Ce document est la version pre-print de l'article paru dans *Socio-économie du travail*, n°15, 2024-1. DOI : https://doi.org/10.48611/isbn.978-2-406-18609-0.p.0117.

ARTICULER QUANTITATIF ET QUALITATIF : METHODES ET PRATIQUES DANS L'EVALUATION DES POLITIQUES DE L'EMPLOI EN FRANCE (2008-2023)[1]

Yves DE CURRAIZE

Université Paris Est Créteil (LIPHA), TEPP-CNRS et CEET

Francesco SERGI

Université Paris Est Créteil (LIPHA)

Qu'est-ce qu'une méthode « qualitative » ? Qu'est-ce qu'une méthode « quantitative » ? Quelle place accorder à ces deux catégories de méthodes dans l'évaluation des politiques publiques ? Quelle articulation proposer entre les deux ? Dans cet article, nous analysons les débats que ces trois questions ont suscités, en France, entre les acteurs impliqués dans l'évaluation des politiques publiques. Notre analyse adopte une perspective d'histoire et de sociologie des sciences, visant à contextualiser, d'une part, les principes méthodologiques énoncés par les acteurs impliqués dans l'évaluation (leurs discours), et, d'autre part, à documenter leurs pratiques concrètes de l'évaluation. Pour être « au plus près » des acteurs, nous nous appuyons sur un corpus original, constitué de « littérature grise » : les guides méthodologiques, les études et les rapports d'évaluation produits dans le cadre de l'évaluation des politiques de l'emploi en France, entre la moitié des années 2000 et 2022 (voir les annexes pour un premier aperçu du corpus)[2].

Notre analyse met en évidence trois résultats. Premièrement, la définition des méthodes « qualitatives » (et, par symétrie, la définition des méthodes « quantitatives ») constitue en elle-même un sujet de débat. Les méthodes qualitatives envisagées dans les guides méthodologiques et mobilisées dans les évaluations sont construites dans des perspectives souvent différentes, par des communautés attachant à ces méthodes différentes significations et objectifs. Par exemple, la méthode de « l'entretien » peut s'insérer à la fois dans une perspective d'étude ethnographique, mais aussi dans une perspective d'« évaluation basée sur la théorie ». Notre analyse apporte donc une clarification du sens de ces deux catégories (méthodes qualitatives / méthodes

[1] Nous tenons à remercier Olivier Baguelin et Laurent Duclos, coordinateurs de ce numéro thématique, ainsi que le comité éditorial de la revue *Socio-économie du travail* et les deux évaluateurs anonymes pour leurs commentaires, qui ont permis d'améliorer significativement notre travail. Nous avons également bénéficié des retours constructifs et attentifs de nombreux collègues. Nous tenons tout particulièrement à remercier les participants au colloque TEPP 2023, les participants aux journées d'étude *Socio-économie du travail* du 17 mars 2023 à l'Université d'Évry, et Eur Frapp-Erudite-Lipha du 24 Juin 2024 à l'Université de Paris Est-Créteil et les participants au séminaire du LIPHA.

[2] La délimitation à un cas d'étude précis nous amène à construire un corpus quasi-exhaustif, permettant d'analyser plus systématiquement les principes méthodologiques et les pratiques de l'évaluation. Ce choix comporte également des limites, notamment la portée générale de nos résultats. Cependant, il rend également possible la comparaison future avec d'autres études de cas.

quantitatives), via leur contextualisation dans les discours et les pratiques des acteurs de l'évaluation.

Deuxièmement, dans le cas français, les méthodes « qualitatives » ont toujours été considérées comme légitimes dans la réflexion méthodologique des acteurs, même si elles n'ont pas été systématiquement mobilisées dans la pratique. Elles ont été principalement conçues et pratiquées soit en complément des méthodes quantitatives, soit sans articulation claire avec ces dernières. Nous identifions, à ce sujet, deux cas de figure généraux. D'une part, le cas de la « juxtaposition » entre méthodes qualitatives et méthodes quantitatives : dans cette configuration, les deux méthodes sont mobilisées en parallèle, le plus souvent par des acteurs différents, sans interaction directe, sans coordination, sans articulation. Dans notre corpus, on identifie une majorité d'occurrences de cette démarche de « juxtaposition ». D'autre part, nous identifions une minorité de cas des « méthodes mixtes », c'est-à-dire des configurations méthodologiques et pratiques proposant une articulation explicite entre méthodes qualitatives et méthodes quantitatives. Notre analyse apporte donc un état des lieux, dans le cas français, de l'articulation (ou, plutôt, du manque d'articulation) entre les deux méthodes.

Troisièmement, on met en évidence une division du travail entre les différents acteurs de l'évaluation. La conduite des « volets qualitatifs » des évaluations semble le plus souvent être confiée à des acteurs privés (cabinets de consultants) ; tandis que le « volet quantitatif » est le plus souvent confié à des acteurs universitaires ou à des acteurs issus des administrations publiques. Cette division du travail peut expliquer en partie la prédominance de la « juxtaposition » entre méthodes qualitatives et quantitatives.

Nos trois résultats illustrent une conclusion, plus générale, présente dans la littérature sur la socio-histoire de la quantification (initiée notamment par Porter, 1995 ; Desrosières, 2000 ; 2008a ; 2008b ; 2014) : à savoir, que les dispositifs de quantification ne se construisent pas exclusivement à partir d'une réflexion méthodologique (ou épistémologique), centrée sur des logiques cognitives (de savoir) ; en effet ces dispositifs résultent également de logiques politiques (d'action), et de la manière dont ces logiques de savoir et d'action s'agencent. Du point de vue de la logique de savoir, nous considérons les évolutions méthodologiques comme des réponses aux obstacles épistémologiques, que l'économie appliquée cherche à dépasser. L'usage du qualitatif permettrait de renforcer la validité tant interne qu'externe des conclusions des évaluateurs. Dans la mesure où le développement de méthodes expérimentales ou quasi expérimentales a contribué à renforcer la validité de l'inférence causale à l'intérieur d'un contexte donné (validité interne) mais au détriment de la validité à l'extérieur de ce contexte (voir Cartwright, 2007), les enjeux méthodologiques de l'usage du qualitatif portent principalement sur la validité externe. Toutefois nous verrons que celui-ci peut aussi répondre à des préoccupations concernant la validité interne. Mais notre analyse met en évidence que la juxtaposition entre méthodes quantitatives et qualitatives s'avère être aussi le résultat d'un choix politique (c'est-à-dire, d'un choix réalisé par les commanditaires), à savoir le choix d'imposer une division du travail entre différents acteurs en charge de l'évaluation des politiques publiques ou plus largement d'une dimension politique renvoyant à l'évolution d'un rapport de force entre les différents acteurs de l'évaluation. La perspective socio-historique adoptée dans cet article peut ainsi nourrir la réflexion des praticiens sur les enjeux politiques dans la construction des dispositifs d'évaluation, en particulier sur la manière d'organiser le dialogue entre différents acteurs – une réflexion développée par d'autres contributions à ce numéro spécial (adoptant des perspectives

Ce document est la version pre-print de l'article paru dans *Socio-économie du travail*, n°15, 2024-1. DOI : https://doi.org/10.48611/isbn.978-2-406-18609-0.p.0117.

différentes de la nôtre), ainsi que de plusieurs travaux existant (voir notamment Labrousse et Zamora, 2013 ; Delahais et al., 2020).

L'agencement entre logiques de savoir et de pouvoir s'explique, plus généralement, par un contexte historique, en particulier par une certaine conception de la relation entre l'action publique, le marché, et les dispositifs de quantification. Dans ses travaux, Alain Desrosières souligne les caractéristiques historiques de différentes « configurations typiques » (voir notamment Desrosières, 2003). Les résultats de notre article renvoient à plusieurs de ces caractéristiques, dans le cas que Desrosières appelle « l'État néo-libéral ». Premièrement, la logique de l'évaluation (qui sous-tend aussi bien la mobilisation de méthodes qualitatives ou quantitatives) relève d'une logique d'action centrée sur les incitations : dans cette optique, l'étude des motivations des individus occupe une place grandissante. Deuxièmement, la gouvernance « polycentrique » (la démultiplication des centres de décisions) dans l'« État néo-libéral » se traduit, dans notre cas d'étude, dans la démultiplication des acteurs de l'évaluation et, in fine, dans la démultiplication des méthodes. Cela donne également lieu, dans le cadre des logiques de savoir, à un plus difficile « mise en équivalence » des procédés d'évaluation – la prévalence de la juxtaposition entre méthodes « qualitatives » et « quantitatives » pouvant se lire comme une conséquence de cette difficulté de « mise en équivalence ».

L'article présente d'abord le corpus au cœur de notre travail, en le contextualisant dans l'histoire de l'évaluation des politiques publiques en France (section I). Nous analysons ensuite les principes exposés dans les guides méthodologiques (section II), puis les pratiques de l'évaluation telles qu'elles sont retracées dans des rapports et études (section III).

I. UNE BREVE HISTOIRE DES REFLEXIONS METHODOLOGIQUES SUR L'EVALUATION DES POLITIQUES PUBLIQUES EN FRANCE

Cette section propose une brève histoire de l'évaluation des politiques publiques en France, sans prétention aucune d'exhaustivité. En effet, l'enjeu est ici plutôt de rappeler le contexte historique et politique dans lequel se déploient les débats méthodologiques et les pratiques d'évaluation qui feront l'objet de notre analyse dans la suite de l'article. Nous commençons par discuter l'importance de l'idée de « tournant expérimental » (I.1), qui structure, pour le cas français, à la fois notre contextualisation dans cette section et la construction de notre corpus. Nous présentons ensuite trois moments clé de l'histoire de l'évaluation des politiques publiques en France (I.2-I.4) : nous suggérons que ces trois moments peuvent s'interpréter comme le résultat d'une dynamique historique et épistémologique, propre au contexte français, dans laquelle le « tournant expérimental » peine d'abord à s'imposer, faisant face à de nombreuses critiques, puis à des difficultés de mise en œuvre.

I.1. LE « TOURNANT EXPÉRIMENTAL » EN FRANCE

La littérature en histoire de la pensée économique et en philosophie économique a consacré, ces dernières années, beaucoup d'attention à l'évolution, depuis les années 1970, de la méthodologie des sciences économiques à l'échelle globale. L'un des aspects

Ce document est la version pre-print de l'article paru dans *Socio-économie du travail*, n°15, 2024-1. DOI : https://doi.org/10.48611/isbn.978-2-406-18609-0.p.0117.

les plus notables de cette évolution méthodologique peut se résumer sous le terme de « tournant expérimental », qui s'insère dans un virage plus général vers « l'économie appliquée » (voir notamment Backhouse et Cherrier, 2017). Dans le contexte de la discipline économique « internationalisée » (Fourcade, 2006), le « tournant expérimental » correspond à la fois à un changement d'échelle et de positionnement de différentes « méthodes expérimentales », à partir dans les années 1970 : le recours à ces méthodes devient substantiellement plus fréquent que par le passé ; et, surtout, ces méthodes trouvent une place au sein du courant dominant en sciences économiques. Les « méthodes expérimentales » protagonistes de ce tournant sont en réalité de différentes natures – et on pourra distinguer notamment deux tendances. D'une part, la montée en puissance de l'expérimentation « en laboratoire », dans le sillage de l'essor de la (nouvelle) économie « comportementale » et de l'économie « expérimentale » (voir notamment Guala, 2005 ; Heukelom, 2014 ; Svorencik, 2015). D'autre part, le « tournant expérimental » s'incarne dans la généralisation des « expérimentations de terrain », portées en particulier par les économistes du développement (voir Favereau, 2021)[3].

Le « tournant expérimental » en sciences économiques entre en résonance avec l'évolution, sur la même période, de la conception et l'évaluation des politiques publiques[4]. Les économistes, sous leur double casquette de « scientifiques » et d' « experts », s'attellent notamment au déploiement des méthodes expérimentales dans le cadre de l'évaluation des politiques publiques. Les méthodes expérimentales ont été ainsi appréhendées, dans la littérature en histoire et philosophie de l'économie, sous le double angle « d'outil de preuve » et « d'outil de gouvernement » (pour reprendre la formule de Desrosières). Ainsi, l'idée d'un tournant expérimental est bien centrale dans les discussions à la fois sur les évolutions de l'économie comme discipline, mais aussi dans l'évolution des pratiques d'expertise des économistes.

Au sein des méthodes « expérimentales » au sens large, l'approche particulière des « *randomised controlled trials* » (traduit en français par « expériences de terrain randomisées » ou, le plus souvent, par « expérimentations aléatoires ») s'est imposée comme un « outil de preuve et de gouvernement » majeur, d'abord aux États-Unis, puis dans d'autres pays ainsi que dans les organisations internationales (voir notamment Jatteau, 2020 pour une histoire plus complète)[5]. Les expérimentations aléatoires jouent ainsi un rôle pivot dans les débats sur l'évaluation des politiques publiques. Comme le remarque, par exemple, Cartwright (2007), certains défenseurs des expérimentations aléatoires s'emploient à élever cette méthode en « standard méthodologique » (« *gold standard* ») pour l'évaluation des politiques publiques. Cette volonté s'appuie notamment sur l'idée que les expérimentations aléatoires produisent une quantification de l'impact

[3] Pour une discussion plus fine des différentes catégories de « l'expérimentation » en sciences économiques, voir Harrison et List (2004) et, pour une perspective plus réflexive, de Curraize et Thoron (2020) et Nagatsu et Favereau (2020).

[4] Pour ce qui concerne plus spécifiquement la conception des politiques publiques, on renverra notamment à l'idée de « *nudge* », qui est le fruit des développements en économie expérimentale et comportementale (voir Thaler et Sunstein, 2008).

[5] Comme le documente Jatteau (2020, p. 1), cette méthode est présentée sous différentes appellations en français (évaluations aléatoires, expériences contrôlées, expériences sociales aléatoires etc.), tandis que la terminologie en langue anglaise est plutôt stabilisée aujourd'hui (*randomized controlled trials*). Dans cet article, nous adoptons la traduction « expérimentations aléatoires ».

Ce document est la version pre-print de l'article paru dans *Socio-économie du travail*, n°15, 2024-1. DOI : https://doi.org/10.48611/isbn.978-2-406-18609-0.p.0117.

causal des politiques publiques, reposant sur une forte validité interne (voir par exemple Banerjee et Duflo, 2009 ou, dans le contexte français, L'Horty et Petit, 2011)[6].

Cette revendication de la nature quantitative des résultats obtenus via les expérimentations aléatoires impliquerait, logiquement, de les classer dans la catégorie des méthodes quantitatives. En réalité, dans la pratique, les choses semblent plus complexes, comme le rappelle par exemple Cartwright et Hardie (2012) : de nombreuses expérimentations aléatoires intègrent des dimensions qualitatives (comme, par exemple, différentes phases d'entretien avec les sujets de l'expérimentations ; voir par exemple Rao et Woolcock, 2003 ; Woolcock, 2019). Certains commentateurs semblent voir dans cet usage des méthodes qualitatives un signe d'une forme de « décalage » entre, d'une part, les principes méthodologiques présentés par les tenants des expérimentations aléatoires et, d'autre part, la mise en pratique concrète de leurs principes, lors de la conduite des expérimentations. En particulier, si la méthode des expérimentations aléatoires est présentée, le plus souvent, comme une méthode strictement quantitative (dont la raison d'être est, comme rappelé plus haut, la quantification des effets causaux des politiques publiques), alors, par conséquent, la pratique de cette méthode ne devrait pas faire appel à des méthodes qualitatives. C'est pour cette raison que, par exemple, Jatteau (2020, Chapitre 3, II) caractérise les expérimentations aléatoires de « bazar épistémologique ».

L'ensemble des critiques vis-à-vis des expérimentations aléatoires renvoient tantôt aux conditions de l'expérimentation, qui sur le terrain ne permettent pas toujours de vérifier les critères de validité interne, notamment la rigueur de l'affectation aléatoire au groupe de traitement, tantôt à la difficulté de généraliser les résultats de l'expérience (validité externe), ce qui s'avère pourtant indispensable dans une perspective d'action publique. Si ce dernier point a notamment été soulevé par Deaton et Cartwright (2018), soulignant l'avantage des modèles structurels qui, à la différence des expérimentations, estiment d'emblée des comportements généraux, cette question est plus largement posée par la philosophie des sciences. Ainsi Cartwright (2007), dans sa typologie des méthodes permettant d'établir une relation causale distingue les moyens d'obtenir une confiance élevée dans une relation causale limitée à un contexte donné, de ceux permettant de rendre compte de relations plus générales, mais que l'on démontre avec moins de certitude. Parmi ces dernières sont mentionnées des méthodes qualitatives telles que l'ethnographie ou encore l'analyse qualitative comparée. Se dessine donc un double enjeu pour une méthodologie articulant des outils quantitatifs et qualitatifs. D'une part les outils qualitatifs tels que les entretiens, l'observation ou l'analyse textuelle permettent, lorsqu'ils consistent à étudier le comportement des agents chargés de mettre en œuvre une expérimentation (dans une perspective de sociologie des professions par exemple), d'améliorer les pratiques de l'évaluation afin de garantir le respect des critères de validité interne. D'autre part lorsque ces outils sont en revanche tournés vers l'études des populations bénéficiaires des mesures évaluées, ils permettent d'expliquer, comprendre et de préciser le contexte dans lequel s'exercent les effets recherchés. Ils mettent donc à jour les conditions dans lesquelles une relation causale établie dans le contexte d'une expérience peut être généralisée hors de celui-ci. Ce deuxième enjeu, de validité externe, est crucial dans une perspective d'action publique.

[6] En deuxième instance, lorsqu'il est impossible de mettre en place des expérimentations aléatoires, les défenseurs de cette approche admettent volontiers la légitimité d'autres méthodes quantitatives, notamment des traitements statistiques (en particulier, les méthodes de différence-de-différences et les méthodes d'appariement statistique).

Ce document est la version pre-print de l'article paru dans *Socio-économie du travail*, n°15, 2024-1. DOI : https://doi.org/10.48611/isbn.978-2-406-18609-0.p.0117.

En nous appuyant sur la littérature en histoire et philosophie de l'économie, on fait apparaître deux éléments de contexte importants pour l'analyse du cas du « tournant expérimental » en France. Premièrement, l'introduction des méthodes expérimentales (au sens large) et des expérimentations aléatoires (en particulier) constitue un point de repère historique (un « tournant ») à la fois pour l'économie comme discipline et pour l'économie comme pratique d'expertise. Cette première conclusion justifie l'organisation chronologique de notre récit historique : nous prenons comme point de départ (ou comme borne temporelle) l'arrivée des expérimentations aléatoires en France, les débats autour de cette méthode occupant une place structurante dans la dynamique de l'évaluation des politiques publiques en France – en écho avec l'idée que les défenseurs de cette méthode l'érigent en « standard méthodologique » (voir aussi Jany-Catrice et al., *SET* 2023-2 n°14). Deuxièmement, la littérature nous indique qu'il existe, en particulier dans le cas des expérimentations aléatoires, une tension à interroger entre les discours d'une part et les pratiques d'autre part. Cette deuxième conclusion détermine notre choix de corpus.

Notre analyse s'appuie en effet sur deux types de sources, qui nous apportent chacune un éclairage différent. D'une part, nous étudions les « guides méthodologiques », les « manuels » et d'autres textes à portée générale et méthodologique, produits par les institutions françaises commanditaires de l'évaluation. Ces sources nous permettent de documenter le rôle et la place des méthodes qualitatives et leur articulation avec les méthodes quantitatives, tels qu'ils ont été conçus et « théorisés » dans les discours des acteurs de l'évaluation des politiques publiques. D'autre part, nous analysons les études, rapports, bilans et publications issues d'évaluations particulières. Ces sources nous permettent de documenter la manière dont les méthodes qualitatives sont mobilisées dans la pratique, c'est-à-dire comment les communautés menant les évaluations utilisent des méthodes qualitatives dans la conduite effective des évaluations. Ces deux types de sources sont originales, dans la mesure où elles sont souvent inexploitées (jusque-là) par les historiens de l'économie s'intéressant soit au « tournant expérimental » soit à l'expertise économique – y compris dans d'autres contextes que celui de la France. Ce corpus nous permet notamment d'interroger de façon critique l'éventuelle dissonance entre les discours et les pratiques[7].

I.2. L'ÉVALUATION DES POLITIQUES PUBLIQUES EN FRANCE : LE TOURNANT EXPÉRIMENTAL « CONTRARIÉ » (1980-2008)

La littérature existante a bien mis en évidence le processus long de la transformation de l'action publique en France et l'évolution, qui l'a accompagnée, des modalités de son suivi, de sa mise en chiffre, de son « évaluation ». On date de la fin des années 1970 l'essor du système contemporain « d'évaluation » des politiques publiques – aboutissant à des dispositifs « permettant d'apprécier l'efficacité des politiques publiques en comparant leurs résultats aux objectifs assignés et aux moyens mis en œuvre » (décret du 18 novembre 1998 instituant le Conseil national de l'évaluation)[8]. Ce processus s'inscrit,

[7] Il faut noter ici aussi que les analyses qualitatives, produites dans les études et rapports, sont souvent « invisibles » en dehors de cette littérature grise. Les volets quantitatifs de ces études (en particulier, les résultats des expérimentations aléatoires) ont en effet souvent une « deuxième vie » sous forme de publication dans des revues à comité de lecture en sciences économiques. Les études qualitatives ont moins aisément ce type de débouché dans la mesure où les sociologues français se sont longtemps tenus à l'écart de la commande publique en évaluation (Delahais et Devaux-Spatarakis, 2018, p. 1).

[8] Cette logique nouvelle vient alors s'ajouter à une logique, beaucoup plus ancienne, de vérification et de contrôle juridique et comptable de l'action des administrations, telle que les incarnes les corps « d'inspection » établis au sein de l'administration dès le XIX^{e} siècle.

Ce document est la version pre-print de l'article paru dans *Socio-économie du travail*, n°15, 2024-1. DOI : https://doi.org/10.48611/isbn.978-2-406-18609-0.p.0117.

dès les années 1970, dans un contexte de rigueur budgétaire et de déclin de la planification « à la française », d'abord avec la Rationalisation des choix budgétaires (RCB), puis, dans les années 2000, avec la Loi organique relative aux lois de finances (LOLF). Les pouvoirs publics enchaînent, sur cette période, la création de nouvelles institutions consacrées à l'évaluation (voir par exemple Perret, 2008, Chapitre IV).

Cette nouvelle conception de l'évaluation et sa mise en place au cœur des institutions étatiques s'accompagne d'une réflexion sur l'évolution possible des méthodes d'évaluation. Dans le cadre de cette réflexion, on verra émerger progressivement une pression de la part des économistes universitaires en faveur de l'introduction des méthodes d'expérimentation aléatoire (voir par exemple ; Fougère, 2000). Les principes de l'expérimentation aléatoire sont connus en France, en raison de leur diffusion, dès les années 1960, aux États-Unis, dans le cadre de l'évaluation des politiques sociales et de l'emploi. « La référence au cas des États-Unis, grands innovateurs en la matière [d'expérimentation aléatoire], a pu constituer un aiguillon appréciable », comme le notait, à la fin des années 1980, l'introduction d'un rapport du Commissariat au Plan sur l'évaluation des politiques publiques (Deleau *et al*., 1986, p. 20). Cependant, le même rapport se montrait sceptique :

> Faut-il en conclure que l'expérimentation avec tirage au sort doit être systématiquement privilégiée pour évaluer les politiques publiques ? En dépit d'exemples significatifs de l'apport de ce plan de recherche, notamment aux États-Unis, il n'est pas réaliste d'aller aussi loin. (Deleau *et al*., 1986, p. 72)

Ce refus, prudent certes, d'engager le développement de l'expérimentation aléatoire est justifié par diverses difficultés – en particulier, la possibilité de généraliser leurs résultats, c'est-à-dire le problème de la validité externe. Le rapport du Commissariat au Plan insiste alors sur la plus grande pertinence d'autres méthodes, notamment de la modélisation économétrique, pour produire des évaluations *ex post* des politiques publiques. Cette position n'est pas sans lien avec le profil général des auteurs du rapport. Il s'agit d'économistes actifs à l'intersection entre le champ universitaire et le champ des administrations publiques. Tous ont contribué à l'évaluation des politiques publiques depuis les années 1960, en particulier dans le cadre du développement de la modélisation macroéconométrique – à l'image notamment de Michel Deleau, premier auteur du rapport, qui a été un contributeur essentiel dans le déploiement de ce type de modélisation dès les années 1970[9].

Durant les années 1990, certains universitaires français (cette fois plutôt ancrés dans le champ universitaire, et dans une mouvance plus institutionnaliste ou « hétérodoxe ») produisent également des critiques vis-à-vis des expérimentations aléatoires, visant à la fois la méthode mais aussi, tout particulièrement, son application au domaine des politiques de l'emploi (voir par exemple Monnier, 1992 ; Morel, 1996 ; Perez, 2000). Ils associent également l'expérimentation aléatoire à une méthode « américaine », séduisante pour « certains » mais ayant déjà fait preuve, aux États-Unis, de beaucoup de limites (voir par exemple Monnier, 1992, p. 40 ; Barbier, 2010). Ces critiques s'accompagnent de propositions alternatives, en particulier d'un élan d'innovation dans les méthodes statistiques – par exemple, la production et l'analyse de données de panel,

[9] Voir par exemple Desrosières (2000) pour une vision plus détaillée du rôle des économistes, et en particulier des économistes-ingénieurs, dans le développement de l'analyse économétrique et macroéconométrique au service du pilotage des politiques publiques en France.

Ce document est la version pre-print de l'article paru dans *Socio-économie du travail*, n°15, 2024-1. DOI : https://doi.org/10.48611/isbn.978-2-406-18609-0.p.0117.

portées à la fois par des acteurs de la statistique publique et par des chercheurs (voir notamment Pénissat, 2011; 2016).

Alors que le paradigme de l'évaluation des politiques publiques évolue entre les années 1980 et 2000, sur cette même période, la méthode des expérimentations aléatoires est l'objet de critiques et réticences, aussi bien dans les administrations que dans le monde universitaire. L'expérimentation aléatoire intègrera cependant les dispositifs d'évaluation des politiques publiques, officiellement avec la modification constitutionnelle de 2003 – introduisant, via l'article 37-1, la possibilité de « dispositions [législatives] à caractère expérimental ». Cependant, comme le souligne par exemple Jatteau (2020, Chapitre 2, III), l'essor véritable des expérimentations aléatoires en France coïncidera tout particulièrement avec la création, en 2008, du Fonds d'expérimentation pour la jeunesse (FEJ).

I.3. La mise en place des experimentations et le Fonds d'experimentation pour la jeunesse (2008) : les premieres reflexions methodologiques

La période qui s'ouvre avec la création FEJ marque certes un « tournant expérimental » en France – dans le sens où cette méthode intègre à part entière (*de jure*) le domaine de l'évaluation des politiques publiques. Cependant, comme nous venons de le rappeler, ce « tournant » n'est point une entrée triomphale ; il s'opère bien dans un contexte de critique, voire de controverse. Ce contexte nous semble contribuer à marquer les débuts du tournant expérimental en France par une réflexion méthodologique, notamment une réflexion sur l'articulation possible entre l'expérimentation aléatoire, d'autres méthodes quantitatives et, enfin, la place accordée aux méthodes qualitatives.

Même si le rôle du FEJ a été central dans le tournant expérimental en France, les premières évaluations des politiques publiques se revendiquant d'une méthode « expérimentale » ont été antérieures au FEJ. Elles ont été menées dans le domaine des politiques de l'emploi, en l'occurrence lors de la mise en place du Revenu de solidarité active (RSA)[10]. Les acteurs de ces évaluations étaient plutôt ancrés dans le champ universitaire, et en particulier gravitant autour de l'École d'économie de Paris et du CREST. Mais il s'agit aussi de profils plus « internationalisés » (par rapport à la génération du rapport « Deleau *et al*. » par exemple). Cette nouvelle génération reste en partie ancrée dans une formation française d' « économiste-ingénieur » (formés à l'ENSAE, par exemple) ; cependant, elle se forme également à l'étranger, et en particulier en Amérique du Nord, où elle tisse des liens internationaux de collaboration, en particulier avec le J-PAL[11]. Nous retrouverons par ailleurs ces mêmes profils au sein du FEJ, en particulier de son conseil scientifique.

Les expérimentations autour du RSA se déroulent en deux temps : d'abord, lors d'une expérience sur l'accompagnement individualisé des chômeurs (Behaghel *et al*., 2008 ; 2013 ; Crépon, 2008), en lien avec la mise en place du RSA ; puis, lors de l'expérimentation du RSA dans certains départements, avant sa généralisation en 2009

[10] Plus largement, c'est la mise en place du Revenu minimum d'insertion (RMI) en 1988 que l'on peut considérer comme la première pierre de ce processus, à deux titres : d'une part, dans la mesure où le principe de l'évaluation est affirmé dans la loi de 1988, et, d'autre part, parce que la mise en place du RSA est une réponse aux effets (dés)incitatifs générés par le RMI – un sujet de débat très présent à la fois dans la littérature universitaire et dans le débat public.
[11] Pour une vision plus générale de la trajectoire « d'internationalisation » des sciences économiques en France, voir par exemple Fourcade (2009).

Ce document est la version pre-print de l'article paru dans *Socio-économie du travail*, n°15, 2024-1. DOI : https://doi.org/10.48611/isbn.978-2-406-18609-0.p.0117.

(Bourguignon, 2008 ; 2009). Dans ces deux cas, les évaluateurs présentent, dans leur méthodologie, la notion de « tirage aléatoire » ; cependant, celui-ci n'a pas pu être mis en œuvre dans le cas de l'expérimentation du RSA[12]. Ces expérimentations comportent également un volet qualitatif, bien que l'on ne trouve pas encore, dans les rapports des comités d'évaluation, d'éléments de méthode stipulant le mode d'articulation entre quantitatif et qualitatif. Les auteurs des rapports présentent cependant clairement les objectifs des enquêtes qualitatives, à savoir de « mieux saisir les mécanismes qui conduisent à la reprise d'emploi, ou, au contraire, les freins à l'intégration sur le marché du travail » (Étude qualitative du CREDOC, Annexe 4, 6, nous soulignons) ; ou, encore de « décrire » les « conditions de mise en œuvre de la mesure », comme le font les auteurs de l'étude qualitative sur l'accompagnement individualisé (Capelier et Mizrahi, 2008). Ici, la méthode de l'enquête qualitative chercherait en particulier à rendre compte de la façon dont la mesure est appliquée par les agents en ayant la charge ou est reçue par le public bénéficiaire. Si les études qualitatives liées à ces premières évaluations ont été moins commentées que leur volet quantitatif, qui inaugure la montée en puissance de l'expérimentation aléatoire en France, elles n'en demeurent pas moins conséquentes[13].

Le FEJ poursuit cette première lancée. Un budget relativement élevé est consacré à un ensemble diversifié d'expérimentations, dont une partie concerne les politiques de l'emploi, au sens large. Plus de 300 expérimentations seront ainsi menées, sous l'égide du FEJ, sur la période 2009-2012 (Aldeghi *et al*., 2012). Deux d'entre elles prennent une place particulière dans cet ensemble : les études sur la mobilité et l'emploi (dont, par exemple, le dispositif « 10000 permis pour réussir ») et le Revenu contractualisé d'autonomie (RCA). Ces expérimentations contribueront ultérieurement au débat et à la mise en place de dispositifs pérennes[14]. Pour cette raison, les études issues de ces deux expérimentations constitueront le premier jalon de notre corpus.

Le fonctionnement du FEJ par appel à projets s'accompagne d'une réflexion méthodologique, incorporée dans le *Guide méthodologique pour l'évaluation des expérimentations sociales à l'intention des porteurs de projets*, édité par le FEJ dès le lancement de son premier appel à projet en 2009[15]. Ce guide définit un cadre méthodologique général pour l'évaluation des politiques publiques[16] : il propose également, pour la première fois dans le nouveau contexte français de cette période, une réflexion explicite sur les liens entre les approches quantitative et qualitative. Ainsi, ce guide méthodologique constitue le deuxième jalon de notre corpus.

---

[12] Ainsi la note d'étape sur l'évaluation de l'accompagnement renforcé des demandeurs d'emploi produit un encadré méthodologique décrivant « un protocole expérimental innovant fondé sur l'affectation aléatoire » (« Les expérimentations d'accompagnement renforcé des demandeurs d'emploi conduites par l'Unédic et l'ANPE », *Note d'étape de l'évaluation*, ANPE, Unedic, Dares, juin 2008). Le rapport final sur l'évaluation du RSA évoque quant à lui le tirage aléatoire comme la méthode de référence, qui n'a toutefois pas pu être appliquée : « Aux vues de la difficulté pratique et éthique de procéder par échantillonnage aléatoire, le principe d'évaluation retenu pour cette expérimentation consiste à procéder par comparaison géographique. » (Bourguignon, 2009).

[13] Par exemple, le volet qualitatif de l'évaluation du RSA s'appuie sur 94 entretiens et sur des monographies menées dans plusieurs départements.

[14] En particulier, nous verrons plus loin comment le RCA a pu contribuer à définir la Garantie jeunes.

[15] La première version de ce guide a été présentée et discutée à Grenoble en novembre 2007, lors des « Rencontres de l'expérimentation sociale ». Le guide sera ensuite réédité, dans des versions similaires, en 2010 et en 2011.

[16] Le FEJ laisse cependant, par des appels d'offre ou des appels à recherche, une place importante aux initiatives des expérimentateurs et des collectivités locales, bien qu'on assiste par la suite à une forme de reprise en main par l'État (Bureau *et al*., 2013).

Ce document est la version pre-print de l'article paru dans *Socio-économie du travail*, n°15, 2024-1. DOI : https://doi.org/10.48611/isbn.978-2-406-18609-0.p.0117.

I.4. ÉVALUATIONS ET EXPERIMENTATIONS (2008-2022) : AJUSTEMENTS METHODOLOGIQUES, DISCOURS ET PRATIQUES

L'activité d'expérimentation impulsée par le FEJ (ainsi que sa prise de position méthodologique à travers le « Guide ») n'épuise pas les débats critiques autour des méthodes de l'évaluation des politiques publiques. Alors que le débat des années 1990 et du début des années 2000 analysait avec scepticisme l'expérience étasunienne, le débat au tournant des années 2000 s'empare des expériences cumulées par le FEJ. Le CREDOC publie ainsi, en 2012, un « Bilan méthodologique des évaluations conduites par le CRÉDOC dans le cadre du Fonds d'expérimentation pour la Jeunesse » (Aldeghi *et al*., 2012). Ce « bilan » porte sur les expérimentations menées par le FEJ dans les trois années précédentes. On trouve ici une discussion explicite des critiques qui traversent, à cette période, les communautés actives dans l'évaluation des politiques publiques, ainsi que diverses propositions « d'améliorations » méthodologiques. Ce document constituera ainsi le troisième jalon de notre corpus, représentatif de la période immédiatement successive au « tournant expérimental » en France[17].

Dans les années 2010, le tournant expérimental en France se nourrit de nouvelles évaluations de politiques de l'emploi et poursuit ainsi son développement. En effet, le processus législatif à l'origine de l'ensemble de ces expérimentations – la loi de finances de 2007, prévoyant l'expérimentation du RSA, puis la loi du 1er décembre 2008 le généralisant (et créant le FEJ) – constitue un point de départ, ne s'achevant pas avec la généralisation du RSA. La législature suivante (2012-2017), au bilan en demi-teinte pour certains en matière d'évaluation des politiques publiques (voir par exemple Bozio et Romanello, 2017), ne représente pas non plus un point d'aboutissement. Au contraire, l'expérimentation de la Garantie jeunes, décidée par le décret du 1er octobre 2013, puis généralisée par celui du 23 décembre 2016, a ouvert un terrain nouveau. En particulier, le cas de la Garantie jeunes montre, cette fois dans la pratique, des positionnements nouveaux quant à l'articulation des approches quantitatives et qualitatives. La documentation liée à l'expérimentation de la Garantie jeune constituera ainsi le quatrième jalon de notre corpus.

À partir de 2017, d'autres lois fixeront le cadre pour un nouvel ensemble d'expérimentations relatives aux politiques de l'emploi. Il s'agit notamment de la loi du 5 septembre 2018 pour la liberté de choisir son avenir professionnel, donnant lieu au « plan d'investissement dans les compétences », qui comportera plusieurs expérimentations. De même, un ensemble d'expérimentations s'inscrit dans la « stratégie nationale de lutte contre la pauvreté ». Par exemple, en lien avec ce plan pluriannuel, la loi du 14 décembre 2008 sur l'inclusion dans l'emploi par l'activité économique prolonge et étend l'expérimentation « territoires zéro chômeurs de longue durée », dispositif initié en 2016 à l'initiative d'ATD quart monde (loi du 29 février 2016). Par son ampleur, ce deuxième cas d'évaluation constituera le cinquième jalon de notre corpus.

Enfin, dans la continuité des débats méthodologiques ayant accompagné le « tournant expérimental » en France et le début des années 2010, France Stratégie publie en 2016

[17] La Direction Générale du Trésor publie également, en 2015, un document de travail intitulé « Évaluation des politiques publiques : expérimentation randomisée et méthodes quasi-expérimentales » (Dupont-Courtade *et al*., 2015). Bien que cela s'apparente plus à une revue de littérature et à une réflexion plus générale, plutôt qu'à un « guide » proprement dit (c'est-à-dire, à destination des acteurs de l'expérimentation), il est intéressant de noter cependant la démarche de la DG qui diffuse ainsi, à destination de son lectorat, une présentation synthétique de ces « nouvelles » méthodes, désormais installées dans le paysage français. Ce document ne mentionne pas les méthodes qualitatives et nous l'avons ainsi exclu de notre corpus, bien que cette absence soit intéressante à noter.

Ce document est la version pre-print de l'article paru dans *Socio-économie du travail*, n°15, 2024-1. DOI : https://doi.org/10.48611/isbn.978-2-406-18609-0.p.0117.

son guide méthodologique « Comment évaluer l'impact des politiques publiques ? Un guide à l'usage des décideurs et praticiens ». Le guide est signé Rozenn Desplatz (France Stratégie) et Marc Ferracci (Université Panthéon Assas, CRED, CREST-ENSAE). Le guide France Stratégie de 2016 a été suivi en 2022 par un complément, édité aussi par France Stratégie, qui s'intéresse cette fois précisément à la question de l'articulation entre méthodes quantitatives et qualitatives. Ce guide est signé par Adam Baïz (France Stratégie) et Anne Revillard (LIEPP). Les deux « guides » de France Stratégie, derniers en date dans la période que nous analysons, constituent le sixième élément de notre corpus. L'ensemble du corpus, contextualisé dans cette section, est résumé dans les deux tableaux en Annexe.

Le profil des auteurs des deux guides nous permet de remarquer la collaboration qui s'est instituée, depuis l'époque du FEJ, entre le champ des administrations publiques et le champ universitaire, en particulier de certains laboratoires ou fédérations de recherche. Ces dernières institutions sont devenues des acteurs bien établis dans le domaine des évaluations, répondant à des appels à projets, réalisant des études et rédigeant des rapports. Parmi ces acteurs, on signalera en particulier le LIEPP (Laboratoire Interdisciplinaire d'évaluation des politiques publiques, créé en 2011 et rattaché à Sciences Po), la fédération de recherche TEPP (Théorie et Évaluation des Politiques Publiques, créée en 2008 et s'appuyant sur plusieurs universités), le J-PAL Europe (créé en 2007 et hébergé au sein de l'École d'économie de Paris) et l'Institut des politiques publiques (IPP, réunissant l'École d'économie de Paris et les entités de recherche rattachées à l'ENSAE/ENSAI)[18]. Ces acteurs s'associent également à la réflexion générale sur les méthodes de l'évaluation, et participent indirectement ou directement (comme dans le cas de Ferracci) à la rédaction de guides méthodologiques.

Un deuxième type d'acteur de l'évaluation prend à cette période une nouvelle ampleur et vient jouer un rôle central dans la conduite des évaluations : les cabinets de conseil, en particulier Asdo Etudes, Itinere Conseil, Kantar, KPMG, Quadrant, et Orseu (auteurs de plusieurs volets des rapports constituant notre corpus)[19]. Certaines de ces structures sont ancrées dans le paysage français (Asdo Études, Itinere Conseil, Orseu, Quadrant), tandis que d'autres ne représentent qu'une branche de cabinets internationaux (Kantar, KMPG) ; certains cabinets ont développé une spécialisation dans la réponse à des commandes d'évaluation des politiques publiques (Asdo Études), tandis que d'autres ont des activités plus diversifiées. Malgré cette hétérogénéité, ces différents cabinets de conseils ont en commun d'entretenir également des liens avec la sphère universitaire et/ou les administrations publiques : en particulier, les évaluateurs actifs dans les cabinets ont parfois une formation de recherche (thèse de doctorat), une affiliation secondaire à des laboratoires de recherche (chercheur associé), ou bien des expériences professionnelles au sein de la statistique publique ou des unités de recherche de l'administration (comme, par exemple, la DARES).

[18] Un certain nombre de contributeurs à l'évaluation sont également issus d'autres laboratoires universitaires (Clersé, Lise CNRS et CEET, Triangle…) ou d'établissements publics comme le Céreq.

[19] On signalera, également l'importance d'organismes « à mi-chemin » entre privé et public, tels que le CREDOC, qui « fait appel essentiellement au marché concurrentiel [mais] bénéficie d'une subvention de l'État et est sous tutelle du ministre chargé de la consommation et du commerce » (site du CREDOC).

Ce document est la version pre-print de l'article paru dans *Socio-économie du travail*, n°15, 2024-1. DOI : https://doi.org/10.48611/isbn.978-2-406-18609-0.p.0117.

II. LES GUIDES METHODOLOGIQUES ET LA REFLEXION SUR LES METHODES MIXTES

Les guides méthodologiques jalonnant le développement de l'évaluation des politiques publiques, présentés dans la section précédente, accordent une place plus ou moins grande aux méthodes qualitatives. Dans cette section, nous analysons ces textes en interrogeant, tout d'abord, la définition des « méthodes qualitatives » qu'ils adoptent (II.1). Nous mettrons ainsi en évidence qu'il existe bien des manières différentes, voir opposées, de concevoir ces méthodes. Dans un deuxième temps, nous abordons la question de l'articulation entre la dimension quantitative et qualitative proposée dans ces guides (II.2). Nous montrerons ici un déplacement, qui s'opère entre le guide du FEJ (2009) et le dernier guide de France Stratégie (Baïz et Revillard, 2022), d'un usage de l'outil qualitatif caractérisé principalement par la juxtaposition des méthodes vers une volonté de plus forte intégration, qui justifierait davantage l'emploi de l'expression « méthodes mixtes ».

II.1. QU'EST-CE QU'UNE METHODE « QUALITATIVE » ?

Le guide méthodologique du FEJ (2009, p. 4) définit les méthodes qualitatives comme « les investigations approfondies sur des petits échantillons raisonnés de l'ordre de quelques dizaines d'individus ». Par symétrie, les méthodes quantitatives constituent alors les méthodes « qui s'appliquent sur des échantillons représentatifs (de quelques centaines à quelques milliers d'individus, voire davantage) » (*ibid.*). Cette ligne de partage, par la taille de l'échantillon, est reprise par France Stratégie (2016, p. 37)[20]. Une distinction ensuite critiquée dans le guide plus récent (Baïz et Revillard, 2022, p. 11), qui suggère plutôt de rattacher les deux types de méthodes à des « cultures » épistémologiques différentes : l'une, ancrée dans la médecine et l'épidémiologie (la « culture quantitative ») et l'autre (la « culture qualitative ») empruntant aux sciences sociales (*ibid.*, p. 37-38)[21]. Cette association « d'emprunt » entre « sciences sociales » et méthodes qualitatives pour l'évaluation est bien présent dans tous les guides (FEJ, 2009, p. 18 ; Desplatz et Ferracci, 2016, p. 38). De nouveau, par symétrie, les méthodes quantitatives sont rangées alors du côté de la médecine, de l'épidémiologie, mais aussi des sciences économiques. Pour ces dernières la « culture » quantitative est doublement revendiquée : comme le résultat, consolidé, d'un emprunt aux méthodes de la médecine, pour ce qui concerne les expérimentations aléatoires ; mais aussi comme une « culture » constitutive de la discipline, pour ce qui concerne les techniques économétriques et les méthodes statistiques plus en général (FEJ, 2009, p. 8).

Les guides établissent un périmètre similaire d'exemples de ce qui constitue une « méthode qualitative »[22]. L'entretien semi-directif est le cas le plus cité, parfois associé à d'autres formes d'entretien – entretiens ouverts (FEJ, 2009, p. 18) ou entretiens « approfondis » (CREDOC, 2012, p. 21) – individuels ou collectifs (dont la méthode des

[20] Mais aussi dans le cas du LIEPP : « alors que les méthodes qualitatives fournissent une information riche et dense (et notamment plus indexée à un contexte) sur un petit nombre de cas, les méthodes quantitatives donnent accès à une information plus limitée sur un grand nombre de cas » (Revillard, 2018).

[21] Les auteurs (*ibid.*) écartent par la même occasion une distinction « formelle » qui ferait reposer les méthodes qualitatives sur " des mots » et celles quantitatives sur « des chiffres » – cependant, ce type de distinction n'a pas été utilisé par d'autres guides.

[22] Respectivement, de ce qui constitue une méthode quantitative. On notera au passage qu'il ne semble pas y avoir de méthode « contestée », classée dans une catégorie pour les uns et dans une autre catégorie pour les autres.

Ce document est la version pre-print de l'article paru dans *Socio-économie du travail*, n°15, 2024-1. DOI : https://doi.org/10.48611/isbn.978-2-406-18609-0.p.0117.

« *focus groups* », mentionnée par le FEJ, 2009, p. 18 et par Desplatz et Ferracci, 2016, p. 39). Les « études de cas et monographies » (FEJ, 2009, p. 19), l'« observation directe inspirée des méthodes de type ethnologique » (*ibid.*) et les « questionnaires » (Aldeghi *et al*., 2012, p. 21) complètent ce périmètre des méthodes qualitatives.

Il y a donc une définition commune à ces guides des « méthodes qualitatives », bien qu'ils choisissent chacun d'insister plus particulièrement sur tel ou tel exemple. Au-delà de ce socle commun, un certain nombre de divergences apparaissent quant à d'autres aspects, qui renvoient au fait que les méthodes quantitatives s'appuient en effet sur des domaines disciplinaires parfois éloignés, ou alors en concurrence. Le FEJ, par exemple, assigne aux deux types de méthodes un rôle et une portée bien différents :

> Même si les choses sont en réalité plus nuancées, on peut dire considérer globalement que les approches quantitatives répondent à la question « est-ce que le programme délivre les effets attendus ? », c'est-à-dire « est-ce que le dispositif est efficace du point de vue de l'objectif poursuivi ? ». Elles fournissent des indicateurs chiffrés de performance du programme. Les méthodes qualitatives apportent des éléments de réponses à des questions du type « comment le dispositif permet d'atteindre ces effets ? ». Elles permettent notamment de repérer les obstacles au bon fonctionnement du programme et les leviers sur lesquels il importe d'agir pour en améliorer les effets […] ce qui est particulièrement important lorsque l'on s'interroge sur la généralisation de l'expérimentation. (FEJ, 2009, p. 8)

Nous retrouvons, à la fin de cette citation, le débat sur le critère de validité externe (la « généralisation » des résultats d'une expérience aléatoire). Nous avons vu plus haut que certains auteurs, épistémologues ou critiques des méthodes expérimentales, pointaient la difficulté des expérimentations aléatoires à vérifier ce critère. Le guide méthodologique du FEJ évoque donc cet enjeu, et assigne aux méthodes qualitatives un rôle déterminant dans la construction de la validité externe des expérimentations.

Une division du travail similaire entre méthodes qualitatives et quantitatives est défendue par le guide de France Stratégie (Desplatz et Ferracci, 2016, p. 11-12). Les méthodes quantitatives permettent de quantifier des effets des dispositifs et d'établir des « liens de causalité » entre eux, notamment grâce au « contrefactuel » permis par les expérimentations aléatoires (ou par les méthodes économétriques alternatives). D'autre part, les méthodes qualitatives ont pour rôle d'analyser les « conceptions », les « représentations », les « perceptions » des acteurs impliqués (*ibid*., p. 38)[23]. Cela placerait d'ailleurs les deux approches sur deux plans épistémologiques différents (*ibid.*) : celui de « l'objectivité » (dans le cas des méthodes quantitatives) et celui de la « subjectivité » (pour les méthodes qualitatives).

Si l'on résume cet ensemble d'éléments, les méthodes qualitatives servent donc à préciser le contexte de l'expérimentation et à expliquer (ou comprendre) les comportements des individus. En ce sens, elles complètent l'approche quantitative, dont l'objet est de prouver une relation de cause à effet (l'identification) et d'estimer son ampleur. À partir de ces définitions, quelle articulation envisager alors entre ces deux méthodes ?

[23] Le principal objectif de ces méthodes est alors, dans une perspective compréhensive inspirée de Max Weber, de mettre à jour « le sens que les individus assignent à leurs actions » (Revillard, 2018).

Ce document est la version pre-print de l'article paru dans *Socio-économie du travail*, n°15, 2024-1. DOI : https://doi.org/10.48611/isbn.978-2-406-18609-0.p.0117.

II.2. LES METHODES MIXTES : QUELLE ARTICULATION DU QUANTITATIF ET DU QUALITATIF ?

Le guide méthodologique du FEJ place d'emblée les méthodes quantitatives et qualitatives sur le même pied de légitimité dans la conduite des expérimentations aléatoires et insiste sur leur complémentarité :

> L'évaluation mobilise un ensemble de méthodes *complémentaires* qui permettent de répondre à des questions différentes. (FEJ, 2009, p. 4)

Le guide du FEJ, davantage centré sur les méthodes quantitatives, comporte néanmoins des parties consacrées au qualitatif, envisagé principalement comme « complément », ou comme substitut au quantitatif. L'usage du qualitatif s'impose lorsque le quantitatif ne peut être utilisé, ou pour des objectifs qu'il ne peut appréhender – ce qui découle assez logiquement de la définition du « partage des tâches » analysée plus haut (II.1). Notons que si la notion de « méthodes mixtes » n'est pas évoquée spécifiquement, le FEJ prône cependant d'utiliser le qualitatif « en combinaison avec une approche quantitative », soit en amont de celui-ci avec une fonction exploratoire (quant à l'hypothèse testée et quant aux variables d'intérêt), soit en aval (dans un but interprétatif). La position du FEJ est donc plutôt celle de la juxtaposition des méthodes, celle-ci n'excluant pas une forme d'articulation entre le qualitatif et le quantitatif, devenus interdépendants.

L'apport du CREDOC (2012) sur les méthodes mixtes (expression qui paraît cette fois à deux reprises) tient en partie à la nature du document, à savoir un bilan méthodologique du point de vue de l'évaluateur, en tant qu'utilisateur du guide méthodologique du commanditaire (le FEJ). Il montre bien l'intérêt du qualitatif dans une approche exploratoire en amont, ou interprétative en aval : « l'apport incontestable » des entretiens dans l'évaluation est souligné (Aldeghi et al., 2012, p. 7). Nous retrouvons de nouveau l'enjeu en termes de validité externe, dans l'esprit des indications du FEJ. Mais une des spécificités du guide du CREDOC est de rappeler également l'enjeu en termes de validité interne. En effet, les auteurs du guide soulignent également le rôle des méthodes qualitatives dans le contrôle de la faisabilité et de l'effectivité des différentes étapes de la mise en œuvre des expérimentations en particulier en garantissant la robustesse de l'assignation aléatoire (*ibid*., p. 40).

Quelques années plus tard, le guide de France Stratégie (Desplatz et Ferracci, 2016, p. 40) vient lui aussi expliciter l'articulation entre différentes méthodes d'évaluation. Partant du constat que « de plus en plus de dispositifs d'évaluation mêlent aujourd'hui à la fois des travaux qualitatifs et quantitatifs » (*ibid*., p. 37), les auteurs supportent les « vertus » de cette démarche « mixte » (*ibid*.). Mais le point le plus développé de ce bilan reste le constat d'un usage nécessaire du qualitatif dans les cas, plus nombreux que prévus initialement, où l'usage du quantitatif s'avère impossible ou limité. Ceci a lieu lorsque le nombre d'individus se révèle finalement trop faible. Les méthodes qualitatives restent donc conçues principalement dans une articulation de substitution aux méthodes quantitatives.

L'intitulé du dernier guide de France Stratégie (Baïz et Revillard, 2022) indique lui précisément qu'il est centré sur le questionnement de l'articulation entre quantitatif et qualitatif. Plusieurs méthodes d'articulation sont développées, comme l'analyse

Ce document est la version pre-print de l'article paru dans *Socio-économie du travail*, n°15, 2024-1. DOI : https://doi.org/10.48611/isbn.978-2-406-18609-0.p.0117.

qualitative comparée, qui systématise la description, à l'aide d'outils qualitatifs, des contextes différenciés permettant aux études quantitatives d'estimer des effets hétérogènes ; ou encore la méthode de « faisabilité-impact » qui s'attache à vérifier si chaque étape d'une chaîne de causalité est bien réalisable et effective. Ce dernier guide développe donc explicitement des méthodes qualifiées de « méthodes mixtes ».

III. DE LA THEORIE A LA PRATIQUE : LES METHODES MIXTES DANS LES RAPPORTS ET LES ETUDES

Penchons-nous maintenant sur la pratique de l'évaluation, par l'intermédiaire de l'analyse des études et rapports produits dans le cadre de l'évaluation de différents dispositifs (voir Annexe 1). Nous commençons par montrer que, lorsque les politiques publiques sont évaluées par des méthodes quantitatives et qualitatives, celles-ci sont davantage juxtaposées qu'articulées. Cependant, nous observons aussi un niveau plus élevé d'intégration entre ces deux dimensions dans deux configurations différentes. La première renvoie à certaines études qui revendiquent l'usage de méthodes mixtes, ou bien articulent de fait ces deux dimensions, en les rendant interdépendantes. La deuxième tient à la façon dont des études juxtaposées sont ensuite plus ou moins recombinées dans des rapports finaux. En effet ces derniers peuvent dans certains cas refléter la juxtaposition des études ou des méthodes ou dans d'autres cas s'efforcer de les synthétiser. Enfin, il arrive que l'articulation entre le quantitatif et le qualitatif s'impose avec le temps. Ceci est notamment le cas lorsque l'évaluation d'un dispositif (ou la succession d'évaluations d'une succession de dispositifs proches) s'inscrit dans la longue durée. Nous prendrons l'exemple de la succession RCA/Garantie jeunes (III.1).

Nous nous penchons ensuite sur la nature des méthodes qualitatives déployées dans les différentes études. Nous mettons en évidence que, si les études partagent les mêmes méthodes qualitatives, elles ne sont cependant pas produites dans le même « esprit ». En nous appuyant notamment sur le cas de l'entretien (et, dans une moindre mesure, sur le cas de l'analyse de texte), nous mettrons en évidence que les objectifs, les interprétations et, finalement, les principes méthodologiques sous-jacent à ces méthodes qualitatives, diffèrent selon la communauté d'acteurs (III.2).

III.1. JUXTAPOSITION DES METHODES OU ARTICULATION ?

Dans la plupart des évaluations, les travaux qualitatifs et quantitatifs sont confiés à des équipes différentes. Cette partition est observable dès les premières expérimentations à la fin des années 2000 (RSA, accompagnement renforcé des chômeurs). On la remarque aussi dans la période suivante, après le lancement du FEJ. Ce sont par exemple des équipes bien distinctes qui assurent les évaluations qualitatives et quantitatives du RCA, et diffusent des rapports intermédiaires puis finaux conçus indépendamment. Ce parallélisme va de pair avec une forme de division du travail : les chercheurs de la statistique publique ou des universités investissent le plus souvent les volets quantitatifs de ces études, en particulier lorsqu'il s'agit d'études d'impact relevant de l'analyse causale, tandis que les volets qualitatifs sont fréquemment confiés à des cabinets de

consultants[24]. Cette répartition du travail est presque systématique, par exemple, dans le cadre des récentes expérimentations régionales relatives au Plan d'investissement dans les compétences (PIC), dont le premier rapport (Bucher *et al*., 2020, p. 23, tableau 3) liste des équipes universitaires (École d'Économie de Paris, Le Mans Université, Université de Caen, Université Gustave Eiffel) en charge des volets quantitatifs de l'évaluation, tandis que les volets qualitatifs ont été confiés à des cabinets privés (Itinere Conseil, Geste, Adso-études, Amnyos). Certes, en principe, des études menées séparément, par des équipes distinctes, ne débouchent pas forcément sur une absence d'interaction entre méthodes quantitatives et qualitatives. Cependant, à partir des rapports finaux de certaines évaluations, on observe bien qu'elle se réduisent dans certains cas à une simple juxtaposition des méthodes.

Le plan du rapport de synthèse de l'évaluation du RCA, par exemple, est structuré en deux sous-parties distinctes, séparant les deux dimensions de l'analyse – celle mené avec des méthodes qualitatives et celle menée avec des méthodes quantitatives. De façon générale, les résultats issus des méthodes quantitatives semblent davantage au centre de l'argumentation du rapport de synthèse : certes, celui-ci reproduit bien les conclusions de chaque volet, mais la conclusion générale reprend essentiellement celles du volet quantitatif. Du volet qualitatif n'est retenu que l'effet positif du RCA sur le lien avec les missions locales, tandis d'autres conclusions sont laissées de côté – par exemple, la forte hétérogénéité des effets de la mesure, un ensemble de comportements favorables à l'insertion professionnelle, ou une forme d'autonomisation plus large (liée notamment à une moindre dépendance financière vis-à-vis des parents).

L'évaluation du projet de " Garantie jeunes », qui succède au RCA témoigne en revanche d'une évolution, vers une intégration finale des volets qualitatif et quantitatif plus accentuée : le plan du rapport de synthèse ne distingue plus le qualitatif et le quantitatif (aucun de ces termes ne figure en titre de partie ou de sous-partie), ce qui traduit formellement une mobilisation des deux dimensions dans l'ensemble du rapport.

Malgré la prédominance d'évaluations où les méthodes qualitatives et quantitatives sont juxtaposées, un certain nombre d'évaluations croisent davantage ces deux dimensions. Cette approche est parfois imposée par le type de politique évaluée. Par exemple, s'il s'agit d'une politique de l'emploi basée sur une réforme du droit du travail, son évaluation impliquera vraisemblablement un travail qualitatif réalisé par des juristes avant toute approche quantitative. Ainsi les études relatives à l'évaluation des « ordonnances travail » relèvent pour certaines d'entre elles de ce type d'intégration, qui peut être associé à un usage revendiqué d'une méthode mixte. Dans leur travail sur le représentant de proximité, Jérôme Pelisse et ses coauteurs cherchent à mesurer la façon dont les salariés se sont approprié cette institution nouvelle. Pour ce faire ils revendiquent l'utilisation « d'un protocole d'enquête interdisciplinaire et des méthodes mixtes » (Pelisse et Wolmark, 2021), précisant la façon dont un travail qualitatif d'analyse de texte (un ensemble d'accords d'entreprises) leur permettra de construire des variables quantitatives :

> Les juristes de l'équipe ont rempli cette grille de saisie (qui) a ensuite été transformée en grille de codage pour ne conserver que quelques éléments, notamment ceux sur lesquels il est pertinent et possible de produire des statistiques, ainsi transformés en variables. (*ibid*., p. 9)

[24] Plus rarement certaines études, telle qu'une des études qualitatives sur la Garantie jeunes (Farvaque et Tuchszirer, 2018) sont réalisées par des équipes combinant des auteurs issus de cabinets privés et de laboratoires universitaires.

On reconnaît dans ce passage une forme de méthode mixte où l'articulation entre le qualitatif et le quantitatif consiste à fonder la construction d'une variable quantitative sur un travail qualitatif préalable – une possibilité évoquée aussi dans le guide du FEJ.

Toujours dans le cadre de l'évaluation des « lois travail », d'autres équipes procéderont également avec une méthodologie de ce type, permettant de mettre en évidence des changements de comportement attribuables à certaines de ces ordonnances. Par exemple, Dalmasso *et al*. (2023) soulignent certaines ruptures de tendances, notamment une augmentation du nombre des licenciements pour faute. Ce résultat est obtenu en croisant « une méthodologie quantitative – statistique ou économétrique – et qualitative – analyse par cas-type ou de texte juridiques – sur des données collectées ou constituées de différentes manières » (*ibid*., p. 297). Une autre évaluation récente employant des méthodes que nous pouvons qualifier de mixtes, bien que les auteurs n'utilisent pas ce terme, a été conduite par Sylvain Chareyron et Yannick L'Horty (2023) au sujet de la discrimination dans l'accès au logement[25]. Cette évaluation, basée sur la méthode quasi-expérimentale du *testing*, consiste à observer la façon dont les employés des guichets d'enregistrement des demandes de logement social répondent à des demandes de renseignement rédigées de manière quasi-identiques, mais émanant de patronymes différents (à consonance européenne pour les uns ou africaine pour les autres). Pour aller au-delà de l'observation de différences de taux de réponse (variable classiquement mesurée dans le cadre des *testing*), les auteurs s'appuient sur une analyse textuelle des réponses reçues des guichets. Sur la base de leur analyse du *contenu* des réponses (notamment l'ajout d'« informations décourageantes » ou encore une différence d' « orientation » ou d' « intensité d'accompagnement » du demandeur), Chareyron et L'Horty construisent une variable quantitative, nommée « discriminants réponse et contenu ».

Les exemples ci-dessus illustrent donc l'usage de méthodes mixtes où les dimensions qualitatives et quantitatives sont intégrées et articulées dans le cadre de l'évaluation d'*une* mesure ou d'*un* dispositif. En élargissant la focale à l'évaluation d'*une suite* de mesures (renvoyant, toutes, à un même objectif de politique publique), on révèle que ce type d'articulation et d'intégration peut apparaître avec le temps long de la succession des études. L'analyse des études sur l'expérimentation du RCA puis celles sur l'expérimentation de la Garantie jeunes (aux objectifs proches, portant sur l'insertion de jeunes éloignés de l'emploi) permettent d'observer une telle suite longue de plusieurs années, jalonnée de rapports intermédiaires puis finaux, qualitatifs, quantitatifs ou de synthèse[26]. L'analyse de cette suite montre une évolution progressive du cadre de l'évaluation, à laquelle contribue, précisément, l'intégration répétée des méthodes qualitatives et quantitatives – avec, par exemple, l'évolution des variables d'intérêt en fonction des résultats intermédiaires.

Le volet qualitatif des études sur le RCA a insisté sur l'hétérogénéité des effets du dispositif, sur son effet positif sur l'autonomie, ainsi que sur le renforcement du lien entre les jeunes et la mission locale ; la partie quantitative concluait à l'absence d'effet global

[25] Bien qu'il ne s'agisse pas d'une évaluation de politique de l'emploi, nous incluons cette étude dans notre corpus car elle illustre les développements récents de la méthode de *testing*, qui en tant que méthode expérimentale, est largement utilisée pour étudier les phénomènes de discrimination sur le marché du travail. Ainsi ce type d'études documente certains rapports d'évaluation de politiques de l'emploi, comme le *Rapport d'évaluation sur l'expérimentation des emplois francs* (Bonnetête *et al*., 2019).

[26] A notre connaissance il n'existe pas d'autres politiques publiques (en dehors peut-être des études sur la mobilité des jeunes) ayant fait l'objet d'une analyse aussi longue et approfondie, à travers plusieurs expériences aléatoires.

du RCA sur l'insertion professionnelle – ce qui était pourtant son but principal. Ces deux derniers résultats appelaient à une réflexion sur les modalités de l'accompagnement par les missions locales, puisqu'un lien plus fort avec celles-ci (conclusion du volet qualitatif) n'améliorait pourtant pas la situation d'emploi (conclusion du volet quantitatif).

Le projet de Garantie jeunes se développe de façon cohérente avec ces résultats[27]. D'une part, il propose justement une modalité d'accompagnement très différente (basée sur un travail en groupes et sur le principe de « l'emploi d'abord »). D'autre part, il clarifie les conditions de cumul avec les revenus du travail, ainsi que la question du respect des engagements par les jeunes, pouvant désormais conduire à la suppression de l'allocation.

Ces nouvelles modalités d'accompagnement visent à limiter les effets désincitatifs décelés par l'évaluation du RCA : à la fois dans son volet quantitatif – qui relève une légère baisse du taux d'emploi en début de période – et dans son volet qualitatif. Ce dernier relaie, par exemple, la parole de certains conseillers qui regrettent de ne pas disposer de « levier pédagogique », c'est-à-dire d'une forme de pouvoir de coercition. Ils regrettent aussi, comme les jeunes eux-mêmes, le manque de lisibilité du système de cumul et de dégressivité de l'allocation. On pourrait donc dire que les deux types d'outils convergent pour questionner certaines caractéristiques du dispositif et, in fine, pour contribuer à les modifier.

L'articulation entre qualitatif et quantitatif se montre ici particulièrement adaptée à la redéfinition du ciblage d'une mesure de politique publique, notamment car les méthodes qualitatives s'avèrent particulièrement aptes à détecter l'hétérogénéité des effets des dispositifs[28]. En effet, l'hétérogénéité des effets du RCA, mise en évidence par le volet qualitatif de l'évaluation, a été le point de départ pour la définition du ciblage de la Garantie jeunes[29]. La catégorie ciblée par la Garantie jeunes (des jeunes sans emploi ni formation, et pour la plupart n'habitant plus avec leurs parents) est bien plus précise que celle concernée par le RCA. Cependant, selon les études qualitatives réalisées lors de l'évaluation du RCA, cette catégorie de jeunes est bien la plus susceptible de réagir positivement au dispositif. La pertinence de ce nouveau ciblage par la Garantie jeunes fera également l'objet d'une nouvelle étude qualitative (Couronné et Sarfati, 2018).

En guise de bilan quant au niveau d'intégration des volets qualitatifs et quantitatifs nous voyons donc que si le cas général est plutôt celui de la juxtaposition des méthodes, celle-ci peut refléter une intégration plus ou moins grande selon la façon dont les rapports de synthèse croisent les deux dimensions de l'analyse. On observe aussi une articulation plus forte dans certains travaux où les variables quantitatives d'intérêt sont construites en s'appuyant sur l'analyse qualitative. Dans certains cas, le terme de méthode mixte est revendiqué explicitement par les auteurs Enfin, nous voyons que dans le cas d'une suite de mesures ou dispositifs correspondant à un même objectif de politique publique (par exemple la série RCA-Garantie jeunes), le degré d'articulation et d'intégration s'accroit avec le temps long de la succession des études.

---

[27] Les rapports concernant la Garantie jeunes n'explicitent pas dans quelle mesure les volets qualitatifs et quantitatifs des études sur le RCA ont été pris en compte pour définir le nouveau projet de la Garantie jeune. Cependant, il est probable que ces études aient contribué à sa définition (Aeberhardt *et al*., 2014).

[28] Les méthodes quantitatives mobilisées dans les expérimentations aléatoires estimant principalement un effet moyen, leur capacité à rendre compte de l'hétérogénéité est très réduite (Favereau, 2021).

[29] L'idée d'un ciblage insuffisant du RCA est évoquée par le volet qualitatif, mais n'est pas vraiment retenue dans le rapport de synthèse.

Ce document est la version pre-print de l'article paru dans *Socio-économie du travail*, n°15, 2024-1. DOI : https://doi.org/10.48611/isbn.978-2-406-18609-0.p.0117.

III.2. QUELLES METHODES QUALITATIVES ? LA LETTRE ET L'ESPRIT

Le périmètre des méthodes qualitatives est généralement désigné, par les guides méthodologiques, comme relevant des « sciences sociales » et recouvre les mêmes éléments, dont notamment l'entretien (sous différentes formes), l'analyse textuelle et l'enquête ethnographique (*Cf. supra*, II.1). Dans la pratique, c'est-à-dire dans notre corpus de rapports et études, on retrouve bien ces éléments et, en particulier, le recours à l'enquête – pratiquement systématique dans la plupart des études. Les acteurs mettant en œuvre ces méthodes appartiennent certes à la vaste catégorie des « sciences sociales », mais cette appellation recouvre en réalité une diversité disciplinaire : enquêtes et analyse textuelle sont pratiquées aussi bien, dans notre corpus, par des sociologues, des économistes, des gestionnaires, des politistes – et, dans deux cas, des juristes[30]. Mais l'étude de ces pratiques met en évidence une différence encore plus fondamentale, qui ne recoupe pas exactement les appartenances disciplinaires des acteurs. En effet, si la « lettre » de l'usage des méthodes qualitatives semble bien commune à toutes les études, l' « esprit » dans lequel ces méthodes sont construites, définies, mobilisées diffère significativement selon les communautés d'acteurs. Cette variabilité vient ainsi nuancer l'idée, avancée dans le guide méthodologique de France Stratégie (Baïz et Revillard, 2022 ; *Cf. supra*, II.2) qu'il existerait une unique « culture qualitative » : nous mettons bien en évidence qu'il existe au fait, dans la pratique, des *«* cultures qualitatives » plurielles, dont les contours sont plus complexes et les ancrages plus « locaux » que les simples catégories disciplinaires (économie, sciences de gestion, sociologie, etc.). Deux clivages généraux se dessinent toutefois : d'une part, sur la capacité des méthodes qualitatives à rendre compte d'une relation causale ; d'autre part, la question du positionnement du chercheur vis-à-vis du commanditaire. Cette dernière question renvoie à deux positionnements : l'un, de type ethnographique (où le cadre interprétatif se construit progressivement) n'exclut pas un positionnement parfois critique ; l'autre, d'un type plus proche de celui du conseil, autorise une plus grande proximité avec le commanditaire.

Si nous examinons en premier lieu l'utilisation des outils qualitatifs au prisme de la question de la causalité, nous voyons que le point de vue le plus courant réserve cette caractéristique aux outils quantitatifs. Néanmoins ce partage des rôles n'est pas exempt d'ambiguïtés. Par exemple, le rapport qualitatif d'évaluation de la Garantie jeunes distingue l'effet « pur » du dispositif, qui ne peut être dégagé qu'en s'appuyant sur un contrefactuel de certaines « inflexions » que l'analyse qualitative pourrait toutefois attribuer au dispositif :

> Nous ne sommes alors pas en mesure de déterminer des effets « purs » du dispositif. Toutefois, le travail d'analyse du matériau d'enquête permet de repérer des inflexions dans les parcours socio-professionnels que l'on peut attribuer à la Garantie jeunes. Nous considérons alors ces inflexions comme étant produites par le dispositif. Elles constituent alors des effets de la Garantie jeunes sur les parcours des jeunes. (Loison Leruste *et al*., 2016, p. 57)

Ce positionnement parfois ambigu vis-à-vis de la causalité peut se révéler lorsqu'il s'agit d'interpréter les typologies produites par certaines études qualitatives. Ces typologies dévoilent souvent les effets différenciés d'une politique publique, et

[30] De plus, comme nous l'avons déjà remarqué (III.1) les volets qualitatifs des enquêtes sont le plus souvent confiés à des cabinets de conseil.

contribuent de ce fait à expliquer ou comprendre les mécanismes à l'origine de la relation de cause à effet. Celle-ci est en revanche estimée en termes d'effet moyen par les volets quantitatifs, selon le partage des rôles évoqué plus haut. Mais ce partage des rôles est moins évident lorsque la typologie se prête elle-même à une interprétation en termes de causalité. La lecture du seul volet qualitatif donne toutefois de nombreux éléments appuyant l'idée d'un effet positif du RCA sur l'emploi. Dès lors, comment interpréter les résultats contraires issus du volet quantitatif selon lequel « tous les effets négatifs envisagés se sont manifestés, et aucun des effets positifs » ? L'étude précise que cette conclusion n'est pas due à un manque de puissance statistique. Faut-il dès lors exclure la question de la causalité du domaine méthodes qualitatives ? Ou au contraire faut-il s'appuyer sur celles-ci pour réévaluer les conclusions des études d'impact, élargir leur champ d'investigation, allonger le terme de l'étude pour estimer ce qui aurait échappé à une première stratégie d'identification ?

Dans d'autres cas, l'analyse qualitative est plus explicitement destinée à établir une relation de cause à effet, notamment si en croisant celle-ci avec la dimension quantitative il s'agit d'améliorer la définition des variables d'intérêt. Nous avons évoqué cette configuration dans la partie précédente, telle qu'elle est utilisée par Chareyron et L'Horty (2023), par Dalmasso *et al*. (2023) ou encore Pelisse et Wolmark (2021).

Enfin, lorsque les auteurs revendiquent l'utilisation de méthodes mixtes, la notion de causalité a fréquemment une place importante dans l'analyse. Elle peut être envisagée à travers une théorie explicitement mobilisée, qui relève d'une des façons dont les méthodes mixtes envisagent la relation causale (Aguilera et Chevalier, 2021). Il s'agit notamment de décrire la chaîne de causalité, c'est-à-dire les différentes étapes qui conduisent une mesure à produire un effet : établir comment et à quel degré le public est informé, puis identifier les contextes associés à l'observation d'un effet, ou d'un changement de comportement, etc. Cette notion de « chaîne de causalité » est propre aux méthodes mixtes relevant de l' « évaluation basée sur la théorie » qui, comme le rappelle Devaux-Spatarakis (2023), s'appuie sur des méthodes qualitatives et quantitatives. Par exemple l'évaluation de « Prépa-projet », une des mesures du PIC, bien que qualitative, mobilise aussi des données quantitatives et comporte de nombreux revois au volet quantitatif de l'évaluation. Nous relevons dans notre corpus ce type d'approche dans plusieurs évaluations relatives au PIC, mais aussi dans d'autres travaux. L'étude Kantar sur l'évaluation des emplois francs s'inscrit dans ce cadre, on trouve également dans les « compléments au rapport d'évaluation sur la stratégie nationale de lutte contre la pauvreté » (2021) plusieurs travaux décrivant une « chaine causale ».

Précisons toutefois que dans ces études, la culture de la causalité, est davantage mobilisée dans une perspective de conseil que dans celle de l'étude d'impact. Nous revenons sur ce point dans le prochain paragraphe où nous examinons comment différentes cultures qualitatives se distinguent par un positionnement et une distance différente vis-à-vis du commanditaire.

Cette question du positionnement de l'évaluateur est très clairement posée par le guide du CREDOC, rendant compte des difficultés de l'évaluateur, en particulier lorsqu'il est sociologue, à mettre en œuvre sa propre culture qualitative :

> La position de l'évaluateur est délicate au sein d'une expérimentation. Une difficulté rencontrée à chaque étape de l'évaluation et, en particulier, au moment de porter un jugement évaluatif, est celle de maintenir une distance et une neutralité dans l'élaboration des conclusions. (Aldeghi *et al*., 2012, p. 55).

Ce document est la version pre-print de l'article paru dans *Socio-économie du travail*, n°15, 2024-1. DOI : https://doi.org/10.48611/isbn.978-2-406-18609-0.p.0117.

On peut toutefois trouver dans les nombreuses études qualitatives qualifiées de « monographie » cette forme de distance et de neutralité, dans la mesure où il s'agit de croiser plusieurs outils qualitatifs (observations, entretiens, analyse de textes et d'images) dont un des objectifs est de décrire de manière approfondie l'objet de l'étude. On trouve par exemple ce type d'approche dans « Territoire Zéro Chômeur à Jouques » (Adso Etudes et ANSA pour la DARES, 2019), étude jalonnée de longs extraits d'entretiens et d'analyses de documents tels qu'un ensemble de photographies.

Ces travaux ne sauraient toutefois rester purement descriptifs : l'exploitation d'un corpus d'entretiens a vocation à être interprété ou examiné à travers certaines hypothèses. La nature de ce travail renverra à un positionnement différent du chercheur selon que, dans une perspective plus ethnographique, le cadre interprétatif sera élaboré progressivement, ou selon qu'il s'agira de tester des hypothèses formulées en amont, notamment lors de la « fabrique » de la politique évaluée. Dans le premier cas, l'évaluateur n'est pas tenu de « rester dans la logique des projets et de ne pas questionner la problématique de l'expérimentateur » (CREDOC, 2012). L'entretien permet alors « d'interroger les prénotions, de faire le travail de construction de l'objet spécifique à la démarche du sociologue et d'adopter une posture critique. » (*ibid.*). Dans le second cas, l'évaluation ne se prêtera pas à ce travail critique autorisant le questionnement sur les principes fondant le dispositif évalué. Cette façon d' « opérer une rupture par rapport au sens commun et mettre en évidence les présupposés tacites de ses sujets d'étude » (*ibid.*) est en revanche bien présente dans l'étude de Couronné et Sarfati (2018) sur la Garantie jeunes. Ils soulignent notamment les limites de la catégorie de jeunes « sans emploi ni formation » utilisée pour définir la population ciblée par le dispositif. Cette catégorie, définie par une double négation, occulterait l'expérience réelle des jeunes, dont le parcours antérieur inclut bien des périodes de formation ou de travail. Les deux auteurs s'appuient sur les entretiens pour montrer que c'est au contraire l' « expérience de la précarité, acquise par leur expérience personnelle, mais également à travers celles de leurs parents et de leurs frères et sœurs, [qui] conditionne leur rapport au travail et à l'emploi » (*ibid.*). En conséquence, certains éléments du dispositif sont inadaptés à leur cible. L'étude rapporte aussi des séances inadaptées aux besoins du public ou encore construites sur des fondements théoriques discutables (Loison Leruste *et al.*, 2016). Ainsi le bilan méthodologique du Crédoc (Aldeghi *et al.*, 2012) oppose l'injonction à « rester dans la logique des projets » et « ne pas questionner la problématique de l'expérimentateur » pour « vérifier l'impact des projets », aux fondements du métier de sociologue, qui au contraire se doit « d'interroger les prénotions, de faire le travail de construction de l'objet spécifique à la démarche du sociologue et d'adopter une posture critique » (*ibid.*). Toutefois ce positionnement ne concerne pas que les sociologues, nous le constatons par exemple lorsqu'une des équipes, constituée uniquement d'économistes, chargée de l'évaluation qualitative de TZCLD consacre un long point de leur rapport aux limites dues aux « impensés de TZCLD » (Fretel *et al.*, 2020).

Un autre point de vue consiste à mener des entretiens dans le but d'observer et de décrire la mise en œuvre d'un dispositif de politique publique, en soulignant parfois les dysfonctionnements, mais en se plaçant toujours en aval de la définition des principes du dispositif. Dans ce cas il s'agit notamment, dans une perspective de conseil, d'aider les concepteurs des dispositifs à provoquer un changement de comportement dans la population ciblée. Par exemple le cabinet Kantar, chargé de l'enquête qualitative auprès des employeurs pour l'évaluation du dispositif des « emplois francs », revendique le cadre théorique du « *behaviour change* », un modèle qui « permet d'aborder l'ensemble des dimensions conscientes ou moins conscientes, susceptibles d'intervenir dans l'adoption

ou non d'un comportement ». Toujours concernant l'évaluation des emplois francs, Thomas Delahais, pour le cabinet Quadrant-conseil, précise le cadre de l' « évaluation réaliste » impliquant de mettre à jour les mécanismes et la « théorie du changement » permettant à la mesure de modifier les comportements : « Pouvoir identifier et caractériser ces processus de décision (appelés 'mécanismes') et les éléments de contexte qui les favorisent ou les obèrent est très précieux pour l'action publique ; et c'est l'objet de l'évaluation réaliste. Celle-ci permet d'aboutir à un 'code de la route' détaillant les modalités, les profils, les éléments de contexte auxquels il faudra être attentif pour assurer un déploiement efficace du dispositif ». Cet extrait d'encadré méthodologique nous montre également que ces outils sont mobilisés pour évaluer l'impact de la mesure mais aussi pour améliorer son efficacité. Une méthode proche est appliquée dans certaines évaluations quantitatives du PIC ; ainsi, avant d'exposer leurs conclusions, les auteurs du rapport qualitatif d'évaluation du dispositif « Prépa projet », de même que l'équipe ayant évalué « Prépa compétence »[31], rappellent la « théorie d'action », c'est-à-dire la chaîne causale reliant les principes du dispositif à son effet final, en passant par des objectifs et des effets intermédiaires. Les recommandations des auteurs sont ensuite présentées au regard de chacune de ces étapes. Il s'agit notamment de savoir dans quelle mesure le dispositif est connu des acteurs, ou bien en quoi sa mise en œuvre est conforme à ses principes, notamment celui de l'individualisation des parcours, ou encore comment les partenaires du dispositif sont réellement mobilisés, conformément aux objectifs affichés. Il ne s'agit donc pas seulement d'évaluer mais aussi de formuler des recommandations.

CONCLUSION : QUELQUES LEÇONS GENERALES

Dans cet article, nous avons étudié un corpus original constitué de « littérature grise » produite par les acteurs de l'évaluation des politiques publiques en France depuis une vingtaine d'année, en particulier dans le cadre des politiques de l'emploi. Notre analyse a contextualisé les positions et les débats de notre corpus dans le contexte d'un « tournant expérimental », qui se produit, en France, au milieu des années 2000. Notre approche nous a permis de mettre en évidence trois résultats.

Premièrement, nous avons mis en évidence que les discours (les guides méthodologiques) et les pratiques (les rapports d'évaluation) accordent une place légitime aussi bien aux méthodes quantitatives que qualitatives. Cependant, les définitions de ces méthodes ne sont pas univoques. Nous avons constaté que la mise en œuvre de méthodes qualitatives s'appuyait sur des cadres théoriques, voire disciplinaires différents. L'usage d'outils tels que les entretiens semi-directifs est largement partagé par un grand nombre d'approches, et sert dans la plupart des études à préciser le contexte dans lequel s'inscrit le dispositif évalué, à expliquer les mécanismes par lesquels il agit, en révélant le plus souvent l'hétérogénéité de ses effets. Nous avons également relevé que, dans plusieurs cas, l'appel aux méthodes qualitatives est conçu comme un moyen de renforcer la validité externe des expérimentations. Mais ces outils sont parfois exploités dans des logiques très différentes. Un des principaux clivages tient à la position du chercheur vis-à-vis des hypothèses sur lesquelles repose le dispositif évalué, notamment selon qu'il s'agisse d'interroger et de discuter les catégories utilisées, dans une tradition plutôt sociologique

[31] Crovella *et al*. (2021 ; 2022). Ces deux études, de nature qualitative, ont été réalisées pour la DARES par les cabinets Orseu et Itinere Conseil.

ou bien qu'il s'agisse d'évaluer l'effectivité ou l'efficacité de leur mise en œuvre, avec la perspective d'établir un certain nombre de préconisations.

Deuxièmement, nous avons mis en évidence que l'une des raisons de la « juxtaposition » des méthodes est la division du travail opérée dans le cadre des dispositifs d'évaluation (ou le « polycentrisme » des dispositifs d'évaluation). Certains acteurs (le plus souvent des cabinets de conseil) sont chargés du volet qualitatif des enquêtes, tandis que le volet quantitatif est l'apanage des acteurs à la frontière entre le monde universitaire et le monde des administrations publiques.

Troisièmement, nous avons observé que les guides méthodologiques (à l'exception du guide récent de France Stratégie ; Baïz et Revillard, 2022) conçoivent plutôt une « juxtaposition » des méthodes qualitative et des méthodes quantitatives, au lieu d'une véritable articulation entre les deux (« méthodes mixtes »). Cette articulation est cependant présente dans plusieurs rapports et études, même si, de nouveau, le procédé exact de cette articulation répond à des logiques différentes.

La prévalence de la « juxtaposition » soulève la question : est-ce que ce type de recours aux méthodes qualitatives permet véritablement de renforcer la validité externe des expérimentations ? Ou est-ce simplement le résultat d'une logique de « division du travail » et, donc, d'une logique de pouvoir plutôt que d'une logique de savoir ? Dans cet article, notre analyse suggère plutôt que les logiques de pouvoir sont prévalentes dans le cas étudié ; cependant, cette conclusion nécessite de compléter notre analyse, dans deux directions. D'une part, par l'approfondissement de l'analyse des pratiques. Une véritable socio-histoire de l'évaluation des politiques publiques en France nécessiterait une exploration plus fine des acteurs impliqués, de leurs interactions, de leurs logiques de savoir et de pouvoir. En particulier, un pas ultérieur consisterait à accompagner l'analyse des textes (telle que nous l'avons menée) par l'analyse d'autres sources, en particulier d'archive (par exemple, l'étude des appels d'offre/appels à projets), ainsi que par le recours à l'histoire orale, voire à l'observation ethnographique des évaluateurs et des commanditaires. La multitude d'acteurs et de lieu de l'expérimentation représentent cependant un défi pour cette démarche, comme le soulignait déjà Barbier (2010). D'autre part, une comparaison du cas français avec d'autres cas d'études permettrait de mieux mettre en perspective les débats présentés ici. Une perspective comparative se justifie par ailleurs dans la mesure où le monde de l'évaluation est perméable aux normes méthodologiques et au débats qui ont lieu à l'échelon européen ou international, en particulier lorsqu'il est question d'acteurs du monde universitaire.

REFERENCES BIBLIOGRAPHIQUES

Asdo études et l'ANSA pour la DARES, 2019, Territoire Zéro Chômeurs à Jouques, Paris DARES/Ministère du Travail.

Aeberhardt R., Chiodi V., Crépon B., Gaini M., Vicard A., 2014, « Du revenu contractualisé d'autonomie à la garantie jeunes : les enseignements d'une expérimentation », *De l'éducation à l'insertion : dix résultats du Fonds d'expérimentation pour la jeunesse*, sous la direction de Bérard J. et Valdenaire M., Paris La Documentation française/ INJEP, p. 133-148.

Aguilera T., Chevalier T., 2021, « Les méthodes mixtes pour la science politique », *Revue française de science politique*, vol. 71, n° 3, p. 365-389.

Aldeghi I., Jauneau P., Muller L., 2012, « Bilan méthodologique des évaluations conduites par le CRÉDOC dans le cadre du Fonds d'expérimentation pour la Jeunesse », Cahiers de recherche du CREDOC, n° 300.

Backhouse R. E., Cherrier B., 2017, « The Age of the Applied Economist: The Transformation of Economics since the 1970s », *History of Political Economy*, vol. 49, Supplement annuel, p. 1-33.

Baïz A., Revillard A., 2022, Comment articuler les méthodes qualitatives et quantitatives pour évaluer l'impact des politiques publiques ? Un guide à l'usage des décideurs et des praticiens, Paris France Stratégie.

Banerjee A. V., Duflo E., 2009, « L'approche expérimentale en économie du développement », *Revue d'économie politique*, vol. 119, n° 5, p. 691-726.

Barbier J.-C., 2010, « Éléments pour une sociologie de l'évaluation des politiques publiques en France », *Revues française des affaires sociales*, vol. 1, p. 25-49.

Behaghel L., Crépon B., Guitard J., Gurgand M., 2008, « Évaluation d'impact de l'accompagnement des demandeurs d'emploi. Rapport intermédiaire », Paris.

Behaghel L., Crépon B., Gurgand M., Kamionka T., Lequien L., Rathelot R., Zamora P., 2013. « L'accompagnement personnalisé des demandeurs d'emploi », *Revue française d'économie*, vol. 131, n° 1, p. 123-158.

Bonnetête F., Firquet S., Rey M., Vignolles B., Audras-Marcy S., Aussedat M., Dieusaert P., Thibuad S., Suster L., Lestienne L., Lamand M., Jouve N., 2019, Rapport d'évaluation sur l'expérimentation des emplois francs, Paris : Ministère du travail.

Bourguignon F. (dir.), 2008, Bilan d'étape sur l'évaluation des expérimentations RSA, Paris Comité d'évaluation des expérimentations.

Bourguignon F. (dir.), 2009, Rapport final sur l'évaluation des expérimentations RSA, Paris Comité d'évaluation des expérimentations.

Bozio A., Romanello L., 2017, « Évaluation des politiques publiques : le bilan contrasté du quinquennat », Les notes de l'IPP, n° 25, Paris : Institut des politiques publiques.

Bucher A., Fretel A., Zamora P., Poujouly C., Vanni N., 2020, Premier rapport du comité scientifique du Plan d'investissement dans les compétences, Paris DARES.

Bureau M.-C., Sarfati. F., Simha J., Tuchszirer C, 2013, « L'expérimentation dans l'action publique. Usages, pratiques et jugements des acteurs », *Travail et emploi*, vol. 135, n° 3, p. 41-55.

Capelier T., Mizrhai R., 2008, « L'évaluation qualitative de la mise en œuvre des expérimentations », L'accompagnement renforcé des demandeurs d'emploi, n° 2, Paris ANPE-Dares-Unédic.

Cartwright N., 2007, "Are RCTs the Gold Standard?", *BioSocieties*, vol. 2, n° 1, p. 11-20.

Cartwright N., Hardie J., 2012, *Evidence-Based Policy. A Practical Guide To Doing Better*, Oxford University Press.

Chareyron S., L'Horty Y., 2023, « Discriminations ethno-raciales dans l'accès au logement social : un test des guichets d'enregistrement », Rapport de recherche TEPP, n° 2023-3.

Couronné J., Sarfati F., 2018, « Une jeunesse (in)visible : les 'Neets vulnérables' de la Garantie jeunes », *Travail et Emploi*, vol. 153, p. 41-66.

Crépon B., 2008, « L'apport des expérimentations dans l'évaluation de l'impact des dispositifs publics », *Informations sociales*, vol. 150, n° 6, p. 56-67.

Crovella E., Farvaque N., Ulmann A.-L., Lextrait V., Marchat Y., Messaoudi D., 2021, « Étude de l'impact du programme 'prépa compétences' sur l'accès à la formation des demandeurs d'emploi dans le cadre du PIC », Rapport d'études de la DARES, n° 007.

Crovella E., Farvaque N., Magnier J., Marchat Y., Messaouid D., Mokeddem L., Bergoend L., 2022, « Évaluation qualitative du dispositive Prépa-Projet mené par la région Bretagne dans le cadre du PIC », Rapport d'études de la DARES, n° 43.

Deaton A., Cartwright N., 2018, "Understanding and misunderstanding randomized controlled trials", *Social Science & Medicine*, vol. 210, n° 18, p. 2-21.

de Curraize Y., Thoron S., 2020, « À propos du statut épistémologique des expériences en économie », *Revue d'économie politique*, vol. 130, n° 4, p. 545-572.

Dalmasso R., Signoretto C., Etiennot A., Etiennot P., Marié R., 2023, Relations de travail et d'emploi, et comportements des acteurs dans un contexte de réformes législatives, Paris : Institut François Gény, Ladyss.

DARES, 2020, Premier rapport du comité scientifique du Plan d'investissement dans les compétences, Paris DARES, https://dares.travail-emploi.gouv.fr/sites/default/files/pdf/dares_pic_comite_scientifique_1er_rapport_oct2020.pdf (consulté le 29/10/2023).

Delahais T., Devaux-Spatarakis A., Revillard A., Ridde V. (dir.), 2020, *Évaluation. Fondements, controverses, perspectives*. Québec : Éditions science et bien commun.

Delahais T., Devaux-Spatarakis A., 2018, « Évaluation des politiques publiques et sociologie : état des lieux d'une relation distanciée », *Sociologies pratiques*, vol. 36, n° 1, p. 47-56.

Deleau M., Nioche J.-P, Penz P., Poinsard R., 1986, *Évaluer les politiques publiques. Méthodes, déontologie, organisation*, Paris, La Documentation française.

Desrosières A., 2003, « Historiciser l'action publique : l'État, le marché et les statistiques » in *Historicités de l'action publique*, sous la direction de P. Laborier et D. Trom, Paris PUF, p. 207-221.

Desrosières A., 2000, *La politique des grands nombres. Histoire de la raison statistique*, 2e édition, Paris, La Découverte.

Desrosières A., 2008a, *Pour une sociologie historique de la quantification. L'argument statistique 1*, Paris, Presses de l'École des Mines.

Desrosières A., 2008b, *Gouverner par les nombres. L'argument statistique 2*, Paris, Presses de l'École des Mines.

Desrosières A., 2014, *Prouver et gouverner. Une analyse politique des statistiques publiques*, Paris, La Découverte.

Desplatz R., Ferracci M., 2016, Comment évaluer l'impact des politiques publiques ? Un guide à l'usage des décideurs et praticiens, Paris, France Stratégie.

Devaux-Spatarakis A. 2023, « L'évaluation basée sur la théorie », LIEPP Fiche méthodologique, n° 27.

Dupont-Courtade L., Chabé-Ferret S., Treich S., Perbet M., 2015, « Évaluation des Politiques Publiques : expérimentation randomisée et méthodes quasi-expérimentales », Les Cahiers de la DG Trésor, n° 2015-01, Paris DG Trésor.

Farvaque N., Tuchszirer C., 2018, « La Garantie jeunes en pratique dans les missions locales : une expérimentation encadrée », *Travail et emploi*, vol. 153, n° 1, p. 15-40.

Favereau J., 2021, *Le hasard et la preuve : apports et limites de l'économie expérimentale du développement*, Lyon ENS éditions.

Fonds d'expérimentation pour la jeunesse (FEJ), 2009, Guide méthodologique pour l'évaluation des expérimentations sociales à l'intention des porteurs de projets, Paris : FEJ.

Fougère D., 2000, « Expérimenter pour évaluer les politiques d'aide à l'emploi : les exemples anglo-saxons et nord-européens », *Revue française des affaires sociales*, vol. 1, p. 111-141.

Fourcade M., 2006, "The construction of a global profession: The transnationalization of economics", *American Journal of Sociology*, vol. 112, n° 1, p. 145-194.

Fourcade M., 2009, *Economists and Societies: Discipline and Profession in the United States, Britain, and France, 1890s to 1990s*, Princeton University Press.

Fretel A., Jany-Catrice F., Semenowicz P., Vatan S., 2020, Territoire zéro chômeur face à la crise sanitaire. L'expérimentation métropolitaine de Lille. Rapport Complémentaire, Paris DARES.

Guala F., 2005, *The Methodology of Experimental Economics*, Cambridge University Press.

Gurgand M., Wargon E., 2013, Synthèse des travaux du groupe présidé par Emmanuelle Wargon et Marc Gurgand, Paris DGEFP.

Harrison G.W., List J.A., 2004, "Field Experiments", *Journal of Economic Literature*, vol. 42, n° 4, p. 1009-1055.

Heukelom F., 2014, *Behavioral Economics: A History*, Cambridge University Press.

Jatteau A., 2020, *Faire preuve par le chiffre ? Le cas des expérimentations aléatoires en économie*, Paris Institut de la gestion publique et du développement économique.

Labrousse A., Zamora P., 2013, « Expérimentations de terrain et politiques publiques du travail et de l'emploi. Apports récents et mises en perspective », *Travail et emploi*, n° 135, p. 5-13.

L'Horty Y., Petit P., 2011, « Évaluation aléatoire et expérimentations sociales », *Revue française d'économie*, vol. 26, n° 1, p. 13-48.

Loison-Leruste M., Couronné J., Sarfati F., 2016, « La garantie jeunes en action », Rapport de recherche du CEET, n° 101, Paris CEET.

Monnier É., 1992, *Évaluations de l'action des pouvoirs publics*, Deuxième édition, Paris Economica.

Morel S., 1996, « L'évaluation des politiques sociales aux États-Unis : l'exemple des programmes d'intégration à l'emploi pour les allocataires de l'assistance sociale », *Politiques et management public*, vol. 14, n° 2, p. 33-58.

Nagatsu M., Favereau J., 2020, "Two Strands of Field Experiments in Economics: A Historical-Methodological Analysis", *Philosophy of the Social Sciences*, vol. 50, n° 1, p. 45-77.

Pélisse J., Wolmark C. (dir.), 2021, Le représentant de proximité : une figure au cœur des enjeux de la nouvelle représentation des salariés. Rapport final, Paris : IRERP, CSO, Institut du travail, Groupe Alpha.

Pénissat E., 2011, « Quantifier l'effet 'pur' de l'action publique : entre luttes scientifiques et redéfinition des politiques d'emploi en France », *Sociologie et Sociétés*, vol. 43, n° 2, p. 223-247.

Pénissat E., 2016, "Quantifying the Effects of Public Action on the Unemployed: Disputes between Experts and the Rethinking of Labour Market (1980-2000)", *The Social Sciences of Quantification. From Politics of Large Numbers to Target-Driven Policies*, dirigé par Bruno I., Jany-Catrice F., Touchelay B., Cham: Springer, p. 83-96.

Perez C., 2000, « L'évaluation expérimentale des programmes d'emploi et de formation aux États-Unis : éléments de critique interne », *Revue française des affaires sociales*, vol. 1, p. 145-163.

Perret B., 2008, *L'évaluation des politiques publiques*, Paris, La Découverte.

Porter T., 1995, *Trust in Numbers. The Pursuit of Objectivity in Public Life*, Princeton University Press.

Rao V., Woolcock M., 2003, "Integrating Qualitative and Quantitative Approaches in Program Evaluation", in *The Impact of Economic Policies on Poverty and income Distribution: Evaluation Techniques and Tools* dirigé par Bourguignon F. et Pereira da Silva L. A., World Bank et Oxford University Press, p. 165-190.

Revillard A., 2018, « Quelle place pour les méthodes qualitatives dans l'évaluation des politiques publiques ? », Sciences Po LIEPP Working Paper, n° 81, Paris, Sciences Po.

Svorenčík A., 2015, *The Experimental Turn in Economics: A History of Experimental Economics*, thèse de doctorat, University of Utrecht.

Thaler R., Sunstein C., 2008, *Nudge*, Yale University Press.

Woolcock M., 2019. "Why does Hirschmanian development remain mired on the margins? Because implementation (and reform) really is 'a long voyage of discovery' ", Center for International Development Working Papers, n° 347, Washington, D.C. World Bank.

Ce document est la version pre-print de l'article paru dans *Socio-économie du travail*, n°15, 2024-1. DOI : https://doi.org/10.48611/isbn.978-2-406-18609-0.p.0117.

ANNEXE

Tableau 1. Résumé du corpus : guides méthodologiques

Année	Titre	Auteurs	Institution
2009	Guide méthodologique pour l'évaluation des expérimentations sociales à l'intention des porteurs de projets	FEJ	FEJ
2012	Bilan méthodologique des évaluations conduites par le CRÉDOC dans le cadre du Fonds d'expérimentation pour la Jeunesse	Isa Aldeghi, Pauline Jauneau et Lara Muller	CREDOC
2016	Comment évaluer l'impact des politiques publiques ? Un guide à l'usage des décideurs et praticiens	Rozenn Desplatz et Marc Ferracci	France Stratégie
2022	Comment articuler les méthodes qualitatives et quantitatives pour évaluer l'impact des politiques publiques ? Un guide à l'usage des décideurs et des praticiens	Adam Baïz et Anne Revillard	France Stratégie

Ce document est la version pre-print de l'article paru dans *Socio-économie du travail*, n°15, 2024-1. DOI : https://doi.org/10.48611/isbn.978-2-406-18609-0.p.0117.

Tableau 2. Résumé du corpus : rapports et études

Année	Titre	Auteurs	Institution / politique évaluée
2008	Les expérimentations d'accompagnement renforcé des demandeurs d'emploi conduites par l'Unédic et l'ANPE. Note d'étape de l'évaluation.	Comité d'évaluation présidé par Claude Seibel Crest, EPP, Amnyos Consultants, Rom & Associés, DARES	UNEDIC / ANPE / Accompagnement renforcé des demandeurs d'emploi
2013	Synthèse des travaux du groupe présidé par Emmanuelle Wargon et Marc Gurgand.	Marc Gurgand et Wargon, Emmanuelle	DGEFP / Garantie jeunes
2014	Rapport d'évaluation du Revenu Contractualisé d'Autonomie	Crest (Romain Aeberhardt, Bruno Crépon, Mathilde Gaini)/EEP (Vera Chiodi, AugustinVicard), CREDOC (Léopold Gilles, Nelly Guisse)	DJEPVA / Revenu Contractualisé d'Autonomie
2018	Rapport final d'évaluation de la Garantie Jeunes	Comité scientifique en charge de l'évaluation de la Garantie jeunes présidé par Jérôme Gautié	Missions locales, DARES / Garantie jeunes
2019	Rapport d'évaluation sur l'expérimentation des emplois francs	Félix Bonnetête, Sandrine Firquet, Marie Rey, Benjamin Vignolles, Sarah Audras-Marcy, Mégane Aussedat, Pascal Dieusaert, Sabine Thibuad, Laurent Suster, Lucie Lestienne, Maéva Lamand et Niclas Jouve	Ministère du travail / Emplois francs
2019	Territoire Zéro Chômeur à Jouques.	Asdo études et l'ANSA pour la DARES	DARES / Territoire zéro chômeurs
2020	Premier rapport du comité scientifique du Plan d'investissement dans les compétences	Anne Bucher, Anne Fretel, Philippe Zamora, Christel	DARES / PIC

Ce document est la version pre-print de l'article paru dans *Socio-économie du travail*, n°15, 2024-1. DOI : https://doi.org/10.48611/isbn.978-2-406-18609-0.p.0117.

		Poujouly, Nicolas Vanni	
2020	Territoire zéro chômeur face à la crise sanitaire. L'expérimentation métropolitaine de Lille. Rapport Complémentaire.	Anne Fretel, Jany-Catrice Florence, Semenowicz Philippe et Sylvain Vatan	DARES / Territoire zéro chômeurs
2021	Le représentant de proximité : une figure au cœur des enjeux de la nouvelle représentation des salariés. Rapport final	Jérôme Pélisse et Cyril Wolmark	IRERP, CSO, Institut du travail, Groupe Alpha / Ordonnances travail
2021	Étude de l'impact du programme " Prépa Compétences » sur l'accès à la formation des demandeurs d'emploi dans le cadre du PIC	Élise Crovella, Nicolas Farvaque, Anne-Lise Ulmann, Victor Lextrait, Yann Marchat et Djamel Messaoudi	DARES / PIC
2022	Évaluation qualitative du dispositif Prépa-Projet mené par la région Bretagne dans le cadre du PIC	Élise Crovella, Nicolas Farvaque, Jeoffrey Magnier, Yann Marchat, Djamel Messaouid, Leïla Mokeddem, Léonard Bergoend	DARES / PIC
2023	Relations de travail et d'emploi et comportements des acteurs dans un contexte de réformes législatives	Raphaël Dalmasso, Camille Signoretto, Agnès Etiennot, Pascale Etiennot, Romain Marié, et al.	Institut François Gény, Ladyss / Ordonnances travail
2023	Discriminations ethno-raciales dans l'accès au logement social : un test des guichets d'enregistrement	Sylvain Chareyron et Yannick L'Horty	TEPP / Logement social