixtes ».

Le soin particulier accordé par les What Works Centres à la diffusion

What Works Centres britanniques : **quels enseignements pour des politiques fondées sur la preuve en France ?**

de leurs enseignements auprès des professionnels qu'ils accompagnent dans la phase d'appropriation et de mise en œuvre retient enfin l'attention du SGMAP. Le ciblage des professionnels de terrain, ainsi que la promotion de méthodes innovantes auprès des territoires, sont des axes à approfondir en France, dans le sillage des réflexions menées par le SGMAP et le CGET sur le « passage à l'échelle ».

Le SGMAP partage le point de vigilance mentionné en fin d'étude concernant l'efficacité du modèle « What Works » dans son ensemble. Cette efficacité, faute d'un recul suffisant et d'un travail évaluatif, n'est pas encore attestée. Dès lors qu'elle le serait, cette efficacité devrait être rapportée aux moyens budgétaires mis en œuvre pour le développement de ces organisations afin d'apprécier l'efficience du modèle. **En l'état, il convient donc d'être mesuré sur l'extrapolation du modèle promu par les What Works Centres au contexte français et de privilégier une démarche d'expérimentations des méthodes qu'ils développent plutôt qu'une adoption sans réserves du modèle.**

Le bref panorama français de l'évaluation réalisé fournit des clés de lecture pour la plupart convergentes avec les réflexions du SGMAP. Des interrogations persistent en France quant à la faisabilité et à l'utilisation des évaluations d'impact. Le nombre d'évaluations d'impact menées en France reste limité, et les travaux étrangers sont rarement transposables en l'état. La diffusion de ces études s'étend rarement au-delà des cercles académiques intéressés et rencontre peu d'intérêt parmi les décideurs.

Les constats de l'étude, loin d'épuiser le débat en France sur la place et les perspectives des politiques fondées sur la preuve, amorcent cependant une dynamique collective à poursuivre sur la question des politiques publiques fondées, sinon sur la preuve, du moins sur la connaissance.

Le SGMAP se positionne en faveur d'une poursuite des réflexions, sur le mode d'une expérimentation ciblée en expérimentant la mise en œuvre de démarches ou de structures visant à outiller l'élaboration

de politiques publiques fondées sur la connaissance. Il ne s'agit pas de trancher ici les questions d'opportunité et de faisabilité relatives à la création de structures ou de réseaux dédiés mais d'avancer en capitalisant les acquis d'un certain nombre de démarches existantes pour traduire en actes cette orientation.

Le SGMAP souhaite aussi que ce travail soit valorisé afin d'animer le débat sur la place des évaluations et la diffusion de leurs résultats aux acteurs chargés de l'élaboration et de la conduite des politiques publiques. L'étude doit également pouvoir être rapprochée des résultats de la « méta-évaluation » que mène le SGMAP sur sa démarche globale d'évaluation des politiques publiques. Cette évaluation préconise une adaptation plus fine des approches aux besoins. Elle souligne que la définition préalable de l'usage attendu d'une évaluation, dès l'élaboration des recommandations (loi, plans d'action, expérimentation, etc.), est une condition nécessaire à la prise en compte de ses enseignements. Sur ce point, le lien avec le modèle What Works est évident.

Des initiatives locales, incarnées notamment par des laboratoires territoriaux d'innovation publique, sont aujourd'hui porteuses de cette volonté de diffuser des expériences qui ont fait leurs preuves.

Cette diffusion de solutions doit être organisée, en particulier par l'État, pour bénéficier à un maximum de nos concitoyens. Le SGMAP entend prendre part à ce chantier collectif, en travaillant dès les prochains mois à la construction opérationnelle d'une telle démarche au service d'une politique publique.

Commissariat général à l'égalité des territoires (CGET)

L'intérêt du CGET pour les What Works Centres britanniques recouvre deux problématiques distinctes : une profusion de travaux scientifiques qu'il devient maintenant difficile de hiérarchiser et de synthétiser ; une difficulté récurrente à évaluer les dispositifs de politiques publiques et à rendre ces évaluations suivies d'effets. Comme le montre très bien le rapport d'étude établi par l'Ansa, la création des What Works Centres apporte, sur ces deux points, des éléments de réponse très intéressants et représente une véritable source d'inspiration pour les acteurs publics français.

Pour ce faire, il est important de considérer les What Works Centres comme un dispositif : de synthèse (dont l'outil principal serait la revue de littérature), de diffusion de solutions probantes (auprès des professionnels) et d'accompagnement méthodologique (d'expérimentations sociales principalement). Insister sur la notion de dispositif a pour intérêt de ne pas réduire la légitimité et la visibilité des What Works à une méthode unique d'évaluation (dite évaluation d'impact dans le rapport). C'est à cette condition que leur intégration dans le paysage français de la production de connaissance et de l'évaluation nous semble possible et souhaitable.

Nous pensons qu'il manque, en France, des structures capables d'organiser (par grands champs thématiques) un suivi systématique des productions scientifiques dans le but d'établir des confrontations et des synthèses, indépendantes et rigoureuses (c'est-à-dire assumant de hiérarchiser les éléments de preuves et la qualité des argumentations) et conçues à des fins d'évaluation et/ou d'aide à la décision (c'est-à-dire mises en forme de telle manière que des options et des alternatives paraissent renseignées).

Dans ce cadre, les évaluations d'impact (telles que définies

et mobilisées par les What Works Centres britanniques) ne sont qu'une source d'information parmi d'autres. L'importance néanmoins que nous leur accordons résulte du fait que nous pensons que ce type d'évaluations nous semble utile et devoir être plus systématiquement déployé. Elles revendiquent un niveau d'ambition scientifique, certes discutable, mais que tous travaux scientifiques ou d'évaluation dignes de ce nom devraient aussi revendiquer. Elles apportent des éléments d'informations que les sciences sociales et les statistiques dites « classiques » ont toujours peiné (ou parfois rechigné) à apporter aux acteurs publics. Elles sont aussi un moyen très évident, et très adapté, d'accompagnement des expérimentations dont nous pensons qu'il est nécessaire d'amplifier l'usage pour tester préalablement, et faire accepter, de nouveaux dispositifs.

Pour autant, ne surestimons pas leur capacité à dire mécaniquement les choses et à être naturellement conclusives. S'agissant des nombreuses réductions auxquelles les conduisent leur méthodologie (et leur souci d'efficacité/simplification à des fins d'évaluation), nous ne pouvons les considérer comme l'unique manière d'évaluer un dispositif de politique publique et, surtout, de déclencher une transformation de l'action publique (qui doit rester, ne le perdons pas de vue, la finalité des What Works)³⁹.

39 Il n'est pas ici l'objet de pointer les limites ou les défauts des évaluations d'impact (et surtout de l'argumentaire politique qui les accompagne). Il suffit de rappeler qu'elles participent chez certains d'une logique de remplacement du jugement par la mesure et de la délibération par la preuve ; à bien des égards, elles tendent à réduire les enjeux de compréhension à l'explication et, ce faisant, à surestimer l'explication dans le passage à l'action (le lien pour le moins erratique entre recours aux évaluations d'impact et généralisation ou passage à l'échelle en constitue une illustration parmi d'autres). Cette confusion est souvent entretenue par une réduction de « l'effet global d'un dispositif » (expression qui, reconnaissons-le, permet de justifier beaucoup de choses et/ou de n'en plus dire grand-chose) à son impact direct et immédiat sur une dimension univoque du bénéficiaire. Cela procède d'un espoir épistémologique – toujours déçu – de traiter strictement une hypothèse (ou une causalité) de manière isolée. Il en résulte souvent un manque d'intérêt des évaluations d'impact pour les caractéristiques organisationnelles, financières ou politiques des politiques publiques... sans compter que cette approche du bénéficiaire n'est pas sans conséquence, à terme, sur la manière de penser la justice sociale, de définir les droits sociaux et les conditions de leur accessibilité (thématique de l'investissement social). À cela s'ajoute leurs fréquentes difficultés de mise en œuvre, d'interprétation des résultats, et l'effet boîte noire évident quand on souhaite comprendre les conditions d'obtention des résultats. Autant de problématiques « scientifico-politiques » complexes, les évaluations d'impacts ne faisant souvent que les cristalliser, et qui

What Works Centres britanniques : **quels enseignements pour des politiques fondées sur la preuve en France ?**

Pour le dire simplement, des What Works Centres, à la française, de la revue systématique de littérature à la discussion publique, devraient contribuer à mettre en scène et à rendre accessible le pluralisme des données et des avis. On peut alors se demander si, au final, des dispositifs du type conférence de consensus, ne devraient pas aussi constituer un mode opératoire privilégié des What Works Centres?

Au final, pour le CGET, ce travail a le mérite de pouvoir ranimer le débat, quelque peu éteint, sur la contribution des scientifiques à l'évaluation et sur les conditions de son organisation à des fins opérationnelles. Nous espérons donc que soit organisée très prochainement sa valorisation (sous forme de mises en perspective contradictoires et pourquoi pas d'une étude de faisabilité) auprès des acteurs concernés.

Emmanuel Dupont.

mériteraient bien sûr d'être plus systématiquement explorées (même si certaines d'entre-elles pourraient être amenées à évoluer rapidement avec le surgissement du « big data »).

Santé publique France

Santé publique France a soutenu le projet de l'Ansa et de ses partenaires car celui-ci rejoint le cœur de ses préoccupations à savoir contribuer à développer sur le territoire des actions de prévention et de promotion de la santé dont l'efficacité potentielle a déjà été démontrée. Deux voies sont possibles pour cela. La première est d'importer, adapter et évaluer dans le contexte français des interventions efficaces dans d'autres pays et d'autres contextes. La deuxième est d'innover, expérimenter et/ou évaluer des interventions créées en France.

C'est ainsi par exemple que Santé publique France a adapté l'opération de *Public Health England* nommé « Stoptober » et qui avait démontré son efficacité en augmentant le nombre de tentatives d'arrêt du tabac outre-Manche (Brown, Kotz et al. 2014). L'adaptation française a été mise en place en novembre dernier en France sous le nom « Mois sans Tabac » et les résultats de l'évaluation de l'opération française seront connus ultérieurement. Santé publique France est également impliqué dans un projet ambitieux expérimental porté par la CNAMTS avec le soutien de la DGS, de prévention de l'entrée dans le diabète de type 2. Cette expérimentation a été construite sur les expériences américaine (Albright and Gregg 2013), finlandaise et chinoise dans ce domaine et permet de réduire de plus de 50% l'entrée dans le diabète par un programme intensif d'aide à l'amélioration de l'alimentation et l'augmentation de l'activité physique.

Santé publique France est également impliqué avec la Mission interministérielle de lutte contre les conduites addictives dans l'évaluation des projets de terrain dans le cadre de la Commission interministérielle de prévention des conduites addictives (CIPCA). Un appel à candidatures a été lancé qui a permis de repérer 80 projets

What Works Centres britanniques : quels enseignements pour des politiques fondées sur la preuve en France ?

de terrain. Cinq projets ont été retenus par un jury interministériel, et des ressources financières et humaines ont été dégagées grâce à la MILDECA pour décrire et évaluer ces cinq projets. Ce travail est en cours.

Au-delà de l'implication de Santé publique France dans la prévention santé qui est son champ spécifique, l'importance de ce partenariat tient au fait des intrications extrêmement étroites entre économie, social et santé.

De ce fait, des actions efficaces dans le domaine de la santé, peuvent provenir des secteurs de l'économie ou du social et réciproquement. Par exemple, l'accès à un logement peut améliorer la santé. Kellam a démontré que l'amélioration du climat de travail dans les premières années de classe de l'école primaire produisait des effets sur la santé au long terme, en particulier un moindre tabagisme 15 ans après (Kellam, Mackenzie et al. 2011). Des programmes d'aide à la parentalité ont montré des réductions de consommation de substances psychoactives (Spath, Redmond et al. 2001). Moffitt et ses collègues ont pu montrer grâce à la cohorte de naissance néozélandaise de Dunedin des relations entre des phénomènes observés dans l'enfance et un ensemble de situations relevant aussi bien du secteur social que sanitaire (Moffitt, Arseneault et al. 2011).

Donc, au-delà de l'intérêt du partage des méthodes autour des synthèses de littérature, des méthodes d'expérimentation et d'essaiage qui ne sont pas à négliger, ce partenariat est surtout important en raison des intrications et interactions des champs d'intervention des partenaires de ce projet.

Albright, A. L. and E. W. Gregg (2013). "Preventing type 2 diabetes in communities across the U.S.: the National Diabetes Prevention Program." *Am J Prev Med* 44(4 Suppl 4): S346-351.

Brown, J., D. Katz, S. Michie, J. Stapleton, M. Walmsley and R. West (2014). "How effective and cost-effective was the national mass media smoking cessation campaign 'Stoptober'?" *Drug and Alcohol Dependence* 135(100): 52-58.

Kellam, S. G., A. C. Mackenzie, C. H. Brown, J. M. Poduska, W. Wang, H. Petras and H. C. Wilcox (2011). "The good

behavior game and the future of prevention and treatment." *Addict Sci Clin Pract* 6(1): 73-84.

Moffitt, T. E., L. Arseneault, D. Belsky, N. Dickson, R. J. Hancox, H. Harrington, R. Houts, R. Poulton, B. W. Roberts, S. Ross, M. R. Sears, W. M. Thomson and A. Caspi (2011). "A gradient of childhood self-control predicts health, wealth, and public safety." *Proceedings of the National Academy of Sciences of the United States of America* 108(7): 2693-2698.

Spoth, R. L., C. Redmond and C. Shin (2001). Randomized trial of brief family interventions for general populations: adolescent substance use outcomes 4 years following baseline. *J Consult Clin Psychol.* 69: 627-642.

Caisse nationale des allocations familiales (Cnaf)

La Caisse nationale des allocations familiales a souhaité soutenir cette étude réalisée par l'Ansa en raison d'un double objectif :

- le premier est de contribuer à alimenter la démarche stratégique de l'institution dont l'un des axes structurants est l'intégration d'une démarche d'investissement social. Cette dernière devrait se révéler riche et opératoire pour définir des orientations de politiques publiques et prioriser des dispositifs et des modalités d'intervention en fonction de leur rendement économique, en capital humain par rapport à d'autres. L'étude sur les What Works peut contribuer à rendre ce concept opérationnel pour la branche Famille dans le cadre notamment, d'une réflexion pour la création d'un fonds d'investissement social qui conduirait des évaluations ;
- la Cnaf a également des missions d'évaluation des politiques publiques portées par sa Direction des statistiques, des études et de la recherche. Le second objectif porte ainsi d'une part, sur une meilleure connaissance des évaluations d'impact et d'autre part, sur la manière dont – de manière générale – les résultats d'évaluation peuvent être mieux appropriés par les acteurs publics.

Le premier intérêt du rapport est d'ordre méthodologique. L'important travail de revues de littérature réalisé par les centres est un élément notable qui nous semble manquer en France. Chaque centre réalise dans son champ d'expertise ce travail qui donne lieu à une bibliothèque mise à disposition pouvant contribuer tant à la réflexion scientifique que politique pour les acteurs de terrain. Organiser l'évaluation par champs thématique, en matière notamment de petite enfance, de jeunesse ou de soutien à la parentalité est un autre apport des centres britanniques.

Les méthodologies d'impact, afin de pouvoir identifier une relation de cause à effet et déterminer ainsi l'efficacité d'un dispositif sont

encore insuffisamment développées et valorisées en France. Or, elles apparaissent comme l'un des outils utiles d'aide à la décision pour justifier d'une intervention, de la dépense induite mais également pour rendre compte de l'action publique auprès des citoyens/bénéficiaires. L'étude souligne bien que ces méthodologies se prêtent davantage aux expérimentations et dispositifs locaux qu'à des politiques nationales. C'est pourquoi, notre intérêt est d'autant plus marqué que le fonds d'investissement social que nous souhaitons créer serait de pouvoir précisément financer des expérimentations et évaluer des dispositifs locaux. La complémentarité de cette méthode avec d'autres, plus classiques, nous apparaît particulièrement présente dans les éléments fournis par l'étude. L'intérêt de cette approche et de sa rigueur scientifique peut permettre de juger de l'efficacité d'un dispositif et, combinée si nécessaire à d'autres types d'évaluations, il devient également possible d'expliquer pourquoi et comment améliorer l'intervention. C'est ainsi également l'amélioration de l'existant qui est en jeu.

Nous manquons en France de lieu ressource pour organiser le dialogue entre scientifiques et décideurs. Le travail qui pourrait être organisé dans le cadre de notre projet de fonds d'investissement social pourrait s'inspirer des centres britanniques comme structures dédiées au transfert de connaissance et à l'accompagnement des acteurs. Ces centres produisent trois types d'outils pour accompagner l'appropriation des résultats de la recherche auprès des acteurs : des bibliothèques et revues systématiques ; des outils d'auto-évaluation ; des descriptifs d'intervention ou guides. Même si l'efficacité des What Works Centres dans cette appropriation n'est pas prouvée, ils offrent des pistes intéressantes pour rendre accessible les résultats de la recherche sous une forme appropriable aux différents acteurs. Tout l'intérêt est de cibler notamment les professionnels de terrain aux prises avec la mise en œuvre des dispositifs. L'enjeu est d'être en capacité de susciter une demande d'évaluation, voire d'expérimentations pour encourager la création de centres en France.

Le rapport fournit de précieux éléments de réflexion sur les conditions

What Works Centres britanniques : **quels enseignements pour des politiques fondées sur la preuve en France ?**

de déploiement. L'indépendance des structures portant l'évaluation est fondamentale pour assurer la rigueur scientifique. Des moyens financiers et humains pour conduire des revues de littérature, des évaluations de qualité et la diffusion des résultats aux acteurs sont également nécessaires. Un partenariat entre différentes institutions pour mutualiser les moyens pourrait s'avérer être l'une des pistes pour concrétiser l'existence de tels centres.

Une évaluation rigoureuse d'une part, et la remontée des expériences de terrain d'autre part, nous paraît être les principaux apports des What Works Centres. Sur ces deux points l'étude menée par l'Ansa alimentera les réflexions de la Cnaf pour la création d'un fonds d'investissement social. Nous invitons également à une large diffusion de ce rapport pour inviter à la constitution d'un réseau institutionnel qui devienne un support de dispositifs d'évaluation analogues.

Sandrine Dauphin.

La Caisse nationale de solidarité pour l'autonomie (CNSA)

En menant une étude sur les What Works Centres britanniques, l'Ansa outille un débat important sur les enjeux et les méthodes pour développer une action publique mieux éclairée par des preuves.

La CNSA est très directement concernée par cette question de l'évaluation probante. Elle participe à l'animation de politiques publiques pour lesquelles la demande sociale se transforme rapidement tant dans le champ du handicap que de la perte d'autonomie des personnes âgées. Les réponses apportées se transforment, et il y a un vrai besoin d'évaluation pour l'existant comme pour les réponses émergentes. À titre d'illustration, deux champs d'application des réflexions de l'Ansa sont très spontanément identifiables pour la CNSA.

Le premier est celui du soutien à l'innovation. La CNSA dans sa fonction d'agence, est régulièrement interrogée ou amenée à soutenir des dispositifs ou des interventions qualifiés d'innovants. Elle manque comme nombre d'autres acteurs, de méthode et de référence pour situer ces actions par rapport à des pratiques existantes, pour juger *a priori* de leur pertinence et quand elle souhaite promouvoir des démarches d'évaluation auprès des porteurs de projets, elle constate qu'ils partagent ce relatif défaut de repères méthodologiques. Enfin, dans ce dialogue avec les porteurs de projets, apparaît très vite la question : qui doit évaluer ? En d'autres termes, sur qui repose la « charge de la preuve » ?

Le second concerne les interventions dites « non médicamenteuses ». Le champ médico-social repose sur une large palette d'interventions auprès des personnes âgées en perte d'autonomie ou de personnes handicapées, qui ne sont qu'en partie évaluées et il n'y a pas

What Works Centres britanniques : quels enseignements pour des politiques fondées sur la preuve en France ?

de cadre de référence pour ces pratiques comme il en existe dans le champ médical. Sans condamner l'empirisme et le besoin légitime des acteurs de terrain d'inventer de nouvelles pratiques d'accompagnement psycho-éducatives ou psycho-sociales, il faut trouver collectivement les moyens de faire la preuve qu'il s'agit de bonnes pratiques et soutenir la diffusion de celles qui sont probantes. Et la question se pose régulièrement : suffit-il d'importer des méthodes éprouvées à l'étranger ? Une meilleure qualification de l'impact de telle ou telle intervention dans le contexte national importe tant pour définir les pratiques d'accompagnement que pour repenser les cadres de financement, car aujourd'hui certaines interventions sont financées facultativement faute de preuve, ou au titre de l'aide à la vie quotidienne faute de certitude sur la nature de l'impact.

En proposant d'envisager une transposition adaptée des What Works Centres britanniques, l'Ansa apporte des éléments de questionnement et de compréhension importants pour envisager des dispositifs permettant une validation collective des pratiques, un accompagnement au changement d'échelle ou à la transposition de pratiques validées à l'étranger, un soutien à l'expérimentation à visée de mesure d'impact, et une diffusion des meilleures pratiques.

L'approche thématique des centres britanniques permet de constater que chaque champ d'action publique est plus ou moins « mûr » pour organiser ce travail de mise en ordre des preuves. L'analyse transversale des différents centres a déjà permis à l'Ansa de mettre en lumière les conditions de réussite d'une telle démarche. L'examen des ressources françaises mobilisables dans le domaine de l'expertise et de l'évaluation suggère que nous ne partons pas de rien, loin s'en faut. Tout est cependant à faire en matière de qualification systématique du niveau de preuve, de mise en ordre des connaissances et, plus encore, de diffusion.

La phase 2 du projet de l'Ansa va permettre sur la base de ce premier

rapport fort stimulant, d'envisager de façon très pragmatique ce que pourraient être des « centres de preuve » thématiques. Pour ce qui concerne la CNSA, il faudra choisir un champ d'action publique relatif à l'accompagnement des personnes âgées ou des personnes handicapées. Et nous avons bien noté, forts de l'expérience britannique, qu'il faudra privilégier une thématique bien circonscrite.

Design graphique : Camille Esayan camilleesayan.fr.



PEFC™

10-32-3010 / Imprimé sur papier issu de forêts gérées durablement / pefc-france.org / Axiom Graphic



IMPRIM'VERT®

Ce rapport a été rédigé par Caroline Allard et Ben Rickey, responsables de projet à l'Agence nouvelle des solidarités actives en partenariat avec France Stratégie, le Secrétariat général pour la modernisation de l'action publique (SGMAP), le Commissariat général à l'égalité des territoires (CGET), Santé publique France, la Caisse nationale des allocations familiales (Cnaf) et la Caisse nationale de solidarité pour l'autonomie (CNSA).



Ansa_Asso

www.solidarites-actives.com

Agence nouvelle des solidarités actives

28, rue du Sentier 75002 Paris

Tél. : 01 43 48 65 24

What Works Centres britanniques : **quels enseignements pour des politiques fondées sur la preuve en France ?**