

**AVANTAGES ET DÉSAVANTAGES DU SONDAGE EN LIGNE
COMME MÉTHODE DE COLLECTE DE DONNÉES : UNE
REVUE DE LA LITTÉRATURE**

Marie-Ève GINGRAS et Hélène BELLEAU

Inédit / *Working paper*, n° 2015-02

**CENTRE
URBANISATION CULTURE SOCIÉTÉ**

**AVANTAGES ET DÉSAVANTAGES DU SONDAGE EN LIGNE COMME
MÉTHODE DE COLLECTE DE DONNÉES : UNE REVUE DE LA
LITTÉRATURE**

Marie-Ève GINGRAS et Hélène BELLEAU

Institut national de la recherche scientifique
Centre - Urbanisation Culture Société
Montréal

Mai 2015

Marie-Ève Gingras et Hélène Belleau

Centre - Urbanisation Culture Société de l'INRS

helene.belleau@ucs.inrs.ca

Centre - Urbanisation Culture Société
Institut national de la recherche scientifique
385, rue Sherbrooke Est
Montréal (Québec) H2X 1E3

Téléphone : (514) 499-4000

Télécopieur : (514) 499-4065

www.ucs.inrs.ca

Ce document peut être téléchargé sans frais à :

<http://www.ucs.inrs.ca/ucs/publications/collections/inedits>

Abstract

In recent years, a growing number of researchers used online surveys as method of collecting data. This document proposes a review of the recent literature concerning the advantages and the disadvantages of the use of this method. Several advantages were found by the researchers: moderate cost, speed of the data collection, decrease of the errors during data capture, freedom of the respondent, etc. Regarding the disadvantages, the literature underlines mainly the difficulties related to the control of the sampling and the limits which are inherent to it (representativeness, confidentiality, anonymity, auto-selection bias, etc.). This document is completed by a synthetic board of the main arguments evoked in the studied scientific papers.

Key Words:

Online Surveys; data collection; methodological limits; Internet

Résumé

Dans le paysage de la recherche actuelle, de plus en plus de chercheurs utilisent le sondage en ligne comme méthode de collecte de données. Ce document propose une revue de la littérature récente concernant les avantages et désavantages de l'utilisation de cette méthode. Plusieurs avantages ont été relevés par les chercheurs : faible coût, rapidité de la collecte des données, diminution des erreurs lors de la saisie des données, plus grande liberté du répondant, etc. Du côté des désavantages, la littérature souligne principalement les difficultés liées au contrôle de l'échantillonnage et les limites qui y sont inhérentes (représentativité, confidentialité, anonymat, biais d'auto-sélection, etc.). Ce document est complété par un tableau synthétique des principaux arguments évoqués dans les textes étudiés.

Mots clés :

Sondage en ligne; collecte de données; limites méthodologiques; Internet

LES SONDAGES EN LIGNE : L'EFFICACITÉ ET LA SIMPLICITÉ À MOINDRE COÛT

Les avantages et les limites des sondages en ligne doivent être évalués à l'aune des diverses méthodes de collecte de données qui ont été retenues car celles-ci peuvent être de natures très différentes. Certains sondages, par exemple, sont diffusés à grande échelle et circulent de manière virale librement dans le cyber espace. Les chercheurs utilisent parfois des relais de diffusion, tels que des associations, des médias, des listes d'individus prêts à faire circuler le sondage dans leurs réseaux. Ces sondages peuvent permettre de rejoindre des milliers de répondants en très peu de temps mais ne sont pas représentatifs de la population, comme nous le verrons. Une autre forme de sondage en ligne consiste à avoir recours à une liste fermée d'adresses courriels, celle d'une association ayant des membres, par exemple. L'accès au sondage peut alors être limité à ces derniers par le biais de mots de passe ou d'un identifiant. Enfin, les chercheurs sont de plus en plus nombreux à avoir recours à des panels de centaines voire de milliers de répondants potentiels sélectionnés de manière à assurer une certaine « représentativité » de la population (Lindhjem et Navrud, 2011).

La littérature sur le sujet révèle que les sondages en ligne présentent de nombreux avantages au plan opérationnel. Leurs coûts d'utilisation sont faibles (Deutskens *et al.*, 2004 : 20; Wang *et al.*, 2013 : 1007; Ganassali, 2008 : 21; Bethlehem, 2008 : 1; Stephenson et Crête, 2011 : 27; Bigot *et al.*, 2010 : 20), et ils permettent de collecter rapidement les données (Deutskens *et al.*, 2004 : 21; Ganassali, 2008 : 21; Bethlehem, 2008 : 1 ; Stephenson et Crête, 2011 : 27; Bigot *et al.*, 2010 : 20). L'envoi de rappel est facilité par l'utilisation du courriel (Deutskens *et al.*, 2004 : 21), et le mode Web élimine le risque d'erreur lors de la saisie des données, plus fréquent dans le mode postal (Stephenson et Crête, 2011 : 27).

Contrairement aux sondages qui nécessitent la présence physique des répondants, les sondages en ligne ont aussi l'avantage de permettre aux sondés d'amorcer les questionnaires au moment et dans le lieu de leur choix, et de le compléter dans un délai qui leur convient (Lindhjem & Navrud, 2011 : 12, Bigot *et al.*, 2010 : 21). De plus, ils constituent généralement une méthode de cueillette de données simple et accessible pour les chercheurs (Bigot *et al.*, 2010 : 6). Grâce aux envois par courriel, ce type de méthode de collecte de données permet aussi de couvrir un bassin géographique de répondants plus large (Wang *et al.*, 2013 : 1007; Bethlehem 2008 : 91 ; Bigot *et al.*, 2010 : 21).

Les limites de l'échantillonnage

La revue de la littérature démontre que les plus grands désavantages du sondage en ligne sont liés à l'échantillonnage ou ce que les chercheurs désignent comme l'erreur de couverture :

Toute étude en ligne qui vise à représenter la population générale connaît un problème de couverture (sauf si l'on équipe les individus non équipés) puisque la population n'est pas entièrement équipée. (Bigot et al., 2010 : 30)

D'une part, les sondages en ligne sont généralement non probabilistes, c'est-à-dire que leur échantillon n'est pas contrôlé : on ne sait pas qui y répond ou qui peut y répondre. Un échantillon probabiliste est constitué d'individus choisis et ciblés de manière à représenter proportionnellement la population générale et les minorités et « chaque individu doit avoir exactement la même chance que les autres de participer à l'enquête » (Bigot *et al.*, 2010 : 24). Les répondants des sondages en ligne ne constituent habituellement pas un échantillon de la population générale, ils sont donc peu fiables. (Lindhjem et Navrud, 2011 : 20; Ganassali, 2008 : 21; Bethlehem, 2008 : 2; Dussaix, 2009 : 152; Yeager *et al.*, 2011 : 737, Stephenson et Crête, 2011 : 27).

D'autre part, les résultats sont limités par le fait qu'ils ne s'adressent qu'aux individus connectés à l'Internet, et que les aînés, les personnes vivant en région éloignée, les minorités visibles et les personnes à faibles revenus sont souvent sous-représentés (Lindhjem et Navrud, 2011 : 20; Bethlehem, 2008 : 3). Certains chercheurs nomment cet écart le *digital divide* (Rivara *et al.*, 2011 : 976; Fricker *et al.*, 2005 : 388).

...en regard des taux de réponse, point de consensus

Les chercheurs ne s'entendent pas sur la qualité des taux de réponse des sondages en ligne, c'est-à-dire sur le nombre de personnes ayant complété le sondage comparé au nombre de personnes contactées. Pour certains, ces taux seraient soit moins élevés que ceux des autres méthodes (Lindhjem & Navrud, 2011 : 19 ; Shin *et al.*, 2012 : 223), soit semblables (Ganassali, 2008 : 21), soit supérieurs (Denscombe, 2006 : 252). Les chercheurs Heerwegh et Loosveldt (2009) se sont intéressés aux facteurs pouvant influencer les taux de réponse des sondages en ligne. Plusieurs résultats d'études ont établi que les éléments techniques (connexion au sondage, nom d'utilisateur, mot de passe, etc.), les récompenses, les rappels, le contenu du courriel d'invitation, le statut et la photo du chercheur et la personnalisation de l'invitation (Couper *et al.*, 2001 ; Crawford *et al.*, 2001; Heerwegh et Loosveldt, 2002; Crawford *et al.*, 2001; Guéguen et Jacob, 2002; Porter et Withcomb, 2003; Matthijs et Loosveldt, 2005; Cook *et al.*, 2000 : 833; Pearson et Levine, 2003¹; Tourkin *et al.*, 2005²) sont des facteurs pouvant influencer la participation au sondage en ligne.

Porter et Withcomb (2003 : 586) se sont pour leur part penchés sur l'impact du type de prise de contact sur les taux de réponse. Les résultats de leur étude ont démontré que l'autorité du chercheur (s'il est connu ou s'il provient par exemple d'une grande université) n'a que peu d'impact sur les taux de réponse. La mention selon laquelle le répondant a été choisi parmi d'autres ainsi que l'ajout d'une date limite pour remplir le questionnaire auraient pour leur part un impact mineur, mais significatif. L'étude de Pan *et al.* (2013 : 64) démontre quant à elle que la

¹ Cités dans Heerwegh et Loosveldt, 2009 : 181.

² Cités dans Farrel et Petersen, 2010 : 121.

position hiérarchique du chercheur responsable du sondage aurait un impact positif sur les taux de réponse, surtout s'il a un nom « américain ». À l'inverse, les noms à consonance étrangère auraient tendance à les diminuer (Pan *et al.*, 2013 : 64).

Pour Bigot *et al.* (2010 : 51), le succès des invitations au sondage dépend de trois facteurs : le lieu où il est diffusé (blogues, sites Internet, etc.), sa visibilité et le moment où il est activé. Ils soulignent aussi l'importance de la page de présentation, de la divulgation des coordonnées du centre ou de l'institut, de la promesse de la confidentialité et des paramètres qui l'encadrent, d'un espace pour les commentaires et les échanges ainsi que de l'utilisation d'un visuel dynamique et accueillant. Les chercheurs ont aussi observé une disparité entre le nombre de « clics » et le nombre de réponses obtenues, qu'ils expliquent par l'allure du sondage ainsi que l'absence d'intérêt à y répondre.

Heerwegh et Loosveldt se sont plutôt intéressés à la théorie du comportement planifié³ pour mieux comprendre ce qui peut influencer un individu à répondre à un sondage en ligne. L'étude a été menée par le biais d'un sondage postal, envoyé à des répondants possédant une adresse courriel. Les résultats de l'étude démontrent qu'une récompense prépayée inciterait davantage les individus à répondre au sondage, en raison d'une certaine obligation morale liée au fait d'accepter un prix à condition de rendre le service qui lui est lié. Les chercheurs soulignent aussi l'importance de non seulement prévoir des procédures de connexion simples, mais aussi de faire en sorte qu'elles apparaissent comme telles aux yeux des répondants. Le fait de communiquer clairement l'utilité du sondage aurait aussi un impact positif sur la coopération des sondés.

Le profil des gens branchés... un biais supplémentaire ?

Ce type de collecte de données exige une familiarité avec les ordinateurs et l'informatique et demande une certaine dextérité dans la manipulation du matériel (Shin *et al.*, 2012 : 223). Différentes études soulignent d'ailleurs qu'il devient de plus en plus difficile pour les répondants de distinguer les courriels des pourriels (*spam*) ou de la publicité des envois sérieux ; certains se montrent donc suspicieux face aux courriels de sollicitation. D'autres affichent également une certaine méfiance liées à la sécurité informatique et craignent les virus (Porter et Whitcomb, 2003 : 579 ; Frippiat et Marquis, 2010 : 48). Ces éléments rendent encore plus difficile la quête de répondants et affectent assurément les taux de réponse. Les répondants aux sondages en ligne ont-ils pour autant un profil particulier ?

Selon Rivara et ses collègues, le fait d'avoir un courriel n'indique pas automatiquement l'aisance web des répondants, pas plus qu'il n'assure que ceux-ci sont des utilisateurs d'Internet (2011 : 976). Le biais d'autosélection joue par contre un rôle important dans la fiabilité des

³ Theory of planned behaviour (« the TPB postulates that behaviour is a function of salient information, ore beliefs, relevant to the behaviour »), Heerwegh et Loosveldt, 2009 :182.

résultats, les répondants étant souvent plus intéressés par le sujet que la population générale (Lindhjem et Navrud, 2011 : 20; Ganassali, 2008 : 21; Bethlehem, 2008 : 2; Stephenson et Crête, 2011 : 27).

Plusieurs auteurs mentionnent que les répondants aux sondages en ligne sont habituellement de plus grands consommateurs et utilisateurs d'Internet, facteur qui a un impact sur le type d'échantillon (plus riche, plus scolarisé ou plus jeune par exemple) (Strabac, Z. et Aalberg, 2011 : 186-187; Stephenson et Crête 2011 : 27; Bosnjak *et al.*, 2013 : 355). Bosnjak *et al.* affirment, pour leur part, que les utilisateurs d'Internet sont généralement plus ouverts à l'expérimentation et plus extravertis, ce qui s'expliquerait en partie par l'âge moyen des répondants en ligne (2013 : 355). Le fait d'être connecté à l'Internet non seulement aurait une influence sur l'attitude des répondants au sujet des technologies en général, mais serait aussi fortement lié au patrimoine. Ce rapport à la technologie révélerait aussi une ouverture plus grande par rapport à certains enjeux (homoparentalité, travail des femmes, etc.) (Bigot *et al.*, 2010 : 9). Soulignons enfin que le caractère généralement non probabiliste des sondages en ligne les rend aussi vulnérables aux groupes d'intérêts lorsque le chercheur ne contrôle pas l'accès des répondants au sondage (Bethlehem, 2008 : 8).

La spécificité des réponses aux sondages en lignes

L'utilisation du sondage en ligne influencerait aussi la nature des données recueillies. D'abord, la qualité des données issues des sondages en ligne serait plus grande que celle des données issues des questionnaires postaux en raison d'un plus faible taux de non-réponse aux questions fermées et ouvertes (Shin *et al.*, 2012 : 223; Denscombe, 2006 : 252). Les répondants auraient moins tendance à laisser tomber des questions lorsque le sondage est sur Internet. Certaines études ont aussi démontré que les répondants ont tendance à rédiger de plus longues réponses aux questions ouvertes dans les questionnaires Web et à se dévoiler davantage que dans les formats postaux (Denscombe, 2006 : 252; Shaefer et Dillman, 1998⁴). D'autres avancent plutôt que les sondages en ligne tendent à défavoriser les questions ouvertes (moins de réponses) (Wolf *et al.*, 2008⁵).

Les réponses comportant un biais d'acceptabilité sociale seraient aussi moins présentes dans les questionnaires en ligne en raison du caractère anonyme de cette méthode (Dillman *et al.*, 2009 : 3; Beck *et al.*, 2009 : 3; Bigot *et al.*, 2010 : 6), et plus spécifiquement lorsque des sujets considérés comme sensibles par les répondants (tels que les problèmes personnels ou les comportements à risque) sont abordés (Wang *et al.*, 2013 : 1008; Lindhjem et Navrud 2011 : 12; Ganassali, 2008 : 21; Nagelhout *et al.*, 2010 : 9; Bigot *et al.*, 2010 : 39).

⁴ Cités dans Bigot *et al.*, 2010 : 22.

⁵ Cités dans Shin *et al.*, 2012 : 213.

La désirabilité sociale reflète la tendance des personnes interviewées à vouloir donner une bonne image d'eux-mêmes ou à répondre de la façon la plus proche possible de ce qu'ils considèrent être la norme sociale. (Bigot et al., 2010: 39)

Dans une étude intitulée « Is Web Interviewing a Good Alternative to Telephone Interviewing? Findings From the International Tobacco Control (ITC) Survey », le chercheur Nagelhout remarque que les répondants au sondage téléphonique se sont généralement montrés plus négatifs à propos de la cigarette que les répondants du sondage en ligne. Une étude visant à comparer les résultats d'une enquête téléphonique et d'une enquête Web a, par ailleurs, démontré d'importantes variations dans les réponses concernant le mode de vie, la santé et l'exercice (Greene *et al.*, 2008 : 245). L'étude menée par Stephenson et Crête (2011 : 27) démontre pour sa part qu'en raison du caractère plus anonyme des sondages en ligne, les répondants seraient moins gênés de choisir « je ne sais pas » ou « je ne désire pas répondre ». Enfin, Frippiat et Marquis soulignent que :

Le fait de répondre par Internet rend les individus moins sensibles au caractère intrusif des questions que s'ils répondaient par d'autres modes, et que cela leur permet de déclarer plus de comportements hors normes ou socialement indésirables (2010 : 63).

À l'inverse, quelques études soutiennent que le format Web des sondages engendrerait parfois une plus grande méfiance des individus par rapport à la confidentialité des données et aux nouvelles technologies en général, ce qui pourrait avoir l'effet recherché contraire et accroître les réponses socialement désirables (Wang *et al.*, 2013 : 1007; Lindhjem et Navrud, 2011 : 12; Frippiat et Marquis 2010 : 62). La présence d'un enquêteur pourrait aussi avoir un impact sur les réponses socialement désirables, les répondants ayant parfois tendance à répondre ce qui est « attendu » d'eux (Bigot *et al.*, 2010 : 11).

Qualité des réponses

Certaines études démontrent par ailleurs que les répondants des sondages en ligne présentent une plus faible tendance au *satisficing*, phénomène qui évoque l'idée selon laquelle un répondant, dans certaines circonstances (long questionnaire, déconcentration, désintérêt pour le sujet), se distancie du questionnaire tout en réduisant l'effort qu'il y met pour répondre correctement (Lindhjem et Navrud, 2011 : 38; Ganassali, 2008 : 21; Stephenson et Crête, 2011 : 27). Le terme se traduirait par « faible implication du répondant » (Frippiat et Marquis, 2010⁶). Heerwegh et Lossveldt (2008) décrivent l'effort nécessaire pour remplir un questionnaire comme suit :

Optimal question-answering takes a considerable amount of cognitive work. A respondent must interpret the meaning and intent of each question, retrieves relevant information from his or her memory, integrates that information into a summary

⁶ Cités dans Bigot *et al.*, 2010 : 41.

judgement, and then reports that judgement taking into account the provided response alternative. (Tourangeau, Rips et Rasinki, 2000, cités dans Heerwegh et Lossveldt, 2008 : 837)

Ce phénomène est influencé par différents facteurs : le degré de difficulté de la tâche demandée à un individu, les aptitudes que cela lui demande (sa *capabilité*) ainsi que sa motivation générale à y participer (Heerwegh et Lossveldt, 2008 : 837; Frippiat et Marquis, 2010 : 63). Le *satisficing* s'observe dans une grande quantité de « je ne sais pas » ou de questions non répondues, dans le choix répété de « mi-échelle » (cocher la case du milieu à chaque question le permettant), dans l'endossement du statu quo ainsi que dans le choix de la réponse la plus simple ou la plus acceptable sans réfléchir⁷ (Heerwegh et Lossveldt, 2008 : 838; Frippiat et Marquis, 2010 : 66).

Autrement qu'à travers l'analyse des données, il est possible de déceler le *satisficing* en intégrant un chronomètre permettant de connaître le temps de réponse. Plus le temps de réponse est court, plus la tendance de *satisficing* augmente (Bigot *et al.*, 2010 : 52). Selon Frippiat et Marquis (2010 : 67), il est possible d'atténuer l'impact du *satisficing* en mélangeant aléatoirement les choix de réponses. Il faut toutefois savoir que

Si cette possibilité semble utile pour tenter de neutraliser l'effet "premier vu, premier choisi", elle pose néanmoins question en ce qu'elle introduit comme une variation supplémentaire entre répondants [...] des individus ayant un même profil ne recevront pas exactement le même questionnaire » (Frippiat et Marquis 2010 : 67).

Il est possible d'atténuer ce phénomène en formulant des questions claires et concises (Bigot *et al.*, 2010 : 41). Bien que parfois difficile à distinguer des réponses socialement désirables, le désengagement serait plus observable dans les sondages en ligne que dans les entretiens menés avec un enquêteur (Bigot *et al.*, 2010 : 11). Ceci s'expliquerait par le fait que les intervieweurs savent capter et retenir l'attention des sondés tout en les motivant; ils favoriseraient aussi une meilleure compréhension des questions (Grenn et Tunstall, 1999⁸; Downes-Le Guin, *et al.*, 2012 : 2; Lindhjem et Navrud, 2011 : 12 ; Holbrook *et al.*⁹).

Holbrook et ses coauteurs affirment par ailleurs que les répondants aux sondages en ligne sont fréquemment soumis au *multitasking*, qui réfère à l'idée de faire plusieurs choses en même temps. Les répondants en ligne ont souvent plusieurs pages Web actives simultanément et cette réalité pourrait engendrer davantage de *satisficing* dans les sondages en ligne que dans d'autres types de sondages. De même, ces chercheurs avancent qu'il est plus difficile de remplir un sondage en ligne puisque cette action exige de comprendre un certain langage informatique en plus de savoir

⁷ « [C]ommonly observed effects are answering "don't know" or refusing (or generally more incomplete answers or item non-response), selecting the first reasonable response alternative, agreeing with assertions (acquiescence), non-differentiation (sticking to the same response category for a sequence of questions), endorsing status quo, "mental coin flipping" (random answers, if "I don't know is not offered as an option", choice of mi-points in rating scales, extremeness, etc. » (Lindhjem et Navrud, 2011 : 16)

⁸ Cités dans Lindhjem et Navrud, 2011 : 10.

⁹ Cités dans Heerwegh et Lossveldt, 2008 : 838.

manipuler le matériel technologique. En effet, le sondage face à face faciliterait le processus de réponse puisque les questions sont lues oralement, alors que le sondage en ligne demanderait plus d'efforts aux individus (Heerwegh et Lossveldt, 2008 : 838).

D'autres auteurs (Chang et Krosnicjk, 2010¹⁰; Skitka et Sargis, 2005¹¹) soutiennent toutefois l'inverse et affirment que les sondages en ligne diminueraient la tendance au désengagement ; ces études viennent confirmer la nécessité d'approfondir les connaissances à ce sujet. Enfin, dans leur recherche visant à comparer les résultats d'un sondage téléphonique à ceux d'un sondage Web, Greene *et al.* (2008 : 245) ont observé que les répondants avaient moins tendance à approuver (*agree*) les questions d'énoncés factuels dans le format Web que dans le format téléphonique, observation qui suggère que les résultats sont plus authentiques dans un format Web.

L'importance du format du questionnaire et de sa durée

En plus du *satisficing*, Bigot *et al.* soulignent que les effets de la fatigue, qui s'observent par « une variance des réponses plus faibles en fin de questionnaire » (2010 : 38), sont des éléments pouvant affecter la qualité des résultats. Ils mentionnent par ailleurs que les enquêtes en ligne devraient idéalement pouvoir se compléter dans un délai de 15 à 20 minutes (2010 : 38).

Sur le plan technique, certains chercheurs notent que la présence d'un support visuel intégré au sondage facilite la compréhension des questions tout en permettant d'en nuancer le propos (Dillman *et al.*, 2009 : 16; Lindhjem et Navrud, 2011 : 12; Stephenson et Crête, 2011 : 27). Citant les observations de variantes importantes dans les réponses selon le choix du visuel¹², Bigot *et al.* affirment qu'il faut toutefois encore être vigilant dans les conclusions à tirer à ce sujet. Par ailleurs, le dévoilement progressif du questionnaire (*pacing*) inhérent au format Web permet d'afficher progressivement les questions à venir dans le questionnaire. Contrairement au format papier où le répondant peut consulter rapidement l'ensemble des questions et choisir celles auxquelles il souhaite ou non répondre, le format Web « force » le sondé à répondre à une question à la fois avant de pouvoir consulter le reste du questionnaire (Frippiat et Marquis 2010 : 56). Les informations importantes devraient aussi se trouver en haut de la page plutôt qu'en bas, parce que les répondants y consacreront plus de temps (Bigot *et al.*, 2010 : 51). Enfin, le dévoilement progressif a aussi pour avantage de « limiter les efforts que le répondant doit fournir pour répondre au questionnaire » puisque les questions filtres lui permettent d'éviter de répondre à des questions moins pertinentes au fil de ses réponses (questions filtres) (Frippiat et Marquis 2010 : 56; Bigot *et al.*, 2010 : 22). Pour Frippiat et Marquis, ces éléments techniques permettent de diminuer les non-réponses liées au questionnaire.

¹⁰ Cités dans Stephenson et Crête, 2011 : 28.

¹¹ Cités dans Ganassali, 2008 : 21.

¹² Kenyon *et al.*, 2001 cités dans Bigot *et al.* : 2010 : 40.

L'enquête par Internet permet par ailleurs d'obtenir des données sur les non-réponses. Dans le cas des sondages postaux, les individus qui ne comprennent pas une question ou qui ne souhaitent pas répondre à toutes les questions ne renvoient pas le questionnaire s'il est incomplet. Dans le cas des sondages en ligne, il est possible pour le chercheur d'établir quelles questions se sont avérées problématiques par le nombre de fois où elles ont été écartées ou de constater à quel moment les répondants ont abandonné le questionnaire puisque ces informations sont à sa disposition. Celles-ci permettent d'une part, de mieux évaluer la qualité d'un questionnaire et, d'autre part, d'établir que « les non-réponses ne peuvent plus être considérées comme totalement aléatoires, puisqu'elles dépendent de la variable même qui fait l'objet de l'enquête » (Fripiat et Marquis, 2010 : 13).

CONCLUSION

Les limites liées aux coûts de production des grandes enquêtes, la moindre qualité des données du dernier recensement canadien mais aussi le fait que les enquêtes existantes ne couvrent pas tous les sujets d'intérêt des chercheurs les conduisent de plus en plus nombreux à se tourner vers d'autres méthodes, dont les sondages en ligne. Ceux-ci permettent à faible coût d'enquêter sur des phénomènes émergents, d'aborder des questions sociales sous de nouveaux angles en s'appuyant sur de vastes échantillons.

En comparaison des méthodes plus classiques, telles que les enquêtes téléphoniques et par courrier, les sondages en ligne présentent de très nombreux avantages mais aussi certains désavantages qui leurs sont spécifiques. En regard de l'échantillonnage et de la représentativité de la population, les cas de figure sont multiples. Les sondages en ligne sont généralement non probabilistes. Les problèmes de couverture ne se limitent plus désormais aux sondages en ligne cependant. En effet, les téléphones fixes sont de moins en moins courants dans les maisonnées, surtout chez les jeunes générations. Ainsi, les sondages téléphoniques probabilistes ne peuvent plus prétendre couvrir l'ensemble des ménages comme autrefois. L'usage d'internet, au contraire, gagne rapidement du terrain et s'avère prometteur à notre avis. Il est intéressant de noter qu'en 2012 au Québec, environ 95% des ménages avec enfant étaient branchés à l'Internet et la haute vitesse est devenue la norme (ISQ, 2013 : 78). Les personnes âgées, celles vivant seules ou en région éloignée sont cependant les moins branchées et c'est une limite importante à considérer dans le choix de cette méthode. Ainsi, la qualité et la pertinence d'un sondage en ligne variera selon la population à l'étude.

En regard de la qualité des réponses aux sondages en ligne, les avis semblent partagés mais plusieurs soulignent néanmoins que comparativement aux enquêtes téléphoniques et aux enquêtes postales, les répondants tendent à se dévoiler davantage devant l'ordinateur et à être moins biaisés dans leur réponse par soucis d'acceptabilité sociale. En somme, cette méthode de collecte de

données, si elle est loin d'être parfaite, constitue à l'heure actuelle un outil de recherche en plein développement qui apparaît très prometteur si ses limites méthodologiques sont considérées et s'il est utilisé avec rigueur et discernement.

BIBLIOGRAPHIE

- BAKER, R., J. M. BRICK, N. A. BATES, M. BATTAGLIA, M. P. COUPER, J. A. DEVER, K. J. GILE et R. TOURANGEAU. 2013. « Summary Report of the AAPOR Task Force on Non-probability Sampling ». *Journal of Survey Statistics and Methodology* 1 (2) : 90-143.
- BARRIOS, M., A. VILLARROYA, Á. BORREGO et C. OLLÉ. 2011. « Response Rates and Data Quality in Web and Mail Surveys Administered to Phd Holders ». *Social Science Computer Review* 29 (2) : 208-220.
- BETHLEHEM, J. 2008. *Peut-on établir des statistiques officielles à partir d'enquêtes en ligne reposant sur le principe de l'autosélection?* : Symposium.
- BEXELIUS, C., H. MERK, S. SANDIN, A. EKMAN, O. NYRÉN, S. KÜHLMANN-BERENZON, A. LINDE et J. E. LITTON. 2009. « SMS Versus Telephone Interviews for Epidemiological Data Collection: Feasibility Study Estimating Influenza Vaccination Coverage in the Swedish Population ». *European Journal of Epidemiology* 24 (2) : 73-81.
- BLAIR, J., R. F. CZAJA et E. A. BLAIR. 2014. *Designing Surveys: A Guide to Decisions and Procedures, 3rd ed.* : SAGE Publications, Incorporated.
- BONNEL, P. et C. BAYART. 2010. « L'impact du mode d'enquête sur la mesure des comportements de mobilité ». *Économie et statistique* : 47-70.
- BÖRKAN, B. 2010. « The Mode Effect in Mixed-Mode Surveys: Mail and Web Surveys ». *Social Science Computer Review* 28 (3) : 371-380.
- BOSNJAK, M., I. HAAS, M. GALESIC, L. KACZMIREK, W. BANDILLA et M. P. COUPER. 2013. « Sample Composition Discrepancies in Different Stages of a Probability-based Online Panel ». *Field Methods* 25 (4) : 339-360.
- BOULIANNE, S. 2013. « Examining the Gender Effects of Different Incentive Amounts in a Web Survey ». *Field Methods* 25 (1) : 91-104.
- CHICHE, J. 2013. « Sondages d'intentions de vote : de la transparence à la confiance ». *Statistique et société* 1 (2) : 29-32.
- COOK, C., F. HEATH et R. L. THOMPSON. 2000. « A Meta-Analysis of Response Rates in Web- or Internet-based Surveys ». *Educational and Psychological Measurement* 60 (6) : 821-836.
- DE BERNARDO, D. H. et A. CURTIS. 2013. « Using Online and Paper Surveys: The Effectiveness of Mixed-Mode Methodology for Populations Over 50 ». *Research on Aging* 35 (2) : 220-240.
- DENSCOMBE, M. 2006. « Web-Based Questionnaires and the Mode Effect An Evaluation Based on Completion Rates and Data Contents of Near-Identical Questionnaires Delivered in Different Modes ». *Social Science Computer Review* 24 (2) : 246-254.
- DEUTSKENS, E., K. DE RUYTER, M. WETZELS et P. OOSTERVELD. 2004. « Response Rate and Response Quality of Internet-Based Surveys: an Experimental Study ». *Marketing letters* 15 (1) : 21-36.

- DILLMAN, D. A., G. PHELPS, R. TORTORA, K. SWIFT, J. KOHRELL, J. BERCK et B. L. MESSER. 2009. « Response Rate and Measurement Differences in Mixed-Mode Surveys Using Mail, Telephone, Interactive Voice Response (IVR) and the Internet ». *Social Science Research* 38 (1) : 1-18.
- DOWNES-LE GUIN, T., R. BAKER, J. MECHLING et E. RUYLE. 2012. « Myths and Realities of Respondent Engagement in Online Surveys ». *International Journal of Market Research* 54 (5) : 1-21.
- DURAND, C. 2013. Les sondages Internet : nouveaux défis et défis revisités. Symposium de Statistiques.
- DUSSAIX, A.-M. 2009. « La qualité dans les enquêtes ». *Revue Modulad* 137 (39).
- FAN, W. et Z. YAN. 2010. « Factors Affecting Response Rates of the Web Survey: A Systematic Review ». *Computers in Human Behavior* 26 (2) : 132-139.
- FARRELL, D. et J. C. PETERSEN. 2010. « The Growth of Internet Research Methods and the Reluctant Sociologist ». *Sociological Inquiry* 80 (1) : 114-125.
- FINE, J. 2013. « Pour une plus grande transparence sur la méthodologie des sondages électoraux ». *Statistique et société* 1 (2) : 23-28.
- FRICKER, S., M. GALESIC, R. TOURANGEAU et T. YAN. 2005. « An Experimental Comparison of Web and Telephone Surveys ». *Public Opinion Quarterly* 69 (3) : 370-392.
- FRIPPIAT, D. et N. MARQUIS. 2010. « Les enquêtes par Internet en sciences sociales : un état des lieux ». *Population* 65 (2) : 309-338.
- GANASSALI, S. 2008. « The Influence of the Design of Web Survey Questionnaires on the Quality of Responses ». *Survey Research Methods* 2 (1).
- GREENE, J., H. SPEIZER et W. WIITALA. 2008. « Telephone and Web: Mixed-mode challenge. » *Health Services Research* 43 (1 P1) : 230-248.
- HEERWEGH, D. et G. LOOSVELDT. 2002. « An Evaluation of the Effect of Response Formats on Data Quality in Web Surveys ». *Social Science Computer Review* 20 (4) : 471-484.
- HEERWEGH, D. 2008. « Face-to-Face Versus Web Surveying in a High-Internet-Coverage Population: Differences in Response Quality. » *Public Opinion Quarterly* 72 (5) : 836-846.
- HEERWEGH, D. 2009. « Explaining the Intention to Participate in a Web Survey: A Test of the Theory of Planned Behaviour ». *International Journal of Social Research Methodology* 12 (3) : 181-195.
- HUNT, K., N. SHLOMO et J. ADDINGTON-HALL. 2013. « Participant Recruitment in Sensitive Surveys: a Comparative Trial of 'opt in' Versus 'opt out' Approaches ». *BMC Medical Research Methodology* 13 (1) : 3.
- INSTITUT DE LA STATISTIQUE DU QUÉBEC. 2013. *L'enquête québécoise sur l'accès des ménages à Internet 2012*, Gouvernement du Québec, <http://www.stat.gouv.qc.ca/statistiques/science-technologie-innovation/utilisation-internet/menages-individus/menage-internet-2012.pdf>
- LEGLEYE, S., G. CHARRANCE, N. RAZAFINDRATSIMA, A. BOHET, N. BAJOS et C. MOREAU. 2013. « Improving Survey Participation: Cost Effectiveness of Callbacks to Refusals and Increased Call Attempts in a National Telephone Survey in France ». *Public Opinion Quarterly* 77 (3) : 666-695
- LINDHJEM, H. et S. NAVRUD. 2011. « Using Internet in Stated Preference Surveys: A Review and Comparison of Survey Modes ». *International Review of Environmental and Resource Economics* (Forthcoming).

-
- MASSEY, D. S. et R. TOURANGEAU. 2013. « Where Do We Go from Here? Nonresponse and Social Measurement ». *The ANNALS of the American Academy of Political and Social Science* 645 (1) : 222-236.
- NAGELHOUT, G. E., M. C. Willemsen, M. E. THOMPSON, G. T. FONG, B. VAN DEN PUTTE et H. DE VRIES. 2010. « Is Web Interviewing a Good Alternative to Telephone Interviewing? Findings from the International Tobacco Control (ITC) Netherlands Survey ». *BMC Public Health* 10 (1) : 351.
- NOVICK, G. 2008. « Is There a Bias Against Telephone Interviews in Qualitative Research? ». *Research in Nursing and Health* 31 (4) : 391-398.
- PAN, B., A. G. WOODSIDE et F. MENG. 2014. « How Contextual Cues Impact Response and Conversion Rates of Online Surveys ». *Journal of Travel Research* 53 (1) : 58-68.
- PORTER, S. R. et M. E. WHITCOMB. 2003. « The Impact OF Contact Type On Web Survey Response Rates ». *Public Opinion Quarterly* 67 (4) : 579-588.
- RAO, K., O. KAMINSKA et A. L MCCUTCHEON. 2010. « Recruiting Probability Samples for a Multi-Mode Research Panel with Internet and Mail Components ». *Public Opinion Quarterly* 74 (1) : 68-84.
- RIVARA, F. P., T. D. KOEPEL, J. WANG, D. DURBIN, K. M. JAFFE, M. VAVILALA, A. DORSCH, M. ROPER-CALDBECK, E. HOUSEKNECHT et N. TEMKIN. 2011. « Comparison of Telephone with World Wide Web-Based Responses by Parents and Teens to a Follow-Up Survey After Injury ». *Health Services Research* 46 (3) : 964-981.
- SHIN, E., T. P. JOHNSON et K. RAO. 2012. « Survey Mode Effects on Data Quality: Comparison of Web and Mail Modes in a US National Panel Survey ». *Social Science Computer Review* 30 (2) : 212-228.
- STEPHENSON, L. B et J. CRÊTE. 2011. « Studying Political Behavior: A Comparison of Internet and Telephone Surveys ». *International Journal of Public Opinion Research* 23 (1) : 24-55.
- STRABAC, Z. et T. AALBERG. 2011. « Measuring Political Knowledge in Telephone and Web Surveys: A Cross-National Comparison ». *Social Science Computer Review* 29 (2) : 175-192.
- VEHOVAR, V., N. BERZELAK et K. LOZAR MANFREDA. 2010. « Mobile Phones in an Environment of Competing Survey Modes: Applying Metric for Evaluation of Costs and Errors ». *Social Science Computer Review* 28 (3) : 303-318.
- WANG, C.-C., K.-S. LIU, C.-L. CHENG et Y.-Y. CHENG. 2013. « Comparison of Web-Based Versus Paper-and-Pencil Administration of a Humor Survey ». *Computers in Human Behavior* 29 (3) : 1007-1011.
- YEAGER, D. S., J. A. KROSNICK, L. CHANG, H. S. JAVITZ, M. S. LEVENDUSKY, A. SIMPSON et R. WANG. 2011. « Comparing the Accuracy of RDD Telephone Surveys and Internet Surveys Conducted with Probability and Non-Probability Samples ». *Public Opinion Quarterly* 75 (4) : 709-747.

Avantages	Auteurs
Faible coût	Deutskens <i>et al.</i> , 2004 : 21 Wang <i>et al.</i> , 2013 : 1007 Ganassali, 2008 : 21 Bethlehem, 2008 : 1 Stephenson et Crête, 2011 : 26 Greene <i>et al.</i> , 2008 : 231 Bigot <i>et al.</i> , 2010 : 21
Rapidité des réponses	Deutskens <i>et al.</i> , 2004 : 21 Ganassali, 2008 : 21 Bethlehem, 2008 : 1 Stephenson et Crête, 2011 : 26 Greene <i>et al.</i> , 2008 : 231 Bigot <i>et al.</i> , 2010 : 21
Envois facilités de rappel	Deutskens et al. 2004 : 21
Le fait de taper les réponses sur un clavier plutôt que le faire à la main.	Shin <i>et al.</i> , 2012 : 213
Meilleure qualité des données (Réponses plus longues; les répondants se livrent plus, Denscombe)	Shin <i>et al.</i> , 2012 : 223 Denscombe, 2006 : 252
Moins de réponses socialement désirables en raison du caractère anonyme (Surtout lorsqu'il est question des problèmes personnels et des comportements à risque, Wang) (Pour les sondages sur le tabagisme, les fumeurs préfèrent le remplir en ligne, Nagelhout) (Possibilité de toucher des sujets sensibles, Stephenson et Kreuter)	Dillman <i>et al.</i> , 2009 : 3 Wang <i>et al.</i> , 2013 : 1008 Lindhjem et Navrud, 2011 : 12 Ganassali, 2008 : 21 Nagelhout <i>et al.</i> , 2010 : 9 Stephenson et Crête, 2011 : 27 Greene <i>et al.</i> , 2008 : 246
Meilleure compréhension des questions et des nuances grâce à la présence de supports techniques	Dillman <i>et al.</i> , 2011 : 12 Stephenson et Crête 2011 : 27
Bassin géographique de répondants plus large	Wang <i>et al.</i> , 2013 : 1007 Bethlehem, 2008 : 91 Bigot <i>et al.</i> , 2010 : 21
Répondant n'a pas la pression de répondre dans un délai prescrit, peut s'y mettre quand il veut	Lindhjem et Navrud, 2011 : 12 Bigot <i>et al.</i> , 2010 : 21

Présence de données socioéconomiques et de renseignements sur les attitudes et affiliations politiques (Possibilité d'identifier les biais de non-réponses)	Lindhjem et Navrud, 2011 : 21
Plus faible tendance au désengagement	Lindhjem et Navrud, 2011 : 38 Ganassali, 2008 : 21 Stephenson et Crête, 2011 : 27
Taux de réponse semblable aux autres méthodes	Ganassali, 2008 : 21
Meilleur taux de réponse que celui du questionnaire papier	Denscombe, 2006 : 252
Surtout utile aux chercheurs voulant rejeter l'hypothèse nulle	Yeager <i>et al.</i> , 2011: 737
Facilite l'échantillonnage des populations marginales	Stephenson et Crête, 2011 : 27
Élimine la possibilité d'erreurs lors de la saisie des données	Stephenson et Crête, 2011 : 27
Favorise l'engagement des répondants	Stephenson et Crête, 2011 : 27
Moins de gêne de la part des répondants pour ce type de réponse : « je ne sais pas » ou « je ne désire pas répondre »	Stephenson et Crête, 2011 : 27
Réponses moins extrêmes qu'avec les autres modes de sondage	Frippiat et Marquis, 2010 : 59

Désavantages	Auteurs
Plus petit taux de réponse (comparativement aux sondages postaux)	Shin <i>et al.</i> , 2012 : 223
Exigent une familiarité avec les ordinateurs (dextérité et littératie)	Shin <i>et al.</i> , 2012 : 213
Réponses ouvertes moins favorisées	Shin <i>et al.</i> , 2012 : 213
Méfiance par rapport à la confidentialité et à l'anonymat, méfiance envers les nouvelles technologies (les réponses apparaissent plus socialement désirables)	Wang <i>et al.</i> , 2013 : 1007 Lindhjem et Navrud, 2011 : 12
Méfiance au sujet de la sécurité (virus)	Porter et Whitcomb, 2003 : 579 Frippiat et Marquis, 2010 : 48

Accès limité à l'Internet	Strabac et Aalberg, 2011 : 187 Lindhjem et Navrud, 2011 : 20 Bethlehem, 2008 : 3
Causalité possible entre l'absence d'intervieweur et le <i>satisficing</i> (même réponse pour toutes les questions et désengagement)	Downes-Le Guin <i>et al.</i> , 2012 : 2 Lindhjem et Navrud, 2011 :12 Stephenson et Crête, 2011 : 27 Heerwegh et Lossveldt, 2008 : 838
Nombre de questions sans réponses plus élevé	Lindhjem, et Navrud, 2011 : 20 Heerwegh et Lossveldt, 2008 : 838
Échantillonnage non probabiliste (Dussaix parle surtout de l'échantillonnage non aléatoire, alors que Stephenson le qualifie de non représentatif)	Lindhjem et Navrud, 2011 : 20 Ganassali, 2008 : 21 Bethlehem, 2008 :2 Dussaix, 2009 :152 Yeager et al. 2011 :731-732 Stephenson & Crête, 2011 :27
Biais d'auto-sélection	Lindhjem & Navrud, 2011 : 20 Ganassali, 2008 :21 Bethlehem, 2008 :2 Stephenson & Crête, 2011 :27
Sondage vulnérable aux groupes d'intérêts	Bethlehem, 2008 :8
Aînés, personnes de régions éloignées et personnes peu scolarisées sous-représentées	Lindhjem & Navrud, 2011 : 20 Bethlehem, 2008 : 3 Fricker et al. 2005 : 388 Rivara et al. 2011 : 976

Caractéristiques neutres	Auteurs
Répondants habituellement de plus grands consommateurs et utilisateurs d'Internet	Strabac et Aalberg, 2011 : 187 Stephenson et Crête, 2011 : 27
Interface HTML traditionnelle plus efficace qu'une interface trop imposante (favorise l'abandon à cause du temps de téléchargement)	Downes-Le Guin, 2012 : 2
Positions plus extrêmes dans les sondages en ligne	Stephenson et Crête, 2011 : 26

Méthodes de rectification des échantillons	Auteurs
<p>Impossibilité d’avoir un échantillonnage probabiliste</p> <ul style="list-style-type: none"> • Pour y remédier, mettre sur pied un panel probabiliste de répondants souhaitant participer en utilisant des méthodes de recrutement aléatoires (appels digitaux aléatoires et RDD). Sinon, recruter des participants prêts à participer avec des moyens non probabilistes à travers la publicité (<i>volunteer panel</i>) qui nécessitent toutefois un échantillonnage par sous-groupes (âge, éducation, revenus) ou une pondération (post sondage), pour réduire les biais de non-représentation de la population générale 	Lindhjem et Navrud, 2011 : 20
<p>Non fiabilité de la pondération appliquée aux échantillons non probabilistes</p>	Lindhjem et Navrud, 2011 : 21 Stephenson et Crête, 2011 : 27 et 48
<p>Utilisation d’une méthode mixte (mais qui pourrait compromettre la qualité des données en raison des effets de mode)</p>	Bethlehem, 2008 : 8
<p>Enregistrement des variables démographiques (pour les comparer au reste de la population et les ajuster à celle-ci)</p>	Bethlehem, 2008 : 8
<p>Correction des sur et sous représentations par la technique de pondération, une stratégie fiable apparemment.</p>	Bethlehem, 2008 : 8 Stephenson et Crête, 2011 : 27

Pour les méthodes de pondération, il faut lire :

- « Les enquêtes par Internet en sciences sociales : un état des lieux » (Fripiat et Marquis 2010)
- « Summary Report of the AAPOR Task Force on Non-probability Sampling » (Baker *et al.*, 2013)