Techniques d'enquête, décembre 2000
Vol. 26, N° 2, pp. 225-234
Statistique Canada, N° 12-001 au catalogue

La conception des écrans et l'ordre des questions dans un module IAO Résultats d'une expérience de facilité d'utilisation menée sur le terrain

Marek Fuchs [1]

Résumé

La conception des écrans et la conception des questionnaires ont une influence sur le comportement des intervieweurs dans un environnement IAO. Des recherches antérieures ont montré que les intervieweurs peuvent travailler de façon plus convenable et efficace si des fonctions et caractéristiques appropriées sont incorporées dans le module IAO. Des expériences de facilité d'utilisation fondées sur la liste des membres des ménages de deux importantes enquêtes gouvernementales ont montré que le recours à des grilles et à des tableaux facilite le travail de l'intervieweur. Ces expériences ont été menées dans des conditions de laboratoire, mais nous avons des résultats d'une première expérience sur le terrain. En mars 1998, une enquête de type ITAO a été menée auprès d'immigrants en Allemagne (taux de réponse 84 %, $n = 501$). Dans cette étude de production, quatre versions d'une liste des membres du ménage ont été comparées afin de vérifier deux conceptions des écrans et deux ordres des questions dans une conception factorielle 2×2. Les quatre versions ont été attribuées de façon aléatoire à des intervieweurs et à des répondants. Des mesures du temps ont été intégrées au programme ITAO, et 234 interviews sélectionnées au hasard ont été enregistrées sur bande magnétoscopique, puis analysées en fonction d'un schéma de codage. À l'aide des données, nous avons évalué la facilité d'utilisation des différentes caractéristiques de conception de l'IAO. Les résultats indiquent que la conception des écrans de même que l'ordre des questions ont une influence appréciable sur la durée de l'interview et le comportement des intervieweurs. En particulier, la version quadrillée et thématique donne les résultats les plus rapides pour ce qui est du temps mis à remplir le questionnaire. Les résultats tirés des données de codage indiquent que les différences entre les versions sont attribuables à des comportements particuliers des intervieweurs et des répondants. Les données indiquent que la version thématique quadrillée facilite chez l'intervieweur un comportement axé sur le répondant et donne lieu au meilleur rendement des intervieweurs pour ce qui est de la durée.

Mots clés : Interview assistée par ordinateur; essai de facilité d'utilisation; expérience sur le terrain; conception des écrans; ordre des questions.

1. Introduction

Les interviews assistées par ordinateur sont en train de devenir une technique d'enquête standard (Couper, Baker, Bethlehem, Clark, Martin, Nicholls et O'Reilly 1998). Pour les enquêtes téléphoniques aussi bien que pour les interviews personnelles, de plus en plus d'études sont menées à l'aide de techniques d'interview assistée par ordinateur (IAO). Plusieurs enquêtes gouvernementales importantes aux États-Unis sont en train de passer à l'IAO, ou cette transition est déjà terminée. Même en Europe, on observe une évolution vers l'interview assistée par ordinateur (Schneid 1991; Fuchs 1994, 1995; Laurie et Moon 1997; Projektgruppe SOEP 1998) - même si, pour le moment, les aspects méthodologiques de ce phénomène ne représentent pas le point central de la recherche européenne.

Les chercheurs et les responsables des enquêtes se fient à l'interview assistée par ordinateur pour plusieurs raisons (toutefois, il semble parfois que de solides arguments sont moins importants que la simple popularité de l'IAO) :

- ils espèrent recueillir des données de qualité supérieure grâce à des vérifications incorporées de la cohérence et des intervalles au cours de l'interview;
- l'IAO permet d'enchaîner les questions automatiquement et de concevoir des questionnaires plus complexes sans imposer un fardeau indu à l'intervieweur;
- ils espèrent consacrer moins de temps et d'argent à l'interview et au post-traitement et réduire les budgets d'enquête une fois les coûts initiaux de matériel et de logiciel absorbés;
- ils espèrent bénéficier de la capacité de l'IAO d'intégrer des données externes à l'interview, fonction particulièrement intéressante pour les études par panel.

La transition générale à l'IAO est vue d'un bon œil. Les chercheurs et les directeurs sur le terrain en tirent des avantages (Nicholls et deLeeuw 1996) et les intervieweurs (Couper et Burt 1994) de même que les répondants (Baker 1992) manifestent beaucoup de sympathie ou du moins de l'acceptation. Par contre, les interviews assistées par ordinateur ont introduit certains problèmes supplémentaires dans le déroulement des interviews : au début, la recherche méthodologique a porté surtout sur des problèmes de matériel et de logiciel (on trouvera des synthèses dans Couper, Groves et Kosary 1989 et dans Weeks 1992). Des études plus récentes ont traité de l'acceptation par les intervieweurs et les répondants, de la

1. Dr. Marek Fuchs, Catholic University of Eichstätt, Department of Sociology, Ostenstrasse 26, 85071 Eichstätt, Germany. Courriel : marek.fuchs@ku-eichstaett.de.

durée des interviews et de la facilité d'utilisation (on trouvera une synthèse dans Couper et coll. 1998). Le présent document fait partie de cette récente discussion des « effets de la technologie » (Fuchs, Couper et Hansen 2000).

2. Contexte théorique

Dans la présente analyse, le point de convergence théorique touche surtout à deux aspects de la facilité d'utilisation : 1) la segmentation du cheminement de l'interview et 2) le manque de souplesse pour l'intervieweur.

1. Segmentation : Dans un environnement IPAO, l'intervieweur se voit imposer un fardeau supplementaire : le processus de saisie a lieu durant l'interview. Normalement, un intervieweur lit une question, reçoit une réponse, entre les données, appuie sur [enter] et voit ensuite apparaître le prochain écran et la question suivante. Comparativement à l'IPC, les intervieweurs ne peuvent pas regarder en avant et anticiper la question suivante tout en inscrivant les réponses à la question précédente, et ils ne peuvent pas commencer à lire la question suivante avant d'appuyer sur [enter] - ils ne peuvent pas s'occuper des deux tâches en même temps. Dans ce genre de procédure, l'interaction entre l'intervieweur et le répondant est segmentée par des touches [enter]. Il n'existe encore aucune indication quantitative que ce genre de segmentation nuit aux données ou au déroulement de l'interview. Par contre, on fait remarquer que l'intervieweur n'a pas de « vue d'ensemble » et que la pertinence des questions et les relations entre elles ne sont peut-être pas claires (House 1985; Groves et Mathiowetz 1984).

Nos résultats, obtenus de plusieurs séries d'essais de facilité d'utilisation menées en laboratoire sur la configuration des écrans d'une liste de membres de ménage (Couper et coll. 1997; Hansen, Couper and Fuchs 1998), ont mis en valeur une configuration d'écran particulière qui permet à l'intervieweur de mieux comprendre le questionnaire, tout en conservant l'interaction avec le répondant et tout en inscrivant les données. Deux versions d'une série de questions ont été mises à l'essai en laboratoire en fonction du temps mis à remplir le questionnaire et de la facilité d'utilisation. Nous avons comparé une conception dite thématique et une conception dite quadrillée. House et Nicholls (1988) ont décrit trois stratégies de conception des écrans pour des modules assistés par ordinateur : axée sur l'élément, axée sur l'écran et axée sur la forme. Un module axé sur l'élément fait afficher une question et un champ de saisie à la fois, et les opérations logiques ont lieu durant la transition d'un écran à l'autre. Facile à programmer, une telle conception dirige l'attention de l'intervieweur sur la question même. Un module axé sur l'écran regroupe plusieurs éléments supposant des réponses séquentielles. Toutes les opérations logiques sont exécutées après chaque élément. Pour ce qui est d'un écran axé sur la forme, plusieurs éléments sont présentés en même temps dans un tableau ou dans une grille, et l'intervieweur peut appuyer sur des touches de curseur pour passer d'un champ à l'autre et les remplir dans n'importe quel ordre.

La version axée sur l'élément vérifiée dans notre expérience correspond aux caractéristiques décrites par House et Nicholls (1988) pour une stratégie fondée sur l'écran. Pour sa part, la conception fondée sur une grille est en réalité un module axé sur la forme. Elle permet aux intervieweurs d'enregistrer l'information dans l'ordre choisi par le répondant; elle donne à l'intervieweur une meilleure vue d'ensemble du module, et les mises à jour ainsi que les reculs en arrière sont plus faciles (on trouvera des détails dans Couper et coll. 1997). De plus, elle permet à l'intervieweur d'avoir plus d'une question à un même écran (avantages : vitesse d'exécution et informations contextuelles). Le graphique ci-dessous illustre la conception des écrans fondée sur l'élément et celle fondée sur la grille pour les IAO.

Nous avons observé que la conception fondée sur la grille permet de réduire la segmentation. Les intervieweurs peuvent commencer à lire la question suivante tout en inscrivant les données de la question précédente. Même un recul en arrière semble plus facile dans la conception fondée sur une grille. Par contre, nous n'avons constaté qu'un faible appui pour la conception fondée sur une grille relativement au temps mis à exécuter la tâche (on trouvera des détails dans Couper et coll. 1997). On peut donc poser la question : Que faire pour réduire la segmentation et améliorer davantage l'efficacité d'une liste de membres de ménage pour ce qui est de la durée?

2. Manque de souplesse : La deuxième caractéristique qui risque de poser des problèmes dans une interview assistée par ordinateur est le manque de souplesse. Un des avantages du module IAO est que l'intervieweur ne peut guère sauter de questions. Même si les modules IAO peuvent faire amplement usage de l'enchaînement des questions et de filtres, l'ordre des questions est déterminé d'avance. Normalement, il faut appuyer sur la touche [enter] pour chaque question avant que le système ne passe à l'écran suivant. On trouve cette rigidité de l'ordre des questions avantageuse car elle évite toute difficulté de la part de l'intervieweur de suivre le déroulement, les questions destinées à des répondants particuliers, les filtres, l'enchaînement des questions, *etc.* L'intervieweur n'a pas à s'en soucier et peut s'en tenir à l'exécution des éléments. Par contre, il en résulte un ordre très rigide des questions et très peu de souplesse, pour l'intervieweur, quant à l'ordre des questions. Voici un exemple qui illustre bien cet effet : la plupart des modules IAO appliquent un ordre des questions à la liste des membres du ménage, toutes les questions adressées à une personne étant posées avant que l'intervieweur ne pose les mêmes questions à la personne suivante (au sujet de la « conception axée

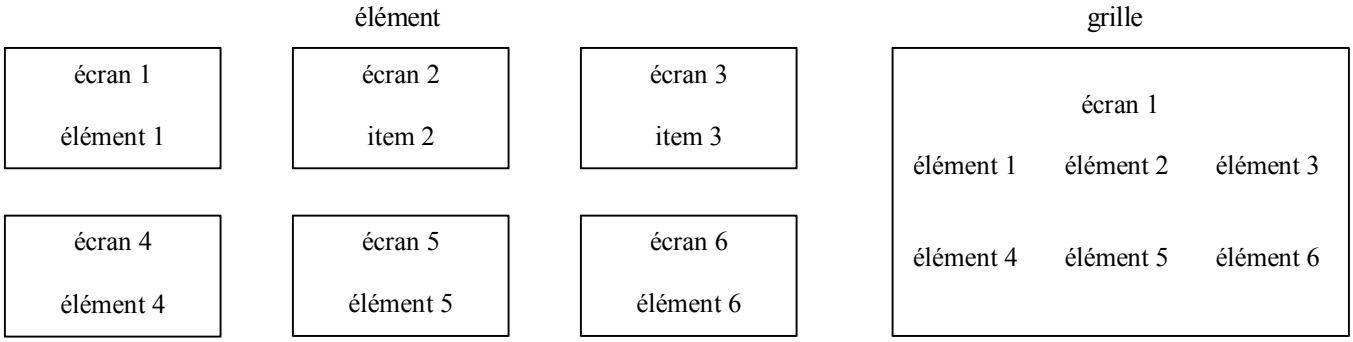

Figure 1. Conception fondée sur l'élément et conception fondée sur la grille.

sur la personne », voir Couper et coll. 1997 et Fuchs 2001; au sujet des « questions groupées », voir Moyer 1996). Ainsi, le module IAO peut exiger que l'on demande l'âge du répondant, son niveau de scolarité et d'autres questions avant de passer à l'âge de son épouse. (Cela s'explique en partie par le fonctionnement des programmes informatiques et des bases de données : les ménages représentent les principales catégories, les personnes ou les autres entités étant traitées comme des sous-catégories.) Lorsqu'on remplit le questionnaire pour une liste de membres du ménage, il arrive (et même assez souvent, comme on le verra ci-dessous) que le répondant fournisse non seulement la réponse à la question posée (par exemple « j'ai 34 ans »), mais également à une question connexe : « j'ai 34 ans et mon épouse en a 32 »; le répondant pourra aussi déclarer : « nous sommes tous noirs » lorsqu'il est question de race (Oksenberg, Beebe, Blixt et Cannell 1992).

En présence d'un questionnaire imprimé, il est facile pour un intervieweur d'utiliser immédiatement l'information supplémentaire fournie par le répondant. Pour une réponse du type « nous sommes tous noirs », par exemple, l'intervieweur peut facilement cocher les cases appropriées pour tous les membres du ménage en même temps. Une personne qui s'intéresse à la conception des questionnaires peut alors se poser la question suivante : vu le manque de souplesse d'un environnement assisté par ordinateur, quel est le meilleur ordre des questions si l'on veut recueillir des renseignements pour tous les membres du ménage?

Moore et Moyer ont présenté les résultats d'une expérience menée sur deux ordres des questions servant à recueillir des informations sur toutes les personnes admissibles d'un ménage (Moore et Moyer 1998a, 1998b). Selon le premier modèle, on pose toutes les questions pour la première personne admissible du ménage, puis on passe à la personne suivante. Un tel ordre des questions est ce que l'on appelle la stratégie axée sur la personne. Selon le deuxième modèle, appelé la stratégie thématique, la première question est posée pour toutes les personnes admissibles, puis on pose la deuxième question pour toutes les personnes, et ainsi de suite. Les résultats de Moore et Moyer appuient fortement une conception thématique : le modèle thématique comporte une non-réponse moins élevée, moins d'interruptions et de refus, et il prend appréciablement moins de temps. Les intervieweurs préfèrent nettement ce modèle.

Dans l'expérience que nous présentons ici, nous avons fait appel aux avantages de la stratégie thématique et de la conception des écrans fondée sur une grille : nous avons combiné les deux modèles d'écran (conception fondée sur l'élément et conception fondée sur la grille) et les deux ordres des questions (ordre axé sur la personne et ordre axé sur le thème) de façon à vérifier les quatre versions résultantes dans une expérience sur le terrain. Pour ce faire, nous avons adopté l'hypothèse qui suit : la facilité d'utilisation d'un module IAO ne relève pas seulement de la programmation, mais elle est liée à la conception du questionnaire et au contexte social de l'interview. Les deux aspects d'un module assisté par ordinateur, c'est-à-dire la conception des écrans et l'ordre des questions, peuvent favoriser ou détruire la bonne marche d'une interview. Les résultats de la recherche antérieure avaient donné lieu à l'hypothèse suivante : la combinaison d'une conception des écrans axée sur une grille et d'un ordre des questions axé sur un thème suscite l'interaction la plus efficace entre l'intervieweur et le répondant.

3. Méthode

L'expérience s'est déroulée en Allemagne en mars 1998. On a interrogé des immigrants d'origine allemande provenant de la Pologne, de la Roumanie et de l'ancienne Union soviétique. À compter du 28 février 1998, et jusqu'au 20 mars 1998, les intervieweurs ont mené $n = 501$ interviews. On a fait parvenir aux répondants une lettre d'introduction à l'avance, et on a communiqué avec eux par téléphone jusqu'à 15 fois. Le taux de réponse a été de 84 %, et la non-réponse à une question a été assez faible. Les interviews ont été menées dans le cadre du programme ITAO CI3. Quelque 95 questions ont été posées sur divers thèmes. En moyenne, les interviews ont duré 23 minutes.

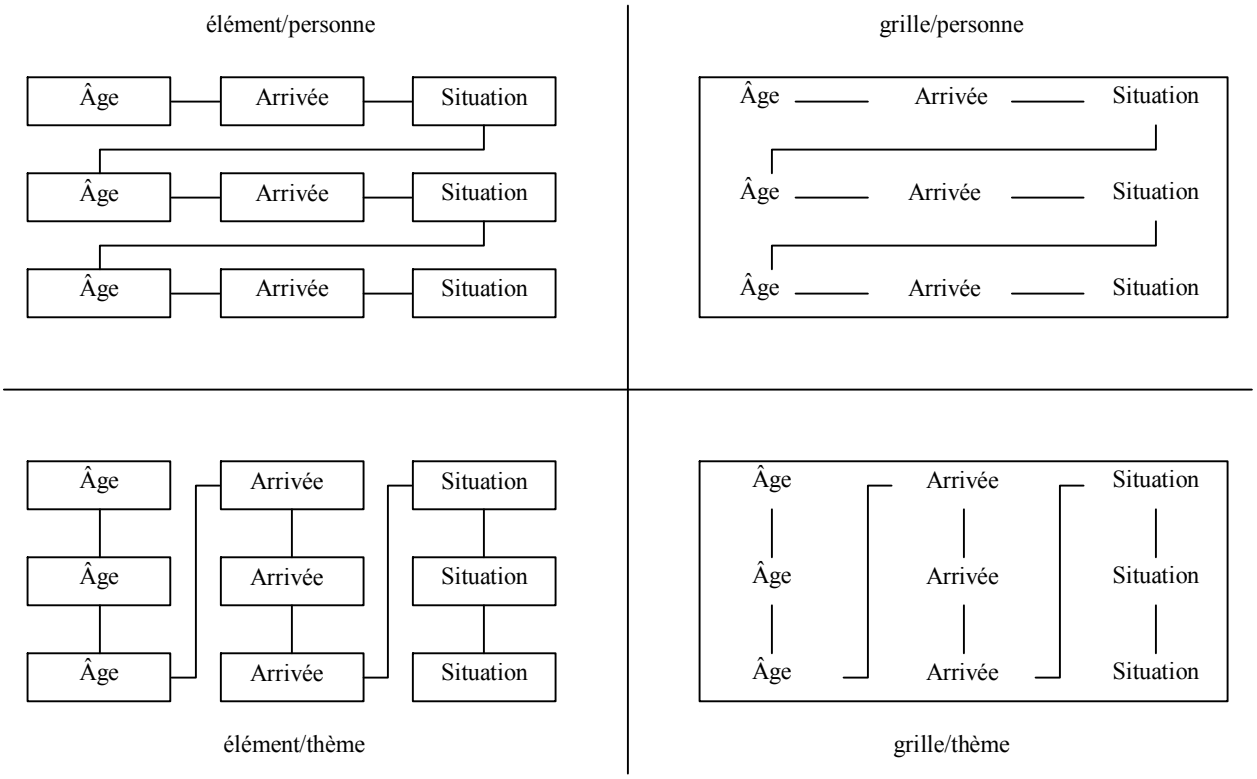

Figure 2. Quatre versions vérifiées dans l'expérience (chaque cases représente un écran).

Quatre versions d'une liste des membres d'un petit ménage, avec trois éléments par personne, ont été incluses dans le module : une version élément/personne, une version grille/personne, une version élément/thème et une version grille/thème. Toutes les versions ont fait appel au même libellé des questions et aux mêmes instructions pour l'intervieweur, mais nous avons modifié la conception des écrans et l'ordre des questions en fonction de la stratégie théorique décrite ci-dessus (figure 2). La version élément/personne est considérée comme la version standard - elle représente la conception des questionnaires appliquée habituellement aux volets socio-démographiques des enquêtes IAO. Une des quatre versions a été attribuée de façon aléatoire à chaque interview - et donc aux intervieweurs et aux répondants. Nous avons mesuré le temps total requis pour la liste des membres du ménage, ainsi que le temps consacré à chaque élément de cette section des 501 interviews. De plus, 234 interviews ont été choisies au hasard, et nous avons enregistré sur bande magnétoscopique l'intervieweur en train de parcourir la section de la liste des membres du ménage. Les segments vidéo ont été codés en fonction du comportement de l'intervieweur et du répondant, et les données résultantes ont été combinées aux mesures du temps.

4. Résultats

La durée a été appréciablement différente pour chacune des quatre versions : les intervieweurs ont consacré 6,6 secondes par élément dans la version élément/personne (considérée comme la version standard). Par contre, chaque élément a supposé 5,5 secondes dans la version grille/thème. C'est là une réduction de 17 % environ pour la version grille/thème. Les deux autres versions se situaient entre les deux premières.

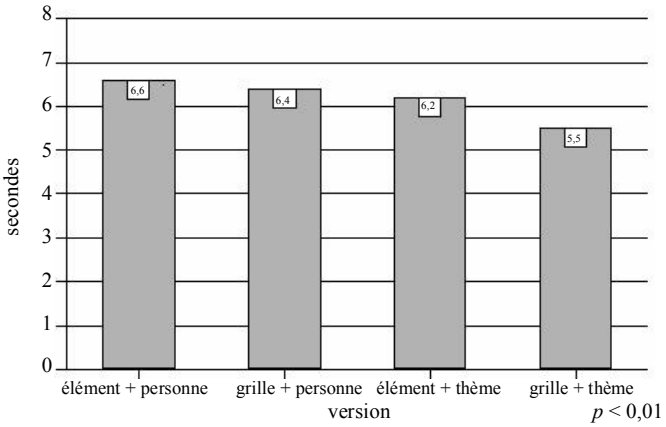

Figure 3. Durée par élément selon la version.

Il importe de mentionner que les deux facteurs semblent jouer un rôle dans la diminution du temps mis à accomplir la tâche. Si nous distinguons les deux facteurs, nous obtenons les résultats suivants : les deux versions axées sur le thème sont appréciablement plus courtes que les deux versions fondées sur la personne, et les deux versions fondées sur la grille supposent un temps appréciablement moindre que les deux versions axées sur la personne. L'effet combiné s'applique à la version grille/thème, et

donne une valeur de 5,5 secondes par élément. (Une analyse de la variance indique que les deux facteurs - c'est-à-dire la conception des écrans et l'ordre des questions - favorisent indépendamment la diminution du temps (conception des écrans : $p < 0{,}01$, un tiers de l'effet total; ordre des questions : $p < 0{,}001$, deux tiers de l'effet total, aucune interaction appréciable).)

Pourquoi la version grille/thème est-elle plus rapide? Une analyse détaillée indique que cette version est particulièrement rapide pour la collecte des renseignements sur la deuxième personne et les personnes subséquentes du ménage - effet appréciable que l'on appelle l'effet de boucle (Fuchs 2001). Cette expression décrit le phénomène suivant : l'intervieweur consacre beaucoup plus de temps à recueillir les renseignements pour la première personne du ménage comparativement aux personnes subséquentes. L'effet de boucle moyen donne au total 3,4 secondes par élément, ce qui représente une réduction de 38 % environ relativement à la première personne (tableau 1).

L'effet de boucle n'est pas particulièrement propre à cette expérience. Nous avons observé des effets de boucle dans nos expériences antérieures liées à la liste des membres des ménages NHIS (Couper et coll. 1997). Toutefois, il est intéressant de noter que l'effet de boucle est appréciablement plus grand pour les versions axées sur le thème qu'il ne l'est pour les versions qui suivent un ordre des questions axé sur la personne (tableau 1). Ainsi, les versions fondées sur le thème permettent d'accélérer le processus pour la deuxième personne et les personnes subséquentes d'un ménage, et comportent donc un effet de boucle plus grand. (Il est possible que la conception de notre expérience a eu comme effet que des intervieweurs ne savaient pas quelle version ils utilisaient, ce qui a pu faire baisser leur rendement pour le tout premier élément. Or cet effet devrait être le même pour toutes les versions, de sorte que les résultats ne devraient pas être touchés.)

Tableau 1
Durée et effet de boucle (en secondes)

Élément	Âge	Arrivée	Situation	Tous les éléments
Durée par élément				
Première personne du ménage	9,4	9,9	7,7	9,0
Toutes les autres personnes du ménage	6,6	5,7	4,5	5,6
Toutes les personnes	8,0***	7,8***	6,1***	7,3***
Effet de boucle				
Différence entre la première et toutes les autres personnes du ménage	-2,8	-4,2	-3,2	-3,4
Effet de boucle selon la version				
Grille + thème	-6,8	-6,4	-4,6	-5,9
Élément + thème	-5,2	-8,3	-4,3	-6,0
Grille + personne	-0,5	-2,7	-3,0	-2,1
Élément + personne	0,3	-0,3	-1,3	-0,4
Effet de boucle moyen	-2,8***	-4,2**	-3,2**	-3,4***

**$p < 0{,}01$; ***$p < 0{,}001$

L'analyse des bandes vidéo nous a permis de fournir des raisons partielles au moins pour ces différences : dans des conditions axées sur le thème, les intervieweurs et les répondants s'adaptent différemment à l'interview comparativement aux versions fondées sur la personne. Lorsque l'on pose les questions pour toutes les personnes du ménage, le répondant reconnaît très rapidement le caractère logique du processus. Dans une forte proportion des cas (30 % environ), la réaction est du type : « nous sommes tous arrivés pendant la même année » (ce qui veut dire : « cessez de me répéter cette question »).

Si le questionnaire suit la conception fondée sur la personne, l'intervieweur doit mémoriser cette information, et au moment d'aborder la personne suivante, il doit se rappeler : « ne répète pas cette question, car le répondant a déjà fourni la bonne réponse! ». Parfois l'intervieweur s'en souvient, mais la plupart du temps il pose la question de nouveau. Cela est particulièrement vrai pour une configuration d'écran axée sur l'élément et ne fournissant aucune indication des réponses à la même question pour les autres membres du ménage. Dans un modèle fondé sur le thème, par contre, l'intervieweur peut facilement s'adapter à cette situation. Il suffit alors d'entrer le même code pour toutes les personnes du ménage sans répéter la question. L'intervieweur aussi bien que le répondant s'habituent aux questions, de sorte que le processus question-réponse comporte moins d'interventions verbales de la part de l'intervieweur comme de celle du répondant. L'intervieweur aussi bien que le répondant peuvent anticiper la prochaine question, et l'interview se déroule plus facilement. Cela est particulièrement vrai lorsque le module IAO comporte une grille et fournit d'autres informations contextuelles, par exemple, les réponses fournies pour d'autres membres du ménage à la même question. (Les résultats figurant dans la partie inférieure de la figure 4 permettent de conclure que la version grille/personne ne bénéficie pas dans la même mesure des avantages de la stratégie fondée sur le thème. Toutefois, à cause de la conception de la grille, l'effet de boucle est appréciablement plus grand que pour la version élément/personne). Par conséquent, le temps consacré par élément est appréciablement plus court, et l'intervieweur peut adopter un comportement orienté vers le répondant semblable à celui dont il est question dans Schober et Conrad (1997).

L'inclusion de remarques par l'intervieweur sert parfois de signal que la réponse à la question précédente a été inscrite de façon à aiguiller le répondant, celui-ci devinant la question suivante et fournissant la réponse appropriée même sans stimulation supplémentaire. Dans un cas extrême, cela peut mener le répondant à fournir l'information au sujet de toutes les personnes du ménage en même temps : « nous sommes tous venus pendant la même année ». Les diverses versions mises à l'essai dans cette expérience favorisent ce genre de comportement à différents degrés. Les résultats nous permettent de conclure que la version grille/thème pousse les intervieweurs et les répondants à

dévier de l'interview prévue de façon plus marquée que les autres versions. Pour ce qui est de la durée, cette version permet à l'intervieweur d'utiliser de façon efficace l'information fournie pour tous les membres du ménage en même temps. L'analyse du codage vidéo appuie notre interprétation de l'occurrence propre à une version de comportements des intervieweurs (1) et de comportements des répondants (2) qui permettent de gagner du temps :

1. L'analyse des bandes vidéo permet d'observer des comportements d'intervieweur ne correspondant pas aux procédures standard : quelque 78 % des éléments sont lus tels quels, et 9,3 % des éléments ne sont pas présentés au répondant par l'intervieweur. De plus, dans 5 % des cas, l'intervieweur ne lit pas la question, mais fournit plutôt un autre stimulus comportant le lien entre la personne suivante et le répondant (par exemple « ... et votre épouse? »). (Il est intéressant de noter que les intervieweurs choisissent les mêmes expressions verbales, spontanément, que Moore et Moyer 1998a, 1998b ont utilisées dans leurs expériences sur l'ordre des questions.) Dans 5,5 % des cas, l'intervieweur ne lit pas la question, mais vérifie plutôt la réponse (« ... et votre épouse a 32 ans? ».) On observe également quelques questions incomplètes et des réponses inscrites incorrectement. Au total, 22 % environ des éléments sont touchés par au moins un comportement d'intervieweur qui ne correspond pas à une séquence d'interview standard - ce qui représente une valeur étonnamment élevée puisque tous les intervieweurs savaient que les interviews étaient enregistrées sur bande magnétoscopique! Toutefois, comparativement à d'autres études sur le comportement des intervieweurs, les valeurs sont appréciablement moins élevées. Oksenberg, Cannell et Blixt (1996), par exemple, ont appliqué un codage du comportement à la National Medical Expenditure Survey, observant de 37 % à 41 % de comportements d'intervieweur de ce genre. Il sera question ci-dessous de la mesure dans laquelle ce genre de comportement aide à obtenir des mesures valables.

De ces résultats particuliers, nous tirons la conclusion suivante : la plupart de ces comportements représentent une espèce de raccourci, par exemple l'intervieweur ne lit pas le texte exact de la question, tente de faciliter la conversation ou de la rendre plus conforme aux règles de la conversation. À notre avis, cela indique que l'intervieweur ne veut pas poser une question à laquelle une réponse a déjà été fournie. L'intervieweur ne veut pas rester indifférent devant l'information fournie par le répondant, cherchant plutôt à respecter les règles de la conversation. Ces comportements prennent donc moins de temps que des comportements d'intervieweur standard. Selon nous, la priorité ne consiste donc pas à épargner du temps, mais bien à personnaliser le processus question-réponse en fonction de comportements du répondant non prévus et ne pouvant pas être intégrés par le module assisté par ordinateur.

Table 2
Comportement de l'intervieweur et comportement du répondant selon la version

	Grille + thème	Élément + thème	Grille + personne	Élément + personne	Total
Proportion moyenne d'éléments touchés par un comportement d'intervieweur ne correspondant pas à l'interview consignée selon le cas	0,48	0,43	0,34	0,21	0,36***
Le répondant fournit l'information pour toutes les personnes du ménage en même temps	38,2 %	44,4 %	29,0 %	10,8 %	29,7 %***

***$p < 0{,}001$

Afin de distinguer la proportion de cas touchés par un certain comportement du répondant et la proportion moyenne d'éléments par cas (!) touchés par un certain comportement de l'intervieweur, nous avons utilisé des pourcentages pour ceux-là et des décimales pour ceux-ci.

Afin de pouvoir comparer les quatre versions du modèle d'écran en fonction des comportements d'intervieweur ne correspondant pas à l'interview standard, nous avons calculé la proportion des éléments par cas touchés par ce genre de comportement. Les quatre versions manifestent des écarts fort différents relativement à la séquence prévue des interviews : l'application de la version grille/thème à une interview entraîne plus de deux fois plus de comportements de ce genre (la proportion moyenne des éléments touchés est de 0,48) que la version élément/personne (0,21) qui représente la norme pour la plupart des études menées jusqu'à présent. (Une analyse de la variance indique que les deux facteurs influencent l'effet global indépendamment (conception des écrans : $p < 0{,}001$; ordre des questions : $p < 0{,}001$; aucun effet appréciable d'interaction). On peut attribuer 25 % environ de l'effet global à la conception des écrans et 75 % environ à l'ordre des questions.) Cela influence aussi le temps consacré à l'interview : les éléments touchés par un comportement d'intervieweur qui diffère de la séquence prévue de l'interview prennent appréciablement moins de temps (4,0 secondes) que les éléments qui respectent la séquence ordinaire (6,8 secondes; $p < 0{,}001$).

2. De plus, une analyse du comportement des répondants indique que le modèle thématique suscite une plus forte proportion de cas (42,3 % comparativement à 19,7 % pour la stratégie personnelle; $p < 0{,}001$) dans lesquels le répondant fournit au moins une fois des renseignements pour toutes les personnes ou pour un groupe de personnes du ménage en même temps (par exemple « nous sommes tous venus pendant la même année »; « nous avons tous le même statut juridique »). Par contre, la différence entre le modèle

axé sur la grille et le modèle axé sur l'élément est appréciablement plus faible (33,6 % contre 26,1 %), sans toutefois devenir significative. L'analyse révèle cependant un effet d'interaction appréciable ($p < 0,05$) : le recours à un ordre des questions axé sur le thème n'entraîne pas une différence appréciable. Toutefois, une conception fondée sur une grille jumelée à un ordre des questions axé sur un thème fait augmenter le nombre de cas dans lesquels le répondant fournit des renseignements pour tous les membres du ménage en même temps.

Il est remarquable que les résultats diffèrent même pour les deux conceptions des écrans lorsqu'on utilise un ordre des questions axé sur la personne. L'étude a été menée par téléphone, les répondants n'étant pas du tout au courant de la conception des écrans. La seule explication possible est que l'intervieweur modifie son comportement en fonction de la conception des écrans, aiguillant le répondant de façon différente. Les répondants aussi bien que les intervieweurs réagissent donc à la conception des écrans et à l'ordre des questions dans le modèle de type grille/personne d'une façon qui facilite l'interaction intervieweur-répondant, ce qui permet à l'interview de se dérouler plus doucement. (Comme nous l'avons vu, les intervieweurs modifient leur comportement même pour un modèle de type grille/ personne (tableau 2), mais l'ordre des questions ne suscite pas chez le répondant un comportement équivalent.)

Un inconvénient possible de ces comportements intervieweur-répondant pourrait être une baisse de la qualité des données à cause de modifications du processus question-réponse prédéterminé; le répondant considère alors la réponse de façon moins intense et rigoureuse. Nous n'observons que de très rares valeurs manquantes pour une question, de sorte que l'analyse de cet indicateur standard de la qualité des données n'est pas efficace. En réalité, nous ne nous attendons pas à une plus forte proportion de valeurs manquantes pour une question dans l'une ou l'autre version. Toutefois, on pourrait s'interroger sur l'homogénéité des réponses fournies par le répondant. Dans une forte proportion des cas, le répondant entend le texte complet de la question une seule fois et il est possible que le répondant soit moins rigoureux au moment de répondre à la même question pour des membres subséquents du ménage. De plus, le fait de répondre pour tous les membres du ménage en même temps (« nous sommes tous arrivés pendant la même année») risque d'accroître l'homogénéité de la réponse et donc de faire baisser la qualité des données.

Afin de pouvoir évaluer cet inconvénient possible, nous avons calculé le nombre de catégories de réponse distinctes choisies par le répondant pour un élément particulier et pour tous les membres du ménage (par exemple pour l'année d'arrivée : le répondant 1985, le partenaire 1987, la fille 1987, le fils 1988 = trois catégories de réponse distinctes). Cela devrait indiquer si les répondants qui adoptent le raccourci (« nous sommes tous venus pendant la même année ») sont seulement ceux pour qui c'est réellement le cas, ou bien si même les autres répondants ont fourni une réponse pour tous les membres du ménage même s'ils auraient dû choisir des catégories de réponse distinctes à cause de la situation dans leur ménage particulier (malheureusement nous n'avons aucune validation externe des réponses fournies). En examinant le nombre moyen de catégories de réponse distinctes (tableau 3), nous ne remarquons aucune différence pour ce qui est de l'homogénéité des données. Pour l'année d'arrivée comme pour le statut juridique (à titre d'Allemand ou à titre d'étranger), il n'existe aucune différence visible entre les versions. Pour toutes les versions, le nombre moyen de catégories de réponse distinctes choisies (une pour chaque personne d'un ménage) ne manifeste aucune différence appréciable

Tableau 3
Nombre moyen de catégories distinctes (homogénéité) par ménage selon la version

Variable	Grille + thème	Élément + thème	Grille + personne	Élément + personne	Total
Année d'arrivée (19 catégories)	1,2	1,2	1,2	1,3	1,2
Statut (4 catégories)	1,3	1,3	1,3	1,3	1,3

Aucune différence appréciable

Ces résultats ne fournissent qu'une faible indication que le modèle fondé sur une grille ne nuit pas à la qualité des données. Il faudra utiliser d'autres indicateurs standard de la qualité des données comportant des ensembles de données plus grands si l'on veut déterminer dans quelle mesure la qualité des données est affectée. Toutefois, d'après les données disponibles, nous ne pouvons déceler un effet sur la validité des réponses.

5. Discussion et conclusion

Les résultats de notre comparaison de quatre versions d'une liste des membres d'un ménage (le même libellé des questions étant utilisé pour toutes les versions) montrent que les intervieweurs tout comme les répondants fonctionnent de façon plus efficace dans le cadre d'un modèle de type grille/thème que pour l'une ou l'autre des trois autres versions. Le fait de combiner un schéma des écrans fondée sur une grille et un ordre des questions fondé sur un thème permet de réduire la durée moyenne de 17 % environ. Deux tiers de cette réduction peuvent être attribués à l'ordre des questions, et un tiers environ à la configuration des écrans. Il importe de mentionner que l'effet de schéma des écrans est moins marqué que celui de l'ordre des questions et encore plus faible, comparativement à l'effet sur la durée, pour le comportement de l'intervieweur et pour le comportement du répondant.

Même si l'effet de la conception quadrillée sur le comportement de l'intervieweur et sur celui du répondant n'est pas considérable, il aide à cerner deux raisons du meilleur rendement de la version grille/thème pour ce qui est de la durée de l'interview : 1) pour la version grille/thème, l'intervieweur aussi bien que le répondant s'adaptent mieux à la logique du processus question-réponse, les deux pouvant anticiper la prochaine question plus aisément et le processus question-réponse se déroulant plus doucement; 2) cette version entraîne plus de cas dans lesquels le répondant fournit des renseignements plus rapidement pour les personnes du ménage, et fournit plus souvent des renseignements pour tous les membres du ménage ou du moins pour plusieurs membres en même temps. Les résultats ne sont pas tout à fait uniformes, mais cette version particulière permet à l'intervieweur de s'adapter plus facilement à la situation, d'inscrire les renseignements et d'aider le répondant à fournir la prochaine réponse appropriée sans répétition du texte complet de la question.

Nos résultats se rapportent à la discussion sur la conception des questionnaires d'enquête pour la collecte des données assistée par ordinateur avec la participation d'un intervieweur. Les résultats décrits dans le présent exposé permettent de conclure que le recours à des grilles facilite l'interaction intervieweur-répondant et accélère la collecte des données. Les expériences que nous avons menées, dans le laboratoire de facilité d'utilisation du Survey Research Center de l'Université du Michigan sur des modèles quadrillés et thématiques, indiquent qu'il est possible d'améliorer le rendement des intervieweurs en leur fournissant des grilles (Couper et coll. 1997). Moore et Moyer (1998a; 1998b) ont montré qu'il est possible d'améliorer l'efficacité des interviews en adoptant également un ordre des questions fondé sur un thème. Le présent exposé indique qu'une combinaison des deux fonctions est encore plus avantageuse pour les interviews.

Le recours à des grilles et à un ordre des questions fondé sur un thème suscite un nombre accru de cas dans lesquels l'intervieweur s'éloigne de l'interview standard. D'un point de vue strictement méthodologique, cela peut représenter un sérieux inconvénient, surtout si l'intervieweur s'éloigne de la séquence d'interview standard en ayant recours à des questions globales pour tous les membres du ménage. Ainsi, Martin (1999) a décelé un accroissement appréciable du nombre de personnes dénombrées dans un ménage lorsque l'on pose d'autres questions. De plus, Kindermann et coll. (1997) ont montré, pour des questions fondées sur une liste hors-ménage, que des signaux supplémentaires sur la victimisation entraînent une augmentation appréciable des crimes dénombrés. Si l'on généralise ces résultats pour des questions globales d'une personne à l'autre, on s'attendrait à une diminution de la qualité des données lorsque les intervieweurs ont recours à des comportements non prévus consistant à appliquer des questions globales. Toutefois, les résultats signalés dans le présent exposé n'indiquent pas que les intervieweurs ont recours à des questions globales, et l'auteur ne recommande pas d'avoir recours à des questions globales lors de la conception d'un questionnaire d'enquête. Les résultats favorisent plutôt une conception des écrans permettant à l'intervieweur d'utiliser les renseignements fournis par le répondant lorsque celui-ci adopte un mode global et fournit des renseignements pour plusieurs membres du ménage en même temps. Il ne s'agit donc pas d'encourager les chercheurs à utiliser des questions globales, et nous ne souhaitons pas que les intervieweurs modifient les questions prévues de façon à poser des questions globales. Toutefois, lorsqu'un répondant fournit plus de renseignements qu'on ne lui en demande, la conception des écrans d'un questionnaire IAO ne devrait pas empêcher l'intervieweur d'en bénéficier.

Il a été démontré qu'un modèle fondé sur une grille facilite la tâche des intervieweurs à cet égard, car il permet à ceux-ci de rajuster leur comportement en fonction des règles générales de la conversation. Les résultats du codage des comportements indiquent que les intervieweurs s'éloignent fréquemment des procédures d'interview standard. Oksenberg et coll. (1992, 1993) ont affirmé que ces changements reflètent souvent des rajustements apportés par les intervieweurs devant les modalités de la situation de façon à mieux enchaîner avec ce qui précède immédiatement ou de respecter la situation particulière du répondant. Cela s'impose plus précisément lorsque le répondant ne se limite pas à la question demandée, mais fournit des renseignements supplémentaires. Comme l'ont souligné Oksenberg et coll. (1992, 1995), afin de ne pas donner l'impression de ne pas porter attention à la réponse, les intervieweurs inscrivent souvent la réponse eux-mêmes sans poser la question ou ne posent la question qu'en partie. Ils adoptent donc un comportement axé davantage sur le répondant afin de ne pas paraître insensibles. Une conception des écrans fondée sur une grille et un ordre des questions fondé sur un thème permettent à l'intervieweur de respecter ces règles de la conversation et de réagir en fonction de la situation. Cela pourrait être acceptable ou même préférable, en présence de questions factuelles, pourvu que le comportement de l'intervieweur ne nuise pas à la qualité des données (par exemple questions tendancieuses).

Ce qu'il faut faire afin d'améliorer un questionnaire assisté par ordinateur du point de vue de l'interaction intervieweur-répondant : nos données indiquent que la version grille/thème suscite un déroulement particulier de l'interview auquel l'intervieweur et le répondant peuvent s'adapter facilement. Jeff Moore (1996; Moore et Moyer 1998a, 1998b) a indiqué que les intervieweurs préfèrent la version axée sur un thème. Par contre, nous connaissons mal la satisfaction des répondants à l'égard de cet ordre des questions. Il est donc important d'évaluer leur opinion au sujet de la version différente. De plus, nous ne savons pas si cette version correspond à la façon dont l'information est mémorisée par les répondants. Il se peut que les répondants puissent facilement s'adapter à cette version, mais que d'un

point de vue cognitif et social (fardeau) ou du point de vue de l'exactitude des réponses, ce ne soit pas la bonne méthode. De plus, il y a lieu de déterminer si nos résultats se laissent transposer d'une liste des membres d'un ménage à d'autres segments d'un questionnaire. Nous menons actuellement des essais sur le terrain afin de comparer différents modèles en fonction d'informations factuelles dans un contexte autre que celui d'une liste des membres d'un ménage et pour des catégories d'attitudes. Les versions étudiées dans la présente expérience comportent différents degrés d'information contextuelle fournis à l'intervieweur dans l'exécution d'un questionnaire (questions précédentes, questions suivantes, *etc.*). La question générale est celle-ci : Que se passe-t-il si nous avons davantage recours à des grilles ou à des écrans axés sur la forme? Dans quelles circonstances est-il utile d'améliorer l'efficacité des interviews et quelles sont les limites d'une telle stratégie? Il est cependant trop tôt pour présenter des résultats.

Il y a également lieu d'aborder des questions restées sans réponse à l'aide de recherches plus poussées. Personnellement, j'aimerais proposer une stratégie particulière servant à évaluer ce genre de question à supposer que la conception de questionnaires assistés par ordinateur ait de l'importance pour différents clients : chercheurs, intervieweurs et répondants. Bien entendu, il est important qu'un questionnaire IAO réponde aux besoins du chercheur et que le processus question-réponse soit bien conçu dans chaque cas. Toutefois, à mon avis, il est également important de tenir compte de la dimension sociale de l'interaction intervieweur-répondant et des comportements qui surviennent entre les questions. Si un questionnaire IAO perturbe la dimension sociale du processus de mesure, il risque de nuire à la qualité même des données. Nous ne savons toujours pas quelle stratégie entraîne les meilleurs compromis entre un processus de mesure valable et fiable d'une part, et une interview qui se déroule doucement, brièvement et sans embarras d'autre part. Afin de déterminer dans quelle mesure une conception particulière des écrans IAO pourrait nuire à la qualité des données et comment elle pourrait aider à épargner du temps, de l'argent et des efforts de la part de l'intervieweur, nous devons mener d'autres études de facilité d'utilisation.

Afin d'évaluer les questions mentionnées ci-dessus, nous avons vraiment besoin d'autres expériences sur le terrain. Puisque nous voulons analyser la dimension sociale de l'interview et ses effets sur le comportement de l'intervieweur de même que sur la durée de l'interview, les expériences de laboratoire ne répondent pas à tous nos besoins. Bien sûr, les expériences de laboratoire se font dans un contexte plus contrôlé, fournissent des renseignements plus détaillés au sujet des participants et, par conséquent, exigent un moins grand nombre de cas. Et pourtant, sans essais sur le terrain, nous ne confronterons jamais nos prototypes et nos solutions en fonction de la véritable interaction intervieweur-répondant et aussi du processus question-réponse. On devrait donc étudier la facilité d'utilisation à l'aide d'une combinaison d'expériences de laboratoire et d'essais sur le terrain.

Remerciements

Une partie des résultats a été présentée à un colloque tenu à l'Institute for Social Research de l'Université du Michigan le 21 mai 1998, et à l'occasion de la 54e assemblée annuelle de l'American Association for Public Opinion Research le 16 mai 1999. Nous remercions les intervieweurs qui ont participé à la présente expérience. Mick Couper, Siegfried Lamnek, Jeffrey Moore et deux examinateurs anonymes ont fourni des suggestions utiles au sujet de versions antérieures du présent exposé.

Bibliographie

Baker, R.P. (1992). New technology in survey research: Computer-assisted personal interviewing (CAPI). *Social Science Computer Review*, 10, 145-157.

Couper, M.P., Baker, R.P., Bethlehem, J., Clark, C.Z.F., Martin, J., Nicholls, W.L. et O'reilly, J. (Éds.) (1998). *Computer Assisted Survey Information Collection*. New York : John Wiley & Sons, Inc.

Couper, M.P., et Burt, G. (1994). Interviewer attitudes toward computer-assisted personal interviewing (CAPI). *Social Science Computer Review*, 12, 38-54.

Couper, M.P., Groves, R.M. et Kosary, C. (1989). Methodological issues in CAPI. *Proceedings of the Section on Survey Research Methods*, American Statistical Association, 349-354.

Couper, M.P., Fuchs, M., Hansen, S.E. et Sparks, P. (1997). CAPI Instrument Design for the Consumer Expenditure (CE) Quarterly Interview Survey. Rapport final. University of Michigan.

Fuchs, M. (1994). *Umfrageforschung mit Telefon und Computer*. Einführung in die computergestützte telefonische Befragung. Weinheim: Psychologie Verlags Union.

Fuchs, M. (1995). Die computergestützte telefonische Befragung. Einige Antworten auf Probleme der Umfrageforschung. *Zeitschrift für Soziologie*, 24, 284-299.

Fuchs, M. (2001). The impact of technology on interaction in computer-assisted interviews. (Éd., D.W. Maynard, H. Houtkoop-Steenstra, N.C. Schaeffer, H. van der Zouwen). *Standardization and Tacit Knowledge*: *Interaction and Practice in the Survey Interview*. Wiley (à paraître).

Fuchs, M., Couper, M. et Hansen, S. (2000). Technology effects: Do CAPI or PAPI interviews take longer? *Journal of Official Statistics* (sous press).

Groves, R.M., et Mathiowetz, N.A. (1984). Computer assisted telephone interviewing: effect on interviewers and respondents. *Public Opinion Quarterly*, 48, 356-369.

Hansen, S.E., Couper, M.P. et Fuchs, M. (1998). Usability Evaluation of the NHIS Instrument. Article présenté à l' Annual Meeting of the AAPOR, St. Louis, MO.

House, C.C. (1985). Questionnaire design with computer assisted telephone interviewing. *Journal of Official Statistics*, 1, 209-219.

House, C.C., et Nicholls, W.L. (1988). Questionnaire design for cati: Design objectives and methods. (Éds., R.M. Groves, P.P. Biemer, L.E. Lyberg, J.T. Massey, W.L. Nicholls et J. Waksberg). *Telephone Survey Methodology*. New York : John Wiley & Sons, Inc., 421-436.

Laurie, H., et Moon, N. (1997). Converting to CAPI in a Longitudinal Panel Study. Document de travail de l' ESRC Research Centre on Micro-Social Change, 97-11, Essex.

Martin, E. (1999). Who knows who lives here? Within-household disagreements as a source of survey coverage error. *Public Opinion Quarterly*, 63, 220-236.

Moore, J.C. (1996). Person- *vs*. Topic-based Design for Computer-Assisted Household Survey Instruments. Article présenté à la InterCASIC '96, International Conference on Computer-Assisted Survey Information Collection, San Antonio, TX.

Moore, J.C., et Moyer, H.L. (1998a). ACS/CATI Person-Based/ Topic-Based Field Experiment - Final Report. Center for Survey Methods Research, U.S. Bureau of the Census.

Moore, J.C., et Moyer, H.L. (1998b). Questionnaire Design Effects on Interview Outcomes. Article présenté à l'Annual Meeting of the AAPOR, St. Louis, MO.

Moyer, L.H. (1996). Which is better: Grid listing or grouped questions design for data collection in establishment surveys? *Proceedings of the Section on Survey Research Methods*, American Statistical Association, 986-990.

Nicholls, W.L., et De Leeuw, E. (1996). Factors in acceptance of computer-assisted interviewing methods: A conceptual and historic review. *Proceedings of the Section on Survey Research Methods*, American Statistical Association, 758-763.

Oksenberg, L., Beebe, T., Blixt, S. et Cannell, C. (1992). *Research on the Design and Conduct of the National Medical Expenditure Survey Interviews*. Rapport final. Survey Research Center, Ann Arbor, USA.

Oksenberg, L., Cannell, C. et Blixt, S. (1996). Analysis of Interviewer and Respondent Behavior in the Household Survey. U.S. Department of Health and Human Services. AHCPR No. 96-N016.

Projektgruppe SOEP (1998). Funktion und Design einer Ergänzungsstichprobe für das Sozio-oeconomische Panel. Diskussionspapiere des DIW, 163, Berlin.

Schneid, M. (1991). Einsatz computergestutzter Befragungssysteme in der Bundesrepublik Deutschland. Ergebnisse einer Umfrage. ZUMA-Arbeitsbericht 91/20. Mannheim : ZUMA.

Schober, M.F., et Conrad, F.G. (1997). Does conversational interviewing reduce survey measurement error? *Public Opinion Quarterly*, 61, 576-602.

Suchman, L., et Jordan, B. (1990). Interactional troubles in face-to-face survey interviews. *Journal of the American Statistical Association*, 85, 45-54.

Weeks, M.F. (1992). Computer-assisted survey information collection: A review of CASIC methods and their implications for survey operations. *Journal of Official Statistics*, 8, 445-465.