



**HAL**  
open science

# L'effet des connaissances du domaine et les modes de raisonnement dans le processus de sélection lors d'une recherche d'information

Aurélie Casado

► **To cite this version:**

Aurélie Casado. L'effet des connaissances du domaine et les modes de raisonnement dans le processus de sélection lors d'une recherche d'information. Education. 2013. dumas-00880304

**HAL Id: dumas-00880304**

**<https://dumas.ccsd.cnrs.fr/dumas-00880304v1>**

Submitted on 5 Nov 2013

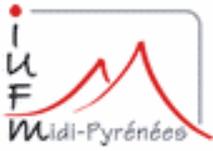
**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

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.



IUFM Midi-Pyrénées, Ecole Interne de l'Université Toulouse II

Le Mirail



# **L'effet des connaissances du domaine et les modes de raisonnement dans le processus de sélection lors d'une recherche d'information.**

Mémoire présenté par Aurélie CASADO

Pour l'obtention du master 2

Mention : Information Communication

Spécialité : Métiers de la formation et de l'enseignement : Documentation

Sous la direction de Mme Nicole BOUBEE, MCF en Sciences de l'information et de la communication.

Toulouse, le 21 juin 2013

## **Résumé**

Ce travail s'intéresse au processus de sélection, lors d'une recherche d'information sur le web, par des utilisateurs experts du domaine de la tâche à effectuer. Dans une observation conduite avec 5 élèves de Seconde d'un lycée général et technologique, nous tentons de mieux comprendre les critères de pertinence employés lors d'une recherche d'information sur le Web, à la fois sur la page de résultat et sur le document primaire. Plus précisément, notre étude vise d'une part, à déterminer les critères de pertinence qui sont employés pour sélectionner des liens dans la page de résultats et des documents primaires, et d'autre part si les connaissances du domaine peuvent influencer les sélections. Nos travaux portent également sur les modes de raisonnement, analytique ou usage d'heuristiques qui sont mis en place lors de ces recherches. Nos résultats ont mis en avant de nouveaux critères de pertinence, évoqués par les experts du domaine, et une différence significative entre les critères évoqués pour la justification des sélections des liens sur la page de résultats et des documents primaires. En ce qui concerne les modes de raisonnements utilisés, ils ne sont pas uniques au cours d'une recherche d'information mais varient selon les différents moments de la tâche.

## **Mots clés**

Recherche d'information, stratégie de recherche, pratiques juvéniles, connaissances du domaine.

# Table des matières

Introduction .....	6
Partie 1 : Etat de la question.....	8
1. L'influence des connaissances du domaine lors d'une recherche d'information.....	8
1.1 Sur la planification .....	8
1.2 Sur les mots-clés .....	8
1.3 Sur le jugement de pertinence.....	9
1.4 Sur les stratégies de recherche .....	10
1.5 Sur les résultats de la recherche .....	10
1.6 L'influence des connaissances techniques et des connaissances informationnelles	10
2. L'évaluation de l'information .....	12
2.1 La nécessité d'évaluer l'information .....	12
2.2 Les différents types d'évaluation .....	13
2.3 Qu'est-ce qu'une information pertinente ? .....	14
2.4 Les critères d'évaluation des pages de résultats .....	18
3. Les raisonnements mis en place lors d'une recherche d'information .....	24
3.1 Le raisonnement heuristique .....	24
3.2 Le raisonnement analytique .....	26
Partie 2 : Dispositif méthodologique.....	28
1. L'échantillon .....	28
2. Les tâches de recherches d'information .....	30
2.1 Le choix du moteur de recherche.....	30
2.2 Présentation des pages de résultats .....	30
2.3 Les sujets des recherches d'informations .....	30
3. La procédure.....	31

4. Le matériel utilisé.....	32
Partie 3 : Résultats.....	34
1. Résultats relatifs à la formulation de la requête .....	34
1.1 Définition du besoin d'information .....	34
1.2 Formulation de la requête .....	34
2. Résultats relatifs à la sélection dans la page de résultats .....	35
2.1 Pages de résultats du moteur de recherche consultées.....	35
2.2 Proportion de références sélectionnées par thème de recherche par les experts et le novice .....	35
2.3 La place des liens sélectionnés dans la page de résultats durant la tâche de recherche d'information .....	36
3. Justification des sélections : les critères de pertinence .....	37
3.1 Experts du domaine et page de résultats .....	37
3.2 Experts du domaine et documents primaires .....	38
3.3 Comparaison des critères de pertinence dans le processus de sélection mentionné par les experts pour la page de résultats et pour le document primaire .....	40
3.4 Comparaison des critères de sélection entre les experts et les novices du domaine.	47
Partie 4 : Discussion.....	49
Conclusion.....	53
Bibliographie.....	54
Annexes.....	57
Annexe 1 : Le choix de Google, top 5 des moteurs de recherche en part de visite sur le web .....	57
Annexe 2 : Grille d'observation .....	58

Annexe 3 : Grille d'aide pour l'entretien .....	59
Annexe 4 : Critères de pertinence des novices du domaine .....	60

## **Index des tableaux**

Tableau 1: Répartition des premiers liens par genre, d'après l'étude de J. Ravestain et al. (2007) .....	20
Tableau 2 : Détail des participants .....	29
Tableau 3 : Thèmes des recherches d'informations par participant .....	31
Tableau 4 : Nombre de formulation de requête par participant .....	35
Tableau 5: Moyenne de liens sélectionnés dans la page de résultats .....	36
Tableau 6 : Justification des sélections sur la page de résultats .....	38
Tableau 7 : Justification des sélections des documents primaires par les experts .....	40
Tableau 8: Expression des justifications des sélections par les experts .....	47
Tableau 9 : Expressions des justifications des sélections par les novices.....	62

## **Index des graphiques**

Graphique 1: Place des liens consultés dans la première page de résultats .....	36
---	----

# Introduction

Après le « *je l'ai lu dans un livre* », représentatif de la graphosphère, puis le « *je l'ai vu à la télé* », symbole de la vidéosphère, « *je l'ai trouvé sur Internet* » est devenu le « *dicton d'autorité personnelle* », selon la formule de Régis Debray (1998)<sup>1</sup>, de la cyber-sphère. Malgré la méfiance qui entoure souvent la fiabilité des informations, Internet est en passe de devenir le nouveau régime de vérité, notamment pour les jeunes générations, qui y puisent une large part de leurs informations sur l'actualité et sur le monde. Cette prédominance de plus en plus affirmée d'Internet, comme principal support et source privilégiée d'information pour les jeunes, atteste la profondeur de la question de l'évaluation de l'information. À partir du moment où n'importe quel type d'interrogation (qu'elle porte sur l'actualité ou qu'elle soit d'ordre documentaire, pratique, professionnel, culturel ou autre) ne suscite désormais qu'un seul réflexe de recherche, « aller sur Internet » ou « googliser », on mesure l'ampleur du problème de l'évaluation : celui de l'autonomie de jugement des jeunes face à l'information, de leur capacité à s'orienter, à identifier, à qualifier les univers informationnels d'Internet, à éviter les pièges et les dangers des diverses « info-pollutions », explique Serres (2005)<sup>2</sup>.

En outre, savoir rechercher rapidement et efficacement de l'information sur Internet fait partie des compétences que doivent nécessairement développer et posséder les élèves dès le Cycle 3, selon les programmes de l'Education Nationale<sup>3</sup> (2011). Pourtant, de multiples études ont révélé que de nombreuses difficultés peuvent expliquer le taux important d'échecs et d'abandons de cette activité chez les jeunes usagers.

En considérant les connaissances impliquées dans la recherche d'information nous pouvons distinguer trois types de connaissances qui conditionnent la réussite ou l'échec dans l'activité chez l'utilisateur. D'abord, les connaissances du domaine qui correspondent à des connaissances déclaratives et concernent le savoir sur le thème de la recherche d'information. Puis, les connaissances informationnelles qui renvoient à l'habileté pour réaliser une recherche d'information. Enfin, les connaissances techniques, c'est-à-dire savoir utiliser

---

<sup>1</sup> Merzeau L. (1998). «Ceci ne tuera pas cela.». *Les cahiers de médiologie*, n° 6, p. 27-39.

<sup>2</sup> Serres, A. (2005). Evaluation de l'information sur Internet : le défi de la formation. *Bulletin des bibliothèques de France*, t.50, n°6.

<sup>3</sup> Ministère de l'Education Nationale Jeunesse et Vie Associative (2011). *Brevet informatique et internet Ecole. Compétence 4 du socle commun, référentiel* [en ligne]. Mis à jour en décembre 2011. [Consulté le 18 juin 2012]. Disponible sur Internet [http://media.eduscol.education.fr/file/Certification\\_B2i/82/6/Referentiel\\_B2i\\_ecole\\_decembre\\_2011\\_202826.pdf](http://media.eduscol.education.fr/file/Certification_B2i/82/6/Referentiel_B2i_ecole_decembre_2011_202826.pdf)

l'outil. Ces trois types de connaissances sont étroitement liées lors d'une recherche d'information rappellent Ihadjadene, Chaudiron et Martins (2003).<sup>4</sup>

Nous avons décidé de travailler sur les connaissances du domaine, celles-ci étant très peu prises en considération dans les études s'intéressant à la recherche d'information des plus jeunes. Notre objectif est de voir si les connaissances du domaine influent sur la sélection des liens dans une page de résultats et sur le choix des documents primaires. Notre choix de considérer les documents primaires mais aussi la page de résultats est également basé sur la quasi-absence d'études sur les évaluations de la page de résultats. De même, nous élargissons aux différents raisonnements notre étude sur l'évaluation de l'information, en nous appuyant sur les deux raisonnements, analytique et usage d'heuristiques, mis en avant par Metzger *et al.* (2010)<sup>5</sup> dans leurs travaux sur la crédibilité de l'information. De façon plus précise, nous supposons que les connaissances du domaine antérieures ont un impact sur la sélection des références lors de la recherche d'information et sur les raisonnements employés.

Pour répondre aux exigences de notre question de recherche, nous avons effectué des observations directes sur des élèves qui effectuaient une tâche de recherche d'information scolaire. Nous avons observé à la fois les critères de sélections employés pour le choix des liens de la page de résultats et pour la justification des documents primaires, puis les raisonnements employés par les participants tout au long de leur tâche de recherche d'information..

Notre mémoire comprend quatre parties. D'abord un état de la question sur les notions de connaissances du domaine, sur celles d'évaluation de l'information et sur les différents types de raisonnement. Puis, une présentation de la méthodologie employée pour le dispositif d'observation, les résultats des données recueillies sur le terrain et enfin une discussion de nos résultats. Nous concluons sur une évocation des implications pratiques de nos travaux.

---

<sup>4</sup> Ihadjadene M., Chaudiron S. (2008). Quelles analyses de l'usage des moteurs de recherche : Questions méthodologiques. *Question de communication*. N°14, p. 17-32.

<sup>5</sup> Metzger, M. J., Flanagin, A. J., & Medders, R. B. (2010). Social and Heuristic Approaches to Credibility Evaluation Online. *Journal of Communication*, 60(3), 413–439.

# Partie 1 : Etat de la question

## 1. L'influence des connaissances du domaine lors d'une recherche d'information

Les connaissances du domaine ont une influence sur l'ensemble d'une tâche de recherche d'information. La planification de la recherche d'information, la production d'une requête, le processus de sélection ainsi que les stratégies de recherche employées dépendent des connaissances du domaine du sujet.

### 1.1 Sur la planification

Les experts du domaine planifient, c'est-à-dire qu'ils passent du temps à définir leur besoin d'information, particulièrement en début de processus. Ainsi, ils lisent les énoncés plusieurs fois et font appel à leurs connaissances antérieures plus souvent que les novices, comme le montrent les travaux de S. Brand-Gruwel, I. Wopereis, et Y. Vermetten (2005)<sup>6</sup>.

### 1.2 Sur les mots-clés

Hembrooke H. et Granka L.A. (2005)<sup>7</sup> ont montré que le niveau de connaissance a une incidence sur la capacité à trouver des mots-clés. En effet, d'après Holscher et Strube<sup>8</sup> (2000), le changement dans la formulation de requêtes est généralement inefficace chez les novices.

---

<sup>6</sup> Brand-Gruwel S., Wopereis I., Vermetten Y. (2005). Information problem solving by experts and novices : Analysis of a complex cognitive skill. *Computer in Human Behavior*, vol. 21, n°3, p. 487-508

<sup>7</sup> Hembrooke H. A., Granka L. A., Gay G. K. (2005). "The effect of expertise and feedback on search term selection and subsequent learning". *Journal of the American Society for Information Science and Technology*, vol. 56, n°8, p. 861-871.

<sup>8</sup> Holscher C., Strube G. (2000). « Web search behavior of Internet experts and newbies ». *Computer Networks*, vol. 33, n°1-6, p.337-346.

Ils constatent que ces derniers font des requêtes plus longues (en nombre de mots) et supposent que les experts du domaine connaissent des termes plus appropriés.

Selon les travaux de Ihadjadène, Chaudiron et Martins (2003)<sup>9</sup>, le nombre de reformulations de requêtes est moins élevé pour les novices du domaine que pour les experts. Au contraire, Wildemuth (2004)<sup>10</sup>, note une grande réitération des requêtes chez les novices du domaine car ils éprouvent des difficultés à fournir les requêtes appropriées.

### **1.3 Sur le jugement de pertinence**

Les observations menées par Brand-Gruwel et ses collègues (2005)<sup>11</sup> mettent en avant le fait que les novices décident très tôt de l'utilité des documents dans le processus de recherche. Au contraire, les experts du domaine examinent le contenu informationnel et font aussi attention à la source. De plus, ces derniers passent peu de temps avec un document du domaine où ils sont experts : c'est-à-dire qu'ils passent moins de temps pour le lire et pour prendre une décision, comme le montrent Holscher et Strube(2000)<sup>12</sup> dans leurs travaux.

Jenkins(2003)<sup>13</sup> met en avant dans ses travaux que les connaissances du domaine sont importantes pour évaluer un document. Aussi, les novices du domaine mais experts en utilisation du web évaluent plus souvent que les doubles novices mais leurs évaluations restent formelles (date de publication etc.) et ne portent pas sur le contenu lui-même.

---

<sup>9</sup> Ihadjadene M., Chaudiron S., Martins D. (2003). « The Effect of Individual Differences on searching the Web », *Proceedings of the 66th Annual Meeting of the American Society for Information Science and Technology*, Long Beach, p. 240-246.

<sup>10</sup> Wildemuth B. (2004). The Effects of Domain Knowledge on Search Tactic Formulation. *Journal of the American Society for Information Science and Technology*, vol. 55, n°3, p. 246-258.

<sup>11</sup> Brand-Gruwel S., Wopereis I., Vermetten Y. (2005). Information problem solving by experts and novices : Analysis of a complex cognitive skill. *Computer in Human Behavior*, vol. 21, n°3, p. 487-508

<sup>12</sup> Holscher C., Strube G. (2000). « Web search behavior of Internet experts and newbies ». *Computer Networks*, vol. 33, n°1-6, p.337-346.

<sup>13</sup> Jenkins C., Corritore C., Wiedenbecks S. (2003). Patterns of information seeking on the Web : A qualitative study of domain expertise and Web expertise. *IT and society*, vol. 1, n°3.

## **1.4 Sur les stratégies de recherche**

Dans leurs travaux de recherche, Brand-Gruwel et ses collègues (2005)<sup>14</sup> ne notent pas de différences dans la façon de chercher chez experts et novices : dans les deux cas, ils utilisent le plus souvent les liens.

D'autre part, dans ses travaux, Wildemuth(2004)<sup>15</sup> met en avant que les stratégies de recherche changent au fur et à mesure que les connaissances du domaine changent.

## **1.5 Sur les résultats de la recherche**

De nombreux travaux comme ceux de Hembrooke, Granka et Gay (2005)<sup>16</sup> ou ceux de Holscher et Strube (2000)<sup>17</sup> ont mis en avant que les experts trouvent plus de documents de qualités ou spécifiques que les novices du domaine. Ainsi, Hirsh<sup>18</sup>, en 1997, conclut que les experts du domaine accomplissent une tâche de recherche d'information avec plus de succès, quel que soit le niveau de complexité de la tâche.

## **1.6 L'influence des connaissances techniques et des connaissances informationnelles**

---

<sup>14</sup> Brand-Gruwel S., Wopereis I., Vermetten Y. (2005). Information problem solving by experts and novices : Analysis of a complex cognitive skill. *Computer in Human Behavior*, vol. 21, n°3, p. 487-508

<sup>15</sup> Wildemuth B. (2004). The Effects of Domain Knowledge on Search Tactic Formulation. *Journal of the American Society for Information Science and Technology*, vol. 55, n°3, p. 246-258.

<sup>16</sup> Hembrooke H. A., Granka L. A., Gay G. K. (2005). "The effect of expertise and feedback on search term selection and subsequent learning". *Journal of the American Society for Information Science and Technology*, vol. 56, n°8, p. 861-871.

<sup>17</sup> Holscher C., Strube G. (2000). « Web search behavior of Internet experts and newbies ». *Computer Networks*, vol. 33, n°1-6, p.337-346.

<sup>18</sup> Hirsh S. G. (1997). "How do children find information on different types of tasks? : Children's use of the science library catalog". *Library trends*, vol. 45, n°4, p. 725-745.

De nombreuses études prennent en compte les dimensions individuelles et cognitives des usagers en s'appuyant par exemple sur une différenciation des usagers selon leur compétence conceptuelle (connaissance du domaine ou non) et leur savoir faire (expertise dans l'utilisation du système). Ainsi, Ihadjadène et Chaudiron (2008)<sup>19</sup> recensent de nombreux travaux sur cette question.

Lazonder *et ses collègues*(2000)<sup>20</sup> et Palmquist et Kim (2000)<sup>21</sup> ont tenté de répondre à cette question dans le cadre d'une expérience comparant une recherche d'information entre deux groupes dont l'un est composé de non-utilisateurs du web et l'autre d'utilisateurs réguliers, les deux étant considérés comme maîtrisant de la même façon le domaine sur lequel portait la recherche d'information. Aucune différence significative entre les deux types de groupes n'a été mise en évidence, les utilisateurs réguliers du web n'ayant pas fait de recherches d'information plus performantes. Les chercheurs ont plus précisément remarqué que les utilisateurs familiers font un usage aussi intempestif qu'inefficace du site web : ouverture fréquente de pages non pertinentes, nombre très grand de liens activés.

Aussi, les travaux de Lazonder *et ses collègues* (2000) montrent que les experts de l'internet parviennent à rechercher de l'information plus rapidement que les novices. Un des résultats intéressants de telles études est que l'expérience de recherche en ligne peut énormément réduire l'effet de dépendance vis-à-vis du domaine.

Ce résultat est corroboré par les travaux de Ihadjadene, Chaudiron et Martins (2003)<sup>22</sup> qui suggèrent l'existence d'un mécanisme de compensation entre les deux types de compétences (dans le domaine et dans la recherche d'information). D'une manière générale, les experts font davantage appel à leurs méta-connaissances, en ce sens qu'ils possèdent les connaissances nécessaires à l'optimisation de leurs stratégies de recherche. Ainsi s'orientent-ils plus souvent vers des sites spécialisés pour lesquels ils font l'hypothèse qu'ils contiennent

---

<sup>19</sup> Ihadjadene M., Chaudiron S. (2008). Quelles analyses de l'usage des moteurs de recherche : Questions méthodologiques. *Question de communication*. N°14, p. 17-32.

<sup>20</sup> Lazonder et al. (2000). "Differences between novice and experienced users in searching in the World Wide Web », *Journal of the American Society for Information Science*, n° 51, pp.576-581.

<sup>21</sup> Palmquist R. A., Kim K. (2000,). « The Effect of Cognitive Style and Online Search Experience on Web Search Performance », *Journal of the American Society o Information Science*, n°51, pp. 558-567.

<sup>22</sup> Ihadjadene M., Chaudiron S., Martins D. (2003). « The effect of individual differences on searching the Web », *Proceeding ASIST 2003 Annual Meeting*, pp. 240-246.

les réponses à leurs questions, comme le montre également Tabatabai dans ses travaux<sup>23</sup> en 2005.

D'autres études, comme celles de Choo et ses collègues(2000)<sup>24</sup> montrent néanmoins que, malgré leur expérience, les experts ont rarement recours aux fonctionnalités avancées telles que les expressions booléennes, la troncature de mots ou les opérateurs de proximité lexicale dans leur interrogation des moteurs de recherche.

Les doubles novices (du domaine et au niveau des connaissances techniques) utilisent la fonction pour retourner sur le navigateur et retourner à la page de résultats plutôt que de changer la stratégie de recherche, selon Holscher (2000)<sup>25</sup>.

Dans les mêmes travaux, il est démontré que les doubles experts c'est-à-dire du domaine et ayant des connaissances informationnelles, se dirigent directement vers des sites spécifiques plutôt que d'utiliser des moteurs de recherche. Ces deux types d'expertise influent donc sur les stratégies de recherche.

## **2. L'évaluation de l'information**

### **2.1 La nécessité d'évaluer l'information**

Dans ses travaux, Ravestein (2007)<sup>26</sup> reprend une étude de Pew Internet et American Life Project de 2005 sur l'utilisation des moteurs de recherche qui indique, d'une part que les utilisateurs sont très satisfaits de leurs moteurs de recherche et d'autre part, qu'ils sont parfaitement ignorants de la manière dont ils fonctionnent et présentent les résultats. Ainsi, pour 68% de la population de l'enquête, les outils de recherche sur internet sont considérés comme des sources d'information légitimes et non biaisées. De plus, seulement 38% des internautes sont conscients de l'affichage de liens sponsorisés et d'informations commerciales dans les résultats fournis et arrivent à les reconnaître.

---

<sup>23</sup> Tabatabai D., Shore B. (2005). « How experts and novices search the Web », *Library & information science research*, 27, 2, pp. 222-248.

<sup>24</sup> Choo *et al.* (2000). « Information Seeking and the Web: An Integrated Model of Browsing and Searching », *First Monday*, n°5.

<sup>25</sup> Holscher C., Strube G. (2000). « Web search behavior of Internet experts and newbies ». *Computer Networks*, vol. 33, n°1-6, p.337-346.

<sup>26</sup> Ravestein J., Ladage C., Johsua S. (2007). Trouver et utiliser des informations sur Internet à l'école : problèmes techniques et questions éthiques. *Revue française de pédagogie*, n° 158, p.71-83.

Tout processus de recherche d'information n'a de sens aux yeux des utilisateurs qu'à travers l'ultime étape qui consiste à visualiser les résultats. Aussi, les profonds changements qui ces dernières années ont marqué le paysage informationnel, ont mis la question de l'évaluation au cœur des processus de traitement de l'information : la multiplication de ressources numériques aisément accessibles mais instables et insuffisamment validées rend cette évaluation plus nécessaire que jamais et en même temps plus délicate à effectuer.

## 2.2 Les différents types d'évaluation

Selon Gerjets, Kammener et Wermer (2010)<sup>27</sup>, deux processus d'évaluation ont lieu lors d'une recherche d'information sur le web : une évaluation des résultats affichés par le moteur de recherche et une évaluation des contenus des pages et/ou sites web.

Cet aspect est important et Rieh (2002)<sup>28</sup> met en évidence le fait que, dans son étude, les universitaires interrogés portent deux types de jugements lorsqu'ils s'informent sur le web : des jugements prédictifs et des jugements évaluatifs. Les jugements prédictifs leur permettent de décider s'ils vont consulter ou non telle source. Les jugements évaluatifs portent sur le contenu des documents ou des ressources et leur permettent de décider d'en utiliser effectivement ou non le contenu. Rieh met en avant que pour effectuer ces jugements, les personnes observées s'appuient sur des hypothèses conscientes ou inconscientes relatives à la valeur de l'information. Ainsi, les évaluations ne sont pas dichotomiques, par exemple adéquat/inadéquat, mais multidimensionnelles, c'est-à-dire qu'elles font intervenir des critères mis en œuvre selon une hiérarchie qui varie en fonction des situations. Rieh montre que les jugements prédictifs s'appuient davantage sur la connaissance personnelle de l'internaute en termes de système ou de domaine du sujet : souvent, les personnes observées se dirigent vers un site spécifique qu'ils connaissent ou qui leur a été conseillé. Les caractéristiques formelles des sources influencent également les jugements prédictifs. Par contre, les jugements évaluatifs s'appuient davantage sur le contenu informationnel, le type d'objet d'information et

---

<sup>27</sup> Gerjets P., Kammener Y., Wermer B. (2010). "Measuring spontaneous and instructed evaluation process during Web search: Integrating concurrent thinking-aloud protocols and eye-tracking data", *Learning and Instruction*, p. 1-12.

<sup>28</sup> Rieh, S. Y. (2002). Judgment on Information Quality and Cognitive Authority in the Web. *Journal of the American Society for Information Science and Technology*, vol.53, n°2, p. 145-161.

sa présentation. Ainsi, elle note que les individus n'accordent pas la même importance aux différents critères à ces deux étapes, l'importance relative attribuée à chacun variant selon la situation. Les critères sur lesquels s'appuient les jugements de pertinence des individus semblent dépendre d'un certain nombre de facteurs notamment l'expérience de l'individu vis-à-vis du SRI et sa connaissance du domaine sur lequel porte l'interrogation.

La plupart des outils électroniques de recherche d'information répond à l'interrogation documentaire de l'utilisateur en proposant une liste, plus ou moins longue, d'informations. Or, le degré de pertinence des informations proposées par l'outil peut être très hétérogène. Aussi, savoir évaluer la pertinence de ces informations est primordiale puisque de cette évaluation dépend la suite de la recherche d'information. Ainsi, l'activité de sélection des informations pertinentes parmi un ensemble d'information est centrale dans toute recherche d'information.

Pour répondre à la requête, le système de recherche d'information présente à l'utilisateur un ensemble de résultats. Simonnot (2008)<sup>29</sup> note que ces résultats sont en général d'abord affichés sous une forme abrégée. Il peut s'agir d'une partie de la notice du document dans un catalogue de bibliothèque ou, dans les pages de résultats des banques de publication, d'une courte suite d'éléments descriptifs (souvent tirés de la référence bibliographique). Ce peut être aussi un *snippet* c'est-à-dire la façon dont les résultats sont présentés dans la page de résultats de recherche, comme dans les moteurs de recherche, où l'on retrouve le titre, l'adresse du document et parfois la date de publication ainsi qu'un bref résumé ou quelques bribes du document contenant les termes de la requête.

## **2.3 Qu'est-ce qu'une information pertinente ?**

### **2.3.1 Le concept de pertinence en recherche d'information**

---

<sup>29</sup> Simonnot B. (2008). « La pertinence en sciences de l'information : des modèles, une théorie ? ». In Papy F., *Problématiques émergentes dans les Sciences de l'Information*, Paris : Hermès-Lavoisier, ch.6, p. 161-182.

Dans ses travaux en 2002<sup>30</sup>, Simonnot rappelle que les chercheurs et les concepteurs de systèmes de recherche d'information informatisés (SRI) distinguaient deux types de pertinence : la « pertinence système », c'est-à-dire l'évaluation par un système de l'adéquation entre des documents et une requête, et la « pertinence utilisateur » qui se traduit par des jugements de pertinence sur les documents fournis en réponse à une requête.

Simonnot présente plusieurs catégories de pertinence, selon les différentes étapes de la recherche documentaire :

- La pertinence des mesures de classement qui mesure le degré auquel le système est capable de retrouver et classer des documents en lien avec une requête.
- La pertinence de l'indexation car avec l'indexation automatique des textes, la difficulté consiste à trouver des termes descripteurs suffisamment discriminants pour caractériser le document.
- La pertinence de la formulation de la requête puisque la qualité des résultats de la recherche dépend fortement de la qualité de la formulation de la requête. Or, pour un utilisateur non averti, formuler une requête qui représente correctement son besoin est un challenge. La plupart des moteurs proposent deux types d'interfaces d'interrogation : la recherche simple et la recherche avancée.
- La pertinence liée à la valeur des documents car un des problèmes essentiels dû à l'absence de chaîne éditoriale sur le Web est la crédibilité et la valeur des informations qui y sont publiées. Mettre un document en ligne est en effet à la portée de chacun. Pour estimer la crédibilité des sources, les moteurs utilisent largement un indice de popularité des pages (IPP), inspiré du facteur d'impact conçu en bibliométrie pour l'information scientifique et technique. Le facteur d'impact d'un article, calculé comme le nombre moyen de citations dont il a fait l'objet durant une période de référence, est considéré comme un indicateur de qualité de la recherche scientifique.

---

<sup>30</sup> Simonnot B. (2002). *De la pertinence à l'utilité en recherche d'information : le cas du Web*. Actes du colloque international MICS-LERASS, Toulouse ; ADBS Éditions, collection Sciences de l'information, série Recherches et Documents, Paris, p. 393-410.

- La pertinence de la présentation des résultats: les résultats d'une requête sont présentés généralement sous la forme d'une page où apparaissent, de manière condensée, un certain nombre de liens vers des documents répondant potentiellement à la demande. Cette présentation succincte doit permettre à l'utilisateur de juger de l'opportunité de consulter l'intégralité de la ressource. Pour la composition de cette présentation succincte, plusieurs approches se distinguent. D'une part, l'approche documentaire classique, comme Alta Vista, qui propose un résumé du document extrait soit de l'en-tête, soit du début du texte. D'autre part, une approche comme le moteur de recherche Google qui sélectionne des extraits du texte où apparaissent en gras les termes de la requête. Si les phrases hachées et les extraits trop brefs ne permettent pas toujours de se faire une bonne idée de la ressource, cette présentation rassure le chercheur d'information sur le bon fonctionnement du moteur : il peut constater que le document proposé contient bien les termes qu'il a utilisés pour sa requête et il est invité implicitement à formuler une meilleure requête s'il n'est pas satisfait des résultats obtenus.
  
- La pertinence du document par rapport au besoin d'information : B. Simonnot rappelle que juger de la pertinence du contenu du document par rapport à la demande fait intervenir différents paramètres liés à l'utilisateur : l'état de ses connaissances sur le sujet de sa recherche, ses savoir-faire - lecture rapide, lecture approfondie - et son niveau d'éducation - niveau de langage, richesse du vocabulaire, aptitude à décrypter des tournures syntaxiques, aptitude à transférer une méthodologie d'un domaine à un autre – ce qui va influencer la pertinence de son interprétation. Aussi, le temps dont il dispose influera sur la longueur des documents choisis.
  
- La pertinence dans le choix de l'outil : Il existe nombre de moteurs ou d'annuaires différents sur le Web. Choisir un outil adéquat pour son type de besoin fait partie d'une bonne démarche de recherche documentaire.

### **2.3.2 Les critères qui interviennent dans le jugement de pertinence**

De nombreux travaux ont tenté d'analyser la façon dont se forment les jugements de pertinence des individus. Simonnot (2008)<sup>31</sup> reprend dans ses travaux ceux de Barry et Schamber (1998)<sup>32</sup> qui ont procédé toutes deux à des études empiriques sur cette question. Tentant de comparer les résultats obtenus dans leurs études respectives, réalisées dans des contextes différents, elles trouvent des catégories de critères plutôt cohérentes. Les principales sont :

- l'étendue, le degré d'approfondissement, la spécificité de l'information ;
- la justesse, la validité ;
- la clarté ;
- le degré d'actualité, le fait que l'information soit récente ;
- le caractère tangible, si l'information comporte des données ou des faits concrets ;
- la qualité ou la réputation de la source ;
- l'accessibilité, l'usabilité ;
- la disponibilité de l'information ou de la source d'information ;
- la vérification, le fait que l'information peut être recoupée par plusieurs documents ou sources ;
- l'émotion induite par l'information.

D'autres critères restent spécifiques à l'une ou l'autre étude, montrant bien l'effet du contexte et de la situation de recherche sur les jugements que les usagers portent sur l'information.

Simonnot reprend également les travaux de Knight et Burn (2005)<sup>33</sup> qui recensent les critères nécessaires à l'évaluation de l'information en ligne. S'appuyant sur d'autres études portant sur la qualité informationnelle, elles identifient de nombreux critères dont la plupart recourent les observations de Barry et Schamber (1998)<sup>34</sup>. Parmi les critères qu'elles retiennent, un plus grand nombre d'entre eux concernent l'utilité de l'information. Elles

---

<sup>31</sup> Simonnot B. (2008). « La pertinence en sciences de l'information : des modèles, une théorie ? ». In Papy F., *Problématiques émergentes dans les Sciences de l'Information*, Paris : Hermès-Lavoisier, ch.6, p. 161-182.

<sup>32</sup> Barry C. L., Schamber L. (1998). « Users' criterias for relevance evaluation : a cross study comparison », *Information Processing and Management*, vol.24, p. 219-236.

<sup>33</sup> Knight S. A., Burn J. (2005). « Developing a framework for assessing information quality on the World Wide Web », *Informing Science Journal*, vol. 8, p. 159-172.

<sup>34</sup> Barry C. L., Schamber L. (1998). « Users' criterias for relevance evaluation : a cross study comparison », *Information Processing and Management*, vol.24, p. 219-236.

soulignent aussi l'influence du SRI sur les jugements des utilisateurs. De plus, l'interface du système joue aussi notamment quant à l'ordre dans lequel sont présentés les résultats.

Enfin, grâce aux travaux de Wang et Soergel (1998)<sup>35</sup>, Simonnot rappelle que les critères tendent à devenir de plus en plus exigeants au fur et à mesure de l'avancement de la session de recherche et que, même si certains documents sont jugés pertinents par un chercheur d'information, ils ne seront pas tous exploités ou utilisés réellement.

## 2.4 Les critères d'évaluation des pages de résultats

La lecture des articles de chercheurs sur la question des critères de jugement de pertinence montre une certaine évolution entre le milieu des années quatre-vingt-dix et le début du 21<sup>ème</sup> siècle. Les critères liés à la présentation de l'information semblent monter en puissance, note Simonnot (2007)<sup>36</sup>. Auparavant, on évoquait uniquement la clarté et la lisibilité des textes. A présent, les critères ergonomiques prennent de l'importance et notamment ceux liés à la navigation au sein des sites. C'est que, en situation de recherche d'information en ligne, la charge cognitive est forte : il faut à la fois piloter le système et analyser rapidement pour décider quel document consulter. Une mauvaise présentation est un obstacle à cette première évaluation.

Déjà en 1999, Hirsh avait recensé dans ses travaux<sup>37</sup> les critères de pertinence appliqués par les jeunes lors de recherches d'information. Ainsi, elle a mis en avant 9 critères de pertinence appliqués par les jeunes : l'adéquation au thème, la nouveauté, l'intérêt personnel, le fait que le document soit intéressant pour d'autres, la qualité de l'information, l'accessibilité, l'autorité et la langue. Cependant d'autres chercheurs en ont relevé d'autres.

Les critères pour des documents secondaires sont moins nombreux. Hirsh note principalement : l'adéquation au thème de recherche, la position du lien dans la page de résultat, la cohérence, le type de document, l'étendue, la conformité avec l'opinion personnelle et la nouveauté.

---

<sup>35</sup> Wang P., Soergel D. (1998). « A Cognitive Model of Document Use during a research project. Study I. Document Selection », *Journal of the American Society for Information Science*, vol.49, p. 115-153.

<sup>36</sup> Simonnot B. (2007). Evaluer l'information. *Documentaliste – Sciences de l'information*, vol.14, n°3, p.210-216.

<sup>37</sup> Hirsh S. G. (1999). Children's relevance criteria and information seeking on electronic resources. *Journal of the American Society for Information Science*, vol. 50, n°14, p. 1265-1283.

La difficulté réside dans le fait que les critères de pertinence appliqués changent au cours de la recherche d'information.

#### **2.4.1 La position de la référence / du lien**

Les travaux de De Cara et ses collègues (2009)<sup>38</sup> soulignent le poids accordé à la position de la référence au sein de la liste, surtout chez les jeunes utilisateurs. Ainsi, chez des élèves de CM2, la référence en position 1 dans la liste de résultat est plus fréquemment sélectionnée lors du 1<sup>er</sup> clic (41,7%) que la référence en position 2 (10,2%).

De la même façon, les travaux de Granka, Joachims et Gay (2004)<sup>39</sup> mettent en avant que les sélections de liens des élèves révèlent les mêmes traits caractéristiques que ceux des usagers adultes, à savoir une sélection effectuée majoritairement sur les liens positionnés en rang 1 et 2. Ces deux rangs constituent 58,6 % des sélections de l'ensemble « collège et lycée ». Toutefois, les observations menées font apparaître un nombre de sélections non négligeable des liens en rang 3 et plus (41 % des sélections). De plus, il est noté que les élèves choisissent rarement plus de deux liens dans la première page et les pages suivantes ne sont que très peu explorées. Ces résultats sont en conformité avec le nombre élevé de soumissions de requêtes : les élèves préfèrent les reformulations à l'examen méthodique des résultats.

Dans leurs travaux en 2007, Ravestein, Ladage et Johsua<sup>40</sup> ont mené des observations sur plusieurs thèmes de recherche d'information. Ils ont notamment étudié la répartition du premier lien, qui est le plus souvent cliqué par les élèves, par genre. Ils ont utilisé trois critères : les sites institutionnels contrôlés qui traitent de la requête (émanation d'une organisation scientifique ou « légale » comme le CNRS, une université, un gouvernement, etc.) ; les sites privés qui sont l'émanation d'une initiative personnelle ou collective et qui traitent de la

---

<sup>38</sup> De Cara B. [et al.] (2009). *La recherche d'information sur Internet par les jeunes usagers : acquisition et développement des stratégies*. Cinquième colloque de Psychologie Ergonomique (EPIQUE 2009), Département Recherche de la Société Française de Psychologie. Nice : 28-30 septembre.

<sup>39</sup> Granka L.A., Joachims T., Gay G. (2004). Eye-tracking analysis of user behavior in WWW search. In *Proceedings of the 27th annual international ACM SIGIR conference on research and development in information retrieval*. Sheffield, Grande-Bretagne.

<sup>40</sup> Ravestein J., Ladage C., Johsua S. (2007). Trouver et utiliser des informations sur Internet à l'école : problèmes techniques et questions éthiques. *Revue française de pédagogie*, n° 158, p.71-83.

requête (pages personnelles, associations, fondations d'entreprises, etc.) et le bruit, c'est-à-dire un site qui ne traite pas du tout de la requête ou qui ne fonctionne pas.

Ils ont obtenu les résultats suivants, toutes requêtes confondues :

Moteurs	Site institutionnel contrôlé	Site privé	Bruit
Google	7	18	2
Yahoo	3	20	4
MSN	3	11	13
Total	23	78	34

Tableau 1: Répartition des premiers liens par genre, d'après l'étude de J. Ravestein et al. (2007)

Nous tirons de ces travaux les conclusions suivantes : premièrement, il est noté une grande quantité de bruit sur les liens en première position, ce qui est relativement gênant pour un jeune internaute ; deuxièmement que la fiabilité des ressources est discutable, puisque le lien n° 1 renvoie trois fois plus souvent sur des sites privés. Si, dans ces sites privés, l'information est souvent en rapport direct avec la requête, les chercheurs citent l'exemple pour le lien « démocratie » où six liens sur neuf renvoient à des sites de partis politiques.

De même, la présence de plus en plus grande du site Wikipedia, qui est l'exemple même d'une source instable, est remarquée. Par ailleurs, la requête « Jacques Prévert », renvoie huit fois sur neuf sur un site « personnel », à l'ergonomie défailante, difficilement exploitable par un élève d'école élémentaire. L'hypothèse qui ressort de ces travaux est qu'il y a une grande difficulté d'exploitation des ressources du Web par des enfants de l'école élémentaire sur des requêtes ordinaires avec les automates habituellement utilisés.

#### 2.4.2 L'examen des pages de résultats

Pourtant, la position du lien dans la page n'est pas le seul critère de pertinence appliqué ni même le premier mentionné, note Boubée (2008)<sup>41</sup>. Le critère le plus mentionné reste celui

---

<sup>41</sup> Boubée N. (2008). Les stratégies des jeunes chercheurs d'informations en ligne. *Question de communication*. Vol 14 p. 33-48.

de l'adéquation au thème de recherche. Cependant, la position du lien dans la page de résultats reste le deuxième critère le plus fréquemment cité. D'autres critères sont évoqués avec une fréquence de mention très réduite : information perçue comme « cohérente », type de document, étendue de l'information, conformité avec l'opinion personnelle, nouveauté (pour le chercheur d'information).

#### 2.4.2.1 Les pages de résultats consultées

Fluckiger C. et Bruillard E. (2008)<sup>42</sup> ont étudié le processus de navigation employé par des collégiens : une fois les mots-clés saisis et que le moteur de recherche affiche les résultats, les adolescents explorent alors systématiquement les différentes pages proposées par Google. Une dizaine de pages suffit généralement à collecter les informations désirées. Lors de ces recherches scolaires, ils notent qu'il est très rare que les jeunes s'aventurent au-delà de chaque page référencée dans le moteur de recherche : une fois cette page parcourue et les informations qui semblent pertinentes notées dans un cahier ou copiées dans un document Word, les élèves utilisent la fonctionnalité « page précédente » de leur navigateur pour retourner à la page de résultat de Google.

Ihadjadène et Chaudiron (2008)<sup>43</sup> reprennent plusieurs travaux portant sur l'appropriation des résultats fournis par les moteurs de recherche et leur validation par les usagers. Beitzel<sup>44</sup>, en 2006, montre que, dans le cas du moteur de recherche AOL, 79% des usagers ne regardent que le premier écran, 15% le second, et 6% trois écrans ou plus.

Lorigo et ses collègues (2006)<sup>45</sup> ont étudié le processus de lecture et de sélection dans les pages de résultats de recherche Google, par des étudiants. Dans 96% des requêtes, les

---

<sup>42</sup> Fluckiger C., Bruillard E. (2008). TIC : analyse de certains obstacles à la mobilisation des compétences issues des pratiques personnelles dans les activités scolaires. In Colloque international « *L'éducation à la culture informationnelle* », Lille 3, 16-18 octobre 2008.

<sup>43</sup> Ihadjadene M., Chaudiron S. (2008). Quelles analyses de l'usage des moteurs de recherche : Questions méthodologiques. *Question de communication*. N°14, p. 17-32.

<sup>44</sup> Beitzel S. M. (2006). *On Understanding and Classifying Web Queries*, PhD Thesis in Computer Science, Illinois Institute of Technology, May,

<sup>45</sup> Lorigo L. [et al.] (2006). The influence of task and gender on search and evaluation behavior using Google. *Information Processing and Management*, vol. 42, p. 1123-1131.

étudiants ne consultent que la première page. Aucun des participants n'a été jusqu'à la troisième page.

Aussi, les deux premiers résultats sont systématiquement regardés en premier et le premier lien est assez souvent cliqué.

Il a aussi été remarqué que les étudiants s'appuient sur les résumés pour décider de la sélection du lien, et lors de l'observation, plus de la moitié des participants ont préféré reformuler la requête sans avoir sélectionné aucun lien pour s'éviter la lecture de documents.

#### **2.4.2.2 La lecture des pages de résultats**

Au niveau de la lecture, Granka, Joachims et Gay (2004)<sup>46</sup> ont montré qu'en moyenne, seuls trois résumés sont lus par page. Ces travaux ont mis en avant que le temps de lecture dépend de la place de la page dans la liste de résultats : les références les plus hautes placées sont les plus lues. Il a aussi noté que les résumés placés de la 6<sup>ème</sup> à la 10<sup>ème</sup> place et qui s'affichent après un déroulement de la page de résultat sont aussi lus.

Ihadjadène et Chaudiron (2008)<sup>47</sup> reprennent les travaux menés par Pan et ses collègues (2007)<sup>48</sup>. Une analyse oculométrique (méthode du *eyetracking*) menée sur MSN et Google montre que si les internautes lisent linéairement les résultats d'un moteur de recherche, le temps de lecture décroît très rapidement entre le premier et le dernier résultat. La lecture du premier dure 2,2 secondes, contre 0,2 seconde à 0,3 seconde pour le dernier. Ce sont donc les premiers résultats qui captent le plus l'attention de l'internaute. La même étude révèle que les utilisateurs de l'internet accordent une confiance exagérée au système de classement des résultats utilisé par le moteur de recherche Google. Ce constat a des conséquences importantes du point de vue des enjeux commerciaux, ce qui explique que certaines techniques de référencement de sites utilisent le *spamming* pour améliorer artificiellement le positionnement de leurs sites dans les pages de résultats.

---

<sup>46</sup> Granka L.A., Joachims T., Gay G. (2004). Eye-traking analysis of user behavior in WWW search. In *Proceedings of the 27th annual international ACM SIGIR conference on research and development in information retrieval*. Sheffield, Grande-Bretagne.

<sup>47</sup> Ihadjadene M., Chaudiron S. (2008). Quelles analyses de l'usage des moteurs de recherche : Questions méthodologiques. *Question de communication*. N°14, p. 17-32.

<sup>48</sup> Pan B., [et al.] (2007). « In Google we trust: Users' decisions on rank, position and relevancy », *Journal of Computer-Mediated Communication, special issues on The Social, Political, Economic and Cultural Dimensions of Search Engines*, n° 12.

Blondel, Kempf et Schwob (2000)<sup>49</sup>, dans leurs observations menées auprès d'élèves de première scientifique lors de travaux pratiques, ont analysé comment les élèves répartissent le temps disponible entre les différentes tâches au cours des deux séances de recherche d'information. Cette analyse indique que les élèves consacrent plus de la moitié de leur temps de recherche (56 %) à la consultation des sites. La formulation des requêtes (16 %) et la consultation des résultats rendus par les moteurs ou les annuaires (20 %) occupent un peu plus du tiers de leur temps (36 % au total). Le reste correspond à des temps morts y compris les délais de réception des résultats ou des documents (8 %). Ils notent cependant que ces moyennes masquent une grande variabilité entre les groupes ; ainsi, le temps de formulation des requêtes (moyenne 16 %) varie entre 3 % et 25 %. Ils notent également que les élèves se contentent très souvent (85 % des requêtes) des premiers résultats, c'est-à-dire la première page retournée par le serveur, soit quinze références de sites pour l'annuaire Yahoo ou dix références de pages pour le moteur AltaVista. La page des résultats suivants n'est consultée que dans 10 % des cas, et les résultats qui suivent, à partir de la troisième, dans 5 % des cas seulement. Ce dernier chiffre est dû à la ténacité d'un des groupes qui est allé explorer en profondeur les résultats de l'une de ses requêtes.

### **2.4.3 Les indices de surface**

#### **2.4.3.1 Le marquage typographique**

Sur Internet, la majorité des moteurs de recherche mettent en valeur les termes correspondant à la requête (mise en gras, en majuscules etc.). Un tel marquage typographique est présent dans tous les moteurs de recherche ou du moins dans les plus utilisés. Les résultats obtenus par Dinet (2006)<sup>50</sup>, lors d'une tâche de sélection de références pertinentes, montrent que plus l'utilisateur est jeune, plus il est sensible au marquage typographique des références, que celles-ci soient pertinentes ou non. En d'autres termes, les participants les plus jeunes sélectionneraient les références, en partie, sur la base d'indices visuels.

---

<sup>49</sup> Blondel M.F, Kempf O., Schwob M. (2000). Recherche d'information sur la toile : Pratiques d'élèves. *Bulletin de l'Union des physiciens*, n°828, vol. 94, p.1819-1846.

<sup>50</sup> Dinet J. (2006). Le choix des sites Web par les enfants et les adolescents : impacts de la typographie. In Piolat A. *Lire, écrire, communiquer et apprendre avec Internet*. Marseille, Solal, p.135-150.

### 2.4.3.2 L'importance du titre et des résumés

D'autre part, de nombreux travaux comme ceux de Bilal (2002)<sup>51</sup> ont mis en avant que les jeunes regardent en premier le titre pour évaluer l'intérêt d'un document. Or, dans les moteurs de recherche, les titres et les résumés sont souvent tronqués, ce qui peut induire en erreur les jeunes usagers.

Les travaux de Blondel, Kempf et Schowb (2000)<sup>52</sup> ainsi que les observations qu'ils ont mises en place montrent que, dans les quelques discussions entre les membres du groupe, la part consacrée à l'analyse des différents critères affichés (indice de pertinence, type de site, pays d'origine...) est réduite. Il semble que la plupart du temps seule la description sommaire fournie par le serveur soit lue rapidement.

En effet, de nombreux travaux ont montré que les jeunes évaluent rapidement. Wallace et ses collègues (2000)<sup>53</sup> notent dans leurs travaux que les élèves de 6<sup>ème</sup> ont une moyenne de 28 secondes pour évaluer le contenu d'un site web.

## 3. Les raisonnements mis en place lors d'une recherche d'information

Différents raisonnements sont mis en place lors d'une recherche et nous pouvons ainsi distinguer deux types de raisonnement.

### 3.1 Le raisonnement heuristique

La procédure heuristique<sup>54</sup> consiste à délimiter un ensemble d'hypothèses possibles. Cette élaboration d'hypothèses repose sur deux processus : la recherche en mémoire de

---

<sup>51</sup> Bilal D. (2002). Children's use of Yahoo! search engine : Cognitive and physical behaviors on fully self-generated search tasks. *Journal of the American Society for Information Science and Technology*, vol.53, n°13, p. 1170-1183.

<sup>52</sup> Blondel M.F, Kempf O., Schwob M. (2000). Recherche d'information sur la toile : Pratiques d'élèves. *Bulletin de l'Union des physiciens*, n°828, vol. 94, p.1819-1846.

<sup>53</sup> Wallace R. [et al.] (2000). Science on the Web : Students online in a sixth-grade classroom. *The Journal of the Learning Sciences*, vol. 9, n°1, p. 75-104.

connaissances relatives au phénomène étudié et l'analyse d'un certain nombre d'éléments de la situation présentée. Les sujets partent de l'analyse de l'énoncé du problème. Certains éléments sont sélectionnés, ce qui amène les élèves à rechercher en mémoire des événements plus ou moins analogues au phénomène étudié, ainsi que des connaissances en rapport avec les concepts évoqués à partir des éléments de la situation.

Ainsi, selon le niveau de connaissances dont il dispose, le sujet se posera plus ou moins de questions sur la description qu'il convient de faire de la situation étudiée.

Souvent, les étapes consistent à créer ou imaginer une situation, puis à rechercher l'information correspondante. Au vu des résultats, le sujet peut alors décider de garder son hypothèse, et éventuellement de la tester par la suite pour vérifier qu'il s'agit bien de la « bonne » hypothèse, soit de changer d'hypothèse

A. Tversky, D. Kahneman (1999)<sup>55</sup> différencie trois heuristiques :

- L'heuristique de représentativité qui consiste à juger un événement particulier en le rattachant à une classe d'événements similaires connus antérieurement.
- L'heuristique de disponibilité où les individus jugeraient la fréquence ou la probabilité qu'un événement se produise selon la facilité qu'ils ont à se rappeler d'exemples de situations où cet événement s'est effectivement produit.
- L'heuristique de l'ancrage-ajustement où les individus jugent d'un événement en le rattachant à un autre événement connu, qui sert d'« ancre ». La valeur initiale de l'ancre sert alors de point de départ pour le réajustement du jugement sur l'événement en question.

Ce mode de raisonnement, intuitif, davantage basé sur l'expérience ou les connaissances antérieures, et qui témoigne d'un manque de logique, peut entraîner des erreurs d'appréciation. Mais il permet d'évaluer rapidement une situation, de la traiter en minimisant l'effort cognitif et en conséquence n'est pas exempt d'efficacité. On pourrait dire qu'il s'agit d'un raisonnement fait d'automatismes.

---

<sup>54</sup> Cauzinille-Marmèche Evelyne, Mathieu Jacques, Weil-Barais Annick (1982). Le raisonnement expérimental des pré-adolescents : stratégies de test d'hypothèses. In: *Enfance*. Tome 35 n°1. pp. 23-37.

<sup>55</sup> A. Tversky, D. Kahneman (1999). *Judgement under uncertainty : heuristic and biases*. Cambridge University Press, p.3-15.

### 3.2 Le raisonnement analytique

Le raisonnement analytique correspond à une procédure de preuves qui a pour but d'aboutir à une conclusion sur l'effet d'un ou de plusieurs facteurs. Il s'agit d'un mode de raisonnement plus réfléchi comme le précisent Metzger et al. (2010)<sup>56</sup>.

La littérature scientifique a également mis en avant l'implication des émotions sur le raisonnement. En effet, l'état émotionnel des individus influence la façon qu'ils ont d'utiliser leurs connaissances dans un raisonnement. De la même façon, l'état émotionnel est présent dans l'évaluation de l'information. Ainsi, Boubée (2012)<sup>57</sup> rappelle les travaux de Lopatovska et Mokros (2008) qui relèvent une dimension affective dans les jugements de pertinence mais cette dernière est non influente dans la sélection de documents, l'utilité pour la tâche de recherche prédominant.

Pour conclure, il existe différents types de connaissances qui sont classées en trois catégories : les connaissances déclaratives, les connaissances procédurales et les connaissances conditionnelles. Les connaissances déclaratives constituent le savoir théorique, les connaissances procédurales renvoient au savoir-faire et les connaissances conditionnelles concernent les conditions de l'action.

Il s'agit ici de s'intéresser à l'influence des connaissances du domaine sur le processus de sélection lors d'une recherche d'information scolaire. L'intérêt de nos travaux réside dans le fait que les connaissances du domaine sont peu prises en compte dans les travaux relatifs à la recherche d'information. Nos travaux reposent sur l'entièreté du processus de sélection c'est-à-dire sur l'évaluation de la page de résultats mais également du document primaire, thématique également peu prises en compte dans les travaux scientifiques.

---

<sup>56</sup> Metzger, M. J., Flanagan, A. J., & Medders, R. B. (2010). Social and Heuristic Approaches to Credibility Evaluation Online. *Journal of Communication*, 60(3), 413–439.

<sup>57</sup> Boubée N. (2012). Le cœur, (la tête), et l'information recherchée : Le rôle des émotions dans l'activité de recherche d'information. In 9<sup>e</sup> Congrès de la Fadben, *Objets documentaires numériques : nouvel enseignement ?*, Paris, 24 mars.

C'est pourquoi, en accord avec le cadre théorique exposé ci-dessus, nous faisons l'hypothèse que les connaissances déclaratives antérieures ont un impact sur la sélection des références pertinentes lors d'une recherche d'information sur Internet. Aussi, nous supposons que les experts du domaine utiliseront plutôt une procédure analytique et les novices du domaine une procédure heuristique.

Hypothèse 1 : les élèves utilisant un raisonnement analytique réfléchissent à leur besoin d'information. *A contrario*, les élèves utilisant un raisonnement heuristique reprennent les mots de la consigne pour formuler leur requête.

Hypothèse 2 : les élèves utilisant un raisonnement analytique lisent les résumés des pages de résultats et ne s'appuient pas sur des critères typographiques. *A contrario*, les élèves utilisant un raisonnement heuristique sélectionnent les différents liens de la page de résultats en s'appuyant sur des critères typographiques.

Hypothèse 3 : les élèves utilisant un raisonnement analytique s'appuie sur des critères de fiabilité de la source lors de la consultation du document primaire (contenu informationnel mais aussi la source (auteur, éditeur)). *A contrario*, les élèves utilisant un raisonnement heuristique s'appuient sur des critères d'évaluation visuels comme l'apparence du site (couleurs, structure etc.).

## Partie 2 : Dispositif méthodologique

Nous souhaitons donc étudier l'influence des connaissances du domaine sur les sites sélectionnés lors d'une recherche d'information ainsi qu'observer les raisonnements mis en place. Dans nos travaux, nous nous focalisons sur l'ensemble du processus d'évaluation et de sélection, c'est-à-dire,

- A la première évaluation, celle de la page de résultats, puisque cette dernière conditionne tout ou du moins une partie de la réussite de la recherche d'information.
- A l'évaluation du document primaire sélectionné.

Après avoir produit une requête et afin de répondre à une question posée, l'élève est amené à sélectionner une ou plusieurs références parmi une liste de références générée par le moteur de recherche Google.

### 1. L'échantillon

Afin de vérifier notre hypothèse, nous avons choisi de travailler avec deux types de chercheurs d'information :

- Des experts du domaine, c'est à dire possédant des connaissances approfondies dans un domaine. Ces derniers, au nombre de 5, se sont évalués comme expert dans le domaine de leur recherche au travers d'une échelle d'évaluation allant de 0 à 5.
- Des novices du domaine, se définissant comme ne possédant aucune connaissance approfondie dans les domaines de leur recherche. Un seul participant s'est évalué comme novice du domaine, grâce à une échelle d'évaluation allant de 0 à 5.

Ainsi, notre échantillon comprend 6 participants c'est-à-dire 5 experts et 1 novice dans différents domaines.

	Garçons	Moyenne d'âge
--	---------	---------------

Experts	5	De 15 à 16 ans
Novices	1	

Tableau 2 : Détail des participants

Le nombre de participants experts et novice est très inégal. De ce fait, nous devrions ne retenir que les données portant sur l'activité des experts. Cependant le comportement du novice s'avérant radicalement différent, nous conservons les données du novice.

Nous savons qu'il est nécessaire de prendre de nombreuses précautions pour choisir les participants et éviter ainsi certains biais.

Dans notre observation, chaque participant a été volontaire, car ceux qui se seraient sentis obligés de participer auraient pu constituer un biais. De plus, ils ne devaient pas avoir de contrainte de temps pour ne pas qu'ils se pressent à finir la tâche que nous leur soumettions.

Aussi, au niveau de la représentativité de l'échantillon, certaines caractéristiques sont plus présentes que d'autres et d'autres sont absentes. Au niveau du genre par exemple, l'ensemble des participants sont de sexe masculin.

D'autre part, afin de rechercher une homogénéité avec les autres connaissances intervenant dans une recherche d'information sur le web, à savoir les connaissances techniques et informationnelles, les participants sont tous issus d'une classe de Seconde d'un lycée général et technologique et possèdent tous entre 15 et 16 ans.

Enfin, nous savons que les participants se seraient sentis valorisés d'être observés et donc ils auraient essayé d'employer des stratégies de recherche adaptées en suivant les conseils donnés par des enseignants.

Nous avons donc choisi d'utiliser plusieurs méthodes, une expérimentation et un entretien donc une observation directe et une observation indirecte car la triangulation des méthodes va permettre de palier les limites de chacune d'entre elles. Enfin, l'auto évaluation des participants sur leurs propres connaissances nous permet de juger leur expertise dans le domaine de la tâche à effectuer.

## **2. Les tâches de recherches d'information**

Pour mettre en place les tâches des différentes recherches d'information, nous avons décidé de ne pas contrôler différents facteurs afin que la situation de recherche d'information soit la plus proche d'une situation naturelle.

### **2.1 Le choix du moteur de recherche**

Nous avons décidé de laisser au participant le choix du système de recherche d'information à utiliser pour leur recherche. La totalité des participants a choisi d'utiliser le moteur de recherche Google. Le Web constitue la source privilégiée d'information et de documentation des adolescents et le moteur de recherche Google est la porte d'accès presque unique à la recherche sur Internet pour les jeunes.

Effectivement, Google est le moteur de recherche le plus utilisé en France, d'après le baromètre d'avril 2012 de AT Internet Institute. Google obtient 90% en parts de visites Web et se place loin devant d'autres moteurs de recherche comme Bing (3,1%) ou Yahoo (1,4%) [Annexe 1].

### **2.2 Présentation des pages de résultats**

La liste des résultats de la requête était conforme au format de présentation propre aux moteurs de recherche les plus utilisés. Ainsi, chaque référence comportait successivement, de haut en bas, les informations suivantes :

- Un titre, affiché et souligné en bleu, correspondant au titre de la page Web
- Une adresse, affichée en vert, correspondant à l'adresse URL de la page Web
- Un résumé, affiché en noir, correspondant à des extraits de textes de la page Web.

Les mots présents dans la requête étaient affichés en gras, dans les titres comme dans les résumés.

### **2.3 Les sujets des recherches d'informations**

Les participants doivent effectuer une recherche en lien avec les programmes scolaires. Ainsi, toutes les recherches portant sur l'orientation ou les domaines de loisir ne sont pas retenues. La totalité des recherches observées sont effectuées à la demande d'un de leurs enseignants, dans le cadre d'un cours. Les participants ont donc des consignes précises, données par l'enseignant, et un travail écrit à rendre. Chaque travail nécessite une recherche d'information pour répondre à la consigne.

<b>Participants</b>	<b>Connaissance du domaine</b>	<b>Tâches de recherche d'information</b>	<b>Type de recherche d'information</b>
Participant 1	Expert (4 sur 5)	ECJS : Recherche sur les réformes judiciaires en France	Recherche exploratoire
Participant 2	Expert (4 sur 5)	ECJS : Recherche sur l'état de la peine de mort dans le monde	Recherche exploratoire
Participant 3	Expert (5 sur 5)	ECJS : Recherche sur le contexte de la mise en place de la loi Badinter	Recherche exploratoire
Participant 4	Expert (5 sur 5)	ECJS : Recherche sur la tuerie en Norvège en 2011	Recherche factuelle
Participant 5	Expert (4 sur 5)	ECJS : Recherche sur la législation en Norvège	Recherche factuelle
Participant 6	Novice (2 sur 5)	Français : Recherche sur la critique de films	Recherche factuelle

Tableau 3 : Thèmes des recherches d'informations par participant

### **3. La procédure**

Nous avons donc observé tâches différentes tâches de recherches d'informations.

Les connaissances du domaine sont évaluées par les participants eux mêmes. En amont de la tâche de recherche d'information, l'élève évalue ses connaissances du domaine sur une échelle allant de 0 à 5. Les participants s'étant évaluant entre 0 et 2 compris sont considérés comme novices du domaine et les participants s'étant évalués de 3 à 5 compris sont considérés comme experts du domaine. Ainsi, parmi les cinq participants experts, trois ont évalué leurs connaissances à 4 sur 5 et deux ont évalué leurs connaissances à 5 sur 5. Le participant novice a évalué ses connaissances à 2 sur 5. Cette méthode comporte un biais, liés au jugement que porte l'élève sur ses propres connaissances. Ce jugement est également influencé par notre présence :

- Certains peuvent s'être surévalués, par crainte d'un jugement de valeur sur leur personne.
- Certains peuvent s'être sous-évalués, par crainte de ne pas représenter un niveau de connaissance assez élevé.

Les consignes sont données avant que la page de résultat ne soit affichée : le participant n'a pas de temps limite pour achever la recherche d'information, il peut aussi retourner sur la page de résultats autant de fois qu'il le souhaite. Il arrête sa recherche d'information lorsqu'il le souhaite.

Chaque recherche d'information se déroulait de la façon suivante :

- l'élève lit la consigne donnée par son enseignant,
- choisit un système de recherche d'information pour répondre à son besoin d'information,
- produit des requêtes,
- sélectionne un ou plusieurs liens de la page de résultats affichée,
- Consulte le document primaire.

Le participant a la possibilité de revenir à la page de résultats si le document consulté ne lui permet pas de répondre à la question posée. Il peut également changer sa requête si nécessaire et consulter plusieurs documents primaires afin de répondre à sa tâche de recherche d'information.

#### **4. Le matériel utilisé**

Durant la tâche de recherche d'information, nous nous sommes concentrés sur plusieurs points à l'aide d'une grille d'observation [Annexe 5]. Au niveau des pages de résultats, nous avons observé le nombre de pages consultées et la place qu'elles occupaient dans la totalité des pages de résultats. Nous avons aussi fait attention au nombre de liens qui étaient sélectionnés avant de pouvoir répondre à la question posée et à la place que ces liens occupaient dans la page de résultats. Nous souhaitions au départ prendre en compte une mesure de temps, mais qui s'est révélée incompatible avec la méthode du *think aloud* décrite ci-dessous.

Pour chaque tâche de recherche, nous avons observé le comportement de navigation de l'élève, comme l'utilisation des aperçus des pages web sur la page de résultats, et pris des notes. Chaque participant était encouragé à « penser tout haut », selon la méthode du *think aloud* et les réflexions ont été enregistrées à l'aide d'un dictaphone puis retranscrites.

A la fin de la recherche, nous avons mené avec chaque participant un entretien semi-directif qui a été également enregistré à l'aide d'un dictaphone et retranscrit. Ces entretiens nous ont permis de questionner sur le choix du système de recherche d'information utilisé, sur l'ensemble des critères de pertinence mis en place lors du processus de sélection et enfin sur les raisonnements et les stratégies employées lors de la recherche d'information. [Annexes 2 et 3].

# Partie 3 : Résultats

Suite aux observations et entretiens menés auprès des participants, nous présenterons les résultats relatifs à la formulation de la requête, à la sélection dans la page de résultats mais également aux critères de pertinence mentionnés lors de la sélection d'un lien dans la page de résultats et d'un document primaire.

Pour identifier les modes de raisonnement, nous nous appuyons des actions spécifiques ayant un lien avec la pertinence comme la réflexion sur le besoin d'information, les requêtes formulées et sur l'identification des critères de pertinence au niveau de la page de résultats et des documents primaires.

## 1. Résultats relatifs à la formulation de la requête

### 1.1 Définition du besoin d'information

Lors de la lecture de la consigne, seuls les experts du domaine ont réfléchi à l'établissement d'une stratégie de recherche. Un des participants explique lors de l'entretien : *« je savais déjà ce que je cherchais dès le début et je me suis rappelé que sur Wikipédia il y avait des informations sur ce sujet là »*. Nous avons également noté que les experts du domaine revenaient sur la consigne au cours de leur recherche *« c'est quoi la consigne déjà ? »*. Au contraire, le novice a commencé sa recherche sans savoir quoi chercher, il nous confie *« je suis tombé sur un site de méthodologie même si je cherchais pas ça au départ, j'y suis tombé par hasard et ça m'a bien aidé. Alors qu'au départ, je voulais trouver un exemple de critique de film, j'avais pas pensé que la méthodologie sur comment faire la critique serait plus utile »*.

### 1.2 Formulation de la requête

Nous avons noté que les experts du domaine ont modifié leur requête moins souvent que le novice du domaine. Les experts du domaine ont, en moyenne, formulé deux requêtes lors de leur tâche de recherche d'information alors que le novice en a formulé quatre.

De manière générale, chaque requête, que ce soit pour les experts ou le novice du domaine, reprenait au moins deux termes de la consigne. Seul un participant, un expert du

domaine, a saisi l'entièreté de la consigne dans la barre de recherche. Il justifie cette stratégie ainsi : « *je veux gagner du temps, au cas où la question ait déjà été posée, sur un forum par exemple, et comme ça je vais pouvoir recopier directement la réponse* ».

Participants	Nombre de requête au cours de la RI
Expert 1	2
Expert 2	1
Expert 3	2
Expert 4	1
Expert 5	4
Novice 1	4

Tableau 4 : Nombre de formulation de requête par participant

## 2. Résultats relatifs à la sélection dans la page de résultats

### 2.1 Pages de résultats du moteur de recherche consultées

Nous avons constaté que tous les participants, quel que soit le niveau de connaissances du domaine exprimé par les participants, n'ont consulté qu'une page de résultats et uniquement la première qui est affichée par le moteur de recherche Google. Les participants ont préféré modifier leurs requêtes plutôt que consulter les pages de résultats suivantes. Lors des entretiens, les participants ont explicité ce choix : « *je reste sur la première page car si on s'éloigne trop ça n'a plus rien à voir* » (Participant 1).

### 2.2 Proportion de références sélectionnées par thème de recherche par les experts et le novice

Nous constatons que le novice a consulté plus de liens sur la page de résultats que les experts. Ces derniers n'ont consulté en moyenne que 3 liens.

	Nombre de liens consultés

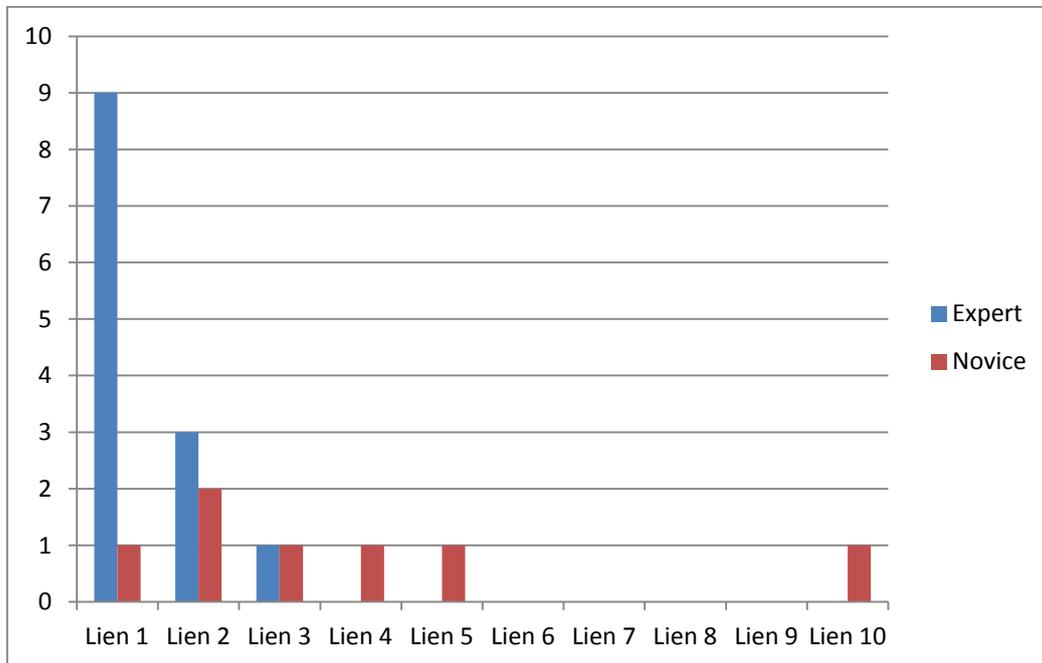
Experts	3
Novices	7

Tableau 5: Moyenne de liens sélectionnés dans la page de résultats

### 2.3 La place des liens sélectionnés dans la page de résultats durant la tâche de recherche d'information

Nous remarquons que les liens les plus consultés sont les liens placés en première, deuxième et troisième position. Effectivement, la totalité des participants a consulté le premier lien de la page de résultats au moins une fois.

De plus les participants experts du domaine ont consulté uniquement les liens apparaissant en première, deuxième et troisième position. Au contraire, le novice du domaine a également consulté les premiers liens mais il est le seul à avoir consultés les liens de la 4<sup>ème</sup> à la 10<sup>ème</sup> position.



Graphique 1: Place des liens consultés dans la première page de résultats

De manière générale, la totalité des participants ont sélectionné en premier le lien apparaissant en première position. Ce choix a été justifié par l'ensemble des participants, le

participant 1 explique par exemple que « *s'il est placé en premier c'est que c'est le plus visité donc le meilleur* ». Ce critère de popularité a été pris en compte par l'ensemble des participants, expert et novice du domaine.

### 3. Justification des sélections : les critères de pertinence

Les entretiens que nous avons eus avec chaque participant ont permis de mettre en avant plusieurs critères qu'ils ont appliqués lors de la sélection d'un lien sur la page de résultats ainsi que dans le document primaire. Les différents critères mentionnés sont issus de la littérature scientifique. Toutefois, nos travaux font apparaître de nouveaux critères, mis en gras dans les tableaux ci-dessous.

#### 3.1 Experts du domaine et page de résultats

Plusieurs critères sont pris en compte lors de la sélection de liens dans la page de résultats. Ces jugements sont prédictifs, c'est-à-dire qu'ils vont permettre de décider si une source doit être ou non consultée, et les participants ont mentionné, après l'adéquation au thème, la position du lien dans la page de résultats. En effet, les liens positionnés en première, seconde et troisième place sont préférés.

La source est également prise en compte. Sont évoqués l'encyclopédie Wikipédia qui est considérée par les participants comme « *site de référence* ». Un d'entre eux explique : « *Wikipédia, c'est complet, et ça me permet de trouver la réponse à ma question presque à chaque fois.* ».

De manière générale, les participants ne témoignent pas d'une réelle stratégie de sélection c'est-à-dire une lecture très précise de la page de résultats, avec une utilisation des aperçus de page web par exemple. Ils expliquent, le plus souvent, choisir les liens « *dans l'ordre* » et de vouloir « *tous les essayer* ». Cette pratique témoigne peut être d'un excès de prudence de la part des participants puisqu'il est difficile de porter un jugement de valeur uniquement à partir du résumé proposé.

Critères	Mentions
Adéquation au thème	4

Position du lien	2
Source	2
Précision du vocabulaire	1

Tableau 6 : Justification des sélections sur la page de résultats

### 3.2 Experts du domaine et documents primaires

Les observations menées auprès de 5 participants experts du domaine ont mis en avant que divers critères permettent de juger un document comme pertinent. Au contraire, un seul critère est utilisé pour éliminer un document.

Les critères les plus cités par les participants experts sont les suivants : l'adéquation au thème, la source, la présence d'image et l'esthétique du site. L'émotion produite par la lecture d'un document s'est également révélée être un critère de pertinence dans la sélection ou non du document mais uniquement pour les tâches de recherche d'information liées à un thème de recherche qui a engendré une émotion chez le participant. Par exemple, un participant qui travaillait sur la fusillade en Norvège en 2011 explique : « *C'est un sujet grave et là il y a des images chocs (cf. victimes de la fusillade), du coup ça me donne envie de continuer sur ce site et de lire l'article* ».

L'étendue du document est également prise en compte. Pourtant, ce critère s'est révélé être soit positif si le document était complet et synthétique et a donc permis la sélection du document ; soit négatif si le document était jugé trop long.

Le niveau de langue est également un critère souvent évoqué. Cependant, nos travaux font apparaître une différenciation : le fait que le document primaire ait un niveau de langue compréhensible n'est pas le seul critère. Les participants prennent également en compte que le niveau de langue soit adéquat à un usage futur du document, que l'élève soit désapprouvé par l'enseignant, à savoir la recopie du document (nous pourrions nommer ce critère « bon à copier-coller »). « *Le niveau de langue est beaucoup trop élevé... La prof verra tout de suite que c'est pas de moi si je le copie* » confie un des participants. En effet, les observations et les entretiens menés ont révélé une forte pratique du copié collé, sans reformulation de l'information prélevée. De cette pratique découle donc ce nouveau critère.

De la même façon, nos travaux font apparaître un même critère relatif à l'exploitation de l'information prélevée et associée à un critère temporel. Les participants jugent un document

primaire pertinent s'il leur permet de gagner du temps dans la production du travail final. Un des participant explique par exemple : « *grâce à ce site, j'ai pu à remplacer ce qui a été fait [cf. une critique de film rédigée] avec ce qui me concerne, comme ça je perdrais pas du temps à formuler mes phrases après* ».

Plusieurs autres critères, évoqués dans la littérature scientifique, comme la cohérence, la conformité avec l'opinion personnelle ou encore la nouveauté, n'ont pas été évoqués lors des observations et des entretiens par les participants experts.

Critères	Mentions
Adéquation au thème	4
Source	2
Type de document	1
Présence d'image	2
Emotion	2
Etendue du document : complet et synthétique	3
Etendue du document : complet et trop développé par rapport au BI	2

Niveau de langue compréhensible	3
<b>Usage futur : bon à copier-coller</b>	<b>2</b>
Précision du vocabulaire	1
Esthétique du site	2
<b>Correspondance avec la consigne donnée et aux attentes scolaires</b>	<b>3</b>
<b>Gain de temps pour la production future</b>	<b>2</b>

Tableau 7 : Justification des sélections des documents primaires par les experts

### **3.3 Comparaison des critères de pertinence dans le processus de sélection mentionné par les experts pour la page de résultats et pour le document primaire**

Tout d'abord, le nombre de critères de pertinence évoqué est inégal entre la page de résultats et les documents primaires. Les experts du domaine ont mentionné 4 critères pour justifier les choix effectués sur la page de résultats alors qu'ils en ont évoqué 12 pour justifier les choix sur les documents primaires.

Certains critères sont cités uniquement pour la page de résultats, comme la position du lien et la précision du vocabulaire ; alors que d'autres critères ne sont évoqués que pour la justification du document primaire :

- Pour tous les experts, les critères relatifs au contenu (type de document, présence d'images, émotion, étendue du document, niveau de langue), au site web

(esthétique) et liés à la consigne donné par l'enseignant et au travail final à rendre ne sont mentionnés que pour les documents primaires.

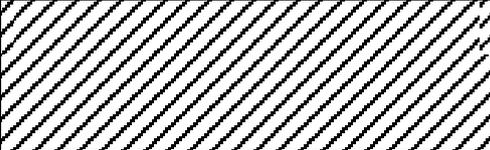
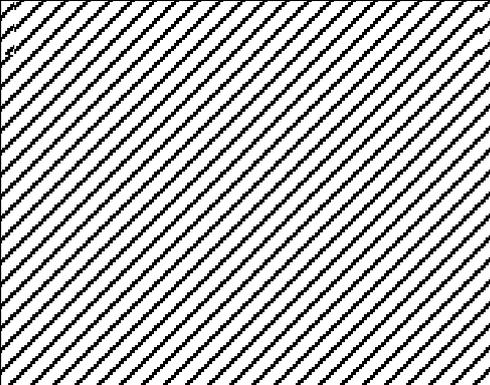
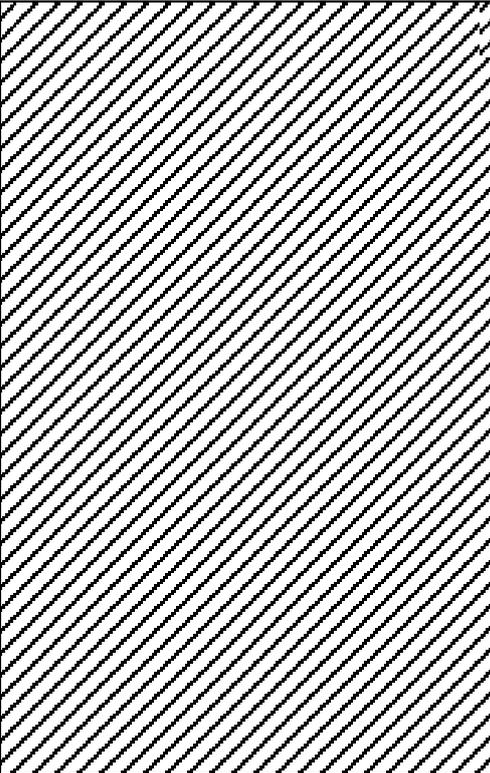
Certains critères apparaissent dans la justification des deux types de sélection, comme l'adéquation au thème et la fiabilité de la source.

En résumé : les experts du domaine mentionnent trois fois plus de critères pour la sélection de document que pour la sélection de la page de résultats. Dans la page de résultats, le critère lien est important puisqu'il est le plus cité après l'adéquation au thème. La fiabilité de la source est citée dans la sélection des 2 types de documents, primaire et secondaire (page de résultat). Le travail de sélection du document primaire ne se fait généralement pas sur des jugements prédictifs puisque les experts du domaine semblent préférer juger le contenu du document primaire plutôt que le résumé proposé par la page de résultats.

Aussi, plusieurs critères évoqués dans la littérature scientifique, comme la cohérence, la nouveauté ou encore la conformité avec l'opinion personnelle n'ont pas été relevés par les experts du domaine lors de nos observations et entretiens.

<b>Expert</b>		
<b>Critères de pertinence</b>	<b>Page de résultat</b>	<b>Document primaire</b>
Adéquation au thème	<p>« <i>Le titre a l'air de correspondre avec ce que je cherche</i> »</p> <p>« <i>Ca parle de mon sujet</i> »</p>	<p>« <i>Ce site correspond bien à ce que je cherche</i> »</p> <p>« <i>Ca n'a rien à voir avec ce que je cherche : c'est une biographie de Christine Taubirat alors que moi je cherche sur la réforme des prisons</i> ».</p> <p>« <i>Ca correspond pas du tout à ce que je cherche, mais alors vraiment pas du tout !</i> »</p> <p>« <i>Ca correspond bien à ma question</i> »</p>

Nombre de mention	4	4
Position du lien	<p>« C'est le premier lien donc je pense que je vais y trouver la réponse ».</p> <p>« S'il y a pas la réponse sur le premier, j'irai sur le second »</p>	
Nombre de mention	2	
Cohérence		
Source	<p>« Ce qui m'a plut ici c'est que ça soit Wikipédia. Pour moi, c'est un site de référence ».</p> <p>« C'est le site du Monde donc c'est bien car c'est un site de référence ».</p>	<p>« C'est un politicien qui parle donc déjà ça correspond à ce que je cherche puisque moi je m'intéresse au contexte politique et du coup lui il est bien placé pour en parler ! »</p> <p>« Wikipédia, c'est complet, et ça me permet de trouver la réponse à ma question »</p>
Nombre de mention	2	2
Type de document		« Il y a une carte et ça reprend en plus le texte donc c'est plus facile à comprendre »

Nombre de mention		1
Présence d'image		« <i>Il n'y a même pas d'image !</i> »  « <i>Les images donnent aussi envie de lire cet article car ce sont des images chocs</i> » (cf Tuerie en Norvège)
Nombre de mention		2
Emotion		« <i>C'est un sujet grave (relatif à la fusillade en Norvège) et là il y a des images chocs (victimes de la fusillade) du coup ça me donne envie de continuer sur ce site et de lire l'article</i> ».  « <i>Cet article est bien parce qu'il y a des témoignages des gens et des proches qui ont vécu ça (recherche sur la fusillade en Norvège). Ça me fait de la peine pour eux et j'ai envie de lire tous les témoignages qu'il y a</i> ».
Nombre de mention		2

Conformité avec l'opinion personnelle			
Etendue		<b>Positif</b>	<b>Négatif</b>
		<p>« Ce site c'est trop précis : ça concerne qu'un pays (sur la peine de mort) alors que moi je cherche tous les pays, un site qui montrera par continent par exemple ».</p> <p>« C'est complet : il y a un classement par continent avec une explicitation pour chaque pays »</p> <p>« C'est complet, il y a tout ce que je cherche qui est expliqué. »</p> <p>« Tout est</p>	<p>« C'est trop long, il y a trop d'écriture »</p> <p>« Je préfère changer car là c'est trop long à lire ».</p>

		<i>marqué ! »</i>	
Nombre de mention		6	
Nouveauté			
Niveau de langue compréhensible		« Sur ce site c'est bien expliqué, on comprend tout ». (x2) « C'est écrit par des lycéens donc on comprend tout, c'est à mon niveau ».	
Nombre de mention		3	
<b>Usage futur : bon à copier-coller</b>		« Le niveau de langue est beaucoup trop élevé... La prof verra tout de suite que c'est pas de moi si je le copie ».	
<b>Nombre de mention</b>		<b>1</b>	
Précision du vocabulaire	« Je choisis le lien 2 car le mot « fusillade » a attiré mon attention et je travaille sur la tuerie qu'il y a eu en Norvège et je sais que ça a été une fusillade donc c'est plus précis que tuerie »		

Nombre de mention	1	
Esthétique		<p>« Le texte est trop serré, il y a pas d'espace entre les lignes et du coup c'est difficile à lire ».</p> <p>« C'est structuré en paragraphe et les titres et sous titres sont en gras ce qui facilite la lecture. »</p>
Nombre de mention		2
<b>Correspondance avec la consigne donnée et aux attentes scolaires</b>		<p>« Ca répond à la consigne : il y a la différenciation entre la peine de mort légale et là où elle est abolie, et avec des exemples »</p> <p>« Ca répond exactement à la consigne de l'exercice : le contexte politique est clairement mentionné »</p> <p>« Sur ce site je reconnais une image qui est tirée d'un film qu'on a vu en classe donc ca montre que ce site c'est bien »</p>
<b>Nombre de mention</b>		<b>3</b>

<b>Gain de temps pour la production future</b>		<p>« <i>Ce site est bien car il reprend toutes les informations que j'avais trouvé sur plusieurs sites alors si je l'avais trouvé plus tôt j'aurais gagné du temps</i> ».</p> <p>« <i>Grâce à ce site, j'ai qu'à remplacer ce qui a été fait [une critique de film rédigée] avec ce qui me concerne comme ça je perdrais pas du temps à formuler mes phrases après</i> ».</p>
<b>Nombre de mention</b>		<b>2</b>

Tableau 8: Expression des justifications des sélections par les experts

### 3.4 Comparaison des critères de sélection entre les experts et les novices du domaine.

Nous pouvons comparer les résultats obtenus lors des observations et des entretiens menés avec les participants experts du domaine avec ceux obtenus auprès du novice du domaine [Annexe 4].

Lors de l'examen de la page de résultats, l'évaluation ne porte pas sur le résumé mais sur le titre : « *Le titre est pertinent, c'est-à-dire qu'il correspond à ce que je cherche* » et la position du lien est plus prise en compte que chez les experts du domaine.

Pour le choix d'un document primaire, les critères de sélection majoritaires cités par le novice du domaine reposent à la fois sur le contenu (adéquation au thème de recherche, cohérence) mais aussi sur des critères d'évaluation superficiels comme l'esthétique du site. Ce dernier explique par exemple : « *J'aime pas la couleur donc ça va pas aller, ça me donne pas envi de continuer sur ce site et de tout lire* ». La conformité avec l'opinion personnelle est également prise en compte, contrairement aux experts du domaine. En ce qui concerne

l'étendue du document, nous avons noté qu'un document jugé trop long par le participant est discriminé, même s'il est en adéquation avec le thème de recherche.

Enfin, un des critères mentionné comme incontournable par le novice est que le document corresponde exactement à la consigne donnée par l'enseignant et correspondent aux critères d'évaluation du travail final. Ainsi, le participant novice nous confie : « *Il faut que ça correspondent avec ce qu'a demandé la prof! Elle nous a donné des consignes précises sur comment on doit rédiger une critique alors faut que je vérifie que cet exemple il corresponde avec ce qu'elle elle nous a demandé* ».

Il convient également de noter que la fiabilité de la source, c'est-à-dire de l'auteur ou de l'éditeur, n'a jamais été mentionnée par le novice du domaine, que ce soit lors de l'examen de la page de résultats ou du document primaire.

## Partie 4 : Discussion

Nos travaux visent à mettre en avant les critères de pertinence et les raisonnements mis en place dans le processus de sélection, lors d'une recherche d'information. Les observations et entretiens menés auprès de 5 élèves s'étant évalués comme experts du domaine et d'un élève s'étant évalué comme novice du domaine ont souligné de nouveaux critères de pertinence. Cependant, les résultats comportent des biais puisque l'échantillon n'est pas représentatif.

Pour identifier les modes de raisonnement, nous nous appuyons des actions spécifiques ayant un lien avec la pertinence comme la réflexion sur le besoin d'information, les requêtes formulées et sur l'identification des critères de pertinence au niveau de la page de résultats et des documents primaires.

Hypothèse 1 : les élèves utilisant un raisonnement analytique réfléchissent à leur besoin d'information. *A contrario*, les élèves utilisant un raisonnement heuristique reprennent les mots de la consigne pour formuler leur requête.

Les experts du domaine ont pris le temps de réfléchir à leur besoin d'information avant de saisir les mots clés, renvoyant ainsi à un raisonnement analytique. Ce constat renvoie aux résultats obtenus par S. Brand-Gruwel, I. Wopereis, et Y. Vermetten (2005)<sup>58</sup> puisque ces chercheurs ont mis en avant que les experts du domaine planifient, c'est-à-dire qu'ils passent du temps à définir leur besoin d'information, particulièrement en début de processus. Les experts du domaine ont en effet fait appel à leurs connaissances antérieures et sont retournés sur la consigne tout au long de leur tâche de recherche d'information plus souvent que les novices. Effectivement, le novice du domaine a eu plus de difficulté à formuler son besoin d'information et donc sa requête et les mots-clés choisis sont peu précis. Le participant novice a donc employé un raisonnement heuristique.

Pourtant, contrairement aux travaux de S. Brand-Gruwel, I. Wopereis, et Y. Vermetten (2005), nous avons constaté que les experts du domaine ont également utilisé un raisonnement heuristique lors de la formulation de la requête. En effet, un des participants expert du domaine a préféré saisir la consigne en entier dans la barre de recherche. Il nous explique vouloir gagner du temps, au cas où la question ait déjà été posée, sur un forum par exemple,

---

<sup>58</sup> Brand-Gruwel S., Wopereis I., Vermetten Y. (2005). Information problem solving by experts and novices : Analysis of a complex cognitive skill. *Computer in Human Behavior*, vol. 21, n°3, p. 487-508

et pouvoir ainsi recopier directement la réponse. Cette stratégie est directement liée à un désir d'économie de temps de la part du participant expert du domaine. Le gain de temps a été un des critères les plus mentionnés au cours des recherches d'information. Il s'est manifesté par une recherche de la réponse déjà rédigée, puis par la recherche d'un gain de temps pour la production finale. La notion d'usage futur a été importante avec une recherche de documents qui soient bons à « copier coller ».

Nos observations ont révélée chez les participants experts à la fois un raisonnement analytique et heuristique. Au contraire, le participant novice a uniquement utilisé un raisonnement heuristique. Ce dernier point confirme notre hypothèse de départ, même si ce résultat n'est pas généralisable avec un seul participant.

Hypothèse 2 : les élèves utilisant un raisonnement analytique lisent les résumés des pages de résultats et ne s'appuient pas sur des critères typographiques. *A contrario*, les élèves utilisant un raisonnement heuristique sélectionnent les différents liens de la page de résultats.

Les participants experts, lors de nos observations, ont mentionné des critères pour la sélection des liens sur la page de résultats qui renvoient à un raisonnement analytique. Ces derniers ont mentionné l'importance que le résumé soit en adéquation avec leur thème de recherche, confirmant ainsi les travaux de Boubée (2008)<sup>59</sup>. Le résumé n'a pas été le seul critère de sélection et le titre affiché sur la page de résultats a également été pris en compte. De plus, les participants experts ont pris en compte la fiabilité de la source lors de la sélection d'un lien sur la page de résultats, confirmant ainsi l'utilisation d'un raisonnement analytique.

Aucun des participants experts comme novice n'a signalé une influence du choix du lien à cause de la typographie (mots en gras par exemple). Les indices de surface ont été peu pris en compte par les participants. Effectivement, ces derniers sont des élèves en classe de seconde et Dinet (2006)<sup>60</sup> souligne que lors d'une tâche de sélection de références pertinentes, plus l'utilisateur est jeune et plus il est sensible au marquage typographique des références, que celles-ci soient pertinentes ou non. Ainsi, dans ce cas là, la totalité des participants à utilisé un raisonnement analytique, confirmant ainsi notre hypothèse pour les experts du domaine mais infirmant celle pour les novices du domaine.

---

<sup>59</sup> Boubée N. (2008). Les stratégies des jeunes chercheurs d'informations en ligne. *Question de communication*. Vol 14 p. 33-48.

<sup>60</sup> Dinet J. (2006). Le choix des sites Web par les enfants et les adolescents : impacts de la typographie. In Piolat A. *Lire, écrire, communiquer et apprendre avec Internet*. Marseille, Solal, p.135-150.

Pourtant, malgré une lecture de la page de résultats, la position du lien a été un critère important pour les experts du domaine, et les liens ont souvent été sélectionnés dans l'ordre, renvoyant ainsi à un raisonnement heuristique, contrairement à notre hypothèse. Ces résultats confirment ceux de Granka, Joachims et Gay (2004)<sup>61</sup> qui mettent en avant que les sélections de liens des élèves révèlent les mêmes traits caractéristiques que ceux des usagers adultes, à savoir une sélection effectuée majoritairement sur les liens positionnés en rang 1 et 2. Ces deux rangs constituent 58,6 % des sélections de l'ensemble « collège et lycée ». Nous pourrions expliquer cette stratégie par un excès de prudence de la part des participants experts du domaine. Egalement, il est possible que les participants se soient évalués comme experts du domaine même si ce n'était pas le cas. Il est en effet difficile de porter un jugement prédictif à partir d'un résumé ou d'un extrait du document primaire. Au contraire, même si la position du lien a été mentionnée comme un critère de sélection important dans la page de résultats, le participant novice n'a pas sélectionné que les liens en première position mais également ceux situés en bas de la page de résultats. Ce constat ne confirme donc pas notre hypothèse puisque le participant novice n'a pas appliqué un raisonnement heuristique comme nous l'avons défini. Ce bilan reste à nuancer puisqu'un seul participant novice n'est pas représentatif.

Hypothèse 3 : les élèves utilisant un raisonnement analytique s'appuient sur des critères de fiabilité de la source lors de la consultation du document primaire (contenu informationnel mais aussi la source et l'auteur). *A contrario*, les élèves utilisant un raisonnement heuristique s'appuient sur des critères d'évaluation superficiels.

Nos observations ont mis en avant que les experts du domaine s'appuient sur de nombreux critères de pertinence avant de sélectionner un document primaire. Les participants experts du domaine ont pris en compte le fait que le document soit en adéquation avec le thème de la recherche, considérant ainsi le contenu informationnel. Ils ont également mentionné l'importance de la source et de l'auteur dans le choix d'un document. Ces résultats renvoient à

---

<sup>61</sup> Granka L.A., Joachims T., Gay G. (2004). Eye-tracking analysis of user behavior in WWW search. In *Proceedings of the 27th annual international ACM SIGIR conference on research and development in information retrieval*. Sheffield, Grande-Bretagne.

ceux Holscher et Strube(2000) <sup>62</sup> qui soulignent que les experts du domaine examinent le contenu informationnel et font aussi attention à la source, et confirment donc notre hypothèse selon laquelle les experts du domaine emploient un raisonnement analytique.

Cependant, les experts du domaine, lors de nos observations, se sont également appuyés sur des critères d'évaluation superficiels, comme la présence d'images non informatives, ce qui témoigne de l'utilisation d'un raisonnement heuristique.

Cette fois encore, nous n'avons pas noté une signification notable entre les deux types de raisonnement puisque certains participants évoquent au cours d'une même recherche à la fois des critères de fiabilité de la source (raisonnement analytique) mais également des critères d'évaluation superficiels (raisonnement heuristique).

En ce qui concerne le novice du domaine, son jugement porte à la fois sur des critères superficiels comme l'esthétique du site mais également sur le contenu informationnel du document. Pourtant, le novice n'a jamais pris en compte la fiabilité de la source et de l'auteur, contrairement aux participants experts du domaine. Ainsi, il semblerait tout de même que le raisonnement heuristique prime lors de la sélection de document chez ce participant novice.

Nos travaux ont révélé de nouveaux critères de pertinence, évoqués par des experts du domaine. Par exemple, ces derniers ont mentionné l'importance que le document primaire corresponde aux exigences de la consigne donnée par l'enseignant, lors d'une recherche d'information scolaire. Aussi, les participants ont porté une attention considérable à la production finale attendue par l'enseignant. La recherche d'information est même guidée par ce travail final : les élèves ont évoqué la nécessité que le document primaire permette de gagner du temps pour achever le travail final. En lien avec cet objectif, les experts du domaine ont révélé l'importance que le document primaire ait un niveau de langue abordable, non pas pour qu'ils puissent comprendre aisément le document, mais pour pouvoir utiliser le copier coller du document de façon rusée, sans avoir besoin de reformuler l'information prélevée.

Nos hypothèses mettaient en avant une différence significative entre les raisonnements analytiques et heuristiques. Pourtant, il semblerait que, lors d'une recherche d'information, un seul type de raisonnement ne soit employé, que ce soit avec un participant expert ou novice du domaine. Ainsi, une recherche d'information sur le web semble basée à la fois sur un raisonnement réfléchi mais également sur des automatismes.

---

<sup>62</sup> Holscher C., Strube G. (2000). « Web search behavior of Internet experts and newbies ». *Computer Networks*, vol. 33, n°1-6, p.337-346.

# Conclusion

Pour conclure, nos travaux ont mis en avant de nouveaux critères de pertinence comme l'importance que le document primaire corresponde aux exigences de la consigne donnée par l'enseignant, lors d'une recherche d'information scolaire, et que le document permette de gagner du temps pour achever le travail final. Ils mentionnent comme nécessaire que le niveau de langue soit abordable, non pas pour pouvoir comprendre aisément le document, mais pour pouvoir utiliser le copier coller du document de façon rusée, sans avoir besoin de reformuler l'information prélevée.

Il est indéniable que les connaissances du domaine opèrent sur les stratégies de recherche, sur la définition du besoin d'information et sur la formulation de la requête. De la même façon, les critères de pertinence appliqués à la page de résultats et au document primaire diffèrent selon les connaissances du domaine.

Ainsi, les raisonnements employés par un individu sont corrélés aux connaissances du domaine.

Il semblerait désormais intéressant de poursuivre nos travaux de recherche en augmentant le nombre de participants de l'échantillon, ce qui permettrait de généraliser les résultats. Aussi, une prise en compte des autres types de connaissance intervenant lors d'une recherche d'information (connaissances techniques et informationnelles) pourrait permettre d'éviter le mécanisme de compensation qui existe entre ces trois types de connaissances.

Enfin, nos travaux permettent de mieux comprendre les pratiques informationnelles juvéniles et les stratégies employées lors de recherches d'informations scolaires. Nos observations ont mis en avant de nombreuses « ruses » pour gagner du temps, à la fois lors de la recherche d'information mais également pour l'exploitation des résultats. Une sensibilisation à la pratique du copier coller, un travail sur la reformulation de l'information ainsi que sur la production de mots clés et l'évaluation de l'information pourraient être des pistes de travail pour développer les compétences informationnelles des élèves.

# Bibliographie

Bilal D. (2002). Children's use of Yahoo!igans! Web search engine: Cognitive and physical behaviors on fully self-generated search tasks. *Journal of the American Society for Information Science and Technology*, vol.53, n°13, p. 1170-1183.

Blondel M.F, Kempf O., Schwob M. (2000). Recherche d'information sur la toile : Pratiques d'élèves. *Bulletin de l'Union des physiciens*, n°828, vol. 94, p.1819-1846.

Boubée N. (2008). Les stratégies des jeunes chercheurs d'informations en ligne. *Question de communication*. Vol 14 p. 33-48.

Brand-Gruwel S., Wopereis I., Vermetten Y. (2005). Information problem solving by experts and novices : Analysis of a complex cognitive skill. *Computer in Human Behavior*, vol. 21, n°3, p. 487-508

Cauzinille-Marmèche Evelyne, Mathieu Jacques, Weil-Barais Annick (1982). Le raisonnement expérimental des pré-adolescents : stratégies de test d'hypothèses. In: *Enfance*. Tome 35 n°1. pp. 23-37.

De Cara B. [et al.] (2009). *La recherche d'information sur Internet par les jeunes usagers : acquisition et développement des stratégies*. Cinquième colloque de Psychologie Ergonomique (EPIQUE 2009), Département Recherche de la Société Française de Psychologie. Nice : 28-30 septembre.

Dinet J. (2006). Le choix des sites Web par les enfants et les adolescents : impacts de la typographie. In Piolat A. *Lire, écrire, communiquer et apprendre avec Internet*. Marseille, Solal, p.135-150.

Fluckiger C., Bruillard E. (2008). TIC : analyse de certains obstacles à la mobilisation des compétences issues des pratiques personnelles dans les activités scolaires. In Colloque international « L'éducation à la culture informationnelle », Lille 3, 16-18 octobre 2008.

Gerjets P., Kammener Y., Wermer B. (2010). "Measuring spontaneous and instructed evaluation process during Web search: Integrating concurrent thinking-aloud protocols and eye-tracking data", *Learning and Instruction*, p. 1-12.

Granka L.A., Joachims T., Gay G. (2004). Eye-tracking analysis of user behavior in WWW search. In *Proceedings of the 27th annual international ACM SIGIR conference on research and development in information retrieval*. Sheffield, Grande-Bretagne.

Hembrooke H. A., Granka L. A., Gay G. K. (2005). "The effect of expertise and feedback on search term selection and subsequent learning". *Journal of the American Society for Information Science and Technology*, vol. 56, n°8, p. 861-871.

Hirsh S. G. (1997). How do children find information on different types of tasks? : Children's use of the science library catalog. *Library trends*, vol. 45, n°4, p. 725-745.

Hirsh S. G. (1999). Children's relevance criteria and information seeking on electronic resources. *Journal of the American Society for Information Science*, vol. 50, n°14, p. 1265-1283.

Holscher C., Strube G. (2000). « Web search behavior of Internet experts and newbies ». *Computer Networks*, vol. 33, n°1-6, p.337-346.

Ihadjadene M., Chaudiron S., Martins D. (2003). « The Effect of Individual Differences on searching the Web », *Proceedings of the 66th Annual Meeting of the American Society for Information Science and Technology*, Long Beach, p. 240-246.

Ihadjadene M. Chaudiron S. (2008). Quelles analyses de l'usage des moteurs de recherche : Questions méthodologiques. *Question de communication*. N°14, p. 17-32.

Jenkins C., Corritore C., Wiedenbecks S. (2003). Patterns of information seeking on the Web : A qualitative study of domain expertise and Web expertise. *IT and society*, vol. 1, n°3.

Lorigo L. [et al.] (2006). The influence of task and gender on search and evaluation behavior using Google. *Information Processing and Management*, vol. 42, p. 1123-1131.

Metzger, M. J., Flanagin, A. J., & Medders, R. B. (2010). Social and Heuristic Approaches to Credibility Evaluation Online. *Journal of Communication*, 60(3), 413–439.

Ravestain J., Ladage C., Johsua S. (2007). Trouver et utiliser des informations sur Internet à l'école : problèmes techniques et questions éthiques. *Revue française de pédagogie*, n° 158, p.71-83.

Rieh, S. Y. (2002). Judgment on Information Quality and Cognitive Authority in the Web. *Journal of the American Society for Information Science and Technology*, vol.53, n°2, p. 145-161.

Serres, A. (2005). Evaluation de l'information sur Internet : le défi de la formation. *Bulletin des bibliothèques de France*, t.50, n°6.

Simonnot B. (2002). *De la pertinence à l'utilité en recherche d'information : le cas du Web*. Actes du colloque international MICS-LERASS, Toulouse ; ADBS Éditions, collection Sciences de l'information, série Recherches et Documents, Paris, p. 393-410.

Simonnot B. (2007). Evaluer l'information. *Documentaliste – Sciences de l'information*, vol.14, n°3, p.210-216.

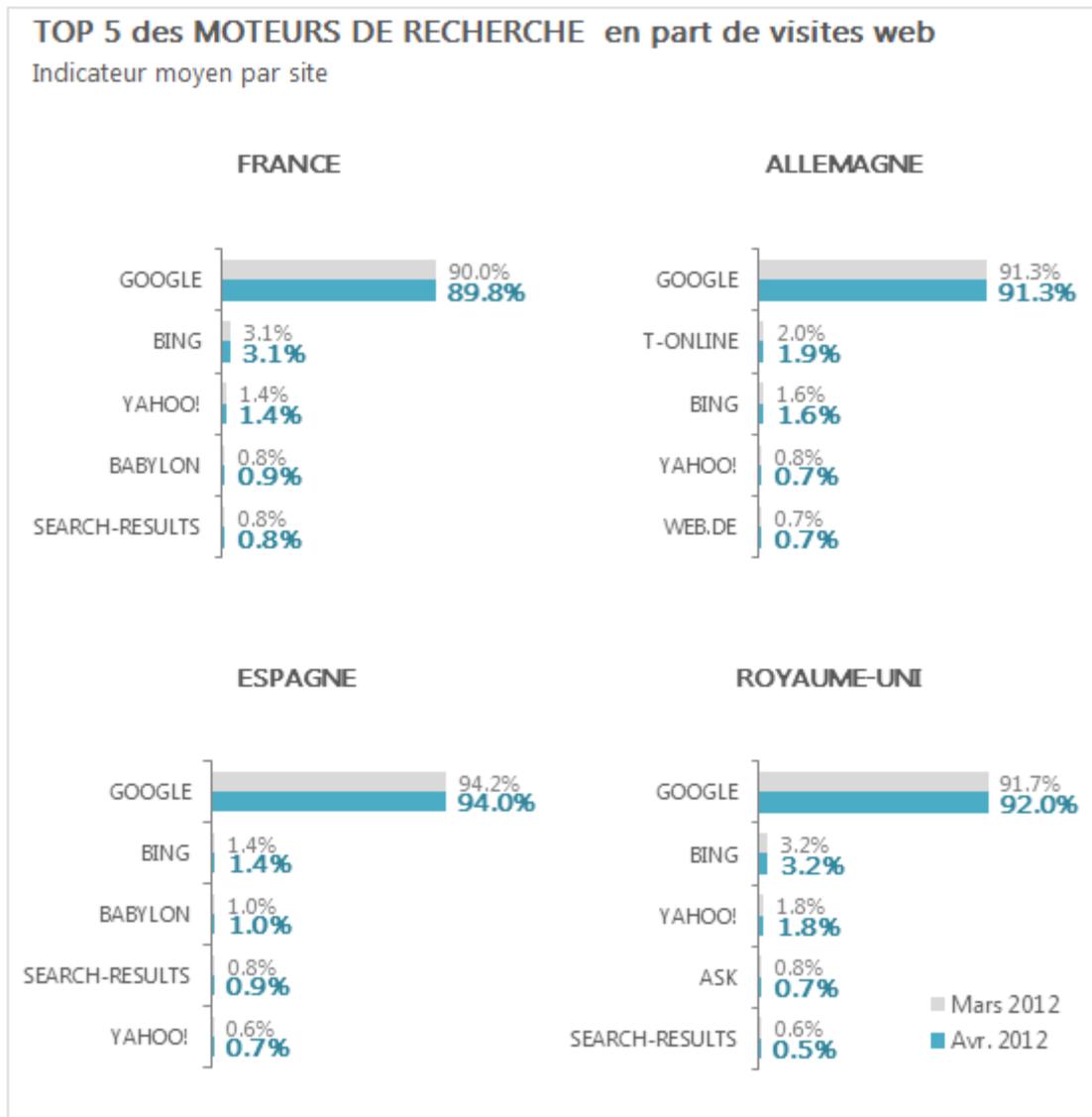
Simonnot B. (2008). « La pertinence en sciences de l'information : des modèles, une théorie ? ». In Papy F., *Problématiques émergentes dans les Sciences de l'Information*, Paris : Hermès-Lavoisier, ch.6, p. 161-182.

Wallace R. [et al.] (2000). Science on the Web : Students online in a sixth-grade classroom. *The Journal of the Learning Sciences*, vol. 9, n°1, p. 75-104.

Wildemuth B. (2004). The Effects of Domain Knowledge on Search Tactic Formulation. *Journal of the American Society for Information Science and Technology*, vol. 55, n°3, p. 246-258.

# Annexes

## Annexe 1 : Le choix de Google, top 5 des moteurs de recherche en part de visite sur le web



Source : AT INTERNET

## Annexe 2 : Grille d'observation

Participant n° ... [expert] [novice]

Classe :

Age :

Sexe :

Connaissance du domaine : 0 – 1 – 2 – 3 – 4 – 5

Sujet / consigne de la RI :

### Observation lors de l'évaluation de la page de résultats :

La page de résultats	Nombre de pages de résultats consultées	
	Numéro / place des pages consultées sur le nombre total de pages	
Les liens consultés	Nombres de liens consultés avant de répondre à la tâche de RI	
	Place des liens consultés dans la liste totale de liens	

### Annexe 3 : Grille d'aide pour l'entretien

- Quel SRI a été choisi ?
- Raisonnement et stratégie ?

#### La page de résultats

Demander pourquoi tel lien a été choisi ou non ?

Critères	Nombre de mentions
Adéquation au thème	
Position du lien	
Cohérent	
Type de document (encyclopédie etc.)	
Conformité avec l'opinion personnelle	
Etendue (recherche du document complet)	
Nouveauté	
Résumé/ titre	
Typographie	
Adresse url	
Niveau de langue (abordable)	
Autre	

## Annexe 4 : Critères de pertinence des novices du domaine

Novices			
Critères de pertinence	Page de résultats	Document primaire	
<b>Adéquation au thème</b>	« <i>Le titre est pertinent, c'est-à-dire qu'il correspond à ce que je cherche</i> »	« <i>C'est super, ça correspond exactement à ce que je cherche !</i> » « <i>Là ça concerne l'art (peinture) donc rien à voir, moi je cherche sur le cinéma et le théâtre</i> »	
<b>Position du lien et popularité</b>	« <i>Il est en premier donc c'est le plus visité et le meilleur</i> » « <i>Plus je m'éloigne du début plus ça devient n'importe quoi</i> »		
<b>Cohérence</b>		« <i>C'est très intéressant car ça explique tout ce qu'il faut pour faire une critique de film</i> » « <i>Je vérifie que ça soit pas des conneries</i> »	
<b>Type de documents</b>			
<b>Présence d'image</b>		<b>Positif</b>	<b>Négatif</b>

			« Il y a des photos mais moi j'en cherche pas »
<b>Conformité avec l'opinion personnelle</b>		« Ca correspond à une critique de Avatar (le film de James Cameron) donc c'est un super film alors la critique doit être bonne.	
<b>Etendue</b>		<b>Positif</b>	<b>Négatif</b>
		« Il y a vraiment tout » « C'est complet » (x3)	« C'est trop long, ça ne me donne pas envie de lire » (x4)
<b>Nouveauté</b>			
<b>Niveau de langue</b>		« Il faut que ça soit compréhensible, pas avec un jargon cinématographique ».	
<b>L'esthétique</b>		« J'aime pas la couleur donc ça va pas aller, ça me donne pas envie de continuer sur ce site et de tout lire »	

<b>Evaluation du système</b>	« <i>Le lien ne s'ouvre pas donc je retourne sur la page d'avant pour en choisir un autre</i> »	« <i>Tous les sites se ressemblent, c'est répétitif!</i> » « <i>C'est encore la même chose</i> »
<b>Temps</b>	« <i>On va tous les essayer car on a le temps</i> »	
<b>Correspondance avec la consigne donnée</b>		« <i>Il faut que ça correspondent avec ce qu'a demandé la prof! Elle nous a donné des consignes précises sur comment on doit rédiger une critique alors faut que je vérifie que cet exemple il corresponde avec ce qu'elle elle nous a demandé</i> ».

Tableau 9 : Expressions des justifications des sélections par les novices