

CHERCHER DES INFORMATIONS DANS LES MENUS WEB : INTERACTION ENTRE TÂCHE, TYPE DE MENU ET VARIABLES INDIVIDUELLES

Jean-François Rouet, C. Ros, G. Jégou, S. Metta

Presses Universitaires de France | « [Le travail humain](#) »

2004/4 Vol. 67 | pages 377 à 395

ISSN 1243-1370

ISBN 2130548067

Article disponible en ligne à l'adresse :

<http://www.cairn.info/revue-le-travail-humain-2004-4-page-377.htm>

Pour citer cet article :

Jean-François Rouet *et al.*, « Chercher des informations dans les menus web :
interaction entre tâche, type de menu et variables individuelles », *Le travail humain*
2004/4 (Vol. 67), p. 377-395.
DOI 10.3917/th.674.0377

Distribution électronique Cairn.info pour Presses Universitaires de France.

© Presses Universitaires de France. Tous droits réservés pour tous pays.

La reproduction ou représentation de cet article, notamment par photocopie, n'est autorisée que dans les limites des conditions générales d'utilisation du site ou, le cas échéant, des conditions générales de la licence souscrite par votre établissement. Toute autre reproduction ou représentation, en tout ou partie, sous quelque forme et de quelque manière que ce soit, est interdite sauf accord préalable et écrit de l'éditeur, en dehors des cas prévus par la législation en vigueur en France. Il est précisé que son stockage dans une base de données est également interdit.

RECHERCHES EMPIRIQUES EMPIRICAL STUDIES

CHERCHER DES INFORMATIONS DANS LES MENUS WEB : INTERACTION ENTRE TÂCHE, TYPE DE MENU ET VARIABLES INDIVIDUELLES

par J.-F. ROUET*, C. ROS*, G. JÉGOU* et S. METTA*

SUMMARY

LOCATING RELEVANT CATEGORIES IN WEB MENUS : EFFECTS OF TASK COMPLEXITY, MENU STRUCTURE AND INDIVIDUAL FACTORS

Most Web portals and other online services use menu structures to present their contents. Given the increasing use of the Web for everyday activities, it is important to assess the design strategies of web menus and their adequacy with the user's needs. The study presented here was carried out within the framework of a longitudinal study over two years and related to the uses of the Internet. Our objective was to examine the influence of individual factors and three different menu designs on the effectiveness of an information search task.

The participants were fifty voluntary adults, representative of the population of a mid-size French town. They were divided into three different age groups (younger, intermediate and older adults). They were asked to perform a series of twelve search tasks using three types of hierarchical web menus. Menu structure (deep, broad-alphabetic and broad-categorized), and question length and implicitness were manipulated in a within-subject design. The subjects' choices for each question, search time and success were recorded via a web database.

Older adults tended to show poorer performance, especially with deep menus. Longer and implicit questions also led to poorer performance for the whole group of participants. Finally, a significant relation was observed between subjects' visuo-spatial capacity and their performance with search tasks using explicit key words.

This experiment provided evidence that locating relevant categories in a hierarchical menu structure is far from a trivial task for general public adult. We discuss possible implications of these results for individual and task-related factors of information search performance.

Key words : *Web, Menu Structure, Search Information Task, Questions Complexity, Age.*

* CNRS FRE 2725, Laboratoire Langage et cognition (LaCo), 99, avenue du Recteur Pineau, F-86000 Poitiers. E-mail : Jean-Francois.Rouet.@univ-poitiers.fr.

I. INTRODUCTION

La plupart des portails commerciaux sur Internet et de nombreux services « *online* » présentent des menus (ou listes d'items) permettant l'accès à quantité de services et de pages d'informations. Ces menus sont en général constitués de mots ou d'expressions thématiques qui renvoient à d'autres menus plus spécifiques, et ainsi de suite jusqu'à trouver la catégorie d'informations désirée. Ces dispositifs permettent, du moins en théorie, à tous types de public d'accéder aux informations correspondant à leurs besoins ou à leurs centres d'intérêt.

La recherche d'informations dans les menus pose des problèmes importants sur le plan théorique comme sur le plan pratique. Sur le plan théorique, la recherche d'informations est une activité encore mal connue, dont les liens avec le langage, la mémoire et le raisonnement restent obscurs. Sur le plan pratique, il semble que les concepteurs de sites web aient quelques difficultés à définir les structures optimales compte tenu du public et du domaine concernés. Certains portails proposent des menus dits « profonds » se caractérisant par un petit nombre de catégories accessibles à chaque niveau de sélection ; et d'autres offrent des menus dits « larges » où un grand nombre de catégories est affiché à chaque niveau. Dans certains cas, les items des menus sont listés alphabétiquement ; dans d'autres, ils sont groupés sémantiquement (par exemple, « finances », « voyages »), avec ou sans intitulé pour chaque catégorie. De plus, l'organisation des portails varie fréquemment dans le temps, ce qui renforce le sentiment qu'il s'agit d'un domaine de conception encore largement dominé par l'empirisme.

Étant donné l'utilisation croissante d'Internet dans les activités quotidiennes, il est important de déterminer l'effet des choix de conception des menus sur les performances d'utilisateurs au cours de la recherche d'informations. Les recherches d'ergonomie menées depuis vingt ans apportent sur ce sujet des informations pertinentes, quant à l'effet de l'organisation et de la structuration des menus sur la performance en sélection. Cependant, il est nécessaire de prendre en compte les caractéristiques individuelles des utilisateurs (âge, expérience avec la technologie, niveau d'études) qui sont susceptibles d'interagir avec certains choix de conception.

II. POSITION DU PROBLÈME

II.1. STRUCTURE DES MENUS, PERFORMANCES ET PRÉFÉRENCES DES UTILISATEURS

L'effet de la structure des menus a fait l'objet d'un grand nombre d'études expérimentales depuis l'apparition des bases de données interactives. L'étude de Miller (1981) a amorcé une série d'expériences sur le

compromis largeur/profondeur dans les menus hiérarchisés. Miller demandait à des sujets adultes de localiser un mot cible (par exemple, « magnétisme », « Espagne ») parmi 64 mots ou expressions présentés sous forme de listes. Ces listes étaient présentées selon différentes organisations hiérarchiques. Dans la version la plus « large », les 64 items apparaissaient sur un seul écran. Dans la version la plus « profonde », les mots étaient organisés hiérarchiquement selon un arbre binaire à 6 niveaux de choix (NB. $2^6 = 64$). L'accès aux mots cibles se faisait alors via des noms de catégories (par exemple, « culture > géographie > pays > Europe > occidentale > Espagne »). Deux versions intermédiaires présentaient respectivement 2 niveaux de choix parmi 8 items et 3 niveaux de choix parmi 4 items. La présentation des items à choisir sur chacun des écrans ne suivait aucune logique particulière. Miller a montré que les temps de sélection étaient plus courts lorsque le menu était organisé de façon intermédiaire (8^2 ou 4^3). Snowberry, Parkinson et Sisson (1983), reprenant le matériel de Miller (1981), ont montré que, dans la condition la plus large (64 items sur un seul écran), les performances étaient significativement améliorées en regroupant les items par catégorie. Par ailleurs, les performances étaient également supérieures dans la condition de largeur maximale. Finalement, Snowberry *et al.* (1983) ont observé une relation entre les capacités visuo-spatiales des sujets et le temps de localisation des cibles. Contrairement à leurs attentes, cependant, cette relation n'était pas spécifique au cas des menus larges, supposés plus exigeants en matière d'attention visuelle (*visual scanning*). D'autres recherches menées par la suite ont confirmé que le processus de sélection dans un menu peut être facilité si les items sont groupés dans des catégories sémantiques (Giroux, Bergeron, & Lamarche, 1987) et/ou si le ratio profondeur/largeur est optimal (Parkinson, Sisson, & Snowberry, 1985).

Entre-temps, l'avènement du World Wide Web a consacré l'utilisation des menus et des liens hypertextes comme moyen d'accès aux banques d'informations en ligne. L'évolution des techniques et leur large diffusion ont suscité un regain d'intérêt pour l'ergonomie des menus, notamment dans le cadre de sites web ou de bases de données multimédias. Dans une étude avec 19 adultes volontaires, utilisateurs expérimentés de l'ordinateur et du Web, Larson et Czerwinsky (1998) ont manipulé la structure des menus dans une base d'informations issues de l'encyclopédie électronique Microsoft Encarta®. Ils ont à leur tour observé que des menus à deux niveaux (de type 16×32 items ou 32×16 items) entraînaient des recherches plus rapides et un moindre sentiment de désorientation qu'un menu plus profond ($8 \times 8 \times 8$ items, soit trois niveaux). L'organisation 32×16 était celle qui recueillait généralement la préférence des sujets.

Plus récemment, Yu et Roh (2002) ont comparé les performances et l'évaluation d'étudiants à travers trois menus différents d'un centre commercial virtuel. Ils ont montré qu'un menu déroulant qui maximisait la visibilité des catégories intermédiaires était le plus performant en termes de temps de recherche, que ce soit pour des tâches spécifiques ou pour des tâches plus globales. Cependant, les évaluations des étudiants, portant sur la qualité de conception et sur leur sentiment de désorientation, ne

variaient pas selon le type de menus. Zaphiris, Shneiderman et Norman (1999) ont répliqué pour leur part les résultats précédemment mis en évidence quant à la supériorité des menus larges sur les menus profonds. Cependant, contrairement aux résultats de Yu et Roh (2002), ils ont montré que la recherche d'informations était généralement plus rapide avec un menu traditionnel qu'avec un menu déroulant, et cela plus particulièrement lorsque la structure des menus est plus profonde. De façon générale, les participants à leur étude (21 étudiants) n'ont montré aucune préférence vis-à-vis des menus proposés. Pourtant, quand la structure des menus était la plus profonde, ils ont préféré le menu séquentiel plutôt que le menu déroulant. Zaphiris et ses collaborateurs ont alors conclu que les menus déroulants étaient acceptables uniquement quand la structure des menus est peu profonde.

Il faut retenir de ces études que, s'il existe aujourd'hui un consensus général en faveur des menus larges et structurés, il n'y a pas totalement convergence entre les résultats expérimentaux. Par ailleurs, la plupart des expériences rapportées dans la littérature ont été réalisées dans des conditions de laboratoire, avec de simples tâches de localisation qui ne reflètent pas la diversité des tâches de recherche d'informations (Rouet & Tricot, 1998). Par ailleurs, ces expériences font généralement appel à des étudiants et/ou des utilisateurs expérimentés du Web. Or les services web sont de plus en plus utilisés par un public large, plus hétérogène en termes d'âge, de niveau d'éducation et d'expérience dans l'usage des ordinateurs. Les possibilités de généralisation des résultats aux menus des services internet pour le grand public sont donc limitées.

II.2. COMPLEXITÉ DE LA TÂCHE ET CARACTÉRISTIQUES DES UTILISATEURS

Les recherches sur l'organisation des menus ont pour la plupart utilisé de simples tâches de sélection, comme, par exemple, localiser le mot « Espagne » dans un arbre hiérarchique. Cependant, les tâches de recherche effectuées par les utilisateurs d'informations en ligne prennent parfois une tournure plus complexe. En effet, lorsqu'on effectue une recherche sur Internet pour des besoins réels de la vie de tous les jours, le contenu de la question initiale peut être implicite et/ou ambigu, et peut donc ne pas correspondre directement aux libellés des items présents dans les menus. Par exemple, l'objectif de la recherche peut être exprimé de façon globale (« chercher un hôtel à Paris ») ou plus précise (« chercher une chambre simple dans un hôtel de luxe près du Panthéon à Paris »). Selon le modèle proposé par Rouet et Tricot (1998), la représentation mentale de l'objectif joue un rôle important dans la recherche d'informations. En effet, cette activité prend la forme d'un cycle de processus qui inclut l'évaluation, la sélection et le traitement des catégories ou des contenus sélectionnés. De la représentation plus ou moins claire et explicite de l'objectif dépend la capacité du sujet à estimer la pertinence des différentes catégories proposées (sélection) et l'adéquation de l'information de contenu à la tâche (traitement). Or la représentation mentale de

l'objectif peut varier selon le type de question reçue (ou formulée) par le sujet. Selon les cas, la phase d'évaluation nécessite donc la production d'inférences afin, par exemple, de spécifier un objectif flou ou d'établir des appariements entre les termes de la question et ceux des menus.

Plusieurs expériences ont montré l'influence de la représentation du but dans la réalisation de tâches de recherche d'informations. Weyer (1982) et Rouet (1994) ont montré que, lorsque la question était formulée dans des termes généraux, les sujets (élèves de lycée) avaient tendance à s'arrêter prématurément, dès les premières informations pertinentes rencontrées, sans approfondir la recherche. Dans une expérience similaire impliquant des étudiants de troisième cycle universitaire, Rouet (2003 *a*) a montré que la précision des choix effectués dans un menu hiérarchique diminuait avec la complexité des questions (questions correspondant à plusieurs catégories cibles). Les questions complexes entraînaient par ailleurs une fréquence plus élevée de relecture en cours de recherche, ce qui suggère un problème de charge en mémoire de travail. La représentation mentale de l'objectif semble donc bien jouer le rôle d'une structure de contrôle. Cependant, l'influence respective de la formulation initiale de l'objectif de recherche, d'une part, et de l'organisation des menus, d'autre part, reste encore à déterminer.

Par ailleurs, dans la majorité des recherches menées jusqu'ici, les participants étaient des étudiants. Les étudiants constituent une population relativement homogène sur le plan de l'âge et d'un niveau d'éducation plus élevé en moyenne que le reste de la population. Or l'on sait que la performance des individus dans des tâches cognitives, y compris la recherche d'informations, diminue avec l'âge (Freudenthal, 2001 ; Westerman, Davies, Glendon, Stammers, & Matthews, 1995). Les personnes vieillissantes semblent surtout désavantagées lors de l'utilisation de hiérarchies profondes, peut-être en partie à cause de l'involution de certaines capacités psychomotrices et cognitives (Freudenthal, 2001). Plus précisément, les résultats suggèrent que le vieillissement entraîne une diminution des capacités visuo-spatiales mobilisées dans l'apprentissage d'environnements nouveaux. Il faut cependant noter avec Marquié (2001) que le vieillissement est un phénomène multidimensionnel (physique, cognitif, affectif, social, etc.), et que chacune des dimensions impliquées peut avoir une influence sur le comportement des individus dans une situation donnée. Par exemple, les personnes vieillissantes développent parfois un sentiment de compétence plus faible à l'égard des nouvelles technologies (Marquié, Jourdan, & Huet, 2002). L'ensemble de ces paramètres est susceptible d'influencer le comportement des personnes dans les tâches de recherche d'informations. Il reste donc à déterminer comment les utilisateurs jeunes ou plus âgés réagissent face à différentes structures de menus, et si des variations individuelles de capacités cognitives (mémoire de travail, notamment) peuvent expliquer ces différences.

III. CADRE DE L'ÉTUDE

La recherche présentée ici a été mise en place dans le contexte d'une étude longitudinale de deux ans portant sur les attentes, les connaissances et les usages de 100 utilisateurs d'Internet (Andréo, Jégou, & Rouet, 2001). Cette étude, réalisée en partenariat avec la Communauté d'agglomération de Poitiers, visait à préfigurer l'arrivée des réseaux numériques à haut débit et à mieux connaître les besoins et les attentes du public dans ce domaine.

Pour la réalisation de cette étude, un panel de 100 individus représentatif de la population de plus de 18 ans (hors étudiants et militaires) a été constitué à partir d'un appel à candidatures. L'échantillon a été constitué selon la méthode des quotas avec l'âge, le sexe et la catégorie socioprofessionnelle comme variables principales. Pour être éligibles, les candidats devaient être novices dans l'utilisation de l'informatique et d'Internet et ne pas déjà posséder d'ordinateur chez eux. Quelques-uns avaient eu une expérience limitée de l'usage d'Internet dans le cadre de leur travail ou de leurs relations personnelles. Chaque participant a été équipé d'un PC et d'un forfait de connexion à Internet de 30 heures par mois. En échange, les participants versaient une cotisation de 100 F par mois et s'engageaient à coopérer avec les enquêteurs.

Les objectifs, la méthode et les résultats d'ensemble de l'étude ont été présentés dans un rapport de synthèse (Rouet, 2003 *b*) et ne seront pas détaillés ici.

L'expérience présentée dans cet article s'est déroulée au cours de la seconde année de l'étude, c'est-à-dire une fois passée la période d'apprentissage des outils. Son objectif principal était d'examiner les facteurs influençant la performance des utilisateurs d'Internet à des tâches de recherche d'informations au moyen de menus web hiérarchiques. Nous avons étudié l'impact de l'organisation des menus et de la complexité des questions de recherche sur les capacités des utilisateurs à localiser des items cibles.

Nos principales hypothèses étaient les suivantes : tout d'abord, nous pensions que la recherche serait plus lente et difficile dans des menus profonds, lorsque la question comporte un plus grand nombre de mots de contenu, ou lorsqu'une inférence est nécessaire pour appairer la question et la catégorie cible. Par ailleurs, nous attendions une diminution des performances avec l'âge des participants. Enfin, nous pensions que cette baisse serait corrélative d'une involution des capacités visuo-spatiales et qu'elle serait plus marquée dans le cas des menus profonds.

IV. MÉTHODE

IV.1. PARTICIPANTS

Cinquante-deux adultes volontaires issus du panel de 100 utilisateurs d'Internet présenté ci-dessus ont participé à cette expérience. Les données de 2 participants ont été retirées car ils ne se sont pas conformés aux exigences de la tâche expérimentale.

Les participants ont été classés en trois groupes d'âge d'effectifs homogènes. L'échantillon final ($N = 50$) comportait donc 9 hommes et 7 femmes âgés de 24 à 36 ans (groupe des plus jeunes ; $n = 16$) ; 5 hommes et 14 femmes âgés de 37 à 53 ans (groupe d'âge intermédiaire ; $n = 19$) ; enfin, 8 hommes et 7 femmes âgés de 54 à 80 ans (groupe des plus âgés ; $n = 15$). Tous les participants étaient des utilisateurs réguliers d'Internet depuis dix-huit mois au moment de l'expérience.

IV.2. MATÉRIEL

À partir de portails web existants, nous avons constitué une structure de menus de 400 items. Cette structure présentait une hiérarchie de catégories et de sous-catégories d'intérêt général (ex. : éducation, voyage, travail, etc.). À partir des 400 items, nous avons développé 3 versions de menus : un menu large catégorisé, un menu large organisé alphabétiquement et un menu profond.

Les versions de menus larges offraient un nombre d'items important sur chaque page (c'est-à-dire avec un maximum de 42 items par page), mais seulement 2 niveaux de sélection (menu principal et sous-menus). La version de menu profond présentait pour sa part un faible nombre d'items par page (6 items maximum par page), mais 4 niveaux de sélection. Dans la version de menu large catégorisé, les items étaient regroupés dans des catégories sémantiques, alors qu'ils étaient classés alphabétiquement dans les 2 autres versions de menus (fig. 1). Les 3 versions de menus développés présentaient le même nombre et les mêmes libellés de catégories et répondaient à une même charte graphique (typologie, couleur, taille, etc.).

Nous avons construit 12 tâches de recherche d'informations permettant de chercher des catégories cibles dans les 3 versions de menus. Pour ce faire, nous avons utilisé pour chacune des questions de recherche de catégories, les mots constituant le chemin d'accès à la catégorie cible. Par exemple, la question « trouver un *hébergement* dans un hôtel de type *Palace* » a été rédigée à partir de la hiérarchie d'items « tourisme > hébergement > hôtel > palace ». Dans la version originale, les termes de la question correspondaient donc littéralement aux termes des menus.

Chaque question a été présentée selon 4 versions de complexité différente. Le premier facteur de complexité des questions correspondait au caractère littéral ou inférentiel de la question. Pour construire le contenu

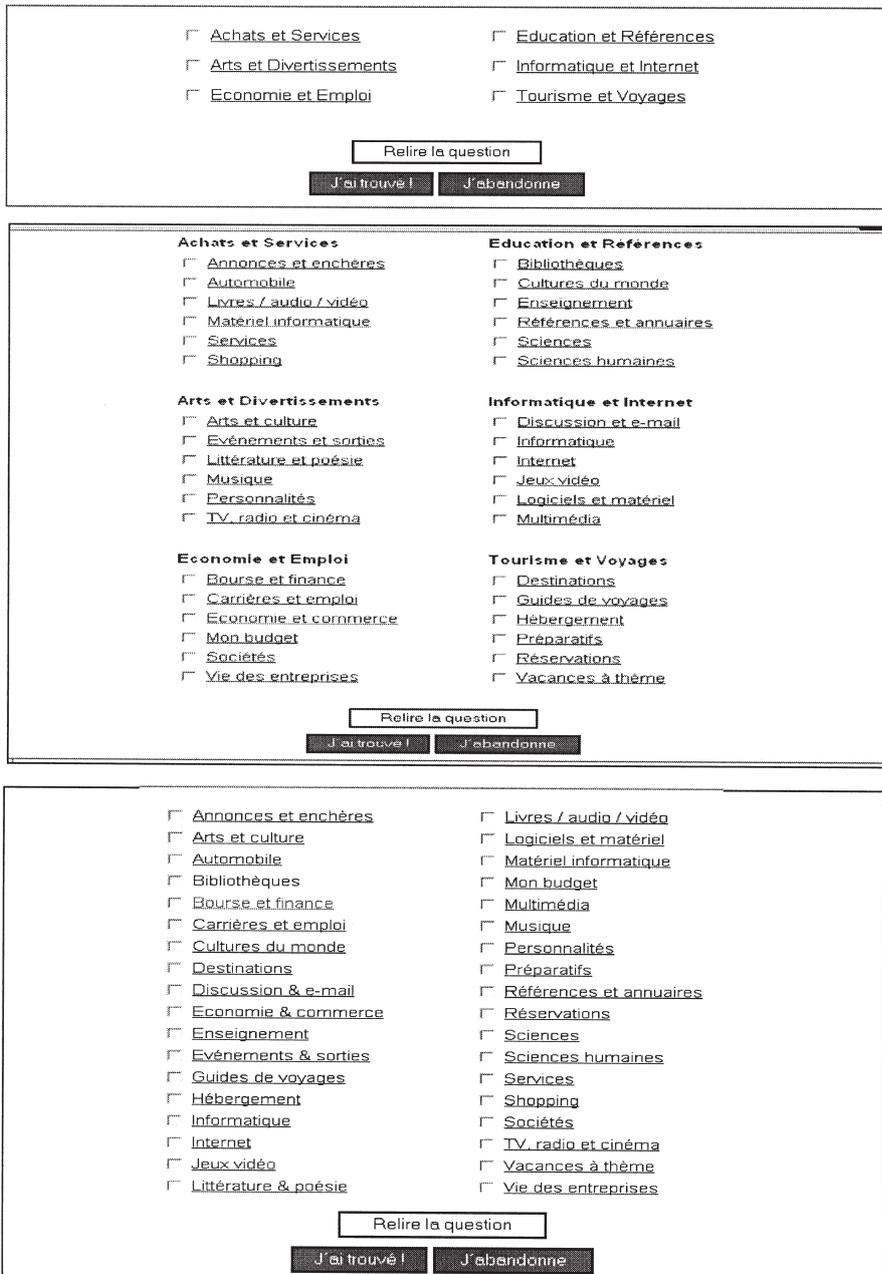


Fig. 1. — Les trois types de menus utilisés dans l'expérience.

En haut, menu profond ; au milieu, menu large catégorisé ; en bas, menu large alphabétique

The three types of menus used in the experiment.

At the top, deep menu ; in the middle, broad-categorized menu ; at the bottom, broad-alphabetic menu

des questions de type inférentiel, les mots de la question originale ont été remplacés par des synonymes, par exemple : « *location d'une chambre dans un hôtel 5 étoiles* ». Ainsi, le sens de la question restait le même, mais le sujet devait produire une inférence pour trouver la correspondance entre la question et les items du menu. Le second facteur de complexité des questions était la longueur de la question. Pour chaque version de questions (littérale et inférentielle), nous avons construit des questions longues en ajoutant 2 à 3 mots à la question initiale (ex. : « trouver un hébergement *occasionnel* dans un *bel* hôtel *parisien* de type *Palace* »). Afin de conserver la signification initiale de la question, les mots ajoutés dans les questions « longues » ne renvoyaient à aucune étiquette de catégorie des menus. Les questions « courtes » étaient composées de 2 à 4 mots de contenu, alors que les questions « longues » comportaient, pour leur part, 5 à 8 mots. Chaque question ne correspondait qu'à une seule sous-catégorie cible des menus.

Nous avons développé un serveur permettant de contrôler l'accès des participants, et de délivrer à chacun d'entre eux l'ensemble des questions et des 3 versions de menus en respectant un contrebalancement expérimental complet des 3 variables indépendantes manipulées. Dans cette expérience, les facteurs versions de menus web, caractère littéral ou inférentiel des questions et longueur des questions, étaient des facteurs intra-groupes. Lors de la passation de l'expérience, le serveur enregistrait le « journal » chronologique des choix faits par le participant. Ces enregistrements ont été utilisés pour extraire les variables dépendantes analysées dans cette étude.

IV.3. PROCÉDURE

Les participants ont été rencontrés à leur domicile sur rendez-vous par l'un des enquêteurs. Ils étaient tout d'abord interviewés sur leurs connaissances d'Internet à partir d'un protocole standardisé. Ensuite, il leur était demandé de réaliser les tâches de recherche de catégories à partir de la liste des 12 questions et en utilisant les 3 versions de menus web. L'équipement informatique de chaque participant était équivalent. Il était constitué d'un ordinateur standard de type PC muni d'un écran de 14 pouces. La connexion au serveur était réalisée via un modem standard de 56 K, avec les fluctuations de débit inhérentes à ce type de technologie.

L'expérimentateur assistait les participants pour la connexion et l'accès au serveur distant. Puis, il leur donnait la consigne standardisée pour la réalisation des tâches de recherche de catégories. Après une courte séance de familiarisation avec l'interface, les participants ont réalisé les 12 tâches de recherche de catégories en utilisant les 3 types de menus web, soit 4 questions de complexité différente par version de menu web. Pour chaque question, les participants devaient trouver et désigner la catégorie cible en utilisant une case à cocher. À tout moment, ils pouvaient relire la question et pouvaient également abandonner la question et passer à la question suivante en appuyant sur un bouton « J'abandonne ». Les participants étaient encouragés à chercher et trouver la catégorie cible

de la façon la plus précise possible. Cependant, si leur recherche était infructueuse après 3 minutes, l'expérimentateur leur demandait d'abandonner et de passer à la question suivante. Pour chaque participant, les catégories sélectionnées et les temps de sélection ont été enregistrés via le serveur web.

Par ailleurs, les participants effectuaient une épreuve d'empan visuo-spatial basée sur la tâche des Corsi Blocks (Gyselinck, Ehrlich, Cornoldi, De Beni, & Dubois, 2000). Dans cette tâche, le sujet doit observer une matrice de 4×4 cases. Une succession aléatoire de cases est brièvement obscurcie et le participant doit reproduire la suite des cases obscurcies dans l'ordre. Cette tâche présente l'avantage d'être indépendante du niveau verbal des sujets.

En nous appuyant sur les résultats des recherches antérieures, nous avons posé l'hypothèse que les sujets les plus âgés devaient être plus lents et moins précis que les 2 autres groupes d'âge. De plus, nous avons supposé que la difficulté à trouver la catégorie cible devait augmenter lors de l'utilisation du menu profond et en présence de questions complexes. Enfin, nous prédisions un effet de l'âge plus marqué dans le cas des menus profonds et une relation positive entre la capacité visuo-spatiale et la performance des sujets, cela à tous les niveaux d'âge.

V. RÉSULTATS

La réussite aux tâches de recherche de catégories cibles, le temps de recherche et le nombre de relectures des questions ont été utilisés comme variables dépendantes. Les données ont été analysées à partir d'une analyse multifactorielle. Dans cette analyse, le facteur groupe d'âge était un facteur intergroupe (à 3 niveaux), et les facteurs versions de menus web, caractère littéral ou inférentiel des questions et longueur des questions, étaient des facteurs intragroupes.

V.1. RÉUSSITE AUX TÂCHES DE RECHERCHE DE CATÉGORIES CIBLES

Le taux moyen de réussite des participants était de 52,5 %. Dans 9 % des autres cas, les participants ont sélectionné une autre catégorie que la catégorie cible attendue, mais leur réponse se situait dans le sous-menu correct. Enfin, dans 38,5 % des cas, ils ont choisi une catégorie incorrecte ou n'ont pas fourni de réponse. Le tableau 1 ci-dessous présente les pourcentages moyens de participants qui sont parvenus à localiser les catégories appropriées en fonction du groupe d'âge, du type de menu web, de la formulation littérale ou inférentielle de la question et de la longueur de la question.

Plusieurs tendances peuvent être observées à partir du tableau 1 (s'agissant de fréquences observées sur de petits échantillons, nous n'en proposons pas d'analyse statistique). Tout d'abord, les performances diminuaient légèrement avec l'âge (55 %, 54 % et 48 % pour les

TABLEAU 1

Pourcentages moyens de réussite en fonction du groupe d'âge, du type de menu web, de la formulation littérale ou inférentielle de la question et de la longueur de la question (C = Courte, L = Longue)

Percentage of correct answers as a function of age group and, menu structure (Profond = Deep ; Large catégorisé = Broad-Categorized ; Large Alphabétique = Broad-Alphabetic) ; question length (c = Short, L = Long) ; and Complexity (Littérale = Literal ; Inférentielle = Inferential)

Type de Menu web	Profond				Large Catégorisé				Large Alphabétique				Moy.
	Litté-rale		Infé-ren-tielle		Litté-rale		Infé-ren-tielle		Litté-rale		Infé-ren-tielle		
Âge	C	L	C	L	C	L	C	L	C	L	C	L	
24-36 (n = 16)	69	44	69	44	88	75	38	44	75	50	25	44	55
37-53 (n = 19)	74	58	53	47	68	68	58	32	63	53	42	26	54
54-80 (n = 15)	40	47	40	13	67	60	40	47	60	60	33	73	48
Moyenne	62	50	54	36	74	68	46	40	66	54	34	46	52,5

groupes 1, 2 et 3, respectivement). Deuxièmement, la version de menu large catégorisé a entraîné un pourcentage moyen de réussite de 57 %, ce résultat est légèrement supérieur à ceux des versions de menu profond et de menu large alphabétique (50 % de réussite dans les 2 cas). Troisièmement, les facteurs formulation littérale ou inférentielle de la question et longueur de la question ont eu un impact sur la réussite aux tâches de recherche de catégories cibles. En effet, les questions courtes et littérales ont été réussies dans 67 % des cas ; à l'opposé, les questions longues et littérales, les questions courtes et inférentielles et les questions longues et inférentielles ont obtenu des pourcentages de réussite plus faibles, soit respectivement 57 %, 45 % et 41 %.

Dans le but de vérifier s'il existait une interaction entre les facteurs groupe d'âge et type de menu web, nous avons réalisé une ANOVA à 2 facteurs (avec le groupe d'âge comme facteur intergroupe et le facteur type de menu web comme facteur intragroupe) sur le pourcentage de questions correctement ciblées par chacun des participants et pour chaque type de menu, tous types de questions confondus¹.

Ni l'âge ni le type de menu n'ont d'effet principal significatif sur la performance en recherche d'informations. En revanche, on observe sur la figure 2 une tendance à l'interaction entre ces deux facteurs. Conformé-

1. Le nombre de questions par condition n'a pas permis de faire d'analyses statistiques plus précises en termes d'interaction entre les facteurs.

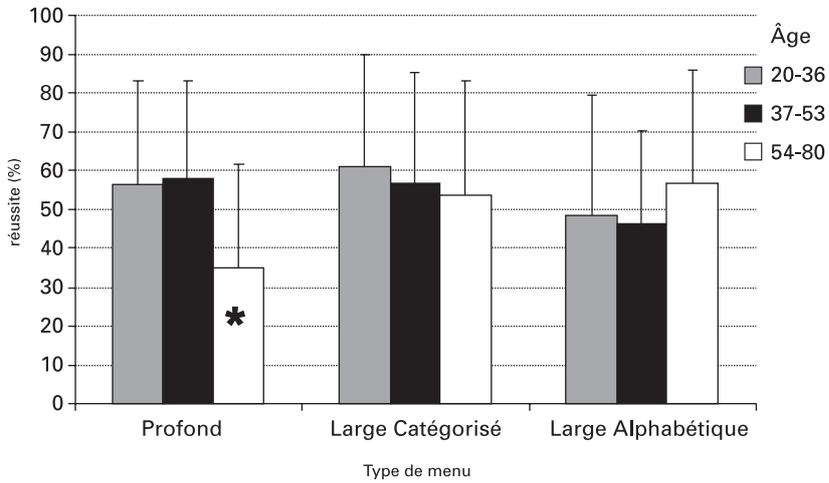


Fig. 2. — Pourcentages moyens de réussite en fonction du groupe d'âge et du type de menu web

Percentage of correct answers as a function of age group and menu structure (Profond = Deep ; Large Catégorisé = Broad-Categorized ; Large Alphabétique = Broad-Alphabetic)

ment à nos attentes, le pourcentage de réussite était le plus faible pour les participants les plus âgés en condition menu profond. L'analyse de la variance ne confirme cependant pas cette conclusion, du fait de la très grande variation intragroupe observable sur la figure 2 ($F(4;94) = 1,81$; NS). Une analyse séparée effectuée sur la condition « menu profond » indique toutefois que les performances des participants âgés étaient significativement inférieures à celles des autres groupes ($F(2;47) = 3,83$; $p < .05$). Dans cette condition, les pourcentages moyens de réussite des groupes d'âge 1, 2 et 3 étaient respectivement de 56,3 %, 57,9 % et 35 %. Dans les deux autres conditions de menu web, le groupe d'âge d'appartenance n'a pas eu d'effet significatif sur les performances des participants.

Afin d'examiner plus précisément notre hypothèse d'un effet négatif du vieillissement sur la performance dans les menus profonds, et considérant l'importante variance des âges dans le groupe 3 (54-80 ans), nous avons subdivisé ce groupe en deux sous-groupes d'effectifs comparables : d'une part, les sujets âgés de 54 à 63 ans ($n = 8$) ; d'autre part, les sujets âgés de 64 à 80 ans ($n = 7$). Nous avons calculé le pourcentage moyen de réussite selon le type de menu dans chacun de ces deux sous-groupes. Nous avons, cette fois, observé un contraste accentué entre le groupe d'âge et le type de menu. La performance des 64-80 ans dans le menu profond n'était que de 17,9 %, contre 50 % chez les 53-64 ans (soit une performance comparable à celle des autres groupes). En revanche, et de manière étonnante, la performance des sujets de 64 à 80 ans dépassait celle des 54-63 ans pour le menu large alphabétique (75 % contre 40,6 %, respectivement). L'interaction entre âge et type de menu était cette fois

significative ($F(2;26) = 7,85 ; p < .01$). Cela confirme donc un effet différentiel des types de menus selon l'âge des participants, mais les limites chronologiques de cet effet ainsi que les mécanismes qui les sous-tendent restent à préciser.

V.2. TEMPS DE RECHERCHE DE CATÉGORIES CIBLES

Il faut tout d'abord noter que la précision des temps enregistrés a souffert des variations de débit inhérentes à la connexion par modem. Pour chaque participant, nous avons calculé la médiane des temps de recherche pour les essais réussis. Compte tenu des performances relativement faibles, nous avons dû réduire le plan en regroupant tous les types d'essais (voir tableau 2). Les données d'un participant du groupe des 53-80 ans ont été retirées de l'analyse car l'enregistrement était incomplet.

TABLEAU 2

*Temps médians de recherche (en secondes)
pour les réponses correctes seulement et en fonction du groupe d'âge*
Median search time (in seconds) for correct answers only
and according to age group

Âge	24-36	37-53	54-80
<i>N</i>	16	19	14
Moyenne (<i>écart type</i>)	42,59 (14,85)	48,87 (14,39)	64,57 (28,20)

À partir de ces données, nous avons réalisé une ANOVA avec le groupe d'âge comme facteur intergroupe. Les résultats statistiques ont confirmé l'augmentation des temps de recherche des catégories avec l'âge ($F(2;46) = 5,01 ; p < .05$). Les participants les plus jeunes et ceux d'âge intermédiaire ne diffèrent pas entre eux, mais ont été bien plus rapides que les participants les plus âgés. Il faut également noter l'importante variance au sein de ce dernier groupe. La comparaison des deux sous-groupes de participants âgés définis dans la section précédente n'a cependant montré aucune différence.

V.3. NOMBRE DE RELECTURES DES QUESTIONS

La figure 3 ci-dessous indique les pourcentages moyens de relectures des questions en fonction du groupe d'âge et du type de menu web.

Une ANOVA à deux facteurs a mis en évidence un effet principal du type de menu web. En effet, les relectures des questions ont été moins fréquentes en condition de menu large catégorisé que dans les deux autres conditions de menus web ($F(2;94) = 3,72 ; p < .05$). De plus, les partici-

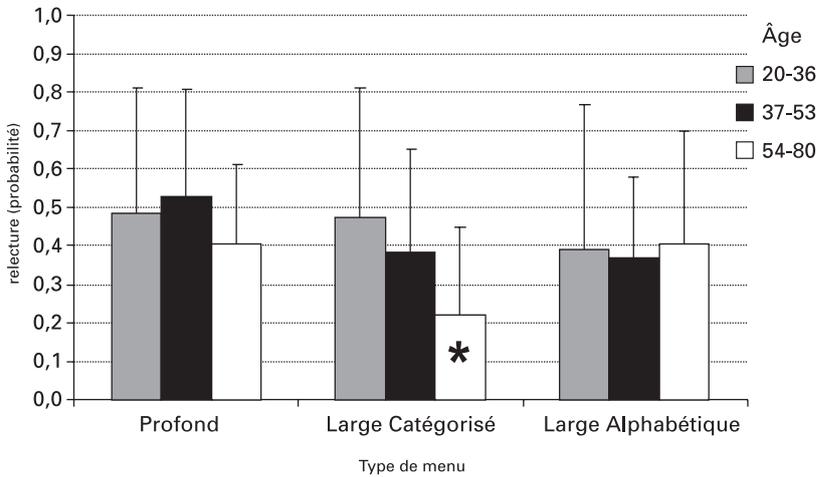


Fig. 3. — Fréquence de relecture des questions en fonction du groupe d'âge et du type de menu web

Frequency of re-reading of the questions according to age group and menu structure (Profond = Deep ; Large Catégorisé = Broad-Categorized ; Large Alphabétique = Broad-Alphabetic)

pants les plus âgés ont eu tendance à effectuer moins de relectures des questions que les deux autres groupes d'âge, et cela plus particulièrement en condition de menu profond et de menu large catégorisé. Une nouvelle fois, l'analyse statistique ne confirme pas ce résultat ($F(4;94) = 1,74$; NS). Une analyse séparée sur le groupe de participants âgés indique cependant que les participants ont relu moins souvent les questions en condition de menu large catégorisé que lors de l'utilisation des autres versions de menus ($F(2;28) = 3,90$; $p < .05$). De façon générale, les questions inférentielles les plus longues ont été relues plus souvent que les questions courtes et littérales (29,3 % de relectures pour les questions courtes littérales ; 38,7 % pour les questions courtes inférentielles ; 44,7 % pour les questions longues littérales ; et 50 % pour les questions longues inférentielles). Ce profil de résultats suggère un impact direct des facteurs linguistiques sur la difficulté de se rappeler le contenu de la question au cours de la recherche de catégories dans des menus web.

V.4. EMPAN VISUO-SPATIAL

L'analyse des scores à l'épreuve visuo-spatiale a montré une baisse significative des performances dans le groupe des plus âgés avec une moyenne de 2,87 contre 3,75 et 3,74 chez les participants jeunes et d'âge intermédiaire respectivement ($F(2;47) = 6,64$; $p < .01$). Les groupes jeunes et intermédiaires ne se différenciaient pas entre eux ($F < 1$) mais différaient significativement des plus âgés. Par ailleurs, nous avons observé sur l'ensemble de l'échantillon une corrélation modeste mais significative entre le score d'empan visuo-spatial et le pourcentage moyen de réussite

($r(49) = .38$; $p < .01$). Une analyse plus détaillée a montré que cette corrélation était significative pour les questions explicites ($r(49) = .48$; $p < .01$) mais pas pour les questions inférentielles ($r(49) = .14$; NS). Ce dernier résultat suggère que, lorsque la cible ne peut pas être localisée au moyen d'un appariement lexical direct, les participants engagent des processus inférentiels qui n'ont pas de rapport direct avec leurs capacités visuo-spatiales.

On peut évidemment se demander si les capacités visuo-spatiales sont prédictives de performances en recherche d'informations chez les sujets jeunes ou d'âge intermédiaire. Pour répondre à cette question, nous avons réparti les 35 sujets âgés de 24 à 53 ans en deux groupes selon leur score au test d'empan visuo-spatial : d'une part, les sujets à faible empan (score inférieur ou égal à 3) ; d'autre part, les sujets à fort empan (score supérieur à 3). Nous avons étudié la réussite de ces deux groupes aux questions littérales, d'une part, et inférentielles, d'autre part. Les résultats sont présentés sur la figure 4.

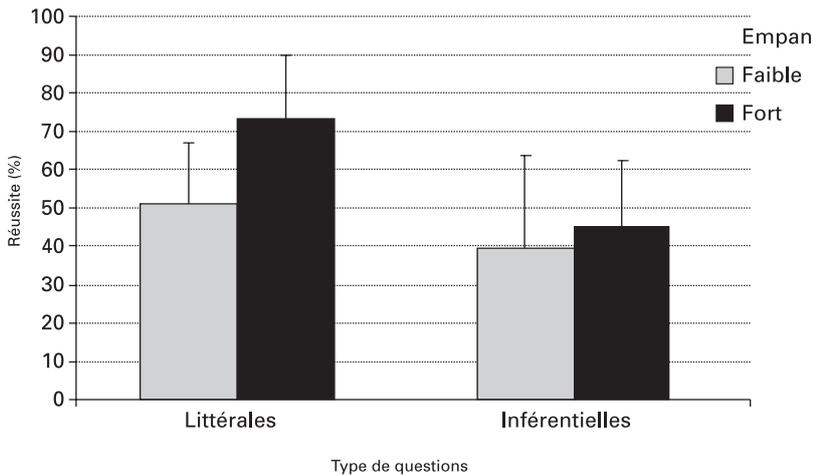


Fig. 4. — Réussite chez les participants de 24 à 53 ans selon l'empan visuo-spatial et le type de questions

Percentage of correct answers for 24 to 53 year-old participants according to the visuo-spatial span (Empan Fort : Good Span ; Empan Faible : Weak Span) and questions types (Littérales = Literal ; Inférentielles = Inferential)

L'analyse de la variance montre un effet significatif du type de questions (littérales > inférentielles, $F(1;33) = 35,79$; $p < .001$), ainsi que de l'empan visuo-spatial ($F(1;33) = 7,10$; $p < .01$). L'interaction entre les deux facteurs est également significative ($F(1;33) = 4,76$; $p < .05$). Les participants à fort empan visuo-spatial ont obtenu une réussite plus élevée que les participants à faible empan pour les questions littérales ($F(1;33) = 14,62$; $p < .01$), mais pas pour les questions inférentielles ($F < 1$). Cela confirme l'existence d'un facteur visuo-spatial dans la per-

formance en recherche d'informations. Ce facteur pourrait expliquer à la fois les moindres performances des sujets âgés ainsi qu'une partie des différences individuelles chez les sujets jeunes et intermédiaires. L'interaction du facteur visuo-spatial avec le type de question suggère que le problème de la représentation du but ne se pose pas dans les mêmes termes selon qu'un appariement lexical entre la question et le menu est possible (questions littérales) ou non (questions inférentielles).

VI. DISCUSSION

À l'heure où Internet tend à s'affirmer comme un ensemble de ressources et de services pour le grand public, il est nécessaire de s'interroger sur la conception des systèmes à menus et sur leur adéquation aux besoins des utilisateurs. Cette étude montre clairement qu'effectuer une recherche dans une structure de menus hiérarchiques n'est pas une tâche triviale pour le grand public. Même si nous étions sûrs que chaque question correspondait bien à une catégorie cible dans la hiérarchie des menus, le pourcentage moyen de réussite n'a été que d'environ 50 %, auquel s'ajoute une réussite « approximative » (choix d'un item incorrect dans une sous-catégorie correcte) dans environ 9 % des cas. Restent donc plus de 40 % de sélections incorrectes.

Nous avons également observé qu'une augmentation de la complexité du libellé des questions, par l'utilisation de synonymes ou par l'ajout de mots supplémentaires, entraînait une nette diminution en termes de réussite à la recherche de catégories cibles. Ces résultats suggèrent que les adultes ont tendance à rechercher une catégorie en se basant sur une similarité lexicale plutôt que sur une représentation sémantique du contenu de la question (c'est-à-dire macropropositions et modèle de situation, Kintsch, 1998). On ne peut donc pas se baser sur les seules tâches de localisation d'un item pour comprendre les processus cognitifs de la recherche d'informations (Rouet, 2003 *a* ; Rouet & Tricot, 1998). L'énoncé de la question fait l'objet d'un traitement cognitif qui intervient dans la complexité de la tâche à effectuer.

Les participants les plus âgés ont été plus lents que les deux autres groupes d'âge. Ce résultat ainsi que leur faible performance en condition d'utilisation d'un menu profond sont cohérentes avec les résultats du domaine (Freudenthal, 2001 ; Westerman *et al.*, 1995). Les sujets les plus âgés auraient plus de difficultés à gérer une série de sous-buts (dans notre cas, une série de sélections), et cela, peut-être, en raison d'une diminution de leur capacité de mémoire de travail. Par ailleurs, il faut noter la relativement bonne performance des participants âgés lorsqu'ils cherchent dans les menus larges et surtout alphabétiques. Par ailleurs, le relatif faible impact de la profondeur des menus chez les sujets jeunes peut paraître contradictoire avec les résultats des études « classiques » sur la recherche dans les menus (Giroux *et al.*, 1987 ; Snowberry *et al.*, 1983). Cependant, ce dernier résultat peut s'expliquer par les valeurs retenues pour configurer les menus larges et profonds. Dans notre expérience, le menu profond

comporte environ six items par niveau, pour une profondeur maximale de trois niveaux. Cela le situe dans une catégorie « optimale » et explique des performances comparables aux menus larges, certes moins profonds mais aussi plus complexes visuellement.

Plusieurs facteurs ont pu affecter les performances des participants les plus jeunes et les plus âgés. Nous avons montré que la capacité visuo-spatiale était l'un de ces facteurs, lié à l'âge mais intervenant également dans la performance des sujets jeunes et intermédiaires. Cependant il ne faut pas sous-estimer l'intervention possible d'autres facteurs, cognitifs ou non. Par exemple, les utilisateurs les plus âgés ont une attitude plus négative vis-à-vis des nouvelles technologies (Czaja & Sharit, 1998) et ont tendance à sous-estimer leurs connaissances vis-à-vis des ordinateurs (Marquié, Jourdan, & Huet, 2002). Dans la présente étude, les enquêteurs chargés de rencontrer les participants ont souvent rapporté une attitude plus anxieuse de la part des participants âgés. Ces différences d'attitudes et d'autres facteurs cognitifs peuvent être responsables des effets observés. Par exemple, le fait que les sujets les plus âgés ont eu des patterns de lectures des questions différents de ceux des deux autres groupes d'âge suggérerait un changement de stratégie plutôt qu'un réel effet dû à leur capacité cognitive. Relire une question est une action stratégique et délibérée qui est probablement dépendante de la représentation de la tâche et des contraintes de l'environnement. Bien que nous manquions de données pour conclure sur cette question, nous pouvons supposer que la confiance vis-à-vis des « affordances » (ici, la représentation externe des informations) de l'environnement puisse expliquer au moins certaines des différences observées.

Enfin, la relation observée entre l'empan visuo-spatial et l'âge des sujets est conforme aux données de la littérature (Freudenthal, 2001). Le fait que cette caractéristique soit en rapport avec la performance aux questions littérales, mais pas aux questions inférentielles, suggère un processus de recherche en deux étapes : les participants essaieraient d'abord d'apparier les mots clés de la question avec ceux des menus puis, en cas d'échec, ils engageraient un processus inférentiel plus coûteux en temps et incertain quant aux résultats. Seule la première étape serait sous la dépendance de leur capacité visuo-spatiale, ce qui expliquerait le pattern de corrélations observé.

Il serait évidemment intéressant d'étudier l'impact d'autres variables différentielles (comme les capacités verbales), ou d'examiner plus précisément l'impact de ces variables sur l'activité de recherche d'informations dans des menus (comme les erreurs de sélection, les relectures de la question, etc.). Cependant, le contexte du travail présenté ici se caractérisant comme une étude « écologique » réalisée avec des utilisateurs grand public, ne permettait ni d'avoir les effectifs suffisants, ni de mobiliser les participants sur une plus longue durée.

En résumé, nos résultats soulignent le fait que la recherche dans des menus repose sur des processus cognitifs plus complexes qu'il n'y paraît et soulignent la facilitation induite par le caractère explicite des catégories et des questions de recherche d'informations. Il paraît donc intéressant d'explorer des interfaces facilitant l'identification terme à terme des caté-

gories cibles et des objectifs de recherche des utilisateurs. Par ailleurs, il faudrait chercher des moyens pour aider l'utilisateur à garder une image précise de l'objectif initial de sa recherche d'informations. Qu'il s'agisse de tâches prescrites ou d'objectifs formulés par l'utilisateur, l'inclusion d'aide-mémoire et d'historique des choix réalisés pourrait faciliter l'activité de recherche d'informations dans des menus.

REMERCIEMENTS

Nous tenons à remercier la Communauté d'agglomération de Poitiers pour le soutien qu'elle a apporté à l'étude « 100 fenêtres sur Internet ». Nous remercions également Céline Delmas, Sami Limam et Aurélien N'Guyen pour leur participation à cette étude.

BIBLIOGRAPHIE

- Andréo, P., Jégou, G., & Rouet, J.-F. (2001). Représentations et usages d'Internet chez des utilisateurs novices : effets de variables individuelles et de l'apprentissage. Communication présentée au *Congrès de la Société Française de Psychologie*. Paris, octobre.
- Czaja, S. J., & Sharit, J. (1998). Age difference in attitudes toward computers. *Journal of Gerontology : Psychological Sciences*, 53, 329-340.
- Freudenthal, D. (2001). Age differences in the performance of information retrieval tasks. *Behaviour and Information Technology*, 20, 9-22.
- Giroux, L., Bergeron, G., & Lamarche, J.-P. (1987). Organisation sémantique des menus dans les banques de données. *Le Travail Humain*, 50, 97-107.
- Gyselinck, V., Ehrlich, M. F., Cornoldi, C., De Beni, R., & Dubois, V. (2000). Visuospatial working memory in learning from multimedia systems. *Journal of Computer Assisted Learning*, 16, 166-176.
- Kintsch, W. (1998). *Comprehension : A Paradigm for Cognition*. Cambridge, MA : Cambridge University Press.
- Larson, K., & Czerwinski, M. (1998). Web page design : Implications of memory, structure and scent for information retrieval. *Proceedings of CHI'98 Conference* (pp. 25-32). New York, NY : ACM.
- Marquié, J.-C. (2001). Changements cognitifs, contraintes de travail et expérience : les marges de manœuvre du travailleur vieillissant. In J.-C. Marquié, D. Paumès, & S. Volkoff (Eds.), *Le travail au fil de l'âge* (pp. 211-244). Toulouse : Octarès.
- Marquié, J.-C., Jourdan, L., & Huet, N. (2002). Do older adults underestimate their actual computer knowledge ? *Behaviour and Information Technology*, 21, 273-280.
- Miller, D. P. (1981). The depth/breadth tradeoff in hierarchical computer menus. *Proceedings of the Human Factors Society 25th Annual Meeting* (pp. 296-300). Santa Monica, CA : HFES.
- Parkinson, S. R., Sisson, N., & Snowberry, K. (1985). Organization of broad computer menu displays. *International Journal of Man-Machine Studies*, 23, 289-297.
- Rouet, J.-F. (1994). Question answering and learning with hypertext. In R. Lewis & P. Mendelsohn (Eds.), *Proceedings of IFIP WG3.3. Workshop : Lessons from Learning* (pp. 39-52). Amsterdam : North Holland.

- Rouet, J.-F. (2003 a). « What was I looking for ? » The influence of task specificity and prior knowledge on students' search strategies in hypertext. *Interacting with Computers*, 15, 409-428.
- Rouet, J.-F. (2003 b). *100 fenêtres sur Internet : rapport de fin de contrat*. Poitiers : Maison des Sciences de l'Homme et de la Société. Disponible : http://www.mshs.univ-poitiers.fr/laco/Pages_perso/Rouet/Textes.htm.
- Rouet, J.-F., & Tricot, A. (1998). Chercher de l'information dans un hypertexte : vers un modèle des processus cognitifs. In A. Tricot & J.-F. Rouet (Eds.), *Hypertextes et hypermédias* (pp. 57-74). Paris : Hermès.
- Snowberry, C., Parkinson, S., & Sisson, N. (1983). Computer display menus. *Ergonomics*, 26, 699-712.
- Westerman, S. J., Davies, D. R., Glendon, A. I., Stammers, R. B., & Matthews, G. (1995). Age and cognitive ability as predictors of computerized information retrieval. *Behaviour and Information Technology*, 14, 313-326.
- Weyer, S. A. (1982). The design of a dynamic book for information search. *International Journal of Man-Machine Studies*, 17, 87-107.
- Yu, B.-M., & Roh, S.-Z. (2002). The effects of menu-design on information seeking performance and users' attitude on the World Wide Web. *JASIST*, 53, 923-933.
- Zaphiris, P., Shneiderman, B., & Norman, K. L. (1999). *Expandable Indexes versus Sequential Menus for Searching Hierarchies on the World Wide Web*. Available : <ftp://ftp.cs.umd.edu/pub/hcil/reports-abstracts-bibliography/99-15html/99-15.html>.

RÉSUMÉ

Cette étude avait pour objectif d'examiner l'influence de variables individuelles et des choix de conception des menus web sur l'efficacité de la recherche d'informations. Cinquante utilisateurs d'Internet volontaires, répartis en trois groupes d'âge différents, ont réalisé une série de douze tâches de recherche de catégories cibles dans trois types de menus web hiérarchiques. La complexité des questions ainsi que la largeur et la profondeur des menus ont été manipulées.

Les résultats montrent que les participants les plus âgés ont réalisé les moins bonnes performances, notamment lors de l'utilisation du menu profond. De plus, les questions longues et inférentielles ont entraîné les moins bonnes performances pour l'ensemble des participants. Nous proposons ici une discussion des implications possibles de ces résultats du point de vue des facteurs individuels et des facteurs liés à la tâche de recherche d'informations.

Mots-clés : Web, Structure des menus, Recherche d'informations, Complexité des questions, Âge.

Manuscrit reçu : décembre 2003.

Accepté par J.-M. Hoc après modification : avril 2004.