

L'ÉVALUATION DES POLITIQUES DE L'EMPLOI AU RÉGIME DE L'EXPERIMENTATION : LA PREUVE DE CONCEPT

Olivier BAGUELIN

Université Paris-Saclay, Univ Evry, CEPS, Tepp (FR 2042)

Laurent DUCLOS

Université Paris-Saclay, ENS Paris-Saclay, IDHES (UMR 8533)

Résumé

La présente discussion porte sur l'évaluation des politiques de l'emploi en France depuis 2000, et sur leur renouvellement dans le cadre d'un « tournant expérimental ». Ce tournant a rapproché les mondes académique et politico-administratif et mis fin à une ère de politiques de l'emploi soumises « au régime de la théorie », dont les résultats se sont avérés décevants. L'état des lieux amène à examiner les promesses d'innovation sociale du couple expérimentation-évaluation. Dans cette perspective, deux usages de l'expérimentation, la preuve de concept et le démonstrateur, doivent être distingués parce qu'ils impliquent des formes d'évaluation différentes. Dans le cadre d'un premier opus consacré à cette question, on se concentre ici sur la preuve de concept et l'évaluation des *expériences collectives* (que l'on oppose aux *expérimentations contrôlées*). Deux expériences collectives sont utilisées comme pierre de touche : la Garantie Jeunes (GJ) et Territoires Zéro Chômeur de Longue Durée (TZCLD). L'évaluation de la GJ a donné lieu à la synthèse d'une variété d'évaluations aux méthodes diverses, quali- et quantitatives : l'effort de mise en cohérence correspondant a permis de tirer parti des avantages comparatifs des différents exercices. L'évaluation de TZCLD est à ce jour plus brouillonne, laissant planer le risque d'une expérimentation « pour rien » *i.e.* dépourvue d'enseignements utiles. En matière de politiques de l'emploi, la fécondité de la démarche expérimentale n'est donc pas assurée. La discussion évoque des approches applicables à l'évaluation des expériences collectives. Deux paraissent spécialement adéquates : l'évaluation basée sur la théorie (*Theory-based evaluation*) et la collecte de preuves d'efficacité basée sur l'observation des pratiques (*Practice-based evidence*). Il s'agit dans les deux cas de ne pas méconnaître la contribution des parties prenantes à la qualité du processus d'évaluation. Cela plaide pour une évaluation des politiques publiques à la fois plus méthodique et moins technocratique.

Le présent opus s'inscrit dans un couple de numéros de *Socio-économie du travail* consacrés à l'expérimentation¹ dans le domaine des politiques de l'emploi et à la façon dont l'approche expérimentale renouvelle l'exercice d'évaluation de ces politiques. L'ambition de ces deux numéros est de dresser un état des lieux de la réflexion socio-économique en la matière. Cette réflexion s'est en effet considérablement développée depuis les années 2000 à mesure que l'application de l'approche expérimentale aux politiques publiques gagnait des soutiens. La réflexion a notamment concerné les méthodes et approches² évaluatives (Stern *et*

¹ Que l'on ne réduit pas ici à l'expérimentation avec assignation aléatoire *contrôlée*.

² En anglais, il est souvent question d'*evaluation design*, expression qui n'a pas trouvé de traduction stabilisée en France : nous suivons l'usage québécois en parlant d'approche évaluative. Une approche évaluative correspond aux choix opérés concernant

al., 2012, p. 15) applicables à un matériau empirique produit dans le cadre d'une expérimentation.

Accompagnant une faveur transnationale pour l'*evidence-based policy*, l'approche expérimentale propose d'adopter telle ou telle intervention ou disposition non sur la foi d'une doctrine politique ou d'une théorie socio-économique mais sur la base de « données probantes ». En lieu et place d'aprioris théoriques ou idéologiques³, une démarche systématique d'évaluation adossée à un instrument essentiel de la méthode scientifique : l'expérimentation. Cette promesse (dont il s'agira d'évaluer la portée) a reçu un écho particulièrement enthousiaste, en France, dans le domaine des politiques de l'emploi (Labrousse et Zamora, 2013).

Sur ces prémices, la présente introduction a trois objets principaux.

Son premier objet est d'indiquer en quoi l'expérimentation pourrait renouveler favorablement les politiques de l'emploi en France. Cela tient largement au rapprochement qu'elle a précipité entre les mondes académique et politico-administratif : l'évaluation des expérimentations fait resurgir des enjeux qui intéressent les deux mondes, notamment en termes de méthode. Cela tient aussi aux déceptions suscitées par les politiques de l'emploi menées en France. Le propos s'appuie sur un encadré consacré à l'analyse des dépenses « en faveur de l'emploi et du marché du travail »⁴ depuis vingt ans. On y interprète l'action menée comme relevant de politiques de l'emploi soumises « au régime de la théorie »⁵ vis-à-vis desquelles l'évaluation apparaît comme un enjeu subalterne, sinon purement formel.

Le deuxième objet de cette introduction est de distinguer deux usages de l'expérimentation relevant, dans une perspective d'innovation sociale (Longtin, 2021), du *démonstrateur* pour l'un et de la *preuve de concept* pour l'autre⁶, et sollicitant chacun des formes adaptées d'évaluation. Ces deux usages permettent de définir le champ de chaque opus : le présent porte sur la conception expérientielle des politiques de l'emploi (la preuve de concept) ; le second, sur leur évaluation expérimentale (la démonstration d'impact). Une façon, on va le voir, de montrer que l'évaluation des politiques publiques n'est pas condamnée à être technocratique.

Enfin, le troisième objet de cette introduction est de présenter les contributions rassemblées dans ce premier opus et d'en donner les enseignements principaux, notamment quant aux conditions de bonne méthode sous lesquelles le tournant expérimental pourrait contribuer à l'amélioration des politiques de l'emploi en France. Une conclusion essentielle tirée de ces contributions est en effet que tirer les fruits d'une mise des politiques de l'emploi « au régime de l'expérimentation » ne va pas de soi : cela suppose la diffusion d'exigences méthodologiques

« les questions [de l'évaluation], la collecte des données, les méthodes d'analyse et l'utilisation des résultats » (Longtin, 2021, p. 45). Sur la variété des approches évaluatives, lire Delahais (2022, p. 155).

³ « Face [aux] oppositions, l'usage des expérimentations permettait d'apaiser la controverse [par le recours à] un protocole présenté [alors] comme non idéologique, empirique plutôt que politique ou théorique » (Arrignon, 2019, p. 1129).

⁴ Des guillemets pour exprimer l'étonnement que pourrait justifier cette terminologie officielle. On aurait pu s'attendre, au nom des conventions internationales ratifiées par la France (ex. convention n°122 sur la politique de l'emploi, 1964), à des dépenses visant « la plénitude de l'emploi et l'élévation des niveaux de vie », « la lutte contre le chômage et la garantie d'un salaire assurant des conditions d'existence convenables », voire à des dépenses destinées à rendre effectif l'article du préambule de la Déclaration universelle des droits de l'homme selon lequel « toute personne a droit au travail, au libre choix de son travail, à des conditions équitables et satisfaisantes de travail et à la protection contre le chômage ». À la place, un projet de politique publique entérinant l'idée que des dépenses en faveur du fonctionnement marché du travail puissent contribuer à l'*emploi décent*. Or cette référence posée par l'Organisation internationale du travail renvoie à un programme exactement opposé : celui de la dé-marchandisation (ou *decommodification*) du travail (Esping-Andersen, 2008).

⁵ Formule que l'on s'efforce d'expliquer ci-après en lui opposant un « régime de l'expérimentation ».

⁶ Les notions sont empruntées à l'analyse des processus d'innovation technologique en sciences de gestion où la preuve de concept correspond à une première étape du cycle de validation de l'innovation, là où le démonstrateur figure l'une de ses dernières étapes. On parle donc, en distinguant ces deux étapes, de degré de maturation très différents.

qui concernent l'exercice d'évaluation en tant que tel, mais aussi l'animation politico-administrative des pratiques.

Les politiques de l'emploi au régime de la théorie

Mesures ciblées ou générales, dépenses passives ou actives, incitation ou subvention : au pays des politiques de l'emploi, règne le nominalisme. De 2000 à 2020, la part du PIB consacrée en France aux dépenses « en faveur de l'emploi et du marché du travail » (hors mesures liées à la crise sanitaire) est passé de 3,9 % à 7,3 %⁷. Moyennant cet effort (3,4 points de PIB, donc), le taux de chômage des 15-64 ans a reculé de 0,5 point de pourcentage et le taux d'emploi progressé de 2,5 points. Cela place le cours du point de taux d'emploi à 1,4 point de PIB, quelque chose de l'ordre de 35 Md€ de 2020. Ce rendement modeste – qui ne tient d'ailleurs pas compte du recul de la qualité moyenne de l'emploi (Picart, 2014) – n'est peut-être pas sans rapport avec un programme de réforme hérité de la *Stratégie européenne pour l'emploi* de 1997 qui se référait à un marché du travail naïvement doté de vertus autonomes (Duclos, 2009). En effet, l'évolution de la répartition de ces dépenses entre les différentes interventions du domaine travail-emploi révèle leur resserrement sur une action, « favoriser le marché du travail », aux effets décevants (L'Horty, 2013 ; Askenazy et Palier, 2018).

Les statistiques de la Dares⁸ permettent de se faire une idée de ce resserrement à partir de la nomenclature Eurostat distinguant entre dépenses dites *ciblées* ou *générales*. Les premières sont réputées s'adresser à des « publics » : personnes inscrites à l'agence pour l'emploi ou « en difficulté sur le marché du travail » (inactifs BIT désirant occuper un emploi, salariés dont l'emploi est menacé) ; le ciblage peut impliquer un critère d'âge (jeunes, seniors) ou une situation déterminée (parent isolé, chômage de longue durée). Les secondes (« dépenses générales ») se définissent par une absence de ciblage⁹.

La FIG. 1 décrit la hausse de ces dépenses, déclinées sur six quadriennats et en neuf catégories principales. Ces catégories de dépenses sont rangées par montants décroissants pour 2017-2020 : (1) soutiens au revenu en l'absence/insuffisance d'emploi (hors crise sanitaire), (2) mesures générales d'exonération hors allègements bas salaires (heures supplémentaires défiscalisées, dispositif « zéro charge » TPE, CICE...), (3) allègements généraux bas salaires et aménagement-réduction du temps de travail, (4) mesures « actives » (détaillées FIG. 2), (5) mesures sectorielles (hôtels-cafés-restaurants, services à la personne, agriculture), (6) incitations financières à l'emploi (prime pour l'emploi, prime d'activité...), (7) services relatifs au marché du travail, (8) mesures géographiques, (9) aides à l'embauche TPE-PME. À ces catégories, la crise sanitaire a ajouté deux dépenses propres au quadriennat 2017-2020 : des mesures d'urgence (exonérations exceptionnelles) et la mobilisation très forte de l'activité partielle.

La FIG. 1 illustre trois aspects des politiques de l'emploi menées en France sur la période : le rôle central des soutiens au revenu (mesure dite « passive ») ; la colonisation du domaine des politiques de l'emploi par les mesures de subvention aux employeurs ; le recul des dépenses consacrées aux mesures dites « ciblées » et « actives ». Chacun de ces aspects appelle une mise en perspective.

Le soutien au revenu en l'absence/insuffisance d'emploi joue un rôle central tout au long de la période. Il s'agit principalement ici de l'indemnisation chômage, levier d'action contra-cyclique mais surtout, notamment depuis 2009, amortisseur essentiel de la fragmentation de l'emploi (Baguelin, 2023). Si la désignation reste la même tout au long de la période, la fonction évolue. Le classement de l'indemnisation chômage parmi les mesures dites « ciblées » et « passives » apparaît de plus en plus inadéquat. Les mesures « ciblées seniors » (préretraites et dispenses de recherche d'emploi)

⁷ À titre de référence, la dépense de l'État en matière d'éducation représentait 4,7 % du PIB en 2000 et seulement 4,1 % en 2020 (Insee, France – Portrait social, édition 2021).

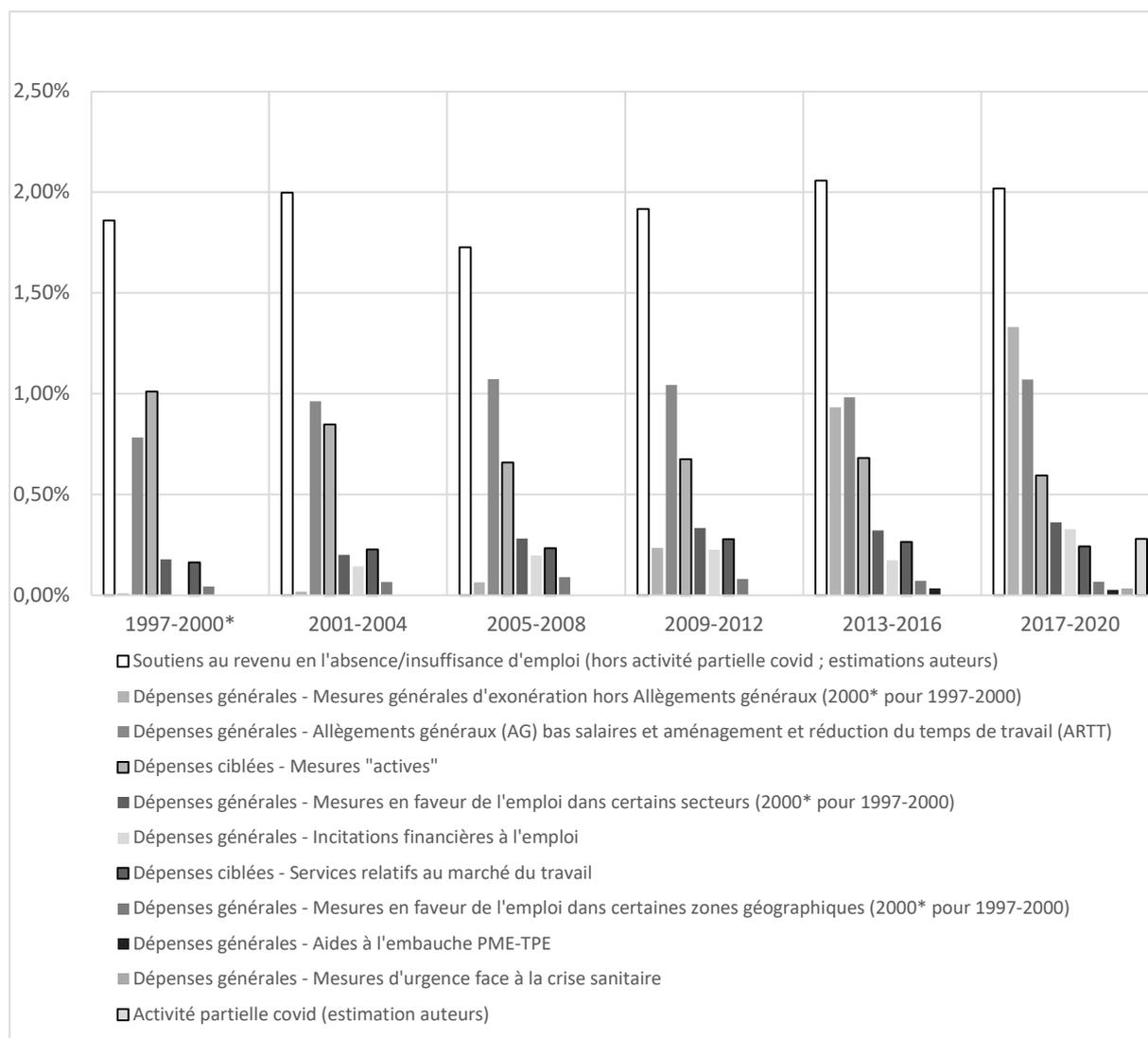
⁸ Statistiques téléchargées depuis la page consacrée aux dépenses en faveur de l'emploi et du marché du travail – <https://dares.travail-emploi.gouv.fr/publication/depenses-en-faveur-de-lemploi-et-du-marche-du-travail-en-2020> (consulté en février 2024).

⁹ L'analyse des dispositifs correspondant aux dépenses dites ciblées ou générales, amène cependant à relativiser la portée de la distinction : l'indemnisation de l'activité partielle, par exemple, est tenue pour une dépense ciblée bien qu'elle ait concerné 8 millions de salariés en avril 2020 (Baron et Guillermin, 2023).

disparaissent et le passage par une inscription à l'agence publique pour l'emploi se banalise. L'épithète « passif » est aussi trompeur, l'indemnisation chômage (et la menace d'en être privé) jouant de plus en plus activement comme aiguillon à accepter les emplois disponibles (Vivès *et al.*, 2023).

En revanche, en un sens interprétatif plutôt que conventionnel, parler d'une dépense « en faveur du marché du travail » devient adapté, l'indemnisation chômage tendant à sortir du registre de la protection sociale (Grégoire *et al.*, 2018). En lieu et place, le couplage indemnisation-contrôle du chômage, dont une manifestation est la consolidation, en particulier à partir de 2009, des dépenses de « services relatifs au marché du travail »¹⁰ (Cf. FIG. 1), au demeurant de plus en plus largement financé par les cotisations chômage (Coquet, 2016a). L'indemnisation chômage vient en soutien de l'activité de l'agence pour l'emploi dans la production d'une offre de travail répondant aux « exigences du marché » (Clouet, 2018).

FIG. 1 – Les dépenses en faveur de l'emploi et du marché du travail (en points de PIB)



Source : Dares (consultation février 2024).

La FIG. 1 illustre en second lieu la place et les formes prises par les mesures générales d'exonération de cotisations sociales. On y distingue les allègements généraux (AG) bas salaires (et aménagements RTT) des autres mesures ajoutées à partir de 2008. La distinction permet d'envisager le bien-fondé des allègements hérités des années 1990. Deux arguments peuvent être avancés en ce sens. D'abord, celui

¹⁰ Services d'information ou d'accompagnement.

de Philippe Askenazy plaçant un rendement marginal décroissant des AG en termes de volume d'emploi (d'autant plus faible que l'on part d'un salaire plus bas), manière de concéder à ses promoteurs que l'intervention ait pu avoir un bon rendement initial (Askenazy, 2011, p. 198). Ensuite, en admettant que le passage aux 35 h sans perte de salaire hebdomadaire justifiait qu'on l'accompagne d'allègements¹¹.

Il paraît plus difficile de tenir pour « favorables à l'emploi » les exonérations ajoutées par la suite, spécialement celles de 2008 qui concernaient les heures supplémentaires et les heures de RTT rachetées. Cela n'a pas empêché une dilatation très forte de ces exonérations hors AG à partir de 2013 : pour le quadriennat 2017-2020, l'ensemble des mesures d'exonération (AG + hors AG) représentent 2,4 points de PIB, quatre fois ce qui prévalait vingt ans plus tôt. Selon une interprétation acritique, on peut certes accepter qu'il s'agit d'intervenir « en faveur du marché du travail » cette fois en stimulant une demande de travail réputée fonction décroissante du prix du travail. Mais classer les mesures générales d'exonération de cotisations sociales parmi les politiques de l'emploi est aussi une façon de subventionner massivement les entreprises (Gautié et Leray, 2024, p. 78).

La plupart des autres dépenses considérées relèvent d'interventions auxiliaires empruntant leur logique de production d'offre ou de demande de travail aux précédentes. Les « mesures en faveur de l'emploi dans certains secteurs » (Cf. FIG. 1) concernent principalement des réductions d'impôt et des exonérations en faveur de particuliers employeurs ou, plus marginalement (essentiellement sur le quadriennat 2005-2008), des aides à l'emploi réservées aux hôtels, cafés et restaurants. Il s'agit de soutenir le développement d'un emploi *présentiel* (selon les termes de l'économie géographique) par nature non délocalisable et permettant (selon ses promoteurs) de procurer un revenu du travail à une main-d'œuvre peu ou pas diplômée (Henneguelle *et al.*, 2023) – d'aucuns parlent plus brutalement d'un retour de la domesticité (Carbonnier et Morel, 2018). Du début à la fin de la période d'intérêt, les dépenses de ce type doublent, passant de 0,18 à 0,36 points de PIB (Cf. FIG. 1).

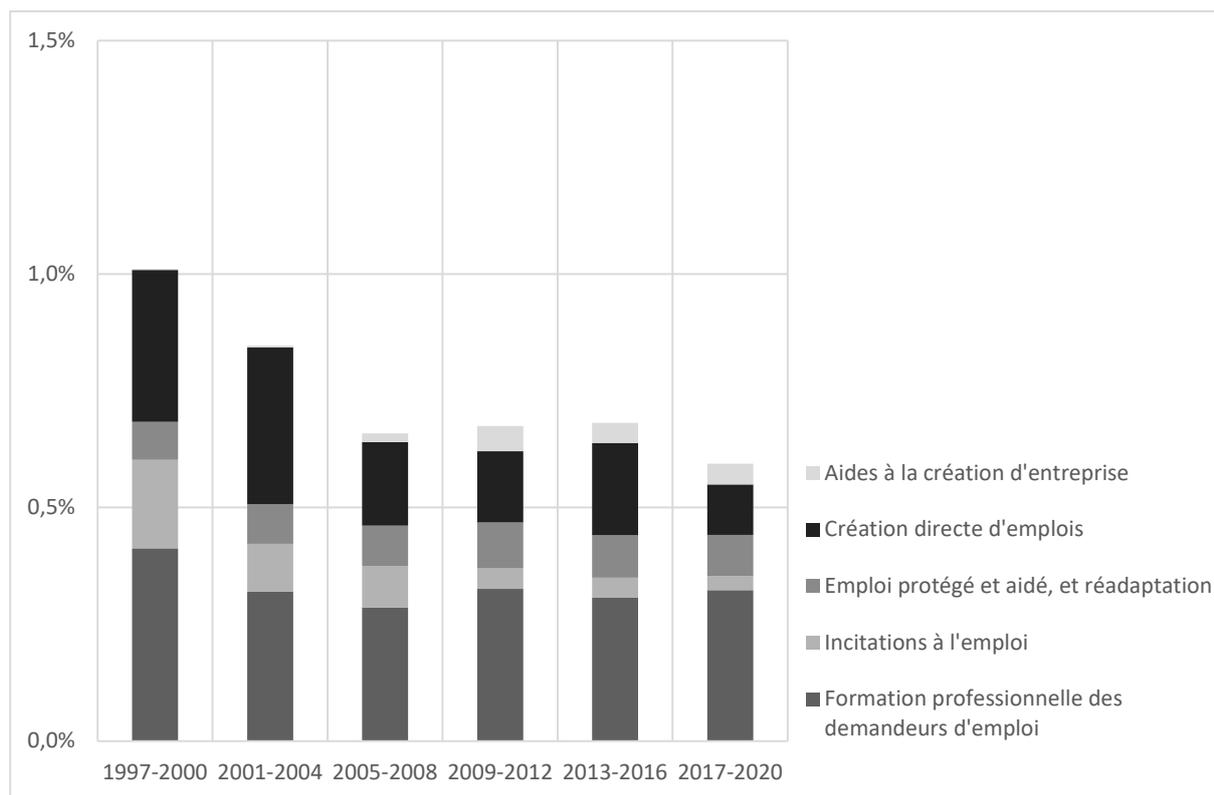
Ces interventions sectorielles contribuent à élargir un segment secondaire d'emploi (Picart, 2017) associé à d'importants mouvements de main-d'œuvre depuis et vers le chômage. La qualité de l'emploi en question (conditions d'emploi, quotité de temps de travail, accès à la formation, etc.) est en effet discutable (Palier, 2023) et son développement nécessite des incitations non salariales à l'emploi : prime pour l'emploi, « RSA activité » puis prime d'activité (Allègre, 2024). Les dépenses correspondantes passent de zéro en 2000, à 0,23 point de PIB pour le quadriennat 2009-2012 ; elles sont ramenée à 0,17 point sur 2013-2016 avant d'être portées à 0,33 point de PIB sur 2017-2020 (Cf. FIG. 1).

La dernière intention de la FIG. 1 est de rendre compte du recul des dépenses ciblées-public consacrées à des mesures « actives » ; de 1,01 point de PIB sur 1997-2000 à 0,59 sur 2017-2020. Sans être une tendance nouvelle (elle est déjà à l'œuvre dans les années 1990), le recul des mesures « actives » détaillé dans la FIG. 2, mérite une analyse spécifique dans le cadre d'une réflexion sur l'évaluation des politiques de l'emploi.

La nomenclature Eurostat range parmi ces mesures la formation professionnelle des demandeurs d'emploi, la création directe d'emplois (contrats aidés du secteur non-marchand), les incitations à l'emploi, l'emploi protégé ou de réadaptation et l'aide à la création d'entreprise. Les deux dernières catégories sont les seules à connaître une certaine vigueur : l'aide à la création d'entreprise devient significative à partir du quadriennat 2009-2012 et le soutien à l'emploi des personnes handicapées conservent sa place sur la période (environ 0,9 point de PIB). La division par six des dépenses d'incitations à l'emploi (de 0,19 à 0,03 point de PIB) reflète quant à elle pour l'essentiel leur « cannibalisation » (Askenazy, 2011, p. 199) par les mesures générales d'exonérations. Mais pour les autres dépenses, dont la justification emprunte moins à l'imaginaire du marché du travail, le recul n'est-il pas révélateur d'une panne d'inspiration ou d'audace ?

¹¹ Notamment au regard de l'augmentation du smic horaire qu'induisait la convergence des garanties minimales de rémunération.

FIG. 2 – Dépenses ciblées en mesures « actives » (en point de PIB)



Source : Dares (consultation février 2024).

La formule du stage de formation professionnelle semble tomber en disgrâce entre 2000 et 2008, même si les dépenses afférentes se stabilisent au cours des années suivantes (probablement s'est-on avisé que le placement en stage retirait à peu de frais les intéressés des chiffres du chômage)¹². La même analyse s'applique de façon encore plus franche à la création directe d'emplois dans le secteur non marchand. Après l'abandon des emplois jeunes, à partir de 2002, les contrats aidés ne constituent plus qu'un outil contra-cyclique – aux propriétés discutables (Baguelin, 2012) – qu'il convient d'écarter dès lors que la conjoncture s'affermi : sur le quadriennat 2017-2020, la dépense correspondante ne représente plus que 0,11 point de PIB contre 0,33 sur 1997-2000. L'impression est qu'il s'agit de dispositifs dont on n'attend pas grand-chose ; les politiques de l'emploi semblent revenues au « on a tout essayé » du début des années 1990.

L'ÉVALUATION : LIEU D'UNE RENCONTRE TARDIVE ENTRE LES MONDES ACADEMIQUE ET POLITICO-ADMINISTRATIF

L'institutionnalisation de l'évaluation des politiques publiques (EPP) en France¹³ est généralement associée aux suites du rapport Viveret sur l'évaluation du Revenu minimum d'insertion (RMI) – Viveret (1989). Ce rapport, élaboré au sein du Commissariat général du Plan (CGP), pose en effet les fondements d'une approche française de l'évaluation dont les ambitions sont d'abord démocratiques. Entre autres garanties – d'indépendance et de rigueur méthodologique – il inscrit l'exigence d'un certain « pluralisme » au cœur de la discipline évaluative. Cette garantie désigne alors un principe d'action publique partagée¹⁴ plus qu'un

¹² Les estimations de l'effet de la formation des demandeurs d'emploi sont par ailleurs décevantes (Chaubaud *et alii*, 2022).

¹³ On met ici de côté l'expérience française de *rationalisation des choix budgétaires* (Huet *et al.*, 1973).

¹⁴ Selon la tradition « d'expertise concertée » chère au CGP.

attribut des méthodes d'évaluation. Sur cette question, le rapport Viveret se démarque¹⁵ cependant des travaux de défrichage réalisés au CGP quelques années plus tôt (Deleau, 1986)¹⁶. Il débouche en 1990 sur la création d'un Comité interministériel¹⁷ de l'évaluation¹⁸ s'adjoignant un Conseil scientifique de l'évaluation auquel succède, huit ans plus tard, un Conseil national de l'évaluation (CNE)¹⁹. Ces premiers éléments de structuration de l'EPP en France tombent pourtant rapidement en désuétude (Gibert et Saussois, 2022).

L'ÉVALUATION DANS LA FRANCE DES ANNÉES 1990 : UNE INSTITUTIONNALISATION À LA PEINE

Un rapport d'information²⁰ produit au Sénat en 2004 témoigne de l'état d'esprit politico-administratif concernant l'EPP dans le courant des années 1990. Le rapport dresse un bilan « en demi-teinte » de la place de l'EPP dans le fonctionnement de l'État déplorant une certaine « hypocrisie organisationnelle » (Gibert et Saussois, 2022) et réclamant que l'EPP soit *enfin* placée « au cœur de la réforme de l'État »²¹. Surtout, il appelle à distinguer l'EPP d'une démarche de recherche²². L'idée est d'affirmer la spécificité de l'exercice évaluatif, dont les critères et objectifs doivent être fixés par le commanditaire (l'État) et non par des équipes de recherche chargée d'évaluation.

C'est ainsi qu'en dépit d'une adhésion progressive des corps de contrôle (Inspections générales, Cour des comptes) au principe d'évaluation, le centre de gravité de l'EPP se constitue jusqu'au milieu des années 2000 au sein de la sphère gouvernementale. En 1993, le décret²³ créant la Direction de l'animation de la recherche, des études et des statistiques (Dares) fait le choix d'instaurer une « évaluation interne » dans le champ du travail et de la formation professionnelle (Adjerad *et al.*, 2002). En 1998, une circulaire²⁴ accompagnant la création du CNE dispose qu'en règle générale « on privilégiera les évaluations visant à répondre à des questions claires, permettant d'aboutir dans des délais brefs à des réponses précises susceptibles de conduire à des décisions opérationnelles [...] Les projets dont les objectifs relèvent davantage de la recherche en sciences économiques et sociales que de la gestion des affaires publiques seront orientés vers d'autres dispositifs ». Cette posture reflète un aspect tôt repéré par les observateurs de l'institutionnalisation de l'EPP en France (Nioche, 1993 ; Spenlehauer, 2004) où administration et recherche académique se conçoivent comme antagonistes²⁵.

¹⁵ Certains parlent même de rupture (Simha, 2015).

¹⁶ Travaux quant à eux explicitement dédiés aux enjeux de méthode mais très marqués par un principe étroit de rationalisation des choix budgétaires (Spenlehauer, 1998).

¹⁷ Censé préfigurer une sorte de *Government Accountability Office* à la française.

¹⁸ Décrets du 22 janvier 1990 relatif à l'évaluation des politiques publiques.

¹⁹ Décrets du 18 novembre 1998.

²⁰ Rapport d'information n° 392 (2003-2004) déposé le 30 juin 2004, fait au nom de la délégation pour la planification du Sénat.

²¹ Le rapport est alors très préoccupé par des enjeux de définition et de renforcement des structures d'évaluation situées au cœur comme à la périphérie de l'État. L'attention porte aussi sur l'articulation de l'impératif d'évaluation avec le dispositif de la *loi organique sur les lois de finances* (Lolf) du 1^{er} août 2001 qui traduisait une volonté forte de l'État d'inscrire l'action publique dans une perspective de performances et de mieux asseoir le contrôle et l'EPP exercés par le Parlement (Arkwright *et al.*, 2007).

²² Reprenant les développements du *Petit guide de l'évaluation des politiques publiques* édité en 1996 par le Conseil scientifique de l'évaluation (CSE), le Sénat considère que « l'évaluation des politiques publiques [n'était] pas assimilable à un travail de recherche ». Alors que l'autonomie de la recherche, implique « le libre choix des questions et des grilles d'interprétation » (CSE, 1996), le chercheur pouvant « légitimement entreprendre de reformuler un questionnement initial » (Sénat, 2004), l'association à des travaux d'évaluation présupposerait « de considérer comme pertinentes – quitte à les discuter ou les reformuler – les questions et valeurs de référence d'un commanditaire politique » (CSE, 1996).

²³ Décret 93-57 du 15 janvier 1993.

²⁴ Circulaire du 28 décembre 1998 relative à l'évaluation des politiques publiques.

²⁵ Ainsi que le relève Jean-Pierre Nioche pour expliquer le lent développement de l'évaluation en France et un certain hiatus avec la « recherche évaluative » : « l'univers de la recherche et l'univers politico-administratif coopèrent mal. Ceci est dû

Ces dispositions ont probablement fait obstacle aux promesses du rapport Viveret. D'abord, elles postulent la capacité du commanditaire à formaliser ses critères d'évaluation à travers l'instrument que constitue le cahier des charges. La pratique suggère au contraire que les critères avancés souffrent souvent d'une indétermination qu'un dialogue avec une équipe de recherche peut permettre de lever²⁶. Ensuite et surtout, la rémanence de l'antagonisme entre administration et recherche a détourné la première des enjeux méthodologiques que comportait l'EPP (comment écrire un cahier des charges sans expertise méthodologique ?) et la seconde d'une tâche à laquelle elle pouvait contribuer à son propre bénéfice. En concevant l'EPP sous l'angle étroit de l'organisation de l'État, reléguant au monde de la recherche les détails fastidieux des différentes techniques de collecte et de traitement de l'information utiles à l'évaluation, la sphère politico-administrative aura finalement freiné son institutionnalisation. La recherche en sciences sociales rechignera à s'occuper d'EPP, parfois tenue pour une pratique servile²⁷ ou méprisée pour ses faiblesses méthodologiques²⁸.

L'AFFIRMATION D'UN LIEN EXPERIMENTATION-EVALUATION

Pour autant, la loi relative au RMI²⁹ ouvrait d'autres perspectives propres à alimenter le développement de l'EPP. Son article 52 instaurait en effet un mécanisme de révision faisant explicitement référence au caractère expérimental de la loi. Il disposait ainsi, qu'au vu des conclusions du rapport d'évaluation qui lui serait adressé trois ans après le vote de la loi, « le Gouvernement déposer[ait] un projet de loi visant à procéder aux adaptations qui lui apparaîtraient nécessaires ». L'option de la loi expérimentale n'est pas inédite (Chevallier, 1996) mais c'est la première fois que l'évaluation se voit accorder un rôle aussi considérable, ce dont témoigne l'importance des moyens mis à disposition³⁰. Ce cadre concerne cependant la seule expérimentation du RMI et reste « déconnecté du dispositif général d'évaluation institué en janvier 1990, ce qui tend (par effet retour) à reléguer celui-ci à une fonction générale d'étude, dégagée de toute emprise sur les processus législatifs concrets » (Chevallier, 1996).

Si le lien expérimentation-évaluation n'en est pas moins posé, il faut attendre la seconde moitié des années 2000 pour qu'une association s'opère véritablement entre expérimentation *de la norme* (au sens de loi d'expérimentation) et expérimentation sociale entendue comme méthode d'évaluation. Le recours à la loi d'expérimentation est porté dès 2003 par une réforme constitutionnelle et le procédé de l'expérimentation sociale importé du monde de la recherche à partir de 2007. L'initiative de cette importation émane du Haut-commissariat aux solidarités actives dans les suites, cette fois, de l'expérimentation du Revenu de solidarité active (RSA). Elle s'enracine dans le paysage politico-administratif (Simha, 2015) par la création du Fonds

principalement aux filières de formation des élites administratives, qui les isolent du monde de la recherche, et au réflexe d'autonomie intellectuelle de la haute administration française » (Nioche, 1993). Analyse partagée quelques années plus tard par Vincent Spenlehauer : « en France, la haute fonction publique tourne le dos au monde académique et à ses méthodologies. Contrairement à ce qui se passe aux États-Unis où les élites administratives sont formées à l'Université et ont une vraie familiarité avec les méthodes de recherche en science sociale » (audition au Conseil économique, social et environnemental, 28 décembre 2013).

²⁶ Cf. le rôle de la phase d'émergence distinguée par Longtin (2021).

²⁷ Les sciences sociales et plus particulièrement la sociologie allant, selon Spenlehauer, jusqu'à produire : « d'imposantes jurisprudences méthodologiques pour juguler sereinement les grandes capacités d'*impairment* attribuables à (...) l'État » (Spenlehauer, 2004).

²⁸ Du côté des économistes, on dénonce : « la faiblesse des évaluations réalisées, du traitement des impacts que ce soit au regard des objectifs affichés par les politiques évaluées ou en termes d'effets non recherchés de celles-ci » (Gibert *et al.*, 2022).

²⁹ Loi n°88-1088 du 1er décembre 1988 relative au Revenu minimum d'insertion.

³⁰ Le dispositif institutionnel comporte une « délégation interministérielle » (décret du 7 décembre 1988), assistée d'un « groupe interministériel de coordination » et d'une « commission nationale d'évaluation relative au RMI » (décret du 6 septembre 1989) elle-même assistée d'un « groupe de coordination scientifique », dotée de crédits spécifiques, bénéficiant d'agents publics mis à disposition et autant que de besoin des services du CGP.

d'expérimentation pour la jeunesse (2008) et l'institution d'un Haut-commissariat à la jeunesse (2009).

La réforme constitutionnelle de 2003 (articles 37-1 et 72-4)³¹ banalise progressivement le recours à l'expérimentation (Cf. l'article de **Vincent Cibois** dans ce numéro). Dans le champ de l'emploi, à partir de l'expérimentation du Contrat de transition professionnelle³² de 2006, le mouvement s'accélère : Garantie jeunes (GJ)³³ en 2013, Territoires zéro chômeur de longue durée (TZCLD)³⁴ à partir de 2016, Emplois francs³⁵ depuis 2018. La loi du 5 septembre 2018 pour la liberté de choisir son avenir professionnel (dite loi « Avenir pro ») comme l'énoncé de la Stratégie nationale de prévention et de lutte contre la pauvreté, en septembre 2019, multiplie les expérimentations.

De par son caractère expérimental, la loi ne bénéficie plus d'un bien fondé de principe ni d'un postulat de pertinence et d'efficacité. Dans le même temps, elle garde un empire en fixant les modalités de sa propre évaluation. Si l'épisode RSA a peut-être le tort de réduire l'évaluation méthodique à l'évaluation contrefactuelle, il a le grand mérite de poser les bases d'une logique d'apprentissage en matière d'action publique. L'expérimentation normative va désormais de pair avec l'établissement d'un cadre spatio-temporel spécifiquement réservé à l'évaluation, une « parenthèse réflexive ». C'est particulièrement clair avec la loi « Avenir pro » et l'expérimentation des *entreprises adaptées*³⁶. En soumettant la liberté d'expérimenter à l'obligation d'évaluer, le droit rejoint sur le plan du « jugement de valeur » la norme scientifique et pragmatique où il ne saurait y avoir expérimentation – ni même expérience (Dewey, 1993) – sans « jugement de fait ». Et l'on voit s'affirmer la promesse expérimentale qui consiste à atteler fermement jugements de valeur et de fait.

UN REGAIN D'INTERET DANS LE MONDE DE LA RECHERCHE

L'affirmation du lien expérimentation-évaluation décrite ci-dessus doit beaucoup à la diffusion mondiale d'un modèle de l'expérimentation sociale entendue comme *expérimentation contrôlée* (Perez, 2000), par référence au principe d'essai avec assignation aléatoire contrôlée – *Randomized controlled trials* (RCT) en anglais. Dans le champ de l'emploi, le premier exercice de ce type porte, dès 2007, sur l'accompagnement des demandeurs d'emploi par les opérateurs privés de placement (Behaghel *et al.*, 2009). Mais le regain d'intérêt du monde de la recherche pour l'EPP prendra plus largement la forme de controverses fécondes, notamment à propos de l'expérimentation du RSA (Gomel et Serverin, 2009).

³¹ La possibilité d'expérimentation est inscrite dans la Constitution aux articles 37 (« La loi et le règlement peuvent comporter, pour un objet et une durée limités, des dispositions à caractère expérimental ») et 72 alinéa 4 (« les collectivités territoriales [peuvent] déroger à titre expérimentale et pour un objet et une durée limités aux dispositions législatives ou réglementaires qui régissent l'exercice de leurs compétences »).

³² Ordonnance du 13 avril 2006.

³³ Décret du 1^{er} octobre 2013.

³⁴ Loi du 29 février 2016.

³⁵ Loi de finance initiale pour 2018.

³⁶ Loi n°2018-771, article 79 : « Un cahier des charges national fixe les critères que doivent respecter les entreprises adaptées candidates à l'expérimentation, notamment les objectifs, les moyens et les résultats attendus en termes de sorties vers l'emploi. Sur proposition du comité de suivi de l'expérimentation, le ministre chargé de l'emploi dresse la liste des structures retenues pour mener l'expérimentation. Un décret précise les modalités de mise en œuvre de cette expérimentation, notamment le montant de l'aide financière susceptible d'être accordée, le contenu de l'avenant au contrat conclu avec l'État ainsi que les conditions de son évaluation en vue de son éventuelle généralisation. Au plus tard douze mois avant le terme de l'expérimentation, est réalisée une évaluation afin de déterminer les conditions appropriées pour son éventuelle généralisation. Au terme de l'expérimentation, le Gouvernement présente au Parlement un rapport sur l'application de la présente disposition au regard de son impact sur l'accès à l'emploi des travailleurs reconnus handicapés, sur les formations suivies ainsi que les conséquences sur les finances publiques ».

Ces premières expérimentations sont le point de départ d'une série de travaux qui marquent, de la part de la recherche, un « *désir de réinvestir* » (Gibert et Saussois, 2022) le champ de l'EPP d'intention scientifique. Ce regain d'intérêt se manifeste dans la plupart des disciplines des SHS : en économétrie où se joue la « révolution de la crédibilité » (Angrist et Pischke, 2010) ; en socio-économie, dans une veine plus critique (Jatteau, 2013) ; en science politique sur le mode traditionnel de l'analyse des politiques publiques (Lacouette-Fougère et Lascoumes, 2013)... mais aussi en droit où l'on questionne la portée juridique de l'expérimentation normative (Marchand-Arvier et Deumier, 2014). Chaque discipline tente d'établir des prétentions évaluatives spécifiques : l'évaluation dite d'impact en économie, l'évaluation des processus parmi les politistes, l'évaluation réflexive chez les sociologues, l'évaluation de la norme du côté des juristes... Dans cette effervescence, un lobbying institutionnel efficace donne toutefois la prééminence à l'expérimentation contrôlée qui va rapidement tenir lieu de méthode de référence dans le monde politico-administratif³⁷.

C'est à ce stade, et dans ce nouveau contexte de l'expérimentation, qu'ont été pensés les deux numéros consacrés aujourd'hui par *Socio-économie du travail* à l'évaluation des politiques de l'emploi. Il y a quarante ans, évoquant l'essai de Donald Thomas Campbell intitulé *Methods for the experimenting society* (1971 ; 1991), Jean-Pierre Nioche considérait que « les projets de société expérimentale souhaités par certains prophètes de l'évaluation où l'expérimentation remplacerait les choix de valeurs [lui paraissaient] de dangereuses utopies » (Nioche, 1982, p. 57-58). Cette perspective est de moins en moins utopique et la responsabilité du monde de la recherche particulièrement engagée : le clivage wébérien du savant et du politique a vieilli (Spenehauer in Ould-Ferhat, 2018) et la « relation distanciée » entre sociologie et EPP mise en cause (Delahais et Devaux-Spatarakis, 2018). Il s'agit de prendre au sérieux le risque que l'*evidence-based policy* ne dégénère en une science réglementaire (Borraz et Demortain, 2015) applicable aux politiques de l'emploi et dont tout jugement de valeur serait remisé.

LES ENJEUX SOCIO-ECONOMIQUES DE L'ATTELAGE EXPERIMENTATION-EVALUATION

Face à ce risque de réduction de l'évaluation à un outil de gestion publique³⁸, *Socio-économie du travail* entend rendre compte de la richesse de la réflexion socio-économique sur l'attelage expérimentation-évaluation, une richesse manifeste dans la littérature internationale (en particulier au Québec³⁹) mais étrangement peu relayée en France. À l'opposé d'une SHS technocratique, cela concerne en premier lieu l'examen du choix des valeurs en évaluation, que ce soit selon des approches pragmatiques⁴⁰ ou dans le cadre de disciplines émergentes telles que les *Valuation Studies* (Vatin, 2013) ou les *Design Studies* (Delahais *et al.*, 2019). Cela concerne aussi l'enjeu de contextualisation prudente des résultats de l'EPP dans le champ des pratiques selon la perspective de l'*Implementation Research* développée en santé publique.

³⁷ Non sans susciter des mises en garde provenant du monde universitaire (par exemple dans le cadre d'une tribune publiée dans la presse en juin 2018 : https://www.liberation.fr/debats/2018/06/20/politiques-publiques-vers-une-evaluation-bien-trop-aleatoire_1660716/).

³⁸ Au sens du *New public management* (Longtin, 2021, p. 6).

³⁹ Deux ressources francophones principales ont nourri la présentation proposée ici : la revue de littérature et des pratiques de Longtin (2021) ; l'ouvrage dirigé par Rey *et al.* (2022). Ces références ont en particulier guidé la terminologie retenue face à l'hétérogénéité lexicale (pour partie liée à des problèmes de traduction de l'anglais) caractérisant le paysage des approches évaluatives. On privilégie ici le lexique élaboré par Marceau et Sylvain (Rey *et al.*, 2022, p. 37).

⁴⁰ Voir notamment les approches évaluatives « concernées par les valeurs » distinguées par Delahais : évaluation participative, *responsive evaluation*, *democratic deliberative evaluation* (Delahais, 2022, p. 163).

De façon plus directe, l'exercice initié par le présent *opus* de *Socio-économie du travail* propose à la communauté de l'évaluation des politiques de l'emploi de redécouvrir une perspective dépassant l'évaluation d'impact privilégiée par l'*evidence-based policy* : l'approche évaluative dite « basée sur la théorie ». Il s'agit de la fameuse *Theory-driven* ou *Theory-based evaluation* élaborée patiemment depuis cinquante ans dans le monde anglo-américain (Chen, 2014). Si feu le Conseil scientifique de l'évaluation lui avait prêté peu d'intérêt, cette approche « alternative » s'avère, depuis 2013, officiellement recommandée par la Commission européenne pour l'évaluation d'impact des fonds structurels (EC, 2013).

POLITIQUES DE L'EMPLOI : LA FIN DE L'INTERVENTION SANS L'EVALUATION ?

Ces dernières années, France Stratégie s'est livré à plusieurs exercices d'*Utilization Studies*⁴¹ (Bono *et al.*, 2018 ; Baïz, 2022). En France, relativement aux autres économies avancées, on a mené peu d'évaluations d'impact méthodiques depuis vingt ans et lorsqu'elles avaient lieu, on en mobilisait peu les résultats. Ce constat n'est peut-être pas étranger à la prééminence des arguments théoriques dans les diagnostics en vogue depuis les années 2000. Les politiques de l'emploi mises au régime de la théorie n'appellent alors plus qu'à un jugement d'adéquation au programme du marché, déniaient son intérêt à l'évaluation. L'assurance chômage allonge-t-elle le temps accordé aux personnes privées d'emploi pour retrouver une situation conforme à leurs attentes ? une « bonne » intervention doit en réduire la générosité. Un segment de main-d'œuvre rechigne-t-elle à occuper des emplois indécents (au sens de l'OIT) ? Une « bonne » intervention consiste à les y « inciter ». La protection de l'emploi freine-t-elle les réallocations de main d'œuvre ? Une « bonne » intervention est celle qui limite cette protection.

Lorsque la carte (la théorie économique) tient lieu de territoire (le réel), l'évaluation n'a plus lieu d'être. La déformation des ressources consacrées aux politiques de l'emploi (encadré) pourrait probablement être interprétée en ce sens. Pourtant, en dépit du recul des ressources dédiées aux mesures « actives ciblées » de politiques de l'emploi (encadré), suggérant un assèchement de la réflexion, s'est imposée une myriade d'innovations, plus ou moins marginales, ayant en commun le recours à l'expérimentation. Caractérisées par leur pragmatisme, ces innovations se sont imposées en bousculant les représentations de l'orthodoxie économique : des diagnostics moins univoques, des hypothèses comportementales moins figées, une meilleure prise en compte des réalités contextuelles, des analyses plus riches de la création d'emploi et de l'intermédiation offre-demande.

LES PROMESSES DU TOURNANT EXPERIMENTAL

Les observations précédentes aident peut-être à comprendre les espoirs suscités par l'expérimentation dans le domaine travail-emploi. Si, depuis les débats relatifs à la loi constitutionnelle du 28 mars 2003, le recours à l'expérimentation est devenu *la* bonne pratique (Conseil d'État, 2019, p. 13), cela a pu paraître spécialement salutaire pour les politiques de l'emploi (en particulier concernant les mesures dites « actives ciblées » – encadré). Accompagnement personnalisé des demandeurs d'emploi, RSA, GJ et, aujourd'hui, TZCLD, SIAE⁴² et entreprises vers l'emploi (SEVE), Revenu de base... autant de promesses d'innovation adossées à une expérimentation. Le référentiel théorique reste certes parfois celui

⁴¹ Étude du recours à l'évaluation et de la mobilisation de ses résultats.

⁴² Structure d'insertion par l'activité économique.

du « marché total » (Supiot, 2010), au moins l'obligation d'évaluation inhérente à l'expérimentation donne-t-elle au réel l'occasion de dire son fait.

À cet égard, l'appel à article lancé par *Socio-économie du travail* posait deux questions principales : comment la multiplication des expérimentations renouvelle-t-elle approches et méthodes d'évaluation des politiques de l'emploi ? Comment renouvelle-t-elle la fabrique de ces politiques ? Tirer les fruits de l'expérimentation pour élaborer des politiques de l'emploi satisfaisantes ne va en effet pas de soi : cela repose très largement sur la sincérité de la démarche d'évaluation et sa capacité à laisser sa « liberté d'expression » au réel (Bunel, 2008). À rebours de l'enthousiasme ambiant, un enjeu du présent exercice est en effet aussi de comprendre en quoi l'expérimentation peut s'avérer stérile et d'en savoir plus sur les moyens de conjurer ce risque.

LES POLITIQUES DE L'EMPLOI MISES AU REGIME DE L'EXPERIMENTATION

Traiter ces questions nécessite d'y voir clair sur l'articulation évaluation-expérimentation. Cette articulation n'est pas univoque parce que la fonction de l'expérimentation dans le processus d'innovation sociale n'est elle-même pas univoque. Ce point est abordé ici à travers la distinction entre l'étape de la *preuve de concept* et celle du *démonstrateur* (test d'hypothèse). La présente section entend montrer en quoi cette distinction est essentielle pour définir de bonnes pratiques en matière d'évaluation. Il s'agit aussi de délimiter l'objet de chacun des numéros tirés de l'appel à articles : un premier opus consacré à la conception expérientielle des politiques de l'emploi (la preuve de concept) et un second, consacré à l'évaluation expérimentale des politiques de l'emploi (la démonstration d'impact).

EXPERIMENTAL OU EXPERIENTIEL ?

Dans son rapport sur l'expérimentation sociale comme levier d'innovation de politique publique, le Conseil d'État déplore l'hétérogénéité des démarches qu'il recense (Conseil d'État, 2019). Une part du constat tient au peu de cas que ce rapport fait de définitions qu'il prend pourtant le soin de poser. La même étiquette, « expérimentation », est en particulier accolée à deux démarches très différentes : l'une consistant à *concevoir* une intervention (perspective d'innovation sociale – Longtin, 2021), l'autre à *tester* l'hypothèse de son effet. Une façon simple de lever la confusion est de jouer sur l'élasticité du terme « expérience », en en retenant l'acception familière de moyen d'apprendre à l'épreuve des choses (apprentissage par la pratique), et de réserver « expérimentation » à une démarche de test d'hypothèse (apprentissage hypothético-déductif). Dans le processus d'innovation, l'expérience se situerait au stade de la preuve de concept (est-ce que *ça peut* marcher ? si oui, sous quelles conditions ?) là où l'expérimentation correspondrait à celui du démonstrateur (est-ce que *ça marche* vraiment ? si oui, comment ?).

L'évaluation expérimentale des politiques de l'emploi (opus 2)

Le démonstrateur figure aux derniers stades du processus d'innovation⁴³. L'intervention est alors bien spécifiée (sa conception, ses ressources, ses modalités de mise en œuvre...) et l'expérimentation vise à en évaluer l'impact. Le type d'intervention considéré a le plus souvent les caractéristiques d'une *prestation*, un traitement suffisamment passe-partout pour qu'on

⁴³ Correspondant peu ou prou à l'évaluation sommative distinguée par Longtin (2021, p. 49-51).

puisse présupposer sa « scalabilité » (Lowenhaupt Tsing, 2012) c'est-à-dire une capacité à être diffusé sans altération. Ces cas de traitement-prestation correspondent précisément à la manière dont les tenants de l'*evidence-based policy* envisagent les politiques publiques.

L'évaluation se réfère alors typiquement au paradigme positiviste et s'efforce d'entretenir un rapport d'extériorité vis-à-vis de l'expérimentation : les RCT forment ce qu'il est convenu d'appeler l'étalon-or des approches évaluatives mobilisées au stade de la démonstration. Ces protocoles a-théoriques sont pourtant confrontés à un certain nombre de limites (Deaton et Cartwright, 2018 ; Jatteau, 2018). Leur analyse et la nécessité d'inscrire les RCT dans des approches évaluatives plus générales constituent l'objet du second opus consacré par *Socio-économie du travail* aux conséquences du tournant expérimental sur l'évaluation des politiques de l'emploi. Pour en donner un avant-goût, disons qu'il s'agit de redéfinir le rôle de la théorie. Le cadrage théorique est certes indispensable pour l'évaluation (Heckman, 2020) mais, au point où en sont nos connaissances sur les processus de mobilisation de la force de travail, la théorie en question ne saurait être générique. Cette proposition est évidemment celle de l'évaluation d'impact *basée sur la théorie*⁴⁴ (White, 2009) mais elle rejoint plus largement la recherche d'un équilibre entre positivisme et pragmatisme (Chateauraynaud, 2004).

Principaux enjeux qu'il s'agira d'aborder dans ce second opus : le problème de l'inférence causale, de l'évaluation d'interventions dites *complexes* ou de l'articulation de méthodes relevant des registres quali- et quanti-tatifs dans l'évaluation des politiques de l'emploi. On trouvera cependant déjà, dans le présent numéro, quelques éléments d'analyse sur ces questions en particulier dans les entretiens menés avec **Mathilde Gaini** et **Jérôme Gautié** (*Cf. infra*) et dans l'article de **Camille Retsin**.

La conception expérientielle des politiques de l'emploi (opus 1)

Reportant les enjeux précédents au numéro à paraître, le présent opus entend se concentrer sur l'expérience comme procédé de conception des politiques de l'emploi (au stade de la preuve de concept, donc). Les contributions rassemblées questionnent le statut de l'évaluation dans le cadre d'une *expérience collective*. Pour cela, le numéro s'appuie principalement sur deux expériences emblématiques, celle de la Garantie jeunes (GJ) et de Territoires zéro chômeurs de longue durée (TZCLD), mais aussi sur une analyse de la Loi « Avenir pro » de 2018. À partir de ces trois exemples, il s'agit de croiser parole d'experts (recueillie à l'occasion d'une journée d'étude) et articles de recherche aux perspectives disciplinaires variées : économie, statistique, sociologie, sciences politiques et droit. La troisième section de la présente introduction en propose une mise en perspective.

Au préalable, il convient d'indiquer en quoi le statut de l'évaluation accompagnant une expérience collective (preuve de concept) diffère de ce qui prévaut pour une expérimentation sociale (démonstrateur). Si l'évaluation est une opération extérieure à l'expérimentation sociale, elle est inhérente à l'expérience collective... au point d'être généralement active ou générative dans le processus d'innovation (passage du préconcept au concept). Le plus souvent *in itinere* ou « embarquée » (Goter et Khenniche, 2022), l'évaluation doit alors négocier continûment son existence dans le déploiement de l'expérience collective parce qu'elle y participe en même temps qu'elle l'observe. On voit la tension que cela représente : la connaissance produite doit à la fois nourrir le processus expérientiel et permettre un jugement éclairé sur un projet d'intervention. La rigueur dans l'évaluation d'expériences collectives

⁴⁴ Au sens de « théorie de l'intervention » ou de théorie « avec un petit t » (Chen et Rossi, 1987) ; on y revient plus bas.

convoque pour cela d'autres approches et méthodes que celles requises par l'évaluation d'expérimentations sociales.

L'ÉVALUATION DES EXPÉRIENCES COLLECTIVES

D'autres méthodes d'évaluation, donc. Mais lesquelles ? Pour saisir l'enjeu d'appliquer une intention scientifique à l'évaluation d'expériences collectives, le préalable est de souligner que le procédé répond à l'incertitude⁴⁵ d'un diagnostic de politique publique. C'est parce qu'on sait qu'on n'en sait pas suffisamment sur les processus sociaux (théorie positive) qu'il s'agit d'infléchir par une intervention, que l'on procède à une expérience. D'où, dans le domaine travail-emploi, le rôle des déceptions associées aux politiques actives du marché du travail (encadré) dans la réception du tournant expérimental : manifestement, ces politiques mobilisaient une théorie fragile.

La question de la théorie n'est pourtant pas absente de l'évaluation des expériences collectives. Mais il s'agit d'une théorie en cours d'élaboration : une *théorie des pratiques* (Dubuisson-Quellier et Plessz, 2013) d'acteurs sociaux cherchant précisément à redéfinir leurs pratiques. La spécificité de l'évaluation d'expérience collective tiendrait alors dans la nécessité d'élaborer une théorie *ad hoc*, la *théorie de l'intervention*⁴⁶ (Funnell et Rogers, 2011). Cela consiste à expliciter les intuitions de personnes porteuses d'un préconception d'intervention ; intuitions pouvant reposer sur des « théories en usage », un référentiel propre aux personnes porteuses, ou mobiliser une littérature académique sur les processus sociaux sous-jacents. Quelle qu'en soit l'origine, la théorie de l'intervention met en commun des présupposés et des attentes.

Ce travail de reconstitution est commun aux méthodes de l'évaluation des expériences collectives. Si ces dernières contrarient par nature l'expérimentation contrôlée, deux approches évaluatives méritent d'être particulièrement discutées : l'évaluation basée sur la théorie (*Theory-based evaluation*, TBE ; Weiss, 1997), déjà évoquée, et la collecte de preuves d'efficacité basée sur l'observation des pratiques (*Practice-based evidence*, PBE ; Schön, 1992).

L'application de la TBE aux expériences collectives

Dans un rapport consacré au rôle dévolu à l'accompagnement dans les « politiques actives du marché du travail », le Conseil d'orientation de l'emploi recommandait d'ouvrir la « boîte noire » de l'accompagnement (COE, 2016). Cet agenda, ouvrir la « boîte noire » des interventions publiques, constitue précisément la proposition générale de la TBE. Dans cette approche, la théorie de l'intervention est une explicitation de la chaîne causale reliant hypothétiquement un dispositif de politique publique à des résultats escomptés (*ex ante*) par les parties prenantes. Il ne s'agit pas d'une théorie générique surplombante mais des hypothèses sur lesquelles se fondent l'intervention.

La première étape d'une TBE est ainsi d'explicitation ces hypothèses (à l'appui d'une documentation préexistante ou d'arguments invoqués par les parties prenantes), typiquement sous forme de logigramme. La décomposition obtenue peut articuler une *théorie du changement* (pourquoi l'intervention est censée fonctionner ?) et une *théorie de l'action* ou *de mise en œuvre*

⁴⁵ Incertitudes pouvant marquer la réalisation ou la mise en œuvre d'une intervention publique. L'expérience collective doit permettre le repérage et la pesée de l'ensemble des facteurs (contextuels, organisationnels, relationnels) susceptibles de produire un effet, conditionnant ainsi l'efficacité de l'intervention.

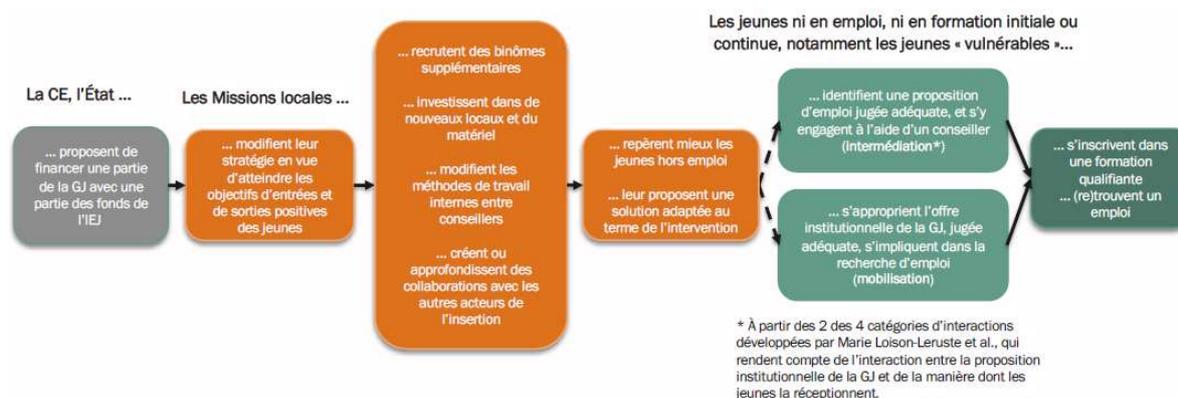
⁴⁶ Voir *infra*. On reprend ici la terminologie de Marceau et Sylvain (2022, p. 219) mais cela recouvre le concept de *théorie de/du programme* couramment utilisée par la communauté de l'évaluation.

(quelles dispositions pratiques permettraient à l'intervention d'opérer ?). Le cadre obtenu est le référentiel de l'évaluation (que celle-ci use de méthodes quali- ou quanti-tative) à l'aune duquel est formulé le jugement attendu.

Les FIG. 3 et 4 offrent l'exemple de logigrammes construits pour l'évaluation de la GJ concernant respectivement : son financement par des fonds européens de l'Initiative pour l'emploi des jeunes, sa mise en œuvre par les missions locales. L'idée du logigramme de la FIG. 3 est que l'ensemble des maillons du chemin causal (symbolisées par des flèches) doivent être activés pour que l'intervention produise les résultats escomptés : l'enquête évaluative consiste à examiner chaque maillon. L'enjeu de la FIG. 4 est d'orienter l'enquête évaluative vers différents moments (amont, mise en œuvre, aval) et niveaux (employeurs potentiels, missions locales, jeunes *neet*⁴⁷) de l'intervention GJ avec des questions évaluatives précises : la GJ a-t-elle permis de mobiliser les jeunes *neet* ? l'orientation vers un accompagnement collectif ou personnalisé a-t-elle opéré ? les missions locales sont-elles parvenues à adapter leurs pratiques ? etc.

Cette évocation (trop rapide mais nous y reviendrons dans le second opus) a pour but de souligner l'adéquation de la TBE à l'évaluation des expériences collectives. La TBE permet en effet de structurer *logiquement* les expériences collectives en amenant les parties prenantes, porteuses d'un *préconception* d'intervention, à expliciter leurs propres objectifs et mieux maîtriser le processus d'innovation. En structurant l'élaboration progressive et itérative de repères communs, la TBE participe à l'identification de points d'achoppement : c'est en particulier le cas des *analyses de contribution* (Delahais, 2023), méthode associée à la TBE, qui permet notamment d'évaluer la cohérence entre moyens et objectifs. La TBE peut éclairer des conditions d'effectivité et de concrétisation d'une intervention, notamment en présence de configurations institutionnelles complexes impliquant un partenariat entre acteurs.

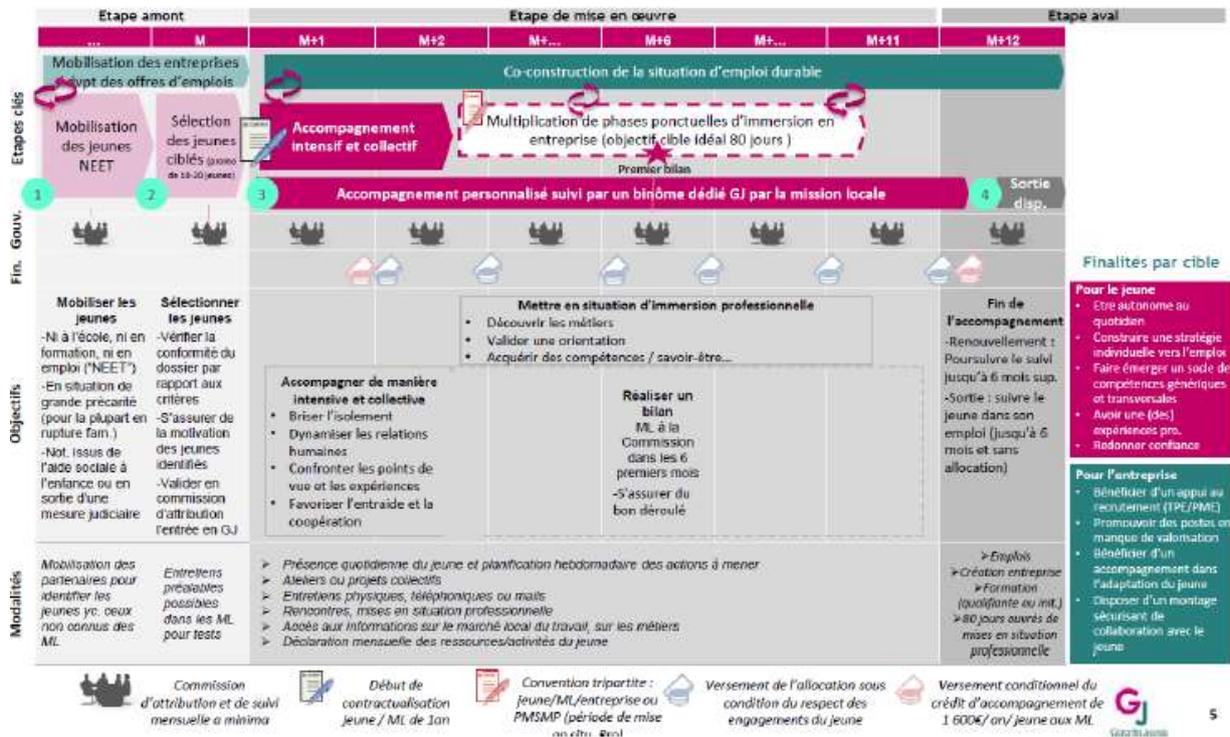
FIG. 3 – Théorie du changement de la GJ financée par l'Initiative pour l'emploi des jeunes



Source : Quadrant Conseil – KPMG (2018), « Étude de cas Garantie Jeunes », *Évaluation de l'IEJ*, p. 3.

⁴⁷ Neither in employment nor in education or training.

FIG. 4 – Théorie de mise en œuvre de la GJ selon le Secrétariat général pour la modernisation de l'action publique



Source : Farvaque N., C. Kramme et C. Tuchsirer (2016), La Garantie Jeunes du point de vue des missions locales, *Rapport Cnam-CEET*, p. 19.

L'approche PBE dans le cadre d'expériences collectives

En médecine, la collecte de preuves d'efficacité par l'observation des pratiques (PBE) propose de partir de l'exercice clinique et se présente comme complémentaire d'un modèle se réclamant d'*evidence-based practices* (Swisher, 2010). Celui-ci consiste à prescrire aux patients, indépendamment des singularités de leur cas, des traitements ayant de bons effets moyens, tels qu'estimés par RCT. La collecte méthodique de PBE répond au contraire à un monde de singularités, un réel complexe que l'on ne peut « transformer en laboratoire ». En lieu et place, les pratiques qui y ont cours sont documentées (étude qualitative) et mesurées (étude quantitative) dans les conditions imparfaites de leur mise en œuvre. L'enjeu n'est pas d'obtenir des résultats contrôlés mais de procéder à l'analyse des pratiques et d'en classer les résultats. La question est moins « est-ce que X cause Y en moyenne ? » que « dans quelle mesure telle intervention a-t-elle modifié le système complexe et singulier auquel on est confronté en pratique ? ». Partant d'un symptôme donné, l'idée n'est pas de sélectionner un traitement moyen validé par un protocole contrôlé mais de documenter la diversité des pratiques et de repérer les contextes dans lesquels elles obtiennent un résultat satisfaisant : dans cette démarche, les « tipologistas » (Cf. l'entretien avec Jérôme Gautié) se substituent aux « randomistas » (Deaton *et al.*, 2018).

La collecte de PBE peut utilement être mise au service de l'évaluation des expériences collectives. Cette démarche est déjà courante en sciences de gestion et en science de l'éducation.

Dans une perspective développée autour du thème de l'apprentissage organisationnel (Schön, 1983) la collecte de PBE peut mobiliser la *pratique réflexive* ce qui paraît particulièrement adapté à l'évaluation des expériences collectives. Comment améliorer les

pratiques à partir de l'expérience ? L'expérience seule ne conduit pas nécessairement à l'apprentissage. Il faut, nous dit Donald Schön, placer praticiennes et praticiens dans une posture de réflexivité, au besoin en encourageant le jugement critique sur les pratiques courantes (les siennes et celles de ses pairs). C'est exactement à cela que peut servir l'évaluation : tendre un miroir le moins déformant possible aux pratiques et s'appuyer sur les pratiques *réflexives* pour leur assigner une valeur. À la « parenthèse réflexive » évoquée *supra*, Schön apporte un régime expérientiel enrichi : plus qu'une simple évaluation des conséquences de la loi (*what we do / what we get*), il propose un apprentissage en « double boucle », permettant d'interroger le fondement même d'une action publique (*Why we do what we do*) – Argyris et Schön, 2002.

À propos de l'évaluation de dispositifs éducatifs états-uniens, l'agenda PBE est défendu avec ferveur par certains auteurs (Bryk *et al.*, 2015). Dans un article qu'il consacre à cet enthousiasme, Pascal Bressoux met pourtant en garde contre une évaluation relevant d'une pure collecte de PBE (Bressoux, 2017). L'enjeu est, selon lui, d'équiper la démarche avec les résultats d'expérimentations contrôlées. Ce point fera l'objet de développement dans l'opus à paraître parce qu'il est essentiel : qu'il s'agisse d'éducation ou de politiques de l'emploi, l'évaluation pluraliste ne saurait être une « évaluation séparatiste ». Dans le cas d'espèce, on fera volontiers le pari que la collecte de PBE est d'autant plus fructueuse qu'elle n'ignore pas les résultats de l'expérimentation contrôlée et, réciproquement, qu'une collecte de PBE peut utilement guider l'expérimentation contrôlée. Mais l'on s'écarte ici de l'évaluation des expériences collectives.

Pour y revenir et se rapprocher des politiques de l'emploi, on indiquera que l'évaluation par la collecte de PBE n'est pas absente des usages nationaux. La collecte de PBE est en fait au principe d'évaluations *participatives* parfois anciennes : l'expérimentation « Contrats aidés-structures apprenantes » (Casa), le projet « Actions de formation en situation de travail » (Afest) ou, dans le cadre du Plan d'investissement dans les compétences (Pic), un dispositif comme *Prépa-compétences*. On est chaque fois dans une détermination *in itinere* de l'intervention (Cf. l'article de **Vincent Cibois**), processus dans lequel les *pratiques* jouent un rôle actif : l'enjeu est alors de ménager à l'évaluation une part d'extériorité, que le miroir tendu aux pratiques ne soit pas déformant.

ÉVALUER LES POLITIQUES DE L'EMPLOI EN LES EXPERIMENTANT : QUELLES TENSIONS ?

Si l'association expérimentation-évaluation est impérative, elle ne va pas sans difficultés pratiques. Le point commun des contributions du présent numéro est de documenter les tensions de cette association et les moyens d'y faire face.

LA GARANTIE JEUNES : RETOUR D'EXPERIENCE

Une première proposition du présent numéro est de questionner l'état de l'évaluation des politiques de l'emploi en France à travers le témoignage de deux observateurs de choix : **Mathilde Gaini**, statisticienne spécialiste de l'évaluation d'impact à la Dares, et **Jérôme Gautié**, économiste, qui préside le Conseil scientifique de France travail (ex Pôle emploi). *Socio-économie du travail* leur a proposé une discussion centrée sur l'évaluation de la GJ, à laquelle ils ont tous les deux pris part (en tant qu'évaluatrice pour l'une, que président du conseil scientifique pour le second). L'intérêt de partir de la GJ est que c'est un bon exemple d'intervention pouvant être qualifiée de *complexe*. La GJ agrège en effet sous un même titre

plusieurs principes d'action (contrat, aide monétaire, accompagnement adossé à des stratégies de médiation du recrutement...) pouvant être diversement articulés d'un déploiement local à l'autre (Cf. FIG. 4). Ce type d'intervention composite s'avère de plus en plus courant dans les politiques de l'emploi expérimentées ces dernières années (par exemple dans le cadre du Plan d'investissement dans les compétences ou de la Stratégie pauvreté) et posent des questions spécifiques à l'évaluation.

Une façon d'y faire face est de diversifier les méthodes d'évaluation, relevant des registres quantitatif, pour la *mesure* d'impact, et qualitatif, pour la *description* de processus, et d'exploiter leurs complémentarités. Mais cela ne va pas de soi. Les expériences récentes suggèrent divers modèles faisant plus ou moins fortement peser l'effort de mise en cohérence de la mesure et de la description sur les instances de coordination (comité scientifique, de suivi, etc.) et articulant plus ou moins les calendriers de travail.

L'évaluation de la GJ est réputée avoir donné lieu à une implication très forte de son comité scientifique pour coordonner « par le haut » enseignements qualitatifs et quantitatifs : la division du travail suit les spécialités disciplinaires mais le comité scientifique anime un dialogue régulier entre les équipes et s'efforce, à l'étape de l'écriture des rapports, de faire communiquer les éclairages. Ce modèle d'intégration « par le haut » assure mois après mois la cohérence d'ensemble de l'évaluation : les différentes équipes ajustent au fil du calendrier d'évaluation l'interprétation des éléments qu'elles mettent à jour.

On peut toutefois s'interroger sur la portée de ce modèle pour tirer tous les fruits de la complémentarité entre *mesure* et *description*. On sait en particulier l'importance que revêt pour l'évaluation contrefactuelle (qu'il s'agisse d'expérimentation contrôlée ou de quasi-expérimentation), une compréhension fine des processus déterminant l'assignation au « traitement ». Cela suggère l'opportunité d'une coordination « par le bas » entre mesure et description, celle-ci pouvant par exemple apporter un éclairage de terrain sur la plausibilité d'une *hypothèse d'exclusion* (dans le cadre d'une estimation d'impact par variable instrumentale) ou sur l'application d'un protocole d'assignation aléatoire. On peut plus généralement envisager une forme d'hybridation par laquelle la statistique accède à une meilleure compréhension de ce qu'elle tient pour ses « données » tandis que l'ethnographie parvient à mieux analyser, dans ses observations, ce qui relève du cas type de ce qui constitue le cas limite.

À ces enjeux relevant largement d'une perspective technique sur l'évaluation, Mathilde Gaini et Jérôme Gautié apportent une autre dimension : celle de l'implication des parties prenantes, que celles-ci soient porteuses de projet ou *ressortissantes* (Warin, 1999). C'est peut-être le point de convergence le plus fort entre leurs témoignages : l'évaluation doit échapper au penchant technocratique prévalant en France ! Non pas seulement au nom de principes démocratiques mais aussi, de façon moins attendue, pour des raisons scientifiques : l'évaluation est plus fructueuse lorsqu'elle s'appuie sur les savoirs indigènes plutôt que de les ignorer. Cela ouvre tout un champ de réflexion quant aux relations entre recherche, administration et « publics » qui va bien au-delà du champ de ces numéros thématiques de *Socio-économie du travail*.

LES TERRITOIRES ZERO CHOMEUR DE LONGUE DUREE : L'EVALUATION IMPOSSIBLE ?

L'expérimentation TZCLD (encore en cours) propose de renouveler la manière de penser les causes du chômage de longue durée : plutôt que de l'imputer aux individus et à leur « employabilité », les promoteurs de TZCLD suggèrent de réfléchir à *partir* des capacités propres à ces derniers pour leur proposer des emplois adéquates. Il s'agit aussi d'une intervention complexe articulant plusieurs principes : « nul n'est inemployable », « ce n'est par

le travail qui manque », « ce n'est pas l'argent qui manque » (l'intervention se voulant neutre pour les finances publiques). Le présent opus rassemble trois articles consacrés à l'expérimentation TZCLD : à l'échelle nationale pour le premier, multi-échelle pour le deuxième, et territorial pour le troisième.

Les deux premiers se livrent à un exercice de sociologie politique du processus d'évaluation de l'expérimentation TZCLD : il s'agit en quelque sorte de faire l'évaluation... d'un processus d'évaluation.

L'article de **Florence Jany-Catrice**, **Anne Fretel** et **Laurent Gardin** s'intéresse à la profusion d'exercices nationaux d'évaluation occasionnés par l'expérimentation TZCLD. Adoptant une perspective de sociologie compréhensive, l'article applique les méthodes de l'analyse de controverse pour inscrire le projet d'une action publique « justifiée par la preuve » dans un jeu d'acteurs nationaux. La controverse s'est déployée dans deux dimensions principales : l'analyse coût-bénéfice de l'intervention ; la question du ciblage des personnes en chômage de longue durée. Autrices et auteur de l'article pointent un « péché originel » des promoteurs de TZCLD dont le plaidoyer reposait excessivement sur l'hypothèse d'une rationalité budgétaire de l'intervention. Cette hypothèse a concentré une bonne part des énergies : au lieu de documenter le plus objectivement possible l'ensemble des effets de l'intervention, nombre d'évaluations ont consisté à tester l'hypothèse de sa neutralité budgétaire, se limitant parfois à un exercice de comptabilité monétaire. L'examen suggère au final une certaine indifférence des acteurs à la qualité (d'ailleurs très variable) des « preuves » rassemblées. À l'instar de ce que l'on a évoqué *supra* à propos de politiques de l'emploi soumises au régime de la théorie, le jeu d'acteurs aurait dévitalisé la démarche évaluative, celle-ci devenant instrument de lutte plutôt que de rationalisation de choix collectifs.

L'analyse multi-échelle d'**Aubin Tantot** confirme largement ce diagnostic en y ajoutant une réflexion sur l'hypothèse que l'expérimentation favorise une *territorialisation* des politiques de l'emploi (CGET, 2016, p. 5). TZCLD se prête en effet particulièrement bien à un tel questionnement. La proximité étant en soi tenue pour source d'efficacité, territorialiser l'intervention publique est essentiel. Mais cela peut vouloir dire deux choses très différentes : territorialiser une initiative nationale (action publique descendante, *top-down*) ; nationaliser une initiative territoriale (action publique ascendante, *bottom-up*). Logique de projection dans le premier cas, de projet dans le second. Aubin Tantot montre que la structuration institutionnelle de l'expérimentation TZCLD a pu hybrider ces logiques. Portée par le monde associatif, l'initiative a échappé à l'État central dans sa phase de conception. Dans le même temps, le Fonds d'expérimentation est parvenu, par la formule de l'appel à projet, à imposer des exigences nationales aux porteurs locaux de l'expérimentation. En somme, si TZCLD (première vague) a pu échouer à faire vivre le tandem expérimentation-évaluation, il a peut-être été plus heureux quant au tandem expérimentation-décentralisation.

C'est aussi l'un des enseignements de l'article de **Camille Retsin**. Mais au-delà d'aspects analytiques, ce troisième article dévolu à l'expérimentation TZCLD s'intéresse aux enjeux méthodologiques propres à l'évaluation d'une expérience collective *territorialisée*.

En matière d'évaluation, l'enquête de terrain (observation ethnographique, entretiens semi-directifs...) est indispensable (*Cf.* le témoignage de **Mathilde Gaini**). Elle permet la collecte d'une information essentielle sur les contextes locaux, les contraintes pesant sur les acteurs ou encore les facteurs de résistance au changement. Pour autant, les méthodes d'enquête qualitative sont diverses, autant que la connaissance que chacune permet d'apporter. Cela s'est manifesté de façon particulièrement nette dans le cadre de l'expérimentation TZCLD. Parmi l'ensemble des méthodes qualitatives déployées, Camille Retsin entend illustrer la portée de l'*analyse processuelle* (Mendez, 2010) dans le cadre de l'évaluation d'une intervention publique

territorialisée. TZCLD sert à l'autrice de pierre de touche : elle montre que l'analyse processuelle apporte des éclairages décisifs sur l'importance de la coordination des opérateurs de terrain (en l'espèce, entre *entreprise à but d'emploi* et *comité local pour l'emploi*), un résultat que les autres perspectives, malgré leur diversité, n'ont pas permis d'obtenir. La formalisation, voire l'institutionnalisation, d'un *droit à l'emploi* (Cf. l'encadre proposé par **Jany-Catrice, Fretel et Gardin**), l'utopie fondant TZCLD, tiendrait ainsi pour partie à une capacité d'organisation locale insuffisamment mise en jeu dans la conception des politiques de l'emploi.

LA FABRIQUE EXPERIMENTALE DU DROIT DE LA FORMATION PROFESSIONNELLE

L'un des écueils de la frénésie d'« évaluations » associée à TZCLD est d'avoir souvent réduit ce projet à une politique de l'emploi. L'*économisme* pointé par Florence Jany-Catrice, Anne Fretel et Laurent Gardin allait en ce sens au-delà d'une prédilection pour l'analyse coût-bénéfice : il tenait au fait de juger du projet TZCLD au seul prisme de l'accès à l'emploi. Dans son article, **Vincent Cibois** part d'un constat analogue à propos de la formation professionnelle. Au fil des années, le sujet en est venu à être de plus en plus souvent jugé à la seule aune de gains en « employabilité » (ce dont témoigne l'inclusion récente de la formation professionnelle dans les publications de la Dares consacrées aux dépenses en faveur de l'emploi et du marché du travail⁴⁸). Une conséquence de ce rapprochement est qu'on évalue volontiers la formation comme s'il s'agissait d'un dispositif de politique de l'emploi... ce qu'en droit, elle n'est pas. La politique de l'emploi relève du livre premier de la cinquième partie du Code du travail ; c'est à la sixième partie que correspond la formation professionnelle. Pour le juriste, celle-ci n'est donc pas une collection de dispositifs mais *une disposition*, un droit (sinon un devoir) attaché au salariat. L'implication logique de cette observation est que juger notre système de formation professionnelle n'est pas qu'un enjeu d'EPP mais aussi d'évaluation de la *qualité du droit* entendue comme norme suffisamment souple, intelligible et cohérente pour inspirer aux agents sociaux des comportements conformes à l'esprit de la loi. La réflexion de Vincent Cibois consiste, dans ce cadre, à examiner ce que pourrait induire le tournant expérimental/expérientiel en matière de fabrication du droit : il n'est question ni de mesure d'impact ni d'évaluation de mise en œuvre, mais de tirer d'une dérogation temporaire et localisée au droit commun, l'occasion d'un jugement de qualité. Cette réflexion ne s'appuie pas par hasard sur le cas de la formation professionnelle : c'est un champ où l'expérimentation est omniprésente. À travers cet exemple, Vincent Cibois examine donc, l'hypothèse selon laquelle le recours à l'expérimentation pourrait alimenter un processus de progression continue de la qualité du droit. La portée de cette hypothèse dépend du cadre de l'expérimentation qui peut, ou non, relever d'une *loi expérimentale* (Chevallier, 1996). La notion de loi expérimentale convoque en effet une méthodologie, alliant expérimentation et évaluation, faisant défaut à nombre d'expérimentations mises en œuvre dans le champ de la formation professionnelle. Lorsque la condition de bonne méthode était vérifiée, l'expérimentation s'est avérée propice à la qualité du droit de la formation professionnelle... mais nombreuses ont été les expérimentations restées sans effet favorables sur la qualité du droit.

Ce premier opus consacré aux politiques de l'emploi mises au « régime de l'expérimentation » comporte pour finir la recension, par **Arthur Jatteau**, observateur averti du tournant expérimental, d'une utile anthologie : *Evaluations : Fondements, controverses, perspectives* dirigée par Thomas Delahais, Agathe Devaux-Spatarakis, Anne Revillard et Valéry Ridde. Il s'agit d'une somme (573 p. !) rassemblant une collection d'articles parus en langue anglaise ces quarante dernières années, sur les questions pratiques et théoriques posées

⁴⁸ Dares Résultats n°49, *Dépenses en faveur de l'emploi et du marché du travail en 2021*, septembre 2023.

par l'évaluation des politiques publiques : à quoi sert l'évaluation ? qui évalue ? comment juger de la valeur des interventions ? l'évaluation est-elle une science ? quelle diversité de paradigmes ? Une façon pour ce numéro d'insister une dernière fois sur l'importance du couple expérimentation-évaluation. Procédant par « coups de projecteurs » sur les thèmes qui l'ont particulièrement intéressé, Arthur Jatteau met en évidence l'opportunité des questions posées, invitant la communauté nationale de l'évaluation à suivre les auteurs de l'ouvrage dans la découverte d'une littérature foisonnante peut-être insuffisamment mobilisée en France. Cette recension directement liée au thème du numéro est complétée par trois comptes rendus de lecture : celui de **Dominique Glaymann**, à propos d'*Employabilité et mutations industrielles* (dirigé par F. Noël et G. Schmidt), de **Chloé Daveux**, à propos de *Gouverner par l'emploi* (de C. Dupuy et F. Sarfati), et de **Marie Neihouser**, à propos de *Sociologie du numérique au travail* (de M. Benedetto-Meyer et A. Bococ).

RÉFÉRENCES BIBLIOGRAPHIQUES

Adjerad, S., Bouygar, F., Charpail, C., Klein, T., (2002), « Intérêts et limites d'une évaluation interne au service de la stratégie publique », Ministère de l'Emploi et de la Solidarité – Dares, Communication au colloque de la Société Française d'Évaluation « L'évaluation au service de la stratégie publique ».

Allègre, G., (2024), « Les nouvelles lois sur les pauvres (1989-202) : L'injonction au travail au risque de la pauvreté ? », *OFCE working paper*, n°2024-01.

Angrist, J. D., Pischke, J. S., (2010), « The credibility revolution in empirical economics: How better research design is taking the con out of econometrics », *Journal of economic perspectives*, vol. 24, n°2, p. 3-30.

Argyris, C., Schön, D. A., (2002), *Apprentissage organisationnel : théorie, méthode, pratique*. Paris, De Boeck Université-Management.

Arkwright, E., de Boissieu, C., Lorenzi, J.-H., Samson, J., (2007), *Économie politique de la LOLF*, Conseil d'analyse économique, Paris, La documentation Française.

Arrignon, M., (2019), « De l'expérimentation sociale en politique ou la preuve par les chiffres. Le cas de la réforme du Revenu de Solidarité Active », *Revue d'anthropologie des connaissances*, vol. 13, n°4, p. 1123-1144.

Askenazy, P., (2011), *Les décennies aveugles – Emploi et croissance (1970-2010)*. Paris, Éditions du Seuil.

Askenazy, P., Palier, B., (2018), « France: rising precariousness supported by the welfare state », *Document de travail du Cepremap*, n°1801.

Baguelin, O., (2012), « Politique conjoncturelle de l'emploi et structures des marchés du travail locaux : le déploiement territorial du contrat d'accompagnement dans l'emploi en 2009 », *Économie et statistique*, n°454, p. 79-107.

Baguelin, O., (2023), « Fragmentation de l'emploi et institutions du marché du travail : une analyse à partir du panel FH-DADS ». <http://o.baguelin.free.fr/> (consulté le 07/06/2024).

Baïz, A., Bono, P. H., Debu, S., Desplatz, R., Lacouette-Fougère, C., Trannoy, A., (2019), « Évaluation d'impact des politiques publiques : que peut apprendre la France des pays les plus avancés ? », France Stratégie, Note de synthèse.

Baïz, A., (2022), « Quelles évaluations des politiques publiques pour quelles utilisations ? », France Stratégie, Rapport.

Baron, A., Guillermin, G., (2023), « L'activité partielle en 2021 et au premier semestre 2022 », *Dares Résultats*, n°36, juin.

Behaghel, L., Crépon, B., Gurgand, M., (2009), « Évaluation d'impact de l'accompagnement des demandeurs d'emploi par les Opérateurs Privés de Placement-OPP et le programme Cap Vers l'Entreprise-CVE », *Paris School of Economics*, working paper.

Berthet, T., Cuntigh, P., Guitton, C., Mazel, O., (2002), « 1982-2002 : la territorialisation progressive des politiques de l'emploi », *Dares PI/PS*, n°24.2, Juin.

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

Borraz, O., Demortain, D., (2015), « Science réglementaire » in Henry, E. (éd.), *Dictionnaire critique de l'expertise. Santé, travail, environnement*, Paris, Presses de Sciences Po, p. 279-285.

Bressoux, P., (2017), « Practice-based research : une aporie et des espoirs », *Éducation et didactique* [En ligne], n°11-3.

Bryk, A. S., (2017), « Accélérer la manière dont nous apprenons à améliorer », *Éducation et didactique* [En ligne], n°11-2.

Bunel, M., (2008), « Comment met-on en œuvre une réforme en France ? Leçons de la loi sur les 35 heures », *La vie des idées*.

Carbonnier, C., Morel, N., (2018), *Le retour des domestiques*. Paris, Éditions du Seuil.

Chaubaud, M., Bucher, A., Givord, P., Louvet, A., (2022), « Quelles sont les chances de retour à l'emploi après une formation ? », Document d'études n°261, Dares.

Chateauraynaud, F., (2004), « L'épreuve du tangible – Expériences de l'enquête et surgissements de la preuve », *La croyance et l'enquête*, Raisons pratiques, Éd. de l'EHESS, p. 167-194.

Chen, H. T., Rossi, P. H., (1987), « The theory-driven approach to validity », *Evaluation and Program Planning*, n°10, p. 95-103.

Chen, H. T., (1990), *Theory-driven evaluations*, Thousand Oaks, CA: Sage.

Chen, H. T., (2014), *Practical program evaluation: Theory-driven evaluation and the integrated evaluation perspective*. Sage Publications.

Chevallier, J., (1996), « Les lois expérimentales » in *L'écriture du droit*, Paris, Éditions Diderot, p. 167-203.

Clouet, H., (2018), « La qualité au guichet : négocier le travail peu qualifié à Pôle emploi », *Travail et emploi*, n°155-156, p. 65-92.

Commission européenne, 2014, *Suivi et évaluation de la politique européenne de cohésion – Concepts et recommandations*. Document d'orientation.

Conseil d'État, (2019), « Les expérimentations : comment innover dans la conduite des politiques publiques ? », *Les études du Conseil d'État*, La documentation Française.

Conseil d'orientation pour l'emploi, (2016), *L'accompagnement dans et vers l'emploi*.

Coquet, B., (2016a), « L'assurance chômage doit-elle financer le Service public de l'emploi ? », *OFCE Les notes*, n°58/22.

Coquet, B., (2016b), « Marché du travail : redonner du sens aux réformes et aux politiques », Contribution au débat France Stratégie du 8 septembre 2016.

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

Delahais, T., Gouache, C., Vincent, S., (2019), « Le design de l'action publique : vers une hybridation entre culture design et culture de l'évaluation », *Sciences du Design*, vol. 2, n°10, p. 83-89.

Delahais, T., (2023), « Le choix des approches évaluatives », in Rey, L., Quesnel, J. S., Sauvain, V. (éd.), *L'évaluation en contexte de développement : Enjeux, approches et pratiques*, Montréal, École nationale d'administration publique, Les éditions JFD, p. 279-285.

Deleau, M., Nioche, J.-P., Penz, P., Poinard, R., (1986), *Évaluer les politiques publiques, rapport pour le Commissariat Général au Plan*, Paris, La Documentation française.

Deaton, A., Cartwright, N., (2018), « Understanding and misunderstanding randomized controlled trials », *Social Science & Medicine*, n°210, p. 2-21.

Dewey, J., (1938), *Logique. La théorie de l'enquête*. Paris, Presses Universitaires de France.

Duclos, L. (2009), « La flexicurité et la question des sécurités adéquates », *La revue de l'Ires*, n° 63, p. 35-62.

Esping-Andersen, G. (2008), *Trois leçons sur l'État-providence*. Paris, Éditions du Seuil.

Funnell, S. C., Rogers, P. J., (2011), *Purposeful program theory: effective use of theories of change and logic models*, San Francisco, Jossey-Bass/Wiley.

Gautié, J., Lerais, F., (2024), « Politiques d'exonération sur les bas salaires : usages et effets potentiels », Rapport de l'Ires pour la CFDT.

Gibert, P., Saussois, J.-M., (2022), « L'évaluation (ex post) de politique en France. Un accent mis sur la frontière de l'objet au détriment d'une incitation à l'objet-frontière ? », *Politiques & management public*, vol. 39, n°4, p. 587-610.

Gomel, B., Serverin, E., (2009), « Expérimenter pour décider ? Le RSA en débat », *Document du travail du Centre d'études de l'emploi*, n°119.

Goter, F., Khenniche, S., (2022), « Évaluation des politiques publiques : vers une pratique intégrée au pilotage de l'action publique », *Gestion et management public*, vol. 10, n°3, p. 35-56.

Grégoire, M., Guergoat-Larivière, M., Vivés, C., (2018), « Discontinuités de l'emploi et indemnisation du chômage – Introduction », *Socio-économie du travail*, n°3, p. 15-30.

Heckman, J. J., (2020), « Randomization and Social Policy Evaluation Revisited », *Iza, Discussion paper*, n°12882.

Henneguelle, A., Jany-Catrice, F., Lamarche, T., (2023), « Économie des services » in Association française d'économie politique (éd.) – Guy, Y., Henneguelle, A., Puissant, E. (coord.), *Grand manuel d'économie politique*, Malakoff, Dunod.

Jatteau, A., (2013), Les expérimentations aléatoires en économie.

Jatteau, A., (2018). « Les expérimentations aléatoires, le 'gold standard' des méthodes d'évaluation d'impact ? », *Regards croisés sur l'économie*, n°22, p. 184-194.

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.

Lacouette Fougère, C., Lascoumes, P., (2013), « L'évaluation : un marronnier de l'action gouvernementale ? ». *Revue française d'administration publique*, n°4, p. 859-875.

Longtin, D., (2021), *Outils d'évaluation en innovation sociale*.

Lowenhaupt Tsing, A., (2012), « On nonscalability: The Living World Is Not Amenable to Precision-Nested Scales », *Common Knowledge*, vol. 18, n°3, p. 505-524.

L'Horty, Y., (2013), « Politiques de l'emploi : pourquoi ça ne marche pas ? », *Regards croisés sur l'économie*, n°13, p. 91-102.

Marchand-Arvier, J., Deumier, P., (2014), « Expérimenter pour bien légiférer ? », *Revue de droit du travail*, p. 305-312.

Marceau, R., Sylvain, F., (2023), « La terminologie de l'évaluation », in Rey, L., Quesnel, J. S., Sauvain, V. (éd.), *L'évaluation en contexte de développement : Enjeux, approches et pratiques*, Montréal, École nationale d'administration publique, Les éditions JFD, p. 37-56.

Mendez, A., (2010), « Processus. Concepts et méthode pour l'analyse temporelle en sciences sociales », *Lectures, Publications reçues*.

Nioche, J. P., (1982), « De l'évaluation à l'analyse des politiques publiques », *Revue française de science politique*, p. 32-61.

Nioche, J. P., (1993), « L'évaluation des politiques publiques en France : 'fast food', recettes du terroir ou cuisine internationale ? », *Revue française d'administration publique*, n°66, p. 209-220.

Nioche, J. P., (2014), « L'évaluation des politiques publiques et la gestion en France: un rendez-vous manqué? ». *Revue française de gestion*, (8), 71-84.

Ould-Ferhat, L., (2018), « Entretien avec Vincent Spenlehauer », *Sociologies pratiques*, n°36, p. 17-24.

Palier, B., (2023), « Le travail *low cost* : Comment les stratégies de baisse du coût du travail ont abîmé le travail en France », *Esprit*, septembre, p. 51-59

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*, n°1, p.145-163.

Picart, C., (2014), « Une rotation de la main d'œuvre presque quintuplée en 30 ans : plus qu'un essor des formes particulières d'emploi, un profond changement de leur usage », in Insee (éd.), *Emploi et salaires*.

Picart, C., (2017), « Trois segments pour mieux décrire le marché du travail », in Insee (éd.), *Emploi et salaires*.

Rey, L., Quesnel, J. S., Sauvain, V. (éd.), *L'évaluation en contexte de développement : Enjeux, approches et pratiques*. Montréal, École nationale d'administration publique, Les éditions JFD.

Schön, D. A., (1983), *The reflective practitioner: how professionals think in action*. New York, Basic Books.

Schön, D. A., (1992), « Designing as reflective conversation with the materials of design situation », *Knowledge-Based Systems*, vol. 5, n°1, p. 3-14.

Simha, J., (2015), *Les expérimentations sociales en France : une sociologie de l'évaluation des politiques publiques*, Thèse de doctorat, Université de Paris Ouest Nanterre.

Spenlehauer, V., (1998), *L'évaluation des politiques publiques, avatar de la planification*, Thèse de doctorat, Université Pierre-Mendès-France, Grenoble.

Spenlehauer, V., (2004), « Pour une déconstruction des légendes sur les rapports Etat/sciences sociales », in Zimmermann, B. (éd.), *Les sciences sociales à l'épreuve de l'action. Le savant, le politique et l'Europe*, Paris, Éditions de la Maison des sciences de l'homme, p. 119-144.

Stern, E., Stame, N., Mayne, J., Forss, K., Davies, R., Befani, B., (2012), « Broadening the range of designs and methods for impact evaluations », Department for International Development, Working Paper, n°38, avril.

Supiot, A., (2010), *L'esprit de Philadelphie, la justice sociale face au marché total*. Paris, Éditions du Seuil, 184 p.

Swisher, A. K., (2010), « Practice-Based Evidence », *Cardiopulmonary Physical Therapy Journal*, vol. 21, n°2, p. 4.

Vatin, F., (2013), « Valuation as Evaluating and Valorizing », *Valuation Studies*, vol.°1, n°1, p. 31-50.

Viveret, P., (1989), *L'évaluation des politiques et des actions publiques – Propositions en vue de l'évaluation du Revenu minimum d'insertion*, Rapport au Premier ministre, Paris, La Documentation française.

Vivès, C., Sigalo Santos, L., Pillon, J.-M., Dubois, V., Clouet, H., (2023), *Chômeurs, vos papiers !, Raisons d'agir*.

Warin, P., (1999), « Les 'ressortissants' dans les analyses des politiques publiques », *Revue française de sciences politiques*, vol. 49, n°1, p. 103-121.

Weiss, C. H., (1997), « Theory-based Evaluation: Past, Present, and Future », *New directions for evaluation*, n°76, p. 41-55.

White, H., (2009), *L'évaluation d'impact basée sur la théorie : Principes et pratique*. Wageningen University & Research.