

Préoccupation des clients pour le respect de la vie privée et réponse à une sollicitation portant sur des données personnelles

Pierre Volle, Caroline Lancelot Miltgen

▶ **To cite this version:**

Pierre Volle, Caroline Lancelot Miltgen. Préoccupation des clients pour le respect de la vie privée et réponse à une sollicitation portant sur des données personnelles. 21ème Congrès International de l'Association Française du Marketing, May 2005, Nancy, France. pp.1-32. halshs-00164835

HAL Id: halshs-00164835
https://shs.hal.science/halshs-00164835

Submitted on 24 Jul 2007

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

PRÉOCCUPATION DES CLIENTS POUR LE RESPECT DE LA VIE PRIVÉE ET REPONSE À UNE SOLLICITATION PORTANT SUR DES DONNÉES PERSONNELLES

Caroline LANCELOT MILTGEN

Université Paris Dauphine, DMSP
Place du Maréchal de Lattre de Tassigny
75775 Paris Cedex 16
caroline.lancelot-miltgen@dauphine.fr

Pierre VOLLE

Université Paris Dauphine, DMSP
Place du Maréchal de Lattre de Tassigny
75775 Paris Cedex 16
pierre.volle@dauphine.fr

PRÉOCCUPATION DES CLIENTS POUR LE RESPECT DE LA VIE PRIVÉE ET REPONSE À UNE SOLLICITATION PORTANT SUR DES DONNÉES PERSONNELLES

RÉSUMÉ

Cette recherche s'intéresse aux perceptions et aux comportements des clients lorsqu'un site marchand leur propose de remplir un formulaire en ligne. Il s'agit à la fois d'analyser la cohérence de la réponse avec le niveau de préoccupation du client pour le respect de sa vie privée et d'étudier l'impact de la situation de collecte, telle que perçue par le consommateur. L'étude permet de montrer qu'environ 40% des répondants auraient rempli le formulaire proposé, certains avouant cependant être tentés de donner une fausse réponse aux questions les plus sensibles. Si deux tiers des répondants ont une réponse cohérente avec leur degré de préoccupation, le tiers restant adopte un comportement *a priori* incohérent (i.e., réponse positive malgré une faible préoccupation ou réponse négative malgré une forte préoccupation). Le contexte de sollicitation s'avère prépondérant dans la décision de se dévoiler ou non, notamment l'attitude globale vis-à-vis de la requête et plus spécifiquement, la sensibilité perçue des données.

Mots clés : respect de la vie privée, informations personnelles, divulgation de soi, Internet

CUSTOMERS' PRIVACY CONCERNS AND RESPONSE TOWARDS A REQUEST FOR PERSONAL DATA

ABSTRACT

Whereas customers are often sollicited for personal data online, we investigate the relationship between general privacy concern, specific perceptions of the request and response behavior (i.e., fill up / do not fill up the form). Our study finds that more than 40% of the respondents agree to fill up the form and supports the idea that customers concerned about privacy issues tend to be less responsive to online data requests. Our study also supports the attitude toward the request and more specific aspects of this attitude (e.g., sensitivity of the data) as predictors of the behavioral response.

Key words : privacy, personal data, self-disclosure, Internet

INTRODUCTION

Pour créer des relations fortes avec leurs clients, un grand nombre d'entreprises cherchent aujourd'hui à améliorer leur stratégie de ciblage, tout en présentant leur offre de manière personnalisée, qu'il s'agisse des produits /services ou du message qui les accompagne. Cette démarche implique de collecter de nombreuses données personnelles sur les consommateurs, celles-ci offrant la possibilité, une fois traitées et analysées, d'accroître l'efficacité des politiques de marketing relationnel, tout en améliorant le service rendu aux clients.

Si les progrès relatifs aux nouvelles technologies ont accru la capacité des entreprises à exploiter ces informations à moindre coût, ces pratiques accroissent dans le même temps les préoccupations des individus quant à une possible violation de leur vie privée, notamment en cas de perte de contrôle quant à la manière dont les informations collectées seront utilisées par la suite (Culnan, 1993). Les sondages d'opinion, réalisés principalement aux Etats-Unis, révèlent que les citoyens sont de plus en plus sensibles aux menaces d'intrusion dans leur intimité, du fait de l'utilisation de ces technologies. Cette préoccupation croissante des individus pour la protection de leur vie privée constitue dès lors l'un des enjeux majeurs des politiques publiques, avec des conséquences importantes pour les entreprises du secteur privé comme du public (e.g., Loi sur l'Economie Numérique promulguée en France en 2004).

La question du respect de la vie privée a été classée comme la préoccupation la plus importante des internautes (Kehoe et al., 1998). Dès lors, une majorité d'entre eux n'utiliseraient pas ce média pour commercer, du simple fait qu'ils craignent que leurs données personnelles soient divulguées (NFO Interactive, 1999). Ce qui change avec Internet c'est l'échelle avec laquelle les invasions sont désormais possibles. Comme le souligne Richards (1997) : « Jamais auparavant un média ne fut l'objet d'un si grand nombre de critiques concernant les menaces qu'il fait peser sur la vie privée des citoyens ». Une étude Business Week et Louis Harris de 1998, citée par N'Guyen, Gray et Kira (2000), indique que 78% des répondants augmenteraient leur utilisation d'Internet s'ils estimaient que leur vie privée était mieux protégée ; la majorité d'entre eux seraient également encouragés à s'enregistrer sur des sites, à fournir des informations personnelles ou à acheter en ligne. Stone et al. (1983) expliquent ce phénomène par le fait que les gens préoccupés estiment généralement manquer de contrôle sur leurs données et sont donc moins enclins à se dévoiler.

A l'opposé, d'autres travaux montrent que les consommateurs sont de plus en plus confiants quant au respect de leur intimité sur Internet. Il est vrai que, malgré une inquiétude annoncée, seuls 6% des internautes interrogés estiment avoir vu leur vie privée envahie en ligne (FTC, 1998). De plus, un grand nombre d'individus réalisent petit à petit qu'ils doivent abandonner un peu de leur intimité pour participer au e-commerce (Gandy, 1993). En effet, le confort et la valeur que fournit l'achat en ligne semblent aujourd'hui en mesure de dépasser, pour certains d'entre eux au moins, leur niveau de préoccupation dans ce domaine.

De plus en plus de recherches récentes menées sur ce thème soulignent que certains individus, une fois en ligne, oublient leurs préoccupations et ont tendance à se dévoiler, même quand les questions s'avèrent personnelles ou en l'absence de raison objective pour le faire. Cet écart entre les croyances et la réponse effective est un phénomène connu en comportement du consommateur, lié aux aspects sociaux et cognitifs qui interviennent lors d'une prise de décision.

La question se pose alors de savoir si, dans le cas d'une sollicitation de données personnelles sur Internet et en présence d'un risque perçu relatif à la sécurité et à la protection de ces informations, le consommateur peut effectivement agir d'une manière différente de ce que son niveau de préoccupation sur le sujet laissait prévoir. Par ailleurs, si les écarts entre la préoccupation vis-à-vis du respect de la vie privée et le comportement réel ont déjà été observés dans la littérature, les raisons de ces divergences sont encore largement méconnues, ce à quoi nous nous proposons de remédier à travers cet article.

Un début d'explication a été fourni par Smith (1995) lorsqu'il indique que les gens peuvent changer de niveau de préoccupation selon la situation. Dommeyer et Gross (2003) soulignent aussi que peu de consommateurs voient toute sollicitation de données personnelles comme une intrusion dans leur intimité et que la majorité d'entre eux se montrent prêts à les divulguer, sous certaines conditions. On peut alors se demander en quoi la situation de collecte modifie la prédisposition des individus à se dévoiler, en les amenant à adopter un comportement opposé à leurs croyances dans ce domaine.

Partant de là, nos deux questions de recherche sont les suivantes :

Y a-t-il ou non cohérence entre le degré de préoccupation de l'internaute vis-à-vis du respect de la vie privée et le comportement effectif, lorsqu'une entreprise le sollicite pour communiquer des données personnelles ?

Quel est l'effet du contexte de collecte, tel que perçu par l'internaute, qui pourrait inhiber ou au contraire l'encourager à communiquer des données personnelles à l'entreprise ?

Par rapport aux recherches antérieures, notre objectif consiste, d'une part, à valider dans un contexte français les études anglo-saxonnes publiées sur le sujet (Turner et Dasgupta, 2003 ; Berendt et al., 2003 ; Acquisti, 2004 ; Acquisti et Grossklags, 2004) et d'autre part, à prendre en compte des éléments contextuels riches et réalistes dans le contexte des sollicitations sur Internet.

Pour répondre à ces deux questions, nous avons mené une enquête exploratoire visant à mesurer, dans quatre situations de collecte différentes, l'écart entre le niveau de préoccupation générale des individus et leur comportement réel, d'une part, ainsi que l'effet des perceptions relatives au contexte de sollicitation, d'autre part. Après avoir passé en revue la littérature sur le thème du respect de la vie privée et de la divulgation de soi, nous présenterons la méthodologie utilisée pour mettre en place et analyser les résultats de cette étude avant d'évoquer les implications, limites et voies de recherche issues de nos conclusions.

PRINCIPAUX CONCEPTS MOBILISÉS DANS LA RECHERCHE

Cette recherche mobilise plusieurs concepts qui sont présentés ci-dessous : droit au respect de la vie privée et degré de préoccupation individuelle pour ce respect ; attributs perceptuels de la sollicitation en termes de sensibilité et de respect de la confidentialité des informations collectées par l'entreprise, ou encore d'équité et de pertinence perçue.

Le droit au respect de la vie privée

Le droit au respect de sa vie privée est un droit fondamental reconnu comme tel par la Déclaration Universelle des Droits de l'Homme, la Convention Européenne pour la protection des individus en regard de la collecte de données personnelles et par d'autres traités nationaux, régionaux et internationaux. Même dans les pays où ce droit n'est pas explicitement inscrit dans la constitution, comme aux Etats-Unis, la jurisprudence ou des lois fédérales en reconnaissent l'existence. En France, il représente un des principes légaux les plus importants, que l'on trouve à l'article 9 du Code Civil : « tout le monde a droit au respect de sa vie privée ».

Le concept initial de respect de la vie privée (*privacy*), tel qu'il fut établi par Warren et Brandeis (1890), reflète le droit à être laissé « seul » ou « tranquille ». Plusieurs définitions ont été proposées par la suite dans la littérature, que l'on peut regrouper en trois catégories principales selon l'aspect mis en avant (informations, interactions ou liberté). La première catégorie met en exergue le contrôle de l'information sur soi. Dans cette perspective, des définitions courantes sont données par Westin (1967) : « Le droit pour les individus de déterminer la nature et l'étendue avec laquelle les informations les concernant peuvent être communiquées aux autres » ou encore, Stone et al. (1983) : « La capacité de l'individu à contrôler personnellement les informations le concernant ». Le deuxième courant insiste plutôt sur la régulation des interactions avec les autres. Dans cette perspective, l'individu cherche à contrôler le nombre de contacts avec les autres, par le biais de différents mécanismes incluant notamment les notions de territorialité et d'espace personnel. Cette définition fait référence à ce que Schwartz (1968) appelle un « mode institutionnalisé de retrait ». Le troisième courant, basé sur l'idée de liberté face au contrôle des autres, se fonde sur les travaux de Goffman (1959), qui montre que lorsque les autres ont des informations sur vous, ils deviennent capables de vous contrôler. Le respect de la vie privée vise alors à assurer l'indépendance et l'autonomie des individus dans les situations où ils pourraient se montrer vulnérables (Kelvin, 1973).

Les définitions du respect de la vie privée sont multiples car les travaux réalisés dans ce domaine sont issus de nombreuses disciplines incluant l'anthropologie, l'histoire, la sociologie, les ressources humaines et plus récemment, le marketing. Cependant, le point commun de toutes ces littératures est le concept de contrôle. Dès lors, une grande partie des théories issues de ce champ de recherche se focalisent désormais sur ce droit à contrôler l'accès à ses données personnelles (i.e., toute information pouvant être associée à un individu identifiable et donc susceptible de violer sa vie privée, en cas de divulgation à autrui sans son autorisation préalable). Reprenant ainsi la conception de McCloskey (1980), nous définirons le droit au respect de la vie privée comme : « Le droit de protéger ses données personnelles et plus précisément, la possibilité de contrôler la collecte et l'utilisation future de celles-ci ». Le débat se focalise dès lors autour de l'opposition entre le droit du client de décider ce qui sera fait des données qu'il décide de divulguer et celui de l'entreprise de les utiliser dans un but commercial.

Plusieurs auteurs, dont Woodman et al. (1982) soulignent que le droit au respect de la vie privée est un phénomène psychologique ayant des composantes à la fois cognitives (croyances relatives, notamment, au type de données considérées comme privées) et affectives (sentiments et émotions liées aux pratiques de sollicitation et de divulgation de soi). La reconnaissance de la dualité de ce concept suppose donc que toute mesure inclut à la fois les croyances des individus relatives à ce type de situation et leurs sentiments, en cas de sollicitation effective.

Préoccupation pour le respect de la vie privée et comportement de divulgation

Plusieurs sondages relatifs au respect de la vie privée, menés notamment aux Etats-Unis [1], montrent que la majorité des individus se disent préoccupés par la question et cherchent à se protéger. En revanche, d'autres recherches semblent prouver que peu de gens se protègent réellement, même quand le coût de la protection s'avère limité voire, totalement gratuit. De plus, Singer et Presser (1996) montrent que l'attitude des citoyens américains face au problème de confidentialité de leurs données personnelles n'est pas un facteur qui explique leur participation aux sondages, soulignant ainsi l'absence de lien entre un niveau général de préoccupation et leur degré de coopération.

Ainsi, des individus *a priori* très préoccupés par le respect de leur intimité et peu enclins à se dévoiler, acceptent de le faire alors qu'ils auraient pu s'y soustraire facilement par le biais de techniques telles que la rétention d'informations, l'utilisation de pseudonymes ou le mensonge (Auteur, 2003).

De tels écarts entre une attitude générale et un comportement, pourtant conceptuellement liés, ont été observés dans de nombreux aspects de la vie humaine et ont été étudiés en psychologie sociale (LaPiere, 1934). Certains de ces écarts pourraient ainsi être attribués aux procédures même de la recherche, telles que le recours à un environnement contrôlé dans les expérimentations, à opposer à un fort biais de désirabilité sociale lors d'un sondage. Une autre explication du phénomène est apportée par Acquisti et Grossklags (2004). Ces auteurs indiquent que, comme tout autre bien, la valeur accordée au respect de sa vie privée, est affectée par des asymétries d'informations, des évaluations ex-post, des interprétations subjectives et des distorsions psychologiques.

[1] Jupiter (1999) et Price Waterhouse (2000), cités dans Acquisti et Grossklags (2004).

Le processus de décision, dans le cadre de la divulgation de données personnelles, pourrait ainsi être l'objet d'une information incomplète et de rationalité limitée. Dans le domaine du respect de la vie privée en effet, les individus ne disposent pas toujours de l'ensemble des informations nécessaires pour faire leur choix, notamment concernant la politique de confidentialité et les pratiques réelles de l'entreprise qui les sollicite. Certains individus ne connaissent pas non plus les risques associés à la divulgation de leurs données, ni l'existence de certaines mesures de protection. Cette information incomplète peut alors affecter l'estimation des risques et bénéfices liés à la divulgation et donc conduire à une décision qui semble incohérente.

La rationalité limitée fait quant à elle référence à l'incapacité de calculer et de comparer toutes les conséquences associées aux différents choix possibles pour l'individu, ainsi qu'à l'incapacité d'analyser toutes les informations disponibles (Simon, 1976). Lors d'une sollicitation de données personnelles, il se trouve souvent dans un contexte d'asymétrie d'information qui l'oblige à baser ses choix sur des estimations difficiles à réaliser, d'autant que certaines des conséquences ne sont connues qu'*ex post*. Les coûts cognitifs associés à l'analyse de la situation sont alors généralement très élevés et obligent la personne à se baser sur de simples heuristiques plutôt que sur une analyse détaillée de l'ensemble des paramètres.

La littérature en psychologie sociale montre également l'existence de distorsions psychologiques susceptibles d'affecter la décision de se dévoiler ou non (Acquisti, 2004). Ainsi, certaines personnes, préférant pourtant se protéger, opteraient pour une gratification immédiate, prenant par exemple la forme d'une récompense offerte par l'entreprise en échange de leurs données personnelles. D'autres individus souffriraient de biais d'optimisme les amenant à considérer les risques qu'ils encourent comme inférieurs à ceux auxquels les autres sont confrontés dans une situation identique. Enfin, la majorité des gens auraient des difficultés à prendre en compte des risques cumulatifs, qui sont pourtant courants dans le domaine de la vie privée. En effet, une fois les informations personnelles divulguées à une entreprise, celles-ci deviennent utilisables par cette dernière et éventuellement par d'autres pendant de longues périodes d'où un risque global associé à la divulgation supérieur à la somme des risques associés à chaque donnée divulguée.

L'ensemble de ces considérations peut donc laisser penser que, contrairement à ce qu'un niveau de préoccupation générale dans le domaine de la vie privée laisse présager, la décision d'un individu de divulguer ses données personnelles, peut être considérée comme une manière rationnelle de réagir aux incertitudes, ce que Lemley (2000) appelle une « ignorance rationnelle ».

D'où notre première proposition (P1a) : **certaines personnes fortement préoccupées par la protection de leur vie privée peuvent pourtant accepter de se dévoiler**. Et son corollaire (P1b) : **certaines personnes peu préoccupées par la protection de leur vie privée peuvent pourtant refuser de se dévoiler**.

Toutefois, si les gens donnent généralement plus de données personnelles que leur attitude générale pour le respect de la vie privée ne le laisse présumer, c'est aussi parce leur comportement se décide sur la base de la situation vécue : type d'informations demandées ; conditions dans lesquelles les informations sont collectées ; possibilité pour que celles-ci soient partagées avec des tiers ; garantie de confidentialité ultérieure ; bénéfices obtenus en échange… autrement dit, un ensemble de facteurs dont l'impact n'est effectif qu'en situation réelle.

L'importance de la situation perçue

Le désir d'intimité est fortement dynamique : les gens s'engagent continuellement dans un processus d'ajustement au cours duquel le désir de vie privée est comparé au désir de divulgation et de communication avec les autres (Kimmel, 1996). Cet ajustement a lieu dans un contexte de forces situationnelles variées telles que la pression des autres ou les normes sociales.

L'étude ethnographique de Gerber (2001) montre que la décision de divulguer des informations personnelles est situationnelle, c'est-à-dire largement dépendante du contexte dans lequel la sollicitation a lieu. Comme le rappellent Stone et Stone (1990), les consommateurs considèrent les demandes des entreprises comme moins intrusives à condition : que les informations soient collectées dans le cadre d'une relation existante ; qu'ils perçoivent pouvoir contrôler les utilisations futures des données (respect de leur confidentialité) ; que les informations collectées soient jugées pertinentes pour la transaction en cours, qu'ils aient le sentiment que les données seront utilisées pour tirer des conclusions fiables les concernant.

Dans une situation de collecte de données personnelles, il peut donc objectivement y avoir invasion de la vie privée, mais ne pas être perçu comme tel par celui censé être offensé. Ainsi, l'analyse doit se référer à un état subjectif, à une expérience vécue par l'individu, pour pouvoir en tirer des conclusions valides. On ne doit donc pas étudier la préoccupation pour le respect de la vie privée en général, mais plutôt le degré d'intimité perçue dans un contexte donné. Hine et Eve (1998) confirment cette idée, en montrant notamment comment la combinaison de différents facteurs (la représentation de soi, de l'organisation et de la technologie impliquée dans la collecte) peut conduire à la notion d'intimité situationnelle. Celle-ci, plutôt que d'être définie de manière absolue comme une sphère inviolable, se focalise surtout sur le contrôle de l'individu eu égard à la confidentialité des données divulguées. Ces auteurs rappellent aussi que, même si certaines informations sont plus sensibles que d'autres (Culnan, 1993), aucune donnée n'est considérée comme personnelle dans toutes les situations et que leur sollicitation peut donc occasionner des réponses différentes selon le contexte dans lequel elle s'opère.

D'où notre seconde proposition (P2) : **la perception du contexte de sollicitation par l'individu influence plus la décision de se dévoiler (ou non) que la préoccupation générale pour le respect de la vie privée**.

Les recherches issues de la littérature, ainsi qu'une étude qualitative menée par l'un des auteurs (Auteur, 2002) montrent que plusieurs facteurs contextuels doivent être pris en compte, notamment : la confidentialité et la sensibilité des données, ainsi que la pertinence et l'équité de la sollicitation (cf. Figure 1, page suivante).

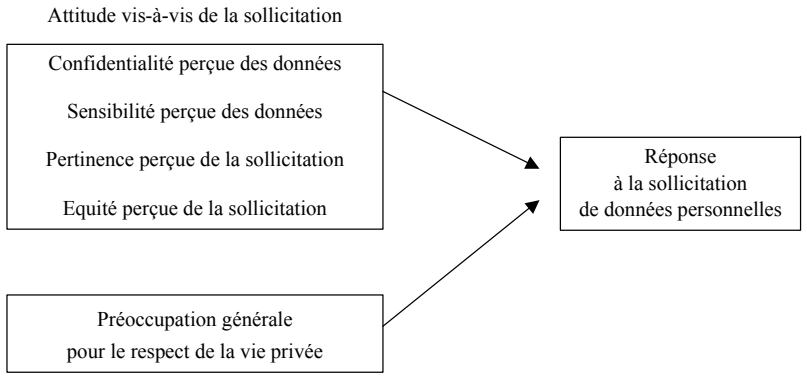
La confidentialité perçue des données divulguées – Ce facteur, lié à la confiance de l'individu quant à la capacité de la firme de maintenir ses données confidentielles, s'avère essentiel. En effet, une fois les données divulguées, celui-ci n'a plus de pouvoir quant à leur utilisation ultérieure (Pavlou et Chellappa, 2001). S'il n'a pas la garantie qu'elles resteront confidentielles, il pourra alors préférer limiter sa divulgation et donc ne pas coopérer.

La sensibilité perçue des données sollicitées – Comme nous l'avons déjà souligné, chaque information a son propre degré de sensibilité, défini par Weible (1993) comme le niveau de préoccupation qu'un individu ressent pour un type de donnée dans une situation spécifique. Cependant, à l'instar d'Acquisti (2004), pour garder le même niveau d'analyse tout au long de notre investigation, nous étudierons le niveau de sensibilité global lié à l'ensemble des données sollicitées.

La pertinence perçue de la sollicitation – Face à une situation de collecte de données, les consommateurs peuvent se poser la question des raisons qui amènent l'entreprise à les solliciter et à leur poser toutes ces questions. C'est surtout le cas lorsque la demande ne semble pas légitime par rapport à la transaction en cours. Ainsi, les sollicitations considérées comme sans motif ou non nécessaires à la conduite de l'action, seraient jugées intrusives et s'accompagneraient généralement d'un refus, voire d'hostilité (Hine et Eve, 1998).

L'équité perçue de la sollicitation (rapport coûts /bénéfices perçus) – D'un côté, quand on s'apprête à donner des renseignements personnels, il existe une perception de vulnérabilité dans la mesure où l'utilisation ultérieure de ceux-ci pourrait affecter de manière négative son avenir. De l'autre, il y a aussi parfois un intérêt personnel à répondre, quand on peut gagner quelque chose en échange. Cette notion de rapport coûts /bénéfices a été particulièrement étudiée dans la littérature sur la divulgation de soi, dans le domaine interpersonnel (Derlega et al., 1973). Or, comme le rappellent Culnan et Armstrong (1999), ces résultats peuvent être facilement rapprochés de ce qui se passe entre une entreprise et un consommateur. Ce dernier accepterait en effet plus facilement de donner des informations personnelles en échange d'offres personnalisées ou de réductions de prix. Il choisirait alors de se dévoiler en cas de résultat net attendu positif, vu comme le solde entre les conséquences positives (bénéfices) et les conséquences négatives (risques) de la divulgation. Cette évaluation de la valeur et de la pertinence des inputs et des outputs se retrouve dans le concept d'équité perçue, que nous retiendrons dans notre recherche, comme mesure du rapport coût / bénéfices perçus.

Figure 1 – Principaux concepts mobilisés dans la recherche



Si ces facteurs contextuels ne constituent pas la totalité des variables susceptibles d'influencer la réponse à la sollicitation – certains auteurs soulignent aussi l'importance du contrôle perçu et de la sécurité perçue – ils présentent cependant un intérêt théorique et managérial certain. En effet, même si leur influence sur le choix de l'individu a été suggérée voire démontrée (Milne et Gordon, 1993 ; Walczuch et al., 2001 ; Dinev et Hart, 2002), celle-ci mérite d'être encore clarifiée. De plus, il s'agit de variables concrètes, sur lesquelles une entreprise peut assez facilement agir. Connaître leur influence présente donc un intérêt managérial évident.

Avant d'analyser les résultats issus de l'étude que nous avons menée auprès d'internautes français, nous allons présenter la méthodologie utilisée pour développer et administrer notre questionnaire.

METHODOLOGIE

Echantillon interrogé et administration du questionnaire

Etant donné la nature exploratoire de l'étude, nous avons interrogé un échantillon de convenance (plus de 300 personnes). Il s'agissait d'internautes, plus ou moins familiers avec Internet, qui furent sollicités par courrier électronique. Le courriel de sollicitation envoyé à ces personnes contenait, outre quelques précisions relatives à la confidentialité des réponses fournies, un lien sur lequel il s'agissait de cliquer pour se retrouver sur la page d'accueil d'un mini site créé pour l'occasion.

Sur ce site, une fois parcourue la première page contenant les instructions et recommandations, la personne était mise dans une situation de sollicitation de données personnelles de la part d'un site de voyages en ligne. Le répondant se retrouvait alors face à un formulaire en ligne comportant des données personnelles qu'on lui demandait de fournir, même s'il ne l'aurait pas fait en situation réelle. Cette instruction offrait en effet l'avantage d'obliger l'individu à réfléchir et à lire attentivement le formulaire. De façon réaliste, celui-ci contenait des questions obligatoires et d'autres facultatives et ne pouvait être validé qu'une fois tous les champs obligatoires renseignés (cf. ex de formulaire, Annexe 1).

Une fois ce formulaire rempli, les participants étaient ensuite dirigés vers un questionnaire en ligne contenant des questions relatives à trois thèmes distincts : 1) attitude et comportement face à la divulgation de leurs données personnelles dans le cadre de la situation présentée ; 2) perceptions relatives à cette sollicitation ; 3) niveau de préoccupation générale en terme de respect de la vie privée. Le répondant devait ensuite préciser son profil (caractéristiques socio-démographiques, ancienneté et fréquence d'utilisation / d'achat sur Internet).

Au final, notre échantillon comprend 308 individus, dont 59,4% de femmes et 42,9% de répondants dans la tranche 25-34 ans. Nous avons majoritairement sondé des enseignants-chercheurs (23,1%), des cadres (22,7%) et des étudiants (16,2%), les internautes français étant plutôt issus de CSP +. Nos répondants sont plutôt plus familiers d'Internet et plus assidus que la moyenne des internautes français (Source Médiamétrie) : 52,3% ont plus de 5 ans d'ancienneté ; 87,6% consultent leur messagerie électronique au moins une fois par jour et 44,8% surfent sur le Web au moins une heure par jour.

Echantillonnage théorique des scénarios

Conformément à la démarche de l'échantillonnage théorique (Yin, 1990), et de façon à faire varier nos quatre variables contextuelles présentées plus haut (confidentialité et sensibilité des données ; pertinence et équité de la sollicitation), différents scénarii étaient aléatoirement proposés, on l'on faisait varier : la politique affichée de respect de la vie privée adoptée par le site ; le degré de familiarité du consommateur avec l'entreprise ; les compensations offertes en échange et le nombre de données demandées. Chaque répondant était amené à ne répondre qu'à un seul des quatre scénarii fournis.

Concernant la confidentialité perçue des données, Foxman et Kilcoyne (1993) suggèrent que ce concept implique à la fois le contrôle de la collecte et de l'utilisation ultérieure des données ainsi collectées. Ces deux dimensions se retrouvent d'ailleurs dans les principes de base des politiques de régulation de ce phénomène, aussi bien aux Etats-Unis qu'en Europe, à travers les notions d'information (*notice*) et de consentement (*choice*) relatif à l'utilisation ultérieure des données collectées.

Dans notre étude, ces deux facteurs ont donc été manipulés conjointement, dans le but d'obtenir un niveau de politique de respect de la vie privée offert par le site. Trois niveaux ont ainsi été distingués : faible (absence d'information et de choix), moyen (informations limitées, choix par « opt-out ») et élevé (informations précises, choix par « opt-in ») [2]. Plutôt que d'opérationnaliser cette politique à travers une charte, dont plusieurs études montrent qu'elles ne sont pas toujours lues (Milne et Culnan, 2004), nous avons choisi de la faire apparaître en bas du formulaire, afin de s'assurer que les répondants la verraient et y seraient donc sensibles.

Concernant la sensibilité perçue des données, plusieurs auteurs suggèrent que les consommateurs sont plus disposés à donner des informations à une entreprise avec laquelle ils ont déjà établi une relation (e.g., N'Guyen et al., 2000). Pour évaluer l'impact réel de l'état de la relation, trois niveaux de familiarité ont ainsi été distingués : faible (site sur lequel ils ne sont jamais allé auparavant), moyen (site consulté à plusieurs reprises par le passé) et élevé (site régulièrement visité, sur lequel ils se rendent toujours pour rechercher un voyage et auprès duquel ils ont déjà réservé un ou plusieurs séjours).

Concernant les bénéfices perçus de la sollicitation, les avantages peuvent être de nature diverse : argent, produits et services personnalisés, accès à des informations spécifiques, etc. Le peu de recherches déjà réalisées dans ce domaine montrent que les conclusions ne sont pas forcément conformes aux résultats attendus. Ainsi, Andrade, Kaltcheva et Weitz (2002) trouvent que le fait de proposer de l'argent en échange de données personnelles, incite les consommateurs à se méfier et les rend réticents à se dévoiler. Nous testerons pour notre part quatre types de circonstances et/ou avantages différents, représentant un aperçu des situations les plus fréquemment rencontrées sur les sites de voyages. Il s'agit d'une newsletter personnalisée, d'un jeu-concours pour gagner un voyage, et d'un catalogue. Nous y avons rajouté le simple accès au site, afin de tester l'impact d'un cas ou le bénéfice est quasiment nul puisqu'il est normalement proposé d'office.

[2] L'opt-out part du principe que sauf accord contraire, le consommateur consent à ce que l'entreprise utilise ses données tandis que dans l'opt-in, il lui appartient d'indiquer sa position.

Concernant à la fois la pertinence et l'équité de la sollicitation, Andrade, Kaltcheva et Weitz (2002) soulignent qu'au-delà du type de données demandées, leur nombre affecte aussi la décision de l'individu de se dévoiler ou pas. Dans notre étude, nous avons choisi de tester trois quantités différentes d'informations demandées, comprenant des données obligatoires et deux données facultatives. Dans leur recherche sur la politique de collecte de données des sites français, Gauzente, Benetteau et Dubreuil (2002) trouvent que ces derniers collectent en moyenne 4,5 informations en plus du nom et de l'adresse email, ce qui constituera la fourchette basse de notre échantillon. Nous testerons ainsi 5, 10 et 15 données obligatoires, en plus des deux données facultatives.

S'agissant d'une étude exploratoire, nous ne cherchions pas à étudier l'ensemble des combinaisons possibles de ces facteurs et donc, l'effet propre à chacun d'eux, mais plutôt à mesurer l'impact intrinsèque des variations contextuelles. Il convenait donc de sélectionner quelques cas particuliers, suffisamment différents pour constater des écarts au niveaux des réactions des individus et ainsi montrer l'impact de la situation perçue. Suivant en cela le principe de réplication théorique (Yin, 1990) et afin d'accroître la validité interne de la recherche, nous avons donc élaboré quatre scénarii différents, représentant un échantillon de cas particuliers (facteurs étudiés et niveaux). Comme le montre le tableau suivant, nous avons ainsi retenu un cas de base, représentatif d'une situation courante de collecte de données personnelles par un site de voyages en ligne, un cas intermédiaire et deux cas extrêmes, opposés sur quasiment chacun des facteurs.

Tableau 1 – Echantillonnage théorique des scénarios

Facteurs	Niveaux	Scénario 1	Scénario 2	Scénario 3	Scénario 4
Politique de respect de la vie privée	Faible	X			
	Moyen		X	X	
	Elevé				X
Familiarité avec l'entreprise	Faible	X			X
	Moyen		X		
	Elevé			X	
Circonstances et avantages liés à l'échange	Accès au site	X			
	Newsletter		X		
	Jeu-concours			X	
	Catalogue				X
Quantité de données demandées	5 obligatoires	X			
	10 obligatoires		X		X
	15 obligatoires			X	

Opérationnalisation des variables

Outre les quatre variables perceptuelles servant à étudier l'impact de la situation proposée (confidentialité et sensibilité des données ; pertinence et équité de la sollicitation), cette recherche nous a amené à mesurer trois autres concepts : le niveau de préoccupation générale de l'individu relatif au respect de sa vie privée ; l'attitude face à la divulgation et le comportement de réponse lui-même.

Les échelles de mesure utilisées proviennent à la fois de la revue de la littérature sur le sujet et de l'étude qualitative menée par l'un des auteurs (Auteur, 2002). Plusieurs items ont ainsi été élaborés pour chacun des concepts retenus, ceux trouvés dans la littérature anglo-saxonne étant adaptés en français. Les items ont ensuite été soumis à cinq experts, pour s'assurer de la validité de contenu des échelles proposées. Cette procédure a permis d'éliminer les items peu représentatifs des concepts étudiés. Au final, une moyenne de 4 à 5 items par concept a donc été administrée aux répondants.

Bien que couramment employé dans la presse et dans les textes réglementaires, le concept de sensibilité n'avait jamais été mesuré en tant que tel dans les recherches passées, amenant les auteurs à élaborer l'ensemble des cinq items retenus après validation par les experts. La pertinence perçue est mesurée à travers quatre items, dont un issu de l'étude de Fusilier et Hoyer (1980).

La mesure de l'équité perçue, concept servant à évaluer le rapport coûts / bénéfices de la divulgation, comprend pour sa part cinq items, dont deux proviennent de Oliver et Swan (1989). Quant à la confidentialité perçue, trois items parmi les quatre retenus ont été élaborés par les auteurs, le dernier provenant de Miyazaki et Krishnamurthy (2002). L'attitude de l'individu face à la divulgation de ses données personnelles, dans le cas proposé, est mesurée à travers quatre items, tous élaborés par nos soins. La mesure du comportement effectif comprend trois items, correspondant aux trois réponses possibles face à un formulaire en ligne : je retourne sur la page d'accueil ou sur une autre page du site sans avoir rempli le formulaire, je quitte le site sans remplir le formulaire, je remplis le formulaire. Enfin, étant donné la diffusion du concept de préoccupation pour le respect de la vie privée, son échelle de mesure en quatre items, a été prise dans la littérature (Dinev et Hart, 2002), non sans avoir au préalable inversé l'un des items afin d'éviter un effet de halo. De façon classique, mis à part la mesure du comportement de réponse, l'ensemble des échelles retenues est au format Likert en 5 points (de « pas du tout d'accord » à « tout à fait d'accord »).

Validité et fiabilité des échelles de mesure

Les échelles de mesure ont fait l'objet d'une procédure de validation quantitative sur la base d'une analyse factorielle exploratoire (analyses factorielles en composantes principales avec SPPS 11.5), doublée d'une analyse confirmatoire (modèle de mesure avec AMOS 5).

Analyses exploratoires – Les analyses factorielles ont tout d'abord été réalisées concept par concept, de façon à vérifier rapidement l'unidimensionnalité de chaque échelle et confirmer l'élimination d'un premier ensemble d'items dont la variance était jugée faible, suite à une analyse descriptive univariée. Les résultats des analyses factorielles sont présentées ci-dessous (Tableau 2). Ils conduisent à accepter provisoirement l'unidimensionnalité et la fiabilité de chaque échelle, avec des coefficients tout à fait respectables pour des échelles en construction (à l'exception de celle relative à la préoccupation pour le respect de la vie privée) et contenant peu d'items non redondants (cf liste des items, Annexe 2). Des analyses factorielles ont ensuite été menées en considérant les concepts deux à deux, afin d'étudier la robustesse et la validité discriminante des échelles de mesure (Tableau 3). Cette étape intermédiaire confirme les résultats de l'étape précédente. L'analyse finalement présentée (Tableau 4) considère simultanément les quatre concepts qui permettent de rendre compte de la situation de collecte, telle que perçue par le répondant.

À l'exception d'un item dont le coefficient de corrélation avec son facteur est faible mais acceptable (Confidentialité 3), les résultats sont tout à fait satisfaisants : la qualité des extractions est systématiquement supérieure à 0,5 et chaque item ne charge que sur un seul facteur.

Tableau 2 – Analyses factorielles avec les concepts pris un par un (étape 1)

Concept	Items initiaux	Items finaux	KMO / Test de Bartlett	% de variance	Qualité de la représentation		α de Cronbach
					Extraction minimale	Extraction maximale	
Confidentialité (CONF)	4	3	0,668 / 0,000	65 ,4%	0,58	0,72	0,73
Sensibilité (SENS)	5	4	0,728 / 0,000	67,1%	0,60	0,74	0,83
Pertinence (PERT)	4	3	0,687 / 0,000	70,5%	0,63	0,73	0,79
Equité (EQUIT)	5	4	0,743 / 0,000	65,7%	0,61	0,73	0,82
Attitude face à la divulgation (ATT)	4	3	0,723 / 0,000	77,7%	0,75	0,81	0,85
Préoccupation pour le respect de la vie privée (PREOC)	4	3	0,729 / 0,000	77,2%	0,74	0,78	0,85

Tableau 3 – Analyses factorielles avec les concepts pris deux à deux (étape 2)

Concept	Items initiaux	Items finaux	KMO / Test de Bartlett	Variance de chaque facteur (*)	Qualité de la représentation		α de Cronbach
					Extraction minimale	Extraction maximale	
Confidentialité (CONF)	3	3	0,760 / 0,000	31,7%	0,55	0,75	0,73
Sensibilité (SENS)	4	4		34,1%	0,63	0,72	0,83
Pertinence (PERT)	3	3	0,818 / 0,000	32,1%	0,66	0,75	0,79
Equité (EQUIT)	4	4		36,5%	0,59	0,78	0,82
Attitude face à la divulgation (ATT)	3	3	0,854 / 0,000	38,9%	0,74	0,81	0,85
Préoccupation pour le respect de la vie privée (PREOC)	3	3		38,4%	0,75	0,80	0,85

(*) % de variance retenus pour la rotation orthogonale

Une analyse complémentaire pour les variables d'attitude vis-à-vis de la collecte (ATT) et de préoccupation pour le respect pour la vie privée (PREOC) permet de réduire chacune de ces échelles à deux items, pour des raisons d'efficacité. Le coefficient de corrélation entre chacun des deux items est respectivement de 0,68 ($p < 0,001$) et 0,63 ($p<0,001$).

Tableau 4 – Analyse factorielle des quatre variables perceptuelles spécifiques (étape 3)

		Facteur 1 (*)	Facteur 2	Facteur 3	Facteur 4	Qualité de la représentation
Sensibilité (SENS)	1	**0,799**				0,757
	2	**0,821**				0,748
	3	**0,667**		-0,371		0,625
	4	**0,753**				0,651
Confidentialité (CONF)	1				**0,831**	0,745
	2				**0,816**	0,723
	3	-0,331	-0,394		**0,560**	0,578
Pertinence (PERT)	1			**0,750**		0,697
	2			**0,757**		0,642
	3			**0,793**		0,730
Equité (EQUIT)	1		**0,609**	-0,330		0,577
	2		**0,858**			0,790
	3		**0,614**	-0,389		0,621
	4		**0,819**			0,736
% de variance		19,3%	14,2%	16,7%	18,5%	

(*) Les coefficients inférieurs à 0,3 n'ont pas été représentés.

Analyses confirmatoires – Des analyses confirmatoires ont ensuite été réalisées. Les variables qui permettent de qualifier la perception de la sollicitation ont été considérées simultanément : confidentialité et sensibilité perçue des données sollicitées ; pertinence et équité perçue de la sollicitation ; attitude générale vis-à-vis de celle-ci. Toutefois, ce modèle comprenant 16 variables n'a pas présenté des indices d'ajustement satisfaisants. Nous l'avons spécifié à nouveau, en éliminant trois variables problématiques (Sensibilité 3, Sensibilité 4 et Equité 2). Dès lors, le modèle final présente des indices d'ajustement tout à fait convaincants : χ^2 / dl = 1,57 ; GFI = 0,957 ; AGFI = 0,928 ; RMSEA = 0,043.

La fiabilité des échelles de mesure est également satisfaisante avec des coefficients de cohérence interne (ρ) respectifs de : CONF = 0,739 ; SENS = 0,875 ; PERT = 0,794 ; EQUIT = 0,757 ; ATT = 0,810. Concernant la validité convergente, les coefficients (ρVC) sont acceptables pour l'ensemble des construits, à l'exception de la confidentialité qui atteint le seuil inférieur : CONF = 0,486 ; SENS = 0,778 ; PERT = 0,564 ; EQUIT = 0,511 ; ATT = 0,686.

Pour finir, concernant la validité discriminante, l'examen des coefficients du modèle de mesure montre que le coefficient de validité convergente est systématiquement supérieur au carré du coefficient de corrélation entre les construits pris deux à deux (r^2_{ij}), à deux exceptions près : le coefficient de corrélation entre l'attitude et la pertinence (r^2_{ij} = 0,672) et entre l'attitude et l'équité (r^2_{ij} = 0,547). Cependant, la validité discriminante est clairement établie entre les concepts qui permettent de qualifier les aspects spécifiques de la sollicitation : confidentialité, sensibilité, pertinence et équité. Seule la validité discriminante entre certaines de ces perceptions spécifiques et l'évaluation globale pose problème. D'une certaine manière, ces perceptions spécifiques peuvent être considérées comme des indicateurs qui permettent d'approcher l'évaluation globale (*proxy measures*).

Finalement, la qualité des échelles étant jugée satisfaisante aux regard des critères habituellement utilisés (α de Cronbach et ρ de Jöreskorg, par exemple), nous avons crée une nouvelle série de variables par simple addition des items.

PRÉSENTATION ET DISCUSSION DES RÉSULTATS

La question se pose maintenant de savoir : (1) si les répondants se sont (ou non) comportés de façon bienveillante face à la demande qui leur était faite de communiquer des informations personnelles ; (2) si cette réaction comportementale s'explique ou non par une prédisposition générale en rapport avec la protection de la vie privée ; (3) si, au contraire, cette réponse est due à la perception du contexte par les répondants, en terme de sensibilité des données ou d'équité, par exemple ? Nous allons successivement aborder ces questions dans les trois sections qui viennent.

Le comportement face à une sollicitation portant sur des données personnelles

Sur la base de notre échantillon, nous pourrions qualifier de « bienveillant », le comportement de réponse face à une sollicitation portant sur des données personnelles. En effet, il apparaît tout d'abord que 42,5% des individus sollicités auraient accepté de répondre au formulaire mis en ligne sur le site factice. Par ailleurs, 57,5% des répondants ont accepté de communiquer leur mail et 48,6% ont accepté de recevoir des offres personnalisées de la part du site de voyage factice qu'ils visitaient [3].

[3] En revanche, le pourcentage de répondants qui déclare accepter de recevoir des offres en provenance d'entreprises partenaires du site de voyage chute à 26,6% (- 22 points)

Ces différents indices traduisent, selon nous, un comportement plutôt ouvert à l'égard des sollicitations portant sur des données personnelles, au moins pour une petite moitié des individus sollicités.

Cependant, dans le même temps, de nombreuses personnes semblent réticentes à l'idée de participer à cette étude, de communiquer des informations personnelles ou encore, de laisser une trace de leur passage sur le site marchand [4]. Tout d'abord, plus de la moitié des répondants auraient refusé de répondre à un tel formulaire si la situation à laquelle ils étaient confrontés s'était présentée dans la réalité (57,5%). Les raisons avancées concernent essentiellement la crainte que ces données soient utilisées par une tierce personne, sans leur consentement (26%). Le manque de temps ne concerne que 9,4% de ces (non)répondants [5]. Par ailleurs, 75% des répondants avouent qu'ils auraient été tentés de donner au moins une fausse réponse à certaines des questions posées (en moyenne 3,4 fausses réponses, avec un écart-type de 2,7). Les répondants sont notamment tentés de donner une fausse réponse quand il s'agit de leur téléphone (53%), de l'adresse mail (26%) ou de leur date de naissance (18%). Finalement, la quasi-totalité de l'échantillon (98,3%) a jugé « sensible » au moins une information demandée. Il s'agissait non seulement des coordonnées, comme le numéro de téléphone (75%), l'adresse mail (36%) ou l'adresse (23%), mais également d'autres données comme la date de naissance (39%), la situation familiale (31%), le nombre d'enfants (26%) ou encore, le niveau d'étude (21%).

En synthèse, le comportement de notre échantillon est contrasté. Un peu plus de la moitié des individus n'auraient pas répondu au formulaire. Cependant, pour ceux qui auraient accepté de répondre, le formulaire contenait au moins une donnée « sensible », dans la quasi-totalité des cas (notamment le téléphone, l'adresse mail ou l'année de naissance) et plus de la moitié auraient donné au moins une fausse réponse.

[4] Rappelons pourtant que notre étude porte sur un échantillon de convenance qui fait plutôt confiance aux chercheurs.

[5] A ce stade, on peut objecter que les individus qui n'avait vraiment pas le temps n'ont même pas pris la peine d'aller si loin dans le questionnaire…

Ces résultats sont difficiles à comparer avec ceux issus de la littérature, dans la mesure où, à notre connaissance, aucune des recherches passées n'a mesuré directement le comportement de réponse. Ils semblent néanmoins confirmer ceux obtenus par Acquisti et Grossklags (2004) qui trouvent que 58,3% de leurs répondants se disent très préoccupés par une collecte de données portant sur leur identité en ligne (adresse mail, notamment) et hors ligne (nom, adresse, téléphone…). Nos résultats portant sur le niveau de sensibilité de certains types de données semblent également confirmer ceux issus d'études passées, notamment en ce qui concerne le caractère privé du numéro de téléphone, pouvant amener certaines personnes à donner un faux numéro si celui-ci est demandé.

Cohérence entre la préoccupation générale et le comportement effectif

Les individus fortement préoccupés par la protection de leur vie privée acceptent-ils pourtant de se dévoiler ? À l'inverse, les individus peu préoccupés par la protection de leur vie privée refusent-ils pour autant de se dévoiler ? Finalement, les convictions sont-elles cohérentes avec les comportements ? Pour investiguer notre première proposition, nous avons croisé le comportement en situation avec la préoccupation générale envers la protection de la vie privée (PREOC) [6]. Nous avons ensuite comparé l'attitude générale face à la sollicitation (ATT) selon le niveau de ces deux variables.

Les résultats montrent (Tableau 5) qu'il existe un lien significatif entre le comportement face à la sollicitation (réponse / non réponse au formulaire) et le niveau de préoccupation (faible / fort), le test du khi-deux de Pearson étant hautement significatif ($\chi^2 = 12,44$; $p < 0,001$). De fait, le taux de non réponse est de 40,3% pour les individus faiblement préoccupés contre 63,2% pour les individus fortement préoccupés. Naturellement, à l'inverse, le taux de réponse est de 59,7% pour les individus faiblement préoccupés contre 36,8% pour les individus faiblement préoccupés. Concernant l'attitude envers la sollicitation, les résultats confirment que le score d'attitude est significativement plus faible auprès des individus préoccupés par le respect de la vie privée (score de 4,2 contre 6,5 ; $F = 100,1$; $p < 0,001$) et auprès des individus qui ne répondent pas au formulaire proposé (score de 4,1 contre 5,7 ; $F = 62,7$; $p < 0,001$).

[6] La variable de préoccupation envers le respect de la vie privée a été recodée en deux catégories sur la base de la distribution observée, avec un seuil de 5 pour la catégorie « faiblement préoccupés » (échelle allant de 2 à 10).

Les résultats permettent donc d'établir l'existence d'un lien significatif entre préoccupation générale vis-à-vis du respect de la vie privée, attitude négative vis-à-vis d'une sollicitation spécifique et comportement de refus face à un formulaire. Ceci dit, le tableau croisé indique qu'un pourcentage non négligeable de répondants ne sont pas « cohérents » : soient ils sont fortement préoccupés, mais ils répondent (n = 85, soit 27,6% de l'effectif total); soit ils sont faiblement préoccupés, mais ils ne répondent pas (n = 31, soit 10,1% de l'effectif total).

En synthèse, si le comportement de réponse semble cohérent avec les attitudes, les résultats mettent également en évidence deux groupes d'individus « incohérents », qui représentent à eux deux, un bon tiers de l'effectif et qui nécessitent très certainement des analyses ultérieures (motivations, profils…).

Cette cohérence établie entre le niveau de préoccupation envers le respect de la vie privée et le comportement de réponse à une sollicitation de données personnelles (au moins pour deux tiers de nos répondants) confirme de nombreux résultats en ce sens dans la littérature. Ainsi, Sheehan (2002) trouve une forte corrélation entre le niveau de préoccupation et plusieurs comportements associés, comme le fait de se plaindre auprès de son fournisseur d'accès à Internet, de ne pas s'enregistrer sur les sites, de fournir des données erronées ou incomplètes. De même, l'incohérence de certaines personnes a également été suggérée dans la littérature (Acquisti, 2004), même si les explications d'un tel comportement sont encore balbutiantes. Ainsi, au delà des distorsions psychologiques, susceptibles d'intervenir dans de telles situations, des raisons plus rationnelles ont aussi été avancées, notamment concernant le rôle de l'expérience passée. Ainsi, un individu ayant récemment vécu une invasion de sa vie privée (réception d'un courriel non sollicité, par exemple) pourra être globalement peu préoccupé, mais ne pas souhaiter laisser ses coordonnées pour éviter de vivre à nouveau ce type d'expérience.

Acquisti et Grossklags (2004) montrent que le niveau de sensibilité associé à la divulgation de données personnelles dans une situation particulière est généralement inférieur au niveau de préoccupation générale dans ce domaine, soulignant ainsi l'impact de la situation, dont nous allons à présent évaluer l'importance dans le cadre de notre étude.

Tableau 5 – Cohérence entre le comportement de réponse et le degré de préoccupation pour le respect de la vie privée

		Degré de préoccupation pour le respect de la vie privée	
		Faible	Fort
Comportement de réponse	Non	40,3%	63,2%
	Oui	59,7%	36,8%

Impact des perceptions situationnelles

La perception du contexte de sollicitation par l'individu influence-t-elle la décision de se dévoiler (ou non) ? Le contexte est-il prépondérant par rapport aux dispositions individuelles, comme la préoccupation générale pour le respect de la vie privée ? Pour investiguer notre seconde proposition, nous avons réalisé une série de régressions logistiques binomiales univariées (avec le comportement de réponse dichotomique en variable dépendante), puis plusieurs régressions logistiques multivariées, en introduisant successivement les variables explicatives (3 blocs) : perception de la situation – confidentialité (CONF), sensibilité (SENS), pertinence (PERT) et équité (EQUIT) – préoccupation générale vis-à-vis du respect de la vie privée (PREOC) – attitude vis-à-vis de la sollicitation (ATT).

Régressions logistiques univariées - Toutes les variables étudiées sont significativement liées au comportement de réponse et ce, dans le sens attendu. Les ratios de log-vraissemblance, qui traduisent bien la capacité explicative du modèle, sont particulièrement faibles pour l'attitude vis-à-vis de la sollicitation (364,6), la pertinence (380,5), la sensibilité (381,5) et l'équité perçue de la sollicitation (392,0). Dans cette première analyse, la préoccupation vis-à-vis de la vie privée (405,9) et la confidentialité (407,1) sont les deux variables les moins explicatives.

Ces résultats, qui restent à confirmer par une analyse multivariée, montrent que la perception globale de la situation (appréhendée avec le concept d'attitude) est la variable la plus explicative du comportement de réponse.

Régressions logistiques multivariées – Les résultats présentés ci-dessous (Tableau 6) confirment que la sensibilité perçue des données est une variable explicative du comportement de réponse (plus les données sont perçues comme sensibles, plus la probabilité de remplir le formulaire est faible). La pertinence des données demandées est également une variable importante, quoi que non significative une fois l'attitude introduite dans le modèle (la corrélation entre ces deux variables est significative).

Le sens de la relation est cohérent avec l'intuition et la théorie : plus la pertinence de la demande est perçue comme élevée, plus la probabilité de remplir le formulaire est élevée. La confidentialité perçue des données et l'équité perçue de la sollicitation ne sont pas significativement liées au comportement de réponse, contrairement à nos attentes. Comme attendu, en revanche, la perception globale de la situation de collecte est fortement significative (ce qui conduit à une baisse du ratio de log-vraisemblance). En revanche, la disposition individuelle générale (PREOC) ne l'est pas.

Tableau 6 – Impact des variables situationnelles et attitudinales sur la probabilité de répondre au formulaire (régression logistique)

	Modèle A		Modèle B		Modèle C	
	B	Sig.	B	Sig.	B	Sig.
Constante	1,663	0,219	0,199	0,892	-0,489	0,759
SENS	**-0,233**	**0,001**	**-0,170**	**0,020**	**-0,179**	**0,015**
CONF	-0,013	0,821	-0,059	0,332	-0,048	0,434
PERT	**0,170**	**0,003**	0,100	0,103	0,099	0,109
EQUIT	-0,128	0,078	-0,089	0,238	-0,098	0,198
ATT			**0,296**	**0,002**	**0,332**	**0,001**
PREOC					0,087	0,272
-2LL	361,9		352,2		351,0	
R^2 de Cox	0,172		0,198		0,201	

En synthèse, la perception globale de la situation domine (ATT : attitude vis-à-vis de la sollicitation). Une variable situationnelle spécifique renforce le modèle : la sensibilité perçue des données demandées (SENS). Compte tenu d'une capacité explicative toutefois modeste, les autres variables situationnelles sont dominées par le jugement global, notamment la pertinence, pourtant fortement significative dans le modèle univarié ainsi que dans le modèle A (Tableau 6). En ce qui concerne la disposition individuelle (PREOC : préoccupation vis-à-vis du respect de la vie privée), on peut conclure que cette variable, fruit de l'expérience et de croyances générales, n'est pas significative, une fois introduite l'attitude dans le modèle. Autrement dit, pour répondre clairement à nos interrogations de départ, la perception du contexte de sollicitation influence significativement la décision de se dévoiler (ou non) et par ailleurs, la perception du contexte est prépondérant par rapport aux dispositions individuelles.

Ces résultats confirment ceux obtenus par Walczuch et al. (2001) qui montrent que les perceptions relatives à la situation (confidentialité et pertinence) ont un lien significatif avec la confiance dans le e-commerce et donc l'intention de se dévoiler [7]. Ces auteurs montrent aussi l'absence de lien significatif entre certaines variables individuelles (traits de personnalité, attitude envers l'ordinateur, Internet et le shopping) et la confiance dans le e-commerce. Le fait que l'attitude vis-à-vis de la sollicitation soit le principal facteur explicatif du comportement de réponse est également conforme à la théorie de l'action raisonnée (Fishbein et Ajzen, 1975) qui prédit que le comportement est déterminé par l'intention comportementale, elle-même déterminée par l'attitude. Pour leur part, Pavlou et Chellappa (2001) montrent que la perception du respect de la vie privée a un impact plus faible sur la confiance dans le e-commerce et donc sur l'intention de se dévoiler que la sécurité perçue, concept plus global. Notre concept de confidentialité perçue étant lui-même une sous dimension de la perception de respect de la vie privée, cela pourrait expliquer en partie son faible impact sur la décision de répondre, dans le cadre de notre étude.

CONTRIBUTIONS, LIMITES ET VOIES DE RECHERCHE

Contributions de la recherche

Les contributions de cette recherche sont à la fois académiques et managériales.

Sur le plan académique, notre étude éclaire la situation française, dans la mesure où, à notre connaissance, aucune investigation spécifique n'avait été menée sur ces questions. En effet, la seule étude publiée portait sur les pratiques des sites et non sur les perceptions des internautes (Gauzente, 2003). Plus largement, cette étude est la première à étudier le comportement de réponse des clients (i.e., le fait de remplir ou non un formulaire) et non pas seulement les perceptions, notamment la confiance envers le site, qui constitue la variable dépendante habituelle. Finalement, nous proposons un ensemble d'échelles validées pour approfondir notre connaissance de ces sujets.

D'un point de vue managérial, plusieurs implications doivent aussi être soulignées. Ainsi, nos résultats doivent encourager les managers à ne pas rendre obligatoire la communication de données perçues comme trop sensibles. En effet, une telle demande peut amener certains

[7] Précisément, les auteurs utilisent les termes de contrôle (i.e., une facette de la confidentialité) et de normalité (i.e. la pertinence).

individus, soit à ne pas répondre, soit à donner de fausses réponses, ce qui est contre-productif dans les deux cas. De plus, l'entreprise devra adapter la demande à l'état de la relation avec ses clients. En effet, certaines données peuvent être perçues comme sensibles à un certain moment, notamment au début de la relation, mais plus après quelques échanges. Ce serait notamment le cas de la date de naissance ou de la situation familiale dont on peut comprendre l'intérêt de les fournir dans le cadre d'un programme relationnel, mais pas lors d'un premier contact, dont on ne sait pas à l'avance s'il y aura une suite.

L'entreprise devra aussi faire comprendre la pertinence de la demande, dont on a vu qu'elle était fortement corrélée à l'attitude générale vis-à-vis de la sollicitation, au sein de notre modèle. Cela passe notamment par le fait d'expliquer aux clients pourquoi telles ou telles données sont collectées, dans quel but et se qui en sera fait à l'avenir.

Si la confidentialité perçue des données collectées ne paraît pas 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 peut être pas sa décision de remplir le formulaire. Mais elle pourrait expliquer d'autres comportements, comme la décision de donner certaines informations plus sensibles, ou le fait de mentir.

De même, si dans cette recherche et compte tenu des limites évoquées ci-dessous, la préoccupation vis-à-vis du respect de la vie privée ne joue pas un grand rôle dans la décision de remplir ou non le formulaire, cette variable peut toutefois donner lieu à des recommandations sur les contenus mis en ligne pour rassurer la cible des consommateurs très préoccupés (environ 25% des individus ont un score très élevé).

Dans une perspective de segmentation / ciblage, nous avons testé également le lien entre la probabilité de remplir le formulaire et quelques variables socio-démographiques et comportementales [8]. L'introduction de ces variables améliore le modèle. Plus spécifiquement, il apparaît que la probabilité de remplir le formulaire est supérieure pour les individus plus jeunes (moins de 35 ans) et moins éduqués. En revanche, de nombreuses variables ne sont pas significatives, notamment le genre et l'usage de l'Internet (ancienneté d'utilisation, nombres d'heures d'utilisation, expérience d'achat en ligne).

[8] Une fois introduites les variables précédemment significatives : attitude générale vis-à-vis de la sollicitation (ATT) et sensibilité perçue des données (SENS).

Limites de la recherche

Malgré les conclusions intéressantes que cette recherche nous permet de tirer, celle-ci n'est pas sans présenter quelques limites, qu'il faudra prendre en compte dans les études futures sur le sujet.

La première limite concerne les répondants et la manière dont ils ont été sélectionnés, ce qui aboutit à la formation d'un échantillon de convenance qui n'est pas représentatif de la population des internautes français (e.g., sur-représentation des femmes ou des 25-34 ans). Par ailleurs, le problème de l'auto-sélection doit également être souligné, dans la mesure où les individus contactés par email avaient le choix de participer ou non à l'étude. Or, comme le suggère Hood et Back (1971), le phénomène de divulgation de soi pourrait être lié au principe du volontariat, notamment pour les hommes. Ainsi, ceux qui se porteraient régulièrement volontaires pour participer à des études ou expérimentations parleraient plus facilement d'eux, ce qui pourrait alors biaiser les résultats. On peut penser que les gens sensibles au thème de la vie privée sont sous-représentés dans l'échantillon, malgré l'assurance d'anonymat et de confidentialité conférée. Plusieurs techniques peuvent alors être utilisées pour limiter la part des non répondants, notamment la participation à une loterie pour gagner un prix. Tuten, Bosnjak et Bandilla (2000) montrent cependant que la part des non répondants est plus faible quand des motivations altruistes sont offertes, comme par exemple la contribution à la recherche, ce qui était le cas dans notre étude, limitant ainsi ce biais de non réponse.

Une deuxième limite concerne la mise en situation et la façon dont les répondants sont plongés dans les scenarii. En effet, dans le domaine du respect de la vie privée, on a souvent affaire à des réponses émotionnelles qui sont difficiles à prédire de façon juste et réaliste. De plus, si la situation étudiée est « distante » de la personne qui l'évalue, celle-ci pourra se montrer psychologiquement détachée et son évaluation sera alors faussée. C'est ce qui semble s'être passé dans notre étude, dans la mesure où l'impact des variables situationnelles s'est avéré plus faible que prévu.

Une dernière limite à considérer concerne le biais de désirabilité sociale, qui est souvent présent lors des enquêtes. Dans ce cas là, l'interviewé cherche à agir comme « un bon répondant », en donnant les réponses qu'il estime qu'on attend de lui, plutôt que ce qu'il pense vraiment. Il existerait aussi un biais de demande, qui inciterait les gens ayant accepté de participer à notre étude à répondre favorablement au formulaire de collecte de leurs données personnelles alors qu'ils n'y auraient pas répondu dans la réalité. Cette tendance se voit notamment dans le fait que certains ont accepté de donner des informations dont ils estiment pourtant qu'elles sont sensibles, voire très personnelles.

Voies de recherche futures

La prise en compte de ces limites offre de nouvelles perspectives d'études. Dans un premier temps, afin de valider les résultats issus de la présente étude, il semble important de mener une recherche équivalente auprès d'un échantillon représentatif de la population des internautes français. De même, il serait intéressant de recourir à un design entièrement expérimental, afin de mesurer de manière plus fiable l'impact des variables situationnelles (nombre et sensibilité des données demandées, degré de familiarité avec l'entreprise, circonstances de collecte et avantages liés à l'échange…), tout en contrôlant au maximum l'effet de variables externes, et notamment des données individuelles (expérience d'invasion de la vie privée, niveau de préoccupation quant au respect de la vie privée…). Comme nos recherches le suggèrent, il serait également bon de tester le rôle modérateur de la préoccupation envers le respect de la vie privée et le rôle médiateur de l'attitude face à la divulgation.

De plus, si cette recherche se penche plus particulièrement sur la collecte de données en ligne, il serait tout aussi utile d'étudier les préoccupations et réactions des consommateurs pour d'autres médias de collecte tels que le téléphone, le courrier ou le face à face, et de les comparer à celles occasionnées par Internet. Une analyse approfondie des freins et des motivations liées à la décision de répondre ou non à une sollicitation de données personnelles serait également nécessaire, notamment en ce qui concerne les deux groupes d'« incohérents » identifiés dans notre étude, à savoir des gens très préoccupés mais qui décident de répondre et des personnes peu préoccupées mais qui ne répondent pas. Une première phase qualitative permettant d'analyser les motivations pourrait se poursuivre par une phase quantitative visant à étudier le profil de ces personnes et les principaux déterminants de leur comportement de réponse.

Annexe 1 – Formulaire correspondant au scénario 2 (newsletter)



Vous souhaitez recevoir tous les mois notre newsletter adaptée à vos goûts et vos préférences ?

Afin de mieux vous connaître et ainsi adapter la lettre à votre profil, merci de bien vouloir remplir les informations ci-dessous.

Les champs de saisie indiqués par un * doivent obligatoirement être renseignés.

Merci de préciser votre format de réception … ◉ HTML ○ TEXTE

et les destinations qui vous intéressent plus particulièrement ☐ Europe - Amérique ☐ Asie - Antilles

Nom :

Adresse :

Code Postal :

Ville :

Adresse Email :

Téléphone :

Date de naissance (jj/mm/aaaa) :

Votre situation familiale : célibataire

Votre situation professionnelle : cadre supérieur - dirigeant

Votre niveau d'études : < ou = au bac

Nombre d'enfants dans votre foyer : 0

Vos loisirs et centres d'intérêt : ☐ Voyages ☐ Lecture ☐ Arts / cinéma ☐ Informatique

Nous partageons occasionnellement notre base de données et donc les renseignements vous concernant avec des entreprises dont les produits /services sont susceptibles de vous intéresser.

☑ J'accepte de recevoir des offres personnalisées de votre part

☑ J'accepte de recevoir des offres en provenance d'entreprises partenaires

Annexe 2 – Liste des items

Concept	Items
Confidentialité perçue (CONF)	*Je pense que l'entreprise respectera la confidentialité des informations que j'envisage de lui donner* *Je crois que l'entreprise considère la confidentialité comme quelque chose d'important* *Je crains que cette entreprise n'utilise ces données sans mon accord préalable*
Sensibilité perçue (SENS)	*Globalement, les données demandées sont trop personnelles* *Je considère que les informations qui me sont demandées sont privées*
Pertinence perçue (PERT)	*Fournir ces informations me semble justifié* *La collecte de ces données ne me semble pas nécessaire* *Il me semble normal que l'entreprise me demande toutes ces informations*
Equité perçue (EQUIT)	*Dans cette situation, l'échange avec l'entreprise n'est pas équilibré* *L'échange de mes données personnelles avec l'entreprise se fait sur des bases équitables* *L'entreprise va retirer beaucoup plus que moi de cet échange*
Attitude face à la divulgation (ATT)	*Cela ne me pose aucun problème de livrer ce genre d'informations* *Je suis favorable à l'idée de donner ces informations*
Préoccupation pour le respect de la vie privée (PREOC)	*Je suis inquiet à l'idée que les entreprises puissent détenir des données que je considère comme privées* *L'idée que mes données personnelles puissent être utilisées sans ma permission préalable ne m'inquiète pas particulièrement*

REFERENCES

Auteur (2002), à compléter.

Auteur (2003), à compléter.

Acquisti A. (2004), Privacy in electronic commerce and the economics of immediate gratification, *Electronic Commerce Conference*, New York, USA.

Acquisti A. et Grossklags J. (2004), Privacy and rationality: evidence from survey data, *third annual Workshop on Economics and Information Security (WEIS 2004).*

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.

Berendt B., Gunther O et Spiekermann S. (2003), Privacy in e-commerce: stated preferences vs. actual behavior, *Communications of the ACM*.

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. et Armstrong P. (1999), Information privacy concerns, procedural fairness and impersonal trust: an empirical investigation, *Organization Science*, 10, 1, 104-115.

Derlega V., Harris M. et Chaikin A. (1973), Self-disclosure reciprocity, liking and the deviant, *Journal of Experimental Social Psychology*, 9, 277-284.

Dinev T. and Hart P. (2002), Internet privacy concerns and trade off factors: empirical study and business implications, *International Conference On Advances In Infrastructure For E-Business*, L'Aquila, Italie.

Dommeyer C. et Gross B. (2003), What consumers know and what they do: an investigation of consumer knowledge, awareness and use of privacy protection strategies, *Journal of Interactive Marketing*, 17, 2, 34-51.

Federal Trade Commission (1998), FTC Releases Report on Consumers Online Privacy, *FTC Press Release*, File N° 954-4807, http://www.ftc.gov.

Fusilier M. et Hoyer W. (1980), Variables affecting perceptions of invasion of privacy in a personnel selection situation, *Journal of Applied Psychology*, 65, 5, 623-626.

Gauzente (2003), Vie privée et Internet : la pratique des sites français, *Décisions Marketing*, 30, 7-13.

Gandy O. (1993), The panoptic sort: a political economy of personal information, New York:Westview Press.

Gerber E. (2001), The privacy context of survey response: an ethnographic account, in C*onfidentiality, disclosure and data access: theory and practical applications for statistical agencies,* New York: North Holland.

Fishbein M. et Ajzen I (1975), *Belief, attitude, intention and behavior*, Reading, MA :Addison-Wesley

Foxman E. et Kilcoyne P. (1993), Information technology, marketing practice and consumer privacy: ethical issues, *Journal of Public Policy and Marketing*, 12, 1, 106-126.

Gauzente C., Benetteau J. et Dubreuil E. (2002), Respect de la vie privée sur Internet : une étude des pratiques des sites français, *1ère journée nantaise de recherche sur le e-marketing,* 117-133.

Goffman E. (1959), *The presentation of self in everyday life*, New York: Doubleday.

Hine C. et Eve J. (1998), Privacy in the marketplace, *The Information Society*, 14, 253-262.

Hoffman D., Nowak T. et Peralta M. (1999), building consumer trust online, *Communications of the ACM*, 42, 4, 80-85.

Hood T. et Back K. (1971), self-disclosure and the volunteer: a source of bias in laboratory experiments, *Journal of Personality and Social Psychology*, 17, 2, 130-136.

Kehoe C., Pitkow J. and Rogers J. (1998), GVU's Ninth WWW User Survey Report, Office of Technology Licensing, *Georgia Tech Research Corporation*, Atlanta.

Kelvin P. (1973), A social-psychological examination of privacy, *British Journal of Social and Clinical Psychology*, 12, 248-261.

Kimmel, A. (1996), *Ethical issues in behavioural research: a study*, Cambridge, MA: Blackwell Publishers.

LaPiere R. (1934), Attitudes vs. action, *Social Forces*, 13, 2, 230-237.

Lemley, M. (2000), Rational ignorance at the patent office, Cahier de recherche, Technical report, Berkeley Olin Program in Law and Economics.

McCloskey H. (1980), Privacy and the right to privacy, *Philosophy,* 55, 17-38.

Milne G. R. et Gordon M. E (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 Culnan M. (2004), Strategies for reducing online privacy risks: why consumers read (or don't read) online privacy notices, *Journal of Interactive Marketing*, 18, 3, 15-29.

Miyazaki A.D. 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.

NFO Interactive (1999), Consumers to e-tailers : don't kiss and tell, http://cyberatlas.internet.com.

Nguyen D., Gray G. et Kira D. (2000), Trust and quality assurance in business-to-consumer electronic commerce: enhancing consumer acceptance and participation, *Proceedings of Pacific Asia Conference on Information Systems* (*PACIS 2000),* Hong Kong.

Oliver R. et Swan J. (1989), Equity and disconfirmation perceptions as influences on merchant and product satisfaction, *Journal of Consumer Research*, 16, 372 – 383.

Oliver R. et Swan J. (1989), Consumer perceptions of interpersonal equity and satisfaction in transactions: a field survey approach, *Journal of Marketing*, 53, 21-35.

Pavlou P. et Chellappa R. (2001), The role of perceived privacy and perceived security in the development of trust in electronic commerce transaction, cahier de recherche, Ebizlab, Marshall School of Business, Los Angeles.

Richards J. (1997), The legal potholes on the information superhighway, *Journal of Public Policy and Marketing*, 16, 2, 319-326.

Schwartz B. (1968), The social psychology of privacy, *The American Journal of Sociology*, 73, 6, 741-752.

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

Simon H. (1976), *From substantive to procedural rationality,* in Method and Appraisal in Economics, Cambridge: Cambridge University Press.

Singer E. and Presser S. (1996), Public attitudes toward data sharing by federal agencies, *Annual Research Conference*, Bureau of the Census.

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

Stone E. et Stone D. (1990), Privacy in organizations: theoretical issues, research findings and protection mechanisms, *Research in Personnel and Human Resources Management*, 8, 349-411.

Turner E. et Dasgupta S. (2003), Privacy on the web : an examination of user concerns, technology and implications for business organizations and individuals, *Information Systems Management*, 8-18.

Tuten T., Bosnjak M. et Bandilla W. (2000), Banner-advertised web-surveys, *Marketing Research*, 11, 4, 17-21.

Walczuch R., Seelen J. et Lundgren H. (2001), Psychological determinants for consumer trust in e-retailing, *Proceedings of the 8th research symposium on emerging electronic markets*, Maastricht, September 16-18.

Warren S. and Brandeis L. (1890), The right to privacy, *Harvard Law Review*, 4, 193.

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.

Woodman R. et al. (1982), A survey of employee perceptions of information privacy in organizations, *Academy of Management Journal*, 25, 3, 647-663.

Yin R. (1990), Case studies research: design and methods, Newbury Park, Sage.