



Société Française de l'Évaluation



PLACE DES APPROCHES QUANTITATIVES DANS LES EVALUATIONS & DECISION PUBLIQUES

*Quand le quantificatif prime, le qualitatif
s'impose*

Synthèse de la journée d'étude du

18 décembre 2015

**SFE, Centre de recherche en économie de
Grenoble,**

**Coordination : Valérie Fargeon, Bruno Lamotte, Fiona Ottaviani, Univ.
Grenoble Alpes, CREG, 38000 Grenoble, France.**



Table des matières

Remerciements	3
Avant-Propos.....	4
Propos introductifs - Anne Le Roy.....	5
1ère session « Palmarès et classement : place et rôle dans l'évaluation »	8
1. Intervention de Gabriel ROMPRE : « Évaluations quantitatives de l'éducation : La problématique de la traduction en politiques publiques efficaces ».....	9
2. Intervention de Cyril RISK : « Évaluation en matière de délinquance : Les conséquences d'un mauvais choix d'indicateurs »	11
3. Débat	13
2è session « Que disent et ne disent pas les chiffres, et le "bon" usage du chiffre. ».....	16
1. Intervention de J. R. BRUNETIERE : « Eclairer la gouvernance par les nombres : de la distance entre les indicateurs usités et les finalités de l'action publique »	16
2. Intervention d'Hélène Clot : « Les enjeux de la construction d'indicateurs de bien-être soutenable territorialisés dans la métropole grenobloise »	19
3. Intervention de Bernard Perret : « L'importance des indicateurs aujourd'hui : raisons et mise en perspective »	20
4. Débat	21
3è Session : LES APPROCHES QUANTITATIVES ET QUALITATIVES DANS LES POLITIQUES DE MAITRISE BUDGÉTAIRE	24
1. Intervention de A. Migus, « L'évaluation des politiques publiques à la Cour des Comptes » ..	24
2. Intervention de Hélène Paris, L'évaluation des politiques publiques au conseil d'analyse économique.....	26
3. Débat	28
Conclusion : Pierre Savignat, Président de la société française de l'évaluation	32
Bibliographie.....	34

Remerciements

Des compte rendus des ateliers ont été réalisés par les étudiants du master Evaluation et Management de politiques sociales de l'Université de Grenoble Alpes.

Atelier 1 : Anne Daudin, Isabelle Planat, Vincent Sindirian, Mohamed Fadiga.

Atelier 2 : Aurélie Perrière, Marlène Coulon, Nathalie Molostoff, Aurélie Costanzo.

Atelier 3 Claire Flascher, Emmanuelle Rivoire, Stéphanie Capizzi, Justine Houssin, Tilate Kombate.

Les organisateurs tiennent à les remercier pour leur travail.

Avant-Propos

Le master 2 « Sciences sociales », parcours « Evaluation et management des politiques sociales », Faculté d'économie de Grenoble (Université de Grenoble Alpes), le Centre de recherche en économie de Grenoble, et la Société Française de l'Evaluation ont établi en 2015 une convention de coopération.

Cette convention instaure un partenariat autour de stages de master 2, une journée nationale d'étude de la SFE à Grenoble. En 2015, celle-ci a porté sur la mesure, la portée et les limites des données quantitatives et leurs articulations avec les méthodes qualitatives. Cette journée a donné lieu à ce document de synthèse. L'ensemble de la promotion, les chercheurs membres du Creg et l'équipe pédagogique du master y ont contribué.

La coopération entre le Creg et la SFE est centrée sur une réflexion concernant la méthodologie de l'évaluation. Une attention particulière est accordée à l'articulation des méthodologies quantitatives et qualitatives dans la réflexion évaluative. On peut en effet s'accorder facilement pour dire que ces méthodologies sont complémentaires. Il est plus difficile de s'entendre sur le mode d'articulation concret des approches et l'élaboration d'une méthodologie intégrée. Il ne s'agit pas seulement de compléter les zones d'ombre d'une approche quantitative par des apports qualitatifs, ou inversement, il s'agit de définir une méthodologie articulant différentes approches dans une démarche améliorant plus profondément les méthodes dans leur unité.

Cette journée d'étude, sur ce thème qui nécessite une réflexion approfondie, sera renouvelée annuellement.

Propos introductifs - Anne Le Roy

Face à l'importance prise par les approches quantitatives dans les évaluations et les décisions publiques, l'idée d'organiser une journée d'étude dédiée au quantitatif a émergé autour de l'hypothèse suivante : « quand le quantitatif prime, le qualitatif s'impose ».

Longtemps, l'évaluation était l'affaire de politologues et de sociologues dont les productions étaient faiblement dotées, voire dépourvues, de données chiffrées. Puis la contrainte budgétaire et la mise en œuvre de la Loi Organique relative aux Lois de Finance vont faire du chiffre une pièce maîtresse des rapports d'évaluation, conduisant les économistes à s'intéresser à l'évaluation en utilisant et en produisant des données chiffrées. De sorte qu'aujourd'hui, toutes les évaluations contiennent des données quantitatives.

Si le recours aux chiffres semble assez « naturel », toute connaissance tendant à s'objectiver et à se préciser dans une mesure chiffrée. Pour autant, il ne s'agit pas de se soumettre au diktat du résultat, à la dictature du chiffre largement encouragée par les dérives d'une vision de l'économie fondée sur l'hypothèse de rationalité et réduisant la réalité à une équation ou un modèle mathématique. La question n'est donc pas « j'aime ou je n'aime pas le chiffre », on peut difficilement s'en passer, mais devient comment « compter ce qui compte afin de conter ce qui compte », afin d'éviter la (dé)mesure de la mesure ?

Aborder un tel questionnement conduit à s'intéresser (1) aux raisons qui poussent à recourir à des indicateurs pour évaluer et décider, ainsi qu'à l'apport d'un tel recours (2). L'objectif est, en effet, de recourir à des indicateurs utiles à l'analyse, et non pour faire « bien » ou parce qu'ils seraient en apparence les garants d'un travail sérieux, rigoureux, voire scientifique.

1. Recourir à un chiffre : pourquoi et à quelles conditions ?

Pourquoi ?

Recourir à un chiffre, c'est recourir à une transformation de la réalité mise sous forme de statistiques. En effet, reprenant le vocable d'Alain Desrosière, un indicateur est une construction sociale représentant, de façon plus ou moins consensuelle, la « réalité » : ils ne font donc pas que les refléter mais la transforment. Qu'elles soient brutes et/ou élaborées, ces données, fruits d'une convention socio-économique, permettent de répondre *a minima* à deux besoins analytiques en :

- Informant *via* une description de la réalité, à partir de la statistique descriptive ou de l'analyse de données.

- Quantifiant un effet ou une cause, confirmant ou infirmant des intuitions *via* des modèles économétriques.

Le chiffre n'est donc pas doté d'une seule fonction, d'un seul usage : il est au contraire multifonctionnel. Aussi tout utilisateur de données quantifiées devrait être amené à préciser et expliciter :

- La fonction analytique qu'il souhaite activer afin que les données sélectionnées et utilisées soient en phase avec les fonctions recherchées.
- La posture adoptée, c'est à dire le processus de quantification mobilisé, afin d'être en capacité de révéler l'information contenue dans le chiffre.

En mobilisant ainsi les données, le risque de mésusage du chiffre se réduit : on lui fait dire ce qu'il peut dire et ce pourquoi il a été construit, et non ce que l'on veut qu'il dise. Dans cette optique, le plus célèbre de nos indicateurs macroéconomiques, le PIB, ayant été créé pour mesurer prioritairement la richesse marchande monétaire, il ne peut mesurer qu'une certaine partie de la richesse : celle qui est monétarisable et, en aucun cas le progrès social dans toutes ses dimensions. Aussi, utiliser le PIB comme un indicateur de progrès revient à valider la corrélation entre la richesse monétarisable et le progrès social. Or, si cette relation statistique était réelle dans les années 1950, elle n'est plus d'actualité (Easterlin, 1974 ; Laurent, 2013). De sorte que ce célèbre indicateur ne peut être mobilisé que pour mesurer un volume (ou montant) de production et rien d'autre !

Comprendre pourquoi et comment les indicateurs sont construits s'avèrent alors être déterminant pour permettre un usage avisé du chiffre.

Maîtriser la construction

Comment éviter la (dé)mesure du chiffre ? Comment développer un usage avisé de la donnée quantifiée ? L'objectif étant de faire en sorte que l'omniprésence du chiffre, en lien avec la rationalisation des politiques, la pratique du *benchmarking* et la culture du résultat qui favorisent la multiplication des indicateurs, ne devienne pas un grand n'importe quoi...

En effet, cette profusion du chiffre présente certains risques, notamment celui d'un recours comme simple habillage, comme outil de communication ou de marketing destiné à attirer des financements. Or, ce risque peut s'accroître.... Le rôle de l'État et des collectivités évoluant, les demandes de données chiffrées changent et gagnent en importance. La maîtrise de leur construction devient capitale, un nouvel enjeu, car c'est un moyen de réduire le risque de mésusage du chiffre.

Aussi, connaître les modes de construction et donc les caractéristiques des chiffres, leur « carte d'identité », permet de mettre en exergue ce qu'ils disent et ne disent pas ; réduisant par la même le risque de mésusage ou de détournement de leur usage conscient de leurs apports et limites.

2. Apports & limites du chiffre

Usage multiple et finalité explicite

On est face à une pluralité d'usages du chiffre : il peut être utilisé à des fins de controverses (rapport Stern, 2006), d'interpellation (rapport sur la dette publique de 2005), de réquisitoire ou plaidoyer (Conseil régional du Nord-Pas-de-Calais), d'appui au diagnostic local (expériences locales d'indicateurs alternatifs), d'outil de communication (rapport sur les 10 indicateurs de richesse, France Stratégie, 2015).

L'existence de cette pluralité d'usages exige qu'on ait en tête la finalité du recours au chiffre pour être en capacité d'adapter sa construction qui devient alors un processus, et non une procédure, répondant à un besoin analytique explicite.

Ainsi, en adoptant cette posture, une démarche d'évaluation pourrait conduire à encourager des « modalités » de production de chiffres permettant de rapprocher davantage évaluation, production de données et analyse de l'intervention politique. Les données chiffrées ainsi construites participeraient à la construction d'un argumentaire étayé, susceptible de créer une connaissance opérationnelle et théorique tout en permettant d'accroître l'utilité des évaluations ainsi conduites.

Des conditions de production et d'utilisation en phase avec la finalité retenue

Les conditions de production et d'utilisation de données chiffrées apparaissent importantes afin que leur mobilisation dans l'évaluation des politiques :

- ne se résume pas à une simple mesure d'objectifs atteints, d'impacts réalisés ;
- mais contribue réellement à l'amélioration de la connaissance des politiques publiques et des mécanismes sous-jacents

La question n'est donc pas chiffre ou pas chiffre, mais comment favoriser :

- L'articulation d'approches quantitatives et qualitatives qui apparaît comme l'un des enjeux méthodologiques forts pour assurer la pertinence et qualité des évaluations.
- Le développement de l'éducation statistique des citoyens afin de développer une culture statistique conduisant à toujours à regarder la source et la façon dont les chiffres sont construits.

Autrement dit, ce n'est pas de l'utilité du chiffre dont il est question durant cette journée, mais de sa construction et de son interprétation ; même si reprenant les propos de Jean Viard (2014), « il faudrait penser l'évolution de la conception de la statistique publique de la même façon que l'évolution du droit face aux métamorphoses sociales (pauvreté, chômage, femme seule...) ».

Et si indéniablement, il importe de faire évoluer les indicateurs, comme on l'a fait avec le PIB puis les indicateurs de développement durable, encore faut-il les utiliser (l'équipe IDDRIS, *Le Monde*, octobre 2015).

Pour conclure ces propos introductifs et introduire la journée dédiée à la place des approches quantitatives, je reprendrai les propos de Jean Paul DELEVOYE, président du CESE prononcé lors d'un colloque organisé en janvier 2014, posant la question suivante : « est-ce que la dictature des chiffres n'est pas un refuge qui impose des choix, qui met en place des mécanismes d'exclusion plutôt que des mécanismes de réflexion ? »

Afin d'alimenter cette réflexion, trois tables rondes sont proposées : (1) « **Palmarès et classement : place et rôle dans l'évaluation** », (2) « **Que disent et ne disent pas les chiffres** » et (3) « **Les approches quantitatives et qualitatives dans les politiques de maîtrise budgétaire** ».

1ère session « Palmarès et classement : place et rôle dans l'évaluation »

Cette première session a été présidée par Bruno LAMOTTE, CREG UPMF, Directeur du Master 2 « Évaluation et Management des Politiques Sociales ». Les intervenants de cette session étaient Gabriel ROMPRE, consultant Unesco, expert dans le domaine de l'éducation et Cyril RISK, responsable des statistiques de l'observatoire de la délinquance et des réponses pénales. Leurs présentations avaient pour objectif d'analyser le rôle des méthodes quantitatives dans l'évaluation des politiques publiques, en se basant sur deux cas l'un dans l'éducation et l'autre concernant la délinquance.

A l'entame de cette session, Bruno LAMOTTE introduit la question des palmarès et des classements et de leur place privilégiée dans l'évaluation. Pour situer le débat, il rappelle l'importance aujourd'hui du classement comme celui dit *de Shanghai* pour les universités. Puis il nous précise que cet indicateur synthétique était, et reste, un outil de pilotage pour des étudiants plutôt scientifiques (chimistes) de Shanghai. La prudence devrait être de mise quant à l'utilisation faite de ces palmarès. Il exprime une certaine perplexité vis-à-vis de l'utilisation de ces indicateurs en citant l'exemple des enquêtes PISA de l'OCDE, qui pour lui très écoutées sur le thème du classement des systèmes éducatifs : L'enquête PISA étant paradoxalement plus prisée par bon nombre d'Etats que celles beaucoup plus complètes réalisées par l'OCDE dans ses « Regards sur l'éducation ». Finalement, Bruno LAMOTTE explique que cette question sera abordée à travers deux champs d'investigations principaux ; le premier présenté par Gabriel ROMPRE sur « **les évaluations quantitatives de l'éducation : La problématique de la traduction en politiques publiques efficaces** », la seconde intervention de Cyril RISK portant elle sur « **l'évaluation en matière de délinquance : Les conséquences d'un mauvais choix d'indicateurs** ».

1. Intervention de Gabriel ROMPRE : « Évaluations quantitatives de l'éducation : La problématique de la traduction en politiques publiques efficaces »

Gabriel ROMPRE a rappelé que le plus souvent le domaine de l'éducation est très sensible à la question d'évaluation. Il explique que l'évaluation de l'éducation repose sur un instrument métrique en apparence simple : le résultat des élèves dans les classes. Puis, de l'évaluation directe des élèves, on peut très facilement réaliser une évaluation multi-niveaux, en changeant de point focal. Ainsi, à partir du résultat des élèves on évalue les enseignants, à partir du résultat des élèves et des enseignants on évalue les écoles, et ensuite jusqu'au système global de l'éducation nationale et régionale. Cet ordonnancement est commun et ancien dans le monde de l'éducation.

Selon cet expert canadien, on peut voir une réelle émergence de la compétition scolaire actuelle. Avant, on allait à l'école de son quartier au sein de son environnement social, sans se poser de question alors qu'aujourd'hui et dans de nombreux pays, on choisit son école. Un phénomène social sous-tend cette évolution, l'importance du classement d'un établissement sur la reproduction sociale de la classe moyenne. Avec la baisse du pouvoir d'achat des classes moyennes après les années 1970, la question de l'acquisition d'un « bon » diplôme dans les « bonnes » institutions devient cruciale. On rentre donc dans un système de quasi « marché » (même si en France le système vise à contraindre le choix des familles en liant le choix de l'école au lieu de résidence). Ce marché introduit une réelle pression sur les pouvoirs publics pour faciliter aux parents un « choix éclairé » sur ce sujet de plus en plus central, donc en faveur de la création d'indicateurs objectifs pour le permettre.

Gabriel ROMPRE déclare qu'historiquement le monde de l'éducation est l'un des premiers secteurs qui a fait l'objet de grandes études. C'est donc un secteur où l'usage du chiffre se fait depuis longtemps, en réalité depuis les années 1960. Dès ses débuts cet usage est extrêmement controversé. La première grande étude anglo-saxonne sur l'éducation est le rapport Coleman (1965). Elle montrait que la question de l'inégalité socioculturelle est bien plus déterminante que la manière dont est gérée l'école. Ce rapport devient séminal pour toutes les études qui vont avoir un lien avec l'éducation. Il déclare que la gestion des écoles apporte une différence quasiment infinitésimale sur les performances des élèves.

Cette vision extrêmement « socialisante » est également démoralisante pour les enseignants et va évoluer dans les années 1980 jusqu'à aujourd'hui, grâce à d'autres études qui vont se focaliser sur une approche plutôt managériale en prônant une « *school-effectiveness* », et la recherche des « *bonnes pratiques* » à travers la comparaison et la création d'indicateurs rigoureux. L'absence de prise en compte de l'origine socioéconomique et ethnoculturelle va dès lors devenir plus la norme que l'exception ce qui d'un point de vue sociologique est absurde.

Aussi, l'évaluation des systèmes d'éducation s'est diffusée au niveau international. Dans l'enseignement secondaire on va avoir trois grandes études internationales¹ dont la plus célèbre est le classement « PISA » depuis les années 2000. Cette tentative d'uniformisation et de synthèse d'éléments était à la fois extrêmement complexe (la performance scolaire) et difficile à uniformiser (la spécificité de chaque système d'éducation). Elle concerne 74 pays avec un échantillon de 525000 participants. L'enquête PISA avec cette comparaison internationale, a un impact concret sur les politiques publiques nationales car elle attire l'attention sur les carences des systèmes éducatifs, aide à l'identification des meilleures pratiques. Ces mesures ont un écho médiatique énorme. L'utilisation qui en est faite offre des avantages mais aussi des risques certains.

D'une part, il permettra d'attirer l'attention et donc de :

- Poser débat sur les carences de certains systèmes d'éducation,
- D'aider à l'identification des bonnes pratiques même si c'est une évaluation pondérée autour de trois grands domaines (lecture, mathématiques et sciences) et qui n'explique pas le « pourquoi » de la performance d'un système face à un autre,
- De suivre l'impact d'une réforme dans le temps de façon objective.

Mais d'autre part, Gabriel ROMPRE montre qu'il peut y avoir certains biais avec cette enquête PISA. Les medias s'en saisissent souvent à des fins spectaculaires. C'est le problème des indices composites, seuls mis en avant en occultant la centaine de pages qui les détaillent. Enfin le risque important de manipulation de ces chiffres a été rappelé, avec par exemple le cas de nos voisins d'outre-manche (Une chute de performance au Royaume Uni dans les années 2000, très médiatisée mais liée à des résultats peu fiables en raison d'une divergence de taux de réponse)².

Gabriel ROMPRE termina son exposé, en prenant quelques exemples de controverse sur des classements d'universités³, expliquant que ces classements sont à relativiser. Attention au choc de l'information, à la récupération politique, aux limites de ces méthodes qui ne sont pas longitudinales. La question est bien de déterminer sur quoi l'on met le curseur dans ces classements. Face à ces problèmes récurrents des critères, et pour éviter de passer du « *teaching for the student* » à « *teaching for the test* », le plus important est de savoir ce que l'on mesure, dans quel but on le fait, en plus d'avoir une modestie dans l'utilisation des résultats.

¹Le "Program for International Student Assessment" (PISA), mais aussi le "Trends in International Mathematics and Science Study" (TIMSS), et finalement le "Progress in International Reading Literacy Study" (PIRLS).

² Titre du 3 décembre 2013 sur *Channel News 4*, « *Low marks: report shows UK education is lagging behind* ».

³ Gabriel Rompré a pris deux exemples : l'Université d'Alexandrie et l'Ecole d'économie de Paris. Dans le premier, un établissement situé mondialement comme 147^e en 2010 (*Times Higher Education Supplément*). Or un seul professeur aurait été responsable de cette hausse en publiant 320 articles dans sa propre revue scientifique. Dans le cas de Paris, PSE est une université virtuelle qui regroupe en fait plusieurs établissements d'enseignement. La création d'une certification commune a fait augmenter leur classement sans qu'il n'y ait eu de modification dans les ressources ou l'enseignement.

2. Intervention de Cyril RISK : « Évaluation en matière de délinquance : Les conséquences d'un mauvais choix d'indicateurs »

Lors de son intervention, Cyril RISK aborde la question de l'évaluation au travers d'exemples concernant la mesure de la délinquance. Faisant une transition avec l'exposé de Gabriel ROMPRE sur l'éducation il entame son intervention par un exemple édifiant d'enquête menée sur les incidents au sein des établissements de l'éducation nationale. Le journal « Le Point » s'était saisi de ces données via la CADA (commission d'accès aux données administratives) et avait publié un classement des établissements scolaires en fonction du nombre d'actes de violences enregistrés. Or les établissements n'ayant pas tous eu la même rigueur pour remplir cette enquête, ceux pointés du doigt à travers l'article étaient ceux qui enregistraient de la manière la plus exhaustive les incidents et non pas ceux où il y avait réellement les plus grandes difficultés. La conséquence de cette utilisation médiatique est que les établissements ont arrêté de remplir l'outil, car ceux qui le faisaient étaient stigmatisés par rapport aux autres.

Un mauvais choix d'indicateurs aura des conséquences négatives sur l'utilité et la qualité de l'évaluation.

Les classements en matière de délinquance sont très nombreux et réguliers. Cyril Risk présente l'observatoire qui a été créé en 2003, remplissant une mission de service public d'information générale, et intervenant dans le champ médiatique-politique. Les médias eux-mêmes font des palmarès.

Au sujet de la fabrique des chiffres, il souligne la nécessité de prendre les chiffres avec réserve. En effet, l'outil statistique de la gendarmerie et de la Police en France a la particularité de ne pas associer les faits au lieu où ils sont commis mais au lieu où ils sont enregistrés ce qui fausse la réalité. Ils ne rendent pas compte de l'évolution de la délinquance. C'est un premier problème.

Il explique les évolutions méthodologiques en cours avec l'exemple de l'utilisation de l'enquête de victimation aux Etats Unis, une enquête directe auprès des victimes. Il s'agit d'une technique statistique, qui va permettre à un moment donné de comprendre la constitution des chiffres et de les pondérer. En effet, on ne peut pas s'appuyer sur les chiffres des plaintes enregistrées par la Police pour mesurer l'évolution de la délinquance si la plupart des victimes ne dépose pas plainte. On va interroger des victimes pour étayer les chiffres de données contextuelles. Ces enquêtes de victimation, non seulement vont fournir des taux qui vont permettre de mesurer la fréquence des phénomènes au sein de la population, éventuellement de mesurer l'évolution de ces phénomènes dans le temps, mais aussi ils vont proposer une estimation de la proportion des victimes à déposer plainte. Cela permettra d'avoir un chiffre de meilleure qualité. C'est un constat qui a été fait dans les années 1960 aux Etats Unis et il a débouché sur un système annuel de victimation, considéré comme un bon outil.

Deuxième problème : L'intérêt de la presse pour les chiffres et les palmarès crée un risque réel d'instrumentalisation de ceux-ci. Un travail pédagogique est à faire autour de leurs usages. Le ministère de l'intérieur en tant qu'acteur principal peut lui-même se retrouver au

cœur de cette instrumentalisation pour justifier d'une politique, il peut ainsi se produire les phénomènes suivants :

- Manipulation des enregistrements pour obtenir artificiellement de bons résultats à l'endroit voulu (exemple des homicides curieusement enregistrés juste à la périphérie de la ville de Chicago).
- Management par le chiffre sur les préfets.
- Dissimulation de faits par le mode d'enregistrement : plutôt en « contravention » qu'en « délit » par exemple ce qui fait baisser artificiellement les chiffres de la délinquance.

En France, le Rapport Caresche-Pandraud de 2002, à l'origine de la création de l'observatoire de la délinquance (ONDRP en 2004) aborde cette question stratégique de la mesure de la délinquance et du débat public :

« La mesure de la délinquance en France fait à l'heure actuelle l'objet de vives polémiques. Utilisés dans le débat sur l'insécurité et les politiques publiques concourant à la lutte contre la délinquance, les chiffres produits annuellement par le ministère de l'Intérieur sont, soit critiqués pour leur manque de fiabilité, soit repris par les acteurs politiques et les médias selon des optiques qui permettent toutes les interprétations. »

Ce rapport pointait déjà du doigt les usages du chiffre de la délinquance dans le débat public permettant toutes les interprétations. On avait une situation d'insatisfaction des usages du chiffre sur la délinquance, produit annuellement par le ministère de l'Intérieur, l'indicateur existant étant détourné de son usage, et n'apportant finalement jamais la réponse à la question posée. L'observatoire a été amené à publier un bulletin mensuel depuis 2006 pour dénoncer les dérives de la mauvaise manipulation des chiffres : les politiques publiques de lutte contre les violences existent et l'un des indicateurs de réussite de ces politiques est justement l'augmentation des plaintes.

Il souligne le rôle joué par les associations de victimes pour faire évoluer la législation et la nécessité de dépasser d'apparentes contradictions. Par exemple, la législation évoluant dans le sens la politique publique de lutte contre les violences faites aux femmes, a fait augmenter le nombre de dépôts de plaintes. On pourrait considérer à tort que l'augmentation des dépôts de plaintes révèle une augmentation des violences signant ainsi l'échec de la politique publique. Alors qu'il s'agit d'un élargissement des « comportements violents » qui sont considérés comme répréhensibles et qualifiées avec des circonstances aggravantes.

Pour finir, il exprime un sentiment d'espoir dans l'élaboration des chiffres, qui est le « data journalisme », c'est le fait que ce champ médiatico-politique de l'utilisation des statistiques sur la délinquance a vu un nouvel acteur, les journalistes, qui vont décrypter l'usage des chiffres. Le bulletin mensuel d'information finit ainsi par porter ses fruits.

3. Débat

La session s'est poursuivie par une série de question, de témoignage et d'échanges avec la salle :

- 1) Jean Pierre Jeantheau chargé de statistique à l'agence nationale de lutte contre l'illettrisme et ex projet manager pour PISA 2000 pour la France et ex responsable de l'évaluation et de la prospective au Ministère de l'Education nationale :

Il se demande si l'on se pose les bonnes questions concernant les enquêtes de victimation. En modifiant son comportement on influence les statistiques ; par exemple concernant le problème des femmes qui se faisaient voler leurs colliers à l'arrachée dans les rues de Marseille, on constate une baisse des dépôts de plainte car les victimes potentielles ont cessé de porter leur bijou dans la rue. L'enquête de victimation n'a dès lors plus de sens.

Dans l'éducation où les personnes qui vont être impliquées dans les résultats des enquêtes sont eux-mêmes dans la construction des indicateurs : une dérive consiste à pousser les enseignants à préparer les élèves à passer les tests PISA.

Pour lui dans l'éducation (avec PISA), on n'est pas sur des mesures métriques, mais sur de l'ordinal. Les personnes sont classées les unes par rapport aux autres localement et non pas en fonction d'un niveau global national.

Autre exemple pour relativiser le classement PISA : au Luxembourg les tests sont passés dans deux langues. Cela n'est pas valorisé par le classement et pourtant ces élèves bilingues ou trilingues seront bien sûr extrêmement recherchés lorsqu'ils arriveront sur le marché de l'emploi.

- 2) Henri Jacot économiste a été élu en charge de l'évaluation des politiques Rhône-Alpes, est une personnalité qualifiée au CNFPT pour la diffusion de l'évaluation, également membre du Conseil de développement de la METROPOLE de Lyon.

Sa question était orientée sur les registres d'évaluation, en se demandant comment faire pour chiffrer les registres tels que « efficacité, efficience, utilité sociale » ? La question du chiffre se pose différemment selon les registres sur lesquels on se place.

Il existe selon lui un débat central et crucial sur le rapport efficacité/efficience, on est alors dans un triptyque : objectifs moyens résultats. Quel rapport de chiffres entre ces 3 éléments ? Puis Henri Jacot cite l'exemple de l'évaluation de l'insertion pour laquelle on a l'habitude de mesurer les « sorties positives », c'est à dire les emplois de plus de 6 mois, les formations, ... Mais cela ne suffit pas car l'insertion représente aussi une utilité sociale importante. Henri Jacot pose les questions suivantes : Comment mesure-t-on l'utilité sociale de l'insertion ? Comment la chiffrer ? Comment l'institution de la métropole de Lyon, comptable de l'utilisation des fonds publics, tient compte de l'utilité sociale ?

- 3) Hicham Bénichi du CREG

Hicham Bénichi affirme que la compétition entre approche quantitative et qualitative n'est pas pertinente. Il fait un parallèle avec les propos d'A. Sen et son intégration de l'indice de développement humain car cela renvoie à la même idée de co-construction du chiffre. Comment intégrer des formes de vies avec un seul indice ? Quelle résolution holiste par la démocratie ? Comment sélectionner les chiffres d'une façon holiste ? Comment sélectionner les données brutes de la réalité ? Quelle relation entre l'expert et les citoyens ?

Selon Hicham Bénichi, l'importation des processus d'évaluation propre à l'entreprise dans le domaine de l'éducation est un principe qui soulève la question de l'application de méthodes venant du privé au secteur de l'éducation. Le rapport le professeur et ses étudiants est un aspect important à prendre en compte, et comment des agrégations sont-elles possibles pour définir des politiques éducatives ? D'après cet intervenant, on est actuellement dans la logique suivante : « Comment équiper les personnes pour intégrer le marché ? » Tout, selon lui, est construit autour de cette analyse.

Hicham Bénichi a terminé son propos en ajoutant que les critères de la question de la justice sociale sont effacés depuis l'érosion de la classe moyenne et que l'on est dans une logique néo-libérale au niveau de l'école.

4) François Journet secrétaire général CEFEDM Rhône-Alpes

François Journet a souligné le risque de l'immédiateté dans le champ politique de l'utilisation des chiffres. Il pose une première question à Cyril RISK lui demandant ce qui permet à l'Observatoire de la délinquance de résister à cette tendance d'instrumentalisation des chiffres et ce qui lui garantit l'autonomie ?

Sa seconde question est destinée aux deux intervenants. Comment utiliser ces chiffres autrement que comme dans une frange de l'immédiateté, en évitant en particulier le classement. Il pose la question de savoir s'il existe des relations entre l'Observatoire de la délinquance et les professionnels de la Police pour utiliser les chiffres de façon plus réfléchie et constructive.

5) Florent Bédécarrats, Agence française de développement

Florent Bédécarrats a fait remarquer que les travers mentionnés pour le quantitatif peuvent être aussi associés au qualitatif. De plus il pose la question « Quels mécanismes mettre en place pour contrer cette frénésie du chiffre ? » Le problème étant posé depuis longtemps selon lui.

Cyril Risk prend la parole le premier pour répondre aux questions.

Revenant sur l'exemple de l'arrachage de colliers, Cyril Risk confirme que les victimes s'adaptent, et ajoute que les systèmes de mesure doivent en tenir compte, en particulier le questionnaire réalisé avec l'INSEE doit évoluer avec les pratiques.

Puis Cyril Risk explique que dans les rapports soumis aux médias, le vocabulaire est très précis par exemple : « victime déclarée » plutôt que « victimes » en général et que c'est une

réponse à une question précise qui contribue à la construction du chiffre. Mais il ajoute que souvent ces précisions « sautent » dans le débat public.

Concernant les relations entre l'Observatoire et les professionnels de la Police, Cyril Risk affirme qu'elles sont assez limitées. Il explique que l'Observatoire essaie de retirer la pression réelle que les chiffres produisent sur les professionnels de la Police et que le Ministère de l'Intérieur et l'Observatoire commentent chacun de leur côté les chiffres. Il ajoute que cette indépendance ne garantit pas l'efficacité.

Pour conclure Gabriel ROMPRE cite l'exemple d'une grande politique éducative américaine qui a dû être abandonnée il y a quelques semaines non pas par manque d'efficacité, mais par perte de contrôle des pouvoirs publics sur les indicateurs, pouvant mener à des absurdités !

2^e session « Que disent et ne disent pas les chiffres, et le "bon" usage du chiffre. »

Cette séance a été présidée par Fiona OTTAVIANI, Docteure en économie et membre associé du CREG, auteure d'une thèse sur la performativité des indicateurs et la construction d'indicateurs alternatifs. Cette seconde session est consacrée à la question de la construction et de l'usage des indicateurs dans les politiques publiques. Alimentant une telle réflexion, l'intervention de Jean-René BRUNETIERE, de la Chaire Economie et Climat à Paris Dauphine, met au jour au travers de nombreux exemples issus de l'examen de la LOLF (Brunetière, 2010), la « distance » entre les indicateurs servant à évaluer les politiques publiques et les finalités poursuivies. Cette réflexion sur la gouvernance par les nombres révèle les nombreux enjeux associés à la construction d'indicateurs et les insuffisances des indicateurs actuellement usités au niveau des politiques étatiques. La seconde intervention, est assurée par Hélène CLOT, chargée de l'évaluation, de l'observation et de la prospective à la métropole grenobloise, et complète ainsi cette réflexion sur les limites des indicateurs existants en soulignant l'intérêt du développement d'indicateurs alternatifs à l'échelle locale en s'appuyant sur l'exemple d'une expérimentation de construction d'indicateurs de bien-être soutenable territorialisés (IBEST) ayant cours depuis de nombreuses années dans la métropole grenobloise. Bernard PERRET, connu pour son rapport sur les indicateurs sociaux (2002) et actuellement socio-économiste au conseil général de l'environnement et du développement durable, apporte ensuite son éclairage sur les raisons profondes pouvant expliquer la tendance « quantophrénique » de nos sociétés.

En préambule, Fiona OTTAVIANI précise que la session est consacrée au « bon usage du chiffre » et qu'un tel objet de réflexion soulève *a minima* trois grandes questions auxquelles les interventions contribueront à répondre. La première question est relative aux finalités de l'action publique et à ce que l'on souhaite mesurer. La seconde question porte sur ce qu'on peut mesurer, tant d'un point de vue technique qu'éthique (Desrosière, 2008, p. 173-174). La troisième question est afférente au public impliqué dans la construction des indicateurs et à la nécessité de mettre au débat les finalités de l'action publique.

1. Intervention de J. R. BRUNETIERE : « Eclairer la gouvernance par les nombres : de la distance entre les indicateurs usités et les finalités de l'action publique »

J.R. BRUNETIERE commence son intervention en rappelant que ses compétences sur le nombre viennent essentiellement de son implication dans l'association *Pénombr* qui s'intéresse à l'usage des nombres dans le débat social. Au centre des réflexions de l'association et de son intervention se trouve donc la question de la manière dont l'existence des nombres et leurs conformations vont orienter l'action publique.

Si l'usage des indicateurs est contraignant, il s'avère pourtant difficile d'objectiver les situations et de les évaluer sans chiffre. J.R. BRUNETIERE souligne alors, à la suite de

Desrosières (Desrosières et Kott, 2005, p. 2), que toute opération de quantification, engage deux opérations : « convenir » et « mesurer ». Trois couches de conventions lui paraissent devoir être distinguées : 1) les fondements idéologiques, qui renvoient à une manière de penser les finalités de l'action publique ; 2) le choix des objets au centre de la mesure ; 3) les détails d'exécution « dans lequel se cache le diable », puisque peut alors apparaître un écart entre les objectifs poursuivis et les indicateurs choisis pour suivre l'objectif. Cet écart, renvoie, selon les mots de P. GIBERT (Gibert et Alilat, 2012) à la « distance » entre ce que l'on veut évaluer et son proxy (le chiffre que l'on va retenir pour en donner une idée).

Contribuant à la suite de P. GIBERT à cette évaluation des distances, J. R. BRUNETIERE propose une graduation de celles-ci.

- Le premier cas de figure évoqué, renvoie à la situation où l'objectif est lui-même l'indicateur comme pour le taux de recouvrement de l'impôt ou le nombre de tués sur les routes. Ce cas est le plus satisfaisant, puisqu'il n'y a pas ici de distance entre le phénomène et l'indicateur.
- Le deuxième cas de figure est celui où l'indicateur est censé contribuer à l'objectif. Par exemple, on a un objectif de pureté de l'eau et on va mesurer le nombre de stations d'épuration construites. Cela suppose une mini théorie de l'action en disant que les stations d'épuration contribuent à la pureté des eaux.
- Troisième cas de figure : l'indicateur a vocation à refléter l'objectif, mais une distance demeure. C'est par exemple le cas des chiffres de la délinquance (*cf.* Session 1).
- Quatrième cas de figure : l'indicateur utilisé est simplement une vague illustration de l'objectif. A titre d'illustration, dans la LOLF (Brunetière, 2010), on a un objectif de rayonnement de la France dans le monde, associé à l'objectif de l'usage du français. La mesure du rayonnement de la France se résume ainsi au décompte de la proportion de documents de l'UE qui sont écrits en français. L'écart ici entre l'objectif et l'indicateur est très grand.
- Le cinquième cas de figure est celui où l'on n'arrive pas à trouver d'indicateurs. La défense de la France dans le monde constitue un bon exemple de cette difficulté : on peut alors avoir une batterie d'indicateurs, mais sans avoir un indicateur pertinent.

J. R. BRUNETIERE précise que la distance est souvent liée au type d'activités : les distances courtes se trouvent dans les activités simples, et notamment dans les industries où les biens et les services sont standardisés. Lorsque les distances sont moyennes alors le chiffrage va supposer des choix de priorité entre les finalités et donc une implication de différentes parties prenantes. L'éducation constitue un bon exemple de ce type d'arbitrage qu'il va falloir opérer sur les objectifs de la politique. La priorité en termes d'éducation, est-ce l'employabilité ? L'acquisition de connaissances ? La formation des comportements ? Pour certaines activités, comme la diplomatie, la distance est rédhibitoire, c'est-à-dire qu'il est très difficile de trouver des proxys proches des phénomènes au centre de l'évaluation.

J. R. BRUNETIERE met en avant que dans la plupart des activités publiques, on rencontre des gammes de distances assez larges. Il démontre la validité de cette affirmation en mettant

en exergue la distance entre les finalités affichées et les indicateurs retenus pour la politique étrangère de la France dans la loi des finances de 2013 et 2014. Les nombreux exemples sur lesquels il s'appuie permettent de révéler que le travail qui consiste à partir de la finalité pour décliner la chaîne de résultats et arriver à des indicateurs convenables n'a pas été fait. Ces nombreux exemples tendent à montrer que beaucoup d'activités essentielles dans le cadre de la politique étrangère ne sont pas chiffrables (la participation à des activités locales par exemple) et que les indicateurs usités pour ce type de politique sont souvent davantage des indicateurs d'activités que de résultats. Les indicateurs existants, fussent-ils valables au plan opérationnel ou local sont en outre difficilement agrégeables au plan politique national.

A travers cet exemple, J. R. BRUNETIERE contribue à montrer qu'en fonction du domaine d'action publique concerné, la quantification s'avère plus ou moins délicate. Dans le cas des politiques de redistribution, les distances peuvent être courtes si les objectifs sont bien définis. Or, force est de constater que la précision de l'objectif politique n'est pas toujours au rendez-vous et qu'une politique peut répondre simultanément à une pluralité d'objectifs. Dans les activités de régulation, la distance est donc courte dans l'activité elle-même, mais longue dès que l'on s'intéresse aux finalités de l'action. Il est en effet aisé de suivre la mise en œuvre d'un décret d'application d'une loi (distance courte), mais bien plus périlleux d'estimer les effets de la loi (distance longue). Dans les activités de représentation, le suivi par les indicateurs s'avèrent illusoire tant l'écart est grand entre les finalités poursuivies et les indicateurs existants. Pour autant, le projet de chiffrage de telles politiques n'a pas été abandonné. Cette volonté de chiffrage est à lire selon J. R. BRUNETIERE au regard des transformations survenues dans l'organisation de l'action publique. Dans les années 1990, l'administration était censée exécuter la volonté du peuple, le parlement donnait des instructions et des moyens et vérifiait que tout se passait conformément à ce qu'il avait dit. Le débat portait alors essentiellement sur les orientations. Balayant cet univers à dominante juridique, l'idée de l'optimisation ayant accompagné l'usage du chiffre a progressé avec la fin des idéologies et le développement du *New Public Management*. Dans les prestations de services, il a alors été question de monitorer l'efficacité, l'efficacités, etc. Des méthodes d'évaluation (méthodes hédonistes, coûts tutélaires, etc.) se sont alors développées pour évaluer l'inévaluable (coût de la vie humaine, etc.)

De telles évolutions questionnent : Quand peut-on employer le nombre ? Qu'est-ce que cela implique ? Où est-ce que cela nous amène ?

J. R. BRUNETIERE souligne qu'employer le nombre, c'est forcément construire des classes d'équivalence. Cela suppose que les choses soient du domaine du fongible, soit par nature soit par convention. Mesurer des objets nécessite donc de passer outre les spécificités de chaque individu étudié.

En conclusion de son intervention, J. R. BRUNETIERE rappelle ainsi la dimension conventionnelle des indicateurs construits en s'appuyant sur l'exemple emblématique de la

contestation du PIB qui mesure « tout, sauf ce qui fait que la vie vaut la peine d'être vécue » (Kennedy, 1968).

2. Intervention d'Hélène Clot : « Les enjeux de la construction d'indicateurs de bien-être soutenable territorialisés dans la métropole grenobloise »

Dans la lignée de ces réflexions sur la remise en cause du PIB, H. CLOT met en avant que deux présupposés méthodologiques ont été au cœur du projet de construction d'indicateurs alternatifs de bien-être soutenable à l'échelle de la métropole grenobloise : 1) la nécessité d'avoir du chiffre pour peser dans le débat ; 2) la nécessité de mettre les indicateurs en construction à la discussion pour participer à l'émergence d'une nouvelle convention sociopolitique selon les termes de J. GADREY (2006).

Après avoir diffusé le webdoc sur la construction d'indicateurs de bien-être soutenable territorialisés (IBEST) dans l'agglomération grenobloise, H. CLOT souligne que deux principales préoccupations sont au centre de la réflexion actuelle sur cette question de la construction d'indicateurs de bien-être :

- 1) la reproductibilité dans l'espace : cette dernière ne peut être que partielle vu que les indicateurs sont construits sur la base de débat ;
- 2) la reproductibilité dans le temps : le choix de construire la donnée manquante par voie d'enquête est couteux, ce qui soulève simultanément la question de la manière de se procurer à moindre coût (*via* l'INSEE ou l'*open data*) ces données pertinentes et celles de la robustesse des données produites.

Une autre question qui lui paraît importante dans le cadre d'une telle expérimentation de construction d'indicateurs alternatifs est celle de l'appropriation et de l'interprétation de la donnée. Les dimensions (« affirmation de soi et engagement », « accès et recours aux services publics », etc.) proposées dans le tableau de bord du bien-être soutenable sont très transversales et montrent bien que la question de l'interpénétration des actions publiques est importante. Afin que les indicateurs soient utilisés, l'instrumentalisation de ces derniers serait finalement souhaitable, car elle permettrait de rendre visible ces travaux encore trop peu connus. Rendre possible cette appropriation dans le système d'action publique suppose de développer des stratégies pour travailler sur l'enjeu de coresponsabilité au cœur des travaux sur le bien-être du Conseil de l'Europe.

3. Intervention de Bernard Perret : « L'importance des indicateurs aujourd'hui : raisons et mise en perspective »

Pour B. PERRET, les interventions de ce matin montrent bien que ce débat quantitatif/qualitatif ne doit pas être conçu en termes d'opposition, mais qu'il est souhaitable de penser l'articulation entre les deux méthodes.

Pour des raisons sur lesquels il reviendra dans la suite de son intervention, B. PERRET fait valoir que les chiffres sont aujourd'hui incontournables dans les politiques publiques.

Résumant les interventions du matin, il insiste sur le double encastrement des chiffres en amont et en aval. Tout d'abord, en amont, les chiffres n'ont de sens que si on a défini un cadre de description de la réalité, des critères, des seuils des conventions, etc. Ces conventions relèvent en partie d'expertise sociologique et en partie de débats démocratiques. Ensuite, en aval, une fois les indicateurs construits, il y a matière à débat sur leur interprétation pour produire une compréhension de la réalité.

B. PERRET émet ensuite trois observations visant à rendre intelligible les raisons pour lesquelles les chiffres prennent aujourd'hui autant d'importance.

Premièrement, une raison fondamentale, bien formulée par H. SIMON (prix Nobel d'économie) à travers sa théorie de la rationalité limitée, est que l'information coûte chère. Elle coûte cher, non simplement financièrement, mais aussi pour celui qui la reçoit : « Dans un monde où l'attention est une ressource majeure des plus rares, l'information peut être un luxe coûteux car elle peut détourner notre attention de ce qui est important vers ce qui ne l'est pas. Nous ne pouvons pas nous permettre de traiter l'information simplement parce qu'elle est là » (Simon, cité dans Léca, 1993).

Cette volonté de pouvoir disposer d'informations simples mène souvent à des choses très perverses, on voit bien une tendance forte de nos sociétés dans lequel s'inscrivent ce développement des indicateurs. Cette tendance, c'est la rivalité c'est-à-dire le besoin d'être en concurrence non violente les uns avec les autres. Les indicateurs peuvent être vus comme l'aboutissement d'un long processus bien décrit par Norbert Elias : le besoin de pacifier la société et de transformer les rapports violents en rapports non violents. Les indicateurs constituent donc des symptômes du processus de pacification de la société, la concurrence s'effectuant au travers de la mise en comparaison. L'exemple de l'IDH est à cet égard parlant : des délégations de pays viennent depuis la publication de cet indice voir le PNUD non pour améliorer leur politique sociale, mais leur classement dans la hiérarchie des pays crée par l'IDH.

Deuxièmement, B. PERRET soutient qu'il y a des bons et des mauvais indicateurs. Il s'avère notamment nécessaire de développer de nouveaux indicateurs car même s'il y a une surutilisation des chiffres dans certains domaines, coexiste parallèlement un sous-développement des indicateurs dans d'autres, comme sur l'impact carbone.

Troisièmement, la loi E. SAS, votée en avril 2015, fait évoluer les indicateurs au centre du débat public et prévoit l'élaboration d'indicateurs complémentaires au PIB publié à l'échelle nationale. Ces indicateurs ont vocation à être utilisés comme critère d'évaluation des politiques publiques.

En conclusion, B. PERRET rappelle que si l'on réfléchit sur la construction d'indicateurs alternatifs, c'est notamment parce que la croissance n'est plus là. Il déclare que si l'on veut redonner un sens à l'action collective, il va falloir se fixer démocratiquement d'autres critères pour mesurer le progrès de nos sociétés.

4. Débat

Question de Monsieur JACOT

Tout d'abord, Monsieur JACOT adresse sa question à J. R. Brunetière. Il s'interroge sur la question de la monétarisation, qui dans le processus d'agrégation des indicateurs, est décisive. Ensuite, en réaction à l'intervention d'H. CLOT, il rappelle que le prix Nobel 2015 a été donné à Angus DEATON, qui validant le paradoxe d'EASTERLIN (1974), a montré qu'aux Etats-Unis, qu'au-delà de 75000\$, le bonheur diminuait.

Réponse de Monsieur BRUNETIERE:

Au sujet de la monétarisation, J. R. BRUNETIERE indique que le chiffre est fait pour rendre comparable des choses différentes, la monnaie aussi. Il y a donc, selon lui, une tendance naturelle à glisser vers le dollar. Cette tendance est étayée par les travaux d'A. SUPIOT (2015) dans la *Gouvernance par les nombres*. Dans cet ouvrage, ce dernier montre que d'un système où des Etats avec des lois et de la violence légitime encadrent des marchés, des activités privées et des entreprises, nous sommes passés dans un système où les marchés fixent les règles internationales et mettent en concurrence les Etats. L'activité privée englobe quelques acteurs qui sont les Etats qui ont comme mission d'assurer l'intendance, de fabriquer les règles pour que le marché se déroule bien. Or, à la lumière des événements récents, on s'est rendu compte que l'Etat n'était pas un prestataire de service. Alors que l'Etat, suite aux privatisations et aux décentralisations, s'occupe essentiellement des activités régaliennes, l'actualité récente a amené la question de la construction d'une nation ouverte vis-à-vis du monde (accord de Paris, terrorisme). Pour les évaluateurs qui travaillent dans le domaine de l'Etat, comment on peut évaluer ce genre de mission dont il devient de plus en plus évident que c'est la seule activité de l'Etat aujourd'hui ?

Réponse d'Hélène CLOT

H. CLOT répond à la remarque relative au paradoxe d'EASTERLIN en évoquant les résultats de l'étude IBEST. Dans le cadre de cette expérimentation, le bien-être a été apprécié comme une tension entre les aspirations des personnes et les moyens dont elles disposent. Tout d'abord, il ressort de cette analyse que ce sont les étudiants et les jeunes retraités qui sont les plus satisfaits dans tous les domaines. Un tel résultat soulève la question du rapport structurant du travail dans la réalisation du bien-être des personnes. Ensuite, la deuxième catégorie de personnes les plus satisfaites de leur existence, sont ces cadres avec un haut niveau d'étude mais qui sont un peu stressés. La question du temps apparaît ici fortement. Enfin, la troisième catégorie de personnes est celle des ouvriers et des employés qui manquent de reconnaissance. Ces résultats montrent bien que la question du rapport à l'argent n'est pas le premier facteur du bien-être. Des indicateurs jouent par ailleurs plutôt en creux, la santé n'est jamais mentionnée comme un critère de bien-être, sauf quand elle n'est pas là. Ainsi, les gens sont malheureux quand ils n'ont pas d'argent ou la santé. Les relations sociales comptent également beaucoup dans le bien-être. Conséquemment, si la sociabilité est un élément de bien-être, les politiques associatives sont importantes, alors qu'en période de restriction budgétaire on a tendance à se recentrer autour des éléments plus économiques. Ces quelques résultats montrent bien en quoi ces indicateurs permettent de questionner les priorités de l'action publique.

Question de Monsieur JOURNET:

M. JOURNET met en lumière que l'évaluation peut être souvent prise, sous son volet qualitatif ou quantitatif, comme une volonté d'objectiver des choses. Dès lors, il constate que trop souvent l'on va très vite à l'objectivation, en naturalisant les phénomènes étudiés, et on ne se rend pas suffisamment compte qu'il y a des questions qui méritent d'être discutées collectivement. La mesure n'est pas un « en soi », ce qui l'amène à formuler une série de questions : Comment associer un maximum de gens à la réflexion ? Comment réintègre-t-on la dimension de choix de société ?

Question d'Anne Le ROY:

A. LE ROY s'interroge sur l'usage du chiffre et l'usage des évaluations. Comment on pourrait éviter autant de détournement tant autour du chiffre que de l'évaluation, que du lien entre les deux ? Est-ce qu'on n'a pas tous un rôle à jouer pour favoriser une éducation à un bon usage du chiffre ?

Réponse d'Hélène CLOT:

H. CLOT indique que l'on a beaucoup parlé de la production et de l'analyse du chiffre, mais peu de la manière dont on le met en scène. Pour les professionnels, cet aspect est essentiel, car ils passent beaucoup de temps à produire, peu temps à analyser et quasiment aucun temps à

valoriser. Dès lors, H. CLOT s'interroge sur la façon de travailler sur cette appropriation, notamment avec les élus, en ayant bien en tête ce que disait Bernard PERRET sur H. SIMON. Sur ce travail d'appropriation, ils ont exploré plusieurs voies : des quizzes et des comparaisons. Hélène CLOT indique que lorsqu'on se compare, contrairement à ce qui a été dit, il n'y a pas qu'une logique de concurrence qui a cours, mais que cela peut également présenter un intérêt pédagogique.

Réponse de Bernard PERRET:

B. PERRET abonde dans le sens d'H. CLOT en précisant qu'il y a des formes de *benchmarking* qui sont tout à fait vertueuses et qui se font sur des bases volontaires de gens qui veulent simplement progresser ensemble en comparant leurs pratiques. Or, au-delà de cet aspect, le chiffre appuie aussi la mise en compétition.

En réaction aux propos d'H. CLOT sur l'appropriation, B. PERRET souligne que la réflexion épistémologique sur les indicateurs lui paraît essentielle. Il soutient que la philosophie de la connaissance, la capacité de réfléchir sur ce qu'est un chiffre et sur son utilité ne sont pas des questions réservées à ces élites, mais que tout le monde devrait un minimum de capacité de décryptage par rapport à toutes ces questions.

Question d'Aline BLANC-TAILLEUR :

A. BLANC-TAILLEUR fait ressortir les difficultés qu'il y a à impliquer les autres élus dans une expérimentation telle qu'IBEST. Il lui semble qu'il est nécessaire de médiatiser ce type d'initiative pour élargir cette dynamique collective. Sur la question des indicateurs de la délinquance, elle rappelle que ce sujet est très sensible.

Réponse de Fiona OTTAVIANI :

En conclusion, F. OTTAVIANI indique que la discussion avec la salle a permis de mettre en avant : 1) la diversité des usages qui ne sont pas forcément concurrentiels entre les indicateurs ; 2) la nécessité de la communication et du débat autour des indicateurs.

3^e Session : LES APPROCHES QUANTITATIVES ET QUALITATIVES DANS LES POLITIQUES DE MAITRISE BUDGÉTAIRE

Introduction de Valérie FARGEON : Comment concevoir les actions publiques ? Comment ces deux institutions envisagent les politiques publiques ? Elles qui ne produisent pas d'évaluation mais qui éclairent le gouvernement sur des décisions économiques. Quelles sont les méthodes et les expertises ? Quel est le lien avec les décisions publiques ? Et comment s'articulent quantitatif et qualitatif pour le Conseil d'Analyse Economique et la Cour des Comptes?

1. Intervention de A. Migus, « L'évaluation des politiques publiques à la Cour des Comptes »

A. Migus, Conseiller Maître à la Cour des Comptes.

La Cour des Comptes se situe à mi-chemin entre le Parlement et le Gouvernement. Le Parlement a pour rôle le contrôle des actions du Gouvernement et gère les politiques publiques.

Le fonctionnement de la Cour des Comptes est normé par des textes législatifs et réglementaires. Depuis la réforme constitutionnelle de 2007, l'évaluation des politiques publiques est une nouvelle mission de la Cour des Comptes, en assistance au Parlement et au Gouvernement. C'est un nouveau métier pour la Cour et son approche est pragmatique.

L'article 47.2 de la Constitution stipule que la Cour assiste le Parlement et contribue à informer les citoyens en publiant les évaluations. Ce dernier point implique que les évaluations doivent être compréhensibles et accessibles.

Les quatre métiers de la Cour des Comptes sont désormais :

- Le contrôle juridictionnel des comptables publics et des agents comptables,
- Le contrôle organique (contrôles ou examens de gestion et des comptes des entités et des actions publiques),
- La certification des comptes publics,
- Les évaluations

Il s'agit de quatre métiers différents L'évaluation consiste à apprécier les résultats et les impacts des politiques publiques, ainsi que les liens de causalité associés. Des normes professionnelles ont été adaptées pour remplir cette nouvelle mission d'évaluation des politiques publiques et sont publiées sur le site de la Cour depuis décembre 2014. Une évaluation peut être demandée par le parlement mais la Cour s'autosaisit aussi pour en faire.

L'évaluation a fait débat dans la Cour des Comptes (c'est moins le cas aujourd'hui) en partie du fait d'un certain conservatisme mais aussi sur la légitimité pour la Cour de conduire des évaluations. En effet, la question est de savoir comment la Cour des Comptes peut aller au-delà de la mise en œuvre d'une politique publique sans empiéter sur les élus, le gouvernement.

Le cycle de l'évaluation (au sens de l'audit de l'ONU) :

Question clé, début du process → estimation de l'évaluabilité → conception → collecte des données → analyse des données → Présentation (rapport) → apprentissage → Utilisation

Tous les 4 ans, la Cour des Comptes revient sur les suites des recommandations de ses rapports d'évaluation.

La méthodologie comporte la mise en place d'un comité d'accompagnement qui est à distinguer d'un comité de pilotage prôné par la SFE. Ce comité est consultatif.

Le travail d'analyse recherche les liens de causalité entre les actions publiques et leurs effets à court ou moyen terme. Il associe les parties prenantes (décideurs, opérateurs et bénéficiaires).

Le rapport d'évaluation comporte l'exposé des opinions des parties prenantes et il y a des procédures de contradiction.

Les difficultés rencontrées dans le travail de l'évaluation sont les délais et la temporalité. Il y a environ quatre mois de procédure et douze mois d'évaluation.

Cela fait cinq ans que la Cour des Comptes a commencé sa mission d'évaluation. Le processus d'évaluation nécessite beaucoup plus de moyens en termes de temps de rapporteurs que les contrôles classiques. Aussi, seulement une à deux évaluations sont réalisées annuellement. On parle « d'enquêtes » lorsque les normes ne peuvent être suivies faute de

temps ou parce que le sujet ne se prête pas à une évaluation complète. C'est souvent le cas pour les demandes du Parlement auxquelles la Cour doit répondre en 12 mois au plus.

La Cour a des avantages par rapport à d'autres institutions pour la conduite des évaluations, en particulier son indépendance. De plus, elle connaît bien les acteurs et activités publiques, car elle travaille en pluridisciplinarité en interne avec des formations inter-chambres. Il y a donc des possibilités de recruter des experts et d'avoir un accès facilité aux données des administrations.

La mission principale de la Cour reste, pour une évaluation comme pour les autres missions, de « s'assurer du bon emploi de l'argent public et d'en informer le citoyen ».

Il y a un besoin croissant de faire de l'évaluation à la Cour, d'acquérir des données, de faire des traitements de statistiques évolués afin d'accroître les compétences de la Cour.

2. Intervention de Hélène Paris, L'évaluation des politiques publiques au conseil d'analyse économique

Secrétaire Générale du Conseil d'Analyse Economique

Le Conseil d'Analyse Economique a deux points communs avec la Cour des Comptes, à savoir l'indépendance et la publication de l'ensemble de ses travaux. Le Conseil d'analyse économique est une instance créée sous le premier Ministre Lionel Jospin en 1997, et a pour mission d'éclairer par les confrontations des points de vue des membres, les décisions. Il a un rôle d'aide à la décision.

En 2012, le Conseil connaît une reconfiguration avec quinze membres économistes universitaires.

Le CAE a produit une note sur « L'évaluation des politiques publiques » en 2013. Selon le texte de loi, l'évaluation vise à répondre à la question suivante: « **dans quelle mesure une politique publique atteint l'objectif assigné et produit les impacts escomptés auprès des publics concernés** ».

Il existe trois finalités à l'évaluation : 1. Connaître, 2. Porter un jugement et 3. Alimenter la prise de décision pour le décideur. Il y a donc ici la portée d'un enjeu démocratique. On se questionne sur la façon de dépenser mieux, ou moins ou encore ailleurs. C'est bien une aide à la décision avec un budget contraint.

L'évaluation s'intéresse aux résultats et aux impacts politiques. On doit dissocier l'évaluation de l'audit et du contrôle. Le contrôle et l'inspection se concentrent sur le respect et la conformité par rapport aux textes, cherchent à analyser le processus interne à l'institution et suivre l'exécution des actions.

En 2012, dans le cadre de la politique de la modernisation de l'action publique (MAP), il est décidé que toute politique doit faire l'objet d'une évaluation (en théorie).

L'accumulation au cours du temps de différents dispositifs dans les politiques publiques pose le problème de la visibilité ou de l'obsolescence, et l'évaluation se donne comme mission de répertorier les objectifs non atteints ou qui ont été déviés. L'un des pièges, est de confondre l'observation d'indicateurs de montée en charge d'un dispositif ou de tests de satisfaction (qui renvoient à la mise en œuvre du dispositif ou de la politique) avec de l'évaluation de politique publique.

Hélène Paris souligne quatre pièges dans l'évaluation :

- Les problèmes de causalité - La causalité inverse : ne pas confondre corrélation et causalité.
- La multiplicité des effets - Le syndrome de ne regarder que « sous le lampadaire », ce qui ne permet pas de voir ailleurs et se limite à une observation trop étroite.
- Minorer les aspects de la mise en œuvre (On peut avoir un dispositif « parfait » en théorie mais dont la mise en œuvre dénature complètement le projet initial).
- Les questions d'incidence - L'échec à toucher la cible des bénéficiaires. Celui à profiter le dispositif, n'est pas forcément celui à qui cela devrait profiter (exemple des aides au logement dont bénéficient in fine les bailleurs).

Comment s'en sortir ?

Il faut identifier le « contre factuel » afin de comparer les situations avec ou sans la politique que l'on cherche à évaluer. Cependant, l'exactitude du savoir qu'on en tire est relative.

On peut aussi utiliser les expériences aléatoires. Ce sont des méthodes expérimentales qui visent la comparaison de deux groupes de population qui ont des caractéristiques identiques et où on tire au hasard une population à qui on fait tester un dispositif, comparée à l'autre population qui n'en bénéficie pas (groupe témoin).

On observe ainsi l'écart des effets dans le temps. Les résultats de cette méthode ne sont pas toujours possibles à traiter, car il y a des difficultés pour repérer ce que la politique crée réellement comme effet.

Enfin, nous pouvons également utiliser les expériences « naturelles ». Pour cela, on repère des discontinuités ou des modifications de paramétrages de dispositifs, de prestations, de différenciation dans les territoires etc...

Sont utilisées des méthodes économétriques pour voir l'effet propre des politiques. Ces méthodes sont les plus rigoureuses pour effectuer une évaluation. La difficulté principale réside dans l'identification du bon groupe de contrôle.

Quelques éléments clés dans l'organisation de l'évaluation.

1/ Disposer de données fiables, 2/ Croiser les expertises techniques et administratives, 3/ assurer une indépendance des évaluateurs, 4/ des publications afin de diffuser les résultats et enfin, 5/ la pluralité des évaluateurs et l'interdisciplinarité.

En effet, il ne faut pas être trop proche ou impliqué dans le sujet, car il y a un risque de conflit d'intérêt et une approche mono-disciplinaire peut conduire à une recherche trop étroite.

Il y a une différence entre le temps de l'évaluation et le temps de la politique, qui a besoin de réponses immédiates. Nous nous devons de toujours contextualiser nos recherches.

Il existe une complémentarité du quantitatif et du qualitatif dans la démarche de l'évaluation. En effet, certains éléments ne sont nécessairement quantifiables et ne peuvent être documentés que via des enquêtes qualitatives, des monographies par exemple.

3. Débat

Question de Monsieur JACOT: N'y a-t-il pas une différence entre les démarches évaluatives et les rapports d'évaluation ?

Réponse d'A. MIGUS :

Il a fallu dix ans entre l'idée des politiques publiques et la normalisation des politiques publiques. L'évaluation est, pour la Cour des Comptes, un métier nouveau qui doit passer dans les mœurs.

L'idée étant que c'est en faisant qu'on s'améliore. On informe depuis peu, les parties prenantes d'une évaluation. Les parties prenantes n'ont pas d'obligation de donner des informations et données, mêmes si ces données sont relatives aux aides financières que leur donne l'Etat.

Pour la Cour des Comptes, la démarche d'évaluation est une activité encore trop récente. La Cour n'a pas la prétention de faire des évaluations de manière académique ni comme des évaluations de politiques publiques.

La Cour est implantée dans différents endroits des territoires, mais n'a pas toujours les moyens d'évaluation car cela coûte cher.

Ainsi un certain mélange se manifeste entre rapports d'évaluation, rapport d'observation et démarches évaluatives.

Pour le Conseil d'Analyse Economique (CAE), les notes et les rapports font partie des démarches évaluatives alors qu'il s'agit en fait de rapports d'analyse avec une démarche d'évaluation mais ce ne sont pas des rapports d'évaluation.

Question: A travers quelles méthodes le CAE met-il en avant la partie évaluation des politiques publiques ?

Réponse d'H. Paris :

Le Conseil d'Analyse Economique (CAE) ne prétend pas réaliser directement des travaux d'évaluation mais il s'appuie sur des travaux effectués par des chercheurs ou des administrations pour produire des rapports d'analyse afin de fournir des préconisations des politiques publiques.

Certains membres du CAE sont en lien avec le comité d'évaluation coordonné par France Stratégie dans la réalisation d'études d'impacts de type évaluation ex-ante. Pour une véritable modernisation des actions politiques, il est très important de s'appuyer sur les travaux d'évaluation, qui permettent d'apprécier l'efficacité des politiques et pouvoir faire des choix éclairés dans un contexte de budgets publics de plus en plus contraints.

Question de Monsieur FAUQUEUR:

La France a longtemps été plombée par l'Inspection Générale des Ministères car pour ces derniers, l'évaluation était déjà faite. Quelle est la place de la Cour des Comptes ?

Réponse d'A. Migus :

Dès que l'on fait de la politique, il est important que l'on regarde ce qui se fait ailleurs dans les autres pays.

En effet, il faut prendre en compte la culture, les normes sociales et juridiques des autres sociétés ainsi que les effets selon les typologies de publics, leur modèle de vie et leurs localisations.

La Cour des Comptes a permis donc de sortir par le haut après des années plombées par l'Inspection Générale.

Question de M. MIGUS à FAUQUEUR

Aujourd'hui atteint-on la limite de l'évaluation par les pairs, est-ce aujourd'hui la limite de la Cour des comptes ?

Réponse de FAUQUEUR à MIGUS

Aujourd'hui il faudrait faire une évaluation de l'évaluation en France.

Intervention de Monsieur PERRET

Deux remarques pour Madame PARIS :

- vous séparez dans l'évaluation analyse de la causalité et processus mise en œuvre ?
- la question du pluralisme des méthodes : les expérimentations aléatoires et les expériences naturelles demandent des conditions très restrictives (intervention stable, population visée, effet particulier) ; elles sont peu adaptées aux politiques aux effets complexes. Dans la littérature anglo-saxonne, on a l'idée que la causalité peut être traitée par d'autres méthodes qualitatives (méthode des contributions, méthode de discussion argumentée, par exemple) (cf. Cahiers de la SFE, *Evaluation des impacts des politiques et des programmes publics*, 2011)

Intervention de M.BELLECARAS

Questions de Monsieur BELLECARAS

- Il y a ambiguïté concernant les évaluations d'impact : polysémie du terme. Qu'en est-il des évaluations d'impact qui doivent désormais accompagner les projets de loi et aider à la maîtrise budgétaire ?

Réponses de Monsieur MIGUS

Il s'agit d'un métier nouveau, cela prend du temps, on y vient par le « big data ».

Concernant la maîtrise budgétaire jusqu'où vous ramenez au coût ?

Il faut se référer aux directives européennes.

Questions de Monsieur JOURNET

Quelles sont les modalités de recrutement des agents de la Cour des comptes ?

Est-ce une pensée univoque ou orientée ?

L'évaluation est une aide à la décision politique, le fruit d'une représentation de la nation.

Mais pourquoi ne s'en préoccupe-t-on pas ?

Question de Monsieur BRUNETIÈRE :

Fait-on un champ de l'évaluation, ou bien tous les champs sont liés à l'évaluation ?

Le terme « évaluation » n'est pas breveté et le point de vue de la Cour des Comptes y fait appui mais pas sur le point de consensus de deux conceptions de l'évaluation.

Réponses de Monsieur MIGUS :

Selon lui, ce serait plutôt la deuxième proposition (tous les champs sont liés à l'évaluation).

La Cour a une casquette d'évaluateur et une casquette de contrôleur. Ceci dit, il ne faut pas que la population la voit comme étant les deux à la fois. La Cour n'a pas l'exclusivité de l'évaluation puisque d'autres personnes le font aussi.

Ainsi pour Monsieur MIGUS, tout le monde fait de l'évaluation sous forme de deux casquettes : le contrôle et l'évaluateur. Toujours selon lui, les administrations ne doivent pas nous regarder selon nos casquettes, et il rappelle que le personnel de la Cour des comptes est constitué de magistrats, et rapporteurs du ministère (ingénieurs).

Conclusion : Pierre Savignat, Président de la société française de l'évaluation

Hier, se tenait à Bercy un colloque sur les techniques quantitatives de l'évaluation⁴. Sans nier l'intérêt de ces dernières, nous y reviendrons, l'objet de nos propos est de montrer non seulement l'articulation entre les données quantitatives et qualitatives mais, plus encore, leurs imbrications. Le fait de les opposer, ou même de les distinguer de façon quelque peu mécanique, ne permet pas d'enrichir les méthodes et outils évaluatifs.

Les processus évaluatifs doivent s'appuyer sur des données suffisamment robustes et fiables pour appuyer les analyses et appréciations sur l'action publique. Ceci étant dit, il convient aussi de questionner les données et leurs modes de production principalement sous deux angles :

-De quelles données avons-nous besoin en évaluation : que veut-on mesurer, quantifier ? Pourquoi ? Comment ?

-Quelle est la nature des informations ainsi produites : utilités ; limites ; pertinence ; portée ; sens ; biais, etc.

Ce questionnement contribue aux controverses relatives à la qualité des données et éclaire les parties prenantes aux processus évaluatifs dans leurs utilisations. Cela permet de comprendre les données comme des constructions socio-économiques et de se défier de la vision d'une scientificité ontologique du chiffre.

Plus largement, les travaux de cette journée, montrent que les données quantitatives, pour utiles qu'elles soient, ne peuvent suffire à rendre compte de la complexité du réel. Cela nous conduit à chercher à mettre les données en perspectives en les discutant dans ce qu'elles disent, en les contextualisant, en les croisant avec d'autres approches, d'autres outils. Ce travail est d'autant plus nécessaire que l'évaluation cherche à établir des relations causales et pas seulement des corrélations.

En résumé, les données quantitatives et qualitatives sont non seulement indispensables mais elles doivent s'articuler, permettre des interrogations croisées et critiques sur les outils, pour conforter les méthodologies d'évaluation. Le caractère pluridimensionnel du processus

⁴ Le 17 octobre 2015.

évaluatif, l'intérêt d'approches pluridisciplinaires, pluri professionnelles permettent ces regards croisés. La mise en débat avec l'ensemble des parties prenantes, et notamment les producteurs et collecteurs de données y contribuent fortement.

Il est temps que des échanges sur ces questions se développent. Il est temps de sortir des querelles pour le moins dogmatiques et peu opératoires entre quantitativistes et qualitativistes. Cette journée en est une étape significative. Les échanges et réflexions seront poursuivis notamment au travers d'une journée d'étude annuelle.

Bibliographie

- Brunetière, J.-R. (2010), « Les objectifs et les indicateurs de la LOLF, quatre ans après... », *Revue française d'administration publique*, 3, n°135, p. 477-495.
- Cahiers de la SFE, *Evaluation des impacts des politiques et des programmes publics*, 2011.
- Caresche, C., Pandraud, R., 2003, « Sur la création d'un Observatoire de la délinquance » *Rapport au Premier ministre*, La Documentation française, Paris.
- Coleman, J. S., et al. (1966) : *Equality of Educational Opportunity*. Washington, U.S. Department of Health, Education and Welfare
- Conseil d'analyse économique, 2013, « L'évaluation des politiques publiques », *Notes du CAE* N° 1, 12p.
- Desrosières, A. et Kott, S. (2005), « Quantifier », *Genèses*, 2005/1, n° 58, p. 2-3.
- Desrosières, A. (2008), *L'argument statistique I*, Paris, Presses de l'École des mines.
- Fargeon, V., Le Quéau, P., Le Roy, A., Offredi, C., Ottaviani, F. (2014), « L'étude Ibest : une mesure du bien-être comme tension », in Lamotte, B., Le Roy, A., Massit, C., Puissant, E. (eds), *Innovations sociales, innovations économiques*, Louvain la Neuve, Presses universitaires de Louvain, Cahiers du Cirtes, hors-série #4, p. 225-242.
- Easterlin, A. R. (1974), « Does Economic Growth Improve the Human lot? », in Paul A. David and Melvin W. Reder, *Nations and Households in Economic Growth: Essays in Honour of Moses Abanovitz*, New York, Academic Press.
- France stratégie, 2015, Les nouveaux indicateurs de richesse, Paris, Rapport du Gouvernement.
- Gadrey, J. (2006), « Les conventions de richesse au cœur de la comptabilité nationale. Anciennes et nouvelles controverses », in Eymard-Duvernay, F. (dir.), *L'économie des conventions. Méthodes et résultats*, Tome I, Paris, La Découverte, p. 311-324.
- Gibert, P. et Alilat, M. B. (2012), « Débat public et jeux sur les indicateurs. L'exemple du « taux d'élucidation » et du « nombre de gardes à vue », *Management international*, Vol. 16, n°3, p. 118-138.
- Leca, J. (1993), « L'évaluation dans la modernisation de l'Etat », *Politiques et management public*, Vol. 11, numéro 2, p. 161-172.
- Perret, B. (2002), *Indicateurs sociaux, état des lieux et perspectives*. Les Papiers du CERC, Rapport au conseil de l'Emploi, des Revenus et de la Cohésion Sociale (CERC), janvier 2002, [en ligne] <http://www.cerc.gouv.fr/doctrav/2002-01.pdf>, consulté le 11/05/2015.
- Stern J., 2006, L'évaluation économique des conséquences du dérèglement climatique, Paris, Rapport.
- Supiot, A. (2015), *La gouvernance par les nombres*, Paris, Fayard.
- Viard J., 2015, "A quoi servent les chiffres ? ", Actes Conférence Cnis et INSEE, 5 février.