HAL
open science

Le rôle de la structuration de la tâche dans la recherche d'informations sur site web

Guillaume Jégou, Pierre Andréo, Jean-François Rouet

To cite this version:

Guillaume Jégou, Pierre Andréo, Jean-François Rouet. Le rôle de la structuration de la tâche dans la recherche d'informations sur site web. Cinquième colloque Hypermédias et apprentissages, Apr 2001, Grenoble, France. pp.291-298. edutice-00000470

HAL Id: edutice-00000470

https://edutice.hal.science/edutice-00000470v1

Submitted on 11 Jun 2004

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.

LE RÔLE DE LA STRUCTURATION DE LA TÂCHE DANS LA RECHERCHE D'INFORMATIONS SUR SITE WEB

Guillaume JÉGOU, Pierre ANDRÉO et Jean-François ROUET

Laboratoire Langage et Cognition - CNRS et Université de Poitiers
99, avenue du Recteur Pineau - 86022 Poitiers Cedex

jegou@mshs.univ-poitiers.fr

Résumé *: Notre étude examine l'effet d'un guidage stratégique sur l'accès à l'information et l'apprentissage à partir d'un site Web chez des étudiants de premier cycle universitaire en fonction de consignes stratégiques. La tâche consistait à utiliser le site Web pour répondre à une série de questions en rapport avec un cours, le tout en temps limité. Les traces de navigation montrent que la consigne stratégique a permis aux participants de mieux planifier leur recherche. Toutefois la consigne n'a pas eu d'influence significative sur la performance lors d'un post-test. Entre autres facteurs, l'expérience préalable des étudiants dans l'usage de sites Web et leur connaissance préalable des contenus pourraient expliquer le profil de résultats obtenus.*

Mots-clés : *ergonomie, hypertexte, enseignement supérieur, métacognition, planification, stratégie de recherche.*

Abstract *: This study investigated the role of strategic guidance on the effectiveness of undergraduate students' search strategies and learning from a Web site. The participants were 65 students assigned to one of two conditions : with (S+) or without (S-) strategic directions. The task consisted in using the Web site in order to answer a series of course-related questions within a limited amount of time. The analysis of computer logs showed a positive impact of guidance on students' search strategies. However, directions were not shown to have any significant influence on students' post-test performance. Several factors, amongst which the participants' level of experience with Internet navigation, or their initial knowledge of the contents, may explain this pattern of results.*

Keywords *: higher education, human factors, hypertext, metacognition, planning, search strategy.*

INTRODUCTION

Le développement d'Internet et les possibilités qui en découlent quant à la publication, la diffusion et l'accès à des ressources documentaires offrent de

nouvelles perspectives pour l'enseignement universitaire. Entre autres, la production de sites d'accompagnement de cours permet aux étudiants de mener des recherches d'informations et de consulter des documents en rapport avec leur discipline. Mais l'utilisation efficace d'Internet fait intervenir différents types de compétences cognitives liées à l'utilisation de documents complexes : lecture, compréhension et recherche d'informations. Par ailleurs, la conception de documents hypertextuels se heurte au problème de la compatibilité cognitive, c'est à dire l'adéquation entre la structure de l'interface et les stratégies naturelles des utilisateurs (Rouet, 1999). Un bon hypertexte se doit donc d'être accessible aux utilisateurs novices et de tenir compte, dans ses choix de conception, de l'état de la recherche en ergonomie cognitive (Rouet & Tricot, 1998). D'autres facteurs individuels, comme les connaissances métacognitives ou les usages et les représentations de l'outil, pourraient moduler l'efficacité des stratégies des apprenants (Fisher & Mandl, 1984 ; Wagner & Sternberg, 1987). L'objectif de cette étude est de déterminer de quelle manière s'articulent variables individuelles et usages effectifs d'un site Internet destiné à aider les étudiants lors de la révision de l'un de leurs cours. On tentera également de voir dans quelle mesure une consigne visant à renforcer la planification d'une recherche d'information peut réduire ces différences individuelles et rendre cette recherche plus efficace pour tous (Cf. Beaufils, 1998 ; Tricot, Pierre-Demarcy & El Boussarghini, 1998).

MÉTHODE

Participants

Les participants ont étés recrutés, sur la base du volontariat, dans le cadre de leur cours de deuxième année de DEUG de psychologie, intitulé « Recherche et applications en psychologie cognitive », à l'Université de Poitiers. Sur 110 étudiants souhaitant participer à cette étude, 80 se sont effectivement présentés et ont rempli des questionnaires (76 étudiantes et 4 étudiants d'une moyenne d'âge de 22 ans).

Matériel

Le site Web. À partir des éléments communiqués par un enseignant, des notes prises par l'un des étudiants et de sources externes, un site Web, nommé « RELiX », a été développé afin de mener cette expérience [1]. Du point de vue de la fonctionnalité, l'une des principales caractéristiques de ce site Internet est la possibilité d'accéder aux contenus par l'intermédiaire d'un mode « plan », restituant l'enchaînement des contenus tels qu'ils ont été vus lors du cours, mais également grâce à un mode « index », permettant d'accéder directement aux mêmes contenus par l'intermédiaire d'une liste alphabétique des principales notions et expériences citées dans le cours. Dans les deux cas, mode « plan » ou « index », les liens de navigation restent apparents sur la partie gauche de l'écran en même temps que s'affiche le contenu. Sur la partie droite du contenu un espace est réservé à des définitions, accessibles par un simple clic sur certains mots du contenu (Figure 1). Un changement de couleur indique les mots dont la définition a déjà été consultée.

---

1 Remerciements à D. Gaonac'h et B. Floquet pour leur collaboration à cette étude

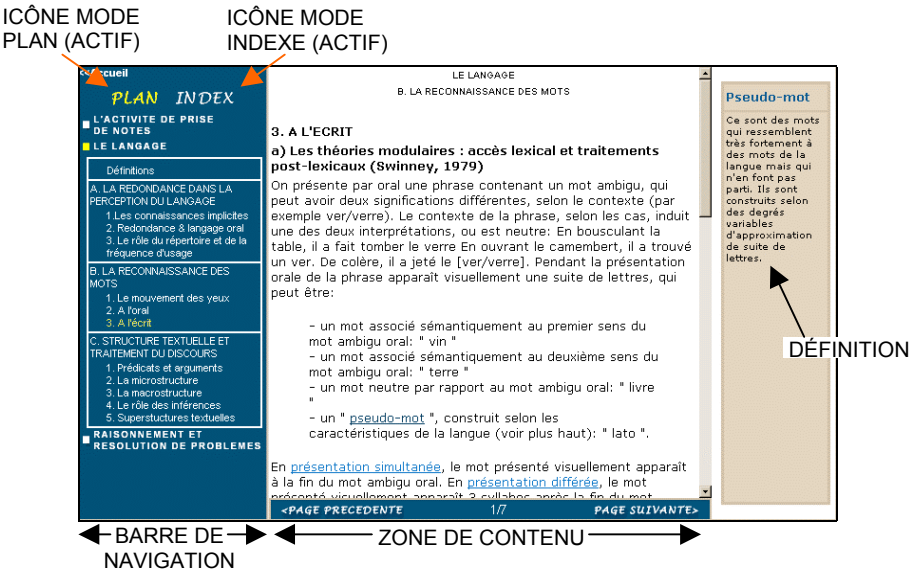

Figure 1. ***L'écran principal du site RELiX.***

Le site RELiX a été conçu pour être accessible à des novices. Les concepteurs ont ainsi veillé à ce que sa mise en forme tienne compte de principes issus des recherches en ergonomie cognitive. Du point de vue des processus de bas niveau (contraste, taille des caractères, type de polices...), les caractéristiques du site sont conformes aux recommandations de Caro et Bétrancourt (1998) sur l'ergonomie des documents informatisés. L'organisation interne de l'hypertexte a également été conçue de façon à respecter quelques principes de compatibilité cognitive (Rouet, 1999) comme la présence d'un plan, d'une séparation entre contenu et espace de navigation, le respect de la structure traditionnelle de textes imprimés (titres, numéros de pages...) et une correspondance explicite avec le cours dont le contenu est issu. Les marqueurs visuels indiquant la présence de liens hypertextes sont clairs, systématiques et conformes aux standards d'Internet (liens sous forme de textes « réactifs » ou soulignés...). Afin d'éviter les situations de surcharge cognitive et de désorientation, très fréquentes chez les utilisateurs novices d'hypermédias (Dillon, 1994 ; Wright, 1991), seules les informations pertinentes sont présentes à l'écran : barre de navigation sur la gauche, afin de repérer sa position dans l'arborescence et d'évaluer les choix disponibles, contenu sélectionné au centre de l'écran et espace réservé aux définitions sur la droite.

Les questionnaires. Plusieurs questionnaires ont étés utilisés. Le premier concerne des questions de recherche. Quatre questions ayant trait à différentes parties du site ont été formulées. Selon les cas ces questions demandaient la localisation d'une information précise ou la révision de plusieurs pages du site. Un second questionnaire (post-test) a servi à évaluer la compréhension des contenus visités lors de la recherche d'informations. En effet, 4 questions sur 6 concernaient les mêmes parties du site que celles visées par les questions de recherche. Seules ces questions sont prises en compte dans la notation. Enfin, un questionnaire de « satisfaction et d'impressions » sur la séance a été distribué en fin de séance aux étudiants. Trois

autres questionnaires portant sur différentes dimensions cognitives, métacognitives et psychosociales ont été distribués mais ne seront pas commentés ici.

Procédure

L'expérience se déroule en une séance de deux heures. Les participants, conviés par groupes de dix au maximum, sont informés qu'ils recevront une initiation à la navigation sur Internet, suivie d'une séance de révisions faisant appel à un document informatisé. Les questionnaires sont distribués à différents moments durant la séance.

Initiation au Web. La séance d'initiation à Internet est essentiellement centrée sur le Web. Après un bref explicatif du fonctionnement général d'Internet, l'initiation porte sur les principales fonctionnalités du navigateur Internet Explorer, la découverte de quelques sites Web et du moteur de recherche Yahoo (www.yahoo.fr). Les étudiants sont ensuite conviés à s'entraîner avec des exercices de recherche d'information sur Yahoo.

Exercices de recherche d'information. La partie expérimentale de la séance commence par une présentation du site RELiX et de ses principales fonctionnalités. L'utilisation du site est expliquée et démontrée à l'aide d'une première question de recherche. Les participants doivent ensuite traiter chacune des trois questions restantes, accessibles par la page d'accueil du site. Pour chaque question la tâche consiste à naviguer dans le site pour trouver la réponse, puis à rédiger celle-ci sur une feuille de réponse. L'ordre de présentation des questions est contrebalancé. Les étudiants sont répartis dans deux conditions expérimentales, qui varient par le degré de structuration de la séance. La proportion de sujets novices ou déjà expérimentés dans l'usage d'Internet est la même dans les deux groupes. Par ailleurs, les sujets ayant suivi le même cours, nous avons fait l'hypothèse que les deux groupes étaient équivalents quant à leurs connaissances préalables du contenu. Les sujets du groupe « fortement structuré » (S+) reçoivent pour chaque question une feuille comportant trois rubriques intitulées « préparation », « sélection », « mémorisation ». Ils ont pour consigne de noter pour chaque question des mots-clés qui leur viennent à l'esprit. Ils doivent ensuite repérer les rubriques pertinentes en s'aidant du « plan » et de « l'index », et noter des mots-clés correspondant dans l'espace nommé « sélection ». Enfin, les sujets ont pour consigne de consulter le contenu correspondant à leur(s) sélection(s) et de prendre des notes dans une rubrique « mémorisation ». Dans cette condition, et afin de renforcer la structuration de la séance, les sujets sont prévenus du temps écoulé toutes les dix minutes (temps normalement alloué à chacune des trois questions). Le groupe « S- » ne reçoit pas d'aide particulière à la planification des recherches d'informations. Une simple feuille leur a été fournie pour prendre des notes, mais sans consigne particulière. De même, ces étudiants ne sont pas informés du temps restant en cours d'exercice mais seulement du délai de 30 minutes maximum pour répondre aux questions. Les sujets des deux groupes peuvent cependant consulter l'heure affichée sur l'ordinateur (en standard, en bas à droite du « bureau » Windows). Au terme de la phase de recherche d'informations, les étudiants reçoivent le post-test de compréhension ainsi que le questionnaire de satisfaction.

Hypothèses

a) Les consignes précises devraient aider les étudiants du groupe S+ à planifier leur recherche d'information. Cette meilleure planification devrait se traduire par un parcours plus efficace et économique en termes de fréquence de « clics », nombre de retour aux questions ou éventuellement temps passé sur les exercices.

b) La condition S+ doit entraîner des sélections plus pertinentes et des résultats plus élevés au post-test de compréhension.

RÉSULTATS

Les parcours de recherche d'informations

Le parcours des étudiants dans le site, durant la recherche d'informations, est enregistré sous la forme de fichiers informatiques, appelés « logs ». Les logs ont ensuite été extraits, filtrés et analysés afin d'en dégager les données utiles et d'éliminer le « bruit » inhérent à ce type d'expérimentation. Suite à cela, seuls 65 « logs » ont pu être conservés. En effet, 15 sujets n'ont pas respecté les consignes, principalement en ne traitant pas les questions dans l'ordre suggéré.

Quatre paramètres ont été retenus pour caractériser les stratégies des participants. Le temps de recherche pour chaque question indique comment les participants ont géré la répartition du temps total de 30 minutes sur les trois questions. Le nombre de lectures de la question indique si la représentation de l'objectif est stable en mémoire. La relecture indique en général le besoin de « rafraîchir » cette représentation. La fréquence des « clics » et le pourcentage de contenus pertinents sélectionnés indiquent dans quelle mesure les choix effectués par les participants correspondent d'emblée à des informations pertinentes. Dans le cas contraire, la fréquence des clics augmente (effet de « papillonnage ») et le pourcentage d'informations pertinentes diminue. Les résultats par groupe et par question sont présentés dans le Tableau 1.

Temps de recherche par question. Les deux groupes passent en moyenne le même temps sur chaque question, ce qui était attendu compte tenu de la limite globale de 30 minutes. Par ailleurs le temps n'évolue pas significativement d'une question à l'autre. En revanche on observe une évolution différente des temps selon les groupes. Dans le groupe S+, le temps de recherche diminue significativement entre la question 2 et la question 3, alors que le temps de recherche du groupe S- diminue de la question 3 à la question 4. L'interaction groupe x question n'est cependant pas significative.

		Question 2		Question 3		Question 4	
		S+	S-	S+	S-	S+	S-
Temps de recherche (sec)	Moy	597	610	509	641	542	487
	Écart	185	216	144	246	212	259
Lecture des questions (N)	Moy	1,8	1,9	1,5	1,7	1,3	1,6
	Écart	0,8	0,9	0,7	0,9	0,6	0,8
Nombre de clics/min	Moy	2,1	2,1	2,1	2,0	2,1	2,6
	Écart	0,6	0,7	0,8	0,9	0,8	1,3
% de contenus pertinents sélectionnés	Moy	0,78	0,86	0,96	0,93	0,93	0,91
	Écart	0,31	0,22	0,13	0,23	0,24	0,24

Tableau 1. ***Paramètres de la recherche d'informations (N.B. la question 1 servant d'entraînement n'est pas représentée ici).***

Nombre moyen de clics par minute. De nouveau, ni le groupe ni la question n'ont d'effet significatif sur la fréquence des clics. Cependant on observe une interaction significative entre ces deux facteurs ($F(2,126)=4.32$, $p<.05$). Les comparaisons partielles montrent que la fréquence des clics augmente dans le groupe S- pendant le traitement de la dernière question.

Nombre et temps de lecture des questions. Le nombre de lecture des questions décroît significativement au fur et à mesure du déroulement de la séance ($F(2,126)=6.58$, $p<.01$). Il n'y a pas de différence globale entre groupes ni d'interaction entre les deux facteurs. Toutefois on observe que les sujets du groupe S+ semblent avoir besoin de moins de relecture de la question en moyenne (1,53 lecture contre 1,72). Afin d'approfondir cette analyse, nous avons calculé le temps moyen de lecture des questions, c'est à dire temps total passé avec chaque question sous les yeux au cours de la recherche. D'une manière globale, on constate que le groupe S- lit significativement moins longtemps les questions que le groupe S+ ($F(1,63)=6.28$, $p<.05$). La différence est d'environ 20% (55 secondes par question en moyenne dans le groupe S+ contre 45 secondes dans le groupe S-). Ainsi, les sujets du groupe S+ étudient plus longtemps la question et la relisent un peu moins souvent, ce qui suggère une meilleure planification de la recherche.

Pourcentage de contenus pertinents sélectionnés. Globalement, le taux de contenus pertinents sélectionnés est assez élevé (entre 78 et 96%, selon les groupes et les questions). Ceci suggère que les participants n'ont eu en général aucune difficulté à localiser les parties du site en rapport avec la question. Ce pourcentage augmente significativement au fur et à mesure de la séance de révision ($F(2,126)=5.91$, $p<.01$). Il ne diffère pas significativement d'un groupe à l'autre.

Les questionnaires

Post-test de compréhension. Les résultats au post-test de compréhension ne font pas apparaître de différence significative entre les groupes S+ et S-, ce qui infirme l'hypothèse b). Dans l'ensemble les résultats sont faibles (de l'ordre de 26% de réussite environ) ce qui indique soit que la recherche n'a pas été un moyen d'apprentissage efficace, soit que le questionnaire final était trop difficile.

Satisfaction et impressions des participants. Les impressions des étudiants sur le site RELiX font apparaître une majorité d'opinions positives. Des termes comme « clair », « complet » ou « facile d'utilisation » ont été cités par une majorité d'étudiants. Du point de vue des opinions négatives, c'est le fait de ne pas pouvoir afficher les questions en permanence qui revient en premier.

DISCUSSION ET CONCLUSION

Les principaux résultats obtenus vont dans le sens de l'hypothèse d'un effet positif de la structuration sur les stratégies de recherche. Le groupe S+ semble avoir mieux réparti son temps de recherche sur les trois questions. Les sujets du groupe S+ ont aussi lu la question plus attentivement, ce qui leur a permis d'éviter de la relire. Enfin, le temps de lecture entre deux clics est plus long ce qui suggère des choix plus judicieux et une conduite moins aléatoire, surtout pour la dernière question. Nos résultats rejoignent en partie ceux de A. Beaufils (1998) qui observait un impact positif d'aides à la planification sur les stratégies de recherche d'élèves de collège.

Cependant l'effet de la structuration reste limité tant en ce qui concerne le pourcentage de contenus pertinents sélectionnés que la compréhension de ces contenus. En effet sur ces deux mesures on n'observe aucune différence entre les deux groupes. Sur ce point notre expérience contraste avec des travaux antérieurs montrant une relation significative entre stratégie d'étude et performance (par exemple, Wagner & Sternberg, 1987).

L'absence d'effet sur la compréhension peut s'expliquer par plusieurs facteurs individuels et circonstanciels propres à cette expérience. D'une part, l'échantillon testé comportait une proportion notable d'utilisateurs novices ou très peu expérimentés en matière d'usage d'Internet. Or une analyse complémentaire montre que dans le sous-groupe des internautes réguliers (utilisation d'Internet plus d'une fois par mois, 15 sujets), le temps moyen de lecture des questions est en corrélation significative avec la note au post-test (.86). Cette corrélation n'est que de -.40 pour les non internautes (moins d'une utilisation par mois d'Internet). Il semblerait donc qu'une bonne planification favorise le rappel, mais seulement chez les utilisateurs expérimentés. Ceci corrobore les résultats d'études précédentes montrant que la familiarisation avec le dispositif influence notablement les stratégies des utilisateurs (cf. Rouet & Tricot, 1998). Par ailleurs l'expérience est intervenue en fin d'année universitaire, alors que les étudiants étaient en période de révision avant l'examen final. Ils disposaient donc en moyenne d'une bonne connaissance globale du cours, comme le suggère l'utilisation majoritaire du plan au détriment de l'index. Le pourcentage élevé de contenus pertinents sélectionnés et les réponses au questionnaire de satisfaction suggèrent également que la tâche était globalement facile, ce qui a pu masquer un éventuel effet de la structuration.

Au total, cette expérience nous a permis de mettre au point une méthode de recueil et d'analyse des parcours individuels des étudiants lors d'une tâche de recherche d'informations dans un site Web. Elle a également montré que des étudiants sans expérience préalable peuvent effectuer des tâches de recherche dans un site Web à condition que celui-ci soit structuré de façon claire et conforme aux connaissances des utilisateurs. Nous envisageons de prolonger ce travail par d'autres

études mettant en jeu des participants plus expérimentés dans l'usage d'Internet, mais disposant de moins de connaissances préalables des contenus traités. Cette situation serait d'ailleurs plus conforme à l'idée que l'on peut se faire de l'usage de sites dans un contexte de formation.

BIBLIOGRAPHIE

Beaufils A. (1998). « Aide à la recherche d'informations dans les environnements hypermédias », *in* J.-F. Rouet & B. de La Passardière (éds), *Hypermédias et Apprentissages 4*, Actes du 4ème colloque Hypermédias et apprentissages, Poitiers, 15, 16, 17 octobre 1998, Paris : INRP/EPI, p. 73-86.

Caro S. & Bétrancourt M. (1998). « Ergonomie des documents techniques informatisés : expériences et recommandations sur l'utilisation des organisateurs para-linguistiques », *in* A. Tricot et J.-F. Rouet (éds), *Les hypermédias, approches cognitives et ergonomiques*, Paris : Hermès, p. 123-138.

Dillon A. (1994). *Designing Usable Electronic Text: Ergonomic Aspects of Human Information Usage*, London : Taylor and Francis.

Fischer P. M. & Mandl H. (1984). « Learner, text variables and the control of text comprehension and recall », *in* H. Mandl, N. Stein & T. Trabasso (éds), *Learning and Comprehension of Text*, Hillsdale, NJ : Lawrence Erlbaum Associates.

Rouet J.-F. & Tricot A. (éds) (1998). « Chercher de l'information dans un hypertexte : vers un modèle des processus cognitifs », *in Les Hypermédias, approches cognitives et ergonomiques*, Paris : Hermès, p. 57-74.

Rouet J.-F. (1999). « Interactivité et compatibilité cognitive dans les systèmes hypermédias », *Revue des Sciences de l'Éducation*, n° 25, p. 87-104.

Tricot A., Pierre-Demarcy C. & El Boussarghini R. (1998). « Définitions d'aides en fonction des types d'apprentissages dans des environnements hypermédias », *in* J.-F. Rouet & B. de La Passadière (éds), *Hypermédias et Apprentissages 4*, Actes du 4ème colloque Hypermédias et Apprentissages, Poitiers, Paris : INRP/EPI, p. 41-58.

Wagner R. K. & Sternberg R. J. (1987). « Executive control in reading comprehension », *in* B. K. Britton & S. M. Glynn (éds), *Executive Control Processes in Reading*, Hillsdale, NJ : Lawrence Erlbaum Associates, p. 1-22.

Wright P. (1991). « Cognitive overheads *vs* cognitive protheses: some issues on evaluating hypertexts », *Proceedings Hypertext'91*, Texas : ACM Press, p. 1-12.