Propension à fournir des données personnelles mensongères sur Internet : une étude exploratoire

Caroline Lancelot Miltgen

--- SIM ---

Propension à fournir des données personnelles mensongères sur Internet : une étude exploratoire

Caroline Lancelot Miltgen
Maître de Conférences en Sciences de Gestion
GRANEM (Groupement de Recherche ANgevin en Economie et Management)
Université d'Angers, UFR Droit – Economie - Gestion
caroline.miltgen@univ-angers.fr

Prix de thèse FNEGE AFM 2007 et CREIS 2007 pour sa thèse sur les réactions des consommateurs face à la collecte de données personnelles.
Responsable scientifique d'un contrat de recherche sur 'vie privée et systèmes d'identification électronique' pour la Communauté Européenne (Nov. 2007-Nov. 2008, budget : 57 000 €).
Responsable du Master 2 'Management et Sécurité des Systèmes d'Information de Santé', Institut Supérieur de la Santé et des Bioproduits d'Angers (ISSBA) – Université d'Angers

Cet article a été publié dans la revue Systèmes d'Information et Management (SIM) n° 14 (3) publié en 2009.

Accès à l'article : http://www.revuesim.org/sim/article/view/247

Propension à fournir des données personnelles mensongères sur Internet : une étude exploratoire

L'essor des pratiques de collecte des données clients, notamment depuis l'arrivée d'Internet, contraint chaque jour davantage de consommateurs à adopter des mesures de protection de leur vie privée. Si certains choisissent de ne pas répondre aux requêtes des sites, d'autres préfèrent fournir des données mensongères. Dans quel cas les individus mentent-ils ? Malgré son intérêt académique et managérial, on observe une absence de réponse à cette question. Cette recherche contribue à combler ce manque en étudiant la propension des internautes à 1) répondre à une sollicitation de données personnelles en ligne et à 2) fournir des données mensongères. Trois catégories d'antécédents – individuels, perceptuels et situationnels - à ces deux types de comportements sont ainsi identifiés et intégrés au sein du modèle conceptuel. Pour le tester, 252 internautes panélistes ont été soumis à un processus expérimental consistant à compléter un formulaire comportant des données plus ou moins sensibles.
Nos résultats viennent enrichir et/ou confirmer – dans un contexte hexagonal - la littérature. La valeur perçue de l'échange – rarement introduite dans les modèles existants – influence ainsi à la fois la probabilité de compléter le formulaire et de fournir des données mensongères. Nous démontrons aussi que la sollicitation de données sensibles, une forte préoccupation pour le respect de la vie privée et l'appartenance à la gente masculine se traduisent par une forte probabilité de mentir. Enfin, la préoccupation aurait un rôle modérateur – que la littérature n'évoque pas - du lien entre sensibilité du formulaire et intention d'y répondre. Nos résultats ont des implications intéressantes à la fois pour les chercheurs et pour les organisations souhaitant améliorer leur dispositif de collecte de données clients.

Mots clefs : Internet, comportement utilisateur, données personnelles, préoccupation pour le respect de la vie privée, mensonge

Internet users' intention to give inaccurate personal data online: an exploratory study

The rise of consumer data-gathering requests has each day forced more consumers to adopt privacy protection strategies, and more specifically since the arrival of the Internet. Some people choose not to answer the requests; others sometimes prefer to provide inaccurate data. Why do some individuals choose to lie? In spite of academic and managerial interest, there is little if no answer to this question. This research contributes to fill this gap by studying the propensity of Net surfers 1) to answer a request for on line personal data and 2) to (or not to) provide untrue data. Three categories of antecedents - individual, beliefs and situational - to both behaviors are identified and integrated within the model. So as to test it, 252 Net surfers panelists were subjected to an experimental process consisting in filling out a form with more or less sensitive data. Our results come to enrich and/or confirm the literature in a French context. The perceived value in this exchange - seldom introduced into the existing models – influences both the probability of filling out the form and of providing untrue data. We also prove that a request of sensitive data, a strong consumer privacy concern and being a male result in a strong probability of lying. Lastly, the privacy concern would have a moderating role - which the literature does not evoke - on the link between the sensitivity of the form and the intention to answer it. Our results have interesting implications for both the researchers and the organizations wishing to improve their consumer data-gathering procedures.

Key words: Internet, user behaviour, personal data, concern for information privacy, lying

1. INTRODUCTION

Les données, notamment sur les clients, jouent un rôle de plus en plus important pour la croissance des entreprises. Ces données ne sont néanmoins utiles que si elles sont exactes, complètes et à jour. Or, une étude du cabinet d'analyse américain Gartner de 2007 souligne l'accroissement des données fausses, incorrectes ou imprécises qui représenteraient environ 25% des informations détenues aujourd'hui par les grandes entreprises. Disposer de données correctes et bien gérées est pourtant un impératif économique aussi bien qu'une obligation désormais légale pour les entreprises, qui doivent se conformer aux réglementations Sarbanes-Oxley aux États-Unis et Basel II en Europe. Au-delà des problèmes légaux que cela pose, ces données erronées engendrent des surcoûts qui deviennent problématiques. Si la qualité des données est médiocre, les équipes commerciales peuvent perdre des contrats voire même des clients. Sans parler des dépenses faites inutilement pour l'envoi, par exemple, de courriers postaux. Au-delà des erreurs liées au stockage, au traitement et à la transmission des données, c'est au moment de la collecte des informations qu'il convient surtout d'être rigoureux. Or, si cette collecte se faisait traditionnellement en face à face, par téléphone ou par courrier, l'arrivée d'Internet a permis une augmentation de la vitesse avec laquelle les données sont collectées, mais aussi du montant d'informations susceptibles d'être récoltées.

Toutefois, cet essor des pratiques de collecte et d'utilisation des données clients ne s'opère pas sans que, parallèlement, les consommateurs y trouvent à redire. Les enquêtes d'opinion réalisées sur le sujet[1] révèlent que les citoyens sont de plus en plus sensibles aux risques que font peser ces pratiques et qu'ils se méfient des situations de sollicitation de leurs données personnelles[2]. D'une part, parce qu'ils n'aiment pas forcément se dévoiler, d'autre part, parce qu'ils ont peur des conséquences possibles d'un tel dévoilement, notamment d'une utilisation ultérieure abusive des données qu'ils auront bien voulu fournir. Les sondages indiquent aussi que les individus sont autant préoccupés par ce que les entreprises peuvent connaître d'eux que par la manière dont elles ont obtenu et utilisent ces informations (Katz et Tassone 1990). En général, les personnes interrogées estiment ainsi qu'on leur demande trop de données personnelles (Chen et Rea 2004 ; Jamal, Maier et Sunder 2005) et qu'elles ont peu de contrôle sur ce qu'il en advient (Chen et Rea 2004 ; Nowak et Phelps 1992 ; Phelps Nowak et Ferrell 2000). La sollicitation de données en ligne engendre donc des réactions variées chez

[1] Enquête IPSOS réalisée en octobre 2008 pour la CNIL et Baromètre de l'intrusion réalisé par ETO en mai 2008 parmi les sondages les plus récents réalisés en France.

[2] Toute information pouvant être associée à un individu identifié ou identifiable et donc susceptible de violer sa vie privée, en cas de divulgation à autrui sans son autorisation préalable.

les consommateurs : d'un dégoût résigné pour certains, à un boycott absolu pour d'autres, en passant par un bouche-à-oreilles négatif pouvant être rapidement relayé par les médias. Si l'une des premières mesures prises par les internautes consiste à ne pas répondre aux demandes des sites, il en est une autre beaucoup plus pernicieuse et de plus en plus répandue : le mensonge. Ainsi, Teltzrow et Kobsa (2004) trouvent que les utilisateurs d'internet, pour 34%, ont déjà menti à propos de leurs habitudes et de leurs préférences. De même, 47% des internautes allemands interrogés dans un sondage cité par Berendt et al. (2005) ont avoué avoir déjà fourni de fausses informations.

Dans quel cas les individus mentent-ils ? Quel est l'impact du niveau de préoccupation pour le respect de la vie privée ? Y a t-il d'autres éléments explicatifs ? Que peut-on faire pour éviter ces cas de mensonge ? Etonnement, malgré leur intérêt à la fois académique et managérial, très peu de recherches - hormis les travaux de Lwin et Williams (2003) et Lwin, Wirtz et Williams (2007) - se sont intéressées à ces questions, encore moins dans le contexte français. Rodhain et Agarwal (2001) soulignent toutefois l'importance d'étudier ces questions éthiques. Nous proposons donc de tester ici un modèle visant à déterminer la propension des utilisateurs à répondre à une sollicitation de données personnelles en ligne et à fournir (ou non) des données erronées. Une donnée erronée[3] peut se définir comme une donnée soit inexacte (comportant des erreurs ou différente de la réalité) soit mensongère (l'utilisateur fournit volontairement une information qu'il sait erronée). C'est au cas des données mensongères[4] que nous nous intéresserons ici. Trois types d'antécédents sont alors testés : 1) individuels, 2) situationnels et 3) perceptuels, à travers le concept de valeur perçue (rapport coûts/bénéfices). Ce dernier est en effet rarement étudié en tant que tel[5] dans la littérature bien qu'il soit unanimement considéré comme important (e.g. Castaneda et Montoro 2007, Chellappa et Shivendu 2007 ; Culnan et Milberg 1998 ; Hoffman, Novak et Peralta 1999). Cette approche permet de prendre en compte le dilemme souvent ressenti par

[3] Nous n'étudierons pas le cas où il s'agit d'une erreur système puisque nous ne nous intéressons qu'à la phase de collecte et au comportement de l'utilisateur.

[4] La distinction entre donnée mensongère et erronée est peu évidente, les deux termes étant souvent considérés comme synonymes dans la littérature anglo-saxonne. La différence entre les deux concerne surtout le point de vue duquel on se place : le consommateur fournit une donnée mensongère (qu'il sait différente de la réalité), celle-ci devenant alors, une fois enregistrée dans le fichier de l'entreprise, une donnée erronée. Le terme « mensonger » souligne donc le caractère volontairement faussé de l'information fournie, il insiste sur le comportement, l'action. Celui d'erroné ne fait que constater le résultat sans s'intéresser outre mesure aux raisons ayant aboutit à ce constat. Il est plus large que le précédent, une donnée pouvant être erronée pour d'autres raisons qu'un mensonge. S'agissant d'étudier le comportement de l'internaute nous emploierons donc majoritairement le terme « mensonger », celui d'erroné étant plus adapté quand nous évoquerons les conséquences pour l'entreprise.

[5] Certains travaux étudient à la fois les coûts (à travers la préoccupation pour le respect de la vie privée) et les bénéfices à fournir des données personnelles (e.g. Hui, Tan et Goh 2005). Mais, à notre connaissance, aucune recherche n'a jusqu'à présent cherché à mesurer directement cette évaluation du rapport coûts/bénéfices - à travers le concept de valeur perçue -. dans le cadre d'une décision de dévoilement de soi.

l'utilisateur quand il doit prendre la décision de fournir les données qui lui sont demandées : ai-je plus à perdre ou à gagner à dévoiler ces informations ?

Notre contribution vise donc à mieux comprendre la réaction de l'utilisateur au regard des menaces liées au dévoilement de renseignements personnels en ligne. Identifier les antécédents de la fourniture de données – a fortiori si elles s'avèrent erronées – pourra servir aux chercheurs qui étudient la complexité du comportement de l'internaute et aider les managers qui cherchent à encourager la fourniture de données et l'achat en ligne. Nous établirons ainsi que la valeur perçue de l'échange a un effet sur la réaction à la sollicitation aussi bien sur l'acceptation de répondre (ou non) que sur la possibilité de fournir (ou non) des données mensongères. Nous démontrerons aussi que la sollicitation de données sensibles, une forte préoccupation pour le respect de la vie privée et l'appartenance à la gente masculine se traduisent par une plus forte probabilité de mentir. Enfin, une moindre intention de remplir le formulaire semble se dégager pour les individus très préoccupés par le respect de leur vie privée et soumis à une collecte de données sensibles. Nos résultats complètent et enrichissent donc largement la littérature sur ce thème en SI et marketing par : 1) l'étude du comportement mensonger (en réalité l'intention de mensonge), rarement étudié jusque là malgré son intérêt managérial et académique évident ; 2) l'étude de l'influence de la valeur perçue, également rarement introduite dans les modèles existants ; 3) le test de relations équivoques (comme le lien préoccupation - intention de répondre) pour lesquels les résultats de la littérature sont parfois contradictoires ; 4) la validation de certains liens dans un contexte hexagonal.

Une première partie examine les notions de préoccupation pour le respect de la vie privée et de valeur perçue dans la littérature. Dans la deuxième partie, nous détaillons les hypothèses de recherche et proposons un modèle explicatif du comportement de l'utilisateur mensonger que nous testons par le biais d'une approche expérimentale auprès d'un échantillon de panélistes français. Dans la dernière partie, nous présentons puis discutons les résultats obtenus.

2. REVUE DE LA LITTERATURE

Cette partie revient tout d'abord sur la définition du concept de préoccupation pour le respect de la vie privée (RVP) et sur le lien entre préoccupation et réactions (ou intentions) comportementales. Elle vise aussi à identifier les travaux sur ce thème en management des systèmes d'information (SI) et en marketing.

2.1 Préoccupation pour le respect de la vie privée : définitions et fondements théoriques

Les définitions du respect de la vie privée sont multiples car les travaux réalisés dans ce domaine sont issus de nombreuses disciplines, incluant les SI et le marketing. Le point commun à ces définitions est la notion de contrôle. De fait, la majorité des travaux sur le sujet se focalisent sur le droit à contrôler l'accès à ses données personnelles. Reprenant la conception de Westin (1967, p.7), nous définirons ainsi la préoccupation pour le respect de la vie privée - dans son aspect informationnel[6] (*"Informational privacy concern"* en anglais) - comme une préoccupation qui renvoie à « l'exigence d'individus, groupes ou institutions de déterminer à quel point et de quelle manière l'information les concernant est communiquée aux autres ». Il existe un lien direct entre préoccupation pour le respect de la vie privée et dévoilement de soi ("self-disclosure" en anglais). Le dévoilement de soi correspond au fait de révéler des informations sur sa personne à d'autres individus (Jourard et Lasakow 1958). Il est donc un moyen - parmi d'autres - de contrôler l'accès à ses données personnelles, et plus généralement de protéger sa vie privée. En ce sens, il est aussi concerné lorsqu'un consommateur est sollicité pour remplir un formulaire comprenant des données à caractère personnel. Dans une telle situation en effet, l'entreprise convie l'individu à fournir des informations, ce qui revient à lui demander de s'engager dans une forme de dévoilement. Pendant longtemps, l'étude du dévoilement de soi a majoritairement, si ce n'est exclusivement, concerné l'aspect interpersonnel c'est-à-dire les relations entre individus. Dans ce cadre, il serait soumis à une triple influence : celle de facteurs culturels, individuels et situationnels. Il serait ainsi fonction du sexe de la personne qui se dévoile, mais aussi de la cible (personne à qui on se confie), du sujet de la discussion (et du degré d'intimité qui lui est lié) et de la relation entre les partenaires (Benner 1968). A l'instar de Moon (2000), cette littérature a ensuite été étendue à l'étude du dévoilement de soi dans un contexte commercial.

Depuis plusieurs décennies, de nombreux travaux à la croisée du Management des SI et du marketing ont été menés à ce sujet. Par exemple, Phelps et al. (2000) montrent que 50% des individus demandent plus d'information et de transparence sur l'utilisation qui est faite de leurs données personnelles une fois celles-ci collectées. De même, la décision d'un individu de fournir ses données personnelles à autrui (ou à une entité commerciale) serait basée sur un « privacy calculus » (Laufer et Wolfe 1977) c'est-à-dire sur une évaluation des coûts et bénéfices à partager l'information. Les recherches ont également étudié les antécédents et les conséquences de la préoccupation pour le RVP (e.g. Phelps, D'Souza et Nowak 2001) ainsi que l'impact du niveau et du type de régulation (Milberg et al. 1995) sur cette dernière.

[6] Pour les autres aspects du concept de préoccupation pour le RVP, nous renvoyons notamment à Solove (2002).

2.2 Sollicitation de données et préoccupation : facteurs d'influence et types de réponse

Nous décrivons ici les éléments individuels, situationnels et perceptuels susceptibles d'influencer les comportements visant à faire respecter sa vie privée.

2.2.1 Des facteurs situationnels aux déterminants individuels

Au-delà des éléments structurels (taux d'équipement, réglementation) et/ou culturels susceptibles d'influencer les réactions individuelles (Rodhain et Agarwal 2001) mais qui n'interviennent pas dans un contexte d'étude homogène (hexagonal), deux types d'éléments sont liés à la volonté de faire respecter sa vie privée : ceux liés à la situation (facteurs situationnels) et à l'individu (individuels).

Les facteurs situationnels

Plusieurs éléments contextuels sont susceptibles d'intervenir dans la décision d'un individu de faire respecter sa vie privée lors d'une sollicitation de ses données personnelles. Lancelot-Miltgen (2006) distingue 4 blocs d'éléments situationnels susceptibles d'influencer la réponse à une sollicitation de données personnelles. On trouve ainsi les facteurs liés : 1) à la politique de RVP mise en place par la firme collectant les données ; 2) au demandeur c'est-à-dire à l'entreprise qui sollicite les données; 3) au dispositif de collecte et, enfin, ceux liés 4) aux « circonstances de collecte ». Les deux premières catégories ont été largement étudiées dans la littérature, aboutissant notamment à montrer l'impact des chartes de confidentialité (e.g. Arcand et Nantel 2005) ou du contrôle fourni au consommateur (Castaneda et Montoro 2007) ainsi que de la réputation de (Andrade, Kaltcheva et Weitz 2002 ; Xie, Teo et Wan 2006) – ou de la familiarité avec - l'entreprise sollicitant les renseignements (Castaneda et Montoro 2007) sur la propension au dévoilement de soi. Les facteurs liés aux circonstances de collecte tels que le moment où la sollicitation est faite et l'humeur dans laquelle se trouve le consommateur au moment où il est sollicité sont des éléments sur lesquels l'entreprise a peu d'emprise. Bien que potentiellement influents, ils seront donc écartés ici au profit d'éléments plus pragmatiques. Les facteurs liés au dispositif de collecte[7] nous intéresseront plus particulièrement ici du fait 1) du manque de travaux empiriques sur ce sujet ou des résultats contradictoires que ces études engendrent[8], 2) de l'intérêt académique à étudier ce type de variables et 3) de leur caractère éminemment opérationnel. Parmi ces facteurs, nous

[7] Par dispositif de collecte, nous entendons tout ce qui concerne le média de collecte, le mode de collecte (directe vs. discrète), le design de l'instrument de collecte, et les données sollicitées (nature, quantité, sensibilité).

[8] A titre d'exemple, si plusieurs auteurs (e.g. Castaneda et Montoro 2007 ; Faja et Trimi 2006) trouvent un lien négatif entre le niveau de sensibilité des données demandées et la propension à répondre, Hui et al. (2007) ne trouvent aucun lien significatif entre les deux.

focaliserons plus particulièrement notre attention sur la sensibilité des données que tous les auteurs s'accordent à considérer comme un élément déterminant dans la décision de fournir des renseignements personnels (Lwin, Wirtz et Williams 2007 ; Margulis 2003 ; Westin 2003). D'une part, les réactions des utilisateurs dépendent majoritairement de la nature des informations requises (Castaneda et Montoro 2007 ; Phelps et al. 2000 ; Sheehan et Hoy 2000 ; Wang et Petrison 1993). Toute chose égale par ailleurs, révéler des informations confidentielles est perçu comme plus risqué que s'il s'agit d'informations moins sensibles (Cranor et al. 1999 ; Milne et Gordon 1993). D'autre part, une étude récente sur le sujet (Hui et al. 2007) vient contredire les conclusions précédentes en montrant que la sensibilité des données n'affecterait pas – toujours – la propension à dévoiler des renseignements sur soi. Cette contradiction amène donc à mener davantage de recherches sur ce point et à intégrer cette variable – la sensibilité des données - dans notre modèle en lui donnant la définition suivante, traduction de celle établie par Weible (1993) : « la sensibilité d'une donnée correspond au niveau d'inquiétude qu'un individu ressent pour cette donnée ».

Les déterminants individuels

Les éléments individuels susceptibles d'influencer la réaction du consommateur face à une sollicitation de ses données personnelles appartiennent à quatre catégories différentes. On distingue : 1) les variables sociodémographiques (genre, âge, CSP …) (e.g. Milne et Boza 1999 ; Phelps, Nowak et Ferrell 2000) ; 2) les variables d'expérience comme la connaissance et l'attitude de l'individu face aux pratiques de collecte et d'utilisation de données (e.g. Culnan 1993 ; Pavlou 2002) ; 3) les variables d'ordre psychologique (e.g. Lee et Turban 2001 ; Smith et al. 1996), liées à la personnalité de l'individu (tendance à se dévoiler, capacité d'extraversion, propension à faire confiance …) et, enfin, 4) les variables d'ordre idéologique (e.g. Dinev et Hart 2006), liées aux valeurs de l'individu (valeur accordée à l'intimité ou préoccupation pour le respect de la vie privée notamment). L'étude de l'influence des trois premières catégories de déterminants individuels est délicate tant les résultats en la matière sont contradictoires. Si, Nowak et Phelps (1992) trouvent que les individus les plus jeunes sont les plus nombreux à demander de retirer leur nom du fichier suite à une sollicitation de données personnelles, Milne et Rohm (2000) trouvent exactement l'inverse (les plus âgés sont davantage demandeurs). Une des raisons à ces résultats contradictoires serait liée à l'effet potentiellement indirect de telles variables sur la réponse des individus, par l'intermédiaire d'une prédisposition à répondre mesurée à travers le niveau de préoccupation pour le respect de la vie privée (Lancelot Miltgen 2006). De par son intérêt majeur dans la littérature, cette dernière sera donc intégrée dans notre modèle en tant que variable de recherche.

2.2.2 Les éléments perceptuels

Les recherches issues de la littérature ainsi qu'une étude qualitative préalable (Lancelot Miltgen 2003) permettent de dénombrer quatre dimensions majeures sur lesquelles les consommateurs évaluent la sollicitation de leurs données : la confidentialité, sensibilité et pertinence perçues des données sollicitées ainsi que l'évaluation du rapport coûts/bénéfices.

Parmi ces perceptions, nous focaliserons notre attention sur le rapport coûts/bénéfices à fournir des données - mesuré à travers, la valeur perçue de l'échange - sachant que cette notion est 1) particulièrement importante pour l'utilisateur, qui accepte rarement[9] de délivrer des informations personnelles s'il n'obtient pas quelque chose en échange (Ashworth et Free 2006), 2) rarement étudiée empiriquement dans la littérature sur la vie privée bien que souvent considérée comme centrale (Hui et al. 2007 ; Xie, Teo et Wan 2006).

Dans une approche économique et traditionnelle, la valeur est une notion subjective qui correspond à un surplus de bénéfices, déduction faite des coûts monétaires et psychologiques liés à la réalisation d'une transaction. Ce concept intervient généralement comme une variable explicative du comportement et est définie comme « une confrontation cognitive (ratio ou différence) entre ce qui est reçu (utilité, qualité, bénéfices) et ce qui est donné (prix, coûts, investissement, sacrifices) » (Bolton et Drew 1991, Wooddruff 1997). Zeithaml (1988) estime ainsi que le consommateur conçoit la valeur comme une fonction positive d'éléments tangibles donnés et comme une fonction négative des sacrifices financiers et/ou non financiers consentis. Ce principe de valeur perçue des données livrées dérive de la justice procédurale (Thibaut et Walker 1975). Selon cette dernière, les individus considèrent une procédure comme juste à partir du moment où ils ont un contrôle sur celle-ci. Ici, nous considérons ce concept dans le cas d'internautes qui souhaitent disposer d'un moyen de contrôle face à une procédure de sollicitation de données personnelles dans laquelle ils investissent, sans pouvoir toujours en mesurer les conséquences à long-terme (e.g. Ashworth et Free 2006 ; Hui et al. 2007). En effet, quand on s'apprête à donner des renseignements personnels, il existe souvent une perception de vulnérabilité dans la mesure où l'utilisation ultérieure de ceux-ci pourrait affecter de manière négative son avenir (Ashworth et Free 2006). A l'opposé, il y a parfois un intérêt personnel à répondre, notamment quand on peut obtenir quelque chose en échange sous forme de cadeaux, bons d'achats ou autres contre-parties commerciales (Berendt et al. 2005 ; Lancelot Miltgen 2003 ; Xie, Teo etWan 2006).

[9] Lancelot Miltgen (2006) montre que seule une frange (20%) de la population française (les « bienveillants ») accepte volontiers de fournir des informations personnelles, sans contrepartie spécifique.

D'un point de vue théorique, la valeur est donc une différence entre des coûts et des bénéfices, raison pour laquelle nous utiliserons ce concept dans notre modèle pour mesurer l'arbitrage opéré par les individus confrontés à une sollicitation de leurs données.

2.2.3 Principaux comportements face à la sollicitation de données personnelles

La préoccupation pour le RVP est liée aux craintes concernant l'accès ultérieur aux données qu'un individu aura bien voulu fournir. Ses conséquences sur le développement du e-commerce sont nombreuses et particulièrement préoccupantes (Berendt, Preibusch et Teltzrow 2008). Le niveau de préoccupation est ainsi la raison la plus fréquemment citée pour expliquer 1) le refus d'utiliser Internet (Westin 2001) et 2) le refus de commercer sur Internet ou avec un e-marchand spécifique (Arcand et Nantel 2005 ; Dinev et Hart 2004, 2005/6 ; Teltzrow et Kobsa 2004). Par conséquent, pour éviter les pertes potentielles liées à l'utilisation non autorisée des données les concernant, les individus peuvent prendre plusieurs mesures, leur réaction dépendant du type et de la gravité des conséquences qu'ils anticipent.

Sheehan et Hoy (1999) ont été les premiers à évoquer les stratégies adoptées par les internautes face à la collecte d'informations d'une part, et à l'utilisation des données pour des contacts marketing, d'autre part (envois d'offres commerciales non sollicitées notamment). Ils identifient six types de comportements susceptibles d'être adoptés par les consommateurs (internautes) en réponse à la collecte[10] et au traitement de leurs données personnelles[11] :

1) fournir des informations incomplètes et/ou mensongères ;

Cette solution semble fournir une réponse idéale au paradoxe soulevé par Cespedes et Smith (1993) sur la manière de participer au e-commerce sans perdre le contrôle de sa vie privée.

2) ne pas s'enregistrer sur le site (ne pas divulguer d'informations personnelles) ;

Cette solution est vue comme le moyen optimal pour protéger sa vie privée mais restreint la participation au e-commerce au strict minimum.

3) demander à se faire rayer des listes (ou fichiers) ;
4) ne pas lire les emails non sollicités ;
5) se plaindre auprès des firmes responsables de l'envoi d'emails non sollicités ;

Cette solution est globalement peu utilisée, mais elle constitue un type de réponse possible, notamment en cas de niveau d'insatisfaction élevé.

6) notifier au fournisseur d'accès la réception d'emails non sollicités

Ceci suggère que les individus cherchent une assistance vers qui se tourner pour se plaindre.

[10] Les réponses à la collecte de données peuvent être considérées comme des stratégies de dévoilement de soi.
[11] Si toute la typologie est présentée, seules les réponses à la collecte (stratégies 1 et 2) nous intéressent ici.

Pour leur part, s'inspirant des travaux de Stone et al. (1983), Stewart et Segars (2002) étudient plus particulièrement quatre comportements, qui reprennent certains de ceux identifiés précédemment : demander à retirer son nom d'un fichier, refuser de divulguer des données, se plaindre auprès de l'entreprise indélicate ou d'organismes gouvernementaux. Quelques auteurs ont étudié depuis les comportements des consommateurs en réponse à une sollicitation de données personnelles. Castaneda et Montoro (2007) distinguent ainsi les réponses directes visant à réduire l'invasion perçue (comme les plaintes) et les réponses indirectes qui visent à affecter la relation avec le demandeur. Nous résumons les principales variables comportementales identifiées dans la littérature dans le tableau suivant (tableau 1). Certaines sont des stratégies de dévoilement de soi (exemple : refus d'enregistrement et fourniture de données mensongères) quand d'autres constituent des réponses indirectes au traitement des données.

Tableau 1 - Principales variables comportementales identifiées dans la littérature

Auteurs	Variables comportementales
Milne (1997)	Inscription sur une liste de marketing direct, autorisation de transfert des données à des tiers
Sheehan et Hoy (1999)	6 comportements de réponse : enregistrement sur le site, fourniture d'informations mensongères ou incomplètes
Jarvenpaa et Tractinsky (1999)	Intention d'achat en ligne
Milne et Rohm (2000)	Désir de retrait du fichier
Phelps, Nowak et Ferrell (2000)	Intention d'achat à distance, désir de retrait du fichier
Miyazaki et Fernandez (2000,2001)	Intention d'achat en ligne
Stewart et Segars (2002)	4 intentions de comportement : retrait du fichier, refus de divulguer des données personnelles, plainte à l'entreprise ou aux organisations gouvernementales
Miyazaki et Krishnamurthy (2002)	Probabilité de divulguer 6 catégories de données, probabilité d'acheter en ligne
Chellappa et Sin (2005)	Probabilité de recours à des services de personnalisation
Castaneda et Montoro (2007)	Réponses directes (ex : plainte) ou indirectes
Lwin, Wirtz et Williams (2007)	Fourniture de données mensongères, utilisation de technologies protectrices et refus d'enregistrement

3. MODELE CONCEPTUEL ET HYPOTHESES

Parmi les stratégies de dévoilement de soi identifiées dans la littérature (cf. tableau 1), les deux comportements utilisateurs qui apportent le moins de valeur à une entreprise sont le fait de refuser de fournir les renseignements (refus de répondre) et le fait de fournir des données mensongères (mensonge). Le premier comportement est préoccupant car il se solde par une information incomplète qui empêche l'entreprise de mettre en œuvre une démarche

CRM efficace. La seconde hypothèse, celle des données mensongères, est rarement explorée dans la littérature - hormis les travaux de Lwin et Williams (2003) et Lwin, Wirtz et Williams (2007) - alors qu'elle est d'une importance critique pour l'entreprise (Krishnan et al 2005) : « un traitement erroné des données peut engendrer des modèles de décision significativement plus mauvais et des problèmes graves dans le e-CRM » (Padmanahan et Tuzhilin, 2005). Pour combler ce manque, nous mènerons notre analyse sur ces deux comportements (en réalité intentions comportementales) de réponse négatifs (refus et mensonge) face à une sollicitation de données personnelles, en focalisant sur le second.

3.1 Formulation des hypothèses de recherche

L'analyse de la littérature qui vient d'être présentée nous permet de proposer un ensemble d'hypothèses selon lesquelles la réaction d'un individu face à une sollicitation de ses données personnelles – à la fois concernant le refus ou au contraire l'acceptation de remplir le formulaire et la fourniture de données mensongères ou non (variables endogènes du modèle) - dépendrait de trois types d'éléments explicatifs (variables exogènes du modèle) :

- des éléments individuels, que nous mesurerons ici à travers la préoccupation de l'individu pour le respect de sa vie privée[12]
- des éléments situationnels dont nous testerons l'impact à travers l'influence du niveau de sensibilité des informations demandées
- des éléments perceptuels, mesurés ici à travers le concept de valeur perçue (rapport coûts/bénéfices) à répondre

Nous développons ci-après les hypothèses correspondant à l'influence de chaque variable exogène retenue sur chacune des variables endogènes choisies, soit 6 hypothèses au total.

3.1.1 Les hypothèses concernant le niveau de préoccupation pour le respect de la vie privée

Dans la littérature sur la vie privée, la préoccupation pour le RVP est souvent liée aux variables aval du processus de réponse, notamment aux intentions comportementales et aux comportements. Ainsi, Smith, Milberg et Burke (1996) notent qu'un faible niveau de préoccupation est associé à de fortes intentions comportementales de dévoiler les informations demandées. Stewart et Segars (2002) trouvent, pour leur part, que les individus à la préoccupation élevée ont une plus forte tendance à retirer leur nom de fichiers et à refuser

[12] A titre exploratoire, seul un élément sera étudié pour chacune des catégories de variables exogènes identifiées. La justification du choix de cet élément est donnée au paragraphe 2.2.

de divulguer des données. Les résultats obtenus par Farag et Krishnan (2003) confirment que la préoccupation pour le RVP est associée négativement à l'acceptation de fournir des informations personnelles. Ils montrent ainsi que les individus les plus préoccupés sont les moins enclins à accepter de partager leurs données dans un but de « profilage » (en anglais, *profiling*). Enfin, Dinev et Hart (2006) confirment l'impact négatif de la préoccupation sur l'intention de commercer en ligne. On remarque toutefois que les variables testées jusqu'ici sont sensiblement différentes d'une étude à l'autre. Du côté de la préoccupation, on distingue préoccupation de vie privée en général de celle spécifique à Internet ou à un marchand. Du côte des intentions, les études s'intéressent rarement à la réponse à un formulaire en général mais plus spécifiquement à la fourniture de données liées à un objectif particulier (i.e. profilage, commerce en ligne, etc …). De plus, les résultats obtenus jusqu'ici concernent des échantillons essentiellement nord-américains voire asiatiques. Comme le niveau de préoccupation dans la population est censé varier selon l'environnement juridique présent dans le pays (Milberg et al. 1995), il s'avère nécessaire de vérifier l'hypothèse du lien entre préoccupation et intention de fournir des données dans un contexte Européen, très protecteur en matière de données personnelles :

H1 : plus la préoccupation pour le respect de la vie privée est élevée, moins l'utilisateur fournira ses données personnelles

Pour Lwin et Williams (2003), dans un environnement où la confiance est faible, mentir est un moyen d'obtenir les bénéfices associés au dévoilement, sans s'exposer aux risques. Un consommateur peu enclin à la communication de ses données (forte préoccupation pour le respect de sa vie privée) peut donc accepter de répondre mais fournir des réponses mensongères, afin de limiter les risques liés à la fourniture de tels renseignements. Les individus très préoccupés s'inquiètent en effet moins de la collecte de leurs données que de l'utilisation qui peut en être faite. En particulier, ils craignent que l'utilisation frauduleuse ou non permise de leurs informations par des entreprises puisse conduire à des conséquences négatives pour eux (Dinev et Hart 2006, Van Slyke et al. 2006). Puisque leur niveau de préoccupation est élevé, leur perception des pertes potentielles liées à l'utilisation de données nominatives les concernant est également élevée. En effet, malgré les conséquences potentiellement positives liées à la fourniture de données personnelles (confort, rapidité …), les pertes tendent à être perçues comme plus élevées que les gains (Kahneman et Tversky 1979), notamment chez les individus fortement préoccupés qui focalisent leur attention sur les éléments négatifs (Ashworth et Free 2006). Par conséquent, pour prévenir un tel opportunisme de la part des entreprises et réduire l'état de tension lié aux conséquences possibles de leurs actes, ils préfèreront fournir des données mensongères qui réduiront toute

possibilité d'utilisation détournée des renseignements qu'ils auront bien voulu fournir. On peut alors formuler l'hypothèse suivante :

H2 : plus la préoccupation pour le respect de la vie privée est élevée, plus l'utilisateur livrera des données mensongères

3.1.2 Les hypothèses concernant l'impact de la valeur perçue

Selon Phelps et al. (2000), une des motivations importantes dans la décision de fournir des données personnelles est l'existence d'un intérêt personnel à répondre. Dans une recherche sur les facteurs explicatifs de la participation à une enquête, Sharp et Frankel (1983) déduisent que les gens qui estiment retirer un bénéfice direct de la participation sont moins nombreux à percevoir la participation comme un fardeau et sont donc plus enclins à répondre. Dans le secteur de la téléphonie, Bolton et Drew (1991) trouvent que la valeur perçue (bénéfices moins coûts) est un déterminant significatif des intentions comportementales des consommateurs à rester fidèle à leur opérateur. Hann et al. (2005) prouvent aussi que les résultats attendus du dévoilement d'informations, sous la forme de récompenses monétaires ou de gain de temps par exemple, sont associés à des valences positives qui accroissent la motivation à répondre. Un surplus perçu de bénéfice pourrait donc encourager le consommateur à répondre, tandis qu'un surcroît de coût pourrait au contraire le décourager. Ceci nous amène à formuler l'hypothèse suivante :

H3 : plus la valeur perçue est élevée, plus l'utilisateur fournira ses données personnelles

D'après l'approche coûts/bénéfices, l'utilisateur serait d'autant plus enclin à livrer des informations qu'il en tirerait un profit et ce, même si cela engendre certains coûts, notamment ceux liés au remplissage du formulaire et aux conséquences de ce remplissage. Ces coûts (de nature monétaire, cognitive ou psychologique) ne doivent cependant pas, du point de vue de l'utilisateur, dépasser le gain qu'il estime pouvoir en retirer. La notion de coût est ici comprise au sens large. De fait, en livrant des données confidentielles, l'utilisateur se met en situation de vulnérabilité (Jarvenpaa et Tractinsky, 1999). Il place sa confiance dans un système et devient alors sensible à la crédibilité du site (Doney et Cannon, 1997) ou à la compétence associée à ce dernier (McKnight, Choudhury et Kacmar, 2002). Dès lors, si l'utilisateur pense pouvoir retirer des gains de la fourniture de ses données personnelles mais estime que les coûts (actuels et futurs) liés au remplissage sont trop élevés ou s'il n'accorde pas suffisamment de confiance au site demandeur, il sera enclin à répondre positivement à la sollicitation mais à ne pas dire la vérité pour éviter toute conséquence dommageable et ainsi limiter les risques. D'où l'hypothèse suivante quant à l'effet sur l'intention de mentir :

H4 : plus la valeur perçue est élevée, moins l'utilisateur livrera des données mensongères

3.1.3 Les hypothèses concernant la sensibilité des informations

Parmi les facteurs importants lors de la prise de décision de dévoiler ou non des informations personnelles, on trouve le fait que les données soient considérées comme trop sensibles (Cranor, Reagle et Ackerman 1999). Une recherche récente (Hui et al. 2007) montre toutefois que la sensibilité des données n'aurait pas forcément d'impact, ni sur la propension à fournir des données, ni sur la propension à mentir. Ce résultat surprenant pourrait être lié aux conditions dans lesquelles cette étude a été menée et notamment au fait d'avoir indemnisé les personnes ayant répondu à l'enquête. Il amène cependant à tester à nouveau ces liens.

La littérature confirme en effet que la collecte d'informations sensibles telles que des données de santé (Kam et Chismar 2006 ; Rohm et Milne 2004) ou financières (Ward, Bridge et Chitty 2005) conduit le consommateur à un niveau de confiance plus faible et à une perception de risques plus élevés (Malhotra, Kim et Agarval 2004). Comme le confirment Lwin, Wirtz et Williams (2007), la sensibilité d'une information est souvent liée au caractère pertinent et approprié de sa sollicitation. Dans le domaine des sondages, considérer que certaines questions ne concernent pas l'entreprise est significativement lié à un comportement de non-réponse ou d'évitement (Singer 1984). Dès lors, les personnes qui estiment que les données demandées sont trop sensibles ou ne concernent pas l'entreprise auront tendance à anticiper des conséquences négatives au dévoilement (Malhotra, Kim and Agarwal 2004). Ceci est de nature à réduire leur motivation à répondre et, selon l'intérêt ou l'obligation que les personnes associent à la réponse, conduire soit à s'abstenir de répondre soit à mentir. Cette dernière solution (le mensonge) serait ainsi révélatrice d'une moindre sensibilité globale du formulaire ou d'un intérêt majeur à conduire la transaction. Le coût d'une renonciation à l'échange est en effet rarement gratuit (perte de confort, de plaisir ou de temps) et peut s'avérer supérieur au risque potentiel d'une utilisation abusive des données. Le mensonge est alors considéré comme un bon compromis permettant de réduire le risque lié à l'échange tout en obtenant les bénéfices proposés. Une telle stratégie est souvent utilisée dans les relations asymétriques pour obtenir un état désiré (Lauer et Xiaodong 2007). D'où les hypothèses suivantes :

H5 : plus la sensibilité des informations est élevée, moins l'utilisateur fournira ses données personnelles

H6 : plus la sensibilité des informations est élevée, plus l'utilisateur livrera des données mensongères

Le tableau 2 qui suit résume les statuts, définitions et mesures utilisés pour les variables de cette recherche.

Tableau 2 – Statuts, définitions et mesures des variables de recherche

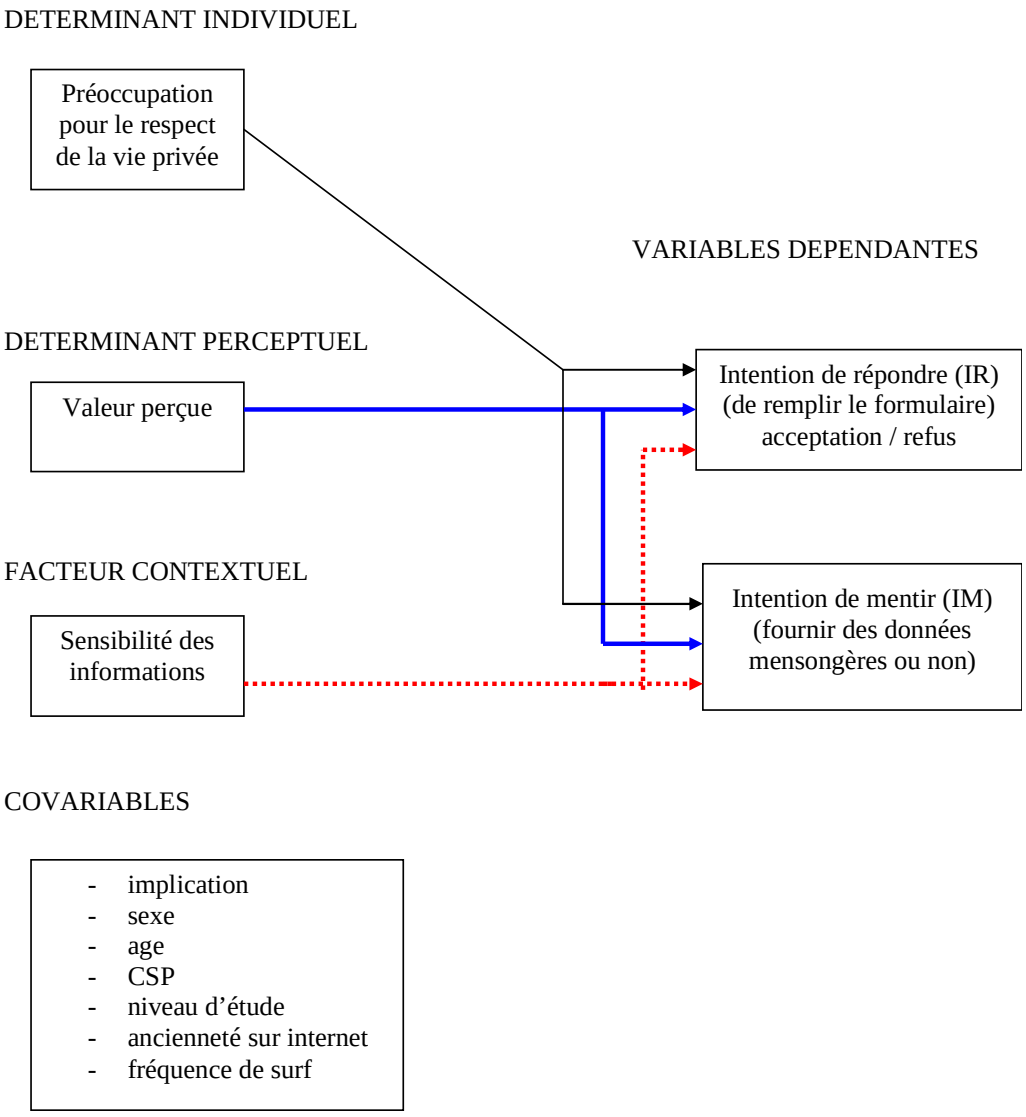
Variables	Définition	Mesure
Variables exogènes		
Préoccupation pour le respect de la vie privée	Exigence d'individus, groupes ou institutions de déterminer à quel point et de quelle manière l'information les concernant est communiquée aux autres (traduite de Westin 1967)	Likert en 7 points 3 items
Valeur perçue	Confrontation cognitive (différence) entre ce qui est reçu (bénéfices) et ce qui est donné (coûts) (traduite de Bolton et Drew 1991)	Sémantique différentiel 7 points 6 items
Sensibilité des informations	Niveau d'inquiétude qu'un individu ressent pour un type de donnée (traduite de Weible 1993)	Variable manipulée (2 niveaux : faible vs. élevée) et non mesurée
Variables endogènes		
Intention de répondre (ou refus)	Probabilité de remplir le formulaire de données personnelles proposé	1 item, 5 point
Intention de mentir	Probabilité de fournir des données mensongères	1 item, 5 point

3.2 Le modèle conceptuel

La figure 1 illustre notre modèle qui vise à tester l'influence de trois catégories d'attributs exogènes - individuels (préoccupation pour le respect de la vie privée), perceptuels (valeur perçue de l'échange), contextuels (sensibilité des données), tout en contrôlant l'effet de facteurs externes (covariables) - sur deux types de comportements (en réalité d'intentions comportementales) de réponse « critiques » : l'intention d'accepter (ou de refuser) de répondre à la sollicitation et celle de fournir (ou non) des données mensongères. Même si dans le cadre de cette recherche nous ne testons - à titre exploratoire - qu'un seul élément pour chaque catégorie de variable exogène identifiée, ce modèle est une intégration des travaux et des modèles existants dans le domaine de la fourniture de données personnelles et du respect de la vie privée. A noter qu'aucune hypothèse n'est émise concernant l'influence des covariables dans la mesure où elles seront simplement contrôlées lors de l'analyse[13].

[13] Ce point sera davantage explicité ultérieurement (cf. point 4 p18-19)

Figure 1 – Le modèle conceptuel



4. METHODOLOGIE

Afin de tester nos hypothèses, l'expérimentation s'est imposée comme la technique la plus appropriée. La démarche expérimentale nous permet en effet de comparer, de la manière la plus contrôlée possible, deux niveaux - chaque répondant n'étant affecté qu'à un de ces niveaux - de sensibilité (faible vs. élevée) des données contenues dans un formulaire, facteur dont nous souhaitons tester l'influence sur la réponse de l'individu (i.e. fourniture de données et exactitude de la réponse), en s'assurant une répartition équitable de l'échantillon sur cette variable. Ce choix du design expérimental vise aussi à mettre les individus interrogés dans une situation de collecte la plus réelle (en réalité simulée) et réaliste possibles, à travers le scénario et le formulaire présentés. Ainsi, la sensibilité est un élément fourni au répondant et non perçu par lui (même si nous vérifions que la manipulation est bien efficace) ce qui

correspond au statut que lui confère notre modèle : variable situationnelle et non élément perceptuel (comme la valeur perçue). Il s'agit d'un élément situationnel puisqu'il est déterminé par la situation, en l'occurrence par l'entreprise sollicitant les données qui définit (à travers le nombre et le type d'informations demandées) le niveau de connaissance client qu'elle souhaite obtenir. A travers ce design expérimental, nous reproduisons donc la situation d'une entreprise qui souhaiterait tester l'impact d'un formulaire contenant des données plus ou moins sensibles[14] sur le taux et la qualité des réponses fournies.

L'expérimentation permet de tester de manière solide les relations causales, en contrôlant l'effet potentiellement dommageable de facteurs externes (Cook et Campbell 1979). Eu égard à la littérature existante, plusieurs caractéristiques du consommateur sont susceptibles d'exercer une influence sur la réponse à une sollicitation de données. Même si l'affectation aléatoire des traitements testés à chacun des 252 individus participant à l'expérimentation constitue la voie fondamentale pour contrôler l'influence de ces variables externes (Brown et Melamed 1990), elle n'équilibre leurs effets qu'en moyenne sur toutes les affectations possibles (Evrard, Pras et Roux, 2003). Aussi, afin de pouvoir vérifier a posteriori la validité interne de l'expérimentation, les caractéristiques du consommateur susceptibles d'exercer (selon la littérature) une influence ont systématiquement fait l'objet d'une mesure.

Dans notre cas, trois types de facteurs externes (covariables) susceptibles d'influencer la réponse à une sollicitation de données ont ainsi mesurés et contrôlés : 1) le niveau d'implication du répondant vis-à-vis de la catégorie de produits/services[15] étudiée ; 2) son origine sociodémographique en terme d'âge, de sexe, de CSP et de niveau d'études (Culnan 1995, Milne et Rohm 2000, Phelps et al. 2000, Wang et Petrison 1993) ; 3) son expérience et utilisation du média Internet (Culnan 1995, Milne et Rohm 2000, Malhotra et al. 2004, Phelps et al. 2000, Wang et Petrison 1993).

En résumé, trois types de contrôles ont été utilisés dans le cadre de cette expérience :

- une affectation aléatoire des individus interrogés aux différents traitements
- un contrôle *a priori* pour les variables de genre (facteur bloqué[16]) et d'âge[17]
- un contrôle statistique *a posteriori* sur l'ensemble des facteurs externes identifiés, qui consiste à les intégrer dans l'analyse statistique en tant que variables indépendantes, de façon à contrôler leur influence (toutes choses égales par ailleurs) sur les variables dépendantes

[14] Il s'agit là d'un enjeu majeur pour une entreprise, un formulaire contenant de nombreuses données sensibles permettant potentiellement d'obtenir des données fructueuses (autorisant par exemple une meilleure personnalisation de la relation par la suite) tout en faisant courir le risque d'un refus de l'internaute.

[15] Van Kenhove et al. (2002) montrent en effet un lien entre l'implication et le décodage de l'information.

[16] Iacobucci (2001) conseille de bloquer des variables comme le genre car la sélection aléatoire des répondants ne garantit pas forcément une bonne répartition sur ce critère.

[17] Il a été fait en sorte que les différentes classes d'âge se répartissent équitablement au sein des traitements.

Le plan expérimental

Parmi les trois variables indépendantes du modèle, seule la sensibilité des données a été manipulée. La préoccupation étant une variable individuelle, elle ne pouvait en effet pas être manipulée. De plus, nous avons préféré mesurer la valeur perçue plutôt que de la manipuler avant de refléter au maximum la réalité (i.e. la perception de l'individu) et d'éviter de créer des conditions « factices » de supériorité des coûts ou des bénéfices liés à la fourniture de données personnelles sur un site web.

La manipulation a consisté à mettre les individus dans une situation de sollicitation de leurs données personnelles par l'intermédiaire d'un formulaire trouvé sur le site web de leur opérateur de téléphonie mobile[18]. Le formulaire proposé était soit peu sensible (champs à compléter correspondants à l'état civil) soit sensible (champs allant de l'état civil au revenu en passant par la situation familiale et professionnelle), correspondant aux deux niveaux de sensibilité (faible [S1] vs. élevée [S2]). Le logo de l'opérateur de téléphonie mobile auprès duquel les répondants déclaraient être abonnés figurait en haut du formulaire, pour mesurer leurs intentions en connaissance de cause. Au total, 6 formulaires ont donc été créés, correspondant aux 2 niveaux de sensibilité et aux 3 principaux opérateurs[19]. Après avoir indiqué le nom de son opérateur, chaque personne interrogée se voyait donc affectée soit à un formulaire « sensible » soit à un formulaire « peu sensible » et devait indiquer sa propension à remplir ce type de formulaire et à fournir (ou non) des données mensongères.

Echantillon interrogé et administration du questionnaire

Afin d'accroitre au maximum la validité externe de cette recherche, nous décidons d'interroger des individus choisis aléatoirement au sein d'un panel d'internautes. Ce choix permet en outre de s'affranchir des limites liées à la sélection d'un échantillon de convenance (notamment d'un échantillon d'étudiants) dont les caractéristiques spécifiques (en particulier en termes d'âge et de niveau d'études) empêchent de pouvoir généraliser les résultats à une population plus conséquente. De plus, cela réduit le biais de couverture puisque l'échantillon est issu du panel, lui-même représentatif des internautes français.

Au total, 252 répondants ont répondu (cf. Annexe A) soit 126 par niveau de sensibilité. L'instrument de mesure utilisé a été administré en ligne. Les répondants recevaient

[18] Ce secteur offre l'avantage d'être à la fois impliquant, important économiquement et réaliste puisque les principaux intervenants sollicitent régulièrement des données personnelles auprès de leurs clients à des occasions diverses (i.e. changement de forfaits, jeux-concours, adhésion à un club de fidélité, …).

[19] Dans la mesure où les formulaires étaient créés à l'avance, nous ne pouvions pas considérer d'autres opérateurs que les trois principaux (Orange, SFR et Bouygues). Tout participant client d'un autre opérateur était donc supprimé de l'échantillon. A l'époque où l'expérience a été menée (2006), le nombre d'opérateurs « annexes » étant encore très limité, le risque de devoir supprimer ces individus était donc minime.

un email proposant un lien sur lequel cliquer pour se rendre sur la page d'accueil du questionnaire. Après avoir indiqué le nom de leur opérateur, ils étaient ensuite affectés à un des deux *scenarii* possibles. On leur expliquait alors la situation amenant à solliciter leurs données personnelles. Suivait une page montrant le formulaire à remplir (sensible ou non). L'ensemble des items correspondant aux variables à mesurer était enfin présenté (Annexe B).

La validité du plan d'expérience

La validité du plan d'expérience a été vérifiée avec soin afin de nous assurer de la qualité des résultats et pouvoir, le cas échéant, les généraliser à d'autres populations. Cette vérification correspond aux conditions d'application de l'analyse de variance (Howell 1998). Nous nous assurons également que les facteurs ont été manipulés avec succès en vérifiant que les différences de moyenne entre les traitements sont significatives[20].

Opérationnalisation des variables

Trois concepts ont fait l'objet d'une mesure multi-items : la préoccupation pour le respect de la vie privée (PREOC) ; la valeur perçue à répondre (VAL) et l'implication (IMPL). Cette dernière sera mesurée en trois items issus de l'échelle de Pertinence-Intérêt-Attirance (Strazzieri 1994)[21]. Pour les deux autres concepts, comme aucune échelle satisfaisante n'existait telle quelle dans la littérature, nous avons procédé à un développement ad hoc des instruments de mesure, sur la base de verbatims tirés d'une étude qualitative préalable et d'énoncés issus d'échelles existantes. Tous les items ont ensuite été soumis à cinq experts, afin de s'assurer de la validité de contenu des échelles proposées. Mis à part les intentions de répondre et de mentir (évaluées chacune par un item unique mesuré en 5 points de « oui certainement » à « non certainement pas »[22]), les autres instruments sont au format Likert en 7 points (de « pas du tout d'accord » à « tout à fait d'accord »). Ces échelles ont ensuite fait l'objet d'une procédure de validation quantitative sur la base d'une analyse exploratoire (sous SPPS 15), doublée d'une analyse confirmatoire (sous AMOS 5).

Analyses exploratoires – Les analyses factorielles ont été réalisées concept par concept, de façon à vérifier l'unidimensionnalité de chaque échelle et à éliminer un premier ensemble d'items dont la variance était jugée faible. Les résultats (cf. Annexe C) conduisent à accepter provisoirement l'unidimensionnalité et la fiabilité de chaque échelle, avec des coefficients tout à fait acceptables pour des échelles en construction (alphas supérieurs à 0,7).

[20] Ainsi, les répondants affectés au niveau « sensibilité élevée » jugent bien le formulaire plus sensible que ceux affectés au traitement « sensibilité faible » (score moyen de 4,98 vs. 2,19 respectivement ; p = 0,000).
[21] Pour ne pas alourdir le questionnaire, nous ne conservons ici qu'un item (sur les deux) par dimension.
[22] Ce format est couramment employé pour mesurer des intentions comportementales, notamment la fidélité.

A noter que l'échelle de valeur perçue (10 items) laisse apparaître 3 dimensions[23] que nous nommons « bénéfices », « coûts d'énergie » et « coûts monétaires et psychologiques »[24].

Analyses confirmatoires – Des analyses confirmatoires ont ensuite été réalisées sous AMOS pour le concept de valeur perçue. Une réplication de l'analyse factorielle conduit à supprimer 4 items sur les 10 initiaux et ne laisse plus apparaître que 2 facteurs (correspondant aux dimensions « bénéfices » et « coûts » mesurées respectivement en 4 et 2 items). Nous testons alors sous AMOS un modèle qui intègre ces deux dimensions. Ce modèle présente des indices d'ajustement satisfaisants (χ^2 / dl = 1,15 ; GFI = 0,985 ; AGFI = 0,965 ; RMSEA = 0,024 ; RMR = 0,084). La fiabilité de chaque dimension de l'échelle de valeur perçue est confirmée par le calcul du Rhô qui est respectivement de 0,86 et 0,76. La validité convergente et discriminante de l'échelle est également établie. Par ailleurs, la préoccupation et la valeur formant les variables exogènes du modèle, nous vérifions le caractère discriminant des échelles correspondantes par le biais d'une analyse factorielle en composantes principales[25] dont les résultats sont présentés en Annexe D. Ceux-ci confirment le caractère discriminant des items mesurant les deux construits testés (i.e. la préoccupation et la valeur), l'AFCP laissant apparaître 3 facteurs correspondant respectivement aux bénéfices (4 items), à la préoccupation (3 items) et aux coûts (2 items) représentant au total 76% de la variance expliquée (soit respectivement 44.7%, 18.8% et 12.5%).

En résumé, la qualité des échelles étant satisfaisante au regard des critères utilisés habituellement (α de Cronbach et ρ de Jöreskog notamment, cf. Annexe E), nous avons créé une nouvelle série de variables par addition des items correspondants. Le score de valeur perçue correspond, ainsi, à la différence entre les scores de « bénéfices » et de « coûts ».

5. PRÉSENTATION ET DISCUSSION DES RÉSULTATS

Pour rappel, nous souhaitons tester les hypothèses suivantes :

(i) la préoccupation pour le respect de la vie privée influence-t-elle l'intention de répondre (IR) à une sollicitation de données (i.e. remplir le formulaire) ? Quelle est son influence sur l'intention de mentir (fournir des données mensongères, IM) ? (Hypothèses 1 et 2) ?

(ii) la valeur perçue à répondre (supériorité des gains par rapport aux coûts) a t-elle un effet sur l'intention de répondre (IR) et l'intention de mentir (IM) ? (H3 et H4)

[23] Puisque nous souhaitions permettre à ces trois facteurs d'être corrélés entre eux, nous avons mené l'analyse factorielle avec rotation oblique de type Promax qui offre l'avantage d'être rapide et conceptuellement simple.
[24] Les coûts d'énergie concernent essentiellement la charge cognitive liée au remplissage du formulaire. Les coûts psychologiques sont liés aux craintes relatives au respect de la vie privée et à la sécurité des transactions.
[25] Rappelons que la sensibilité, autre variable exogène du modèle est un facteur manipulé (à deux niveaux) et non une variable mesurée et, qu'à ce titre, elle ne peut donc pas être introduite dans l'analyse factorielle.

(iii) la sensibilité des données demandées a t-elle un effet sur l'intention de répondre à la sollicitation (i.e. remplir le formulaire) ? Qu'en est-il de son influence sur l'intention de mentir (i.e. de fournir des données mensongères) ? (Hypothèses 5 et 6)

Le traitement statistique de données issues d'une expérimentation renvoie logiquement à l'analyse de variance (dans notre cas de covariance). Ce choix est renforcé ici par le fait d'avoir des variables indépendantes à la fois métriques (préoccupation et valeur) et nominales (variables de contrôle). Les deux variables dépendantes (IR et IM) n'étant pas forcément liées entre elles (on peut remplir des formulaires mais ne jamais mentir), nous décidons de mener une analyse de variance pour chacune d'elle[26]. Bien que n'ayant pas proposé d'hypothèse sur de possibles effets d'interaction entre les trois variables indépendantes, nous testons dans un premier temps cette possibilité en conduisant une analyse de covariance (ANCOVA) avec effets d'interactions, pour chaque variable dépendante testée. Dans ces analyses, l'intention – de répondre ou de mentir – est la variable dépendante, la sensibilité le facteur manipulé, les autres variables indépendantes (préoccupation et valeur) et les variables de contrôle des covariables. L'analyse menée sur l'intention de mentir montre l'absence d'effet d'interaction entre les trois variables indépendantes prises deux à deux et nous conduit à mener une seconde analyse de variance sans effet d'interaction, dont les résultats sont présentés plus loin. Les résultats concernant l'intention de répondre sont présentés ci-dessous : ils confirment l'existence d'un effet d'interaction significatif (sur les trois possibles).

Test des hypothèses sur l'intention de répondre

Les résultats de l'analyse de covariance menée sont fournis au tableau 3.

Contrairement à ce que nous avions supposé, ni la préoccupation pour le respect de la vie privée ni la sensibilité du formulaire ne semble influencer – chacune prise individuellement - l'intention de l'individu de compléter celui-ci ($p = 0.984$ et 0.162 respectivement), conduisant ainsi à ne pas valider les hypothèses H1 et H5. Une explication à ce résultat contraire à nos hypothèses réside toutefois dans l'existence d'un effet d'interaction entre ces deux variables ($p = 0.072$). Ainsi, pour les individus les plus préoccupés (niveau supérieur à 5 sur l'échelle de 1 à 7), la sollicitation de données sensibles engendrerait une forte probabilité de refus, celle-ci s'avérant plus faible pour les individus moyennement ou peu préoccupés. Ce n'est donc pas la sensibilité ou la préoccupation qui ont un impact direct sur l'intention de remplir

[26] Par mesure de précaution et à des fins de vérification, nous réalisons une analyse multivariée de variance (MANOVA) en considérant simultanément les deux variables dépendantes. Les résultats sont sensiblement identiques à ceux présentés ultérieurement pour chaque variable dépendante prise séparément.

le formulaire mais plutôt l'effet conjoint de ces deux variables, suggérant ainsi un effet potentiellement modérateur – et non direct – du niveau de préoccupation, résultat qui vient sérieusement enrichir la littérature sur le respect de la vie privée. La variable de préoccupation est en effet souvent considérée comme un antécédent des intentions comportementales dans les recherches existantes (par exemple, Farag et Krishnan 2003, Stewart et Segars 2002) alors qu'elle aurait plutôt, d'après les résultats exposés ici, un effet modérateur. D'autres travaux devront être menés pour valider ce résultat dans un contexte plus large et autoriser ainsi une nouvelle prise en compte de cette variable majeure et incontournable dans les modèles.

Tableau 3 – ANCOVA sur l'intention de répondre (IR)

Source	Somme des carrés de type III	ddl	Moyenne des carrés	F	Sig.	Hypothèses et conclusion
Modèle corrigé	142,405 [a]	13	11,724	11,022	,000	
Constance	55,909	1	55,909	52,563	,000	
Sexe	,003	1	,003	,003	,955	
Age	,154	1	,154	,144	,704	
Csp	,561	1	,561	,528	,468	
Niveau d'études	1,720	1	1,720	1,622	,204	
Exp. d'Internet	,807	1	,807	,759	,385	
Util. (surf web)	2,389	1	2,389	2,246	,135	
Implication	,009	1	,009	,009	,925	
SENSIBILITE	2,089	1	2,089	1,964	*,162*	*H5 (ns.)*
PREOCCUPATION	,000	1	,000	,000	*,984*	*H1 (ns.)*
VALEUR	10,667	1	10,667	10,028	**,002** ***	**H3 (sig. à 1%)**
SENS * PREOC	3,465	1	3,465	3,257	**,072** *	**Sig. hors hyp.**
SENS * VAL	,488	1	,488	,459	,499	
PREOC * VAL	,023	1	,023	,022	,883	
Erreur	252,496	238	1,064			
Total	3645,000	252				
Total corrigé	394,901	251				

a. R deux = ,361 (R deux ajusté = ,326)

Conformément à ce que nous avions supposé, la valeur perçue influence bien l'intention de répondre au formulaire (p = 0.002). En outre, plus la valeur de l'échange est perçue comme élevée, plus l'intention de répondre est forte, **validant ainsi l'hypothèse 3** au seuil de 1%. Ainsi, un individu confronté à un formulaire de données en ligne qui estime avoir plus à gagner qu'à perdre en répondant (forte valeur perçue) a une forte probabilité de remplir le formulaire, et ce quelque soit son niveau de préoccupation (p = .883) et le niveau de

sensibilité du formulaire (p = .499). Notons qu'aucune variable parmi celles contrôlées n'est en mesure d'influencer le phénomène (i.e. la décision de répondre ou non à la sollicitation).
Une entreprise qui souhaiterait mieux connaître ses clients devra donc sérieusement réfléchir avant de demander des renseignements à ses prospects/clients au risque de les voir refuser. Outre la sensibilité du formulaire pour les individus les plus préoccupés, elle devra être particulièrement attentive aux coûts engendrés par le remplissage du formulaire ainsi qu'aux bénéfices susceptibles d'être proposés en échange. Qu'en est-il de l'intention de mentir ?

Test des hypothèses sur l'intention de mentir

Les résultats de l'analyse de variance sur l'intention de mentir sont fournis au tableau 4 et commentés ci-après.
Comme nous l'avions supposé, la préoccupation pour le respect de la vie privée et la valeur perçue influencent bien de manière directe l'intention de fournir des données mensongères (p = 0.069 et 0.000 respectivement). De plus, plus la préoccupation est élevée, plus l'utilisateur a l'intention de fournir des données mensongères. De même, plus la valeur de l'échange est perçue comme élevée, moins l'intention de mentir est forte, **validant ainsi les hypothèses 2 et 4**, aux seuils respectifs de 7% et 1%. Ainsi, confronté à un formulaire de données personnelles en ligne, un individu attentif au respect de sa vie privée a de grandes chances de mentir sauf s'il estime avoir plus à perdre (par exemple ne pas recevoir sa commande ou son cadeau) qu'à gagner (forte valeur perçue) en fournissant des données mensongères. En outre, d'après les seuils de significativité, il apparait que l'effet lié à la valeur perçue est supérieur à celui lié à la préoccupation, indiquant que le premier critère (élément perceptuel) supplante largement le second (disposition personnelle), ce qui a déjà été suggéré dans la littérature sans avoir reçu jusqu'ici de réelle validation empirique.
Concernant l'impact de la sensibilité, les résultats sont particulièrement intéressants. L'analyse confirme en effet que la sensibilité des données influence l'intention de fournir des données mensongères dans le sens attendu (plus les données sont perçues comme sensibles, plus l'intention de mentir est forte) mais avec un risque d'erreur de 9%, **validant ainsi l'hypothèse 6** (à condition de considérer ce taux de 9% comme un risque acceptable). Autrement dit, lorsque l'individu estime que l'entreprise n'a pas à demander ce type de renseignement (forte sensibilité perçue des données demandées), il peut préférer mentir plutôt que de rompre la transaction en cours. L'arbitrage se fait sur la base de deux critères. Le premier concerne la valeur perçue : quand l'intérêt à répondre est fort (forte valeur perçue), l'individu peut s'abstenir de mentir. Au contraire, si les risques perçus à répondre sont élevés, l'individu aura tendance à fournir des données mensongères pour réduire cette prise de risque.

Le second critère est individuel, lié au genre de la personne interrogée. En effet, lorsqu'on étudie l'impact des variables de contrôle, il apparait que le genre influence également l'intention de mentir (p = 0.000) et que son effet semble similaire - en poids – à celui de la valeur perçue. Une analyse bivariée entre le genre et l'intention de mentir laisse apparaitre que les femmes auraient moins tendance à mentir que les hommes, ce qui avait déjà été démontré dans un échantillon américain (Sheehan 1999) et qui est confirmé ici sur un échantillon français.

Au total, ces résultats montrent que l'intention de fournir des données mensongères en réponse à une sollicitation de données personnelles sur Internet dépend surtout de l'estimation par l'individu du gain net à répondre et de son appartenance à la gente féminine ou masculine, l'influence de la préoccupation pour le respect de la vie privée et de la sensibilité des données demandées – bien que plus faible et à la limite du seuil de significativité acceptable - ne devant pas être négligée.

En comparaison avec l'intention de répondre à la sollicitation, on voit donc apparaitre ici l'effet du genre (variable individuelle) dont l'influence vient s'ajouter à celle de la variable perceptuelle (valeur perçue). De plus, la préoccupation et la sensibilité exercent un effet direct (bien qu'à la limite de la significativité) alors qu'elles exerçaient un effet conjoint (sous la forme d'un effet modérateur de la préoccupation) dans le cas de l'intention de répondre.

Tableau 4 – ANOVA sur l'intention de mentir (IM)

Source	Somme des carrés de type III	ddl	Moyenne des carrés	F	Sig.	Hypothèses et conclusion
Modèle corrigé	47,512[a]	10	5,239	6,064	,000	
Constance	18,435	1	18,435	21,338	,000	
Sexe	15,044	1	15,044	17,413	**,000** ***	**Sig. à 1%**
Age	,069	1	,069	,079	,778	
Csp	1,010	1	1,010	1,169	,281	
Niveau d'études	,357	1	,357	,412	,522	
Exp. d'Internet	1,770	1	1,770	2,049	,154	
Util. (surf web)	1,663	1	1,663	1,925	,167	
Implication	,066	1	,066	,076	,783	
SENSIBILITE	2,543	1	2,543	2,943	**,088** *	**H6 (sig. à 10%)**
PREOCUPATION	2,887	1	2,887	3,341	**,069** *	**H2 (sig. à 7%)**
VALEUR	12,298	1	12,298	14,235	**,000** ***	**H4 (sig à 1%)**
Erreur	208,722	241	,864			
Total	1085,000	252				
Total corrigé	256,234	251				

a. R deux = ,185 (R deux ajusté = ,152)

6. CONCLUSION

Grâce à cette recherche nous montrons que :

(1) le niveau de préoccupation pour le respect de sa vie privée influence à la fois (au seuil de 10%) l'intention de répondre à une sollicitation de données en ligne (en interaction avec la sensibilité des données demandées) et l'intention de fournir des données mensongères, ce qui confirme le rôle de la préoccupation comme variable centrale dans les modèles de décision de dévoilement de soi ;

(2) le genre de la personne interrogée est la seule variable individuelle contrôlée dont l'effet est significatif. Elle apparait en outre comme une variable explicative majeure de l'intention de mentir prouvant qu'il s'agit là – au moins en partie - d'une décision individuelle et sexuée.

(3) la valeur perçue (avantages plus élevés que les coûts) influence également (au seuil de 1%) l'intention de compléter (ou non) le formulaire et l'intention de mentir, indiquant que la décision du consommateur est fortement influencée par la réponse à la question suivante : ai-je plus à gagner qu'à perdre en fournissant ces données ?

(4) la sensibilité des données demandées influence (au seuil de 10%) l'intention de répondre à la demande (de compléter le formulaire) - en interaction avec la préoccupation - et l'intention de fournir des données mensongères.

Ce dernier résultat bien qu'évident et sous-tendu par plusieurs travaux (notamment ceux de Lwin, Wirtz et Williams en 2007) n'avait cependant jamais été démontré empiriquement. D'ailleurs, ces résultats diffèrent de ceux de Hui et al. (2007) qui trouvent que la sensibilité des données n'a aucun impact sur le dévoilement de soi (i.e. la propension à révéler des données personnelles) dans un contexte commercial. Ceci est peut être dû aux différences de contexte dans lequel ont été menées ces recherches. Notre étude a en effet été réalisée en France (vs. Singapour pour Hui et al. 2007) auprès de 250 panélistes (vs. 109 étudiants).

De façon générale, les résultats (1), (2), (3) et 4) montrent que les facteurs individuels, situationnels et perceptuels testés influencent la décision du consommateur de fournir (ou non) des données personnelles complètes et justes. Ils confirment aussi que les individus sont prêts à perdre un peu d'intimité en échange de certains avantages[27]. La notion de rapport coûts/bénéfices ou de « privacy calculus » (Dinev et Hart 2006, Laufer et Wolfe 1977) est ainsi empiriquement démontrée, à travers le concept de valeur perçue.

[27] En anglais : "consumers can give up some privacy in exchange of some advantages".

6.1 Implications théoriques et managériales

Cette recherche offre des implications importantes pour les chercheurs et les managers. Elle propose également des implications méthodologiques.

D'un point de vue théorique, nous contribuons à la littérature existante concernant le lien entre préoccupation pour le respect de la vie privée et dévoilement de soi, plus particulièrement ici dans le cadre d'un environnement numérique. Nous confirmons que le souci d'intimité influence l'intention de divulguer des données personnelles et montrons qu'il influence aussi l'intention de fournir des données mensongères, variable dépendante rarement étudiée par le passé, malgré son intérêt stratégique tant d'un point de vue académique que pratique. Nous trouvons toutefois que l'influence de la préoccupation sur l'intention de remplir un formulaire ne se fait pas de manière directe - comme la littérature le laisse penser - mais indirectement, en modérant le lien entre sensibilité du formulaire et intention de le remplir. Plus précisément, la sollicitation de données sensibles aurait un effet négatif (et donc dissuasif) plus important pour les individus les plus préoccupés. Cet effet modérateur ne se retrouve pas s'agissant de l'intention de mentir. Dans ce cas là en effet, la préoccupation pour le respect de la vie privée influence directement la probabilité de fournir des données mensongères, au seuil de 10%.

A l'impact d'une telle variable individuelle, s'ajoute celui du genre. La littérature indique que le genre a un impact majeur sur le dévoilement de soi dans le domaine interpersonnel, et ce, à trois niveaux : 1) les femmes valorisent davantage le dévoilement que les hommes (Jourard 1971), 2) les femmes se dévoilent plus que les hommes (Jourard et Lasakow 1958, Cosby 1973), 3) les femmes fournissent des informations plus intimes (sensibles) que les hommes (Pedersen et Breglio 1968). Face à une sollicitation de données personnelles dans un contexte commercial et aux conséquences potentiellement risquées, il n'est alors pas étonnant de trouver que les hommes ont une plus forte tendance à mentir. Ce résultat vient donc à la fois confirmer la littérature sur le dévoilement de soi, tout en l'appliquant à de nouveaux contextes (commercial et hexagonal) et comportements (fourniture de données mensongères).

Nos résultats confirment enfin l'applicabilité du concept de valeur perçue comme élément majeur de prise de décision de fournir (ou non) des données complètes et vraies. Bien que sous-entendu dans de nombreuses recherches (à travers la notion de « privacy calculus »), cet élément n'avait jamais été testé empiriquement auparavant. Nous étendons donc l'applicabilité du concept de valeur perçue, jusque là majoritairement réservé à l'étude de l'achat (voire du réachat) de produits ou de services (voir par exemple : Dodds et al. 1991; Teas et Agarwal 2000), à l'étude des critères d'influence du dévoilement de soi.

En étudiant les facteurs d'influence du dévoilement de soi, cet article s'inscrit dans les travaux récents visant à analyser la construction d'identités numériques. Fournir (ou non) des

données personnelles (vraies ou mensongères) sur Internet contribue en effet à façonner son ou ses identité(s) numérique(s), voire sa réputation numérique. En fournissant des données personnelles[28] (mensongères ou non) à des partenaires commerciaux[29], chacun de nous laisse des traces, parfois neutres et insignifiantes, souvent non négligeables en termes d'impacts sur notre vie numérique mais aussi non-numérique (destruction de réputation, vol d'identité). Face aux risques de laisser des traces qui nuisent à notre réputation, nous pourrions décider de ne rien publier ou de le faire de manière entièrement anonyme. Mais ce serait alors se priver de services (numériques) très utiles. Face aux dangers du vol d'identité, nous pourrions aussi choisir de ne plus dévoiler d'éléments de son identité sur Internet ou de fournir des données mensongères (par exemple : faux nom-prénom, fausse adresse, fausse date de naissance). Mais, quel serait alors l'intérêt de se construire une vraie réputation sur une fausse identité ? Données personnelles et identité sont deux concepts étroitement liés dont l'analyse demande donc à être approfondie, de par l'impact du numérique sur les liens qu'ils entretiennent.

D'un point de vue managérial, nos résultats devraient aider les entreprises à améliorer la qualité de leur système d'information, et en particulier celle des données collectées auprès des clients et prospects. En outre, connaître les facteurs susceptibles d'entrainer la fourniture de données mensongères permettra aux managers de réduire les pertes financières liées au recours à des stratégies de conquête et de fidélisation inefficaces de par l'utilisation de données erronées. Plusieurs implications émergent. D'abord, nos résultats encouragent les managers à prêter attention à la manière avec laquelle ils demandent des données à leurs clients, et particulièrement à la sensibilité des données contenues dans le formulaire. Une demande jugée trop sensible, bien qu'elle fournisse une connaissance approfondie des clients, peut en effet s'avérer contre-productive parce qu'elle n'incite pas les personnes à répondre – en particulier les plus préoccupées d'entre elles – et peut même les conduire à fournir des données mensongères. Une telle requête peut aussi amener certains clients à se demander si ces données sont vraiment nécessaires et conformes à la transaction en cours, ce qui peut entrainer une certaine méfiance, peu propice à l'établissement d'une base de données saine. Pour éviter cette perception, les sites web devraient limiter la collecte aux données vraiment nécessaires pour une transaction (par exemple ne demander l'adresse physique que lorsqu'il y a une livraison). En second lieu, les compagnies devraient tenir compte de la valeur perçue de l'échange, qui influence de manière importante l'intention des consommateurs de remplir le formulaire et de fournir (ou non) des données justes. Les managers sont alors encouragés, soit

[28] Rappelons que nous faisons référence ici à des données liées à une personne identifiée ou identifiable.
[29] L'identité numérique va bien au-delà des interactions commerciales mais c'est le cas qui nous intéresse ici.

à 1) offrir à leurs clients des avantages substantiels en échange des données (par exemple davantage de confort lié à la reconnaissance automatique de son profil ou l'envoi d'offres toujours plus personnalisées) et à communiquer clairement ces avantages, soit à 2) réduire les coûts (cognitifs, en énergie, en temps) et les risques (perte de contrôle, utilisation frauduleuse, spams …) inhérents au remplissage du formulaire et à la fourniture de données. Les individus ayant une attirance différente pour les avantages pouvant leur être offerts en échange (Lancelot Miltgen et Gauzente 2006), plusieurs types de bénéfices (par exemple : bons d'achat, cadeaux, informations personnalisées) pourront être proposés en fonction des spécificités de la clientèle. Par ailleurs, le choix d'une ergonomie facilitant le remplissage du formulaire (cases à cocher, listes déroulantes, pré-remplissage, champs facultatifs, …) pourrait réduire les coûts et donc inciter davantage les consommateurs à répondre. Enfin, l'adhésion à une politique de confidentialité précise et clairement affichée peut constituer une autre stratégie de réduction des risques. L'existence de mentions de respect de la vie privée et l'apposition de labels « confidentialité » ont en effet un impact sur la propension des consommateurs à dévoiler des renseignements personnels (Arcand et Nantel 2005 ; Hui et al. 2007 ; Miyazaki et Krishnamurthy 2002), même si les résultats en la matière sont controversés, d'autres auteurs (e.g. Head et Hassanein 2002 ; Mauldin et Arunachalam 2002) concluant en l'absence d'influence de ces méthodes. Toutefois, cette stratégie offre l'intérêt, tout en réduisant les coûts/risques liés au remplissage du formulaire (effet sur la valeur perçue), de rassurer le consommateur quant à la confidentialité de ses données. En effet, si la préoccupation pour le respect de la vie privée est faiblement significative dans notre étude, il ne faut pas en déduire pour autant que le répondant ne cherche pas à être rassuré sur l'utilisation ultérieure, notamment externe, de ses données. La perception par le répondant de la façon dont les données seront ensuite utilisées par l'entreprise n'explique pas directement sa décision de remplir le formulaire. Mais, d'une part, son influence sur l'intention de remplir le formulaire est indirecte (lien modérateur). D'autre part, elle expliquerait d'autres comportements tels que la décision de mentir (ou non).

Enfin, dans une perspective de contrôle, nous avons testé le lien entre la probabilité de fournir des données et 1) certaines variables sociodémographiques ou 2) des éléments liés au comportement de l'internaute. L'introduction de ces variables a peu amélioré le pouvoir explicatif du modèle, hormis en ce concerne l'effet du genre sur l'intention de mentir. Plus spécifiquement, il apparaît que la probabilité de fournir des données mensongères est supérieure pour les individus de sexe masculin. Ce résultat encourage par conséquent à segmenter le formulaire selon le genre de l'individu interrogé, à condition toutefois de le

connaître auparavant[30]. Pour décourager les hommes de mentir, le formulaire les concernant pourrait par exemple contenir davantage de champs facultatifs que celui destiné aux femmes.

D'un point de vue méthodologique, les résultats de notre étude prouvent que les réactions des utilisateurs exigent des chercheurs d'adopter une approche contextuelle plutôt que générale, notamment dans la mesure où l'intimité de l'utilisateur est concernée (Malhotra et al. 2004). Ceci réclame des méthodologies spécifiques telles que les méthodes expérimentales qui examinent les réactions réelles des individus - quoique simulées – face à un cas concret de collecte de données personnelles.

6.2 Limites et voies de recherche

Les limites de cette étude constituent autant de pistes d'amélioration de ce travail. D'abord, parmi les nombreux facteurs situationnels qui pourraient influencer la décision de fournir des données personnelles, nous considérons seulement la sensibilité des données sollicitées. Le contexte dans lequel la demande est faite (par exemple participer à un jeu-concours, souscrire à une newsletter ou à un programme de fidélité), la longueur du formulaire, la connaissance de la société demandant les données, la présence (ou non) d'une charte de confidentialité pourraient également avoir une influence sur la décision de fournir (ou non) des données personnelles complètes et justes. Des recherches complémentaires devraient donc se pencher sur l'impact de ces autres facteurs, tout en testant différentes manière de les manipuler.

Une deuxième limite se rapporte au biais du répondant (lié à la manière de se comporter face à l'expérimentateur). Cette limite est d'autant plus préoccupante qu'elle est renforcée ici par les caractéristiques mêmes de l'échantillon, composé de panélistes. Nous avons en effet en partie affaire à des « professionnels » des sondages. Or, le simple fait d'être inscrit à un panel modifierait le comportement du répondant. En particulier, les motivations pour lesquelles les participants ont souhaité s'inscrire à un panel auraient une incidence sur la manière de se comporter face aux enquêtes (Deutskens, Ruyter et Wetzels 2005). Mener cette expérience auprès d'un échantillon de clients réels permettrait de limiter substantiellement ce biais.

Enfin, même si les répondants interrogés ici l'ont été par le biais d'un questionnaire en ligne, ce type de mise en situation n'est pas tout à fait identique à un contexte réel, d'autant que les participants n'avaient pas à naviguer sur un vrai site. Là encore, mettre les individus dans une situation réelle constituerait une démarche plus adéquate pour tester de manière efficace l'impact de l'exposition à une sollicitation de données sur la réponse des individus.

[30] Cette connaissance peut être directe (l'individu a indiqué être de sexe masculin) ou indirecte (par exemple par le biais du prénom).

Une attention particulière devra être exercée avant d'interpréter et d'extrapoler les résultats obtenus ici, notamment parce que la notion de préoccupation pour le respect de la vie privée et son impact sur la propension à se dévoiler sont censés varier selon les contextes politiques, légaux, et culturels étudiés (Harris, Van Hoye et Lieven 2003 ; Rustemli et Kokdemir 1993). Par exemple, des niveaux plus élevés de préoccupation tendraient à émerger dans les environnements de régulation modérés (Milberg et al. 1995). Dans certains cas cependant, aucune différence n'apparaît concernant le niveau d'attente en termes de respect de la vie privée entre des salariés américains et français (Rodhain et Agarwal 2001).

Cette étude a montré l'influence de facteurs individuels, situationnels et perceptuels sur le comportement des internautes qui mériterait d'être validée et confortée par des recherches supplémentaires. D'abord, il s'agirait de répliquer cette étude sur une population plus importante et plus diversifiée (autre nationalité par exemple) afin d'identifier l'impact potentiellement différent des facteurs testés ici (i.e. sensibilité des données et valeur perçue notamment). Cette recherche pourrait également être étendue à d'autres secteurs que la téléphonie mobile. Enfin, limiter les facteurs perceptuels à la seule valeur perçue est assurément restrictif ; d'autres éléments tels que la confidentialité perçue des données, ou la pertinence de la demande pourraient également avoir des effets importants.

Bibliographie

Andrade E., Kaltcheva V. et Weitz B. (2002), Self disclosure on the web: the impact of privacy policy, reward and company reputation, *Advances in Consumer Research*, 29, 350-353

Arcand M. and Nantel J. (2005), a website's privacy policy: a double-edged sword: risks and benefits when consumers read your privacy statements, Cahier de recherche HEC Montreéal n°05-10-2

Ashworth L. et Free C. (2006), Marketing dataveillance and digital privacy: using theories of justice to understand consumers' online privacy concerns, *Journal of Business Ethics*, 67, 107-123

Bagozzi R. (1977), Structural equation models in experimental research, *Journal of Marketing Research*, 14, 209-226

Benner H. J. (1968), Self-disclosure as a construct, Doctoral dissertation, Michigan State University.

Berendt B., Gunther O. et Spiekermann S. (2005), Privacy in e-commerce: stated preference vs. actual behaviour, *Communications of the ACM*, 48, 4, 101-106

Berendt B., Preibusch S. et Teltzrow M. (2008), A privacy-protecting business-analytics service for on-line transactions, *International Journal of Electronic Commerce*, 12, 3, 115-150

Bolton R. et Drew J. (1991), A longitudinal analysis of the impact of service changes on customer attitudes, *Journal of Marketing*, 55, 1, 1-10

Brown S.R. et Melamed L.E. (1990), *Experimental design and analysis*, Sage University, Series Quantitative Applications in the Social Sciences, 07-074, Newbury Park, CA: Sage.

Castaneda A. J et Montoro F. J (2007), The effect of Internet general privacy concern on customer behaviour, *Electronic Commerce Research*, 7, 117-141

Cespedes F. et Smith J. (1993), Database marketing: new rules for policy and practice, *Sloan Management Review*, 7-22

Chellappa R. K et Shivendu S. (2007), An economic model of privacy: a property rights approach to regulatory choices for online personalization, *Journal of Management Information Systems*, 24, 3, 193-225

Chellappa R. et Sin R. (2005), Personalization versus privacy: an empirical examination of the online consumer's dilemma, *Information Technology and Management*, 6, 2-3, 181-202

Chen Kuanchin and Rea Alan I Jr (2004), Protecting personal information online; a survey of user privacy concerns and control techniques, *Journal of Computer Information Systems*, 85-92

Cook T. et Campbell D. (1979), *Quasi-Experimentation, Design & Analysis Issues for Field Settings*, Houghton Mifflin: Boston

Cranor L., Reagle J. et Ackerman M. (1999), Beyond concern: understanding net users' attitudes about online privacy, AT&T Labs, Research Technical Report, 99.4.3

Cozby P. (1973), Self disclosure: a literature review, *Psychological Bulletin*, 79, 2, 73-91

Culnan M. (1993), How did they get my name ? an exploratory investigation of consumer attitudes toward secondary information use, *MIS Quarterly*, 17, 3, 341-363

Culnan M. (1995), Consumer awareness of name removal procedures: implications for direct marketing, *Journal of Direct Marketing*, 9, 2, 10-19

Culnan, M. J. 2000. Protecting privacy online: Is self-regulation working? *Journal of. Public Policy and Marketing* 19(1) 20–26.

Culnan, M.J., et Milberg, S.J. (1998), The second exchange: managing customer information in marketing relationships, unpublished manuscript, Georgetown University

Deutskens L., Ruyter K. et Wetzels M. (2005), Are online access panels the future of marketing research? A motivational perspective on respondent participation in online panels, *34th European Marketing Academy Conference,* Milan, Italy: Bocconi University.

Dinev, T., et Hart, P. (2006), An Extended Privacy Calculus Model for E-Commerce Transactions, *Information Systems Research*, 17 (1), 61- 80

Dinev T. et Hart P. (2004), Internet privacy concerns and their antecedents: measurement validity and a regression model, *Behaviour and Information Technology*, 23, 6, 413-422

Dodds, W. B., Monroe K. B et Grewal D. (1991), Effects of Price, Brand and Store Information on Buyers' Product Evaluation, Journal *of Marketing Research,* 28 (August): 307-319.

Doney P. et Cannon J. (1997), An examination of the nature of trust in buyer – seller relationships, *Journal of Marketing*, 61, April, 35-51

Evrard Y., Pras B. et Roux E. (2003), *Market : études et recherches en marketing*, Dunod, Paris

Faja, S. et Trimi, S. (2006), Influence of the Web Vendor's Interventions on Privacy-Related Behaviors in E-Commerce, *Communications of the AIS*, Volume 17,

Farag N. et Krishnan M. (2003), An empirical evaluation of information features and the willingness to be profiled online for personalization, http://misrc.umn.edu/workshops/2003/spring/farag_030703.pdf

Greenberg J. (1990), Organizational justice: yesterday, today and tomorrow, *Journal of Management*, 16, 2, 399-432

Hann et al. (2005), Analysing online information privacy concerns: an information processing theory approach, working paper, http://www.comp.nus.edu.sg/~lung/privacy_conjoint.pdf

Harris M.M., Van Hoye G. et Lieven F. (2003) Privacy and attitudes towards Internet-based selection systems: a cross cultural comparison, *International Journal of Selection and Assessment*, 11, 2/3, 230-236

Head, M.M. et. Hassanein K (2002), Trust in E-commerce: Evaluating the Impact of Third-Party Seals, *Quarterly Journal of Electronic Commerce*, (3)3, 307- 325.

Hoffman, D. L., Novak T. P. et Peralta M.. (1999), Building consumer trust online. *Comm. ACM,* 42(4), 80–85.

Howell D. (1998), Méthodes Statistiques en Sciences Humaines, DeBoeck Université, Paris

Hui K-L., Tan C. Y. and Goh C. Y. (2005), Online Information Disclosure: Motivators and Measurements, *ACM Transactions on Internet Technology*, November 2006, Volume 6, Number 4, pp. 415-441.

Hui Kai-Lung, Teo Hock Hai, Sang-Yong Tom Lee (2007), The value of privacy assurance; an exploratory field experiment, *MIS Quarterly*, 31(1), 19-33.

Iacobucci D. (2001), Should I treat a participant attribute as a blocking factor? *Journal of Consumer Psychology*, Special Issue: Methodological and Statistical Concerns of the Experimental Researcher, 10, 1 et 2, 18-20

Jamal K., Maier M. et Sunder S. (2005), Enforced standards versus evolution by general acceptance; a comparative study of e-commerce privacy disclosure and practice in the United States and the United Kingdom, *Journal of Accounting Research*, 43, 1, 73-96

Jarvenpaa S. et Tractinsky N. (1999), Consumer trust in an Internet store: a cross-cultural validation, *Journal of Computer Mediated Communication*, 5, 2

Jourard S. (1971), *The transparent self*, New York: Van Nostrand

Jourard S. et Lasakow P. (1958), Some factors in self-disclosure, *Journal of abnormal and Social Psychology*, 56, 91-98

Kahneman D. et Tversky A. (1979), Prospect Theory: An Analysis of Decision under Risk, *Econometrica*, 47(2), 263–291.

Krishnan R., Peters J., Padman R. et Kaplan D. (2005),On data Reliability Assessment in Accounting Information Systems, *Information Systems Research*, Vol. 16, N°3, 307-326.

Lancelot Miltgen C. (2003), Vie privée et Internet : influence des caractéristiques individuelles et situationnelles sur les attitudes et les comportements des internautes face à la collecte de données personnelles, Congrès de l'Association Française du Marketing, Tunis

Lancelot Miltgen C. (2006), Dévoilement de soi et réponses du consommateur face à une sollicitation de ses données personnelles : application aux formulaires sur Internet, Thèse de Doctorat en Gestion, Université Paris Dauphine

Lancelot Miltgen C. et Gauzente C. (2006), Vie privée et partage de données personnelles en ligne : une approche typologique, Congrès de l'Association Française du Marketing, Nantes

Lauer T. W et Xiaodong D. (2007), Building online trust through privacy practices, *International Journal of Information Security*, 6, 323-331

Laufer R. et Wolfe M. (1977), Privacy as a concept and a social issue: a multidimensional developmental theory, *Journal of Social Issues*, 33, 22-42

Lee M.K.O et Turban E. (2001), A Trust Model for Consumer Internet Shopping, *International Journal of Electronic Commerce*, 6, 1, 75-91

Lwin M. et Williams J. (2003), A model integrating the multidimensional developmental theory of privacy and theory of planned behavior to examine fabrication of information online, *Marketing Letters*, 14, 4, 257-272

Lwin M., Wirtz J. et Williams J. D (2007), Consumer online privacy concerns and responses; a power-responsibility equilibrium perspective, *Journal of the Academy of Marketing Sciences*, 35(4), 572 – 585

Malhotra N.K., Kim S.S. et Agarwal J. (2004), Internet user's Information Privacy Concern (IUPC): The construct, the scale and a causal Model, *Information Systems Research*, Vol. 15, December, 336-355.

Margulis S. (2003), Privacy as a social issue and behavioral concept, *Journal of Social Issues*, 59, 2, 243-261

Mauldin, E. et V. Arunachalam (2002) An Experimental Examination of Alternative Forms of Web Assurance for Business-to-Consumer e- commerce, *Journal of Information Systems*, 16 (Supplement), 33-54.

McKnight H., Choudhury V. et Kacmar, C. (2002), Developing and validating trust measures for e-Commerce: An integrative typology, *Information Systems Research*, 13 (3), 334-359.

Milberg S., Burke S., Smith J. et Kallman E. (1995), Values, personal information, privacy and regulatory approaches, *Communication of the ACM*, 38, 12, 65-74

Milne G. (1997), Consumer Participation in Mailing Lists: A Field Experiment, *Journal of Public Policy and Marketing*, 16(2) Fall, 298-309.

Milne G. et Boza M. (1999), Trust and concern in consumers' perceptions of marketing information management practices, *Journal of interactive Marketing*, 13, 1, 5-24

Milne G. et Gordon M. (1993), Direct mail privacy efficiency trade-offs within implied social contracts framework, *Journal of Public Policy and Marketing*, 12, 2, 206-219

Milne G. et Rohm A. (2000), Consumer privacy and name removal across direct marketing channels: exploring opt-in and opt-out, *Journal of Public Policy et Marketing*, 19, 2, 238-249

Milne G. R. (2000), Privacy and ethical issues in database/interactive marketing and public policy: a research framework and overview of the special issue, *Journal of Public Policy and Marketing*, 19, 1, 1-6

Miyazaki, A. D. et Fernandez A. (2000), Internet privacy and security: An examination of online retailer disclosures, *Journal of Public Policy* and *Marketing* 19(1) 54–61.

Miyazaki, A. D. et Fernandez A. (2001), Consumer perceptions of privacy and security risks for online shopping. *The Journal of Consumer Affairs*, 35(1) 27–44.

Miyazaki A. et Krishnamurthy S. (2002), Internet seals of approvals: effects on online privacy policies and consumers perceptions, *The Journal of Consumer Affairs*, 36, 1, 28-35

Moon Y. (2000), Intimate exchanges: using computers to elicit self disclosure from consumers, *Journal of Consumer Research*, 26, 323-339

Nowak G. et Phelps J. (1992), Understanding privacy concerns, *Journal of Direct Marketing*, 6, 4, 28-39

Padmanahan B. et Tuzhilin A. (2005), On the Use of Optimisation for Data Mining : Theoretical Interactions and eCRM opportunities, *Management Science*, Vol. 49, N° 10, 1327-1343.

Pedersen D. et Breglio V. (1968), Personality correlates of actual self-disclosure, *Psychological Reports*, 22, 495-501

Pavlou, P. A. (2002), Institutional Trust in Interorganizational Exchange Relationships: The Role of Electronic B2B Marketplaces, *Journal of Strategic Information Systems*, 11(3/4), 215-243

Phelps J., D'Souza G. et Nowak G. (2001), Antecedents and consequences of consumer privacy concerns : an empirical investigation, *Journal of Interactive Marketing*, 15, 4, 4-17

Phelps, J., Nowak G., Ferrell E. (2000), Privacy concerns and consumer willingness to provide personal information, *Journal of. Public Policy and Marketing*, 19(1), 27–41.

Regan P. (1995), Legislating privacy: technology, social values and public policy, Chapel Hill: University of North Carolina Press

Rodhain F. et Agarwal R. (2001), Le message électronique: une propriété privée? Perceptions des salaries quant à la propriété de leurs courriels et au respect de leur vie privée sur leur lieu de travail, *Systèmes d'Information et Management*, 6, 4, 49-72

Rohm A. J et Milne G. R (2004), Just what the doctor ordered: the role of information sensitivity and trust in reducing medical information privacy concern, *Journal of Business Research*, 57, 1000-1011

Rustemli A. et Kokdemir D. (1993), Privacy dimensions and preferences among Turkish students, *Journal of Social Psychology*, 133, 6, 807-815

Sharp L. et Frankel J. (1983), Respondent burden: a test of some common assumptions, *Public Opinion Quarterly*, 47, 36-53

Sheehan K. (1999), An investigation of gender differences, online privacy concerns and resultant behaviors, *Journal of Interactive Marketing*, 13, 4, 24-38

Sheehan K. (2002), Toward a typology of internet users and online privacy concerns, *The Information Society*, 18, 21-32

Sheehan K. et Hoy G. (1999), Flaming, complaining and abstaining: how online users respond to privacy concerns, *Journal of Advertising*, 28, 3, 37-51

Sheehan, K. et Hoy G. (2000), Dimensions of privacy concern among online consumers. *J. Public Policy Marketing* 19(1) 62–73.

Singer E. (1984), Public reactions to some ethical issues of social research: attitudes and behaviour, *Journal of Consumer Research*, 11, 501-509

Smith, H. J., S. J. Milberg, S. J. Burke. (1996). Information privacy: Measuring individuals' concerns about organizational practices, *MIS Quarterly*, 20(2) 167–196.

Solove D. (2002), Conceptualizing privacy, *California Law Review*, 90, 1087-1155

Stewart, K.A. et Segars A.H. (2002), An empirical examination of the concern for information privacy instrument, *Information Systems 13(1) 36–49.*

Stone E., Gueutal H., Gardener D. et McClure S. (1983), A field experiment comparing information privacy values, beliefs and attitudes across several types of organizations, *Journal of Applied Psychology*, 68, 3, 459-468

Strazzieri A. (1994), Mesurer l'implication durable vis-à-vis d'un produit indépendamment du risque perçu, *Recherche et Applications en Marketing*, 9, 1, 73-92.

Teas, R. K. et Agarwal S. (2000), The Effects of Extrinsic Product Cues on Consumers' Perceptions of Quality, Sacrifice, and Value, *Journal of the Academy of Marketing Science, 28* (2): 280-292.

Teltzrow M. et Kobsa (2004), A. Impacts of user privacy preferences on personalized systems: a comparative study. In Karat C.M., Blom J., Karat J., eds., Designing Personalized User Experiences in eCommerce, Kluwer Academic Publishers, Dordrecht, Netherlands, 315–332.

Thibaut, J.W. et Walker, L. (1975). Procedural justice: A psychological analysis. New York: Wiley.

Van Kenhove P., Wijnen K. et De Wulf K. (2002), The influence of topic involvement on mail survey response behaviour, *Psychology and Marketing*, 19, 3, 293-301

Van Slyke, C., Shim, J. T., Johnson, R., et Jiang, J.(2006), Concern for Information Privacy and online consumer purchasing, *Journal of the Association for Information Systems*, 7(6), 415-444.

Wang P. et Petrison L. (1993), Direct marketing activities and personal privacy, *Journal of Direct Marketing*, 7, 1, 7-19

Ward S., Bridges K. et Chitty B. (2005), Do incentive matter? An examination of online privacy concerns and willingness to provide personal and financial information, *Journal of Marketing Communications*, 11, 1, 21-40

Weible R. (1993), Privacy and data: an empirical study of the influence of types of data and situational context upon privacy perceptions, Doctoral Dissertation, Mississippi State University.

Westin A. (1967), *Privacy and Freedom*, New York: Atheneum

Westin A. (2001), Opinion Surveys: What Consumers Have To Say About Information Privacy, Prepared Witness Testimony, The House Committee on Energy and Commerce, W.J. Billy Tauzin, Chairman, May 8.

Westin A. (2003), Social and political dimensions of privacy, *Journal of Social Issues*, 59, 2, 431-453

Woodruff R. (1997), Customer value: the next source for competitive advantage, *Journal of The Academy of Marketing Science*, 25, 2, 139-153

Xie En, Teo Hock Hai et Wan Wen (2006), Volunteering personal information on the Internet; effects of reputation, privacy notices and rewards on online consumer behaviour, *Marketing Letters*, 17, 61-74

Zeithaml V. (1988), Consumer perceptions of price, quality, and value: a means-end model and synthesis of evidence, *Journal of Marketing*, 52, 2, July, 2-22

Annexe A - Caractéristiques de l'échantillon interrogé dans le cadre de l'expérience

Variables	Valeurs	Fréquences	Pourcentage
Variables sociodémographiques			
Sexe	Masculin	128	50,8%
	Féminin	124	49,2%
Age	15-24 ans	79	31,3%
	25-34 ans	72	28,6%
	35-44 ans	44	17,5%
	45-54 ans	33	13,1%
	55-64 ans	20	7,9%
	65 ans et plus	4	1,6%
CSP[31]	Employés / ouvriers	98	38,9%
	Elèves / étudiants	57	22,6%
	Autres inactifs	40	15,9%
Niveau d'études	< Bac	59	23,4%
	Bac à Bac + 2	122	48,4%
	Bac + 3 ou 4	46	18,3%
	Bac + 5 et plus	25	9,9%
Variables expérientielles (expérience et utilisation d'Internet)			
Expérience d'Internet	Moins de 2 ans	26	10,3%
	Entre 2 et 5 ans	86	34,1%
	Plus de 5 ans	140	55,6%
Fréquence d'utilisation de la messagerie	Plusieurs fois / jour	173	68,7%
	1 à 2 fois / jour	61	24,2%
	Plusieurs fois par semaine	18	7,1%
	Moins souvent	0	0,0%
Fréquence de surf	Plusieurs fois / jour	176	69,8%
	1 à 2 fois / jour	46	18,3%
	Plusieurs fois par semaine	27	10,7%
	Moins souvent	3	1,2%
Nombre d'achats déjà réalisés en ligne depuis qu'ils utilisent Internet	Aucun	17	6,7%
	Moins de 5	64	25,4%
	De 5 à 20	88	39,4%
	Plus de 20	83	32,9%
Variables liées à la téléphonie mobile			
Nom de l'opérateur	Orange	92	36,5%
	SFR	98	35,9%
	Bouygues télécom	62	24,6%

[31] Nous ne faisons figurer ici que les catégories les plus représentées. Par ailleurs, certaines catégories étant très peu représentées, nous avons dû les regrouper avec d'autres afin d'obtenir des effectifs suffisants. Ainsi, les « autres inactifs » correspondent aux personnes à la retraite, au foyer ou à la recherche d'un emploi.

Annexe B – Instruments de mesure

Implication (3 items, 7 points)
A propos de la téléphonie mobile (téléphones portables, offres des fournisseurs, services associés), quelle est votre proposition par rapport aux propositions suivantes :
C'est un sujet auquel j'attache une importance particulière
On peut dire que c'est un sujet qui m'intéresse
Je me sens particulièrement attiré par tout ce qui a trait à la téléphonie mobile

Intention de répondre (1 item, 5 points)
J'aurais rempli puis validé ce formulaire

Intention de mentir (1 item, 5 points)
J'aurais menti à certaines des questions posées

Préoccupation pour le respect de la vie privée (3 items, 7 points)
Je suis préoccupé par le fait que les informations que je donne aux entreprises puissent être utilisées à mon insu
L'idée que mes données puissent être utilisées sans ma permission préalable m'inquiète
Je suis inquiet à l'idée que des entreprises puissent détenir des données que je considère comme privées

Valeur perçue (6 items, 7 points)
Répondre aux questions de ce formulaire vous parait :
Risqué/pas risqué
Avoir des avantages/n'avoir aucun avantage
Bénéfique/pas bénéfique
Pouvoir avoir des effets négatifs/pouvoir avoir des effets positifs
Utile/inutile
Pouvoir vous apporter beaucoup/ne rien pouvoir vous apporter

Annexe C - Résultats des analyses factorielles exploratoires (AFE)

Concepts	Items initiaux (avant AFE)	Items finaux (après AFE)	KMO / Test de Bartlett	% de variance	Qualité de la représentation		α de Cronbach
					Extraction minimale	Extraction maximale	
Valeur perçue (VAL)	12	10	0,723/0,000	63,3%	0,58	0,98	-
Préoccupation pour le respect de la vie privée (PREOC)	4	3	0,729/0,000	77,2%	0,74	0,79	0,85
Implication / catégorie de produits-services (IMPL)	3	3	0,682/0,000	66,2%	0,61	0,70	0,74

Annexe D - Résultats de l'analyse factorielle en composantes principales (AFCP)

Matrice des composantes après rotation[a]				**Qualité de représentation**
Items	Composante			
	1	2	3	
VAL_benef1	**,835**			,738
VAL_benef2	**,822**			,727
VAL_benef3	**,807**			,692
VAL_benef4	**,800**			,689
PREOC3		**,923**		,863
PREOC1		**,889**		,831
PREOC2		**,677**	-,209	,712
VAL_cout1		-,282	**,858**	,835
VAL_cout2	-,299		**,774**	,779
% de variance expliquée	44,7 %	18,8 %	12,5 %	

Annexe E - Résultats des analyses confirmatoires

Concepts	Nb d'items	Fiabilité		Validité convergente (ρvc)	Validité Discriminante	Validité prédictive
		α	Rhô			
Valeur perçue	4 + 2	0,86 0,70	0,864 0,759	0,614 0,619	Oui	Oui
Préoccupation pour le RVP	3	0,87	0,876	0,702	Oui	Oui