



N° 11-522-XIF au catalogue

**La série des symposiums internationaux
de Statistique Canada - Recueil**

Symposium 2003 : Défis reliés à la réalisation d'enquêtes pour la prochaine décennie

2003



Statistique
Canada

Statistics
Canada

Canada

Recueil du Symposium 2003 de Statistique Canada
Défis reliés à la réalisation d'enquêtes pour la prochaine décennie

LA RECHERCHE EN SONDAGES AU COURS DE LA PROCHAINE DÉCENNIE : PERSPECTIVES D'AVENIR ET DÉFIS

Graham Kalton¹

RÉSUMÉ

Durant la deuxième moitié du siècle dernier, la recherche en sondages a connu un grand essor en ce qui a trait au nombre d'enquêtes, à la gamme de sujets considérés comme pouvant faire l'objet d'études et à la méthodologie. L'utilisation de données d'enquête pour l'élaboration des politiques et la recherche empirique continue de croître. En outre, les analystes se servent de plus en plus de ce genre de données pour appuyer leurs analyses approfondies. Au cours de la prochaine décennie, il est probable que l'on assistera à de nouveaux développements importants, particulièrement en ce qui concerne les enquêtes internationales, les enquêtes par panel, l'estimation pour petites régions, les études par observation, les analyses secondaires et les modes de collecte des données. Le plus grand souci est que la coopération des membres du public aux enquêtes diminuera et que les taux de non-réponse augmenteront. Les méthodes de pointe utilisées aujourd'hui en recherche en sondages ont donné lieu à une spécialisation dans les diverses disciplines sur lesquelles s'articulent le processus d'enquête. La création d'un organisme professionnel international regroupant ces disciplines est examinée. Certaines idées concernant l'établissement d'un programme de formation permanente en statistique d'enquête sont présentées.

MOTS CLÉS : Analyses secondaires, enquêtes internationales, enquêtes par panel, estimation pour petites régions, formation permanente, modes de collecte des données, non-réponse, organisme professionnel.

1. INTRODUCTION

L'objectif ambitieux du présent article est de faire l'inventaire des principaux changements qui surviendront vraisemblablement dans le domaine de la recherche en sondages au cours de la prochaine décennie. Dans certains cas il s'agira de progrès en méthodologie, dans d'autres, de l'augmentation de la fréquence de certains types d'enquêtes et dans d'autres encore, de conséquences de l'évolution externe des technologies et de la société. Il est, bien sûr, dangereux d'essayer de prévoir ce genre de faits nouveaux. Comme l'a fait remarquer Nils Bohr « prévoir est un art très difficile, surtout pour le futur ». Je vais malgré tout tenter de le faire. En empruntant la langue des météorologistes, mes prévisions sont les suivantes « ensoleillé, avec un nuage noir menaçant pointant à l'horizon ».

La présentation de l'article est la suivante. La prochaine section est consacrée à l'examen de certains changements importants survenus au cours du dernier quart de siècle, dans le but de donner une perspective historique et de dégager les tendances qui pourraient exister. La section 3 donne des prédictions pour la prochaine décennie en ce qui concerne les types d'enquêtes et d'estimations d'après des données d'enquête dont la demande augmentera vraisemblablement, les défis méthodologiques que posent ces enquêtes et les progrès vraisemblables dans le domaine de la collecte des données. La section 4 est consacrée à l'avenir de la profession de la recherche en sondages, sa structure organisationnelle et les besoins de formation des statisticiens d'enquête. Enfin, la section 5 contient certaines conclusions.

2. PROGRÈS PASSÉS

Dans un article antérieur (Kalton, 2000), j'ai passé en revue les progrès réalisés dans le domaine de la recherche en sondages durant le dernier quart de siècle. Les ordinateurs y occupent une place importante, à cause de leur incidence énorme sur de nombreux aspects de la méthodologie d'enquête. Les ordinateurs ont été introduits, et sont

¹ Graham Kalton, Westat, 1650 Research Boulevard, Rockville, MD, 20850, U. S.A. (GrahamKalton@Westat.com).

maintenant généralement acceptés, dans le domaine de la collecte des données par interview personnel assistée par ordinateur (IPAO), par interview téléphonique assistée par ordinateur (ITAO), par auto-interview assistée par ordinateur (AIAO) et par audio-AIAO pour la collecte de données délicates. Les logiciels nécessaires pour traiter ces applications ont été développés, et les premières versions simples ont fait l'objet de mises à niveau importantes durant cette période. Tous les programmes d'activités statistiques ont profité de la puissance accrue de calcul pour élaborer et appliquer des méthodes plus puissantes. Donc, la croissance de la puissance de calcul a stimulé l'élaboration de méthodes plus poussées de pondération pour tenir compte de la non-réponse et pour l'étalonnage, ainsi que le recours à l'imputation et la mise au point de méthodes d'imputation plus perfectionnées. Conjuguée à la création des logiciels appropriés, l'augmentation de la puissance de calcul a rendu presque routinier le calcul des erreurs d'échantillonnage pour des statistiques simples ainsi que complexes provenant de plans d'échantillonnage complexes, y compris l'utilisation de méthodes de rééchantillonnage exigeant une grande capacité de calcul. Enfin, l'application de méthodes analytiques plus complexes (régression linéaire multiple, régression logistique, analyse de données catégoriques et modélisation multiniveaux) à des données d'enquête a pris de plus en plus d'importance.

La facilité croissante avec laquelle les données d'enquête peuvent être analysées a poussé les analystes à étendre leurs horizons. Cette expansion a notamment eu pour résultat l'accroissement de la portée des bases de données d'enquête grâce, en partie, à l'ajout de questions au questionnaire et, en partie, à des couplages avec des données administratives. L'intérêt des analystes pour les analyses dynamiques ainsi que statiques a donné lieu à l'expansion rapide de l'utilisation d'enquêtes par panel au cours du dernier quart de siècle. Cette expansion a également eu pour résultat de stimuler la demande d'estimations pour de petits domaines, particulièrement pour de petites régions géographiques. En général, la taille d'un échantillon d'enquête ne permet pas de produire des estimations pour petites régions standards basées sur le plan de sondage d'une précision suffisante. Pour contourner ce problème, beaucoup d'efforts ont été consacrés en recherche à l'élaboration de méthodes d'estimation pour petites régions basées sur un modèle. Ces dernières années, l'utilisation de ces méthodes est devenue plus généralement reconnue. Un troisième résultat de l'expansion a été de reconnaître de plus en plus que les ensembles de données d'enquête fournissent souvent les données nécessaires pour toute une gamme d'analyses qui ne correspondent pas à l'objectif principal de l'enquête. L'analyse secondaire des données d'enquête ayant pris beaucoup d'ampleur, des moyens de faciliter l'utilisation des données d'enquête à cette fin ont été recherchés.

Un autre domaine méthodologique qui a suscité une attention croissante au cours du dernier quart du siècle est celui des erreurs non dues à l'échantillonnage, une importance plus grande ayant été accordée aux concepts d'erreur totale d'enquête et de plan de sondage total. Le regain d'intérêt pour la conception des questionnaires, suscité par le mouvement consacré aux aspects cognitifs de la méthodologie d'enquête, est un progrès qui a été bien accueilli.

Parallèlement aux progrès en analyse des données d'enquête, à l'usage plus fréquent du couplage avec des données administratives et à la demande d'accès aux ensembles de données aux fins d'analyse secondaire, les inquiétudes en ce qui a trait à la protection des renseignements personnels fournis par les répondants et de la confidentialité se sont accentuées. Ces préoccupations ont donné lieu à l'élaboration de méthodes de limitation de la divulgation des renseignements personnels dans le cas des fichiers de données à grande diffusion, et de sites sécurisés pour l'analyse par des chercheurs autorisés des fichiers de données non destinés à la grande diffusion.

Tous les faits nouveaux susmentionnés représentent des progrès pour la recherche en sondages. Le seul fait nouveau négatif est la moins grande disposition des membres du public à participer aux enquêtes. Cette diminution de la participation est compensée en partie par de plus grands efforts en vue d'obtenir leur coopération, mais malgré ces efforts, les taux de réponse ont baissé, du moins aux États-Unis. Il s'agit du « nuage noir » que j'ai évoqué plus haut. Comme je le mentionne plus loin, j'ai le sentiment que la non-réponse pourrait devenir un problème très préoccupant pour les spécialistes de la recherche en sondages.

Enfin, un événement important du dernier quart de siècle a été l'émergence manifeste de la discipline de la recherche en sondages (Kalton, 2000). Durant cette période, le lancement des revues *Techniques d'enquête* et *Journal of Official Statistics* a offert un véhicule valable pour les articles traitant de méthodologie d'enquête qui, avant la création de ces revues, étaient publiés de façon très dispersée. En outre, plusieurs associations professionnelles à l'intention des méthodologistes d'enquête ont été fondées durant cette période, durant laquelle ont aussi vu le jour nombre de conférences et de symposiums, nationaux ainsi qu'internationaux, axés sur la méthodologie d'enquête. Je reviendrai sur le sujet de la profession de la recherche en sondages à la section 5.

3. LA PROCHAINE DÉCENNIE

Commençons par l'exposé de mes prévisions en ce qui concerne l'évolution de la recherche en sondages au cours de la prochaine décennie. La plupart de ces prévisions sont simplement des projections des tendances récentes. Mais, dans certains cas, il s'agit plutôt de domaines de premier plan dans lesquels j'espère voir se réaliser des progrès méthodologiques dans les années à venir que de prédictions quant à ce qui se produira vraiment.

Enquêtes internationales et transnationales. L'intérêt pour la réalisation d'une enquête qui soit essentiellement la même dans plusieurs pays existe depuis de nombreuses années. La World Fertility Survey (WFS), lancée en 1972 grâce à un financement des Nations Unies, est un des premiers exemples de programme d'enquête internationale de grande envergure. En plus de fournir des données comparables sur le comportement lié à la fécondité pour de nombreux pays en développement, la WFS a contribué de façon extrêmement utile et durable à l'établissement des capacités requises pour réaliser des enquêtes dans ces pays. Les Demographic and Health Surveys (DHS) en cours, qui sont financées par l'Agence américaine pour le développement international (USAID), peuvent être considérées comme le prolongement de la WFS. Divers organismes des Nations Unies parrainent des enquêtes sur d'autres sujets, comme la Multiple Indicator Cluster Survey (MICS) de l'UNICEF et la World Health Survey (WHS) mise sur pied récemment par l'OMS. La Banque mondiale parraine l'Étude sur la mesure des niveaux de vie. Dans le domaine de l'éducation, l'OCDE parraine le Programme international pour le suivi des acquis des élèves (PISA), l'Association internationale pour l'évaluation du rendement scolaire parraine la Troisième enquête internationale sur les mathématiques et les sciences (TEIMS) et un certain nombre de pays ont participé à l'Enquête internationale sur l'alphabétisation des adultes (EIAA). Depuis les années 1980, les organismes d'enquête d'un nombre croissant de pays ont collaboré à la réalisation de l'International Social Survey Program (ISSP), qui est un programme annuel d'enquêtes sociales internationales. L'Enquête sociale européenne (ESS pour *European Social Survey*) a été lancée récemment par un certain nombre de pays européens. Eurostat a établi des enquêtes sur la population active (EPA), des enquêtes sur le budget des ménages (EBM) et des enquêtes fournissant des statistiques sur le revenu et les conditions de logement (enquêtes-ménages communautaires par panel) normalisées pour les pays membres de l'Union européenne.

Cette liste incomplète donne une idée de la diversité des auspices sous lesquels sont menées les enquêtes internationales et transnationales, auspices qui incluent des organismes internationaux, des collaborations entre bureaux nationaux de la statistique et des collaborations entre des chercheurs spécialisés dans le domaine social ne travaillant pas pour des organismes gouvernementaux. Outre la répétition d'enquêtes complètes dans les divers pays, il convient de souligner l'adoption et la mesure de concepts normalisés aux échelons international et transnational. Un bon exemple est le Système de classification des industries de l'Amérique du Nord (SCIAN) qui a été créé pour assurer la comparabilité des statistiques sur les activités des entreprises du Canada, du Mexique et des États-Unis.

Les enquêtes internationales et transnationales sont manifestement déjà bien établies et je prévois une croissance considérable dans ce domaine. Ces enquêtes posent de nombreux problèmes méthodologiques importants qui, selon moi, n'ont pas encore reçu l'attention qu'ils méritent. Pour obtenir des résultats comparables pour divers pays, il faut examiner minutieusement tous les aspects de la méthodologie d'enquête. Les défis considérables que pose la résolution des problèmes de comparabilité des traductions lors de l'utilisation de plusieurs langues ou de la comparabilité des concepts sous-jacents d'un pays à l'autre sont bien connus². Moins bien reconnue est la nécessité de considérer d'autres aspects de la méthodologie, comme la couverture des enquêtes, l'échantillonnage, la formation des intervieweurs et la qualité, les taux de non-réponse totale et de non-réponse partielle, la vérification et la pondération. Dans le cas des enquêtes internationales à financement externe, des normes de qualité uniformes peuvent être spécifiées et appliquées. Cependant, il est plus difficile d'assurer la qualité des données dans le cas d'enquêtes collectives transnationales. Les problèmes que suscitent les différences entre les méthodes d'enquête coutumières des pays qui collaborent peuvent être épineux. J'espère que l'on accordera plus d'attention à ces questions dans l'avenir. L'ouvrage sur les méthodes d'enquêtes transculturelles publié récemment par Harkness, Van de Vijver et Mohler (2003) est un signe encourageant. Le lecteur peut aussi consulter Verma (2002) sur la

² En à-côté, je soulignerais qu'il est moins bien reconnu que ces problèmes s'appliquent aussi au niveau national dans de nombreux pays et qu'ils prendront de plus en plus d'ampleur à mesure que la mobilité entre pays augmentera.

comparabilité internationale des statistiques d'enquête et Holt (2003) sur les questions que soulève la production d'indicateurs statistiques aux fins de comparaisons internationales.

Enquêtes par panel. Comme je l'ai mentionné à la section précédente, les enquêtes par panel sont devenues très en vogue au cours du dernier quart de siècle. Des enquêtes-ménages par panel conçues sur le modèle de la Panel Study of Income Dynamics (PSID) de l'Université du Michigan ont été mises sur pied dans de nombreux pays et des enquêtes comparables, comme la Survey of Income and Program Participation (SIPP) des États-Unis et l'Enquête sur la dynamique du travail et du revenu (EDTR) du Canada ont également été lancées. Nombre d'enquêtes par panel sont réalisées dans le domaine de l'éducation, dont les Early Childhood Longitudinal Studies, la National Education Longitudinal Study de 1988 et la Baccalaureate and Beyond Longitudinal Study réalisées aux États-Unis. Des études de cohorte de naissances ont été menées dans de nombreux pays, notamment en Grande-Bretagne, où la première, appelée National Survey of Health and Development, a été lancée en 1946 et se poursuit encore à l'heure actuelle. Aux États-Unis, on est en train de planifier une enquête par panel importante, la National Children's Study, en vue d'évaluer les effets environnementaux sur la santé et le développement. L'étude a pour but de suivre un échantillon de plus de 100 000 enfants d'avant leur naissance jusqu'à l'âge de 21 ans. Au Canada, l'Enquête nationale sur la santé de la population est réalisée tous les deux ans auprès d'un panel de Canadiens depuis 1994. Et je pourrais citer de nombreux autres exemples d'enquêtes portant sur toute une gamme de sujets. Les enquêtes par panel sont bien établies et sont même devenues à la mode aujourd'hui.

Comme d'autres enquêtes, les enquêtes par panel portent généralement sur un sujet primaire unique. Cependant, la répétition des interviews que comporte ce genre d'enquête donne la possibilité de recueillir des données sur toute une gamme de sujets connexes durant l'existence du panel. Souvent, des modules thématiques sont ajoutés lors de divers cycles pour recueillir des données supplémentaires en plus des données de base recueillies lors de chaque cycle. L'ajout de ces modules présente de grands avantages, le premier étant l'accroissement du potentiel analytique des données de panel et le second, la possibilité d'obtenir, grâce à la collecte de données supplémentaires, des fonds supplémentaires afin de couvrir le coût élevé d'une enquête par panel. En revanche, la collecte de données supplémentaires rend l'enquête plus complexe et peut occasionner un fardeau de réponse excessif et l'érosion anormalement forte du panel. Même sans les données supplémentaires, les enquêtes par panel sont complexes et le fardeau de réponse, lourd. Il convient donc de résister à la tentation de surcharger une enquête par panel en voulant répondre à un trop grand nombre de demandes.

Le principal point fort d'une enquête par panel est la capacité d'analyser les relations entre des variables mesurées à divers points dans le temps, comme la variation brute d'une caractéristique (p. ex., l'état de santé) d'un point dans le temps au suivant. Cependant, le fait que les erreurs de mesure puissent gravement fausser les résultats de ce genre d'analyse est un problème sérieux. Il faudra poursuivre les travaux de recherche en vue de découvrir des moyens de traiter l'erreur de mesure dans ce genre d'analyse longitudinale. Dans le passé, une critique courante des enquêtes par panel était que peu d'analyses longitudinales sont effectivement réalisées au moyen des données qu'elles produisent. Grâce aux grands progrès réalisés en informatique, la situation s'est améliorée considérablement, mais il n'en reste pas moins qu'en général, on pourrait réaliser rentablement des analyses longitudinales des données de panel plus complètes qu'elles ne le sont à l'heure actuelle. Offrir de l'aide aux analystes secondaires des fichiers de données d'enquêtes complexes et promouvoir l'utilisation de méthodes d'analyses longitudinales contribueraient à résoudre le problème de la sous-utilisation des données de panel recueillies à grand frais.

Estimation pour petites régions. Rao (2003) a fait un recensement des progrès importants réalisés ces dernières années en ce qui concerne les techniques statistiques d'estimation pour petites régions. Étant donné la volonté des décideurs d'acheminer les ressources jusqu'aux personnes nécessiteuses au niveau local, je pense que la demande d'estimation pour petites régions restera forte aussi bien dans les pays développés que dans les pays en voie de développement. Par exemple, la cartographie de la pauvreté demeure une préoccupation importante pour les pays en développement. Lors de futurs travaux, j'espère qu'on accordera plus d'attention à la collecte de données sur des variables auxiliaires mesurées uniformément pour les petites régions en vue de les utiliser comme prédicteurs dans la modélisation, ainsi qu'à l'évaluation minutieuse de la qualité des estimations pour petites régions produites d'après les modèles.

Études par observation. Les enquêtes sont utilisées à grande échelle dans des études conçues pour évaluer les effets de programmes d'intervention. Les interventions peuvent être de nature globale, comme les campagnes télévisées contre la drogue, ou concentrées dans un certain nombre d'emplacements d'essai, comme l'introduction

de techniques spéciales pour augmenter le taux de vaccination dans quelques régions choisies. Les emplacements d'essai sont parfois choisis aléatoirement, mais, le plus souvent, de façon intentionnelle. Les études peuvent comprendre des plans d'enquête par panel et des groupes témoins. Selon moi, l'utilisation de méthodes d'enquête dans le cadre d'études par observation — quasi expérimentales — de ce genre ou d'études expérimentales augmentera vraisemblablement dans l'avenir. L'analyse des données provenant de ces études pose des défis auxquels je pense qu'il faudra accorder plus d'attention. Des problèmes analytiques comparables se posent dans le cas des études épidémiologiques cas-témoin, où les groupes témoins sont sélectionnés par des méthodes d'échantillonnage, souvent par la méthode téléphonique de composition aléatoire.

Analyses secondaires. À mon avis, la demande de données d'enquête pour la réalisation d'analyses secondaires continuera de croître. À l'heure actuelle, les principaux véhicules utilisés pour répondre à cette demande sont les fichiers de données à grande diffusion et les sites sécurisés. Je prévois la poursuite de la recherche de méthodes de protection contre la divulgation de renseignements personnels et de l'étude des effets de ces méthodes sur les estimations d'enquête. Le nombre d'analystes secondaires effectuant leurs analyses en direct augmentera sans doute rapidement, et les analyses fourniront non seulement les estimations requises, mais aussi les erreurs-types de ces estimations calculées en tenant compte du plan de sondage complexe de l'enquête. À titre d'exemple de cette approche générale, je mentionnerai le système d'analyse de données en ligne disponible à l'heure actuelle pour certaines enquêtes dans les archives de données sur les sciences sociales de l'Inter-University Consortium for Political and Social Research des États-Unis.

Modes de collecte des données. Aujourd'hui, les ordinateurs sont utilisés à grande échelle de multiples façons, pour la collecte des données d'enquête (voir, p. ex., Couper et coll., 1998). D'autres progrès techniques en informatique et en communications donneront vraisemblablement lieu à des changements dans les méthodes de collecte des données au cours de la prochaine décennie (Nathan, 2001). Ainsi, l'interview enregistrée assistée par ordinateur (IEAO) vient de faire son apparition. Les progrès réalisés dans le domaine des ordinateurs de poche, ou ordinateurs tablettes, s'avéreront particulièrement intéressants pour la collecte des données dans certaines circonstances, comme les enquêtes auprès de populations nomades et les interviews au pas de la porte. Une fois perfectionnées, les méthodes de reconnaissance automatique de la parole auront plusieurs applications, la messagerie textuelle au moyen de téléphones mobiles trouvera certains usages et les méthodes de saisie de données au moyen d'un téléphone à clavier pourraient devenir plus populaires.

En revanche, certains changements risquent de nuire à la recherche en sondages au lieu d'offrir des possibilités d'amélioration. L'usage croissant de téléphones mobiles pour remplacer le service téléphonique conventionnel à fil pose des problèmes, du moins aux États-Unis, puisqu'une personne qui reçoit un appel sur un téléphone mobile doit payer l'appel. Les répondeurs automatiques et l'identification de la ligne appelante sont d'autres causes d'inquiétude, en raison de leurs effets négatifs éventuels sur les taux de réponse. La capacité de garder le même numéro de téléphone à un nouvel emplacement causera vraisemblablement des problèmes lors des enquêtes locales, puisque les numéros de téléphone pour un secteur donné ne seront pas limités à certains codes de bureau. Enfin, les effets qu'aura sur la participation aux enquêtes téléphoniques aux États-Unis l'introduction de la liste « Ne pas appeler » destinée aux télévendeurs ne sont pas certains : ils pourraient être nuisibles ou bénéfiques.

À l'heure actuelle, les enquêtes sur Internet et les enquêtes par courriel suscitent un vif intérêt et font l'objet de nombreux travaux de recherche, et leur utilisation augmentera sans aucun doute considérablement dans l'avenir, tant pour les enquêtes auprès des entreprises que pour les enquêtes sur la population. Ces méthodes offrent des possibilités d'utiliser des formes innovatrices de communication (p. ex., l'affichage d'images) et de faire des vérifications directes qui continueront sans aucun doute de faire l'objet d'études assidues. Pour les enquêtes-entreprises, la question fondamentale est de savoir comment persuader les répondants d'utiliser ces modes de réponse plutôt que les questionnaires papier et crayon. Le secret pourrait être de rendre la méthode informatique plus conviviale pour le répondant, autrement dit, peut-être d'adapter l'instrument de collecte des données de sorte qu'il concorde avec les dossiers de l'entreprise. Pour les enquêtes-ménages, le problème fondamental est la pénétration limitée du raccordement à Internet à domicile. Certaines personnes travaillant dans ce domaine semblent être séduites par le fait que la nouvelle technologie permet de générer un grand nombre de réponses à faible coût, en une période très courte. Elles ne tiennent pas compte du manque de représentativité des échantillons ainsi obtenus. Cette erreur a été commise lors du Literary Digest Poll en 1936 et l'on ne devrait pas oublier les leçons tirées de ce sondage.

J'estime qu'au moins pendant un certain nombre d'années, la collecte de données par Internet dans le cadre des enquêtes-ménages devrait demeurer l'un des modes de collecte proposé parmi d'autres dans des enquêtes à mode de collecte mixte auprès d'un échantillon probabiliste de l'ensemble de la population. Des études devraient être réalisées dans ces conditions d'utilisation en vue de déterminer le moyen d'intégrer le mode Internet et de s'assurer que des données comparables soient recueillies selon les divers modes de collecte offerts aux répondants.

D'autres formes de collecte des données qui, selon moi, prendront de l'ampleur dans l'avenir consistent à passer des déclarations faites par les répondants aux données d'observation. À titre d'exemple, je mentionnerai l'installation d'instruments au domicile des ménages pour enregistrer le niveau de pollution environnementale, l'utilisation d'enregistrements sur bande vidéo pour observer les interactions parents-enfants, le prélèvement d'échantillons de sang, de cheveux et d'urine pour le dépistage des drogues, la collecte de données d'observation sur les quartiers où vivent les ménages échantillonnés, et l'utilisation de systèmes de positionnement global (GPS) pour déterminer l'emplacement exact d'une unité échantillonnée aux fins du couplage à d'autres données.

Non-réponse. Il semble évident que la population générale soit de moins en moins disposée à collaborer aux enquêtes. De plus grands efforts en vue d'obtenir la coopération des personnes sélectionnées, particulièrement le suivi de celles qui n'étaient pas à leur domicile, ont atténué les effets sur les taux de réponse, mais ceux-ci sont malgré tout à la baisse (Groves et Couper, 1998). En raison des préoccupations suscitées par cette diminution des taux de réponse, de nombreuses études ont été entreprises en vue de rechercher des moyens d'obtenir la coopération du public, d'une part, et de corriger les données pour la non-réponse dans l'analyse, d'autre part (voir, p. ex., Groves et coll., 2002). Je crains personnellement que la volonté de collaborer aux enquêtes ne continue de s'éroder dans l'avenir.

Aux États-Unis, les taux de réponse aux enquêtes téléphoniques auprès des ménages par la méthode de composition aléatoire (CA) sont déjà tombés à un niveau qui menace sérieusement la qualité des estimations produites d'après les données de ces enquêtes. Une réduction supplémentaire risque de rendre cette forme d'échantillonnage inacceptable pour de nombreuses enquêtes. Le cas échéant, la mise au point d'autres méthodes pourraient être nécessaire afin de retenir les avantages économiques de l'interview téléphonique, particulièrement pour le sondage de populations rares. Au Royaume-Uni, étant donné le faible taux de réponse aux enquêtes téléphoniques, Collins et Sykes (2003) préconisent d'abandonner l'échantillonnage probabiliste au profit d'échantillons qu'il est convenu d'appeler « access panels » de personnes disposées à participer et de recueillir les données auprès de celles-ci selon le mode qu'elles préfèrent. Comme le souligne Nathan (2003) dans sa discussion de l'article de Collins et Sykes, cette prescription fait resurgir les problèmes du Literary Digest Poll de 1936.

En Allemagne, les faibles taux de réponse aux enquêtes-ménages sont l'une des raisons qui ont poussé le Bureau fédéral de la statistique à examiner l'utilisation d'access panels sélectionnés à partir de son micro-recensement. Ces panels comprendraient un échantillon de personnes indiquant qu'elles sont disposées à participer aux enquêtes (Körner et Nimmergut, 2003). L'un des avantages de cette approche est que les données du micro-recensement sont disponibles aussi bien pour les personnes qui sont disposées à participer aux enquêtes que pour celles qui ne le sont pas, ainsi que pour les répondants et les non-répondants à une enquête particulière. Ces données peuvent être utilisées pour essayer de corriger pour la non-réponse et pour évaluer le biais qu'elle engendre. D'autres types d'access panels ont également été utilisés, comme les panels sur Internet où on demande aux ménages participants de répondre à une série d'enquêtes. Pour que l'échantillon d'une enquête fondée sur un access panel soit de haute qualité, ce dernier doit être sélectionné par des méthodes scientifiques d'échantillonnage probabiliste à partir d'une base de sondage dont le taux de couverture est élevé et les encouragements à participer doivent être suffisants pour obtenir un taux de réponse élevé lors du recrutement initial du panel, puis un taux de réponse élevé lors de l'enquête en question.

Autres points. Les méthodes d'échantillonnage sont, selon moi, bien établies et je ne m'attends pas à de grands changements au cours de la prochaine décennie. De nouvelles techniques d'échantillonnage pourraient découler de nouvelles méthodes de collecte des données, mais il s'agira vraisemblablement de variantes des techniques existantes. Parmi les progrès auxquels on peut s'attendre, je mentionnerai la recherche de meilleures méthodes pour traiter la non-réponse partielle multivariée et l'estimation de la variance en présence de données imputées, ainsi que la simplification des méthodes d'estimation de la variance en cas de plan d'échantillonnage complexe grâce à l'intégration de l'estimation de la variance dans les logiciels d'analyse standards.

La réduction des erreurs non dues à l'échantillonnage, particulièrement les erreurs de réponse, demeure l'objectif le plus important en recherche en sondages. Il est absolument nécessaire de poursuivre les travaux de recherche sur les méthodes de collecte des données et la conception des questionnaires, y compris les travaux suscités par le mouvement des aspects cognitifs de la méthodologie d'enquête (Sirken et coll., 1999; Tourangeau, Rips et Rasinski, 2000). Bien qu'on puisse s'attendre à des progrès, des percées importantes dans ce domaine complexe semblent peu probables. L'application des théories de diverses sciences sociales peuvent nous aider à comprendre certains problèmes. Ainsi, la recherche sur la mémoire en psychologie peut nous aider à comprendre les erreurs dans les réponses aux questions de remémoration, la recherche en linguistique peut nous aider à comprendre les effets de l'énoncé des questions et les théories sociologiques des interactions entre personnes peuvent nous aider à comprendre les motivations de la réponse aux enquêtes. Toutefois, la question de savoir si cette compréhension peut être convertie en prescriptions permettant d'améliorer considérablement les méthodes d'enquête demeure ouverte.

Le concept de l'erreur totale d'enquête a été introduit il y a un certain nombre d'années pour mettre l'accent sur le fait que des compromis doivent être réalisés entre un certain nombre de sources d'erreurs afin d'obtenir le meilleur plan de sondage possible. Récemment, une attention plus grande a été accordée à ce concept et il est vraisemblable qu'il continue d'en être ainsi. Biemer et Lyberg (2003) discutent de la qualité des enquêtes en ce qui concerne ces sources d'erreurs, les compromis entre elles et les processus d'enquête que l'on peut utiliser pour contrôler les erreurs de sources diverses.

Une vue plus générale de la qualité des données statistiques a été articulée récemment par un certain nombre d'organismes statistiques. Par exemple, Brackstone (1999) a décrit les dimensions suivantes de la qualité : pertinence, exactitude, actualité, accessibilité, intelligibilité et cohérence (voir aussi Statistique Canada, 2003). Cette vue générale de la qualité, qui tient compte des compromis nécessaires entre les diverses dimensions, fournit un cadre intéressant pour la planification et l'évaluation.

La dimension d'actualité des données est importante pour de nombreuses enquêtes. Le besoin de données à jour mène souvent à la publication de résultats provisoires, suivie par des révisions ultérieures, dans le cas des enquêtes-entreprises. On aurait pu s'attendre à ce que l'introduction des ordinateurs dans toutes les phases des opérations d'enquête accélère fortement l'entièreté du processus. Il en est ainsi dans une certaine mesure, mais l'effet a été partiellement compensé par la complexité plus grande des enquêtes contemporaines, elle-même rendue possible par les ordinateurs. Il semble qu'on puisse en faire davantage en ce qui concerne l'actualité des données, grâce à une intégration plus poussée des processus d'enquête. Je pense que nous assisterons probablement à ce genre de progrès dans les prochaines années.

4. LA PROFESSION DE LA RECHERCHE EN SONDAGES

Il y a environ un quart de siècle, l'ensemble des connaissances sur les méthodes de recherche en sondages était suffisamment limité pour qu'un spécialiste des enquêtes arrive à maîtriser raisonnablement tous les aspects du sujet. Les progrès importants survenus depuis ont entraîné le remplacement du méthodologiste généraliste par une équipe de spécialistes pour la réalisation d'une enquête. L'équipe doit compter des statisticiens d'enquête, des méthodologistes d'enquête, des gestionnaires d'enquête et des informaticiens. (Naturellement, toute enquête nécessite aussi la participation de spécialistes du domaine visé par l'enquête.) L'expansion de l'ensemble de connaissances et la spécialisation résultante ont une incidence sur la structure de la profession et la formation dont je discute brièvement à la présente section.

La profession. Les origines de la recherche en sondages remontent aux nombreuses disciplines fondamentales qui ont recouru à des méthodes d'enquête pour réaliser leur recherche. Comme les méthodes d'échantillonnage étaient d'une importance capitale, les sociétés statistiques sont devenues un point de concentration. À mesure que le domaine de la statistique a pris de l'expansion, on a assisté à une spécialisation et des sections spéciales consacrées à la recherche en sondages ont été créées par les sociétés statistiques (p. ex., l'Association internationale des statisticiens d'enquête (AISE) de l'Institut international de statistique, la Section on Survey Research Methods (SRM) de l'American Statistical Association et la Social Statistics Section de la Royal Statistical Society). Ces sections ont embrassé tous les aspects de la méthodologie d'enquête, mais, étant donné la vocation des sociétés auxquelles elles appartenaient, elles se sont évidemment concentrées en grande partie sur les aspects statistiques des

enquêtes. D'autres sociétés, comme l'American Association for Public Opinion Research (AAPOR), se sont intéressées davantage aux aspects non statistiques de la méthodologie d'enquête et à la gestion d'enquête. Au Royaume-Uni, l'Association for Statistical Computing étudie les nombreuses utilisations des ordinateurs en recherche en sondages.

Les groupes qui étudient divers aspects des méthodes d'enquête sont tous très actifs et contribuent considérablement à l'expansion du domaine. Cependant, cette subdivision de la recherche en sondages en ses diverses disciplines n'est pas idéale. La recherche en sondages demande la collaboration étroite d'équipes multidisciplinaires et le partage des idées entre les diverses disciplines est essentiel. Certaines collaborations fort appréciées et très précieuses existent, comme celle entre la SRM, l'AISE et l'AAPOR en vue d'organiser des conférences internationales, mais je pense que la mise en place d'une structure plus officielle en vue de réunir les spécialistes de la recherche en sondages de toutes les disciplines serait bénéfique aujourd'hui et peut-être encore plus dans l'avenir. De surcroît, j'estime que ce genre de structure devrait être internationale, en harmonie avec la façon dont le domaine s'est développé. Je tiens à insister sur le fait que cette société internationale de la recherche en sondages proposée devrait s'ajouter aux sociétés existantes. Les chercheurs appartenant à chacune des disciplines de la recherche en sondages doivent entretenir des rapports étroits avec ceux qui travaillent dans d'autres domaines de leur discipline afin de profiter des progrès que ceux-ci réalisent. Un modèle de société internationale de la recherche en sondages fort séduisant consisterait à donner à cette société la forme d'une collaboration officielle évolutive entre sociétés ou sections existantes appartenant à diverses disciplines dans un certain nombre de pays.

Formation. Il existe aujourd'hui, pour chaque domaine de la recherche en sondages, un grand ensemble de connaissances que les méthodologistes travaillant dans ces domaines devraient maîtriser. Partout dans le monde, il y a pénurie de statisticiens d'enquête et de méthodologistes d'enquête possédant une formation approfondie et les possibilités de formation sont limitées. La plupart des universités sont mal équipées pour enseigner les méthodes de recherche en sondages en tant que sujet multidisciplinaire fondé sur une combinaison de théories et de pratiques. Néanmoins, la création de programmes de maîtrise dans quelques universités est un signe encourageant. J'espère qu'un plus grand nombre de ces programmes verront le jour au cours de la prochaine décennie.

Les programmes de maîtrise en recherche en sondages sont, certes, importants, mais je crains que, du moins dans le domaine de la statistique d'enquête, on ne puisse s'attendre à ce qu'ils enseignent dans les détails requis toutes les méthodes dont les statisticiens d'enquête ont besoin pour accomplir leur travail. Comme j'en ai discuté plus en détail ailleurs (Kalton, 2002), j'estime qu'un programme de maîtrise en statistique d'enquête devrait comprendre des cours sur la théorie et les méthodes statistiques, la théorie et la pratique de l'échantillonnage, les méthodes générales de recherche en sondages, ainsi que l'acquisition de compétences en calcul et en communications. Idéalement, le programme devrait permettre aux étudiants de bien se familiariser avec les méthodes courantes et d'acquérir les connaissances de base nécessaires pour l'apprentissage de nouvelles méthodes qui verront le jour durant leur carrière. Les méthodes de pointe utilisées aujourd'hui à grande échelle en statistique d'enquête nécessitent une connaissance fondamentale solide de la théorie et des méthodes statistiques, ainsi que l'enseignement détaillé de nombreux sujets abordés en statistique d'enquête. Je pense, personnellement, que ces exigences ne peuvent être satisfaites adéquatement même par un programme de maîtrise très exigeant. Les étudiants de maîtrise peuvent être exposés à une vaste gamme de sujets en statistique d'enquête, mais pas avec la profondeur requise pour une application pratique éclairée.

Un moyen éventuel de résoudre ce problème consisterait à offrir la formation de recherche en sondages et en statistique d'enquête à un stade plus précoce des études, par exemple au niveau du programme de premier cycle, peut-être sous forme d'un grade collectif. Notre approche consisterait à donner la possibilité d'obtenir une formation plus approfondie après l'achèvement d'un programme de maîtrise. (J'exclus ici les étudiants qui poursuivent un programme de doctorat.) Dans les grands organismes, comme Statistique Canada et l'Institut national de la statistique et des études économiques (INSEE) en France, des cours de formation permanente peuvent être, et sont, offerts à l'interne, mais la plupart des statisticiens d'enquête exerçants ne bénéficient pas d'une telle approche. Les brefs cours intensifs d'une demi-journée et à deux journées offerts lors des réunions professionnelles et par certaines universités sont fort utiles, mais sont généralement limités à une bonne vue d'ensemble du sujet étudié. Les participants font rarement le travail supplémentaire nécessaire pour assimiler et comprendre complètement la matière qui leur a été présentée. Par conséquent, j'entrevois un besoin de programmes officiels de formation permanente à l'intention des statisticiens d'enquête, assortis d'une forme d'agrément par les universités ou une société professionnelle. Un programme pourrait, par exemple, comprendre un ensemble de cours sur des sujets

particuliers, avec des devoirs de lecture et d'exercices, et une forme d'examen sur lequel fonder l'octroi d'un certificat attestant que chaque cours a été suivi avec succès. Les cours seraient étalés sur une certaine période, dont la longueur dépendrait de la matière à couvrir. Les méthodes d'enseignement à distance pourraient être utilisées pour permettre la participation à partir de n'importe quel emplacement. La liste qui suit illustre le genre de sujets qui pourraient être inclus dans les programmes : correction pour la non-réponse par calage et repondération; méthodes de vérification et d'imputation; conception et estimation pour les enquêtes qui s'étendent dans le temps; modèles d'erreur de mesure, estimation pour petites régions, échantillonnage pour les enquêtes auprès des établissements; conception des études par observation et analyse des résultats; techniques de limitation de la divulgation. Les étudiants en maîtrise pourraient étudier ces sujets dans une certaine mesure durant le programme de deuxième cycle, mais il est peu probable que leur exposition soit suffisante pour leur permettre de les traiter en pratique. En outre, les progrès rapides qui ont eu lieu récemment en statistique d'enquête se poursuivront sûrement dans l'avenir. Des programmes de formation permanente sont nécessaires pour permettre aux statisticiens d'enquête exerçants de se tenir au courant des faits nouveaux.

5. CONCLUSION

À part la menace de non-réponse croissante, l'avenir de la recherche en sondages semble fort prometteur. La demande de données statistiques pour l'élaboration des politiques et pour la recherche empirique augmente depuis de nombreuses années et cette tendance à la hausse se poursuivra vraisemblablement. L'accroissement de la gamme de domaines spécialisés pour lesquels des données d'enquête sont nécessaires, la complexité croissante des analyses que les analystes souhaitent réaliser et les progrès technologiques qui ont lieu dans l'industrie des télécommunications nécessiteront la mise en place d'un programme dynamique permanent de recherche méthodologique. Je prévois une augmentation de la demande de spécialistes de la recherche en sondages possédant une formation solide pour exécuter ces activités. Les défis qui se posent sont, d'abord, de savoir comment attirer des gens dans ce domaine excitant, puis, comment offrir la formation nécessaire pour leur permettre d'appliquer comme il convient les méthodes de pointe utilisées de façon généralisée et de développer des méthodes innovatrices pour résoudre les nouveaux problèmes qui se poseront sans aucune doute.

REMERCIEMENTS

Je remercie mes collègues Mike Brick, David Cantor et Brad Edwards pour les discussions constructives durant la préparation de cet article. Les opinions exprimées ici n'engagent, naturellement, que moi.

RÉFÉRENCES

- Biemer, P.P. et Lyberg, L.E. (2003), *Introduction to Survey Quality*, New York: Wiley.
- Brackstone, G. (1999), "Managing Data Quality in a Statistical Agency", *Survey Methodology*, 25, 139-149.
- Collins, M. et Sykes, W. (2003), "What is the Future for Telephone Surveys?" *Proceedings of the Berlin Session of the International Statistical Institute* (CD-ROM).
- Couper, M.P., Baker, R.P., Bethlehem, J., Clark, C.Z.F., Martin, J., Nicholls, W.L., et O'Reilly, J.M. eds. (1998), *Computed Assisted Survey Information Collection*, New York: Wiley.
- Groves, R.M., et Couper, M.P. (1998), *Nonresponse in Household Interview Surveys*, New York: Wiley.
- Groves, R.M., Dillman, D.A., Eltinge, J. L., et Little, R.J.A. eds. (2002), *Survey Nonresponse*, New York: Wiley.
- Harkness, J.A., Van de Vijver, F.J.R., et Mohler, P.P., eds. (2003), *Cross-Cultural Survey Methods*, New York: Wiley.

- Holt, D. (2003), "Methodological Issues in the Development and Use of Statistical Indicators for International Comparisons", *Survey Methodology*, 29, 5-17.
- Kalton, G. (2000), "Developments in Survey Research in the Past 25 Years", *Survey Methodology*, 26, pp. 3-10.
- Kalton, G. (2002), "Training Needs for Survey Statisticians in Developed and Developing Countries", 2002 *Proceedings of the American Statistical Association* (CD-ROM).
- Körner, T. et Nimmergut, A. (2003), "An Access Panel for Official Statistics," contributed paper at the Berlin Session of the International Statistical Institute.
- Nathan, G. (2001), "Telesurvey Methodologies for Household Surveys—A Review and Some Thoughts for the Future", *Survey Methodology*, 27, 7-31.
- Nathan, G. (2003), "Discussion of New Challenges with Future Tele-Surveys", *Proceedings of the Berlin Session of the International Statistical Institute* (CD-ROM).
- Rao, J.N.K. (2003), *Small Area Estimation*, New York: Wiley.
- Sirken, M.G., Herrmann, D.J., Schechter, S., Schwarz, N., Tanur, J.M. et Tourangeau, R., eds. (1999), *Cognition and Survey Research*, New York: Wiley.
- Statistics Canada (2003), *Statistics Canada Quality Guidelines*, Ottawa: Ministry of Industry. <http://www.statcan.ca/cgi.bin/downpub/freepub.cgi>.
- Tourangeau, R., Rips, L.J., et Rasinski, K.A. (2000), *The Psychology of Survey Response*, Cambridge, U.K.: Cambridge University Press.
- Verma, V. (2002), "Comparability in International Survey Statistics," International Conference on Improving Surveys, Copenhagen, August, 2002, http://www.icis.dk/ICIS_papers/Keynote3_0_5.pdf.