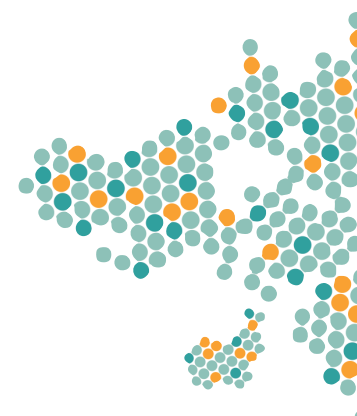


Quels systèmes d'évaluation pour quels usages de l'évaluation ?

Comment concevoir ou reconcevoir un système d'évaluation qui soit utile aux organisations ? La mise en place de dispositifs formalisés d'évaluation représente un travail important, en termes de mise en place de processus et de rôles spécifiques, de montée en capacité, de négociation avec les parties prenantes internes et externes... **Et pourtant, trop souvent, ces systèmes qui ont demandé tant d'effort sont jugés peu utiles et finissent par ne représenter pour les uns et les autres qu'un fardeau supplémentaire.** En effet, il ne suffit pas de produire des évaluations pour qu'elles soient utilisées ! Dans cette note, nous proposons de repartir des usages et d'usages possibles de l'évaluation pour réfléchir à la forme que les systèmes d'évaluation peuvent prendre. **Six critères sont principalement mis en avant : l'orientation interne ou externe, le focus stratégique ou opérationnel, les types d'usage recherchés, le positionnement dans l'organisation, la direction de transformation suivie, et le type de profil des évaluateurs/rices.**



DE L'ÉVALUATION INTERNE AUX SYSTÈMES D'ÉVALUATION

Une riche littérature propose depuis plus de 50 ans des réflexions sur l'évaluation interne, depuis J. Wholey jusqu'à A. Love (1991). R. Rist et N. Stame (2006) ont redéfini dans les années 2000 la notion de fonction évaluative, pour envisager une dimension systémique et parler plutôt de systèmes d'évaluation.

Les systèmes d'évaluation :


- ▶ sont conçus au travers de **questions d'usages et d'influence** plutôt que de la production d'évaluations
- ▶ n'envisagent pas l'évaluation comme une question ponctuelle, mais comme un **flux continu d'information** en lien avec les pratiques dans l'organisation
- ▶ permettent de penser **la fonction évaluative au sein de ces flux : les commander, les modeler, les redistribuer, s'y laisser porter, y résister, etc.**

Dans cette logique, il est admis qu'aucune évaluation seule ne peut répondre à tous les besoins d'information qui existent dans une organisation, d'une part parce que les besoins sont généralement multiformes, et d'autre part parce que le temps de l'évaluation et celui de l'action et de la décision ne sont jamais réellement synchrones. **On apprend rarement d'une évaluation seule,**

mais au travers de processus évaluatifs ou de l'utilisation d'informations à des moments où elles sont utiles. Il faut donc partir de ces besoins pour agir sur les flux d'information (en « remontant le courant »), mais aussi savoir diriger l'information vers celles et ceux qui en ont besoin (en « suivant le courant »).

En ce sens, toute réflexion sur la mise en place d'un système d'évaluation commence par une réflexion sur les origines et les destinations des flux :

- ▶ **Quels usages peuvent être faits de démarches évaluatives**, à quels moments, pour quoi faire ? Que doivent-ils influencer en priorité ?
- ▶ **Quelles sources d'information, quelles pratiques peuvent contribuer le mieux à ces usages ?** Cela inclut des sources d'information internes, mais non utilisées (par exemple dans le cas de fonctionnement en silo, ou de barrières institutionnelles) et externes.
- ▶ **Quelle place y a-t-il pour une intermédiation entre cette « offre » et cette « demande » ?** Avec notamment des questions de gestion de la connaissance, mais aussi de sa « traduction » en direction des utilisateurs/rices.

 -Penser en termes de systèmes amène à considérer toutes les sources de connaissance, théoriques ou empiriques, sans se limiter aux données de suivi ou d'évaluation.

LES SYSTÈMES D'ÉVALUATION EN RÉPONSE AUX TENSIONS INTERNES OU EXTERNES

L'évaluation est rarement une pratique « naturelle » des organisations. Souvent, la mise en place de systèmes d'évaluation (au-delà de démarches ponctuelles) répond plutôt à des tensions ou intérêts externes ou internes, qui diffèrent d'une organisation à une autre (pression du financeur, contractualisation, nouvel exécutif, nouveau cadre managérial...).

Un système d'évaluation conçu pour répondre à une pression externe sera ainsi souvent organisé de façon à assurer son indépendance, pour que les usagers externes puissent avoir confiance dans les résultats – limitant de facto les interactions avec le reste de l'organisation. Inversement, un système qui serait une émanation de pratiques professionnelles internes aura des chances d'être plus imbriqué dans l'administration, mais serait potentiellement soucieux de ne pas attirer l'attention extérieure.

Cette question originelle affecte la façon dont les systèmes d'évaluation sont constitués et dont ils peuvent donner lieu à des usages effectifs, parfois longtemps après que la demande initiale a disparu. Il n'est pas impossible de faire virer de bord de tels systèmes – par exemple passer d'une évaluation indépendante à usage de redevabilité envers les financeurs à une évaluation à usage d'apprentissage interne, suscitée par une demande interne – mais cela reste compliqué. En fait, il n'est pas rare que les systèmes d'évaluation disparaissent lorsque la pression ou l'intérêt qui les avait suscités s'est évanouie – surtout en l'absence d'usagers effectifs, ou lorsque ces usagers ne sont pas assez influents dans la sphère de l'action publique. Cela donne alors l'occasion de remettre en place un système sur de nouvelles bases.

Le simple fait d'être tourné vers l'interne ou l'externe ne préjuge pas de l'usage effectif qui est fait des évaluations. Pour cela, il convient de prendre en compte et de s'inscrire dans des pratiques existantes – et/ou d'engager un travail de long terme de formation ou de changement des pratiques professionnelles de façon à ce que les usagers intègrent l'évaluation comme une

composante à part entière de leurs métiers. Ce travail sous-jacent apparaît ainsi comme une fonction importante des systèmes d'évaluation, qui contribue également à leur survie sur le long terme.

QUELLES FINALITÉS ET USAGES CONCRETS POUR LES SYSTÈMES D'ÉVALUATION ?

Tant la forme générale que l'activité des systèmes d'évaluation dépendent largement des finalités qui lui sont attribuées. Pour en parler, Olejniczak et al. (2023) ont proposé de distinguer deux grands axes, d'un côté la distinction classique entre redevabilité (extérieure) et apprentissage (interne), et d'autre part le ciblage opérationnel ou stratégique, ce qui donne quatre grandes raisons d'être pour un système d'évaluation :

- ▶ la redevabilité au regard des objectifs stratégiques
- ▶ la redevabilité vis-à-vis de la bonne mise en œuvre
- ▶ l'apprentissage stratégique : qu'est-ce qui marche et pourquoi ?
- ▶ l'apprentissage opérationnel : comment mieux mettre en œuvre les actions menées ?

Il est possible cependant de penser ces systèmes d'évaluation à partir d'usages plus concrets. T. Delahais, A. Devaux-Spatarakis et F. Jégou (2023) ont ainsi mené une analyse des usages effectifs des évaluations menées au sein de l'Agence française de développement. Celle-ci (partiellement transposable à d'autres situations) fait apparaître 4 grandes catégories :

- ▶ **les usages stratégiques** permettent d'appuyer l'élaboration stratégique et de la réinterroger. Ces usages demandent une forte capacité critique des évaluations approfondies, ainsi qu'un travail de synthèse important pour prendre une vue d'ensemble sur les travaux de l'organisation en général ou dans un secteur donné (Résout-on tel problème ? Tel choix stratégique s'avère-t-il pertinent ?).
- ▶ **les usages au service du pilotage** sont centrés sur les services opérationnels. Il s'agit d'aider à mener à bien les programmes et les projets et d'aider les équipes à faire preuve de réflexivité sur leur activité. Ces usages demandent de multiplier les boucles d'apprentissage, et de mettre en place des outils permettant des prises de décision rapides.
- ▶ **les usages pour l'amélioration continue** visent à une prise de distance vis-à-vis de l'action, pour plutôt faire apparaître des enseignements utilisables dans le cadre d'interventions futures. Ces usages appellent à une réflexion sur la généralisabilité des résultats, qui peut prendre la forme d'une logique « what works » (c'est-à-dire visant à stabiliser des instruments répliquables) ou au contraire de leçons ou de « bonnes pratiques » utilisables dans le cadre de l'activité habituelle.
- ▶ **les usages pour le dialogue**, enfin, font de l'évaluation un outil de collaboration et de délibération entre les acteurs. L'évaluation est alors vue comme une parenthèse dans l'action lors de laquelle il est possible d'échanger sur les problèmes à résoudre, les finalités de l'action, etc.

Dans cette typologie, il est possible d'utiliser les résultats ou le processus de l'évaluation pour apprendre ou pour assurer la redevabilité, sans distinction fondamentale de nature quant aux évaluations menées.

POSITION DANS L'INSTITUTION

Les finalités assignées aux systèmes d'évaluation se retrouvent dans la position occupée dans l'institution : au plus près du conseil d'administration ou de la direction générale ; parmi les fonctions support, au côté des finances, des ressources humaines ou de la communication ; dans les services sectoriels ; ou encore au plus près du terrain (service au public, agence...). Ce choix est notamment guidé par des considérations relatives à la demande d'indépendance fonctionnelle. Lorsque l'évaluation naît d'une pression externe ou hiérarchique, elle est plus souvent extérieure aux programmes effectivement engagés (afin de ne pas être « juge et partie »).

Kupiec et al. ont ainsi proposé, à partir d'une observation des systèmes d'évaluation nationaux dédiés aux fonds européens (2019, 2023), une typologie en trois parties : 1) **Une unité centrale**, au plus près de la décision, gère l'ensemble des fonctions évaluatives. 2) **Des unités décentralisées**, généralement étroitement imbriquées dans les services opérationnels. 3) Et enfin, des unités décentralisées, mais reliées par une **unité de coordination**. Les unités décentralisées peuvent être territoriales ou thématiques, elles peuvent être intégrées aux services opérationnels ou y être simplement associées.

Les auteurs/rices mettent en évidence les différences entre ces différents systèmes :

- ▶ **Les fonctions centralisées sont généralement composées de spécialistes** (méthodologiques, techniques) à plein temps. Elles réalisent moins d'évaluations, mais elles ont le potentiel d'être d'un bon niveau de qualité méthodologique et technique. L'utilité des travaux dépend largement de l'orientation, qui peut être d'assurer une fonction de redevabilité vers l'extérieur ou vers le conseil d'administration, ou bien d'aide à la décision (dimension stratégique). Il y a un risque que cette fonction soit cependant coupée du terrain et peu en prise avec les besoins et les attentes des acteurs en charge de la conception et de la mise en œuvre de l'action.
- ▶ **Les fonctions très décentralisées correspondent souvent à des temps partiels d'agents ayant d'autres métiers**. Elles réalisent potentiellement beaucoup plus d'évaluations, très opérationnelles (dont on peut imaginer qu'elles peuvent être utiles au pilotage, mais dont le contenu ne remonte pas toujours à la direction). Elles sont souvent moins ambitieuses d'un point de vue méthodologique, et il y a un risque qu'elles perdent leur dimension évaluative pour se fondre dans les demandes des services dans lesquelles elles s'inscrivent (qui peuvent relever du suivi, de l'étude voire de l'audit).

Il n'y a donc pas de solution parfaite, mais différents compromis à effectuer. Dans tous les cas, la question du positionnement dans l'institution demande à agir sur différents leviers, qui sont en particulier :

- ▶ **La capacité des équipes** (en charge de l'évaluation, mais aussi des usagers potentiels qui doivent apprendre à s'en saisir et à en tirer parti),
- ▶ **Les cadres institutionnels** (évaluation obligatoire ou non, circuit de l'évaluation, rôle formalisé de l'évaluation dans les prises de décision et dans les processus administratifs et techniques, etc.)

CONNEXION AUX ENJEUX DE TRANSFORMATION DES ORGANISATIONS

La façon dont l'évaluation est pensée dans les organisations est également liée à la façon dont les organisations pensent leur rapport aux transformations qu'elles veulent voir advenir.

La gestion basée sur les résultats et la focalisation, à partir des années 2000, sur les notions de performance et d'impact, ont incontestablement renforcé le rôle de l'évaluation dans de nombreuses organisations, notamment lorsqu'elles font face à de fortes pressions de redevabilité (organisations internationales, certaines organisations philanthropiques dépendantes du renouvellement des dons, agences de développement, économie sociale et solidaire...).


Ce renforcement s'est notamment fait dans trois sens : prouver la performance de gestion ; démontrer un impact attribuable à l'organisation, si possible de façon irréfutable ; protéger l'organisation de la critique externe (en la dotant d'une critique interne). Cette évolution a amené notamment à :

- ▶ **formaliser la fonction évaluative.**
- ▶ **la faire monter en compétence évaluative**, notamment méthodologique.
- ▶ **investir largement dans des systèmes d'information** sophistiqués, dont la gestion est un métier/une occupation à part entière

Dans les années 2000 également, l'innovation est propulsée au rang de manière privilégiée de résoudre les problèmes épineux auxquels les organisations s'attellent, avec deux corollaires : le développement du « mode projet », d'une part, et le renforcement de l'expérimentation d'autre part. L'idée sous-jacente est qu'il est possible de développer des solutions globales par un processus de « mise à l'échelle » progressif (Riddell & Moore 2015). Cela a également affecté l'évaluation, qui se renforce en activité de *proofmaking* (Béjean et al. 2022) à une échelle micro, mais aussi d'accompagnement à la montée en généralité à travers le suivi de programmes expérimentaux (*scaling out, up, deep*, Moore et al. 2015).

Les deux tendances présentées ci-dessus sont toujours à l'œuvre. Mais on peut aussi considérer qu'elles sont aussi largement remises en cause. Ainsi, les démarches d'institutionnalisation de l'évaluation ont amené des pratiques nouvelles, mais aussi créé une forte complexité, et parfois une pratique de façade, qui se fait au détriment des usages effectifs. De plus, les organisations doivent faire face à une triple transformation sur le front de la redevabilité (Bovens 2010) :

- ▶ La simple existence d'un dispositif de redevabilité ne suffit plus à faire confiance ; on attend de plus en plus des organisations concernées qu'elles s'engagent à prendre des mesures pour résoudre les problèmes identifiés par les audits, évaluations, ou autres études engagées.
- ▶ Les organisations ne sont plus uniquement redevables à leurs financeurs ou tutelles, mais également aux partenaires, aux publics, ou plus précisément à ceux qui subissent les conséquences des problèmes difficiles que les organisations n'arrivent pas à résoudre.

 - Une politique et un plan d'évaluation, des rapports annuels, des comités dédiés, des plans de renforcement de capacité... sont des outils de formalisation, considérés comme des signes extérieurs de bonne performance

- ▶ Enfin, le sujet de la redevabilité évolue, et ne porte plus seulement sur la performance ou l'impact, mais aussi sur la façon de conceptualiser ou estimer cet impact (le bien-être ou la dignité des personnes plutôt que le revenu, par exemple), ou encore sur des aspects éthiques (« Ne pas nuire »).

Le développement de l'innovation et des pratiques « agiles » n'est pas sans problème non plus. Le mode projet, loin de casser les silos, s'y associe et produit des effets délétères en termes de fragmentation et de manque de continuité de l'action (Graan 2002). De même, les démarches expérimentales se multiplient sans effet apparent en termes de changement dans la fabrique des politiques publiques. Dans les deux cas, les évaluations centrées sur le micro-projet et l'impact à court terme donnent une illusion d'efficacité qui s'évanouit lors de la généralisation, alors que les impacts des interventions ne croissent pas au même rythme que leur mise à l'échelle.

Enfin, **les grandes organisations tendent aussi à prendre conscience qu'elles s'inscrivent dans des systèmes de changement complexe, hautement incertains.** Dans ces systèmes, elles peuvent certes avoir une contribution significative, mais qui sera nécessairement limitée. Or, la volonté de s'attribuer ces impacts crée un certain niveau de défiance (notamment dans un climat de post-vérité : en fait, si tout le monde a un tel impact, pourquoi les problèmes sociaux ne sont-ils pas résolus ?) et empêche de discuter des échecs.

Toutes ces évolutions amènent les systèmes d'évaluation à se transformer en réponse aux demandes nouvelles qui leur sont faites (GEO, 2009). On observe ainsi certaines des évolutions suivantes, qui cohabitent avec les logiques héritées des années 2000 :

- ▶ **Un besoin de répondre à des demandes de redevabilité alternatives,** qui passe notamment par la réalisation d'évaluations au regard du genre, du point de vue des enfants, des plus démunies, etc. ou à des pratiques d'évaluation plus participatives. Certaines fonctions d'évaluation vont plus loin et prennent un rôle de « vigie morale », voire d'aiguillon de la transformation, dans laquelle ces nouveaux critères de jugement et pratiques servent à engager un dialogue sur la conception et la mise en œuvre des politiques avec les services sectoriels.
- ▶ La mise en place de cadres, d'espaces de discussion ou de processus structurés pour penser les évolutions stratégiques. Ces espaces, considérés comme nécessaires pour naviguer dans la complexité et l'incertitude, sont alimentés en continu par des données empiriques et des analyses de terrain et sont ainsi détachés du cycle des politiques publiques.
- ▶ **une évolution vers plus de continuité et de recherche d'amélioration à long terme,** ce qui amène d'une part à investir dans la synthèse et l'analyse stratégique de l'ensemble des données collectées à travers des évaluations individuelles ; et d'autre part à envisager le recours à des évaluateurs/rices faisant appel à leur posture évaluative (Delahais 2023) et présentes sur le long terme en soutien des finalités poursuivies.
- ▶ la réalisation que ce sont les acteurs/rices de terrain financés, plutôt que les grandes organisations, qui sont les premiers moteurs de changement, qui appelle à un rapprochement de l'évaluation et de l'accompagnement technique, et à un investissement dans la réflexion sur la généralisabilité des interventions ;

- ▶ le besoin de réflexivité par rapport à la pression externe pour prouver son impact appelle à donner le champ libre à une critique vue comme vecteur d'échange et d'évolution stratégique (plutôt que comme « vérification indépendante de la performance »).

L'évaluation peut enfin prendre une dimension particulièrement collaborative, qui reflète certaines des évolutions engagées au sein des organisations elles-mêmes. Elle devient ainsi un pivot de « communautés d'apprentissage » orientées vers le changement (complexe, systémique), plutôt qu'un outil ponctuel de retour de terrain ou de redevabilité. L'enjeu, au final, est de disposer d'un socle partagé de connaissances et d'interprétations pour analyser les difficultés rencontrées et préparer les évolutions futures.

DIFFÉRENTS PROFILS D'ÉVALUATEURS/RICES DANS LES SYSTÈMES D'ÉVALUATION

Face à cette multiplication des enjeux, la question dès lors est de savoir si les profils d'évaluateurs/rices (ou de chargées de suivi/évaluation) dans les organisations sont ceux qui conviennent face à ces évolutions. Traditionnellement, les évaluateurs/rices internes étaient soit formés en sciences sociales, soit juristes (en charge de marchés d'évaluation), soit tout simplement des fonctionnaires positionnés à ce poste sans avoir à faire valoir une qualification spécifique. Love, en 1991, identifie 7 rôles que les évaluateurs/rices peuvent endosser avec succès : Consultant·e en management, spécialiste de l'aide à la décision, spécialiste de la gestion de l'information, généraliste des systèmes, expert dépanneur, chargé·e de plaider, et planificateur/rice systématique. Ces rôles restent tout à fait pertinents aujourd'hui.

Nous proposons ci-dessous des rôles différents qui correspondent plus largement à la vision qui est donnée de la fonction évaluation dans les systèmes actuels (Love, 1991; Mark et al. 1999; Olejniczak et al. 2016) :

- ▶ **Le/la commanditaire** vise d'abord à faire en sorte d'avoir une réponse de qualité aux demandes exprimées par les acteurs/rices au sein de l'organisation. Cela passe en particulier par la capacité à élaborer des Termes de référence de qualité, sécurisés juridiquement, et à assurer le suivi et le contrôle qualité des prestations réalisées. Il/elle doit donc aussi avoir une bonne expertise méthodologique dans l'évaluation ou la conduite d'études, pour assurer une relation équilibrée avec le prestataire nécessaire à l'obtention de la qualité voulue.
- ▶ **Le/la spécialiste des systèmes d'information** est d'abord concerné par la disponibilité de données pour les acteurs opérationnels et stratégiques de l'organisation. Il/elle s'intéresse à la définition d'indicateurs ou de descripteurs utiles à la décision, et mobilise les technologies de l'information pour créer des tableaux de bord, des répertoires, des wikis... et les infrastructures permettant de les abonder en données récentes et utiles.
- ▶ **L'expert·e thématique** est intéressé·e par la production de connaissance elle-même, en particulier en réponse à des problèmes complexes ou sur des connaissances lacunaires. Cela signifie un travail permanent de veille et de synthèse de la connaissance sur les sujets qui intéressent l'organisation, et de diffusion vers les acteurs internes et externes à l'organisation. Il/elle peut commanditer des études ou évaluations selon les besoins, mais les réintroduit généralement dans une vision plus large, de façon à donner du sens.

- ▶ **Le/la courtier·e en connaissance** fait un travail permanent de mise en relation entre celles et ceux qui produisent de l'information, et ceux qui l'utilisent. Il/elle comprend et reformule les besoins d'information des différents acteurs impliqués et les retraduit en direction des producteurs possibles, quels qu'ils soient. Les courtier·es sont aussi généralement investi·es dans la diffusion de connaissances issues des évaluations en dehors de l'organisation.
- ▶ **L'évaluateur/riche interne** réalise des évaluations dans l'organisation, généralement à la demande de la direction ou du conseil d'administration. Il/elle peut avoir une indépendance statutaire (rend compte au CA) ou fonctionnelle (rend compte à la direction). L'évaluateur/riche interne doit pouvoir mobiliser tous les outils de l'évaluation, et adopte généralement une posture d'utilité maximale des travaux pour l'instance à laquelle il/elle réfère.
- ▶ **L'ami·e critique** est d'abord investi en soutien aux équipes opérationnelles et stratégiques impliquées dans la conception et la mise en œuvre de l'action. Il/elle utilise sa posture évaluative et les outils de l'évaluation pour rendre l'action plus pertinente, cohérente, efficace... tout cela dans un échange permanent et constructif avec les équipes. L'ami·e critique réalise rarement des évaluations complètes, mais sait mobiliser les outils d'évaluation autant que de besoin.

En pratique, la fonction évaluation doit souvent gérer plusieurs de ces différentes fonctions, et il s'agit alors de disposer de plusieurs profils pour y arriver.

POUR CONCLURE : UN « MECCANO » ÉVALUATIF

Nous avons présenté un certain nombre d'enjeux et d'ingrédients pour un système évaluatif et le rôle de la fonction d'évaluation en son sein. Il n'existe pas d'approche parfaite, mais à partir des éléments indiqués (orientation interne ou externe, focus stratégique ou opérationnel, types d'usage recherchés, positionnement dans l'organisation, direction de transformation, type de profil), il est possible de constituer différents scénarios.

Par exemple, on peut imaginer un système à orientation externe, avec une fonction centralisée rendant compte de la bonne mise en œuvre des interventions engagées à travers des tableaux de bord librement consultables par le public, mis en œuvre par un profil type spécialiste des systèmes d'information. Une autre possibilité est un système à orientation interne, tourné vers le pilotage et la montée en compétences des équipes, dans lequel des évaluateur/rices « ami·es critiques », embarqué·es dans les services pour améliorer les projets engagés.

Dans tous les cas, il n'existe pas de solution toute faite ou de « bonne façon » d'élaborer un système d'évaluation. C'est en se saisissant des opportunités que présente une situation existante qu'il est possible de mettre en place ou de faire évoluer le système d'évaluation.

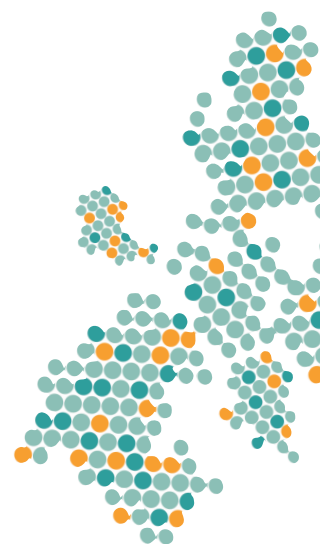
BIBLIOGRAPHIE SÉLECTIVE

Delahais, T., Devaux-Spatarakis, A., Jégou, F. (2022). Évaluation d'impact, Cartographie des usages. Quadrant Conseil - SDS pour l'Agence française de développement.

Kupiec, T., Wojtowicz, D., & Olejniczak, K. (2023). Structures and functions of complex evaluation systems: Comparison of six Central and Eastern European countries. *International Review of Administrative Sciences*, 89(1), 202-220.

Mark, M. M., Henry, G. T., & Julnes, G. (1999). Toward an Integrative Framework for Evaluation Practice. *American Journal of Evaluation*.

Rist, R. C., & Stame, N. (2006). From studies to streams: Managing evaluative systems (Vol. 12).



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



Réutilisation possible sans permission en citant la source: « Thomas Delahais, Quadrant Conseil, 2024 »