



**FRANCE STRATÉGIE**

ÉVALUER. ANTICIPER. DÉBATTRE. PROPOSER.

# Évaluation d'impact des politiques publiques : que peut apprendre la France des pays les plus avancés ?

L'évaluation d'impact vise à apprécier, à partir de méthodes statistiques, les effets propres d'une politique publique au regard des objectifs qui lui sont assignés. Développée aux États-Unis dès les années 1960, elle apparaît en Europe à la fin des années 1990, portée par la volonté de rationaliser l'action publique et de rendre le processus de décision politique plus transparent.

Alors que la France connaît actuellement, avec retard, un essor des évaluations d'impact, France Stratégie a entrepris de recenser les bonnes pratiques dans cinq des pays considérés comme les plus avancés en la matière : les États-Unis, le Canada, le Royaume-Uni, la Suède et l'Allemagne<sup>1</sup>. Sur la base d'entretiens, de rapports institutionnels et d'articles universitaires, trois facteurs clés favorisant la production et l'utilisation des évaluations dans la décision sont analysés : l'existence de dispositifs formels promouvant les évaluations d'impact, le degré de diffusion et d'influence de ces évaluations sur le public et sur les décideurs, et enfin l'ouverture du milieu administratif aux chercheurs en économie.

La comparaison de la France avec cinq pays parmi les plus avancés montre qu'il n'y a pas de modèle « clé en main » pour encourager l'évaluation d'impact. Ces dernières années, chaque pays a développé son modèle de « politiques publiques fondées sur la preuve » en s'appuyant sur ses propres leviers d'action, au premier rang desquels la vitalité académique, l'impulsion politique et l'accès aux données administratives. Avec l'ouverture du milieu administratif aux acteurs de la recherche aux États-Unis, la formation des agents publics aux méthodes d'évaluation d'impact au Royaume-Uni ou encore la place des agences exécutives d'évaluation des politiques publiques en Suède, le panorama réalisé souligne qu'il existe des déterminants différents de réussite dans chaque pays.

Loin donc de plaider pour l'importation d'un modèle institutionnel unique, les bonnes pratiques repérées à l'étranger dessinent plutôt un ensemble de pistes à envisager et à expérimenter pour soutenir l'élan de l'évaluation d'impact en France. Elles mettent l'accent sur trois enjeux majeurs : la nécessité d'une meilleure articulation entre la demande et la production d'évaluation d'impact ; la définition de principes communs garantissant notamment l'indépendance, la crédibilité et la transparence des évaluations ainsi que le partage des enjeux, des pratiques et des résultats de l'évaluation d'impact au sein d'une alliance réunissant la communauté la plus large possible d'acteurs intéressés.

Plusieurs bonnes pratiques pourraient inspirer le déploiement de l'évaluation d'impact en France :

- aux États-Unis, l'*Intergovernmental Personnel Act Mobility Program* favorise la mobilité de carrière entre l'administration publique et le monde de la recherche ;
- les guides méthodologiques sur l'évaluation d'impact édités par le Trésor britannique et largement diffusés garantissent un cadre commun ;
- au Canada, le titre d'évaluateur accrédité reconnaît des compétences d'évaluation et garantit une déontologie commune ;
- les *What Works Centres* et les *Clearinghouses* des pays anglo-saxons centralisent les résultats des évaluations pour classer les dispositifs publics en fonction de leur efficacité et pour les rendre accessibles au plus grand nombre ;
- au Royaume-Uni, depuis 2015, la *What Works Team* encourage les ministères à publier des *Areas of Research Interest (ARI)* à destination des chercheurs, qui indiquent les domaines où des besoins de preuves scientifiques existent.

1. Cette note offre la synthèse du document de travail intitulé « *Vingt ans d'évaluations d'impact en France et à l'étranger. Analyse comparée des pratiques dans six pays* », également disponible sur le site de France Stratégie. Ce document rassemble lui-même les principaux enseignements tirés de monographies consacrées à l'évaluation des politiques publiques dans six pays : Allemagne, Canada, États-Unis, France, Royaume-Uni, Suède.

Adam Baïz  
Département Économie

Sandra Debu

Rozenn Desplat  
Département Économie

Pierre-Henri Bono  
Conseiller scientifique

Clément Lacouette-  
Fougère  
Sciences Po

Alain Trannoy  
Conseiller scientifique

La *Note de synthèse* est publiée sous la responsabilité éditoriale du commissaire général de France Stratégie. Les opinions exprimées engagent leurs auteurs et n'ont pas vocation à refléter la position du gouvernement.

## INTRODUCTION

Pratiquées aux États-Unis dès les années 1960, les évaluations d'impact des politiques publiques ne sont apparues que récemment en Europe, d'abord au Royaume-Uni à la fin des années 1990 puis dans les autres pays à partir des années 2000. Elles reposent sur le principe qu'il est possible de mesurer *ex post*, au moyen de méthodes statistiques et économétriques rigoureuses, « l'effet propre » d'une intervention publique. À ce titre, elles sont souvent considérées comme la branche quantitative, voire scientifique, de l'évaluation des politiques publiques<sup>2</sup>.

L'évaluation d'impact est aujourd'hui une pratique bien établie au sein des administrations américaines et britanniques. D'autres pays comme l'Allemagne, la Suède ou le Canada ont mis en place des dispositifs institutionnels – structures, financements ou règlements – pour encourager la production d'évaluations d'impact et leur prise en compte dans la gouvernance publique.

En comparaison, la France a connu un développement tardif de ce type d'exercice, même si depuis quelques années on constate un mouvement de rattrapage. Plusieurs comités d'évaluation ont été mis en place au niveau national et souvent confiés à France Stratégie : les comités d'évaluation du CICE (Crédit d'impôt pour la compétitivité et l'emploi), des politiques d'innovation (CNEPI), de la réforme de la fiscalité du capital, des ordonnances Travail, de la loi Pacte, de la stratégie de lutte contre la pauvreté ou du plan Très Haut Débit recourent tous à des degrés divers à des évaluations d'impact. Les services statistiques ministériels pilotent par ailleurs des évaluations d'impact, par exemple la DEPP pour le dédoublement des classes de CP ou la DARES pour le Plan d'investissement dans les compétences<sup>3</sup>. Plus récemment, mais dans une moindre mesure, la Cour des comptes a commencé à faire appel à la communauté académique pour réaliser des évaluations d'impact dans le cadre de ses évaluations de politique publique<sup>4</sup>. Le mouvement devrait s'amplifier en France dans les prochaines années, d'autant que les chercheurs ont accès à un volume toujours plus grand de données administratives *via* le CASD (Centre d'accès sécurisé aux données).

Dans ce contexte d'une pratique en plein essor, France Stratégie a souhaité dresser un panorama complet de l'évaluation d'impact dans cinq pays économiquement

proches et considérés comme les plus avancés en la matière – l'Allemagne, le Canada, les États-Unis, la Suède et le Royaume-Uni – afin d'identifier les bonnes pratiques pouvant inspirer notre pays.

Ce travail fait suite à une analyse bibliométrique publiée en décembre 2018<sup>5</sup>, qui avait permis de quantifier la production d'évaluations d'impact en France par rapport aux autres pays, sur le champ restreint des publications dans des revues à comité de lecture. En prolongement, le panorama proposé ici repose sur la réalisation de six monographies<sup>6</sup> examinant pour chaque pays, dont la France, la place et l'influence des évaluations d'impact dans le processus de décision. En comparant les bonnes pratiques en matière de collaboration entre administration et recherche, de mise en place de dispositifs formels mais aussi de diffusion et d'utilisation des résultats, ce panorama met en évidence une douzaine d'enseignements clés visant à accompagner le développement de l'évaluation d'impact en France et à favoriser son appropriation par les décideurs publics et les citoyens.

## UN DÉVELOPPEMENT RÉCENT À L'INITIATIVE DES EXÉCUTIFS NATIONAUX

L'évaluation d'impact a connu ses premiers développements dans les années 1960 aux États-Unis, dans le contexte des grands programmes de lutte contre la pauvreté initiés par l'administration Johnson. En Europe, c'est le Royaume-Uni qui lance les premières évaluations d'impact des politiques publiques à la fin des années 1990. La plupart des pays du continent – dont la Suède, l'Allemagne et la France – attendront les années 2000 pour voir se concrétiser les premiers exercices nationaux. Dans la quasi-totalité des pays, ces travaux marquent une rupture tant dans les méthodes mobilisées que dans la manière de concevoir les politiques publiques.

D'autres pays, comme le Canada, connaissent cependant une institutionnalisation plus lente des pratiques d'évaluation d'impact. Celles-ci prennent corps progressivement dans un dispositif formalisé de longue date – à partir des années 1970 – pour assurer le contrôle et la performance des programmes fédéraux.

Dans les deux pays précurseurs que sont les États-Unis et le Royaume-Uni, les impulsions politiques émanant du pouvoir exécutif ont joué un rôle décisif dans la recon-

2. L'évaluation d'une politique publique, de manière générale, vise à apprécier l'efficacité de cette politique en comparant les résultats obtenus aux objectifs assignés et aux moyens mis en œuvre. Elle peut prendre diverses formes et recourt parfois simplement à des suivis d'indicateurs ou à des méthodes qualitatives. Par ailleurs, les évaluations d'impact ne doivent pas être confondues avec les études d'impact, qui accompagnent tous les projets de loi depuis 2009 et qui correspondent à une évaluation *ex ante* de la politique considérée.

3. DEPP = Direction de l'évaluation, de la prospective et de la performance, ministère de l'Éducation ; DARES = Direction de l'animation de la recherche, des études et des statistiques.

4. Voir le chapitre V du rapport annuel 2019 de la Cour des comptes.

5. Bono P.-H., Debu S., Desplatz R., Hayet M., Lacouette-Fougère C. et Trannoy A. (2018), « Vingt ans d'évaluations d'impact en France et à l'étranger. Analyse quantitative de la production scientifique », Document de travail, n° 2018-06, France Stratégie, décembre.

6. Chaque monographie s'appuie sur une revue de la littérature scientifique et de la documentation institutionnelle ainsi que sur une série d'entretiens auprès des acteurs nationaux de l'évaluation d'impact.



naissance des évaluations d'impact. C'est le cas du gouvernement Blair au Royaume-Uni avec la création du NICE (National Institute for Health and Care Excellence), premier *What Works Centre* mis en place dans le champ de la santé au tournant des années 2000. Aux États-Unis, c'est l'adoption par l'administration Obama de la loi FEBPA (*Foundations for Evidence-Based Policymaking Act*) qui a servi d'accélérateur, en renforçant l'accès aux données ainsi que la capacité du gouvernement fédéral à produire et à utiliser les éléments de preuve. Ces élans politiques ont surtout permis, dans ces deux pays, d'imposer l'évaluation d'impact comme un instrument incontournable d'aide à la décision gouvernementale et d'amélioration de la transparence vis-à-vis des citoyens dans les principaux domaines de politique publique (santé, social, emploi, éducation, etc.).

L'étude sur un temps long des pratiques dans les six pays permet d'observer que les grandes séquences réformatrices nationales ont souvent été des moments d'accélération de la production d'évaluations. Outre les exemples précités des États-Unis et du Royaume-Uni, c'est le cas avec l'évaluation des réformes Hartz en Allemagne.

En revanche et de manière surprenante, le rôle du Parlement en matière d'évaluation d'impact se révèle modeste dans l'ensemble des pays. La reconnaissance du rôle des parlements nationaux dans l'évaluation des politiques publiques – parfois consacrée dans la Constitution comme en France ou en Suède – ne s'est dans les faits pratiquement jamais traduite par un développement des pratiques d'évaluation d'impact dans ces enceintes<sup>7</sup>.

## LES UNIVERSITÉS ET LES CENTRES DE RECHERCHE, PRINCIPAUX PRODUCTEURS

De par la complexité des méthodes qu'elle mobilise, l'évaluation d'impact est principalement réalisée par des chercheurs en économie. Dans les pays les plus matures, les principaux producteurs sont les départements d'économie des grandes universités nationales. Ainsi, en France, sont en première ligne des structures spécialisées rattachées au monde universitaire, telles que l'Institut des politiques publiques (IPP), J-PAL Europe, le Laboratoire interdisciplinaire d'évaluation des politiques publiques (LIEPP) de Sciences Po, le Centre de recherche en économie et statistique (CREST) ou encore la fédération de recherche Théorie et Évaluation des politiques publiques (TEPP). Ces structures réalisent des évaluations d'impact en réponse à des commandes du gouvernement ou de leur propre initiative.

La vitalité académique en la matière est aussi le fruit d'acteurs indépendants, présents aux côtés des universités,

notamment dans les pays anglo-saxons. Au Royaume-Uni, plusieurs petites structures comme l'Institute for Fiscal Studies (IFS) produisent des travaux d'évaluation tandis qu'aux États-Unis, les ministères et agences fédérales s'appuient régulièrement sur les études du MDRC ou du Mathematica Policy Research Group, deux organismes de recherche indépendants spécialisés dans l'évaluation d'impact. L'Allemagne se distingue quant à elle par la prédominance de ces instituts de recherche indépendants par rapport aux universités. Deux instituts sont ainsi les principaux foyers de production d'évaluations d'impact : l'Institut allemand pour la recherche économique (Deutsches Institut für Wirtschaftsforschung, DIW) et l'Institut d'économie du travail, plus connu sous le nom d'IZA (Forschungsinstitut zur Zukunft der Arbeit).

Dans ce contexte, les évaluations d'impact produites « en propre » par les administrations constituent plus l'exception que la règle au niveau international. Pourtant, en France, l'Insee et les services statistiques ministériels comme la DARES (ministère du Travail), la DREES (ministère de la Santé) ou la DEPP (ministère de l'Éducation nationale) sont des producteurs majeurs d'évaluations d'impact, en raison d'un accès longtemps exclusif aux données administratives et d'une tradition du service statistique public (SSP) à ne pas se cantonner à la production statistique. En Suède, les agences exécutives d'évaluation des politiques publiques, indépendantes des ministères, produisent également des évaluations d'impact et sont pour certaines de véritables organismes de recherche, en particulier l'Institute for Evaluation of Labour Market and Education Policy (IFAU) sur les politiques d'emploi et d'éducation. Au Canada également, de nombreuses évaluations sont produites par les administrations, mais ne mobilisent que rarement des méthodes économétriques de mesure de l'impact.

## LES ADMINISTRATIONS, PRINCIPAUX COMMANDITAIRES

La demande d'évaluations d'impact de politiques publiques provient, dans les pays étudiés, principalement des ministères, en particulier dans les domaines du marché du travail, de l'éducation et de la santé. Les moyens humains et financiers qui y sont consacrés sont cependant souvent très inégaux selon les ministères, même dans les pays les plus avancés. Le choix d'évaluer peut être du ressort de l'administration elle-même (autosaisine), du gouvernement (des ministres), du Parlement ou de la loi (lorsqu'elle en prévoit le principe).

Aucun des pays étudiés dans le cadre de notre benchmark ne possède à ce jour de critères explicites et formalisés permettant de décider le lancement d'une évaluation.

7. En France, depuis la réforme de 2014, l'article 145-7 du Règlement de l'Assemblée nationale prévoit, trois ans après l'entrée en vigueur d'une loi, la présentation par deux députés, dont un issu de l'opposition, d'un rapport d'évaluation sur l'impact de la loi. Jusqu'à présent, les évaluations cherchant à mettre en évidence l'impact causal de la loi ont cependant été peu nombreuses.

Plusieurs critères entrent en ligne de compte, comme le caractère expérimental de la réforme, les enjeux économiques (les programmes d'envergure tels que le Programme des investissements d'avenir en France ont ainsi été adossés à un dispositif d'évaluation), la mobilisation préalable des acteurs de la recherche sur le sujet ou encore le calendrier général d'élaboration et de mise en œuvre de la réforme. La demande sociale d'évaluations émanant de la société civile – Parlement, partenaires sociaux, médias, citoyens – joue cependant un rôle important dans certains pays comme les États-Unis, le Royaume-Uni, la Suède et l'Allemagne. En France, le rôle du Parlement dans l'impulsion des évaluations est plus récent.

Dans les pays anglo-saxons, la demande d'évaluations d'impact est également soutenue par des fondations privées. C'est le cas de l'association Nesta au Royaume-Uni, qui promeut l'utilisation des preuves scientifiques dans la décision publique, notamment grâce à l'organisation de conférences. Aux États-Unis, la Mac Arthur Foundation et la Laura & John Arnold Foundation financent et promeuvent presque exclusivement les travaux reposant sur des méthodes expérimentales mobilisant des RCT (*randomized controlled trials*).

Enfin, dans les six pays étudiés, il existe une société professionnelle de l'évaluation, c'est-à-dire une organisation regroupant les praticiens de la sphère administrative, académique et du privé – en France, la Société française d'évaluation (SFE). Ces sociétés professionnelles ne développent cependant pas toujours des initiatives spécifiques en faveur des évaluations d'impact. Seules la société américaine d'évaluation (American Evaluation Association) et la société allemande d'évaluation (Deutsche Gesellschaft für Evaluation) ont ainsi formulé des propositions visant à renforcer la place de l'évaluation d'impact dans les pratiques et les décisions.

## UN CADRE STRUCTURÉ SEULEMENT AU CANADA ET AUX ÉTATS-UNIS

Un cadre structuré et coordonné peut contribuer à faciliter le recours à l'évaluation d'impact au sein de l'administration. Bien qu'il ne s'agisse pas d'évaluation d'impact, il est à noter que le Canada est le pays ayant le plus fortement structuré la pratique évaluative, au sens large, sur le plan institutionnel. Une directive de 2001 du **Secrétariat du Conseil du Trésor (SCT)**, actualisée en 2009 et 2011, rend obligatoire l'évaluation pour tous les ministères et organismes publics. Elle a donné lieu à la création, dans la plupart des ministères, d'unités chargées de formaliser un plan d'évaluation ministériel (souvent quinquennal) permettant d'identifier, programme par programme, les besoins attendus et les méthodes à considérer. Le Centre d'excel-

lence en évaluation (CEE), placé au sein du SCT, coordonne quant à lui ces travaux et établit chaque année un rapport sur l'activité de ces unités. Périodiquement, il commande aussi des « méta-évaluations » permettant de vérifier la qualité et l'utilité d'ensemble des travaux produits.

Aux États-Unis, l'Office of Management and Budget (OMB) au sein de l'administration de la Maison blanche coordonne également l'évaluation et son utilisation par des directives à destination de l'ensemble des agences fédérales. Certains ministères sont en outre dotés d'unités d'études et d'évaluation particulièrement développées, ayant un rôle avancé de pilotage et de coordination des évaluations. Le ministère du Travail dispose ainsi d'un chef évaluateur (*Chief Evaluation Officer*, CEO) qui a pour mission de concevoir et de superviser les projets d'évaluation du ministère, d'accroître la qualité de ces exercices, d'améliorer l'utilisation des résultats et de faciliter l'accès aux données. Suivant cet exemple, la loi FEBPA de 2018 impose la nomination dans chaque agence fédérale d'un chef évaluateur chargé de définir et de mettre en œuvre une politique d'évaluation. Cette loi prévoit également la mise en place dans toutes les agences de *Learning Agendas* qui recensent les questions politiques et de recherche les plus importantes et les hiérarchisent en fonction de contraintes politiques et budgétaires. Leur élaboration donne généralement lieu à une large consultation de parties prenantes, incluant les hauts responsables de l'administration, les directions en charge des programmes, les chercheurs externes et le Congrès.

Dans les autres pays, la pratique de l'évaluation d'impact est souple et peu coordonnée, ce qui ne l'empêche pas d'être généralisée au Royaume-Uni, en Suède et en Allemagne. Au Royaume-Uni, la What Works Team, hébergée par le Cabinet Office, promeut la pratique des évaluations d'impact au sein des ministères, en plus de coordonner le réseau des *What Works Centres*. En revanche, elle n'assure pas de réelle coordination de l'évaluation au-delà d'une coordination des formations, des actions et des séminaires.

## DES FORMES MULTIPLES DE GOUVERNANCE

En matière de pilotage des évaluations, aucun des pays étudiés n'a établi de règle explicite. Partout, la pratique varie fortement d'une évaluation à l'autre : les instances évaluatives sont quasiment chaque fois constituées de manière *ad hoc*, en fonction du périmètre de la politique ciblée et des questions posées à l'équipe chargée de l'exercice. Cependant, un principe est souvent réaffirmé : celui du pluralisme au sein des instances d'évaluation, notamment dans le choix des membres des comités de suivi ou directement dans le choix des acteurs en charge



de l'évaluation. Du fait des méthodes mobilisées, la présence d'un haut degré d'expertise scientifique est également un trait commun que l'on retrouve dans les instances d'évaluation de tous les pays étudiés.

Aux États-Unis et au Royaume-Uni, l'évaluation de dispositifs structurants est souvent confiée à un consortium d'acteurs (centres de recherche, cabinets d'études et think tanks) afin de pouvoir couvrir des travaux de différente nature (sur la mise en œuvre, la mesure des effets à différents horizons, l'analyse coûts-bénéfices) et généralement conduits sur plusieurs années. L'externalisation y est perçue comme une garantie d'indépendance des évaluateurs vis-à-vis des commanditaires.

En France, France Stratégie, rattachée aux services du Premier ministre, pilote plusieurs évaluations d'impact de politiques publiques pour le compte de comités d'évaluation pluralistes, tels les comités d'évaluation de la loi Pacte et de la stratégie de lutte contre la pauvreté, tous deux installés à l'automne 2019. Ces travaux ont la particularité de s'inscrire dans la durée : à titre d'exemple, le comité d'évaluation du CICE, installé en 2013, a rendu depuis lors six rapports annuels. France Stratégie mène en outre des réflexions sur les méthodes d'évaluation *via* l'élaboration de guides méthodologiques ou l'organisation d'événements et séminaires dédiés.

L'ouverture des instances d'évaluation aux parties prenantes, qu'il s'agisse des opérateurs de la politique évaluée ou de représentants des bénéficiaires, est une dynamique nouvelle qui semble à l'œuvre dans tous les pays et qui va de pair avec les enjeux de communication et de diffusion des évaluations d'impact.

## DES COLLABORATIONS ORIGINALES ENTRE ADMINISTRATION ET RECHERCHE : LES *NUDGE UNITS* ET LES PANELS D'EXPERTS

Des modalités novatrices rendent également possible la collaboration entre les administrations et les chercheurs sur les questions d'expérimentation et d'évaluation d'impact. C'est le cas des *Nudge units* qui ont essaimé dans la quasi-totalité des administrations étudiées. Ces unités spécialisées dans l'application des sciences comportementales aux politiques publiques apportent généralement leur soutien aux administrations pour tester de nouvelles solutions et en évaluer rigoureusement les résultats à l'aide de méthodes expérimentales (RCT). Il s'agit par exemple de la Behavioural Insights Team (BIT)

au Royaume-Uni, de l'Office of Evaluation Sciences (OES) aux États-Unis ou de l'équipe en charge du développement des sciences comportementales appliquées aux politiques publiques au sein de la Direction interministérielle de la transformation publique (DITP) en France. La DITP a par exemple confié en 2018 à plusieurs équipes de recherche, dont le CEPREMAP et la Behavioural Insights Team britannique, l'accompagnement et l'évaluation de sept projets d'application des sciences comportementales aux politiques publiques proposés par des administrations<sup>8</sup>.

Une autre modalité offrant des points de contact entre chercheurs et administrations repose sur la mise en place de panels d'experts pour accompagner la conception et la mise en œuvre de programmes publics. La création en 2015 au Royaume-Uni du Cross-Government Trial Advice Panel (CG-TAP), composé d'experts de l'administration et du monde académique, en fournit un bon exemple. Le CG-TAP incite les différents ministères à mettre en œuvre des évaluations d'impact de leurs programmes, et les y accompagne en formant et en conseillant techniquement les agents concernés. Le ministère du Logement, des Communautés et du Gouvernement local (MHCLG) a par exemple bénéficié de l'expertise de ce panel en mars 2018 pour une RCT étudiant l'impact d'un programme d'enseignement de l'anglais sur l'intégration sociale<sup>9</sup>.

## LA PRÉSENCE DE PROFILS ACADÉMIQUES DANS L'ADMINISTRATION : UNE CONSTANTE DES PAYS LES PLUS AVANCÉS

La possibilité pour des chercheurs d'intégrer la sphère administrative est un marqueur fort des pays disposant des pratiques d'évaluation les plus matures. Aux États-Unis, au Royaume-Uni, en Allemagne et en Suède, il est fréquent que des universitaires ou des chercheurs soient recrutés au sein de l'administration centrale pour des postes exigeant une expertise scientifique.

L'administration américaine est certainement celle qui a poussé le plus loin cette possibilité avec le programme de mobilité IPA (*Intergovernmental Personnel Act*) qui permet les échanges de personnel qualifié entre les agences fédérales et les universités et centres de recherche. Des fonctionnaires et des chercheurs peuvent ainsi changer de structure pour une période déterminée (jusqu'à quatre ans) avec la garantie de retrouver leur poste, ou un poste équivalent, dans leur organisation d'origine. Grâce au programme IPA, l'Office of Evaluation Sciences (OES) au sein de la General Services Administration (GSA), spécialisé dans l'application des résultats des sciences comportementales

8. Les projets retenus portent sur des sujets divers : promouvoir une consommation raisonnée des antibiotiques, favoriser le recours aux paiements dématérialisés, réduire l'impact des stéréotypes et lutter contre les discriminations à l'embauche, améliorer les politiques de prévention du manque de sommeil chez les enfants de CP, encourager la pratique sportive, favoriser l'adoption de modes de chauffage éco-responsables, création d'un dispositif d'information des consommateurs pour promouvoir l'économie circulaire. Voir le site du ministère.

9. Voir le blog *What Works* sur le site du gouvernement britannique.

aux politiques fédérales, est ainsi composé d'une équipe interdisciplinaire de chercheurs issus des principales universités américaines (Harvard University, Reed College, Washington University, etc.).

## DES HAUTS FONCTIONNAIRES FORMÉS AUX MÉTHODES D'ÉVALUATION D'IMPACT

La sensibilisation et la formation des hauts fonctionnaires à l'évaluation d'impact est un autre levier de développement des pratiques. La plupart des pays proposent des sessions de formation de ce type, mais qui s'adressent essentiellement aux agents publics spécialisés, c'est-à-dire en position de piloter ou de mener des évaluations d'impact. Le Royaume-Uni est le seul pays à avoir inclus l'évaluation dans les standards de compétences utilisés pour le recrutement, l'évaluation individuelle et l'évolution professionnelle des agents impliqués dans la conception des politiques publiques (environ 18 500 personnes). Ces standards de compétences, appelés *Policy Profession Standards*, définissent les compétences permettant d'utiliser à bon escient les preuves dans la conception des politiques publiques. Afin d'appliquer ces standards, les ministères britanniques font régulièrement appel à des prestataires extérieurs, comme l'Alliance for Useful Evidence (A4UE) ou la What Works Team, pour former leurs fonctionnaires à l'utilisation des preuves scientifiques.

## DES CORPUS ET STANDARDS MÉTHODOLOGIQUES AUX ÉTATS-UNIS ET AU ROYAUME-UNI

Les publications éditées depuis une quinzaine d'années par le Trésor britannique offrent la collection la plus complète de méthodologies d'évaluation à destination des administrations. Ensemble, le *Green Book*, le *Magenta Book* et le complément *Quality in policy impact*, couvrent toutes les dimensions de l'évaluation des politiques publiques, de l'analyse de leur mise en œuvre à la mesure de leur impact en passant par les estimations coûts-bénéfices. La portée de ce corpus méthodologique s'étend à toute la fonction publique britannique : ces publications sont intégrées aux programmes d'évaluation des ministères ainsi que dans les standards de compétences des fonctionnaires. Elles servent également de référence lors de la mise en œuvre ou la commande d'évaluations d'impact par les ministères<sup>10</sup>.

Plus largement, dans les pays anglo-saxons, la réflexion sur la qualité des évaluations passe par la diffusion de standards de preuve (« *standards of evidence* ») qui permettent de juger de la robustesse des résultats des évaluations en fonction des méthodes utilisées. Les « *standards of evidence* » édités par l'association britannique Nesta ou l'échelle de Maryland (*Maryland Scientific*

*Methods Scale*) sont les plus courants. Les méthodes expérimentales sont considérées comme les plus fiables pour produire des résultats robustes, devant les méthodes quasi expérimentales. Ces standards sont également mobilisés pour réaliser des revues systématiques (synthèses des enseignements tirés des évaluations répondant à un standard de qualité élevé), construire des bases de données de résultats probants issus des évaluations, ou tout simplement mettre en évidence les manques en matière d'évaluation rigoureuse dans tel ou tel domaine d'action publique.

Au Canada, une initiative singulière a été développée en 2009 par la Société canadienne d'évaluation : le titre d'évaluateur accrédité. Ce programme de certification professionnelle est construit sur trois piliers : un code de déontologie, des normes professionnelles et un ensemble de 36 compétences théoriques et pratiques. Les compétences attendues ne mettent pas l'accent sur une discipline en particulier mais sur les dimensions professionnelles générales. Le titre n'est délibérément pas prescriptif du point de vue des méthodes et des techniques utilisées (quantitatives, qualitatives) afin d'encourager la pluridisciplinarité des évaluations. Sur la base de ce référentiel, le Canada compte, en 2018, 395 évaluateurs accrédités dont la liste est disponible sur demande à la société canadienne de l'évaluation ; les chercheurs en économie y sont relativement minoritaires.

## DES FINANCEMENTS DÉDIÉS À L'ÉVALUATION ET UN SOUTIEN FINANCIER PAR LES ACTEURS PRIVÉS DANS LES PAYS ANGLO-SAXONS

Dans tous les pays étudiés, les évaluations d'impact sont d'abord financées par les crédits budgétaires récurrents des administrations qui décident elles-mêmes de leur répartition entre leurs différentes activités d'études et d'expertise.

Dans certains cas cependant, des fonds spécifiques peuvent être réservés pour l'évaluation d'impact d'un programme ou d'un dispositif, ou pour soutenir l'activité d'une instance d'évaluation en particulier. Aux États-Unis par exemple, jusqu'à 0,5 % des crédits alloués au titre de la loi *Every Student Succeeds Act* de 2015<sup>11</sup> peuvent être mobilisés pour son évaluation. De même, depuis 2012, le Chief Evaluation Office au ministère du travail est doté de financements représentant jusqu'à 0,75 % du montant total des crédits affectés aux agences opérationnelles de ce même ministère. En France, quelques dispositifs ou programmes ont ainsi bénéficié de fonds spécifiques pour leur évaluation, par exemple le Programme des investissements d'avenir (PIA) lancé en 2010 ou plus récemment le Plan d'investissement dans les compétences (PIC).

10. En France, la pratique des guides méthodologiques est restreinte aux évaluations socioéconomiques *ex ante* des projets d'infrastructures publiques.

11. Voir sur le site du ministère américain de l'éducation.



Dans les pays anglo-saxons, aux côtés des fonds publics, les fondations privées participent aussi au financement d'évaluations d'impact. Cette contribution prend la forme de dotations pérennes attribuées à des structures spécialisées, sur le modèle du Big Lottery Fund au Royaume-Uni, ou de financements directs de travaux d'évaluation, comme pour la fondation Laura and John Arnold aux États-Unis.

## TOOLKITS, CLEARINGHOUSES ET RÉPERTOIRES : DES OUTILS DE CAPITALISATION ET DE DIFFUSION DES RÉSULTATS

Afin de faciliter la diffusion des résultats et leur appropriation par les utilisateurs, plusieurs pays développent des dispositifs plus élaborés que les simples synthèses des résultats d'évaluations. Les initiatives les plus remarquables sont les *toolkits* mis en place par les *What Works Centres* au Royaume-Uni et les *Clearinghouses* aux États-Unis<sup>12</sup>. Ces deux outils ont en commun de rassembler les études disponibles dans un domaine de la politique publique, de les synthétiser dans un format lisible par les non-spécialistes et de les classer selon les méthodes qu'elles utilisent et les résultats auxquels elles conduisent. Ils facilitent notamment la comparaison des interventions publiques en fonction de leurs effets, de leurs coûts et de la robustesse des preuves scientifiques associées. Ces démarches innovantes ont commencé à essaimer dans les autres pays. Grâce au financement de l'université McMaster (Ontario), le Canada a depuis 2017 développé un moteur de recherche recensant les politiques sociales et leurs évaluations (près de 15 000 références). En France, Santé Publique France a lancé en 2018, avec le soutien de la Direction générale de la Santé, un répertoire des interventions efficaces ou prometteuses – et de leurs évaluations – dans le domaine de la prévention et de la promotion de la santé.

## DES APPROCHES INCITATIVES QUI ÉMERGENT AUX ÉTATS-UNIS ET AU ROYAUME-UNI

Comment s'assurer que la connaissance produite soit la plus utile pour la décision publique ? C'est la question à laquelle plusieurs pays cherchent à répondre en mettant l'accent, ces dernières années, sur l'enjeu de l'utilisation

des résultats des évaluations par les décideurs et les citoyens. Depuis 2013, le gouvernement britannique a par exemple confié à la *What Works Team* l'organisation régulière d'« audits sur les preuves » visant à identifier ministère par ministère les domaines de politique publique où le recours aux éléments objectifs de preuves issus de la recherche pourrait être renforcé. Cette démarche s'accompagne de suggestions pour améliorer la diffusion et la prise en compte des évaluations d'impact.

Le développement depuis une dizaine d'années des obligations « à impact social » dans les pays anglo-saxons constitue une autre modalité privilégiée pour inciter les décideurs et opérateurs des politiques publiques à mettre en application les résultats de la recherche. Reposant sur un mécanisme d'incitation financière, les *Social Impact Bonds* (SIB) apparus en 2007 au Royaume-Uni permettent ainsi à des investisseurs privés de financer des interventions délivrées par des organisations non gouvernementales ou à but non lucratif, et de récupérer leur investissement, avec intérêt, si et seulement si ces interventions atteignent leurs objectifs. Les interventions financées par SIB font dans certains cas l'objet d'une évaluation d'impact *ex post* qui permet d'établir si le dispositif a ou non produit les effets attendus et donc si l'investisseur privé sera remboursé par le décideur public. La version américaine de ce type de dispositif, connue sous le nom de *Pay for Success* (PFS), a par exemple permis de soutenir des projets dans les domaines de l'éducation, de la petite enfance ou de la lutte contre la récidive. Le premier projet a été lancé en septembre 2012 par la ville de New York, afin de réduire la récidive des jeunes détenus de la prison de Rikers Island (programme *Adolescent Behavioral Learning Experience-ABLE*)<sup>13</sup>. Dans le même esprit, sous l'administration Obama, ont été déployées à partir de 2010 des « *tiered-evidence grants* », également appelées « *innovation fund grants* » : les subventions versées par le gouvernement fédéral aux autorités locales et aux organisations à but non lucratif pour la mise en œuvre de programmes sociaux affichent un montant d'autant plus important que ces dispositifs sont appuyés par des preuves solides de leur efficacité. Ces mécanismes financiers incitent les acteurs à choisir des dispositifs efficaces pour bénéficier de subventions plus élevées, tout en garantissant que les projets soient évalués rigoureusement.

12. Voir par exemple le *Crime Reduction Toolkit* du What Works Centre for Crime Reduction britannique, et la *What Works Clearinghouse* du Department of Education américain.

13. Ce programme était financé par Goldman Sachs à hauteur de 9,6 millions de dollars. La ville de New York n'était tenue d'effectuer aucun paiement si le taux de récidive était abaissé de moins de 8,5 % grâce au programme. Au vu des évaluations disponibles avant la mise en place du programme, il était vraisemblable d'attendre une réduction du taux de récidive d'environ 11 % du fait de ce programme. Toutefois, les résultats d'évaluation obtenus à l'aide d'une méthode quasi expérimentale ont montré que le projet n'avait pas eu d'effet sur le taux de récidive au terme de la première année. Sur la base de ces résultats, Goldman Sachs a décidé de mettre fin au projet, et la ville de New York n'a eu à procéder à aucun paiement (voir le document de travail portant sur les États-Unis).

## CONCLUSION

Tous les pays étudiés ont renforcé leurs pratiques d'évaluation sans que se dégagent un cadre commun ou un modèle « clé en main ». Le Canada et les États-Unis disposent d'un cadre général de coordination contrairement aux pays européens, même si la pratique d'évaluation est généralisée au Royaume-Uni, en Suède et en Allemagne. La France se distingue par le rôle des services statistiques ministériels dans la production d'évaluations d'impact. Dans les autres pays, ce sont plutôt des institutions académiques qui s'en chargent. Partout cependant l'administration est le principal commanditaire, même si le rôle de la société civile est croissant, y compris sous forme de financements privés dans certains pays. Nulle part il n'existe de règle en matière de pilotage des évaluations, cependant le principe du pluralisme des parties prenantes est souvent affirmé. Des modalités novatrices voient le jour, comme des unités spécialisées dans l'application des sciences comportementales aux politiques publiques ou des panels d'experts pour accompagner la conception et la mise en œuvre des programmes publics.

Les bonnes pratiques repérées à l'étranger dessinent un ensemble de pistes à envisager et à expérimenter pour soutenir l'élan de l'évaluation d'impact en France. À cette fin, trois enjeux majeurs se dégagent : une meilleure articulation entre demande et production d'évaluations d'impact ; un cadre commun garantissant l'indépendance, la crédibilité et la transparence des évaluations ; enfin l'utilité d'une alliance réunissant la communauté la plus large possible d'acteurs intéressés, afin de partager les enjeux, les pratiques et les résultats de l'évaluation d'impact<sup>14</sup>.

Mots clés : évaluation d'impact, politiques publiques, *ex ante*, *ex post*, comparaison internationale, administration

14. Les auteurs remercient Vincent Aussilloux pour son implication active, leurs collègues de France Stratégie ainsi que les administrations et les experts interrogés dans le cadre de ce projet.

Directeur de la publication : Gilles de Margerie, commissaire général ; directeur de la rédaction : Cédric Audenis, commissaire général adjoint  
secrétaires de rédaction : Olivier de Broca, Valérie Senné ; impression : France Stratégie ; dépôt légal : décembre 2019 ; N° ISSN 2609-701X  
contact presse : Matthias Le Fur, directeur du service Édition-Communication-Événements, 01 42 75 61 37, matthias.lefur@strategie.gouv.fr

## RETROUVEZ LES DERNIÈRES ACTUALITÉS DE FRANCE STRATÉGIE SUR :



[www.strategie.gouv.fr](http://www.strategie.gouv.fr)



[@Strategie\\_Gouv](https://twitter.com/Strategie_Gouv)



[france-strategie](https://www.linkedin.com/company/france-strategie)



[FranceStrategie](https://www.facebook.com/FranceStrategie)



[@FranceStrategie\\_](https://www.instagram.com/FranceStrategie_)



[StrategieGouv](https://www.youtube.com/StrategieGouv)

Institution autonome placée auprès du Premier ministre, France Stratégie contribue à l'action publique par ses analyses et ses propositions. Elle anime le débat public et éclaire les choix collectifs sur les enjeux sociaux, économiques et environnementaux. Elle produit également des évaluations de politiques publiques à la demande du gouvernement. Les résultats de ses travaux s'adressent aux pouvoirs publics, à la société civile et aux citoyens