

Quels sont les effets de la qualité ergonomique d'un site web et de l'expérience des utilisateurs sur la charge cognitive et le temps de navigation ?

Aline Chevalier et Maud Kicka
Université de Paris X - Nanterre
Laboratoire des Processus Cognitifs et
Conduites Interactives (EA 1588)
U.F.R. SPSE
200, avenue de la République
92000 Nanterre cedex

Aline.Chevalier@u-paris10.fr
maud.kicka@club-internet.fr

Julien Cegarra
IRCCyN-CNRS (UMR 6597)
Équipe PsyCoTec
1, rue de la Noë
B.P. 92101
44321 Nantes cedex 3

Julien.Cegarra@irccyn.ec-nantes.fr

Mots clés

Sites web, charge cognitive, utilisabilité, recherche d'informations.

1. Cadre de l'étude

Le nombre de sites web ne cessent de s'accroître de façon extrêmement rapide, sans pour autant que la qualité en termes d'utilisabilité accompagne cette évolution. En effet, les travaux montrant que les sites web ne sont pas satisfaisants pour les utilisateurs sont toujours aussi nombreux (cf., par exemple, Lee, 1999 ; Teo, Oa, Liu, & Wei, 2003 ; Bhatt, 2004). Aussi, est-il indispensable de conduire des travaux sur le fonctionnement cognitif des utilisateurs, mais également de mieux comprendre les difficultés que rencontrent les concepteurs de ces sites. Plus précisément, nous défendons l'idée selon laquelle les recherches sur les activités de conception et d'utilisation de sites web doivent être mises en œuvre de façon conjointe, si l'on souhaite pouvoir aider les concepteurs à développer une activité centrée utilisateurs. Une première étude conduite par Park, Choi et Kim (2004) a tenté justement de mettre en relation le point de vue des utilisateurs avec celui des concepteurs sur les aspects esthétiques de sites web. Cette étude visait à déterminer si les concepteurs anticipent efficacement les préférences esthétiques (seulement) des utilisateurs (point important dans les sites web à visée commerciale – cf. Kim & Moon, 1998).

Afin de contribuer à cette thématique, plusieurs études expérimentales ont été conduites auprès de concepteurs et d'utilisateurs de sites web (cf. Chevalier & Ivory, 2003 ; Chevalier, 2004 ; Chevalier, sous presse). Ces études ont montré que les concepteurs débutants et professionnels se focalisent sur les besoins du commanditaire au détriment de ceux de l'utilisateur lorsqu'ils réalisent des sites. En d'autres termes, en situation de réalisation de sites web, les concepteurs ont tendance à développer une activité centrée commanditaire, plutôt qu'une activité centrée utilisateurs. Cela a alors des répercussions négatives sur les sites réalisés par les concepteurs, en ce sens qu'un grand nombre de problèmes d'utilisabilité sont introduits par les concepteurs.

Lorsque l'on demande à des concepteurs et à des utilisateurs d'évaluer un site web pré-existant, on note que les concepteurs prennent toujours peu en compte les aspects liés à l'utilisateur (cf. Chevalier, 2004). Ils identifient également moins de problèmes d'utilisabilité que les utilisateurs. Sur la base de ces résultats, plusieurs questions ont été soulevées. L'une d'entre elles a particulièrement retenu notre attention : les problèmes identifiés explicitement par les concepteurs et les utilisateurs perturbent-ils réellement leur activité de recherche d'informations ? Dans le but d'apporter des éléments de réponse à cette question, une nouvelle expérience a été conduite. Sa présentation est l'objet de la partie suivante.

2. Étude expérimentale

Cette étude vise à comparer des concepteurs et des utilisateurs lors d'une activité de recherche d'information au sein d'un web réalisé pour les besoins de l'étude. Dans cet article, nous présentons certains résultats consacrés aux utilisateurs.

2.1. Procédure expérimentale

14 utilisateurs novices et 13 utilisateurs expérimentés ont accepté de participer à cette étude. La passation expérimentale se déroule en trois étapes successives :

Étape 1 : Dans un premier temps, il est nécessaire de déterminer le niveau de connaissances des utilisateurs dans le maniement de l'outil Internet. Pour cela, un questionnaire a été proposé aux utilisateurs avant d'effectuer l'activité de recherche d'informations. Ce questionnaire reprend certaines questions proposées par Strahm (2002). Les questions posées permettent de distinguer deux niveaux d'expérience dans l'utilisation de l'Internet : les utilisateurs considérés comme débutants et les utilisateurs considérés comme expérimentés ou avertis.

Un utilisateur est considéré comme expérimenté dès lors qu'il satisfait les critères suivants : (1) il s'auto évalue comme averti dans le maniement du Web (utilisation et configuration de la messagerie, navigation aisée sur le Web et sur les outils de recherche), (2) qu'il se connecte au moins une fois par jour et que chaque connexion dure au minimum 1 heure (ce critère peut se compenser si le sujet répond qu'il se connecte plusieurs fois par jour au moins ¼ heure par connexion) et (3) qu'il dispose d'une connexion haut débit (par exemple l'ADSL – ce critère est moins important que les deux précédents).

Étape 2 : Afin de déterminer si les problèmes d'utilisabilité identifiés par les concepteurs et utilisateurs, lors de l'étude conduite précédemment (Chevalier, 2004), nuisent à leur recherche d'informations, deux versions d'un même site web sont proposées :

- Une version dite "ergonomique", qui a fait l'objet des deux évaluations par huit étudiants en DESS de psychologie cognitive : association de la méthode du Cognitive Walkthrough (Blackmon, Polson, Kitajima et Lewis, 2002) avec l'évaluation de conformité aux recommandations proposées par Bastien, Leulier et Scapin (1998). Cette version se compose de dix-huit pages liées entre elles par des liens hypertextes.
- Une version dite "non ergonomique", qui comporte les principaux problèmes d'utilisabilité identifiés par les concepteurs et les utilisateurs lors de l'étude conduite par Chevalier (2004). Cette version se compose également de dix-huit pages liées par des liens hypertextes. Chaque page, à l'exception de celles consacrées à la commande passée, comporte le même nombre de problèmes d'utilisabilité, soit 14 problèmes par page web.

Les problèmes d'utilisabilité introduits dans cette version du site web se répartissent de la façon suivante (tableau 1) :

Critères ergonomiques	Description des problèmes introduits
1. Guidage. Incitation.	Le menu ne se trouve pas sur toutes les pages du site.
2. Guidage. Groupement/distinction entre items. Par la localisation.	Les différents items ne se distinguent pas par leur emplacement sur la page du site.
3. Guidage. Groupement/distinction entre items par le format.	Dans le menu, ne figure aucun repère sur l'item correspondant à la page où se trouve l'utilisateur.
4. Guidage. Groupement/distinction entre items par le format.	Le lien de la page où se trouve l'utilisateur est toujours actif.
5. Guidage. Groupement/distinction entre items par le format.	Des textes sont soulignés, alors que ce ne sont pas des liens.
6. Guidage. Lisibilité.	La police de caractère utilisée n'est pas standard.
7. Guidage. Lisibilité.	Le contraste texte/fond n'est pas soutenu, au contraire les couleurs sont complémentaires.
8. Guidage. Lisibilité.	Les textes sont longs (plus de 50 lignes).
9. Charge de travail. Brièveté. Actions minimales.	Certaines pages sont très longues dans le site.
10. Charge de travail. Brièveté. Actions minimales.	Toutes les pages du site ne disposent pas d'un lien de retour vers la page d'accueil.
11. Charge de travail. Densité informationnelle.	Il y a des informations superflues qui risquent de distraire l'utilisateur (par exemple, des publicités).
12. Homogénéité/Cohérence.	Toutes les pages du site n'ont pas la même structure.
13. Homogénéité/Cohérence.	Le lien vers la page d'accueil n'a pas toujours le même intitulé.
14. Signification des codes et dénominations.	Les liens activables ne reflètent pas systématiquement le contenu des pages de destination.

Tableau 1 – Répartition et description des problèmes d'utilisabilité introduits dans la version "non ergonomique" du site web.

Les deux versions de ce site présentent un même magasin fictif qui commercialise des produits liés à la musique (CDs, billets de spectacles, etc.). En guise d'illustration, voici la page d'accueil du site "ergonomique" (figure 1) et celle du site "non ergonomique" (figure 2) :



Figure 1 – Page d'accueil du site "ergonomique"

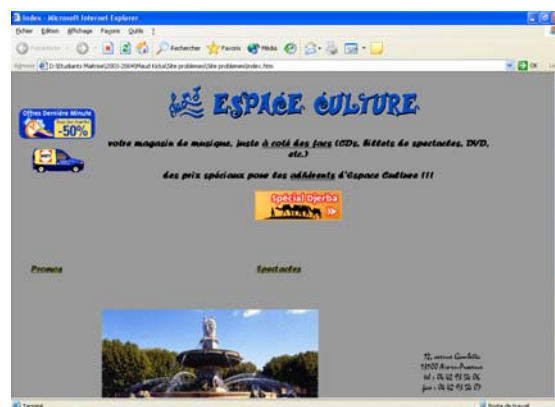


Figure 2 – Page d'accueil du site "non ergonomique"

La moitié des utilisateurs est confrontée à la version du site "ergonomique" et l'autre moitié à la version "non ergonomique".

Les utilisateurs doivent rechercher au sein de ces deux versions des informations pour répondre à trois questions, présentées les unes à la suite des autres (en contre-balançant l'ordre de présentation des questions).

Ces trois questions sont les suivantes¹ :

Question A : Quelle est la principale différence entre un enregistrement analogique et un enregistrement numérique ?

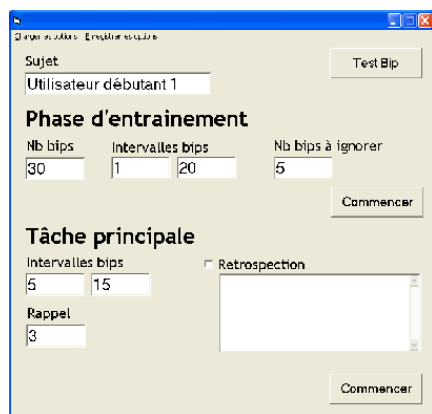
Question B : Vous souhaitez vous rendre dans le magasin "Espace Culture". Vous êtes étudiant à la fac d'Aix et sans voiture. Quels sont les bus que vous devez prendre ?

Question C : La semaine prochaine, c'est l'anniversaire de votre grand-mère. Celle-ci apprécie particulièrement la musique classique. Aussi, avez-vous décidé de lui acheter le Cd d'Hector Berlioz (avec Charles Munch et Suzanne Danco Boston). A quel prix est-il vendu sur le site d' "Espace Culture" pour les adhérents du magasin et pour les autres ?

Une fois que les utilisateurs ont pris connaissances de la question et de la tâche qu'ils doivent effectuer (chercher et trouver la réponse dans le site web), l'expérimentateur leur présente la page d'accueil du site sur lequel ils devront naviguer. La navigation (ouverture des pages, changement de pages, etc.) est enregistrée grâce au logiciel Find-Out[®] (Applied Logic Corporation) directement installé sur l'ordinateur PC utilisé par les participants. Ce logiciel permet d'identifier pour chaque utilisateur et chaque question le temps nécessaire pour y répondre (ce temps est calculé à partir du moment où le sujet navigue sur le site et jusqu'à ce qu'il indique avoir trouvé la réponse), ainsi que les pages consultées.

Pendant que les utilisateurs recherchent les informations, ils doivent réagir à des bips sonores en appuyant avec le pied sur une pédale (leurs mains restant libres pour l'utilisation de l'ordinateur). Cette dernière composante de la passation permet de recueillir des temps de réaction moteurs (principe de la "double tâche" initialement proposée par Posner & Boies, 1971), grâce au programme informatique Tholos pour PC inspiré du programme Scriptkell² (pour Macintosh) et, sur cette base de déterminer la charge cognitive mobilisée par les utilisateurs. La charge cognitive est déduite des temps de réaction pondérés – TRp – (il s'agit de TR moteurs enregistrés pendant la phase expérimentale auxquels un TR de base, propre à chaque utilisateur mesuré avant la phase expérimentale, a été soustrait) : plus le temps de réaction est long, plus le traitement cognitif en cours est considéré comme requérant d'importantes ressources cognitives³.

L'interface du programme Tholos est la suivante (cf. figure 3) :



L'interface est composée de quatre parties :

Partie 1 : pour chaque participant, on choisit le nom. On peut lui faire écouter un bip en cliquant sur le bouton "test Bip".

Partie 2 : au-dessous du nom du sujet, se trouve une partie de l'interface consacrée à la saisie des paramètres lors de la phase d'entraînement. Cette phase permet de déterminer le TR moteur moyen de base. L'expérimentateur indique le nombre de bips à présenter au sujet, l'intervalle entre la présentation de ces bips – intervalle déterminé de façon aléatoire par le programme – et le nombre de bips à ignorer dans le calcul du TR moteur moyen (ici, les cinq premiers bips). L'expérimentateur clique sur le bouton "Commencer" pour lancer l'entraînement. Une fois terminée, on revient automatiquement sur cette interface.

Partie 3 : cette partie de l'interface est consacrée à la saisie des paramètres spécifiques à la tâche principale (dans notre étude, rechercher des informations). La fonction "Enregistrer les options" (dans la barre de menu) permet de sauvegarder les paramètres, ce qui évite de devoir les saisir à chaque nouvelle passation.

Partie 4 : la partie en bas à droite intitulée "Rétrospection" est en cours de programmation. Elle n'est pas utilisée dans le cadre de l'étude présentée ici.

Une fois la tâche terminée, il y a création d'un document MS Excel dans lequel l'ensemble des données pour chaque participant est présenté.

Figure 3 – Présentation du logiciel expérimental de calcul de la charge cognitive – Tholos

¹ Afin de se conformer aux recommandations de Bastien *et al.* (1998) et de Nielsen (2000), deux clics de souris (soit ouverture de 2 pages en plus de la page d'accueil) sont nécessaires pour trouver l'information quelle que soit la version du site.

² cf. Piolat, Olive, Roussey et Thunin (1999).

³ Il est important de noter que des travaux conduits par Levy et Ransdell (1995) et Piolat et Olive (2000) ont montré que cette procédure expérimentale de recueil de données ralentit l'activité des sujets sans pour autant perturber les processus cognitifs mis en œuvre.

Étape 3 : Les utilisateurs doivent, à l'issu de la passation, naviguer librement sur le site auquel ils ont été confrontés lors de l'étape 2 et répondre ensuite à un questionnaire de satisfaction (le WAMMI – Web site Analysis and Measurement Inventory – Kirakowski, Claridge, & Whitehand, 1998).

Nous avons donc deux facteurs expérimentaux :

- Niveau d'expérience des utilisateurs : $C_2 = \{\text{utilisateurs débutants ; utilisateurs expérimentés}\}$
- Version du site web : $V_2 = \{\text{site "ergonomique" ; site "non ergonomique"}\}$

Cela permet d'obtenir le plan d'expérience suivant : $S < C_2 * V_2 >$

2.2. Objectifs

Cette étude vise à déterminer si les problèmes d'utilisabilité identifiés par les concepteurs et les utilisateurs lors de l'étude conduite précédemment (cf. Chevalier, 2004) nuisent réellement à l'activité de recherche d'informations d'utilisateurs de différents niveaux d'expérience. Plus précisément, les objectifs poursuivis sont les suivants :

- Déterminer le temps nécessaire pour trouver les informations en fonction de la version du site et de l'expérience des utilisateurs ;
- Mesurer la charge cognitive des utilisateurs selon la version du site et l'expérience des utilisateurs ;
- Évaluer le niveau de satisfaction des utilisateurs selon la version du site et l'expérience des utilisateurs.

2.3. Résultats

Tous les utilisateurs parviennent à trouver les réponses correctes aux trois questions.

2.3.1. Temps nécessaire pour trouver les informations⁴

Les analyses statistiques montrent que le temps pour trouver les réponses est plus long lorsque les utilisateurs sont confrontés au site "non ergonomique" (NE) et cela aussi bien pour les débutants que pour les expérimentés ($F(1,23)=48.325$; $p<.0001$ – cf. figure 4). On observe également un effet significatif du niveau d'expérience des utilisateurs : les débutants ont besoin de plus de temps que les expérimentés quelle que soit la version du site ($F(1,23)=11.266$; $p=.0027$).

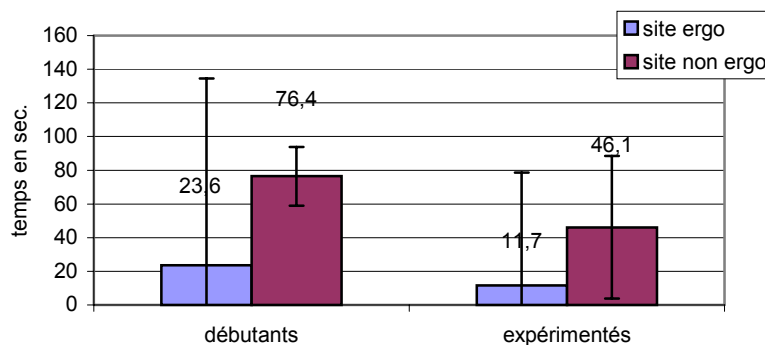


Figure 4 – Temps moyen en sec. (et écart-type) nécessaire pour répondre aux trois questions selon la version du site et le niveau d'expérience des utilisateurs

2.3.2. Charge cognitive (temps de réaction pondérés)⁴

On observe un effet significatif de la version du site sur les temps de réaction pondérés moyens (TRp) des utilisateurs (cf. figure 5). Contrairement à nos attentes, la version E engendre une charge cognitive plus importante que la version NE ($F(1,23)=4.41$; $p=.0469$), malgré une forte variabilité entre les sujets.

⁴ Nous présentons les temps (de recherche et de réaction) globaux pour les trois questions.

Par contre, on n'observe pas d'effet significatif dû au niveau d'expérience des utilisateurs ($F(1,23)=0.261$; $p>.1$).

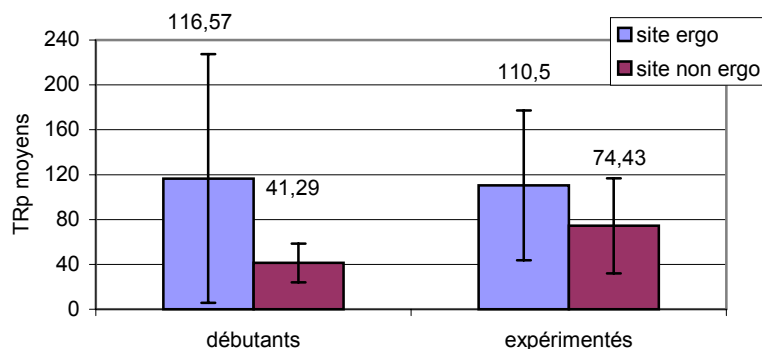


Figure 5 – Temps de réaction pondérés moyens en msec. (et écart-type) pour répondre aux trois questions selon la version du site et le niveau d'expérience des utilisateurs

2.3.3. Satisfaction vis-à-vis des deux versions du site

Les analyses statistiques mettent en évidence un effet significatif de la version du site en ce qui concerne la note attribuée (échelle en 5 points) par les utilisateurs au questionnaire de satisfaction (cf. figure 6). Le site E obtient un score significativement plus important que la version NE ($F(1,23)=7,929$; $p=.0098$). Plus précisément, pour les utilisateurs débutants, on observe un effet significatif de la version du site à $p=.0315$; pour les utilisateurs expérimentés il n'y a pas de différence significative entre les deux versions.

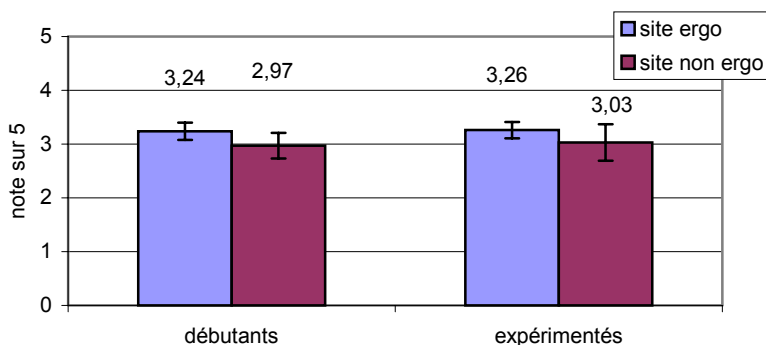


Figure 6 – Notes moyennes sur 5 points (et écart-type) au questionnaire de satisfaction selon la version du site et le niveau d'expérience des utilisateurs

3. Conclusion et perspectives

Les résultats obtenus montrent que la version du site "ergonomique" requiert des temps de recherche d'informations significativement plus courts que la version "non ergonomique" quel que soit le niveau d'expérience des utilisateurs. En ce qui concerne la charge cognitive, on note également un effet significatif, mais inverse, *i.e.* que c'est la version "ergonomique" qui requiert une charge cognitive plus importante.

Par conséquent, les problèmes d'utilisabilité que les concepteurs et utilisateurs ont identifiés dans l'étude conduite par Chevalier (2004) ralentissent certes leur activité de recherche d'informations, sans pour autant engendrer une charge cognitive plus importante. Ce résultat surprenant, mais néanmoins intéressant, pourrait s'expliquer de deux façons complémentaires l'une de l'autre :

- (1) Les utilisateurs confrontés au site "ergonomique" alloueraient davantage de ressources cognitives à leur activité de recherche d'informations que ceux confrontés à la version "non ergonomique", car ils ne seraient pas distraits par des informations non pertinentes. Ce faisant, ils trouveraient en un temps plus court les informations recherchées, en mobilisant pour cela une quantité plus importante de ressources cognitives que ceux confrontés à la version "non ergonomique".

- (2) Une explication complémentaire à la première pourrait être que les sujets seraient plus enclins à mobiliser plus de ressources cognitives lorsque le site est "ergonomique" et qu'ils économiseraient leurs ressources dans le cas contraire. En effet, lorsqu'ils sont confrontés à la version "non ergonomique", ils éprouveraient plus de difficultés à focaliser leur attention sur la tâche demandée par l'expérimentateur et, de ce fait, mobiliseraient moins de ressources cognitives.

Pour tester ces deux hypothèses explicatives, une analyse plus détaillée et davantage qualitative des protocoles est nécessaire. Cette analyse visera à déterminer (1) le nombre de pas requis pour répondre aux trois questions, et (2) quels sont les problèmes d'utilisabilité qui génèrent un temps de recherche plus long.

L'étude présentée ici utilise deux techniques expérimentales, la capture des démarches des utilisateurs et le calcul de la charge cognitive. Les résultats obtenus par ces deux techniques seront complétés par l'utilisation de la méthode des mouvements oculaires (cf. sur ce point Baccino & Colombi, 2000). Cela permettra de déterminer plus finement quels sont les problèmes d'utilisabilité qui ralentissent et distraient les utilisateurs.

En ce qui concerne le niveau de satisfaction indiqué par les utilisateurs, on note un effet significatif en faveur de la version "ergonomique" pour les utilisateurs débutants. Ces derniers jugent la version "ergonomique" comme répondant davantage à des critères de facilité d'utilisation que la version "non ergonomique". En revanche, il n'y a pas de différence significative pour les utilisateurs expérimentés due à la version du site. Aussi, peut-on avancer l'hypothèse selon laquelle les utilisateurs expérimentés, ayant davantage l'habitude de naviguer sur le Web, ont rencontré plus souvent ce type de problèmes. De ce fait, ils ont pu développer certaines procédures qui leur permettent de trouver plus rapidement que les débutants les informations recherchées, sans avoir l'impression que le site "non ergonomique" soit forcément difficile à utiliser.

Remerciements

Nous remercions Maeva Strahm (Université de Nice-Sophia Antipolis) pour nous avoir permis d'utiliser son questionnaire afin de déterminer le niveau d'expérience des internautes.

Merci également aux participants-internautes qui ont accepté de consacrer un peu de leur temps à cette recherche.

Références bibliographiques

- Baccino, T., & Colombi, T. (2000). L'analyse du mouvement des yeux sur le web. *Revue d'Intelligence Artificielle*, 14, 127-148.
- Bhatt, G. (2004). Bringing virtual reality for commercial Web sites. *International Journal of Human Computer Studies*, 60, 1-15.
- Blackmon, M. H., Polson, P. G., Kitajima, M., & Lewis, C. (2002). Cognitive Walkthrough for the Web. *CHI Letters*, 4, 463-470.
- Chevalier, A., & Ivory, M. Y. (2003). Web Site Designs: Influence of Designer's Experience and Design Constraints. *International Journal of Human-Computer Studies*, 58, 57-87.
- Chevalier, A. (2004). Concepteurs de sites web : est-il plus facile d'évaluer un site que d'en concevoir un ? *Revue d'Interaction Homme-Machine*, 5.
- Chevalier, A. (sous presse). L'orientation du cahier des charges constitue-t-elle une aide à la conception de sites web ergonomiques ? *Revue d'Intelligence Artificielle*.
- Kirakowski, J, Claridge, N., & Whitehand, R. (1998). *Human Centered Measures of Success in Web Site Design*. Paper presented at the *Human Factors and the Web Workshop*, Basking Ridge, NJ.
- Lee, A. (1999). Web usability. A review of the research. *SIGCHI Bulletin*, 31, 38-40.
- Levy, C. M., & Ransdell, S. (1995). Is writing as difficult as it seems ?. *Memory and Cognition*, 23, 767-779.
- Park, S., Choi, D., & Kim, J. (2004). Critical factors for the aesthetic fidelity of web pages: empirical studies with professional web designers and users. *Interacting with computers*, 16, 351-376.
- Piolat, A., & Olive, T. (2000). Comment étudier le coût et le déroulement de la rédaction de textes ? La méthode de la triple tâche : Un bilan méthodologique. *L'Année Psychologique*, 100, 465-502.

Posner, M. I., & Boies, S. (1971). Components of attention. *Psychological Review*, 78, 391-408.

Piolat, A., Olive, T., Roussey, J.-Y., & Thunin, O. (1999). SCRIPTKELL : an experimental tool assisted by computer for analysis cognitive effort and time processing during writing and other tasks., *Behavior Research Methods, Instruments, and Computers*, 31, 113-121

Strahm, M. (2002). *Élaboration d'un outil d'évaluation permettant la discrimination du niveau d'expertise dans le domaine de la navigation Internet*. Mémoire de Maîtrise de Psychologie sous la direction de Thierry BACCINO. Université de Nice Sophia Antipolis.

Teo, H-H, Oa, L-B, Liu, C, & Wei, K-K. (2003). An empirical study of the effects of interactivity on Web user attitude. *International Journal of Human-Computer Studies*, 58, 281-305.