

Typologie des approches d'évaluation

Trois perspectives : méthode, cycle, instrument

Note de référence pour la conception du cœur de Themiros — Septembre 2026

Cadre de classification

Les approches d'évaluation se classent selon trois dimensions largement indépendantes. Toute évaluation réelle combine les trois.

Dimension	Question	Catégories principales
Méthodologique	Comment établit-on les constats et la causalité ?	Expérimentale, quasi-expérimentale, fondée sur la théorie, participative, économique, mixte
Cycle de la politique ou du projet	Quand évalue-t-on ?	Ex-ante, évaluabilité, formative, à mi-parcours, finale, ex-post
Instrument	Qu'évalue-t-on ?	Projet, programme, politique, stratégie, secteur, thème, portefeuille, institution

Exemple de combinaison : évaluation *ex-post* (cycle) de l'impact d'un *programme* (instrument) par un dispositif *quasi-expérimental* (méthode).

Deux distinctions transversales complètent le cadre : **qui conduit l'évaluation**, et **selon quels critères elle juge l'intervention**. Elles sont traitées en fin de document.

1. Perspective méthodologique : comment établir la causalité et la valeur

La ligne de partage principale oppose **l'attribution** et **la contribution**. L'attribution demande si l'intervention a causé le changement, et dans quelle mesure. La contribution demande comment

et pourquoi l'intervention y a plausiblement contribué, parmi d'autres facteurs (**Mayne 2001, 2012 ; Befani & Mayne 2014**).

Le choix du design dépend des attributs de l'intervention et des données disponibles, selon la matrice proposée par **Stern et al. (2012)**. La distinction entre aspects *compliqués* et *complexes* structure l'ensemble de cette section (**Rogers 2008**).

1.1 Approches d'attribution (contrefactuelles)

Famille	Principe	Atouts	Limites
Expérimentale	Affectation aléatoire des unités aux groupes traité et témoin. Toute différence de résultat est attribuable à l'intervention.	Validité interne maximale.	Coût élevé. Contraintes éthiques et politiques. Peu adaptée aux politiques nationales et aux interventions complexes.
Quasi-expérimentale	Construction d'un contrefactuel crédible sans randomisation.	Applicable <i>a posteriori</i> et à l'échelle d'un territoire.	Dépend de la qualité des données et des hypothèses d'identification.

Méthode	Famille	Principe	Référence
Essai contrôlé randomisé (ECR)	Expérimentale	Randomisation individuelle. Variantes : ECR en grappes, essais multi-bras ou factoriels.	Duflo, Glennerster & Kremer 2007
Doubles différences	Quasi-expérimentale	Compare l'évolution dans le temps entre groupe traité et groupe non traité.	Bertrand, Duflo & Mullainathan 2004
Régression sur discontinuité	Quasi-expérimentale	Exploite un seuil d'éligibilité, par exemple un score de pauvreté.	Imbens & Lemieux 2008
Appariement sur score de propension	Quasi-expérimentale	Associe chaque bénéficiaire à un non-bénéficiaire statistiquement similaire.	Rosenbaum & Rubin 1983

Méthode	Famille	Principe	Référence
Contrôle synthétique	Quasi-expérimentale	Construit un « jumeau » pondéré à partir d'unités de comparaison. Utile pour les politiques régionales ou nationales.	Abadie, Diamond & Hainmueller 2010
Variables instrumentales	Quasi-expérimentale	Isole une source de variation exogène de l'exposition à l'intervention.	Angrist, Imbens & Rubin 1996
Séries temporelles interrompues	Quasi-expérimentale	Détecte une rupture de tendance au moment de l'intervention.	Bernal, Cummins & Gasparrini 2017

Vue d'ensemble des exigences de données et des cas d'usage : **Gertler et al. (2016)**.

Références — §1.1

Abadie, A., Diamond, A. & Hainmueller, J. (2010). Synthetic Control Methods for Comparative Case Studies. *JASA*, 105(490), 493-505. doi:10.1198/jasa.2009.ap08746

Angrist, J. D., Imbens, G. W. & Rubin, D. B. (1996). Identification of Causal Effects Using Instrumental Variables. *JASA*, 91(434), 444-455. doi:10.1080/01621459.1996.10476902

Bernal, J. L., Cummins, S. & Gasparrini, A. (2017). Interrupted time series regression... *Int. J. Epidemiology*, 46(1), 348-355. doi:10.1093/ije/dyw098

Bertrand, M., Duflo, E. & Mullainathan, S. (2004). How Much Should We Trust Differences-in-Differences Estimates? *QJE*, 119(1), 249-275. doi:10.1162/003355304772839588

Duflo, E., Glennerster, R. & Kremer, M. (2007). Using Randomization in Development Economics Research: A Toolkit. *NBER TW 333*. nber.org/papers/t0333

Gertler, P. J. et al. (2016). *Impact Evaluation in Practice*, 2e éd. Banque mondiale & BID. doi:10.1596/978-1-4648-0779-4

Imbens, G. W. & Lemieux, T. (2008). Regression discontinuity designs: A guide to practice. *J. Econometrics*, 142(2), 615-635. doi:10.1016/j.jeconom.2007.05.001

Rosenbaum, P. R. & Rubin, D. B. (1983). The central role of the propensity score... *Biometrika*, 70(1), 41-55. doi:10.1093/biomet/70.1.41

1.2 Approches de contribution (fondées sur la théorie)

Ces approches ouvrent la « boîte noire » en testant la chaîne causale décrite dans une théorie du changement (**Weiss 1997 ; White 2009**). Elles conviennent aux interventions complexes, aux

petits nombres de cas et aux politiques sans contrefactuel possible (**Rogers 2008 ; Stern et al. 2012**).

Approche	Origine	Question centrale	Référence
Analyse de contribution	Mayne	L'intervention a-t-elle contribué de façon plausible au changement observé ?	Mayne 2001, 2012
Évaluation réaliste	Pawson et Tilley	Qu'est-ce qui fonctionne, pour qui, dans quel contexte et pourquoi ?	Pawson & Tilley 1997, 2004
Traçage de processus (<i>process tracing</i>)	Sciences politiques	Quels mécanismes relient l'intervention aux effets observés ?	Collier 2011 ; IEG Banque mondiale
Analyse qualitative comparée (QCA)	Ragin	Quelles combinaisons de conditions produisent le résultat ?	Ragin 2008 (<i>hors MVP</i>)
Méthode d'élimination générale	Scriven	Quelles explications alternatives peuvent être écartées ?	Scriven 2008
Récolte des effets (<i>outcome harvesting</i>)	Wilson-Grau	Quels changements ont eu lieu et comment l'intervention y a-t-elle contribué ? Démarche inductive.	Wilson-Grau & Britt 2012/2013
Changement le plus significatif	Davies et Dart	Quels changements les parties prenantes jugent-elles les plus importants ? Démarche inductive.	Davies & Dart 2005

La combinaison *process tracing* + analyse de contribution est exécutée par le moteur Themiros (**Befani & Mayne 2014**), avec mise à jour bayésienne pour la confiance corpus (**Befani & Stedman-Bryce 2017**).

Références — §1.2

Befani, B. & Mayne, J. (2014). Process Tracing and Contribution Analysis. *IDS Bulletin*, 45(6), 17-36. doi:10.1111/1759-5436.12110

Befani, B. & Stedman-Bryce, G. (2017). Process Tracing and Bayesian Updating for impact evaluation. *Evaluation*, 23(1), 42-60. doi:10.1177/1356389016654584

Collier, D. (2011). Understanding Process Tracing. *PS: Political Science and Politics*, 44(4), 823-830. doi:10.1017/S1049096511001429

Davies, R. & Dart, J. (2005). *The 'Most Significant Change' (MSC) Technique: A Guide to Its Use*. man.de.co.uk

Independent Evaluation Group, Banque mondiale. *Process Tracing Method for Program Evaluation*. ie.g.worldbankgroup.org

Mayne, J. (2001). Addressing attribution through contribution analysis. *Canadian Journal of Program Evaluation*, 16(1), 1-24. doi:10.3138/cjpe.016.001

Mayne, J. (2012). Contribution analysis: Coming of age? *Evaluation*, 18(3), 270-280. doi:10.1177/1356389012451663

Pawson, R. & Tilley, N. (1997). *Realistic Evaluation*. Sage. — (2004). *Realist Evaluation*. betterevaluation.org

Ragin, C. C. (2008). *Redesigning Social Inquiry: Fuzzy Sets and Beyond*. University of Chicago Press. ISBN 978-0-226-70275-9.

Scriven, M. (2008). A Summative Evaluation of RCT Methodology... *Journal of MultiDisciplinary Evaluation*, 5(9), 11-24. doi:10.56645/jmde.v5i9.160

Weiss, C. H. (1997). Theory-based evaluation: Past, present, and future. *New Directions for Evaluation*, 76, 41-55. doi:10.1002/ev.1086

White, H. (2009). Theory-based impact evaluation. *J. Development Effectiveness*, 1(3), 271-284. doi:10.1080/19439340903114628

Wilson-Grau, R. & Britt, H. (2012/2013). *Outcome Harvesting*. Ford Foundation. usaidlearninglab.org

1.3 Autres familles

Famille	Approches	Usage	Référence
Participative et orientée usage	Évaluation axée sur l'utilisation (Patton). Évaluation développementale. Évaluation par l'autonomisation (<i>empowerment</i>).	Programmes innovants ou adaptatifs. Appropriation des constats par les parties prenantes.	Patton 2008 (<i>hors MVP</i>)
Économique	Analyse coûts-bénéfices (ACB). Analyse coût-efficacité (ACE). Analyse coût-utilité. Retour social sur investissement (SROI). Analyse multicritère.	Rapport entre coûts et résultats. L'analyse multicritère s'impose quand les bénéfices ne sont pas monétisables, cas fréquent en adaptation climatique.	OCDE 2021 ; Gertler et al. 2016
Études de cas	Cas unique, cas multiples, cas comparés.	Compréhension approfondie d'un contexte.	Yin (études de cas, hors MVP)

Famille	Approches	Usage	Référence
Systémique	Approches sensibles à la complexité. Cartographie des systèmes.	Interventions à effets non linéaires et à acteurs multiples.	Rogers 2008
Synthèse de données probantes	Revue systématique. Méta-analyses. Revues rapides.	Capitalisation transversale à partir d'évaluations existantes.	—
Méthodes mixtes	Combinaison de méthodes quantitatives et qualitatives.	Norme actuelle : estimation quasi-expérimentale de l'effet, complétée par une explication des mécanismes fondée sur la théorie.	White 2009 ; Stern et al. 2012

Références — §1.3

OCDE (2021). *Applying Evaluation Criteria Thoughtfully*. OECD Publishing, Paris. doi:10.1787/543e84ed-en

Patton, M. Q. (2008). *Utilization-Focused Evaluation*, 4e éd. Sage.

Rogers, P. J. (2008). Using Programme Theory to Evaluate Complicated and Complex Aspects of Interventions. *Evaluation*, 14(1), 29-48. doi:10.1177/1356389007084674

Stern, E., Stame, N., Mayne, J., Forss, K., Davies, R. & Befani, B. (2012). *Broadening the range of designs and methods for impact evaluations*. DFID Working Paper 38. gov.uk

2. Perspective cycle : quand évaluer

Chaque étape du cycle de la politique ou du projet appelle un type d'évaluation et une finalité propres.

Étape	Type d'évaluation	Finalité principale
Conception	Évaluation ex-ante. Étude de faisabilité (<i>appraisal</i>). Étude d'impact. Analyse d'impact de la réglementation (AIR).	Justifier et concevoir l'intervention.

Étape	Type d'évaluation	Finalité principale
Avant lancement	Évaluation de l'évaluabilité. Étude de référence (<i>baseline</i>).	Vérifier que l'intervention est évaluable. Fixer les valeurs de référence.
Mise en œuvre	Suivi. Évaluation formative ou de processus. Évaluation en temps réel (courante dans l'humanitaire).	Suivre l'avancement. Améliorer l'exécution.
Mi-parcours	Revue ou évaluation à mi-parcours.	Réorienter. Réallouer les ressources.
Clôture	Évaluation finale. Rapport d'achèvement (par exemple l'ICR de la Banque mondiale).	Apprécier les résultats au regard des objectifs.
Après clôture	Évaluation ex-post. Évaluation d'impact. Évaluation de la durabilité.	Mesurer les effets durables. Redevabilité. Apprentissage.
Transversal	Méta-évaluation. Évaluation de synthèse.	Apprécier la qualité des évaluations. Agréger les enseignements.

Distinction	Contenu
Formative / sommative	La formative améliore l'intervention en cours d'exécution. La sommative juge sa valeur globale, généralement à la fin.
« Évaluer d'abord » (Better Regulation, UE)	L'évaluation ex-post des règles existantes précède toute nouvelle analyse d'impact ex-ante.

Références — §2

OCDE (2021). *Applying Evaluation Criteria Thoughtfully*. doi:10.1787/543e84ed-en

Commission européenne. *Better Regulation Guidelines*. commission.europa.eu

3. Perspective instrument : qu'évaluer

L'objet évalué s'élargit progressivement du projet à la politique. Plus il s'élargit, moins la randomisation est faisable (**Stern et al. 2012**).

Objet	Périmètre	Remarque méthodologique
Projet	Intervention unique, bornée dans le temps, aux extrants définis.	Terrain privilégié des méthodes contrefactuelles.
Programme	Ensemble de projets liés partageant des objectifs communs.	Combine souvent plusieurs méthodes.
Politique publique	Politique, loi ou réglementation, souvent d'échelle nationale.	Contrefactuel difficile. Approches fondées sur la théorie et contrôle synthétique fréquents.
Stratégie	Cadre stratégique d'une organisation ou d'un gouvernement.	Analyse de cohérence et de positionnement.
Secteur	Secteur entier (agriculture, santé), y compris les approches sectorielles.	Multiplicité d'acteurs et d'instruments.
Thème	Thème transversal (genre, adaptation climatique) à travers plusieurs interventions.	Démarche de synthèse.
Programme-pays ou régional	Ensemble de l'appui d'un bailleur à un pays ou une région.	Pertinence et cohérence au premier plan.
Portefeuille	Groupe d'investissements ou d'opérations.	Analyse de performance agrégée.
Institution ou organisation	Performance, pertinence et gouvernance d'une organisation.	Critères organisationnels.
Instrument spécifique	Modalité d'aide ou instrument de politique : appui budgétaire (approche en trois étapes du CAD de l'OCDE), financement basé sur les résultats, subventions, taxes.	Cadre méthodologique propre à chaque instrument.
Bilan de qualité (fitness check, UE)	Groupe de politiques liées.	Cohérence, chevauchements, lacunes.

4. Distinctions transversales

Critère de classement	Catégories
Qui évalue	Évaluation interne. Évaluation externe ou indépendante. Évaluation conjointe (plusieurs bailleurs ou partenaires). Auto-évaluation. Revue par les pairs.
Critères appliqués	Critères du CAD de l'OCDE : pertinence, cohérence (ajoutée en 2019), efficacité, efficience, impact, durabilité.

Les définitions des six critères et leur application raisonnée sont précisées par **OCDE (2021)**.

Références — §4

OCDE (2021). *Applying Evaluation Criteria Thoughtfully*. OECD Publishing, Paris. doi:10.1787/543e84ed-en

5. Articulation des trois dimensions

Le choix de l'approche dépend de quatre facteurs.

Facteur	Implication
La question évaluative	« Combien ? » appelle un dispositif expérimental ou quasi-expérimental. « Comment et pourquoi ? » appelle une approche fondée sur la théorie.
La nature de l'intervention	Les projets simples et les politiques complexes ou adaptatives relèvent de méthodes différentes (Rogers 2008).

Facteur	Implication
La disponibilité des données	Une situation de référence exploitable et un groupe de comparaison élargissent les options.
Le budget, le calendrier et l'usage prévu	Ces contraintes bornent le dispositif réalisable.

À mesure que l'objet passe du projet au programme, puis à la politique, la randomisation devient moins faisable. Les approches fondées sur la théorie, les méthodes mixtes et les synthèses deviennent alors la norme (**Stern et al. 2012 ; White 2009**).

Bibliographie générale

Socle théorique et méthodologique

Befani, B. & Mayne, J. (2014). Process Tracing and Contribution Analysis. *IDS Bulletin*, 45(6), 17-36. doi:10.1111/1759-5436.12110

Befani, B. & Stedman-Bryce, G. (2017). Process Tracing and Bayesian Updating for impact evaluation. *Evaluation*, 23(1), 42-60. doi:10.1177/1356389016654584

Collier, D. (2011). Understanding Process Tracing. *PS: Political Science and Politics*, 44(4), 823-830. doi:10.1017/S1049096511001429

Davies, R. & Dart, J. (2005). *The 'Most Significant Change' (MSC) Technique: A Guide to Its Use*. man de.co.uk

Mayne, J. (2001). Addressing attribution through contribution analysis. *Canadian Journal of Program Evaluation*, 16(1), 1-24. doi:10.3138/cjpe.016.001

Mayne, J. (2012). Contribution analysis: Coming of age? *Evaluation*, 18(3), 270-280. doi:10.1177/1356389012451663

OCDE (2021). *Applying Evaluation Criteria Thoughtfully*. OECD Publishing, Paris. doi:10.1787/543e84ed-en

Pawson, R. & Tilley, N. (1997). *Realistic Evaluation*. Sage. — (2004). *Realist Evaluation*. betterevaluation.org

Ragin, C. C. (2008). *Redesigning Social Inquiry: Fuzzy Sets and Beyond*. University of Chicago Press. ISBN 978-0-226-70275-9.

Rogers, P. J. (2008). Using Programme Theory to Evaluate Complicated and Complex Aspects of Interventions. *Evaluation*, 14(1), 29-48. doi:10.1177/1356389007084674

Scriven, M. (2008). A Summative Evaluation of RCT Methodology... *Journal of MultiDisciplinary Evaluation*, 5(9), 11-24. doi:10.56645/jmde.v5i9.160

Stern, E., Stame, N., Mayne, J., Forss, K., Davies, R. & Befani, B. (2012). *Broadening the range of designs and methods for impact evaluations*. DFID Working Paper 38. gov.uk

Weiss, C. H. (1997). Theory-based evaluation: Past, present, and future. *New Directions for Evaluation*, 76, 41-55. doi:10.1002/ev.1086

White, H. (2009). Theory-based impact evaluation. *J. Development Effectiveness*, 1(3), 271-284. doi:10.1080/19439340903114628

Wilson-Grau, R. & Britt, H. (2012/2013). *Outcome Harvesting*. Ford Foundation. usaidlearninglab.org

Méthodes contrefactuelles

Abadie, A., Diamond, A. & Hainmueller, J. (2010). Synthetic Control Methods for Comparative Case Studies. *JASA*, 105(490), 493-505. doi:10.1198/jasa.2009.ap08746

Angrist, J. D., Imbens, G. W. & Rubin, D. B. (1996). Identification of Causal Effects Using Instrumental Variables. *JASA*, 91(434), 444-455. doi:10.1080/01621459.1996.10476902

Bernal, J. L., Cummins, S. & Gasparrini, A. (2017). Interrupted time series regression... *Int. J. Epidemiology*, 46(1), 348-355. doi:10.1093/ije/dyw098

Bertrand, M., Duflo, E. & Mullainathan, S. (2004). How Much Should We Trust Differences-in-Differences Estimates? *QJE*, 119(1), 249-275. doi:10.1162/003355304772839588

Duflo, E., Glennerster, R. & Kremer, M. (2007). Using Randomization in Development Economics Research: A Toolkit. *NBER TW 333*. nber.org/papers/t0333

Gertler, P. J., Martinez, S., Premand, P., Rawlings, L. B. & Vermeersch, C. M. J. (2016). *Impact Evaluation in Practice*, 2e éd. Banque mondiale & BID. doi:10.1596/978-1-4648-0779-4

Imbens, G. W. & Lemieux, T. (2008). Regression discontinuity designs: A guide to practice. *J. Econometrics*, 142(2), 615-635. doi:10.1016/j.jeconom.2007.05.001

Rosenbaum, P. R. & Rubin, D. B. (1983). The central role of the propensity score... *Biometrika*, 70(1), 41-55. doi:10.1093/biomet/70.1.41

Hors MVP

Independent Evaluation Group, Banque mondiale. *Process Tracing Method for Program Evaluation*. ie.g.worldbankgroup.org

Patton, M. Q. (2008). *Utilization-Focused Evaluation*, 4e éd. Sage.