

1. ENQUETE WEB : QUEL POTENTIEL POUR LES ENQUETES DEPLACEMENTS ?

Patrick BONNEL, Laboratoire d'Economie des Transports, ENTPE, Lyon, France; Tel +33 (0)4 72 04 70 48; fax +33 (0)4 72 04 70 92; patrick.bonnel@entpe.fr

Le développement des enquêtes dans de nombreux domaines est très rapide. Il est donc important d'analyser le potentiel du web pour collecter des données sur les déplacements. Nous proposons tout d'abord une définition de ce que nous entendons par enquête web (section 1). Nous analysons ensuite les avantages et limites des enquêtes web pour collecter des données sur les déplacements (section 2). Nous identifions enfin les pistes de recherche pour améliorer la pertinence du web pour ces enquêtes (section 3).

1.1 Définition et classification

Compte tenu de la diversité des enquêtes web que l'on peut trouver dans la littérature, nous avons préféré adopter une définition très générique qui intègre les possibles développements dans le futur. Notre définition correspond à « toute enquête pouvant être conduite sur le web ». Cette définition se réfère explicitement au web et exclues donc les enquêtes par e-mail afin d'insister sur le potentiel d'interactivité offert par le web (Alsnih, 2004). Cette définition est volontairement non restrictive. Cela permet d'insister sur le caractère de média support du web au même titre que les enquêtes téléphoniques, postales...

Certaines classifications sont déjà proposées dans la littérature (Dillman et Bowker, 2001). Elles s'appuient le plus souvent sur les trois dimensions suivantes :

- Enquête web exclusive / enquête combinant les media d'enquête. Cette première dimension soulève le problème de la construction de l'échantillon d'enquête et de la comparabilité des données issues de plusieurs média d'enquête (sections 2 et 3) ;
- Accès restreint / accès libre. Ce point est fondamental par rapport au contrôle de l'échantillon. Seul un accès restreint et contrôlé permet d'envisager de développer des méthodes d'inférence statistique afin de généraliser les résultats à l'ensemble de la population cible (Bradley, 1999) ;

- Méthode de contact et de recrutement. Les méthodes sont variées. Bradley (1999) identifie 15 méthodes différentes pour le web. Elles n'offrent toutefois pas toutes les mêmes possibilités de contrôle de l'échantillon et ne permettent donc pas toutes l'application des méthodes d'inférence statistique.

1.2 Avantages et limites du Web en tant que média d'enquête

1.2.1 Principaux avantages des enquêtes web

Notre analyse est centrée sur les avantages du web en tant que media d'enquête comparativement aux autres média d'enquête.

□ Faible coût d'enquête

Les coûts variables d'une enquête web sont beaucoup plus faibles que ceux d'une enquête administrée, car il n'est pas nécessaire de recruter des enquêteurs. De plus, par rapport à une enquête postale, les coûts de codification et de saisie sont nuls. La situation est plus complexe pour les coûts fixes car elle dépend des coûts de développement nécessaires à l'implantation de l'enquête sur le web. Le développement de logiciel d'enquêtes web devrait toutefois réduire ces coûts.

□ Interactivité

Cet avantage est partagé avec les enquêtes administrées assistées par ordinateur (CATI et CAPI pour les enquêtes en face à face et téléphonique). Le web permet ainsi d'étendre cet avantage à des enquêtes auto-administrées (CAWI). La gestion centralisée de l'enquête web offre toutefois davantage de latitude en la matière.

□ Limitation du caractère intrusif

La réalisation d'une enquête administrée, surtout en face à face, est souvent perçue comme une intrusion dans l'espace individuel. L'absence de l'enquêteur permet de limiter fortement cette perception de la part des enquêtés. De plus, l'enquêté peut choisir de répondre au moment qui lui convient et non plus au moment proposé par l'enquêteur (même si la prise de rendez-vous tend à réduire la gêne de l'enquêté).

- Possibilité de contacter les personnes qui ont une faible participation aux enquêtes conduites à l'aide d'autres média

Les personnes qui se déplacent le plus sont généralement les plus difficiles à joindre car ... elles sont moins souvent à leur domicile. Les enquêtes auto-administrées offrent davantage de souplesse pour que les enquêtés puissent répondre quand et où ils le souhaitent.

Par ailleurs, les enquêtes doivent faire face à des taux de réponse qui ont tendance à décroître dans le temps. La combinaison de modes d'enquêtes offre des perspectives pour faire face à cette difficulté.

- Saisie directe et contrôle automatique des données

Comme les autres modes d'enquêtes administrées, le web permet une saisie directe des données qui réduit les coûts et permet un contrôle de cohérence des données en temps réel.

- Temps de réponse plus court

Le web permet une forte réduction des temps nécessaires à la réalisation d'une enquête sur le terrain.

1.2.2 Principales limites des enquêtes web

Si les avantages sont nombreux, il ne faut pas oublier que les enquêtes web connaissent aussi de fortes limites qui sont généralement organisées autour de 4 sources majeures de biais (Dillman et Bowker, 2001) erreur de couverture, d'échantillonnage, non-réponse et erreur de mesure.

- Erreur de couverture : Pénétration d'internet

Bien que le taux de pénétration d'internet évolue rapidement sur tous les continents, il est encore trop faible pour que la population d'internaute soit représentative de l'ensemble de la population (tableau 1). De plus la population des internautes apparaît fortement corrélée à l'âge des individus (plus jeune), au revenu et au niveau d'éducation (plus élevé) (Alsnih, 2004). Les hommes semblent également répondre davantage aux enquêtes web que les femmes.

TABLEAU 1 : STATISTIQUE MONDIALE D'ACCES A INTERNET

	Population (2004 Est.)	Accès à internet, (2000)	Accès à internet, 31 mars 2005	Croissance (2000-2005)	Pénétration (% Population)
Afrique	900 465 411	4 514 400	13 468 600	198,3 %	1,5 %
Asie	3 612 363 165	114 303 000	302 257 003	164,4 %	8,4 %
Europe	730 991 138	103 096 093	259 653 144	151,9 %	35,5 %

Moyen Orient	259 499 772	5 284 800	19 370 700	266,5 %	7,5 %
Amérique du Nord	328 387 059	108 096 800	221 437 647	104,9 %	67,4 %
Amérique Latine - Caraïbes	546 917 192	18 068 919	56 224 957	211,2 %	10,3 %
Océanie	33 443 448	7 619 500	16 269 080	113,5 %	48,6 %
Total	6 412 067 185	360 971 012	888 681 131	146,2 %	13,9 %

Source : <http://www.internetworldstats.com/stats.htm>, 31 mars 2005

De plus, la population ayant accès à internet n'est pas toujours suffisamment familière de son usage pour répondre à des enquêtes web. Bradley (1999) distinguent 13 classes d'internautes selon leur type d'équipement et d'usage d'internet. Les individus appartenant à la moitié de ces classes ne sont pas à même de répondre à une enquête web. Même si l'on n'a pas de statistiques sur la répartition des individus selon ces classes, il apparaît peu probable que l'on puisse conduire des enquêtes web représentatives de l'ensemble de la population dans un avenir proche.

Il est donc fortement recommandé de ne pas conduire d'enquête web exclusive auprès de l'ensemble de la population si l'on veut obtenir des données représentatives de cette population. En revanche, la littérature présente de nombreux exemples d'enquêtes de qualité auprès de populations spécifiques ayant un accès important et contrôlable à internet, comme par exemple auprès de salariés d'une entreprise ou d'un campus universitaire...

□ Erreur d'échantillonnage : construction de l'échantillon

A la différence d'autres média d'enquêtes, il n'existe pas pour le moment de bases de sondage pour des enquêtes web auprès de l'ensemble de la population. La construction de l'échantillon est donc souvent plus complexe pour une enquête web. L'absence de base d'adresses e-mail nécessite donc le recours à d'autres media (face à face, téléphone, postal...) pour construire l'échantillon. L'utilisation d'accès restreint et contrôlé à l'enquête est donc généralement indispensable. De plus, dans tous les cas, il est nécessaire d'analyser la représentativité de l'échantillon.

Toutefois, cette limite ne concerne pas les enquêtes auprès de population spécifiques pour lesquelles les bases d'adresses e-mail existent. Les techniques usuelles d'échantillonnage peuvent alors s'appliquer comme pour les autres média. Enfin, lorsque le web est utilisé en complément d'autres media, par exemple pour enquêter les non-répondants à une enquête en face à face, il permet d'accroître la représentativité des données.

□ Erreur de non-réponse : un faible taux de réponse

De faibles taux de réponse sont souvent présentés dans la littérature pour les enquêtes auto-administrées (Alsnih, 2004). De nombreuses recherches ont

montré qu'il n'est pas possible de postuler que la non-réponse n'est pas corrélée aux comportements de mobilité. Dès lors, la non-réponse est source potentielle de biais. Il est donc nécessaire de développer des protocoles d'enquête permettant de réduire au maximum ce biais (lettre introductive, campagne de communication, relances, design de l'enquête...).

□ Erreur de mesure : des problèmes techniques

Les erreurs de mesure sont une source permanente de biais particulièrement dans les enquêtes visant à collecter les déplacements ou les activités des individus. Certaines de ces erreurs sont communes à tous les média, mais d'autres sont liées à des aspects techniques du média web (Dillman et Bowker, 2001 ; Alsnih, 2004):

- Non disponibilité du serveur, qui de plus peut survenir sans que l'administrateur de l'enquête en soit immédiatement informé ;
- La présentation du questionnaire peut dépendre du browser utilisé ;
- Un temps de chargement excessif du questionnaire.

Certaines recommandations peuvent être faites pour limiter ces problèmes :

- Choisir un langage informatique standard tel que HTML ou XML, et utiliser les fonctionnalités les plus simples possibles ;
- Utiliser le design le plus simple possible (Dillman et Bowker, 2001) ;
- Utiliser deux serveurs afin de pouvoir pallier la défaillance de l'un des deux ;
- Et comme pour toutes enquêtes conduire des tests.

1.3 Questions de recherche

Pour faire face aux principales limites évoquées précédemment, nous avons identifié trois axes de recherche privilégiés :

- Combinaison de modes d'enquêtes (section 3.1);
- Echantillonnage et représentativité (section 3.2);
- Questions de design (section 3.3).

1.3.1 Combinaison de modes d'enquêtes

La question de la combinaison de modes d'enquêtes n'est pas nouvelle (Bonnell, 2003). Toutefois, le développement des enquêtes web amène à reconsidérer cette question compte tenu du bénéfice potentiel offert par la combinaison du web avec d'autres média. La combinaison permet le plus souvent d'accroître le taux de réponse.

La première question qui mérite d'être posée concerne la comparabilité des données recueillies. Est-ce que les réponses fournies à l'aide du web sont similaires à celles des autres media. Les recherches effectuées en marketing

indiquent une plus grande propension des enquêtés à fournir des réponses extrêmes dans les questions de satisfaction ou à donner des réponses socialement désirables ou à acquiescer dans les modes oraux (téléphone et face à face) que visuels (postal et web) (De Leeuw, 1992). Par ailleurs, une des rares expériences de comparaison dans le domaine des enquêtes déplacements a montré une mobilité plus importante des répondants au web par rapport au mode postal (Resource Systems Group, 2002). Toutefois, ces différences peuvent être en partie expliquées par les caractéristiques socio-économiques des répondants au web par rapport au postal.

Le choix du design du questionnaire qui n'est généralement pas identique selon le mode (par exemple les filtres ne sont pas possible dans une enquête postale) peut également influencer les réponses (Dillman et Bowker, 2001).

Il serait donc nécessaire de développer des tests afin d'analyser la comparabilité des données produites dans différents contextes. Enfin afin d'optimiser la combinaison des modes d'enquêtes, il est souhaitable d'analyser les facteurs influençant les préférences des individus quant au media d'enquête.

1.3.2. Echantillonnage et représentativité

Nous avons déjà évoqué le problème de la représentativité de l'échantillon dans une enquête web lorsqu'il s'agit d'enquêter l'ensemble de la population. Plusieurs approches peuvent être identifiées pour tenter de réduire ce problème :

- Le taux de pénétration d'internet est évidemment un indicateur important, mais il serait utile de le compléter par un taux "d'aisance" dans l'utilisation du web ;
- Quelle est la meilleure stratégie pour construire un échantillon représentatif dans une enquête web. Certaines techniques ne permettent pas toujours de s'assurer de la représentativité de l'échantillon ou encore de calculer un taux de réponse... De nombreux auteurs reconnaissent que la recherche est insuffisante sur ces questions ;
- L'utilisation des motivations financières génère de nombreuses discussions au niveau international du fait des différences culturelles entre les pays. Toutefois, pour les enquêtes web, l'expérience manque pour connaître l'incidence des différentes formes de motivations financières sur les taux de réponse ou la qualité des données produites ;
- Le nombre de relances nécessite des considérations spécifiques aux enquêtes web. Certains auteurs recommandent leur limitation pour certains média, est-ce également le cas pour les enquêtes web ?

Comment le nombre de relances affectent les réponses fournies, en termes de nombre de déplacements recueillis par exemple ?

1.3.3 Questions de design

Un des avantages du media web est d'offrir de grande potentialité d'interactivité. Dans le même temps, il est recommandé d'avoir un design le plus simple possible. Il est donc nécessaire d'identifier ce qui est possible afin d'améliorer la qualité de l'enquête sans accroître la charge pour l'enquêté. Par exemple, des techniques existent pour géocoder des données de localisation en temps réel en face à face ou par téléphone. Ces techniques sont-elles applicables pour le web ?

Les enquêtes auto-administrées conduisent fréquemment à un taux de non-réponse à certaines questions plus important que les enquêtes administrées. Le web permet de rendre la réponse obligatoire ou bien d'assurer une relance en cas de non-réponse. Le risque est toutefois d'accroître la lourdeur de l'enquête pour l'enquêté qui pourrait abandonner ou fournir des réponses biaisées. De nouveau, des expérimentations sont nécessaires afin d'identifier ce qui est possible.

Les enquêtes déplacements cherchent fréquemment à enquêter l'ensemble des membres du ménage pour le même jour donné. Faut-il rendre la réponse obligatoire pour l'ensemble des membres du ménage au risque d'accroître les refus ou les biais dans les réponses fournies ? Faut-il assurer des relances ? et combien ? ...

Les enquêtes déplacements ont souvent une durée assez longue (1h30 en moyenne, en France, pour les enquêtes ménages déplacements urbaines, par exemple). Cette durée ne semble pas envisageable pour le web, mais jusqu'où peut-on aller en termes de durée sans accroître fortement le risque d'abandon ou de refus ou encore une faible qualité de réponse ? Cet aspect est particulièrement important pour la combinaison de modes d'enquêtes, la durée « admissible » n'étant pas identique selon le média d'enquête.

1.4 Conclusions

L'utilisation du web pour les enquêtes déplacements va très probablement s'accroître à mesure de sa diffusion dans la population, comme on peut déjà l'observer pour les enquêtes marketing. Compte tenu des taux de pénétration actuels et de l'aisance d'utilisation des internautes, l'utilisation du web en tant que media unique pour enquêter l'ensemble de la population n'est pas souhaitable. Son utilisation doit être limitée à des contextes spécifiques dans lesquels il est possible de contrôler la construction de l'échantillon afin de pouvoir utiliser les méthodes usuelles d'inférences statistiques. Toutefois, son utilisation en tant que média complémentaire aux autres médias d'enquête

semble prometteuse, notamment pour faire face à la croissance des refus de réponse. Toutefois, de nombreuses recherches sont encore nécessaires afin de préciser l'univers du possible en matière de design et d'organisation d'enquêtes web. De plus, la comparabilité des données produites par le web avec celles produites par d'autres médias reste encore à analyser.

Bibliographie

- Alsnih R (2005), Characteristics of Web-based surveys and applications in travel research. CD-Rom of the *ISCTSC conference*, Costa Rica, August 1-6, 2004, <http://www.its.usyd.edu.au/isctsc/costaricaconference.asp>.
- Bonnel P (2003), Postal, telephone and face-to-face surveys: how comparable are they? In *Transport Survey Quality and Innovation*. Stopher PR and Jones PM (Eds.), Pergamon, Oxford, pp. 215-237.
- Bradley M (1999), Sampling for internet surveys; an examination of respondent selection for internet research. *International Journal of Market Research*, Vol. 41, N°4, pp. 441-452.
- De Leeuw E (1992), *Data Quality in Mail, Telephone and Face to Face surveys*, TT Publikaties Amsterdam, Vrije Universiteit, 168p.
- Dillman DA and Bowker DK (2001), The Web Questionnaire Challenge to Survey Methodologists in *Dimensions of Internet Science*. Edited by Ulf-Dietrich Reips & Michael Bosnjak. Pabst Science Publishers, Lengerich, Germany.
- Resource Systems Group Inc. (2002), *Documentation for SBIR Phase II Final Report: Computer-Based Intelligent Travel Survey System*, DTRS57-00-C-10030., Prepared for the FHWA, US Department of Transportation, Washington, D.C. (<http://www.fhwa.dot.gov/ohim/trb/sbir/sbir.htm>).