

**Intervention au séminaire de l'unité Migrations Internationales et Minorités,
INED, séance du 5 juin 2007 : « Testing - avancées et limites d'une méthode »
Eric Cédiey (ISM-Corum)**

À partir de l'expérience de l'enquête par tests de discrimination conduite sur la France en 2006 par ISM-Corum pour le Bureau International du Travail et le Ministère de l'Emploi ¹, et sur l'exemple aussi des autres testings scientifiques, je vais tenter d'ouvrir une discussion plus générale sur la portée et les limites de cette technique de mesure. Ce qui bien sûr va nous amener à l'évaluer en regard d'autres techniques de mesure, ce qui bien sûr va nous situer dans les débats qui courent sur ces questions en France.

Face aux différentes techniques disponibles, la première question est celle de **l'objet** : qu'est-ce que l'on prétend mesurer ? Cette question je trouve manque de clarification dans les débats que je viens d'évoquer, je vais donc m'y attarder.

Les testings prétendent mesurer des discriminations. De quelles définitions disposons nous pour cet objet ? Ma réflexion, et les travaux menés à ISM-Corum, se tournent d'abord vers les définitions qu'en donne le droit. Pour diverses raisons, dont au moins les deux suivantes :

- les définitions juridiques de la discrimination ne procèdent pas d'abord d'une modélisation théorique, mais d'abord d'une visée opératoire empirique, et cette opérationnalité est sans cesse précisée et affinée par la jurisprudence ;
- sinon dans la vie intellectuelle, du moins dans la vie sociale pratique, les définitions de la discrimination par le droit doivent bien à un moment donné faire référence commune, précisément parce qu'il s'agit des règles de notre droit.

Ce choix ne signifie pas que le droit épuise toutes les analyses et toutes les actions possibles sur les discriminations, parce que c'est tout à fait faux. Il se fonde simplement sur le constat

¹ Cédiey E. et Foroni F. (2007), *Les discriminations à raison de "l'origine" dans les embauches en France. Une enquête nationale par tests de discrimination selon la méthode du Bureau International du Travail*, rapport complet et synthèse téléchargeables sur www.ismcorum.org et sur www.ilo.org/public/french/bureau/inf/download/discrim_france.pdf

que le droit offre des définitions précises, opératoires, et destinées à faire référence commune pour tout le monde, et pas seulement pour les chercheurs, ou pour tels agents économiques par exemple.

Notre droit en la matière est communautaire européen. Il identifie quatre formes de discriminations. Il en est deux qui n'apparaissent guère dans les débats sur la mesure, ce sont le harcèlement discriminatoire et l'injonction de discriminer ; il y aurait peut-être là une occasion d'analyser les points aveugles de nos débats, mais laissons cela de côté. Les deux autres qualifications juridiques sont en revanche mobilisées dans les débats sur la mesure, il s'agit de la discrimination directe et de la discrimination indirecte. Rapidement :

- une discrimination directe se produit lorsqu'une personne, sur le fondement d'un motif prohibé (l'origine, le sexe, etc.), est traitée de manière moins favorable qu'une autre ne l'est, ne l'a été ou ne le serait dans une situation comparable ;
- une discrimination indirecte se produit lorsqu'une disposition, un critère ou une pratique apparemment neutre est susceptible d'entraîner un désavantage particulier en fonction d'un motif prohibé, sauf si cette disposition, critère ou pratique remplit certaines conditions (elle est objectivement justifiée par un objectif légitime et elle est un moyen approprié et nécessaire de réaliser cet objectif).

Il me semble que dans les deux cas (et le harcèlement et l'injonction peuvent ici être réintégrés), nous sommes placés sur le même registre : les discriminations résident dans des pratiques, dans des traitements ou des processus, des dispositions ou des règles, des décisions ou des critères opératoires. Elles ne résident certainement pas dans les personnes ou les populations que visent ces pratiques. Ce qui fait que les exercices qui sont développés en termes de « mesure de la diversité » proposent tout autre chose qu'une mesure des discriminations ; et qu'en termes de préconisations d'action, le discours de la promotion (quantitative) de la diversité et celui du refus des discriminations peuvent entrer en contradiction directe (le « management de la diversité » est une autre question).

Mais je crois que les discriminations ne résident pas non plus, en tant que telles, dans les positions différenciées occupées par tel ou tel type de personnes ou de populations. Les exercices statistiques ou économétriques classiques qui analysent les positions de différentes populations sur le marché du travail, par exemple, peuvent assurément proposer une approche

des discriminations, mais une approche qui reste indirecte. Si des écarts de positions sont constatés entre différentes personnes en fonction d'un motif de traitement prohibé, si ces écarts sont systématiques, d'une ampleur significative, et sont mesurés toutes choses égales d'ailleurs, cela autorise pleinement à poser l'hypothèse que des pratiques discriminatoires participent à la formation de ces écarts. Mais il reste entièrement à identifier les pratiques ou les processus discriminatoires qui peuvent être en cause. Y compris (et peut-être surtout) si l'on prétend montrer l'existence de discriminations indirectes, puisqu'aussi bien une pratique apparemment neutre mais qui ne l'est pas dans les désavantages qu'elle entraîne peut quand même être non discriminatoire si elle remplit certaines conditions. Il me semble donc que les statistiques de positions ne sont jamais suffisantes à l'identification des discriminations, même indirectes.

Pour compléter le propos entre statistiques et discriminations indirectes, qui sont souvent associées dans les débats sur la mesure, je pense que l'on peut ajouter que les premières ne sont pas forcément nécessaires à l'identification des secondes. Il ne manque pas d'exemples où des discriminations indirectes sont identifiées sans qu'il ne soit besoin de recourir à nulle statistique. Ce qui fait que l'argument selon lequel la consécration de la qualification de discrimination indirecte en droit français *réclame* le développement de statistiques de positions différenciées selon les motifs prohibés, parce que ces statistiques seraient complètement nécessaires à la révélation des discriminations indirectes et le seul instrument en mesure de le faire, me semble quelque peu partiel.

Ce que j'ai appelé les statistiques de positions ne réalisent donc pas une mesure des discriminations au sens exact du terme, ni ne suffisent à leur identification. Elles mesurent plutôt des inégalités de situations, dont les causes sont loin de pouvoir être rapportées, d'ailleurs, aux seules discriminations, et qui n'en constituent pas moins un objet et un problème en tant que telles. **À trop confondre les discriminations avec les inégalités de situations, je crois en fait qu'il faut redouter que nous ne traitions de façon inadéquate et les unes, et les autres.** À tirer les discriminations vers les inégalités de situations, nous risquons de rater la cible spécifique des politiques anti-discriminatoires, à savoir les pratiques ou les dispositions qui discriminent, leurs ressorts et leurs acteurs. Mais à ramener les inégalités de situations selon le genre ou l'origine, par exemple, à la seule cause des discriminations, nous risquons d'être extrêmement réducteurs et de laisser de côté bien

d'autres causes de ces inégalités : les socialisations inégalitaires, les inégalités de structure entre les positions sociales disponibles, les inégalités de conditions en divers types de ressources, les auto-sélections, les mécanismes de reproduction des inégalités, les inégalités des opportunités et des possibles, les effets des rapports de domination en général... Pour prendre en compte toute l'ampleur et la complexité des inégalités, je crois qu'il faut veiller à proposer des techniques de mesure et des traitements appropriés pour chacun de leurs facteurs ou mécanismes, les discriminations et les autres, sans les réduire et les confondre entre eux.

Revenons au testing. Le testing ne peut proposer qu'une mesure des discriminations, certainement pas des inégalités de situations en général. Mais la pertinence spécifique de la technique du testing en la matière, c'est de prendre pour objet les pratiques et les procédures concrètes, et de pouvoir saisir ainsi les discriminations de façon directe quand elles se manifestent. On peut alors considérer que pour identifier et caractériser effectivement des discriminations, à l'embauche par exemple, un testing sur les processus de recrutement sera plus fiable et plus précis que des statistiques sur les différentiels de chômage (ce qui ne veut pas dire que ces statistiques sont inutiles).

Il me semble que le testing, à travers l'approche directe qu'il propose, peut concerner les deux formes de discriminations, directes comme indirectes. Même si la première cependant, la discrimination directe, est sans doute plus fréquemment rencontrée dans les tests et plus identifiable au seul vu de leurs résultats. Mais il me semble que l'on opère des testings qui rencontrent des discriminations indirectes relativement à l'origine, par exemple, lorsque l'on teste des offres d'emploi qui réclament la maîtrise d'un dialecte local (critère apparemment neutre pouvant être légitime dans certaines conditions), ou lorsque l'on teste la variable du lieu de résidence (même commentaire).

Les statistiques de positions de leur côté, à travers leur approche seulement indirecte des discriminations, peuvent néanmoins concerner les deux formes de discriminations elles aussi, directes comme indirectes. Les statistiques classiques du marché du travail le font en général indistinctement, et il est difficile d'y trouver des indications sur chacune des formes discriminatoires, parce qu'il faudrait remonter, entre autres, aux choix concrets de recrutement : au vu de certains écarts dans les taux de chômage ventilés par sexe ou par origine on est fondé à présumer l'existence de discriminations *directes ou indirectes* dans les

embauches, mais il est difficile de construire à partir des taux de chômage des indices sur la présence et la part de chacune de ces formes de discriminations. Dans d'autres domaines néanmoins, à condition de produire des statistiques de positions très précisément ajustées, l'identification de la discrimination et de sa forme, indirecte ou directe, sera bien plus robuste. Ainsi, l'essentiel du contentieux de la discrimination *directe* pour raison syndicale dans les déroulements de carrière en France, s'appuie sur des mesures d'écart de positions entre groupes-témoins, en analyse longitudinale par étapes, qui permettent y compris de chiffrer le préjudice salarial subi.

En résumé : dans les débats sur la mesure des discriminations, on voit circuler l'idée que le testing ne serait utile qu'à la mesure des discriminations directes, et les statistiques à la mesure des discriminations indirectes. Je ne suis pas sûr que le grand partage soit là, et il me semble que les exemples montrent que toutes les combinaisons sont possibles.

Avec le dernier exemple évoqué, celui des déroulements de carrière, nous passons du choix des objets (les discriminations, ou les populations, ou leurs positions relatives) à celui des **domaines d'investigation**.

Si l'on veut étudier des discriminations, alors il faut pouvoir rattacher l'exercice de mesure à une pratique ou des pratiques déterminées. Ce qui signifie également que *tout* domaine correspondant à un ensemble de pratiques ou de dispositions identifiables est a priori éligible pour un exercice de mesure des discriminations : par exemple des procédures d'orientation scolaire, ou un dispositif d'accompagnement vers l'emploi, ou l'ensemble des actes de gestion des ressources humaines chez un employeur, ou un système d'allocation de logements sociaux, ou un système judiciaire, etc.

Il me semble ici que le testing rencontre clairement ses limites. Comme les tests de discrimination sont des expériences artificiellement contrôlées mais que l'on prétend plonger dans des situations de vie réelle, ils n'apparaissent guère praticables que sur des processus d'entrée — recrutements, inscriptions scolaires, accès aux biens et services — qui mettent en présence des gens qui ne se connaissent pas. Il est difficile d'imaginer, en revanche, ce que seraient des tests de discrimination en matière de promotion professionnelle ou de formation

continue, par exemple. Je ne crois donc pas, en particulier, que l'on puisse soutenir que le testing est une technique qui suffise à la mesure des discriminations au sein des entreprises ou chez les employeurs publics.

Des statistiques de positions peuvent-elles aider à l'identification de ces types de discriminations internes qui se développent au sein des organisations ? Assurément, **mais à certaines conditions**, que l'on a commencé d'évoquer sur l'exemple des déroulements de carrières. Je suggèrerais qu'il vaut mieux :

- que ces statistiques soient précisément ajustées aux systèmes organisationnels qu'elles veulent étudier ;
- qu'elles construisent des analyses longitudinales pour essayer de reconstituer des processus internes ;
- qu'elles le fassent par étapes, dans des exercices de comparaison avant-après resserrés autour de pratiques identifiables (par exemple tel moment d'orientation scolaire, ou l'accès à tel échelon hiérarchique, etc.) ;
- et qu'elles le fassent sur les positions occupées par différents groupes-témoins ou cas-témoins caractéristiques des motifs prohibés, et constitués de façon réfléchie sur les variables à contrôler (par exemple l'ancienneté, ou les évaluations, etc.).

Avec le type d'études qui se dessinent là, on s'écarte quelque peu des productions classiques de la statistique publique, à savoir des statistiques socio-démographiques ou économiques en populations générales². Et on se rapproche plus, mutatis mutandis, de techniques de mesure qui sont utilisées dans le champ judiciaire précisément pour identifier l'existence de discriminations dans les déroulements de carrière, et qui, dans ce champ, ont été appelées les « comparaisons de panels ».

² On peut noter en même temps une concordance, dans l'esprit de certaines des propositions ci-dessus, avec ce que réalisent les enquêtes « Générations » du Céreq, mais dans le domaine de *l'accès* à l'emploi.

Ceci amène, je crois, à préciser l'existence, pour finir, de deux **logiques de mesure** différentes. Et j'en profite d'être à l'INED pour remercier Benoît Riandey³ d'avoir attiré ma réflexion sur ce point.

Il me semble que nos débats sur la mesure gagneraient à savoir distinguer entre les deux logiques suivantes :

- d'une part, les exercices de mesure statistique qui ont d'abord pour propos de caractériser des populations, que ce soit dans leurs attributs, leurs comportements, leurs positions, leurs trajectoires : ces exercices vont, soit s'appuyer sur un recensement exhaustif de ces populations, soit, le plus souvent, sur la théorie de l'estimation pour constituer des échantillons représentatifs des populations en question ;
- et d'autre part, les exercices de mesure statistique qui ont d'abord pour propos de vérifier l'existence et de caractériser les formes et les conditions de manifestation d'un risque différencié que peuvent plus ou moins rencontrer les personnes en fonction d'un critère donné : ces exercices vont pouvoir s'appuyer sur la théorie des tests statistiques pour constituer des séries de cas-témoins ou différents groupes-témoins stéréotypés en fonction du critère en question, avec des effectifs suffisants pour s'assurer de la significativité des écarts mesurés, mais sans avoir à se soucier que ces séries ou que ces groupes soient représentatifs de populations plus générales.

Ces deux logiques de mesure ont chacune leur valeur scientifique. Elles ne sont ni exclusives ni nécessairement associées, tout simplement parce qu'elles n'ont pas le même objet. **Je pense que lorsque l'on veut traiter effectivement de l'objet *discrimination*, la seconde logique de mesure est la plus appropriée.** C'est cette logique que le testing pousse à son maximum ; c'est aussi celle que développent les « comparaisons de panels ».

Si l'on s'intéresse à une entreprise par exemple, ou toute autre organisation ou dispositif, une mesure prenant pour objet les pratiques ou les processus discriminatoires ne réclame pas que l'on compte et que l'on classe ou que l'on représente tout le monde. L'exercice va pouvoir

³ Statisticien à l'INED, ancien responsable du Service des enquêtes, président du groupe Enquêtes de la Société Française de Statistique.

s'appuyer sur la constitution de groupes-témoins stéréotypés ou d'une série de cas-témoins stéréotypés selon le(s) motif(s) prohibé(s). Si les situations et les parcours de ces groupes ou cas-témoins, analysés de façon ajustée aux pratiques et aux processus qui les affectent, révèlent des écarts de positions entre eux, alors ces pratiques ou processus peuvent être considérés comme discriminatoires. On peut alors essayer de qualifier ces discriminations dans les formes qu'elles prennent, les moments précis où elles interviennent, les éléments des situations où elles se manifestent avec plus ou moins d'acuité, ce qui peut bien sûr s'avérer utile pour les corriger. La question du dénombrement ou de l'estimation de toutes les personnes qui les subissent est une autre question, qui peut avoir, ou non, son intérêt, mais à laquelle, quoi qu'il en soit, il n'est pas besoin de répondre pour repérer les discriminations et les traiter.

En guise de conclusion :

Les trois questions que j'ai essayé d'introduire, celle de l'objet, celle du domaine d'investigation et celle de la logique de mesure, me semblent devoir être posées et devoir recevoir des réponses si l'on veut être clair et cohérent sur l'exercice de mesure dont on parle. Ces trois questions, pourtant, peuvent paraître négligées dans les débats qui courent en la matière en France. Ces débats s'obnubilent d'une quatrième question : celle du mode de construction, du libellé et du niveau d'application des catégories relatives à certains des motifs prohibés, en particulier les catégories de « l'origine ». La question de la catégorisation est certes d'importance, sur les registres scientifique comme politique. Mais elle mérite d'autant moins de faire oublier les trois autres qu'il se pourrait qu'elle en dépende, sur les deux registres.