

Évaluer les expérimentations : Enjeux et postures de l'évaluateur/rice

Depuis une décennie, la conduite d'expérimentations connaît un développement sans précédent en France et a pris des formes multiples ^{CONSEIL D'ÉTAT, 2019}. L'évaluation des expérimentations a donné lieu à de nombreuses controverses et débats méthodologiques. Cependant, malgré la diversité des pratiques, il existe certains enjeux transversaux autour desquels les pratiques évaluatives se retrouvent.

Cette note propose une synthèse de ces enjeux propres à l'évaluation des expérimentations, ainsi que des pistes méthodologiques pour concevoir une démarche évaluative à même d'y répondre.

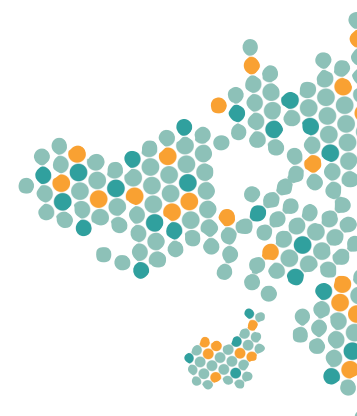
QU'EST-CE QU'UNE EXPÉRIMENTATION ?

L'expérimentation recouvre une **diversité de pratiques** autant dans le type de projets expérimentés que dans la manière de les évaluer. Selon la définition proposée récemment par Huitema²⁰¹⁸, l'expérimentation se caractérise par une :

mise à l'essai temporaire et contrôlée, dans des conditions de terrain, d'une innovation pertinente pour une politique donnée, et qui produit des enseignements utiles aux décisions publiques à venir.

Cette définition fait apparaître **4 critères de base** discriminants pour savoir si l'on est en présence d'une expérimentation :

- ▶ **La mise à l'essai d'une intervention susceptible de résoudre un problème**, ici appelée innovation, mais qui peut être comprise de façon très large : « une idée, une pratique, ou un objet perçu comme nouveau par un individu ou une autre unité d'adoption » ^{ROGERS 1983}.
- ▶ **La durée** : une expérimentation est par nature temporaire, et généralement sa durée est fixée à l'avance, même si elle peut être ensuite raccourcie ou prolongée ;
- ▶ **Les conditions de mise en œuvre**, qui peuvent être plus ou moins contrôlées (mise en œuvre directive ou progressive, dans son entièreté ou partielle, avec comparaison de modalités différentes ou pas, etc.) ;
- ▶ **La volonté, enfin, de produire des enseignements utiles à la décision publique** : identifier, amender, valider des hypothèses sur les effets des interventions (des « théories d'action »), ou identifier des



conditions de mise en œuvre qui permettront la traduction des intentions politiques sur le terrain.


Plusieurs approches de production de connaissance plus ou moins formelles peuvent être déployées dans le cadre des expérimentations. L'évaluation des politiques publiques en est la plus structurée.

ÊTRE ASSOCIÉ EN AMONT DE LA MISE EN ŒUVRE DU PROJET

L'expérimentation étant conçue pour produire des enseignements, il est très utile de **réfléchir à son évaluation avant même le déploiement sur le terrain**. Fréquemment, l'équipe d'évaluation est en contact avec les porteurs de projet dès l'amont, lors de la conception, ou pendant les premières phases du déploiement du projet. Cette particularité est une opportunité pour que l'évaluateur puisse **co-construire un système de suivi du projet** opérationnel et adapté au projet dès ses premiers moments de déploiement.


Une particularité de l'expérimentation est que dans certains cas **l'équipe d'évaluation** peut elle-même concevoir le projet ou **être étroitement associée à sa conception** :


- ▶ **Sur le contenu même du projet** : l'évaluateur peut proposer de tester certains mécanismes, ou modes d'intervention, à la lumière de son expérience ou expertise dans le domaine d'intervention. Il peut aider le porteur de projet à **identifier des besoins de connaissance** (*Knowledge Gap*) pouvant être renseignés par l'expérimentation.
- ▶ **Sur les modalités de mise en œuvre du projet** : l'évaluateur peut demander au porteur de projet de déployer son dispositif de manière à **s'inscrire dans un protocole scientifique particulier**. Par exemple, si l'évaluateur veut avoir recours à la méthode expérimentale par assignation aléatoire, il peut demander au porteur de projet de tirer au sort en amont les participants à l'expérimentation afin de comparer un groupe bénéficiaire avec un groupe de non-bénéficiaires.

 Cette « ingérence » possible de l'évaluateur dans le projet a été identifiée comme une **brèche potentielle de la neutralité** de la posture de l'évaluateur. Ce dernier étant « **embarqué** » dans le projet et risquant de devenir juge et partie de l'expérimentation DUBOIS ET PODEVIN 2011. Cette situation demande alors de penser différemment le rôle de l'évaluateur afin qu'il se place davantage dans une position d'accompagnateur de l'innovation (approche de l'évaluation évolutive) que de jugement rétrospectif sur les effets du projet.

ÉVALUER UN PROJET INNOVANT AMENÉ À ÉVOLUER

Une expérimentation se définit par l'aspect innovant du projet qu'elle propose. De fait, lorsqu'une idée ou une théorie n'a jamais été appliquée auparavant (ou pas dans ce contexte, ou avec ces publics, ou de cette façon, etc.), il est très probable que les **hypothèses initiales qui guidaient sa mise en œuvre évoluent** en cours de route. C'est la raison d'être même d'une expérimentation de tester ces hypothèses. Mais faut-il en tirer des conséquences dans la mise en œuvre des projets ?

 Dans l'expérimentation le système de suivi et évaluation occupe une place centrale dans la conduite du projet. On parle d'ailleurs de plus en plus de système de suivi, évaluation et apprentissage (en anglais MEL : Monitoring & Evaluation & Learning)

 Pour comprendre les conditions d'application de la méthode expérimentale par assignation aléatoire dans le cadre des expérimentations consulter : Bernard Godel et Évelyne Serverin, « L'expérimentation sociale aléatoire en France en trois questions », Travail et Emploi, 135 | 2013, 57-71.

À ce sujet, plusieurs écoles s'affrontent : certain·es évaluateur·rices vont maintenir les modalités initiales, même si elles s'avèrent initialement un échec, pour pouvoir les écarter définitivement d'expérimentations à venir. D'autres au contraire vont se servir de l'expérimentation pour tâtonner dans la mise en œuvre, jusqu'à obtenir si possible une mise en œuvre satisfaisante.

Les modifications peuvent typiquement porter sur :

- ▶ **Les publics ciblés** (s'il s'avère que ce public était mal défini au départ);
- ▶ **La manière de mobiliser les publics** (en cas de faible recours au projet, les porteurs peuvent envisager de nouveaux modes de sollicitation);
- ▶ **Les actions conduites** (supprimer des actions dont les acteurs ont du mal à se saisir, modifier des actions à la marge après leur premier test par les opérateurs, ajouter de nouvelles actions);
- ▶ **Les partenaires envisagés** (de nouveaux partenaires peuvent être identifiés ou des partenariats envisagés peuvent s'avérer inopérants);
- ▶ etc.

Dans certains cas, c'est la théorie sous-jacente elle-même qui change en cours d'expérimentation.

Dans tous les cas, cela complique le travail de l'évaluateur/riche, si celui-ci est pensé comme la vérification, en fin d'expérimentation, de l'atteinte des objectifs fixés initialement et devenus obsolètes avec ces évolutions, ou lorsque la rigueur de sa démonstration dépend (comme dans le cas d'une démarche longitudinale ou contrefactuelle) de la stabilité de l'intervention évaluée (voir la fiche :). Nous proposons alors deux stratégies distinctes, selon la situation.

Dans le premier cas, l'évolution du projet lors de la mise en œuvre **est assumée par les porteurs de projet dès sa conception**. Alors le rôle de l'évaluateur.rice peut être **d'organiser, d'animer le processus de (re)conception progressive du projet avec les opérateurs et porteurs de projet** (cette approche est appelée developmental evaluation, ou l'évaluation évolutive). Il/elle mobilise alors ses outils (structuration du questionnement autour de la cohérence, la pertinence, les premiers résultats, collecte de données, médiation et animation des échanges entre partie prenantes). Un premier système de suivi peut être alors déployé pour accompagner l'apprentissage autour du projet mais la structuration d'un système de suivi-évaluation finalisé uniquement une fois le projet stabilisé. Le projet peut être stabilisé au cours de l'expérimentation (au bout d'un an par exemple), voire uniquement à la fin de l'évaluation.

Dans un second cas de figure, les porteurs n'avaient pas initialement envisagé que des évolutions pourraient être nécessaires. L'équipe d'évaluation peut alors adopter plusieurs stratégies :

- ▶ **Alerter sur les implications de ces évolutions sur le déploiement de l'évaluation** (difficulté d'évaluer l'impact d'un projet si celui-ci est modifié) et solliciter le porteur de projet afin qu'il cadre la mise en œuvre par les opérateurs de terrain (proposer un kit d'intervention, formaliser les protocoles d'intervention) ;

☛ Pour une présentation détaillée des principes de la Developmental evaluation, voir : Michael Quinn Patton (2010) « Developmental Evaluation: Applying Complexity Concepts to Enhance Innovation and Use ».

Identifier, formuler et tester des hypothèses causales

QUADRANT CONSEIL

Identifier, formuler et tester des hypothèses causales

Plus rigoureux que dans d'autres approches d'évaluation, développer des hypothèses causales permet de clarifier les mécanismes d'action et de mieux comprendre les processus de changement. Elles sont essentielles pour concevoir des interventions efficaces et évaluer leur impact réel.

Les hypothèses causales sont des affirmations sur la manière dont les choses fonctionnent, en particulier :
 • les effets que peut produire l'intervention (par exemple, pour un public cible) dans un contexte donné (situation, lieu, temps, etc.)
 • les effets sur les bénéficiaires (dans le cas d'un projet social, l'ensemble des effets sur les bénéficiaires envisagés)
 • le lien entre l'intervention et les résultats attendus (en particulier, les résultats à court terme et les résultats à long terme)

Une évaluation d'impact s'intéresse à tester l'existence d'une relation de cause à effet entre une intervention et un impact recherché. Mais, au-delà de cet impact, il y a une multitude de facteurs susceptibles d'influencer l'impact. C'est pourquoi, en intervenant sur une hypothèse causale, c'est à dire sur un mécanisme de la façon dont se manifeste une relation de cause à effet.

Il s'agit donc de mieux connaître les mécanismes de l'intervention et de son impact sur les bénéficiaires, et de tester l'existence d'une relation de cause à effet entre l'intervention et l'impact recherché.

LA THÉORIE DU CHANGEMENT ET SES COMPOSANTES

Le diagramme de la théorie du changement rassemble l'ensemble des hypothèses qui sont faites quant à la façon dont une intervention agit pour produire des résultats attendus. C'est un processus continu qui évolue au fil du temps et qui est constamment révisé et amélioré.

Plus l'intervention est complexe, plus elle nécessite une compréhension plus détaillée des hypothèses causales. Les hypothèses causales peuvent être de plusieurs types. Elles peuvent porter sur :

- sur le lien causal (le chemin) : « Grâce à l'intervention, les bénéficiaires acquerront les compétences nécessaires ». Ces hypothèses peuvent porter sur :

- ▶ **Cibler l'évaluation sur une partie fixe du projet**, opter pour une évaluation partielle du projet en se concentrant sur l'une de ses composantes (concentrer l'évaluation sur la sollicitation des publics, ou une action d'accompagnement).

Dans les deux cas, dans le cadre d'une expérimentation, un rôle central de l'évaluateur sera de **documenter et d'explicitier les modifications apportées** au projet tout au long de la période.

PRODUIRE DE LA CONNAISSANCE DANS DES DÉLAIS CONTRAINTS SUR UN NOUVEAU PROJET

Estimer les résultats de l'expérimentation

Dans le cadre des expérimentations, il est souvent demandé à l'équipe d'évaluation de produire des connaissances sur le programme, sa mise en œuvre, ses résultats et ses impacts sur les bénéficiaires dans des délais très contraints. Lorsqu'il y a un **hiatus entre le temps nécessaire à l'obtention des résultats et celui de l'expérimentation**, l'enjeu est d'adapter la réponse à ce qui peut être attendu dans les délais impartis.

Il faut alors définir quelles conséquences peuvent être attendues sur les cibles de l'expérimentation, ce qui peut être fait en s'appuyant sur la **théorie du changement de l'intervention**. Si celle-ci n'a pas été explicitée lors de la conception de l'expérimentation, l'équipe d'évaluation doit débiter son travail par l'élaboration de celle-ci avec les parties prenantes de l'expérimentation. Cette théorie du changement permet de décomposer le contenu des actions menées, les hypothèses que l'on fait sur la manière dont les publics cibles s'en saisissent, et in fine sur comment ces résultats contribuent à générer les impacts escomptés.

Cette démarche permet dans le cadre de l'expérimentation d'apporter des éléments de connaissance sur les premiers éléments de cette théorie en renseignant les hypothèses liées à la mise en œuvre ainsi qu'aux premiers résultats sur les publics.

■ Par exemple, si l'on évalue l'impact d'une formation proposée par Pôle Emploi sur le retour à l'emploi durable des travailleurs en situation de handicap, il est possible que, dans le cadre temporel de l'expérimentation, l'équipe d'évaluation ne puisse observer l'effectivité du retour à emploi, et sa durabilité pour tous les participants.

L'évaluation peut néanmoins apporter des éléments sur les étapes précédentes de la théorie du changement :

- ▶ Sur la **mise en œuvre** (Les formations sont-elles en cohérence avec les autres dispositifs proposés à ces publics? Répondent-elles à leur besoin à celles des employeurs? Leurs conditions d'organisation sont-elles adaptées?)

- ▶ Sur les **conditions préalables** à l'atteinte des impacts (Ces formations génèrent-elles une montée en compétences des stagiaires? Ont-ils les moyens de les mobiliser dans leur recherche d'emploi?)
- ▶ Et enfin sur les **premiers impacts** (les stagiaires ont-ils été en contact avec davantage d'employeurs? Ont-ils eu plus de succès? Quels types de contrats leur ont été proposés? Ces propositions correspondaient-elles à leurs attentes?)


Établir la validité des résultats dans d'autres contextes

L'évaluation des expérimentations doit aussi relever le défi de produire des connaissances sur la mise en œuvre et les effets d'un projet tout en précisant le périmètre de leur validité dans d'autres contextes.

Une expérimentation est conduite dans un contexte particulier (un territoire, un échantillon des publics etc...) et a vocation à générer des connaissances mobilisables dans d'autres contextes et dans une perspective de changement d'échelle.

Plusieurs changements d'échelle peuvent être envisagés, principalement :

- ▶ **L'essaimage** (répliquer le projet sur d'autres territoires ou auprès d'autres publics)
- ▶ **La généralisation** (transformer le projet en dispositif de droit commun déployé sur l'ensemble du territoire)
- ▶ **La fusion** (utiliser une partie des enseignements du projet pour la programmation d'autres dispositifs)

 - Pour aller plus loin sur les différentes modalités de changement d'échelle et leurs enjeux voir le rapport du SGMAP/CGET : [Ensemble accélérons : accompagner les acteurs de l'innovation dans leur changement d'échelle, décembre 2015.](#)



Les résultats de l'évaluation doivent donc prendre la forme d'un « code de la route » définissant leur périmètre de validité PAWSON ET TILLEY 2004. Ainsi on peut conclure qu'un dispositif déployant des formations auprès des demandeurs d'emploi en situation de handicap a contribué à insérer dans l'emploi les $\frac{3}{4}$ des participants. On précisera alors :

- ▶ **Qu'est ce qui « marche »** dans les formations (l'échange entre pairs permet de se sentir moins isolé et de se motiver, la montée en compétence est source de confiance pour postuler, etc.);
- ▶ **Pour quels publics** (les publics bénéficiant le plus de ces formations sont les plus éloignés de l'emploi, notamment les jeunes);
- ▶ **Dans quelles conditions** (les effets de ce dispositif reposent sur une connaissance fine par les conseillers des besoins des publics ciblés, l'établissement au préalable d'une relation de confiance, et la prédisposition des entreprises locales à l'insertion des publics en situation de handicap).

Tirer des apprentissages utiles

Cette préoccupation est commune à toutes les approches d'évaluation ^{PATTON 1997; WEISS 1998; LEVITON ET HUGHES 1981}, néanmoins elle est **centrale dans l'évaluation des expérimentations puisque c'est bien l'objectif de production d'enseignements qui légitime la conduite même du projet.**

Cependant, plusieurs types d'apprentissages peuvent être organisés dans le cadre de l'évaluation d'expérimentation et **les différentes parties prenantes de l'expérimentation** (évaluateurs, porteur de projet, financeurs/décideurs publics) peuvent avoir des attentes distinctes en la matière.

Nous identifions **3 types d'apprentissages qui peuvent parfois cohabiter** dans le cadre d'une expérimentation. Cependant, si ceux-ci ne sont pas explicités entre les différentes parties prenantes d'une expérimentation ils peuvent être source de « malentendus d'apprentissage » pouvant perturber la bonne conduite de l'expérimentation ainsi que la production de connaissances utiles.

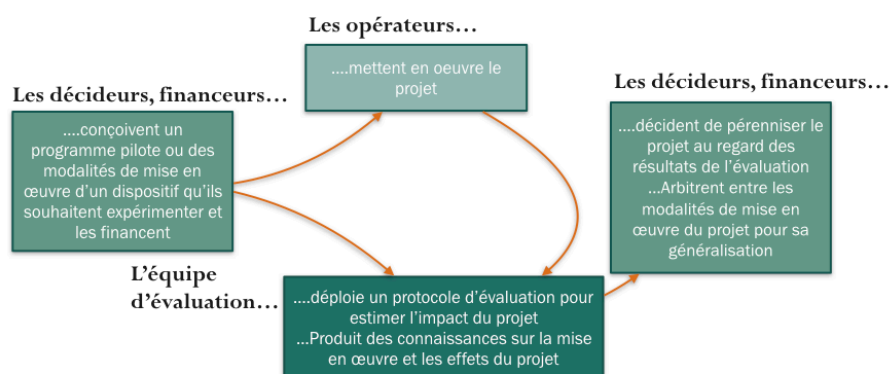
L'apprentissage programmatique

Ce mode d'apprentissage désigne la vision traditionnelle de l'expérimentation, au sens où celle-ci est **conduite afin de « tester » une version pilote d'un programme** et d'apporter des informations sur sa mise en œuvre et ces effets permettant ainsi de statuer sur la pertinence de son déploiement.

Ce type d'apprentissage requiert pour garantir l'utilité de l'évaluation:

- ▶ D'identifier au préalable les **destinataires du rapport** (décideurs);
- ▶ De concevoir une évaluation permettant de **répondre aux interrogations précises** des acteurs de la prise de décision;
- ▶ De produire des résultats dans les **délais de la prise de décision.**

Ce processus d'apprentissage programmatique, ainsi que le rôle de chacune des parties prenantes de l'expérimentation peut être représenté schématiquement de la manière suivante :



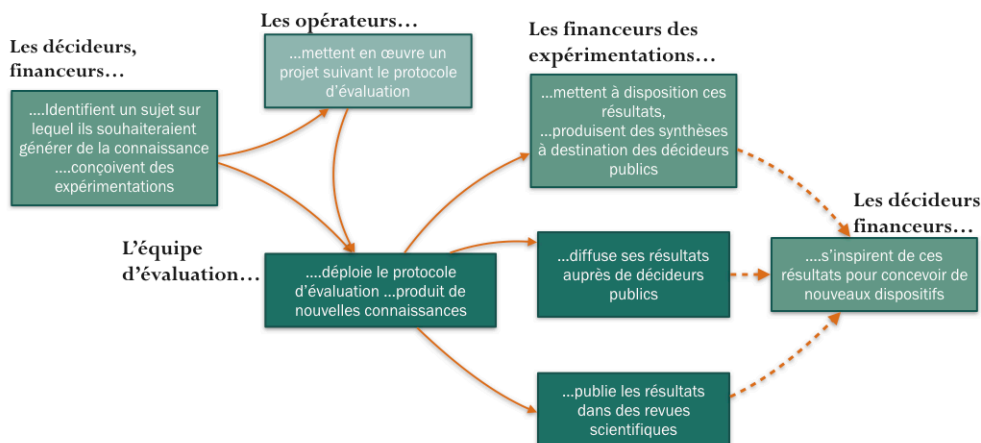
L'apprentissage conceptuel

Ce mode d'apprentissage défini par Carol Weiss ^{WEISS 1999} porte sur des connaissances, des analyses, des perspectives nouvelles sur un enjeu donné que peut générer l'expérimentation, « **qui modifie la façon dont les décideurs considèrent un problème et envisage des solutions** » ^{DELARUE, NAUDET, ET SALIVAT 2009}, sans toutefois qu'elles soient reliées à une prise de décision spécifique.

Ce type d'apprentissage requiert pour garantir l'utilité de l'évaluation :

- ▶ La conduite d'une évaluation déployant des protocoles s'approchant des **standards d'une production scientifique**;
- ▶ Une identification en amont entre les parties prenantes de l'expérimentation des composantes du projet sur lesquelles seront produits les éléments de connaissance;
- ▶ Un respect par les opérateurs du terrain des **contraintes potentielles du protocole d'évaluation**;
- ▶ Que certaines des parties prenantes de l'expérimentation (l'équipe d'évaluation ou les porteurs de projet) **diffusent les connaissances produites** afin que des décideurs, financeurs s'en saisissent.

Ce processus d'apprentissage conceptuel, ainsi que le rôle de chacune des parties prenantes de l'expérimentation peut être représenté schématiquement de la manière suivante :



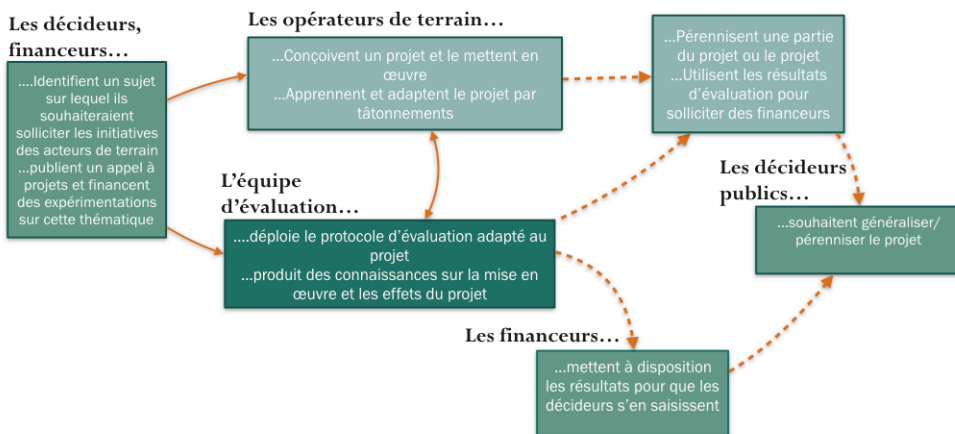
L'apprentissage pragmatique

Ce mode d'apprentissage est issu de la pratique concrète du projet expérimenté. Il s'inscrit dans l'idée pragmatiste que **pratique et connaissance sont consubstantiellement liées** : « *Ce n'est que lorsqu'une solution proposée est mise en œuvre que la valeur de la solution proposée peut être établie. [...] Pour acquérir le savoir, nous avons besoin d'action.* » ^{DEWEY, 2005 ; ELLIOTT, 2007; GILLBERG ET VO 2011}. Ces apprentissages font appel à l'intelligence des acteurs de terrain et à leur tâtonnement pragmatique ^{PRAWAT 1999}.

Ce type d'apprentissage requiert pour garantir l'utilité de l'évaluation :

- ▶ Un **protocole d'évaluation dédié à l'accompagnement de l'apprentissage** des opérateurs de terrain ou porteurs de projet au moins dans un premier temps de l'expérimentation (voir plus haut la notion d'évaluation évolutive) ;
- ▶ L'organisation de **boucles d'apprentissage courtes** pendant l'expérimentation afin d'ajuster le projet chemin faisant ;
- ▶ La possibilité pour les opérateurs de terrain d'être décisionnaire dans les ajustements de pratiques issus des apprentissages pragmatiques.

Ce processus d'apprentissage pragmatique, ainsi que le rôle de chacune des parties prenantes de l'expérimentation peut être représenté schématiquement de la manière suivante :



Ces trois types d'apprentissages ne sont pas exclusifs les uns des autres. D'ailleurs, **la plupart des expérimentations poursuivent des objectifs d'apprentissage hybrides**. L'enjeu est alors, pour l'équipe d'évaluation, d'amener les parties prenantes à expliciter ces différents objectifs d'apprentissage et à préciser, via le référentiel d'évaluation, lesquels seront nourris par l'évaluation et de quelle manière.

BIBLIOGRAPHIE

Conseil d'Etat. 2019. « Les expérimentations: comment innover dans la conduite des politiques publiques? » Les études du Conseil d'Etat. Paris.

Delarue, Jocelyne, Jean David Naudet, et Véronique Sauvat. 2009. « Les évaluations sont-elles utiles? Revue de la littérature sur “connaissances et décisions” ». 03. Série Notes méthodologiques Ex Post. Département de la Recherche, Division Evaluation et Capitalisation, Agence Française de Développement.

Dubois, Jean-Marie, et Gérard Podevin. 2011. « Entre indépendance et connivence, la délicate posture de l'évaluateur embarqué ». Bref, Céreq, 1 (294).

Gillberg, Claudia, et Linh Chi Vo. 2011. « Approche pragmatiste de la connaissance et de l'apprentissage dans les organisations ». Management & Avenir 43 (3): 410.

Huitema, Dave, Andrew Jordan, Stefania Munaretto, et Mikael Hildén. 2018. « Policy Experimentation: Core Concepts, Political Dynamics, Governance and Impacts ». Policy Sciences 51 (2): 143–159.

Leviton, Laura C., et Edward F.X. Hughes. 1981. « Research On the Utilization of Evaluations: A Review and Synthesis ». Evaluation Review 5 (4): 525–48.

Patton, Michael Quinn. 1997. Utilization-Focused Evaluation: The New Century Text. Thousand Oaks; London; New Delhi: Sage Publications.

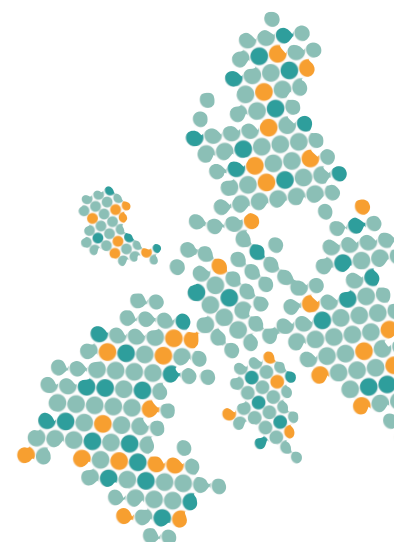
Pawson, Ray, et Nick Tilley. 2004. « Realist Evaluation ». British Cabinet Office.

Prawat, Richard S. 1999. « Dewey, Peirce, and the Learning Paradox ». American Educational Research Journal 36 (1): 47–76.

Rogers, Everett M. 1983. Diffusion of Innovations. 3rd ed. New York: London: Free Press; Collier Macmillan.

Weiss, Carol H. 1998. « Have We Learned Anything New About the Use of Evaluation? » American Journal of Evaluation 19 (1): 21–33.

Weiss, Carol H. 1999. « The Interface between Evaluation and Public Policy ». Evaluation 5 (4): 468–86.



Pour en savoir plus

Retrouvez des guides, des outils sur l'évaluation sur notre site <http://quadrant-conseil.fr/ressources.php>

Impliquez-vous dans l'évaluation en vous rendant sur le site de la SFE : www.sfe-asso.fr



Cette fiche a été rédigée dans le cadre d'une commande de Pôle Emploi. Réutilisation possible sans permission en citant la source: « Quadrant Conseil, 2020 »