

**Tests d'utilisabilité :
comparaison de deux méthodes appliquées au
site e-rara.ch**

Mémoire de recherche réalisé par :

Valérie MEYSTRE

Raphaël REY

Sous la direction de :

René Schneider, professeur HES

Carouge, le 10 janvier 2014

**Master en Information documentaire
Haute École de Gestion de Genève (HEG-GE)**

Déclaration

Ce mémoire de recherche est réalisé dans le cadre du Master en information documentaire de la Haute école de gestion de Genève. Les étudiants acceptent, le cas échéant, la clause de confidentialité. L'utilisation des conclusions et recommandations formulées dans ce travail, sans préjuger de leur valeur, n'engage ni la responsabilité des auteurs, ni celle de l'encadrant.

« Nous attestons avoir réalisé le présent travail sans avoir utilisé des sources autres que celles citées dans la bibliographie. »

Fait à Carouge, le 10 janvier 2014

Raphaël Rey

Valérie Meystre

Remerciements

Nous tenons à remercier chaleureusement tous ceux qui ont participé aux tests à distance, à savoir les 160 personnes qui ont les débutés et encore davantage les 84 qui ont eu la patience et la détermination pour achever l'ensemble des tâches et répondre à toutes les questions.

Nous remercions aussi les sept personnes qui ont bien accepté de se livrer au test modéré. Nous avons particulièrement apprécié leur engagement, d'autant plus que l'exercice du think-aloud peut paraître un peu déroutant au début. Pour des raisons d'anonymat, nous ne les nommerons pas, mais leur contribution nous a été des plus précieuses.

Nous remercions aussi tout particulièrement Mme Eliane Blumer, dont la disponibilité et l'expérience nous ont beaucoup aidés pour anticiper de nombreuses difficultés.

Pour finir, nous adressons tous nos remerciements à M. René Schneider qui a dirigé ce projet de recherche et su nous proposer une thématique à la fois stimulante et riche en apprentissages.

Résumé

L'utilisabilité est un aspect primordial d'un site Internet. Selon la norme ISO 9241-11 (1998 p. 2), elle concerne l'efficacité, l'efficience et la satisfaction dans l'interaction entre un utilisateur et un système en fonction de buts. Garantir cette utilisabilité ne va pas de soi, cela requiert d'être testée régulièrement et il existe à cet effet plusieurs méthodes. L'objet principal de ce travail est la comparaison de deux d'entre elles : une méthode dite classique (avec un modérateur présent et recours au think-aloud) et une méthode automatisée à distance. Nous avons choisi le site e-rara.ch comme théâtre de cette comparaison, avec également pour objectif de proposer des recommandations d'amélioration. Nous avons élaboré un test comprenant un questionnaire préalable récoltant des données sur les compétences des participants, cinq tâches et un questionnaire final portant sur l'évaluation générale du site en plus d'une question ouverte où les participants pouvaient proposer directement des améliorations. Pour la mise en place du test, nous avons utilisé le logiciel Loop11 et réalisé une version francophone et germanophone de celui-ci. Dans un premier temps, nous avons lancé le test à distance, en recrutant nos participants majoritairement via la liste de diffusion Swiss-lib. 84 personnes ont effectué la totalité du test. Dans un deuxième temps, nous avons réalisé le test modéré en présentiel avec 7 participants, 5 francophones et 2 germanophones. Il s'agissait de déterminer si le test à distance était interprétable sans recours aux données récoltées en présentiel. Pour chaque tâche, nous avons examiné le taux de succès et le premier clic, mais surtout les stratégies employées, les commentaires et les erreurs, qui nous ont permis d'identifier 25 problèmes d'utilisabilité, tous tests confondus. Pour l'ensemble d'entre eux, nous avons expliqué comment ils ont été décelés ou non par chacune des deux méthodes, puis nous avons émis des recommandations pour leur apporter des solutions. Nous avons ainsi pu comparer les deux types de tests en termes d'efficacité et d'efficience. Il est apparu que l'analyse des tâches via le test automatisé à distance permet de relever les problèmes d'utilisabilité les plus sévères, mais pour beaucoup d'autres problèmes, le repérage nécessite une question ouverte où les participants peuvent s'exprimer librement (propositions d'améliorations, difficultés éprouvées, etc.). Il a aussi l'avantage d'appuyer ses découvertes par des données statistiques grâce à son nombre plus élevé de participants. Le test modéré en présentiel peut réussir à repérer des problèmes de n'importe quel type, y compris des questions de manque de visibilité de certains éléments, même s'il ne permet pas d'en évaluer de manière chiffrée leur impact. Le think-aloud est également précieux : il permet de suivre au plus près les éventuels manques de compréhension ou doutes des personnes observées. Le petit nombre de participants présente toutefois le risque de passer à côté de problèmes marginaux. En conclusion, l'idéal serait évidemment de pouvoir recourir au deux méthodes, qui sont largement complémentaires. Si ce n'est pas possible, le choix dépendra des finalités du test.

Mots-clefs : expérience utilisateur, test d'utilisabilité, test à distance, test modéré, think-aloud, étude comparative, étude de cas, bibliothèque numérique, bibliothèque digitale, imprimés anciens, bibliothèques suisses, e-rara.ch, Loop11, SUS.

Table des matières

Déclaration.....	i
Remerciements	ii
Résumé	iii
Liste des tableaux	vi
Liste des figures.....	vi
1. Introduction.....	1
2. Etat de l’art.....	3
2.1 Notion d'utilisabilité.....	3
2.2 Bref historique.....	3
2.3 Evaluation de l'utilisabilité	4
2.3.1 Généralités.....	4
2.3.2 Etudes Qualitatives et études quantitatives	5
2.3.3 La méthode du think-aloud.....	5
3. Méthodologie	8
3.1 Planification générale de l'étude.....	8
3.2 Choix du logiciel Loop11	8
3.3 Création du test.....	10
3.3.1 Considérations générales.....	10
3.3.2 Message d'accueil.....	11
3.3.3 Questionnaire prétest.....	11
3.3.4 Tâches à réaliser.....	12
3.3.5 Questionnaire posttest	14
3.3.6 Prétests.....	17
3.3.7 Ordre des deux méthodes de tests	18
3.3.8 Réalisation du test automatisé à distance	19
3.3.9 Réalisation du test en présentiel	19
3.3.10 Méthode de comparaison des deux tests	20
3.3.10.1 Efficacité au repérage des problèmes d'utilisabilité.....	20
3.3.10.2 Probabilité du repérage des problèmes d'utilisabilité	20
3.3.10.3 Efficience du repérage des problèmes d'utilisabilité	21
3.3.10.4 Evaluation de la gravité du problème	22
3.3.10.5 Facilitation de la construction de propositions d'amélioration	22
3.3.10.6 Processus d'évaluation.....	23
4. Analyse des données	24
4.1 Remarques générales	24
4.2 Test automatisé à distance.....	24
4.2.1 Généralités.....	24
4.2.2 Données personnelles sur les participants	24

4.2.3	Questionnaire d'évaluation générale du site.....	25
4.2.4	Analyse des tâches	26
4.2.4.1	Erreurs	26
4.2.4.2	Temps de réalisation	27
4.2.4.3	Projection de la difficulté de la tâche.....	27
4.2.5	Question ouverte.....	27
4.2.6	Bilan des tests automatisés à distance.....	28
4.3	Test modéré.....	249
4.3.1	Généralités.....	29
4.3.2	Données personnelles sur les participants	29
4.3.3	Analyse des tâches	29
4.3.4	Questionnaire d'évaluation générale du site.....	30
4.3.5	Question ouverte.....	31
4.3.6	Bilan des tests modérés.....	32
5.	Comparaison des méthodes.....	34
5.1	Généralités	34
5.2	Données personnelles sur les participants.....	34
5.3	Analyse des tâches.....	35
5.4	Questionnaire d'évaluation générale du site.....	37
5.5	Question ouverte.....	39
5.6	Fiabilité des tests	41
5.7	Capacité à comparer différentes versions d'un site.....	42
5.8	Attitude des participants	42
6.	Conclusion	44
	Bibliographie	46
	Annexe 1 : Analyse du test automatisé à distance	50
	Annexe 2 : Comptes rendus des tests modérés	71
	Annexe 3 : Problèmes d'utilisabilité et recommandations	98
	Annexe 4 : Comparaison des méthodes en fonction des problèmes d'utilisabilité	123
	Annexe 5 : poster de présentation du projet	150

Liste des tableaux

Tableau 1 : Les quatre éléments de l'expérience utilisateur selon Guo	16
Tableau 2 : Questionnaire posttest et axes analysés.....	17
Tableau 3 : Variable 1	20
Tableau 4 : Variable 2	21
Tableau 5 : Variable 3	22
Tableau 6 : Variable 5	23
Tableau 7 : Liste des problèmes repérés par le test automatisé	29
Tableau 8 : Liste des problèmes repérés par le test modéré	32
Tableau 9 : L'essentiel en bref : exploitation des données personnelles.....	35
Tableau 10 : Liste des problèmes suspectés (test automatisé)	36
Tableau 11 : L'essentiel en bref : analyse des tâches	37
Tableau 12 : L'essentiel en bref : exploitation du questionnaire d'évaluation générale	39
Tableau 13 : Liste des problèmes repérés via la question ouverte (test automatisé) ...	39
Tableau 14 : L'essentiel en bref : exploitation de la question ouverte	41

Liste des figures

Figure 1 : Exemple de heat map défectueux (tâche 3)	9
Figure 2 : Extrait d'un clickstream (tâche 3).....	9
Figure 3 : Résultats d'une recherche sur e-rara.ch.....	13
Figure 4 : Affichage des pages sous forme de mosaïque	14
Figure 5 : Echelle d'interprétation du score SUS	15
Figure 6 : Scores moyens du questionnaire posttest (test automatisé)	25
Figure 7 : Scores moyens du questionnaire posttest (test modéré)	31
Figure 8 : Nombre de problèmes repérés via l'analyse des tâches.....	35
Figure 9 : Nombre de problèmes repérés par méthode (sans question ouverte pour le test automatisé).....	40
Figure 10 : Nombre de problèmes repérés par méthode	40
Figure 11 : Comparaison du temps et du nombre moyen de pages vues entre les participants du test automatisé et ceux du test modéré	42

1. Introduction

On ne compte plus les ouvrages qui nous donnent des conseils pour la création de sites web à la fois fonctionnels et ergonomiques. Toutefois, ces ressources ne suffisent pas à modéliser l'ensemble des facteurs qui font qu'une interface est facile d'utilisation ou non. Cela suppose une quantité de paramètres à prendre en compte bien trop importante et spécifique à chaque contexte (comme la disposition, la taille, le nombre des éléments, la logique de la structure du site, le choix des couleurs pour n'en citer que quelques-uns). Un expert en ergonomie ne peut que difficilement détecter toutes les faiblesses d'une interface ou en proposer une parfaitement claire et facile d'utilisation dès sa création. Son expérience l'aide à repérer ce qui pose habituellement problème, mais le recours à des tests se révèle indispensable pour observer directement les interactions entre système et utilisateurs, dans le but de confirmer ou infirmer les hypothèses et les choix à la base de la création de celui-ci.

Ce travail de recherche poursuit deux objectifs : le premier consiste à comparer deux méthodes de test d'utilisabilité en terme d'efficacité et d'efficience : l'une à distance et complètement automatisée, l'autre basée sur le recours au think-aloud (ou pensée à voix haute), dans laquelle un modérateur invite la personne effectuant le test à verbaliser ses pensées et intentions. La comparaison s'appuie sur une série de tests avec pour cible le site e-rara.ch. Il s'agit d'une bibliothèque numérique mettant à disposition des imprimés anciens provenant de plusieurs bibliothèques suisses. Ce site avait déjà été l'objet d'un premier test réalisé en 2011 par Eliane Blumer et dont les propositions ont été la source de nombreuses améliorations (voir Blumer, 2011). Un nouveau test garde cependant tout son sens puisqu'il est intéressant d'évaluer l'impact des mesures prises et de vérifier si certains problèmes n'avaient pas échappé lors de cette première démarche. Proposer une nouvelle série de recommandations aux responsables du site e-rara.ch constitue par conséquent le second objectif de ce projet de recherche.

Lors de la comparaison des deux méthodes (objectif 1), l'enjeu sera de déterminer à quel point il est possible de repérer les problèmes indifféremment avec l'une et l'autre ou s'il y a des cas spécifiques pour lesquels l'une d'entre elles est de loin la plus adaptée. En plus de la capacité à détecter les éléments impactant négativement l'utilisabilité, il sera aussi pertinent d'évaluer l'effort nécessaire (temps, moyens) pour y parvenir ainsi que la fiabilité et la précision de l'analyse. Est-ce que l'on comprend facilement où se situe exactement la faiblesse ou faut-il encore d'autres investigations ? Peut-on estimer la gravité du problème ? Déterminer s'il s'agit d'un cas isolé ou qui touche la plupart des usagers ? En répondant à ces questions en fonction des types de problèmes rencontrés, nous parviendrons à définir les avantages et faiblesses des deux méthodes de test utilisées ici.

La première partie de ce travail définira un cadre théorique général. Après une définition du terme « utilisabilité », il décrira l'évolution des pratiques dans ce domaine à travers un bref historique et s'arrêtera sur la notion de think-aloud, fondamentale pour notre travail. Arrivera ensuite une partie qui expliquera plus en détail la méthodologie, d'une part comment ont été réalisés les deux tests, mais aussi les critères sur lesquels se fondera la comparaison des deux méthodes. La section suivante offrira un panorama extrêmement sommaire des résultats des tests, mais suffisant pour comprendre les enjeux de la dernière partie qui sera le lieu de la comparaison proprement dite. La conclusion établira un bilan sur les avantages et inconvénients de chacune des deux méthodes. La plus grande parties des données que

nous avons utilisées sont disponibles parmi les annexes. Celles-ci regrouperont les résultats détaillés des différents tests (annexes 1 et 2), les recommandations (objectif 2 ; annexe 3) et la comparaison des méthodes problème par problème (annexe 4).

2. Etat de l'art

2.1 Notion d'utilisabilité

Le concept d'utilisabilité (ou parfois usabilité ; usability en anglais) désigne la propriété qualitative d'une interface d'être facilement utilisable. Il s'agit d'une qualité parmi d'autres, qui implique que « les personnes utilisant ce produit peuvent le faire rapidement et facilement pour accomplir leurs propres tâches » (Dumas et Redish 1999, p. 4).

Jakob Nielsen, un des acteurs les plus éminents dans le domaine, définit l'utilisabilité comme « un attribut qualité qui évalue à quel point les interfaces utilisateurs sont faciles à utiliser » (Nielsen 2012a, p. 1). Voici la définition plus complète de l'Organisation internationale de normalisation qui en détaille plusieurs composantes :

Le degré selon lequel un produit peut être utilisé, par des utilisateurs identifiés, pour atteindre des buts définis avec efficacité, efficience et satisfaction, dans un contexte d'utilisation spécifié. (Organisation mondiale de normalisation 1998, p. 2)

L'utilisabilité comporte plusieurs attributs, en nombre et en désignations légèrement différents selon les auteurs. Dans la définition citée ci-dessus, l'Organisation internationale de normalisation mentionne efficacité, efficience et satisfaction. Nielsen (1993 et 2012a) en dénombre cinq : apprentissage (learnability), efficience, rémanence (memorability), taux d'erreur et satisfaction. Ce sont les références les plus citées, mais d'autres auteurs les nomment un peu différemment (Jeng 2005, p. 48).

L'utilisabilité se focalise sur le point de vue des utilisateurs¹. Dumas et Redish (1999, p. 5-6) notent que ce sont les utilisateurs qui déterminent si un produit est facile à utiliser.

2.2 Bref historique

Issu de l'étude des interactions humain-ordinateur (computer-human interaction ou HCI), le champ de l'utilisabilité s'en distingue par ses méthodes. En effet, en utilisabilité, il ne s'agit pas d'appliquer les méthodes de recherche des sciences du comportement, ce qui a été souvent le cas dans les programmes d'enseignement du HCI (Dumas 2007), mais de proposer de nouvelles approches, qui lui sont spécifiques.

Le concept d'utilisabilité voit le jour dans les années 1980 avec les travaux de John Whiteside et de John Bennett et notamment la publication d'un article en 1988 sur « Usability engineering » (Dumas 2007). Ils proposent une approche pratique et « engineering » pour le design de produits, mettant entre autres en avant l'utilisation de prototypes, le processus d'évaluation itérative et l'importance d'avoir un produit utilisable pour être productif.

Les cinq années qui suivent voient un très fort intérêt pour cette nouvelle approche, aidée par l'apparition du Macintosh, « symbole d'un produit avec un bon design » (Dumas 2007, p. 56). L'utilisabilité entre ainsi parmi les critères essentiels d'un nouveau produit, et il s'ensuit une explosion de la demande en travaux de ce type. Des personnes de tous horizons se tournent vers l'utilisabilité, apportant des idées nouvelles et des valeurs provenant de leur propre discipline. Durant ces cinq années, une véritable profession s'est constituée et dégagée « des racines académiques en psychologie (...) et a embrassé les concepts

¹ Il existe également d'autres méthodes d'évaluations comme l'évaluation heuristique qui recourt à des experts (Barnum 2011), mais nous n'y avons pas eu recours dans le cadre de ce travail et n'en parlerons par conséquent pas plus en détail.

d'engineering et d'utilisabilité » (Dumas 2007, p.57). L'apparition d'Internet a permis un nouveau champ d'application de ces méthodes. Toutefois, le laboratoire est encore bien présent en tant que cadre des tests.

Au cours des années 1990, de nouvelles techniques visant justement à sortir les tests d'utilisabilité des laboratoires ont vu le jour (Tullis et al. 2002), dans le but de réduire les coûts en termes de personnel et d'installations, et de pouvoir augmenter le nombre de personnes réalisant les tests. Dans un premier temps, il s'agit simplement de séparer géographiquement les participants aux tests et les observateurs, les seconds examinant les premiers par vidéo interposée ou via un logiciel spécifique. Au final, ce dispositif ne permettait pas vraiment de réduire le temps et le personnel. Une autre forme de test à distance émergea donc, dissociant totalement la réalisation du test et l'observation des données (collectées automatiquement et pouvant être analysées plus tard).

2.3 Evaluation de l'utilisabilité

2.3.1 Généralités

Tester l'utilisabilité d'un produit (plus spécifiquement d'un site ou une partie de site web comme c'est le cas dans cette étude) consiste à demander à des utilisateurs représentatifs du public cible de réaliser des tâches, elles aussi représentatives des fonctionnalités et buts du site en question. L'observation des erreurs commises lors de la réalisation de ces tâches permet de mettre le doigt sur des problèmes pouvant préteriter la bonne utilisation du site. Or, l'on sait que sur le web, un site difficile à utiliser engendre de la frustration chez les utilisateurs qui finissent par le quitter et se rendre ailleurs (Krug 2006, p. 9 ; Bolt et Tulathimutte, 2009). Il est donc important de s'assurer que le visiteur trouve facilement et rapidement ce qu'il cherche.

Il s'agit d'un procédé itératif : un test permettra de mettre au jour un certain nombre de problèmes à régler, mais ces solutions peuvent également engendrer d'autres problèmes. Et des problèmes antérieurs, masqués par ceux que l'on a résolus, peuvent encore subsister. Il est donc préférable de tester souvent et dès le début du processus de conception (Bolt et Tulathimutte, 2009 ; Krug, 2006).

Nielsen (2000) affirme qu'un test avec cinq participants suffit à identifier 85% des problèmes d'utilisabilité d'un produit. Au-delà, les faits collectés deviennent redondants. Toutefois, le nombre de participants conseillé varie selon les auteurs. Pour Rubin et Chisnell (2008), cela dépend des ressources à disposition et des besoins en représentativité, mais pour trouver un grand nombre de problème en un temps relativement court, 4 à 5 personnes suffisent (p.126). Dumas et Redish (1999, p. 128) recommandent 6 à 12 participants. Spool et Schroeder (2001) quant à eux remettent en question le nombre de 5 participants proposé par Nielsen. Dans leur étude, les 5 premiers participants n'ont mis en évidence que 35% des problèmes d'utilisabilité des sites proposés à l'analyse. Ceux-ci étant différents de ceux examinés par Nielsen, le type site testé pourrait avoir un effet sur le nombre optimal de participants.

Les tests d'utilisabilité se déclinent dans une grande variété, tant pour les tests classiques en « laboratoire » que pour ceux à distance, que ce soit en terme d'organisation ou des éléments relevés : think-aloud, eye-tracking, flux de clics, heat maps, intervention active, co-découverte, etc. Elles sont présentées dans de nombreux manuels (notamment Nielsen

1993 ; Dumas et Redish 1999 ; Rubin et Chisnell 2008 ; Bolt et Tulathimutte 2009 ; Barnum 2011).

Des travaux de recherche comparant plusieurs de ces méthodes ont été menés dès les débuts de l'histoire des tests d'utilisabilité (Dumas, 2007). Il s'agissait notamment de vérifier que les tests à distance permettent bien de détecter efficacement les points posant problème (Tullis et al. 2002).

Schmidt (2012) a également comparé le test classique (avec modérateur) et le test automatisé à distance. Dans un premier temps, elle a effectué le test en présentiel avec modérateur, puis le test à distance. Il en est notamment ressorti que le test à distance permettait plus difficilement de signaler des problèmes de gravité moyenne à faible (voir aussi point 3.7).

2.3.2 Etudes Qualitatives et études quantitatives

De manière générale, les premiers tests d'utilisabilité ont été qualitatifs, du fait du petit nombre de participants employés. C'est avec l'apparition des tests à distance, utilisant les technologies du web tant pour la diffusion que pour la réalisation à proprement parler, que des analyses plus quantitatives ont pu être possibles grâce à un plus grand nombre de participants. Elles ont permis ainsi d'assurer une meilleure représentativité et de mettre des chiffres sur certains phénomènes observés. Aujourd'hui, les deux aspects cohabitent dans la littérature, les deux ayant leur place en fonction des buts spécifiques et de la portée des travaux de recherche.

2.3.3 La méthode du think-aloud

La méthode du think-aloud (ou pensée à voix haute) est l'une des plus utilisées pour les tests d'utilisabilité. Il s'agit de demander aux participants de verbaliser leurs pensées, interrogations et réactions, de manière à ce que l'observateur et le modérateur aient accès à leurs processus de réflexion.

Cette méthode possède de nombreux avantages (Nielsen 2012b). Elle est peu coûteuse, car demande peu de matériel. Elle est suffisamment robuste pour obtenir des résultats intéressants, même s'il y a des ratés au niveau méthodologique et suffisamment flexible pour être utilisée tant autour d'un prototype papier (au début de la conception du produit) que d'un site totalement fonctionnel. Enfin, cette méthode est relativement facile à apprendre et permet de confronter les concepteurs ou décideurs aux mots et réactions directes des utilisateurs.

Parmi ses défauts, nous pouvons mentionner le fait que cette verbalisation n'est pas naturelle pour la plupart des gens et que cela peut leur poser quelques difficultés. De plus, certains participants voudront prendre un temps de réflexion avant de verbaliser leurs actions, de peur de dire quelque chose qui pourrait être perçu comme une bêtise. Or, c'est justement les réflexions et pensées à chaud, formulées aussitôt qu'elles ont été pensées, qui intéressent les chercheurs utilisant cette méthode. En outre, les interventions et demandes de clarification du modérateur risquent de biaiser quelque peu la façon dont se comportent les participants au test.

Les manuels de tests d'utilisabilité entrent rarement dans les détails du protocole utilisé lorsqu'ils parlent du think-aloud (Boren et Ramey, 2000, Olmsted-Hawala et al. 2010). Or, il en existe plusieurs. Ericsson et Simon (1993), que reprennent ces auteurs, préconisent le

moins d'intrusions possible de la part du modérateur (le simple « n'oubliez pas de parler »²). Les manuels de test proposent souvent des modérations plus intrusives, telles les demandes de clarification citées plus haut (voir par exemple l'intervention active de Dumas et Redish 1999, p. 31-32) pour s'assurer de ne manquer aucun fait pertinent.

Boren et Ramey (2000) relèvent cette différence entre les principes théoriques annoncés et la pratique. L'application de techniques même légèrement différentes pose le problème de la comparaison des résultats des différentes recherches. Ils proposent donc un cadre théorique alternatif prenant en compte les concepts d'Ericsson et Simon tout en étant adapté aux besoins des tests d'utilisabilité. Ils demandent une écoute active de la part du modérateur (sans toutefois verser dans l'interview ; p. 276). Ils voient cette méthode comme un dialogue et se basent sur le courant « Speech communication », présentant notamment le fait que, du point de vue du langage, celui qui parle demande des interventions de la part de son interlocuteur pour s'assurer qu'il écoute. Le « continuez à parler » souvent préconisé ne constitue pas, dans ce contexte, une contribution très naturelle.

Khramer et Ummelen (2004) ont employé ce nouveau cadre et l'ont comparé à celui préconisé par Ericsson et Simon. S'ils n'ont pas relevé de différence au niveau de la volubilité des participants (nombres de mots et de clics égaux), ils ont en revanche observé que ces derniers étaient moins perdus dans le cadre de Boren et Ramey, et qu'ils y achevaient plus de tâches. Le nombre de problèmes détectés et l'évaluation du site étaient également comparables entre ces deux tests. D'autres tests comparatifs seraient toutefois nécessaires pour sortir des conclusions générales.

Deux principaux types de protocoles émergent de la littérature (Dumas et Redish, 1999, p. 279, Van den Haak, De Jong et Schellens, 2003, 2004 et 2009) : le « concurrent think-aloud » (CTA), où les participants verbalisent leurs pensées en même temps qu'ils réalisent la tâche, et le « retrospective think-aloud » (RTA), où les participants réalisent leurs tâches en silence en étant filmé, puis revoient l'enregistrement et verbalisent à ce moment-là. Le CTA est le plus commun. C'est d'ailleurs sur ce protocole que se base la définition du think-aloud de Nielsen (2012b) et les présentations de la méthode dans divers manuels (par exemple Rubin et Chisnel 2008).

Van den Haak, De Jong et Schellens (2003, 2004, 2009) ont comparé ces différents protocoles en situation de test. Ils ont montré qu'ils étaient comparables (même nombre de problèmes d'utilisabilité relevés, mêmes types et même pertinence), mais qu'en revanche, la façon dont ces problèmes étaient reportés variait. Le CTA relèverait plus de problèmes pouvant être observés durant la réalisation des tâches (non-verbal, etc.), tandis que le RTA relèverait davantage de problèmes uniquement détectables grâce à la verbalisation. Des différences au niveau du taux de succès et du temps de réalisation des tâches ont été remarquées dans certaines études (Van den Haak, De Jong et Schellens, 2003), mais pas dans d'autres (Van den Haak, De Jong et Schellens, 2009). Les auteurs concluent que ce sont des méthodes équivalentes, mais ayant leurs spécificités propres (2003) et qu'elles peuvent fonctionner de manière différente en fonction du type de site testé (2009).

² De leur point de vue, des données obtenues en « interaction avec l'expérimentateur ne sont pas fiables » (Khramer et Ummelan 2004, p.2).

Toutefois, ces aspects méthodologiques du think-aloud nécessitent encore des études plus approfondies, les comparaisons entre les différentes variantes n'étant finalement que relativement récentes.

3. Méthodologie

3.1 Planification générale de l'étude

Après la définition du cadre général de ce projet de recherche, sa planification et la préparation d'un état de l'art, nous avons fait des recherches pour trouver un logiciel adapté à la réalisation de notre étude. Pour plusieurs raisons qui seront détaillées plus bas, notre choix s'est arrêté sur Loop11. Nous avons ensuite commencé à préparer le test proprement dit en définissant tâches et questionnaires. Une fois celui-ci achevé, nous l'avons testé en présentiel avec deux volontaires, puis nous avons lancé nos deux grandes séries de tests : la première à distance automatisée et la seconde avec modérateur.

Les données récoltées nous ont permis de repérer un certain nombre de problèmes d'utilisabilité. Pour chacun d'eux, nous avons comparé les deux méthodes en analysant laquelle a été la plus à même de le mettre en évidence. Dans les sections qui suivent, nous détaillerons chaque étape du processus.

3.2 Choix du logiciel Loop11

Pour réaliser notre test automatisé à distance, nous avons besoin de choisir un logiciel adapté proposant certaines caractéristiques bien précises. Il devait naturellement assurer des fonctions de base, comme permettre de proposer des tâches à réaliser et de faire répondre à des questions. Pour faciliter l'exécution du test, il nous paraissait important que les participants n'aient rien à installer sur leur machine et que l'accès se fasse simplement via url.

Nous souhaitions aussi intervenir aussi peu que possible sur le site e-rara.ch. N'étant pas les gestionnaires, nous n'avions par exemple pas la garantie d'un accord pour l'ajout de code javascript, pourtant un élément souvent requis pour ce genre de test.

Une autre contrainte était le coût: nous ne pouvions y consacrer qu'un montant relativement modeste, ce qui nous fermait des possibilités comme Whatusersdo qui coûte 30£ par utilisateur.

Finalement, la solution retenue devait faciliter un traitement quantitatif des données, c'est-à-dire permettre aisément une analyse statistique.

Suite à des recherches dans diverses sources comme le livre de Bolt et Tulathimutte (2010, p. 202-213) ou des billets de blogs comme ceux de Tomlin (2009) ou de Cameron Chapman (2011), nous avons exploré les fonctionnalités de nombreux logiciels et sur la base de ces critères Loop11 et Usabilla pouvaient convenir. Au final, nous avons choisi le premier, qui a déjà servi lors de plusieurs tests réalisés par des étudiants et chercheurs de la HEG de Genève. Cet outil coûte 350\$ par projet ou 1900\$ annuel. Au niveau de l'exploitation des résultats, il propose plusieurs fonctionnalités aidant à la visualisation, comme les heat maps ou les clickstreams. Les premiers permettent de voir en un coup d'œil les endroits les plus cliqués sur une page (plus une zone est colorée vers le rouge, plus elle a reçu de clics). Toutefois, ce système n'est guère fiable. Dans l'exemple ci-dessous, nous voyons la carte représentant le premier clic de la troisième tâche, qui consistait à repérer le plus ancien ouvrage d'Erasme mis à disposition par e-rara.ch. L'image donne l'impression que l'outil de recherche a été l'élément le plus utilisé alors qu'il n'en est rien. En réalité, l'index des auteurs a davantage retenu l'attention (31 sur 54 clics) et pourtant aucune couleur ne le met en

évidence. Ayant constaté cette faible fiabilité et ayant environ les mêmes informations disponibles par d'autres canaux, nous avons très peu exploité ces heatmaps.

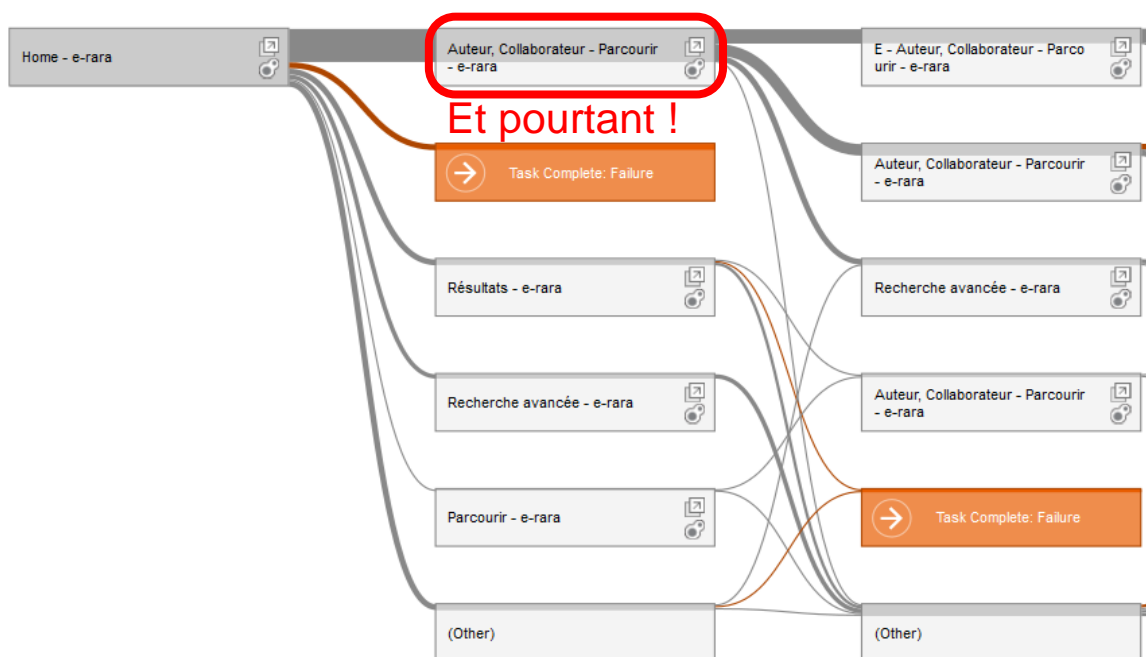
Figure 1 : Exemple de heat map défectueux (tâche 3)³



(Loop11, 2013)

Concernant les clickstream, il s'agit en fait d'une vue sous forme d'arborescence synthétique de l'ensemble des différents parcours des participants au test. L'information ici nous a paru plus fiable. Dans le cas particulier de la tâche 3, il confirme d'ailleurs l'imprécision du heat maps, puisque le premier clic le plus courant apparaît manifestement comme celui de l'index.

Figure 2 : Extrait d'un clickstream (tâche 3)



(Loop11, 2013)

Le logiciel offre la possibilité d'indiquer des pages de succès. Cette fonctionnalité peut se révéler utile, mais il faut pour cela anticiper l'ensemble des url permettant de résoudre la

³ Les notes en rouge consistent en des ajouts afin d'attirer l'attention du lecteur sur certains éléments qui illustrent notre propos. De telles annotations se retrouveront fréquemment en particulier sur les captures d'écran de sites Internet.

tâche. De plus, si une personne passe par une des pages en question, mais accède encore à une autre avant de valider, le système arrivera à la conclusion d'un échec.

Pour la tâche 3, nous avons donné comme url de succès la page de la notice du livre recherché. Dans le test francophone, personne n'a cliqué sur le bouton « tâche accomplie » en visitant cette page spécifique, alors que nous avons malgré tout pu enregistrer près d'un quart de succès.

Par conséquent, les statistiques données par Loop11 sont difficiles à exploiter. Pour dépasser ce problème, il y a essentiellement deux solutions : vérifier le parcours intégral de chaque participant via la liste des pages visitées ou alors poser une question à la fin de la tâche pour vérifier si le participant a bien trouvé l'information recherchée. En général, pour un examen approfondi, il convient de passer par ces deux moyens, qui relèvent d'ailleurs des fonctionnalités de base de la plupart des outils destinés à la réalisation de tests d'utilisabilité. Comme nous souhaitons que les résultats de notre étude soient relativement peu influencés par le choix d'un logiciel en particulier, nous avons donc essentiellement basé nos investigations sur deux sources de données, à savoir les réponses aux questions posées et la liste des pages visitées. Ce sont, de plus, les données les plus robustes et exploitables fournies par Loop11.

3.3 Création du test

3.3.1 Considérations générales

Pour comparer deux méthodes de test, le plus simple est de tester exactement le même objet en changeant aux conditions de déroulement seulement ce qui est directement impliqué par le choix méthodologique. Cela signifie que les questions et les tâches proposées sont strictement identiques dans les deux cas. Il n'y a donc que les modalités de récolte et d'analyse des données qui diffèrent : en partie quantitatives avec le test automatisé à distance et essentiellement qualitatives avec le test modéré en présentiel et le recours au think-aloud. Un tel choix peut sembler en partie discutable. En effet, on pourrait nous opposer l'idée que les tâches auraient dû être adaptées en fonction de l'une et l'autre méthode. Il n'est pas exclu, et même probable, que certains types de tâches se prêtent mieux à l'une des deux formes de test. Par conséquent, même pour repérer des problèmes similaires, il faudrait idéalement proposer des tâches distinctes pour l'une et l'autre méthode. Nous avons préféré négliger cette question, ou plutôt l'aborder de manière inverse. Le fait de proposer un même test est justement le moyen de mettre en évidence ce qui convient le mieux à chacune des manières de procéder. En créant notre dispositif, nous voulions tester le site e-rara.ch non pas dans sa globalité (ce qui en général n'est pas réaliste pour un test d'utilisabilité), mais malgré tout couvrir une part importante et diversifiée des fonctionnalités offertes : comment savoir si une bibliothèque met à disposition certains de ses livres sur cette plateforme, comment s'abonner à un flux rss, comment retrouver un livre spécifique et s'y repérer afin de trouver une information précise à l'intérieur.

Comme e-rara.ch est disponible en plusieurs langues, nous avons souhaité faire le test au moins en français et en allemand. En plus d'élargir le bassin de recrutement des participants, ce choix présente l'avantage de pouvoir comparer des versions similaires d'un même site. En effet, une des applications des tests d'utilisabilité est la confrontation de plusieurs versions afin d'en retenir la meilleure. La différence linguistique permet en quelque sorte de

simuler ce cas de figure et de déterminer laquelle des méthodes est la plus adaptée pour ce type d'objectif.

Dans ce qui suit, nous allons détailler les différentes parties du test et exposer les motivations qui nous ont guidés dans son élaboration. Nous ne mentionnerons que la version francophone, sauf dans les cas où une distinction est pertinente à ce niveau.

3.3.2 Message d'accueil

Une fois que le participant accède au test, il reçoit tout d'abord un message d'accueil qui lui explique comment va se dérouler le processus. Rien ne garantit qu'il va le lire, surtout dans la version automatisée. Il s'agissait donc pour nous d'aller à l'essentiel (on lit plus volontiers ce qui est bref), mais de mettre malgré tout l'ensemble des informations que pourrait attendre un lecteur attentif, comme l'utilité de test ou les remerciements. Même si ce type de contenus ne concerne pas directement l'exécution du test, leur absence pourrait engendrer une perception négative. Nous avons donc établi un compromis entre ces diverses exigences, que nous reproduisons ici :

Bonjour ! Merci d'avoir accepté de participer à ce test d'utilisabilité. Vos retours nous seront précieux et permettront de remplir deux objectifs :

1) l'amélioration du site e-rara.ch ;

2) la réalisation d'un projet de recherche sur la comparaison de deux méthodes de test d'utilisabilité : test automatisé en ligne et test accompagné par un modérateur. Voici encore quelques informations avant de vous lancer dans le test :

- Le système rencontre parfois des lenteurs, cela est dû au dispositif de test et est tout à fait normal.

- Lors de la réalisation des tâches, une barre de navigation verte apparaît en haut de votre écran. Elle permet d'abandonner la tâche ou de passer à la suite du test une fois celle-ci effectuée.

- Ne faites le test que dans une seule fenêtre, pour que le système puisse vous suivre et enregistrer vos actions.

- Le test est complètement anonyme et les données récoltées ne serviront qu'aux besoins de l'étude.

Vous pouvez dès maintenant vous plonger dans le test, et nous vous remercions encore une fois pour votre participation. Si vous avez des questions ou souhaitez être informé des résultats du test, n'hésitez pas à nous contacter.

Valérie Meystre (valerie.meystre@etu.hesge.ch) Raphaël Rey (raphael.rey@hesge.ch)

3.3.3 Questionnaire prétest

Comme le fait observer Barnum (2011, p. 173) dans son manuel, un questionnaire prétest est très utile pour approfondir le niveau de connaissance sur les participants. Cette partie se doit toutefois d'être courte pour ne pas les décourager. Nous nous sommes contentés de quatre questions à évaluer sur une échelle de 1 à 5, concernant des points qui nous semblaient pouvoir avoir un impact direct sur les performances des utilisateurs :

1. Comment évaluez-vous votre expertise dans l'utilisation d'Internet ?
2. Comment qualifieriez-vous votre aisance dans l'utilisation d'une bibliothèque numérique ?
3. Quelle est votre expérience avec le travail et la consultation de livres anciens (en ligne ou sous leur forme originale) ?

4. Avez-vous déjà visité le site e-rara.ch ?

A l'exception de la dernière, effectuer le test aurait pu influencer directement les réponses, puisque les difficultés rencontrées auraient remis en question la perception qu'ont les participants de leurs compétences.

Concernant la dernière, il nous semble plus naturel de la poser avant l'exécution des tâches, mais le choix inverse aurait aussi été envisageable. Notre objectif est de vérifier si les compétences revendiquées par les participants, ou si le fait d'avoir déjà fréquenté le site, est corrélé avec les résultats du test. Pour l'une et l'autre des méthodes analysées dans notre étude, quelle utilisation de ce type de données peut-on faire ? Peut-on, par exemple, profiler les participants sur cette base et émettre par exemple des recommandations différentes en fonction du public que l'on souhaite favoriser ? En définitive, est-ce pertinent de poser ces questions ? Pour répondre à ces interrogations, nous avons voulu trouver différentes corrélations et nos conclusions sont disponibles dans la partie consacrée aux résultats de la comparaison des deux méthodes.

3.3.4 Tâches à réaliser

Avant le début de chaque tâche, nous avons affiché sa description et demandé aux participants d'en évaluer la difficulté sur une échelle de 1 à 5. Une fois celle-ci accomplie, une autre question demandait à nouveau d'évaluer la difficulté de cette même tâche par rapport à ce à quoi s'attendait initialement le participant. A travers ces deux questions, un des buts est d'évaluer à quel point l'utilisation du site peut se révéler frustrante. En effet, quelqu'un qui s'attend à une tâche facile et qui répond ensuite que la tâche était bien plus complexe que prévue a selon une certaine vraisemblance éprouvé un peu d'agacement en réalisant ce qui lui a été demandé. Ce n'est pas forcément le signe d'un problème d'utilisabilité, mais montre à tout le moins que le système ne répond pas aux attentes préalables et que cela mériterait d'être approfondi si possible. Nous verrons dans la suite l'usage que nous avons pu faire de ces couples de questions.

Concernant la formulation des énoncés des tâches, nous avons choisi de mettre en situation les participants au sein de petits scénarios, comme le recommande certains auteurs comme Barnum (2011, p. 128). Lui-même donne peu de justification, mais cela nous semblait renforcer le sens des tâches si elles pouvaient correspondre à des besoins réels, en plus d'accentuer peut-être le caractère ludique du test. Cela implique toutefois des descriptions de tâches plus longues avec des informations non directement nécessaires à leur réalisation.

Afin de vérifier si l'utilisateur a bel et bien réussi les tâches, à la fin de chacune d'entre elles sauf pour la seconde, la plus simple (un seul clic suffisait), nous avons posé une question de contrôle. Par exemple, au terme de la troisième, nous avons demandé l'année de publication d'un ouvrage spécifique. Cette pratique permet de s'assurer facilement si le participant a bel et bien compris ce qu'on attendait de lui, et mieux que si on se contentait d'analyser le parcours suivi. En effet, il est tout à fait possible de visiter une page comportant la solution au problème posé sans en prendre conscience. De plus, ce moyen est rapide pour déterminer si une tâche a été réussie ou non.

Le test dans sa version finale comporte cinq tâches. Une sixième a été supprimée suite à une phase d'évaluation (voir ci-dessous). Pour les premières, proposer aux participants des tâches relativement simples et invitant à la navigation les aide à se familiariser avec le site. Il s'agit aussi de les mettre en confiance pour qu'ils puissent aborder plus sereinement des

manipulations plus complexes. Dans cette optique, nous avons demandé, pour commencer, de vérifier si la Bibliothèque des Pasteurs de Neuchâtel proposait une partie de ses collections sur e-rara.ch ou non. Comme l'ont montré les résultats, de nombreux chemins peuvent mener à la solution.

En deuxième, nous avons demandé de trouver la liste des flux rss. Cela reprenait en partie une des tâches du test réalisé par Eliane Blumer : « Sie möchten gerne über die Neuzugänge informiert werden. Gibt es eine Möglichkeit, dies automatisch zu tun ? » (Blumer 2011, p. 87). Comme un nouveau bouton a été rajouté suite aux résultats de cette étude, il nous paraissait intéressant d'en évaluer l'impact.

A partir de la troisième tâche, nous entrons dans des activités plus complexes avec le repérage d'un ouvrage spécifique : trouver le document le plus ancien dont Erasme est l'auteur présent sur e-rara.ch. Plusieurs chemins permettaient de trouver la solution, soit en passant par l'index des auteurs ou via une recherche (simple ou avancée). Pour cette tâche, il y avait une différence importante entre la version francophone et alémanique du test. En effet, le système ne reconnaît pas l'orthographe française et ne retourne des résultats que si on entre le nom en latin « Erasmus », dénomination qu'utilisent aussi les germanophones. Par conséquent, le test francophone présente une difficulté spécifique supplémentaire et offre la possibilité de comparer les deux méthodes de tests analysées dans cette étude en fonction de leur capacité à évaluer l'impact de cette différence.

Une fois les ouvrages d'Erasme trouvés (via les facettes par exemple), repérer le plus ancien demandait encore de parcourir les résultats (plus de 150) ou alors, plus efficacement, de recourir aux filtres.

Figure 3 : Résultats d'une recherche sur e-rara.ch

HOME

153 Titres

Recherche: erasmus

Trier par: Pertinence Résultats: 10 Ordre: croissant Résultats suivants: 1 - 10

Auteurs, Collaborateurs

Erasmus, Desiderius	153
Leo, Leo	10
Estienne, Jean	9
Gwalther, Rudolf	6
Isocrates	4

Afficher 45 autres résultats

Périodes

1501-1600	149
1601-1700	3
1701-1800	1

Langues

Latin	134
Grec ancien	25
Allemand	17
Hébreu	2

Lieux

Basel	103
Zürich	34
Genève	12
Amsterdam	1
Augsburg	1

Afficher 4 autres résultats

(e-rara.ch, 2013)

Les deux dernières tâches impliquent de consulter directement des documents. Pour la quatrième, nous demandons de trouver quel genre d'oiseau était le griffon pour Buffon dans son *Histoire naturelle des oiseaux* (1770, vol. 1, p. 151). Le participant débutait avec la

notice de l'ouvrage en question à l'affichage. Il devait ensuite trouver un moyen d'atteindre le chapitre consacré à cet animal soit via l'index proposé par le livre lui-même, ou plus simplement par la page des contenus. Une fois le chapitre atteint, il lui suffisait de lire les premières lignes pour trouver la réponse.

Dans ce cas, les germanophones étaient désavantagés puisque le livre de Buffon est en français et que, de plus, nous avons traduit le terme « griffon » en « Greif » dans l'énoncé de la tâche. Il fallait donc dans un premier temps retrouver le terme français avant de pouvoir le rechercher dans l'ouvrage en question. A nouveau, cette différence, même si sa pertinence concernant l'utilisabilité du site est limitée, permet de mettre en évidence la capacité des deux méthodes considérées dans cette étude de mettre en parallèle des versions différentes d'un test similaire.

Quant à la dernière tâche, il ne s'agissait plus de retrouver une information textuelle, mais une image, à savoir celle d'un faucon sort. Le but, ici, était d'inviter les participants à recourir à la possibilité d'afficher des miniatures (mosaïque) des pages afin de parcourir rapidement un chapitre. Pour cette tâche, nous n'avons pas traduit le nom de l'oiseau si bien que la différence linguistique n'a pas vraiment d'autre impact que peut-être psychologique, du fait que la langue du document est moins familière.

Figure 4 : Affichage des pages sous forme de mosaïque



(e-rara.ch, 2013)

3.3.5 Questionnaire posttest

A la fin des tests, on propose, en général, aux participants un questionnaire afin d'évaluer leur vision globale de l'utilisabilité du site web ou du produit testé. Pour ce faire, il existe plusieurs questionnaires standardisés, et en particulier celui de John Brooke, qui, lorsqu'il l'a créé en 1986, ne devait nullement s'attendre à rencontrer un succès tel que celui qu'il a reçu, puisque le SUS (System Usability System) est aujourd'hui le plus utilisé dans ce contexte. Pour analyser l'utilisabilité, ce questionnaire s'inspire des catégories proposées par la norme

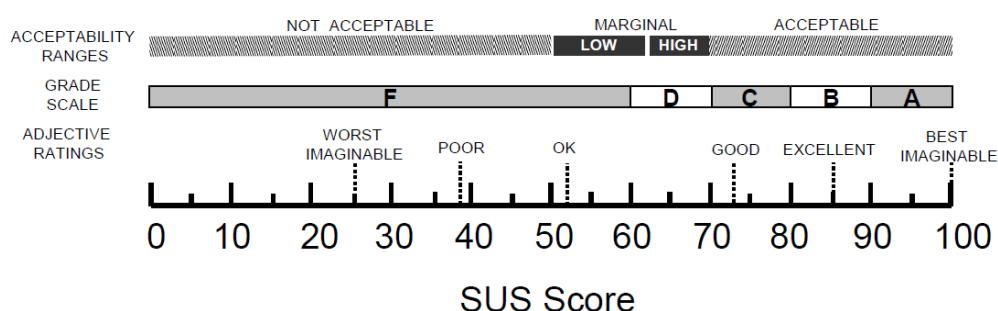
ISO 9241-11, à savoir l'efficacité, l'efficience et la satisfaction liées à l'utilisation d'un système.

Il se présente sous la forme de dix affirmations à évaluer sur une échelle de Likert à 5 niveaux, allant de « strongly disagree » à « strongly agree » et dont voici le détail dans sa version originale (Brooke 1986, p. 4) :

1. I think that I would like to use this system frequently
2. I found the system unnecessarily complex
3. I thought the system was easy to use
4. I think that I would need the support of a technical person to be able to use this system
5. I found the various functions in this system were well integrated
6. I thought there was too much inconsistency in this system
7. I would imagine that most people would learn to use this system very quickly
8. I found the system very cumbersome to use
9. I felt very confident using the system
10. I needed to learn a lot of things before I could get going with this system

Les raisons justifiant l'adoption de ce questionnaire sont nombreuses. D'une part, il se caractérise par sa brièveté. D'autre part, ce test ayant été très largement utilisé, sa robustesse est démontrée. De plus, il permet d'analyser le score de l'objet étudié en le comparant aux résultats obtenus dans d'autres tests. A ce titre, un article coécrit par Bangor, Kortum et Miller (2009, p. 121) définit une série d'adjectifs pour venir interpréter le score du questionnaire SUS :

Figure 5 : Echelle d'interprétation du score SUS



(Bangor, Kortum et Miller 2009, p. 121)

Malgré ces avantages indéniables, nous avons malgré tout choisi de prendre le risque de construire notre propre questionnaire, tout en nous inspirant assez largement de ce dernier. Plusieurs raisons ont guidé notre choix : tout d'abord, la première question nous semblait peu pertinente dans le cas où le test est effectué auprès d'un public large et non spécifiquement le public cible du site, ce qui était inévitable vu notre mode de recrutement via Swiss-lib. En effet, si je n'ai pas de raison particulière de consulter des livres anciens, je pourrais difficilement être complètement d'accord avec l'affirmation « I think that I would like to use this system frequently » (Brooke 1986, p. 4) et ceci complètement indépendamment de toute problématique liée à l'utilisabilité.

Outre cette faiblesse par rapport à notre contexte, le questionnaire SUS nous paraissait encore trop long. Nous posions déjà de nombreuses questions au fil du test et nous ne voulions pas en rajouter un trop grand nombre au terme de celui-ci avec le risque d'ennuyer complètement les participants, risque aggravé si on tient compte d'une certaine redondance qui peut facilement être ressentie lorsqu'on répond à SUS, tant certaines affirmations semblent parfois aborder la même problématique à travers des formulations différentes. Par exemple, on pourrait très difficilement imaginer qu'on puisse être totalement d'accord à la fois avec l'affirmation 5 « I found the various functions in this system were well integrated » et l'affirmation 6 « I thought there was too much inconsistency in this system » (Brooke 1986, p. 4). La corrélation entre les différentes affirmations est manifeste et même recherchée afin de garantir la significativité du score global.

Par ailleurs, un tel score permet surtout d'établir une forme de diagnostic général, mais on observe que, prises séparément, les différentes affirmations du questionnaire SUS ne sont pas très significatives, comme en témoignent les travaux de Bangor et de ses associés (2009).

De plus, la contribution à la découverte des problèmes d'utilisabilité eux-mêmes reste mince comme le fait remarquer Sauro (2011) dans son blog : ce questionnaire ne permet pas de diagnostiquer des problèmes, mais plutôt de faire du benchmarking :

SUS was not intended to diagnose usability problems. In its original use, SUS was administered after a usability test where all user-sessions were recorded on videotape (VHS and Betamax). Low SUS scores indicated to the researchers that they needed to review the tape and identify problems encountered with the interface. SUS can be used outside of a usability test for benchmarking, however, the results won't shed much light on why users are responding the way they are.
(Sauro 2011)

Notre but était donc de créer un questionnaire plus directement exploitable pour détecter de potentiels problèmes. Pour ce faire, nous avons certes pris en compte la norme ISO 9241-11, qui divise l'utilisabilité en trois éléments (l'efficacité, l'efficacité et la satisfaction), mais nous avons également repris des éléments du modèle de Guo (2012), proposé sur le blog UXmatters :

Tableau 1 : Les quatre éléments de l'expérience utilisateur selon Guo

Four Elements of User Experience	
<p>Value</p> <p>Is it useful?</p>	<p>Usability</p> <p>Is it easy to use?</p>
<p>Adoptability</p> <p>Is it easy to start using?</p>	<p>Desirability</p> <p>Is it fun and engaging?</p>

(Guo 2012)

Ce modèle nous permet d'élargir notre perspective à travers la réorientation de certaines parties de SUS et la création de nouvelles affirmations afin de couvrir l'ensemble des quatre champs proposés par Guo et notamment le domaine de l'utilité que le questionnaire SUS ignore ou du moins n'aborde pas directement.

Au final, nous avons demandé aux participants de notre test de se situer par rapport aux six affirmations suivantes, que nous mettons en parallèle avec la matrice de Guo :

Tableau 2 : Questionnaire posttest et axes analysés

Affirmations à évaluer sur une échelle de 1 à 5	Critères selon Guo	Critères retenus
1. J'ai réussi à obtenir des résultats pertinents aux tâches proposées	Usability	Efficacité
2. Ce site est facile d'utilisation (repris de SUS 3)	Usability	<i>Général</i>
3. Ce site est inutilement compliqué (repris de SUS 2)	Usability	Efficiency
4. Je recommanderai ce site, si une de mes connaissances s'intéresse aux livres anciens	Value	Utilité
5. J'imagine que, pour la plupart des gens, apprendre à utiliser ce site ne demande que peu de temps (repris de SUS 7)	Adoptability	Adoptability
6. L'utilisation de ce site m'a paru très agréable	Desirability	Satisfaction

Nous affinons encore ce modèle en précisant la notion d'utilisabilité en reprenant la distinction efficacité-efficience de la norme ISO 9241-11. L'affirmation 1 permettrait donc de recueillir une information concernant l'efficacité alors que la troisième se rapporterait à l'efficience. Quant à la seconde, elle conserve une portée plus générale autour de l'utilisabilité. Au final, nous aboutissons à cinq axes d'analyse : efficacité, efficience, utilité (que nous préférons à une traduction littérale « valeur »), adoptabilité, satisfaction (nous retenons le terme de la norme ISO plutôt que « désirabilité » qui serait la traduction de ce que propose la matrice de Guo).

Ce questionnaire nous permet donc d'établir un diagnostic un peu plus précis qu'un test SUS classique et constitue un indice supplémentaire sur le type de faiblesses que le site risque de comporter en fonction des axes qui semblent présenter un faible niveau de satisfaction.

En plus de ces questions, nous avons également laissé un espace pour que les utilisateurs puissent nous donner leur avis sur le site ou recommander des améliorations. Nous avons pour finir demandé une indication concernant l'âge des participants et s'ils étaient impliqués dans le projet e-rara.ch. Ces derniers ont d'ailleurs été exclu de l'analyse de ce questionnaire (leur implication aurait pu compromettre l'objectivité de certaines réponses).

3.3.6 Prétests

Afin de vérifier la faisabilité, la pertinence et la longueur de notre test, nous avons effectué deux tentatives en présentiel, qui nous ont permis de faire quelques ajustements. Quand nous avons constaté que l'un des participants avait eu besoin de plus d'une demi-heure pour réaliser l'ensemble du processus, nous avons décidé de supprimer une des tâches qui en comptait alors six. Nous avons éliminé la recherche d'un livre d'astrologie du 17^e siècle, qui faisait en partie doublet avec la recherche du plus ancien ouvrage d'Erasmus.

Quant à la tâche demandant de découvrir ce qu'est un griffon dans un livre de Buffon, initialement nous attendions des participants qu'ils retrouvent l'ouvrage dans un premier temps. Nous avons supprimé cette partie, qui rendait la tâche longue et ne centrait pas

suffisamment sur sa part la plus spécifique, à savoir retrouver une information précise à l'intérieur d'un ouvrage.

Ces prétests nous ont aussi amenés à découvrir quelques problèmes d'utilisabilité. Parmi eux, nous avons noté un dysfonctionnement au niveau des outils de tri, qui ne permettaient pas de faire remonter l'ouvrage le plus ancien au sein d'une liste de résultats.

Par ailleurs, un des participants n'a pas trouvé le bouton « mosaïque » et a commencé à parcourir de nombreuses pages dans le but d'atteindre l'image demandée. Même si à ce stade, il est encore trop tôt pour en tirer des conclusions, on peut admettre qu'il y a probablement là un problème à définir.

Ces prétests ont également été l'occasion d'ajuster des questions organisationnelles. Nous avons, par exemple, constaté qu'il était utile de mettre papier et crayon à disposition des participants.

3.3.7 Ordre des deux méthodes de tests

Comme le principal objectif de cette étude est d'analyser dans quelle mesure deux méthodes d'investigation permettent de repérer un même type d'information, à savoir les problèmes d'utilisabilité, il y a un fort risque que la première qui est mise en œuvre influence largement les résultats de la seconde. Celle-ci pourrait se réduire à une tentative de retrouver des informations déjà connues.

Eviter complètement ce biais aurait requis de mener les deux séries de tests de manière complètement indépendante et par des personnes distinctes. Une telle solution aurait toutefois entraîné d'autres inconvénients avec notamment la perte d'une vision globale. En effet, pour le chercheur, avoir mené les deux types de tests permet sans aucun doute de les comparer plus efficacement avec un regard reposant sur une pratique réelle de l'une et de l'autre.

Comme la méthode automatisée à distance permet surtout d'enregistrer des données factuelles sur le comportement des participants, elle requiert une part d'interprétation relativement importante, vu que sur cette base, il faut en déduire les problèmes d'utilisabilité.

Par contre, dans les tests en présence d'un modérateur invitant au think-aloud, les utilisateurs disent explicitement ce qui motive leurs actions et ce qui les dérange. Cela réduit donc le besoin d'interprétation.

En conséquence, l'exploitation des données récoltées par la méthode automatisée à distance se trouve grandement facilitée si on a déjà effectué une étude en présentiel avec un modérateur et que l'on dispose donc d'une connaissance préalable des problèmes que l'on cherche. L'analyse des données risquerait alors de se résumer à trouver la confirmation de problèmes déjà constatés, alors que le but est de déterminer si la méthode permet de les détecter de manière indépendante.

L'inverse est beaucoup moins vrai et le fait d'avoir réalisé une étude à distance automatisée n'influence que peu les conclusions qu'on tire d'un test en présence d'un modérateur, étant donné que la part d'interprétation et de déduction requise pour reconnaître les problèmes est moindre.

Pour ces motifs, nous avons donc décidé de commencer par le test automatisé à distance, y compris l'analyse des données avant de nous lancer dans les tests en présentiel.

3.3.8 Réalisation du test automatisé à distance

Une fois que le test est planifié sur la plateforme Loop11, il suffit de distribuer le lien permettant d'y accéder et d'attendre que les participants l'exécutent.

Concernant le recrutement, nous avons utilisé la liste de diffusion Swiss-lib en faisant un appel à la fois en français et en allemand. Pour le reste, nous avons recouru à nos réseaux personnels, notamment Facebook.

Par notre première source, nous avons essentiellement atteint des spécialistes de l'information, ou du moins des personnes intéressées par les domaines qui gravitent autour, donc pas forcément le public cible du site constitué plutôt de chercheurs et d'amateurs de livres anciens. Toutefois, les lecteurs de Swiss-lib, qui comptent notamment des bibliothécaires, des archivistes et des documentalistes, ne sont évidemment pas étrangers à ces problématiques. Au contraire, on peut s'attendre à ce qu'ils soient habitués à utiliser des catalogues et des bibliothèques digitales, au point que cela pourrait constituer un biais par rapport au public réel d'e-rara.ch.

Conscient de ce problème, nous avons toutefois choisi de mener l'étude avec ce public parce qu'il nous a semblé suffisamment pertinent et facilement accessible.

Quant aux personnes recrutées via Facebook, il s'agit pour l'essentiel de personnes spécialisées en sciences de l'Antiquité, par conséquent très proches de la catégorie des chercheurs qui constituent une part importante des visiteurs d'e-rara.ch.

Lors de l'analyse des données, comme dit précédemment, nous avons essentiellement exploité les réponses aux questions et la liste des pages visitées. Le principal moyen de repérer les problèmes d'utilisabilité est de passer en revue le parcours des participants les uns après les autres et d'identifier les erreurs commises. Avec le nombre, il devient assez aisé d'établir des statistiques rendant compte de leur fréquence et de leur gravité (une erreur qui empêche le plus souvent de trouver finalement la solution est plus problématique qu'une qui ne cause qu'un léger contretemps).

Une fois les erreurs détectées, il s'agit encore de reconnaître les problèmes d'utilisabilité qui en sont à l'origine. Sans interaction directe, il n'y a guère que la question ouverte finale qui offre un espace où l'utilisateur peut exprimer des recommandations ou une opinion par rapport au site. La part d'interprétation est donc importante.

Concernant l'analyse des questionnaires, nous avons essentiellement procédé de manière quantitative. Nous reviendrons sur ce point dans la partie consacrée à l'analyse des résultats (point 4).

3.3.9 Réalisation du test en présentiel

Afin de limiter au maximum les différences parasites entre le test automatisé à distance et le test modéré en présentiel, nous avons également recouru au logiciel Loop11 dans les deux cas. Le modérateur avait essentiellement pour rôle d'inviter les participants à verbaliser leurs pensées, intentions et ressentis au moment où ils accomplissaient les tâches. En plus d'une prise de notes pendant le test, nous avons effectué des enregistrements à l'aide d'un

dictaphone afin de garder une trace des paroles prononcées. Pour le reste, nous pouvions bénéficier des mêmes données que pour le test automatisé à distance vu que Loop11 était également en fonction. Pour synthétiser les résultats, nous avons rédigé des comptes rendus sous forme de fiches disponibles dans l'annexe 2.

Nous avons recruté des participants que nous estimions pouvoir appartenir au public cible d'e-rara.ch, c'est-à-dire trois historiens, une historienne de l'art, une philologue, un écrivain et une spécialiste des banques de données du domaine audiovisuel.

Côté représentativité des langues, nous avons effectué cinq tests en français et deux en allemand.

En principe, sept participants constituent un nombre souvent considéré comme suffisant, même si le fait que le test soit multilingue puisse remettre en cause cette idée. Comme un de nos objectifs est aussi de mesurer l'efficacité des méthodes expérimentées, nous ne voulions pas que le temps consacré à celle-ci soit largement plus important que celui consacré à l'autre.

3.3.10 Méthode de comparaison des deux tests

Les tests d'utilisabilité ont pour fonction essentielle de repérer les problèmes d'utilisabilité dans le but de proposer ensuite des mesures d'amélioration. Afin de comparer les deux méthodes examinées dans cette étude, le plus simple est de partir des problèmes qu'elles ont permis de déceler et de mettre en parallèle chacune d'elles en fonction d'une série de critères définis ci-dessous.

3.3.10.1 Efficacité au repérage des problèmes d'utilisabilité

Par efficacité, nous entendons la capacité de la méthode à repérer un problème d'utilisabilité donné. Pour ce critère, nous nous restreignons aux résultats de nos tests : soit un problème a été clairement identifié, soit observé de manière douteuse, soit complètement ignoré.

Tableau 3 : Variable 1

Capacité au repérage : Echelle identique pour les deux tests
1. Problème passé inaperçu
2. Problème soupçonné
3. Problème identifié

A partir des résultats de nos tests, nous pourrions évaluer assez simplement cette variable pour l'une et l'autre méthode.

3.3.10.2 Probabilité du repérage des problèmes d'utilisabilité

Ici se pose la question de la garantie qu'une méthode repère un problème donné. En général, ce critère est très étroitement lié avec le précédent. Il cherche toutefois à prendre un peu de recul par rapport aux données issues de nos tests en essayant d'évaluer la probabilité de chacune des deux méthodes de découvrir les problèmes passés en revue. Quand les deux tests permettaient de repérer le problème, nous avons fait l'hypothèse que les participants du test à distance représentent un échantillon représentatif pour analyser l'échantillon plus restreint du test modéré. Par exemple, un problème très rare qui n'apparaît

que chez deux participants du test à distance (sur la quasi centaine) a très peu de chance d'apparaître dans le nombre restreint du test modéré. Nous avons défini notre variable en nous appuyant sur un ensemble d'approximations qui n'a d'autre but que de donner une estimation aussi représentative que possible. Si on considère un test avec une cinquantaine de participants, dans le cas où 5% des utilisateurs rencontrent un problème donné, celui-ci a déjà plus de 90% de chance d'apparaître. Par contre, pour atteindre un tel niveau avec seulement six ou sept participants, il faut que le problème soit rencontré par environ 30% des utilisateurs. Dans la construction des paliers de notre variable, nous avons choisi des nombres qui correspondent très approximativement à une probabilité de 90% pour l'échelon très fiable, de 30 à 90% pour le probable et de moins de 30% pour le peu probable. Un échelon « repérage problématique » a été ajouté pour tous les cas où au-delà de la probabilité, c'est la possibilité même du repérage qui est en question. Par ailleurs, dans certains cas, aucun calcul ne peut être effectué si, par exemple, le test automatisé ne détecte pas un problème et qu'une seule personne en rend compte dans le test modéré. Comment évaluer dans ces conditions les probabilités de repérage du test modéré ?

Tableau 4 : variable 2

Probabilité de repérage	Test automatisé	Test modéré
1. Repérage ou calcul problématique	Raison à définir	Raison à définir
2. Repérage peu probable	1 à 3 participants du test automatisé	Moins de 2 participants du test modéré et moins de 5 dans le test automatisé
3. Repérage probable		2 participants ou au moins 5 dans le test automatisé
4. Repérage très probable	Au moins 3 participants du test automatisé	Au moins 3 participants du test modéré ou au moins 15 dans le test automatisé

Le lecteur constatera que pour le test modéré, le fait que le problème ait été détecté ou non par au moins un participant n'est pas fondamental pour le calcul de la probabilité (à moins qu'on puisse montrer que le repérage est problématique pour une raison ou une autre). Nous partons du principe que lors de nos différentes passations, nous avons pu jouer de malchance.

3.3.10.3 Efficience du repérage des problèmes d'utilisabilité

Idéalement, il serait souhaitable de comparer l'efficience des méthodes en calculant le temps nécessaire à chacune d'entre elles pour déceler tel problème et en déduire une proposition d'amélioration, mais ceci n'est que peu envisageable. En effet, le traitement des données, surtout pour celles enregistrées par Loop11 dépend largement de cet outil. Par ailleurs, ce temps repose également en bonne partie sur l'expérience des personnes qui dirigent les tests et n'est pas forcément de la même nature pour les deux procédés. Dans le test automatisé, l'analyse des données requiert beaucoup d'énergie, alors que dans le test modéré, c'est plutôt l'organisation qui exige un investissement. Les problèmes eux-mêmes sont en général détectés via observation directe ou les commentaires des participants.

Par conséquent, nous avons décrit les moyens utiles à la détection du problème. Cela suffit à mettre en évidence les plus gros décalages d'efficience séparant les deux méthodes. Pour

les cas où l'une d'entre elles ne serait pas parvenue à détecter un problème, des conjectures seront émises sur les chemins possibles qui auraient éventuellement abouti à ce résultat. Nous proposons une variable relativement simple : soit le problème peut être repéré via un traitement standard des données (observation directe ou commentaire de la part d'un participant pour le test modéré et analyse des questionnaires ou des flux de clics pour le test automatisé), soit le problème requiert une méthode plus complexe (par exemple le recoupement de plusieurs types de données).

Tableau 5 : Variable 3

Efficienne du repérage : Echelle identique pour les deux tests
1. Repérage requérant une manipulation particulière des données
2. Repérage ne requérant pas de manipulation particulière des données

L'évaluation de la variable sera accompagnée d'une note concernant le mode de détection afin de permettre une comparaison plus fine.

3.3.10.4 Evaluation de la gravité du problème

Résoudre certains problèmes peut se révéler extrêmement coûteux et demander des modifications assez profondes dans le fonctionnement du site. L'intérêt ne manque donc pas à une évaluation de l'impact des problèmes détectés. S'il est fréquent et empêche complètement la réalisation d'une tâche (et ce régulièrement), il nécessite probablement plus d'attention qu'un élément qui n'entraîne le plus souvent qu'une simple perte de temps. Comme les différents critères sont ici indépendants, nous n'avons pas proposé d'échelle, mais plutôt une liste dont chaque élément peut recevoir une validation ou non en fonction de la méthode. Voici les quatre facteurs retenus permettant d'évaluer la gravité d'un problème d'utilisabilité :

- Fréquence du problème
- Niveau d'insatisfaction soulevé
- Taux d'échec de la tâche causé
- Temps perdu ou effort nécessaire pour dépasser le problème

Pour constituer malgré tout une variable (la quatrième), nous avons attribué un point à chaque élément validé ce qui nous amène à une échelle de 0 à 4.

3.3.10.5 Facilitation de la construction de propositions d'amélioration

Quand un problème est détecté, cela ne signifie pas encore que l'on sait comment le résoudre. La variable suivante va évaluer dans quelle mesure une méthode fournit des indications utiles pour proposer une amélioration. Voici l'échelle retenue pour notre cinquième variable :

Tableau 6 : Variable 5

Facilitation de la construction de propositions d'amélioration Echelle identique pour les deux tests
1. Le problème est repéré mais avec peu d'indice sur la manière de le résoudre
2. Une solution est plus ou moins induite

3.3.10.6 Processus d'évaluation

L'analyse de chaque problème d'utilisabilité via ces critères nous a ensuite permis de construire une typologie des problèmes d'utilisabilité en fonction des méthodes les plus adaptées pour les repérer et les analyser. On peut s'attendre à ce que la plupart d'entre eux puissent être détectés avec l'une et l'autre, mais ceux qui restent sont très intéressants parce qu'ils fixent des limites à ces méthodes, limites dont il est capital d'être suffisamment conscient pour évaluer à sa juste valeur les résultats d'un test d'utilisabilité et les failles qu'il est susceptible de laisser invisibles selon la manière dont il a été mené.

4. Analyse des données

4.1 Remarques générales

Dans cette partie, nous mentionnerons les points saillants de l'analyse des résultats de nos tests, en lien direct avec la détection de problèmes d'utilisabilité. Pour une vision plus en détail, nous renvoyons aux annexes 1 pour le test automatisé et 2 pour le test modéré en présentiel.

4.2 Test automatisé à distance

4.2.1 Généralités

94 participants francophones et 66 germanophones ont entamé notre test à distance. Certains n'ont pas dépassé le pré-questionnaire et beaucoup ne l'ont pas terminé. Une part de ces derniers ont toutefois fourni des données pour certaines tâches et ont donc été pris en considération dans les analyses. Parmi les 84 qui l'ont rempli jusqu'à la fin (46 francophones et 38 germanophones), un petit nombre (3 francophones et 6 germanophones) a réalisé le test en dehors de Loop11 (peut-être dans une autre fenêtre ?) et, bien qu'ils aient répondu correctement aux questions, n'ont pas laissé de chemin de clic, donnant l'impression qu'ils avaient simplement cliqué « terminer la tâche » sans parcourir le site. Ils n'ont donc pas été comptés dans les analyses, leurs données n'étant pas exploitables. En revanche, leurs propositions d'améliorations (question ouverte) ont été prises en considération.

4.2.2 Données personnelles sur les participants

En ce qui concerne le pré-questionnaire (compétences des participants), nous prendront en compte ici les 98 participants ayant réalisé au moins une des tâches du test complètement.

Aucun d'entre ne s'est considéré comme novice dans l'utilisation d'Internet. 30,61% se voient comme experts (niveau 5) et près de la moitié (45,92%) donne un niveau de 4. Nous arrivons à une moyenne de 3,86 pour les francophones et 4,26 pour les germanophones sur l'échelle de Likert à 5 points..

La majorité de nos participants estime avoir une aisance moyenne à bonne dans l'utilisation des bibliothèques numériques (21,43% au niveau 3, 50% au niveau 4 et 22,45% au niveau 5 – très grande aisance). La moyenne des francophones se situe à 3,6 et celle des germanophones à 4,2.

La majorité de nos participants n'ont qu'une faible expérience du travail avec les livres anciens : 14,29% et 28,57% pour les niveaux 1 (aucune expérience) et 2 respectivement. Seuls 17,35% se considèrent comme très expérimentés et 21,43 comme bien expérimentés (niveau 4). Les 18,37% restants se donnent le niveau 3. Cela représente une moyenne de 2,6 pour les francophones et de 3,4 pour les germanophones.

Il est important de rappeler que ces estimations sont subjectives et par conséquent difficilement comparables.

Au niveau de la fréquentation du site e-rara.ch, 50% des participants français n'avaient jamais consulté le site auparavant, contre 17,3% des germanophones. Seuls 16,3% de nos participants (9,6% pour le test français et 23,9% pour le test allemand) sont des utilisateurs réguliers. Près de la moitié (48,98%) visite e-rara.ch une fois ou l'autre.

Nous n'avons pas pu effectuer de comparaisons probantes entre les différentes classes d'âge que nous avons délimitées, et ce pour deux raisons. La première est que certaines de ces classes ne comportaient pas suffisamment de représentants pour en tirer des observations significatives. La seconde est la position de la question dans le test. Comme nous l'avons posée à la fin pour ne pas surcharger le pré-questionnaire et entrer rapidement dans le vif du sujet, les participants ayant abandonné le test en court de route n'y ont pas répondu. Il nous manque ainsi des données pour 23 des 98 participants ayant réalisé le test et fourni des données exploitables.

4.2.3 Questionnaire d'évaluation générale du site

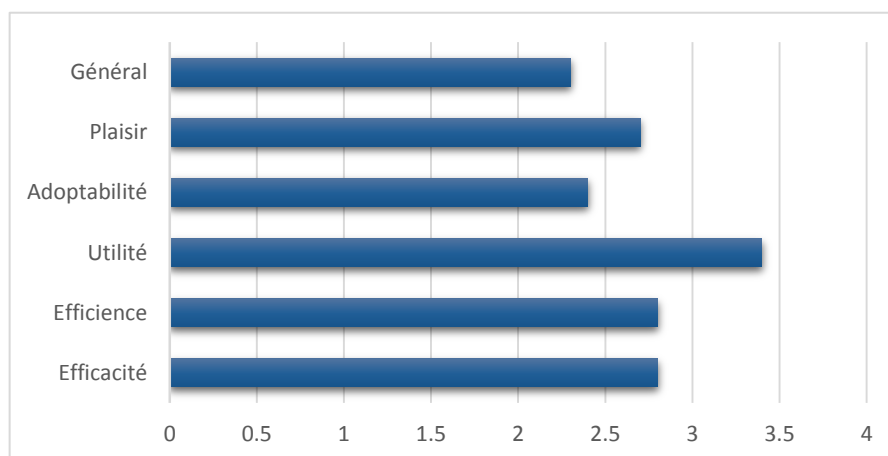
Comme expliqué dans la partie méthodologique, une part importante de ce questionnaire ne sert en réalité qu'à établir un diagnostic général. Pour ce faire, nous avons retiré les résultats des personnes impliquées dans le projet e-rara.ch qui n'avaient à notre avis pas une position suffisamment neutre pour donner une évaluation du site. Cela concerne cinq participants, un francophone et quatre alémaniques.

Les réponses sont en général supérieures au seuil médian, ce qui signifie une appréciation plutôt positive du site. Si, sur le modèle de SUS, nous calculons un score de 0 à 100 avec 0 si tous les participants ont donné chaque fois la pire évaluation et 100 si au contraire ils ont donné chaque fois la meilleure, on obtient un score de 58. Un tel chiffre, si on l'estime avec l'échelle proposée par Bangor, Kortum et Miller (2009) correspond à un score à la limite de l'acceptable.

Il n'y a pas de différences notables entre les tests francophones et germanophones. Notons cependant les différences au niveau de l'adoptabilité, où les francophones se sont montrés plus sévères avec une note moyenne de 2.2 contre 2.7 pour les germanophones sur une échelle de 0 à 4. On observe exactement le même décalage concernant le facteur plaisir : 2.5 pour 3.0. Difficile d'y voir une raison précise liée directement au test et ce d'autant plus que les francophones ont le plus souvent l'impression d'avoir mieux réussi (note auto-attribuée de 2.9 en efficacité contre 2.6 pour les germanophones).

Si on détaille les différents axes comme nous avons projeté de le faire dans la méthodologie, on obtient les résultats suivants :

Figure 6 : Scores moyens du questionnaire posttest (test automatisé)



Tout d'abord, facile de constater que même si le site n'est pas parfait, il est presque toujours considéré comme utile voire très utile. Cela signifie que ses éventuels problèmes

d'utilisabilité ne sont pas tels qu'ils empêcheraient dans la majorité des cas une bonne utilisation.

On prendra acte aussi que l'élément le plus bas est l'adoptabilité. On en déduit que le site n'a pas toujours paru très facile à aborder. Cela se confirme aussi par la note recueillie par l'affirmation générale « ce site est facile d'utilisation »: 2.3 en moyenne (2 étant le choix médian).

Le site semble donc avoir posé d'assez sérieux problèmes aux utilisateurs qui n'ont pas toujours eu le sentiment de trouver aisément leur chemin pour mener à bien les tâches. D'ailleurs, les taux d'échec, on l'a vu, sont assez importants. On observe cependant une assez bonne tolérance de la part des participants, qui considèrent malgré tout le site comme plutôt efficient. La difficulté est présente, mais assez bien acceptée.

Ce questionnaire nous incite à nous montrer particulièrement attentifs à la question de l'adoptabilité. Faut-il mettre en place davantage d'aide, des tutoriels, des bulles informatives, d'autres mesures ?

4.2.4 Analyse des tâches

4.2.4.1 Erreurs

Les différentes erreurs ont été relevées à partir des chemins de clics des participants. Certains ayant réalisé leurs tâches en plusieurs onglets, il a parfois été difficile, voire impossible, de bien comprendre les stratégies guidant leurs choix. Nous n'avons donc relevé que les éléments et actions clairement erronées.

Les tâches 1 et 2 n'ont présentés que peu d'erreurs : la première à cause de la diversité des chemins possibles pour parvenir à l'information demandée, et la seconde du fait de sa brièveté (un clic suffisait, c'est pourquoi les erreurs se sont lues au niveau de l'analyse du premier clic).

En ce qui concerne la 3ème tâche, l'erreur la plus emblématique est sans doute l'échec de la recherche avec le terme « Erasme », qui a conduit 36,9% des participants du test francophone à l'erreur. Les facettes qui diffèrent selon la lettre choisie de l'index auteur a également été cause de confusion pour 25% des participants, et ce pour les deux langues.

Les erreurs de la tâche 4 ont surtout concerné la difficulté à atteindre le chapitre adéquat (27,6% n'y sont pas parvenus) tandis que pour la tâche 5 s'y ajoutaient des difficultés à identifier et utiliser les fonctions proposées (à savoir la fonction « mosaïque »). 33% de nos participants ne l'ont pas utilisée (40% des francophones et 25,7% des germanophones), 20% ont rencontré des difficultés lors de son utilisation (20% dans les deux tests) et 28% (27,5% des francophones et 28,5% des germanophones) ne l'ont utilisée que bien après avoir entrepris la réalisation de la tâche.

Certaines de ces erreurs relèvent clairement des problèmes d'utilisabilité du site. D'autres ne le sont peut-être pas ou indirectement. Sans accès direct aux intentions des participants, il est parfois difficile de trancher. La question ouverte (voir 4.2.5) a souvent permis de réduire ce problème, mais pas entièrement. Pour la présentation des différentes erreurs (sous forme de graphiques), nous renvoyons à l'annexe 1.

4.2.4.2 Temps de réalisation

De manière générale, la question du temps mis à réaliser les tâches ne nous a concrètement pas aidés à identifier des problèmes d'utilisabilité, dans la mesure où l'on ne peut être certain que le temps indiqué par Loop11 ait été entièrement consacré à la réalisation du test. Les participants ont très bien pu être interrompus à un moment ou à un autre, ou avoir des problèmes de connexion. De plus, l'exécution du test en lui-même entraînant des lenteurs, c'est un facteur qui ne nous paraît pas particulièrement pertinent dans le cas du test à distance.

Toutefois, nous avons pensé que le temps moyen de réalisation aurait pu donner des indications intéressantes pour les tâches 3 et 4 puisqu'elles contenaient des éléments de difficultés différents entre le test en français et le test en allemand. Nous avons donc comparé la moyenne francophone et la moyenne germanophone des participants ayant réussi ces deux tâches. Nous avons obtenu respectivement 123,67 secondes contre 140,83 pour la tâche 3 et 118,55 contre 157,78 pour la tâche 4. Ce ne sont donc forcément ceux qui ont la plus grande difficulté qui prennent le plus de temps. Toutefois, il se trouve que seul un petit nombre de participants ont réussi ces tâches, trop peu pour que ces moyennes soient réellement significatives.

Le critère de temps n'est donc pas pertinent pour la seule analyse de notre test à distance. Nous avons trop peu de données et elles ne sont pas assez fiables pour en tirer quelque chose.

4.2.4.3 Projection de la difficulté de la tâche

Nous avons demandé à nos participants d'estimer la difficulté de chaque tâche avant de la réaliser, puis d'ajuster si besoin cet estimation à la fin de la tâche. Nous espérions pouvoir dégager de ces questions des corrélations intéressantes.

Malheureusement, ce ne fut pas du tout le cas. Nous n'avons pas pu retirer d'informations très pertinentes à l'aide de ces questions, les corrélations ne donnant pas de résultats probants. Leur manipulation s'est révélée très complexe et elles ne nous ont pas permis de détecter de problèmes, aussi n'avons-nous pas particulièrement creusé cette piste.

4.2.5 Question ouverte

36 participants ont répondu à la question ouverte et proposé des améliorations, ainsi qu'un certain retour sur d'éventuelles difficultés qu'ils auraient rencontrées. Parmi ces réponses, certaines étaient laconiques (par exemple le participant 21 du teste francophone qui nous dit juste « recherche plein texte » ou le simple « plus rapide » du participant 44), tandis que d'autres donnaient un retour détaillé sur les points les plus problématiques qu'ils avaient rencontrés.

Cette question ouverte est une partie très importante pour la détection des problèmes d'utilisabilité dans un test à distance. En effet, c'est le seul point de contact avec le ressenti et le retour d'expérience des participants. Cela nous donne un éclairage sur leur prestation durant le test, sur des points qui ne sortent pas clairement avec les flots de clics uniquement⁴.

⁴ Par exemple le participant 11 : « Bei den doppelten Seitenzahlen (ursprünglich/neu) muss ich, wenn ich vom Inhaltsverzeichnis ausgehe, jeweils überlegen, welche ich jetzt auswählen soll,

De par le libellé de la question, nos participants ont donc proposé plusieurs recommandations d'amélioration, en fonction des problèmes qu'ils avaient rencontrés. Si quelques personnes signalent directement le problème (par exemple le participant 14 : « une page d'explications complètes pour l'utilisateur »), dans la majorité des cas, il faut remonter au problème à partir de la proposition avancée, voire des interrogations. Par exemple, la question du participant 40 (« gab es eine Möglichkeit nach der Bibliothek zu suchen ohne einen Suchbegriff einzugeben? ») nous montre qu'il n'a pas vu le lien « à propos d'e-rara » dans le texte de présentation du site. Le chemin de clic seul n'aurait pas suffi à obtenir cette certitude.

Autre exemple ; le participant 9 nous dit « J'ai trouvé le chapitre faucon en faisant un ctrl+find dans la page contenu du document. Selon la taille de cette page, il peut être fastidieux de rechercher un mot-clé. ». Nous avons ainsi le rendu d'une manœuvre qui n'est pas visible de par le chemin de clic. Cela nous permet d'en déduire un problème de recherche dans la liste des contenus.

Dernier exemple: le commentaire « Da ich auf einem kleinen Bildschirm suchte, hatte zum Beispiel den rss-feed nicht im Blickfeld » du participant 12 nous permet de remonter à un problème de dynamisme du site.

Onze problèmes d'utilisabilité ont pu être révélés grâce la question ouverte uniquement. Celle-ci est donc primordiale pour exploiter au maximum les possibilités du test automatisé à distance en tant qu'outil. De plus, il permet aussi les retours positifs : « De manière générale, très bon portail, avec une excellente structuration des documents. Bravo! » (participant 50).

4.2.6 Bilan des tests automatisés à distance

De manière générale, la totalité des tests automatisés à distance nous a permis de repérer 18 problèmes d'utilisabilité. Ils sont listés dans le tableau suivant, par ordre de gravité. Nous avons utilisé 3 niveaux de gravité :

- Niveau 1 : problèmes mineurs n'ayant eu que peu d'impact sur le taux de réussite et les pourcentages d'erreurs des participants
- Niveau 2 : problèmes entraînant erreurs et perte de temps (allongement des chemins de clics)
- Niveau 3 : problèmes ayant compromis la réalisation de la tâche

um zum Treffer zu kommen. Bei der Frage nach dem Faulcon sort habe ich eine Weile lang nach dem besten Weg suchen müssen (der beste wäre dann der kürzeste) und bin dabei die Inhaltsverzeichnisse aller Bände durchgegangen. Das dauerte aber zu lang. Nachher merkte ich, dass ich direkt das Faulcon-Kapitel anklicken kann. »

Tableau 7 : Liste des problèmes repérés par le test automatisé

Problèmes repérés	Gravité
1 Absence de renvois entre les différentes dénominations des auteurs	3
2 Fonctionnalité « parcourir » peu claire avec l’affichage des notices de documents	3
3 Navigation et repérage peu aisé dans les miniatures	3
4 Absence d'outil de recherche pour les images	2
5 Accès difficile à l'information à l'intérieur des documents	2
6 Affichage des résultats trop peu clair (pas d'étiquettes comme titre, auteur, etc.)	2
8 Confusion liée à la double pagination	2
9 Fenêtre de recherche peu visible	2
11 Manque d'aide générale	
13 Scroll latéral fréquent sur des écrans étroits	2
15 Limitation dans la recherche des variantes orthographiques au sein des équations de recherche	1
17 Notion de lien permanent peu claire	1
18 Pas de possibilité de classer alphabétiquement les facettes disponibles	1
20 Recherche avancée impossible sans entrer un terme	1
21 Recherche difficile dans la liste des contenus	1
23 Troncature "*" non acceptée	1
24 Visibilité des bibliothèques participantes insuffisante	1
25 Visibilité des liens hypertextes	1

4.3 Test modéré

4.3.1 Généralités

Après avoir effectué le test automatisé et en avoir analysé l'essentiel des données, nous avons effectué le test avec cinq participants francophones et deux alémaniques. Des comptes rendus des différents tests sont disponibles dans l'annexe 2. Ce chapitre va seulement décrire de manière très sommaire et synthétique les résultats des tests. Pour davantage de détails, le lecteur pourra se référer aux fiches détaillées jointes en fin de travail (annexe 2).

4.3.2 Données personnelles sur les participants

Cette partie du test sur les compétences et l'âge des participants nous a en réalité été de peu d'utilité dans la mesure où nous connaissions au préalable les sept personnes choisies pour effectuer ces tests. Nous avons cependant pu constater l'importance de cet élément pour interpréter les actions des participants. Par exemple, nous avons observé que ceux qui ont une grande habitude de travail avec les livres physiques avaient une grande tendance à reproduire exactement les mêmes stratégies alors qu'une plateforme comme e-rara.ch offre des outils permettant de simplifier certaines manipulations (par exemple pas besoin de consulter la table des matières de l'ouvrage lui-même). Dans le design des solutions, penser aux habitudes du public cible du site devient alors intéressant pour essayer de leur faciliter la vie.

4.3.3 Analyse des tâches

L'utilisation d'un dictaphone a rendu quelque peu complexe l'analyse des données, vu que le lien entre le flux de clics enregistré par Loop11 et la parole est déconnecté, mais cela nous a

uniquement quelque peu ralenti dans l'analyse. Une écoute attentive permet de suivre aisément le fil. En plus, recourir à un logiciel comme Morae nous aurait amenés à demander aux participants d'exécuter les tâches dans un environnement particulier alors que là, comme pour le test automatisé, ce sont les participants qui ont pu choisir le matériel qui leur convenait le mieux, sauf les deux premiers du test francophones qui l'ont exécuté dans un environnement non familier : ordinateur portable avec un navigateur ancien et un écran de faible résolution.

Le moyen le plus simple de repérer les problèmes revient à analyser les erreurs. Par exemple, lorsque plusieurs participants entrent « Erasme » dans la fenêtre de recherche et ne trouvent pas de résultats alors que la tâche leur demande de trouver des livres d'Erasme, on peut aboutir à la conclusion qu'il y a là un problème. En effet, le système n'accepte que la dénomination latine « Erasmus ». Dans ce cas, l'ajout d'un système de renvois parviendrait à éviter ce type d'erreur qui a empêché un des utilisateurs du test modéré d'accomplir la tâche et en a sévèrement gêné un autre⁵.

Les commentaires des participants sont également une source précieuse : pour cette même tâche, un participant a choisi plutôt de parcourir les auteurs. Ce lien donne accès à une liste de notices de documents dont au moins un des auteurs commence par A. Un des participants⁶ explique ne pas comprendre la logique de cet affichage : « J'ai déjà demandé de trier par auteur et après il te redemande de trier par auteur, et avec tout ça tu n'as toujours pas le nom de l'auteur à mettre quelque part⁷. » Ce commentaire, souligne le manque de clarté de cette fonctionnalité et une confusion entre les différents outils permettant de filtrer les résultats. Par exemple, si au lieu des notices, on avait une liste d'auteurs commençant par A, le sens de l'affichage serait immédiatement patent. L'impression de confusion des outils existe surtout en raison de cet affichage de notices inadéquat, ce qui constitue en réalité le problème de fond.

Parfois ce sont les propos seuls, qui aident à déceler des problèmes d'utilisabilité comme dans la situation suivante : un participant a décidé de rechercher dans l'index à la fin du dernier volume de l'Histoire naturelle des oiseaux de Buffon le chapitre consacré au griffon. Voyant la longueur de la table, il s'écrie : « ich weiss nicht, ob ich das Wort eingeben kann⁸ ». Par cette exclamation, le participant, qui certes n'a pas choisi la méthode la plus efficiente, émet le désir de pouvoir recourir à une recherche plein texte qui lui faciliterait grandement la tâche.

Pour tous les autres problèmes repérés, nous renvoyons aux fiches situées parmi les annexes qui retracent le déroulement de la passation des tests. Quant aux questions sur l'évaluation de la difficulté avant et après la tâche, elles permettent surtout de s'assurer que la tâche est bien comprise, puis d'obtenir un retour et une analyse de la part du participant une fois la tâche achevée.

4.3.4 Questionnaire d'évaluation générale du site

Les questions qui font suite au test sont l'occasion d'échanger avec les participants sur leurs perceptions du site cible. Ce type d'entretien peut rappeler une forme de « retrospective

⁵ Respectivement les participants 5 et 4 du test francophone, voir annexe 2.

⁶ Plus précisément le participant francophone 3, voir annexe 2.

⁷ Propos du participant 3 au test francophone, voir annexe 2.

⁸ Propos du participant germanophone 2, voir annexe 2.

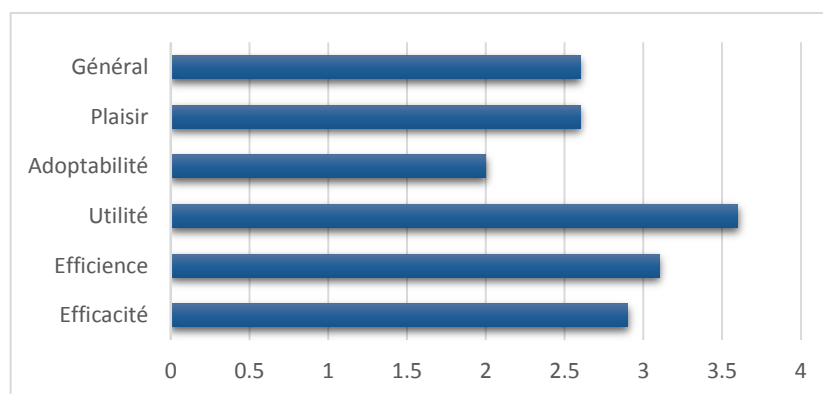
think-aloud » où, même si on ne repasse pas une vidéo des interactions lors de la réalisation des tâches, on demande cependant de revenir sur le test et de le commenter.

Nous avons pu constater que la valeur des commentaires était ici très variable. Tous les participants se sont montrés motivés par les tâches. Par contre, quand il s'agissait après coup d'analyser leurs perceptions durant le test, plusieurs se sont montrés peu riches en renseignements intéressants.

Dans les cas où le participant se montrait volontaire, l'entretien tournait assez vite sur la question des améliorations potentielles et le questionnaire posttest avait surtout pour utilité de fournir un stock de questions pour stimuler la réflexion des participants. Les informations recueillies par ce moyen sont difficilement dissociables de la question ouverte et seront donc analysées dans la section suivante qui y est consacrée. Par contre, si l'utilisateur n'avait que peu de commentaires à donner, le questionnaire en venait à être complété plus ou moins en silence. Quoi qu'il en soit, même si les tests modérés sont essentiellement qualitatif, le questionnaire posttest constitue malgré tout un élément où un peu de quantitatif peut s'insérer.

Sur les sept participants, si on calcule le score global des différentes questions sur le modèle du SUS, on aboutit à un score moyen de 64, ce qui correspond à la limite de l'acceptable selon l'échelle fixée par Bangor, Kortum et Miller (2009, p. 121). Ce résultat est un peu supérieur à celui du test automatisé, qui avait reçu la note moyenne de 58. Voici le détail de la répartition des différents axes :

Figure 7 : Scores moyens du questionnaire posttest (test modéré)



On constate que l'utilité du site n'est pas du tout remise en question, mais que l'adoptabilité en constitue un réel point faible. Ce score faible amène à prendre particulièrement au sérieux tout type de commentaires des participants concernant des carences au niveau de l'aide à disposition et des récriminations concernant des éléments implicites.

4.3.5 Question ouverte

Cette question a été le lieu de commentaires de plusieurs ordres. Certains participants sont revenus sur des éléments qui leur ont posé problème. Par exemple, l'un⁹ d'eux a mentionné un élément jusqu'ici passé inaperçu dans le test. Dans les notices bibliographiques, il n'a pas compris immédiatement que l'illustration du livre était cliquable et permettait de lire le livre. Sur cette base, il sera possible ensuite de proposer une amélioration.

⁹ Participant francophone 3, voir annexe 2.

Plutôt que de mentionner un problème, certains émettent directement des recommandations comme « rendre l'aide plus visible¹⁰ », « améliorer les contrastes¹¹ » ou encore mettre en place des tutoriels vidéo¹². Il est en général utile de remonter au problème afin de s'assurer que la proposition est pertinente et répond de manière appropriée à une difficulté rencontrée.

4.3.6 Bilan des tests modérés

Au final, l'ensemble des tests modérés nous a permis de dresser une liste de 19 problèmes d'utilisabilité dont voici la liste que nous avons classés selon leur degré de gravité :

Tableau 8 : Liste des problèmes repérés par le test modéré

Problèmes repérés	Gravité
1 Absence de renvois entre les différentes dénominations des auteurs	3
2 Fonctionnalité « parcourir » peu claire avec l'affichage des notices de documents	3
3 Navigation et repérage peu aisé dans les miniatures	3
4 Absence d'outil de recherche pour les images	2
5 Accès difficile à l'information à l'intérieur des documents	2
6 Affichage des résultats trop peu clair (pas d'étiquettes comme titre, auteur, etc.)	2
8 Confusion liée à la double pagination	2
9 Fenêtre de recherche peu visible	2
10 Interaction entre les différents menus des outils de tri peu claire	2
11 Manque d'aide générale	2
12 Manque de visibilité du caractère cliquable de l'image du livre des notices	2
13 Scroll latéral fréquent sur des écrans étroits	2
14 Intitulé « provenance du document original » trop peu clair	1
16 Menu horizontal (titre, contenu, etc.) peu visible	1
17 Notion de lien permanent peu claire	1
19 Position du lien vers les fichiers pdf trop à droite	1
22 Termes « croissant » et « décroissant » indifféremment appliqué à des chiffres qu'à des lettres.	1
23 Troncature "*" non acceptée	1
24 Visibilité des bibliothèques participantes insuffisante	1

Les problèmes de gravité de niveau 3 entraînent des erreurs qui empêchent assez souvent l'utilisateur de mener à bien sa tâche. Ils sont source de frustration importante. Ceux de niveau 2 entraînent des erreurs et causent par conséquent des pertes de temps. Ils causent également une certaine insatisfaction. Quant à ceux que nous avons classés au niveau 1, ils constituent des éléments perfectibles, mais qui n'ont eu que peu d'impact sur les participants.

La richesse des résultats dépend grandement de la qualité des participants. Certains se prêtent très volontiers au jeu du think-aloud et s'expriment beaucoup, alors que d'autres ont besoin qu'on leur rappelle régulièrement de verbaliser. Si l'utilisateur reste muet malgré les invitations du modérateur, alors la méthode modérée perd une grande partie de son intérêt. Il suffit d'ailleurs de consulter les fiches de l'annexe 2 pour constater que certains tests ont

¹⁰ Participant francophone 5, voir annexe 2.

¹¹ Participant francophone 2, voir annexe 2.

¹² Participant francophone 1, voir annexe 2.

permis de mettre en évidence beaucoup plus de problèmes que d'autres. Comme le prétend Nielsen (2000) cinq participants peuvent suffire, encore faut-il tomber sur les « bons ». Une part des déconvenues peut s'éviter via un recrutement soigneux en visant toute la diversité du public cible de l'objet à tester, mais il faut vraiment une connaissance approfondie de la personne pour déterminer s'il sera un participant actif et constructif lors de la passation.

5. Comparaison des méthodes

5.1 Généralités

Pour comparer les deux méthodes, nous allons parcourir l'ensemble des parties du test et pour chacune d'elles opposer les apports de l'un et l'autre test. La fin du chapitre est consacrée à deux points particuliers : le premier sera la question de la fraude et des données faussées dans les tests automatisés et la seconde abordera sommairement la problématique de la comparaison de deux versions d'un site afin d'en retenir la meilleure.

5.2 Données personnelles sur les participants

Dans cette étude nous avons récolté des informations sur les compétences dans l'utilisation d'Internet, des bibliothèques digitales et dans la consultation d'ouvrages. A la fin du test, nous demandions également une indication concernant l'âge.

Au fil de nos analyses, nous avons constaté que ces informations n'apportent que très rarement une contribution solide à la détection des problèmes d'utilisabilité. On pourrait alors se demander pourquoi récolter de telles informations si on ne parvient pas que très peu à les exploiter pour proposer des améliorations ?

Concernant le test modéré, mieux connaître les participants permet de contextualiser et de mieux comprendre leurs actions. Nous avons observé que plusieurs personnes dont nous savions qu'elles étaient familiarisées avec la consultation de livres anciens préféraient parfois recourir aux index du livre lui-même plutôt qu'à ceux mis à disposition par e-rara.ch.

Analyser le profil des participants permet aussi de donner une validité à l'échantillon. Il ne s'agit donc plus directement de découvrir des problèmes, mais de s'assurer que les personnes qui ont été retenues pour le test sont représentatives d'une certaine population cible du site.

Dans le cas du test automatisé, la récolte de données remplit essentiellement cette fonction, surtout si le recrutement des participants est relativement aléatoire et basé sur du volontariat, comme dans notre étude. Récolter ces données personnelles permet ensuite de démontrer la représentativité et la diversité par rapport à ce qu'on estime du public cible du site. Dans notre cas, nous pouvions par exemple déplorer la faible participation de personnes au-delà de 65 ans, qui risquent davantage d'être confrontés à des problèmes en lien également avec l'accessibilité.

En de rares occasions, lorsqu'un participant adopte un comportement qui sort de la norme, il peut être intéressant de se pencher sur les données personnelles pour tenter d'y trouver des indices permettant une meilleure compréhension.

Un des avantages du test automatisé réside dans la quantité de données récoltées. Techniquement, il est tout à fait envisageable de calculer des facteurs de corrélation entre plusieurs paramètres. Se pose ensuite la question de l'utilité de découvrir, par exemple, que ceux qui se déclarent moins familiarisés avec Internet commettent davantage telle erreur. Il en va de même avec les questions liées à l'âge. Au mieux ces données aident à comprendre certains phénomènes, mais nous n'avons pas rencontré de cas où elles étaient décisives pour repérer un problème. Rappelons-le : il s'agit là du but du test d'utilisabilité et non d'établir des statistiques, aussi intéressantes soient-elles.

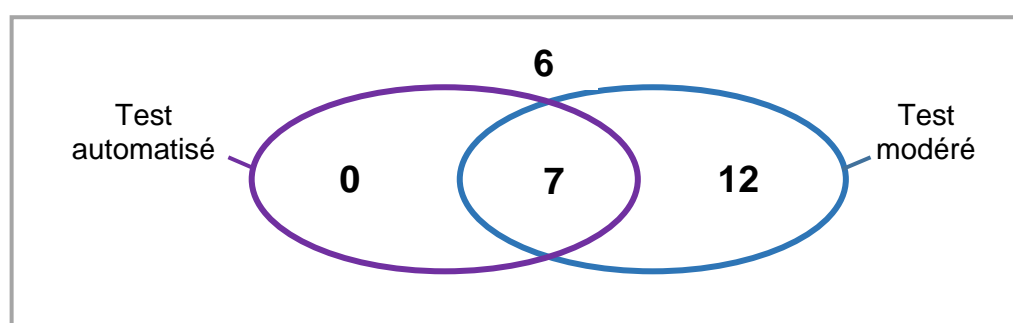
Tableau 9 : L'essentiel en bref : exploitation des données personnelles

Test automatisé	Test modéré
<ul style="list-style-type: none"> + Démontrer la diversité et la pertinence des divers participants. • Apporter dans une moindre mesure un éclairage sur les interactions des participants lors de l'exécution des tâches. • Etablir des corrélations entre ces données et des observations du test est possible, mais ne contribue que peu à la découverte de nouveaux problèmes. 	<ul style="list-style-type: none"> + Démontrer la diversité et la pertinence des divers participants. + Apporter un éclairage sur les interactions des participants lors de l'exécution des tâches.

5.3 Analyse des tâches

L'analyse des tâches est un point faible du test automatisé. Pour s'en convaincre, il suffit de comparer le nombre de problèmes qu'il permet de détecter par rapport à l'analyse des tâches du test modéré :

Figure 8 : Nombre de problèmes repérés via l'analyse des tâches



Il reste 6 problèmes à découvrir par d'autres moyens comme la question ouverte par exemple. D'ailleurs, l'un d'eux n'a été observé que lors des prétests et a passé complètement inaperçu lors des tests réguliers. Il s'agit du problème 7, le menu de l'outil de tri qui passe automatiquement à relevance dès qu'on modifie l'ordre.

Même si les problèmes principaux sont repérés par les deux méthodes, il ne fait pas de doute que le test modéré est largement avantage pour l'analyse des tâches. Première constatation : une plus grande quantité de participants ne compense pas la perte du think-aloud et de l'observation directe concernant le repérage des problèmes lors des tâches.

Nous pouvons toutefois nuancer ce constat si nous prenons en considération que le test automatisé laisse encore suspecter 6 problèmes (tous constatés par le test modéré). Ces derniers présentent des points communs qui mettent en évidence une faiblesse de la détection des problèmes via le test automatisé :

Tableau 10 : Liste des problèmes suspectés (test automatisé)

Problèmes repérés	Gravité
6 Affichage des résultats trop peu clair (pas d'étiquettes comme titre, auteur, etc.)	2
9 Fenêtre de recherche peu visible	2
10 Interaction entre les différents menus des outils de tri peu claire	2
12 Manque de visibilité du caractère cliquable de l'image du livre des notices	2
14 Intitulé « provenance du document original » trop peu clair	1
24 Visibilité des bibliothèques participantes insuffisante	1

Les problèmes 9, 12 et 24 concernent des questions de manque de visibilité. En effet, dans le test automatisé, mis à part constater qu'un certain nombre d'utilisateurs n'a pas utilisé une fonctionnalité, il est difficile de décider s'ils ont vu ou non un certain élément, surtout si plusieurs chemins peuvent mener à la solution. Dans le cas du problème 24, qui concerne surtout la première tâche où le taux de succès est assez élevé (plus de 75% entre le test germanophone et francophone), il n'est pas évident de conclure qu'il y a un problème de visibilité, puisque les fausses pistes suivies sont assez diverses (confusion entre bibliothèque et imprimeur, confusion entre plusieurs bibliothèques, etc.¹³). On peut néanmoins supposer que cette dispersion est le fait d'un problème d'utilisabilité, mais le test automatisé ne permet que difficilement de le définir précisément et de confirmer son existence.

Les éléments 6, 10 et 14 impliquent un défaut de compréhension. Sans le think-aloud, l'accès aux pensées des participants et à la motivation de leurs actions est très limité. L'ensemble des problèmes qui n'entraînent pas directement une erreur spécifique et qui relèvent d'une mauvaise interprétation de certains éléments de la page est très difficile à détecter via l'analyse des tâches d'un test automatisé.

Une autre catégorie qui pose des difficultés à cette méthode est celle des problèmes invisibles dans le flux de clics et qui se manifestent essentiellement à l'affichage. Par exemple, le bug du problème 7, qui cause le retour du critère de tri à l'option « relevance » lorsqu'on modifie l'ordre. Le problème du scrolling horizontal (élément 13) n'est non plus pas aisément repérable si on ne voit pas soi-même le participant interagir avec l'interface.

De son côté, le test modéré semble pouvoir contribuer à repérer n'importe quel problème. Il présente toutefois une limite lui aussi : le nombre réduit de ses participants. Le risque de passer à côté de problèmes marginaux ou peu fréquents par rapport aux tâches proposées est relativement élevé. Par exemple, le bug du problème 7 (qui n'est pas si marginal) n'a été repéré que dans les prétests.

Nous avons constaté également une grande disparité dans les contributions des divers volontaires. Certains se sont montrés très peu à l'aise avec le think-aloud et n'ont finalement communiqué que peu d'informations sur leur manière d'envisager les tâches et de les exécuter¹⁴.

Si le test automatisé paraît désavantagé par rapport au repérage des problèmes, il présente malgré tout d'autres atouts. Dans le cas où l'analyse des tâches permet de révéler le problème, avec la quantité de participants, il est également assez souvent possible d'en

¹³ Voir les types d'erreurs dans l'annexe 1, tâche 1.

¹⁴ Par exemple, le participant francophone 5 a apporté une.

déterminer la fréquence, le taux d'échec engendré et l'effort nécessaire pour surmonter l'obstacle (surtout pour les problèmes sévères qui occasionnent des erreurs dont il est facile de compter les occurrences). Ces données sont particulièrement intéressantes pour évaluer l'opportunité d'apporter une amélioration, surtout si celle-ci entraîne un coût important (dans notre étude, nous proposons par exemple l'ajout de notices d'autorité avec des renvois pour améliorer le repérage des noms d'auteur dans différentes langues).

Les données récoltées par le test modéré sont d'une autre nature. Par exemple, à moins que presque tous les participants rencontrent une difficulté (problème 3 avec la navigation dans les miniatures), il n'est en général pas possible de se faire une idée précise de la fréquence ni du taux d'erreur. Par contre, à travers l'observation, on peut évaluer au cas par cas les efforts nécessaires pour dépasser une difficulté et surtout constater comment le participant l'appréhende : on remarquera éventuellement un état de doute ou d'agacement par exemple.

Au final, le type de données qu'on souhaite produire au terme de l'étude est un critère fondamental dans le choix d'une méthode de test.

Tableau 11 : L'essentiel en bref : analyse des tâches

Test automatisé	Test modéré
<ul style="list-style-type: none"> + Repérage essentiellement des problèmes sévères ou qui engendrent des erreurs. + Evaluer la fréquence des problèmes sévères (et de quelques autres). + Evaluer le taux d'échec causé par les problèmes sévères (et quelques autres). + Evaluer l'effort nécessaire pour dépasser le problème (problèmes sévères et quelques autres). • Repérage difficile des problèmes liés à une incompréhension d'éléments de l'interface. • Repérage difficile de problèmes liés à un manque de visibilité d'éléments de l'interface. • Repérage difficile de problèmes d'affichage non décelables dans le flux de clics. 	<ul style="list-style-type: none"> + Repérage efficace de n'importe quel type de problème. + Evaluer régulièrement la manière dont le participant appréhende les problèmes d'utilisabilité. + Evaluer l'effort nécessaire pour dépasser le problème. • Repérage hasardeux des problèmes marginaux ou rares.

5.4 Questionnaire d'évaluation générale du site

Ce questionnaire comporte dans notre étude six affirmations à évaluer sur une échelle de Likert à cinq niveaux. Lors du test automatisé, les participants n'ont pas la possibilité de commenter leurs réponses aux questions. Ces données permettent donc surtout d'établir un bilan général de l'utilisabilité du site étudié. Dans une certaine mesure, il est possible de détailler différents axes et d'analyser séparément plusieurs facettes de l'utilisabilité. Ces informations ne contribuent pas directement à la découverte de problèmes d'utilisabilité, mais facilite la compréhension des forces et faiblesses du site et favorise la pratique du benchmarking, en particulier si plusieurs études sont menées successivement sur le même

site. Quant au score moyen, il permet d'établir une note globale d'utilisabilité au site (58/100). Tous ces éléments sont utiles pour établir un diagnostic général de la situation.

L'utilité est bien plus manifeste dans le test modéré où chaque question est l'occasion de revenir éventuellement sur un ou plusieurs éléments du déroulement du test. Il suffit au modérateur de demander de justifier les réponses en demandant des exemples. Ces approfondissements sont particulièrement efficaces pour l'affirmation 2 (« ce site est inutilement compliqué ») et l'affirmation 6 (« l'utilisation de ce site m'a paru très agréable »), où en fonction de la réponse donnée, on demande ce qui a déplu.

A titre d'exemple, voici les propos d'un participant auquel on demandait d'expliquer pourquoi il ne manifestait que partiellement son désaccord par rapport à l'affirmation : « ce site est inutilement compliqué ». La raison était qu'il trouvait les pages un peu surchargées avec beaucoup d'options, certaines peut-être superflues selon les choix précédents de l'utilisateur¹⁵. On ne peut pas en déduire directement un problème d'utilisabilité précis, mais donne pour le moins des indications pertinentes : impression de surcharge et que tout ne paraît pas utile au même titre.

Lorsqu'on en est venu à la question de la satisfaction, nous avons demandé à ce même participant ce qui lui a déplu. Il a mentionné, les couleurs et les contrastes trop faibles. Nous avons récupéré cette idée de contraste insuffisant pour nos propositions sur la fenêtre de recherche notamment. Comme on peut le constater, sur la base de cet exemple, avec l'affirmation 2, on obtient des informations plutôt sur l'architecture et le fonctionnement du site tandis qu'avec la dernière cela touche en général davantage le visuel.

Dans une moindre mesure, il est aussi possible de calculer un score global d'utilisabilité à partir de ces questions pour le test modéré. Toutefois, vu le nombre réduit de participants, la fiabilité d'une note moyenne reste limitée (64/100). Les perceptions des utilisateurs peuvent être très diverses, et leurs opinions ne reflètent pas toujours ce que nous avons pu observer pendant les tests. Par exemple, un des participants germanophone (le premier) a connu passablement de difficultés dans certaines parties du test et a décidé que c'était lui qui n'avait pas trop l'habitude de ce genre de système, mais que celui-ci était parfait. Au final, il a attribué les notes maximales à chaque affirmation du questionnaire posttest. Sur un grand nombre de données, une moyenne prend son sens, mais sur un petit nombre, cela permet surtout de comprendre comment les utilisateurs individuellement perçoivent le site. On s'éloigne donc du diagnostic général établi à l'aide du test automatisé, qui favorise une analyse quantitative des données.

¹⁵

Voir annexe 2 participant francophone 2

Tableau 12 : L'essentiel en bref : exploitation du questionnaire d'évaluation générale

Test automatisé	Test modéré
<ul style="list-style-type: none"> + Etablir un diagnostic de l'utilisabilité d'un site. + Pratiquer du benchmarking par rapport à des états antérieurs ou des sites similaires. 	<ul style="list-style-type: none"> + Contribuer à la découverte des problèmes en demandant au participant de commenter les réponses. + A travers une série d'exemple, mieux comprendre les différentes perceptions des individus par rapport au site. • Dans une moindre mesure, établir un diagnostic de l'utilisabilité d'un site.

5.5 Question ouverte

La question ouverte offre un espace pour les participants qui souhaiteraient émettre des recommandations. En voici le texte francophone : « Quelles améliorations pourraient être apportées au site e-rara.ch afin d'augmenter son confort d'utilisation ? ». Les informations que livrent les participants dans cet espace se sont révélées capitales pour découvrir bon nombre de problèmes, en particulier ceux de faible gravité.

Dans le test automatisé, cela représente 11 problèmes repérés essentiellement grâce à cette question et qui n'ont pas du tout été repérés via l'analyse des tâches ou alors seulement suspectés. Certes, ces problèmes sont relativement peu importants, mais cela montre malgré tout à quel point l'efficacité de ce mode de test repose sur la question ouverte¹⁶.

Tableau 13 : Liste des problèmes repérés via la question ouverte (test automatisé)

Problèmes repérés	Gravité
4 Absence d'outil de recherche pour les images	2
6 Affichage des résultats trop peu clair (pas d'étiquettes comme titre, auteur, etc.) (suspecté par l'analyse des tâches)	2
9 Fenêtre de recherche peu visible (suspecté par l'analyse des tâches)	2
11 Manque d'aide générale	2
13 Scroll latéral fréquent sur des écrans étroits	2
15 Limitation dans la recherche des variantes orthographiques au sein des équations de recherche	1
18 Pas de possibilité de classer alphabétiquement les facettes disponibles	1
20 Recherche avancée impossible sans entrer un terme	1
21 Recherche difficile dans la liste des contenus	1
24 Visibilité des bibliothèques participantes insuffisante (suspecté par l'analyse des tâches)	1
25 Visibilité des liens hypertextes	1

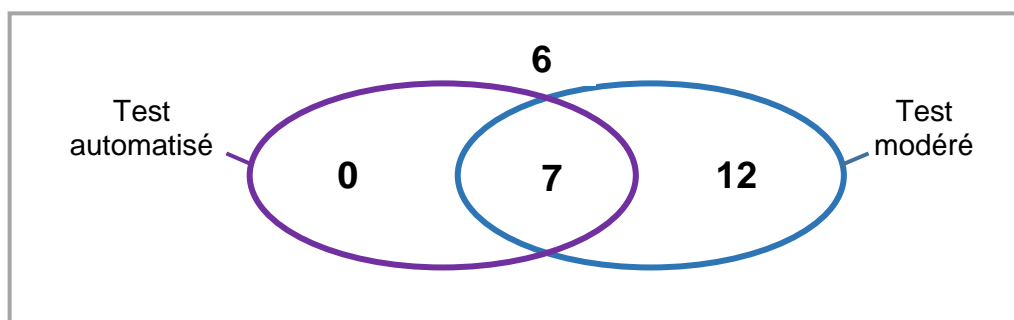
Comme nous l'avons vu dans la section précédente, l'analyse des tâches des tests automatisés ne parvient que très difficilement à repérer certains problèmes d'utilisabilité. Cette question vient donc comme une sorte de filet de sécurité pour récupérer une partie de

¹⁶ Schmidt (2013, p. 56) souligne aussi la grande valeur des commentaires pour la détection des problèmes.

ce qui n'a pas pu l'être auparavant, ce qui rétablit un équilibre approximatif avec le test modéré.

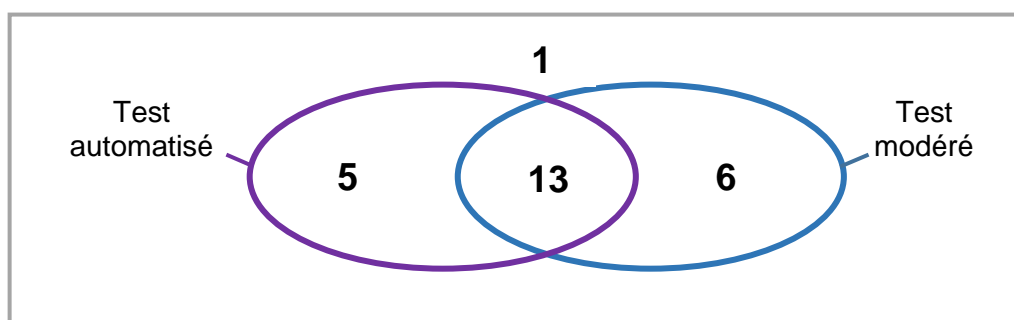
Sans prendre en compte cette question pour le test automatisé, la différence entre les deux méthodes serait écrasante, même si les trois principaux problèmes sont repérés par l'une et l'autre :

Figure 9 : Nombre de problèmes repérés par méthode (sans question ouverte pour le test automatisé)



Grâce à cette question, on rétablit une certaine parité, qui nous amène au compte suivant :

Figure 10 : Nombre de problèmes repérés par méthode



Du côté du test modéré cette question revêt moins d'importance. Elle n'a permis de découvrir qu'un seul nouveau problème : le manque de visibilité du caractère cliquable de l'illustration du livre dans les notices (problème 12). En général, les participants disent surtout qu'il n'y a pas assez d'aide, ce qui est probablement vrai étant donné sa quasi absence. Même si certains participants voient peut-être là le principal problème, le bon sens veut que si on ajoute un manuel à un système, on pourra peut-être mieux l'utiliser, mais cela ne va pas le rendre beaucoup plus pratique s'il comporte à la base des déficiences.

Il n'est pas étonnant que certains participants soient relativement avares d'information pour cette partie du test. N'oublions pas que pour la plupart, ils ont déjà une trentaine de minutes de test derrière eux et qu'ils ont déjà eu plus que largement l'occasion de s'exprimer. Notons aussi que seuls 36 participants du test automatisé sur les 86 qui en sont arrivés à ce point ont laissé un commentaire dans la question ouverte (nous ne tenons pas compte ici de la pertinence), soit environ 40%. De plus, il s'agissait pour eux de la seule opportunité pour formuler une pensée plus élaborée.

Sur sept participants, même si le modérateur incite davantage à s'exprimer qu'une question seule sur un écran, on peut s'attendre à ce que plus de la moitié n'ait pas vraiment de commentaire particulier à rajouter. Ils utilisent alors une réponse par défaut du type « il n'y a pas assez d'aide » (revient cinq fois) ou alors évitent de se prononcer à l'image du

participant germanophone 1 (voir annexe 2), qui ne s'estimait pas suffisamment expert pour dire quoi que ce soit.

Revenons encore brièvement au test automatisé. Avec le nombre élevé de ses participants, cette question ouverte augmente les chances de trouver des problèmes marginaux. Cela signifie que des éléments qui ne se présentent que rarement ou qui ne sont pas directement examinés par les tâches planifiées ont passablement de chance de ressortir malgré tout. On peut penser au problème 20, qui empêche de lancer une recherche avancée sans équation, par exemple uniquement en sélectionnant une institution. Il arrive aussi qu'une remarque dénonce un problème, mais qu'aucun autre élément ne permette de le confirmer ou de l'infirmer. Une faiblesse de ces commentaires réside dans le fait qu'ils sont déconnectés de la spontanéité d'une action et relèvent déjà d'une démarche d'analyse. Il y a par conséquent un risque de s'embourber dans de faux problèmes, surtout quand il s'agit de cosmétique. Peut-être est-ce le cas avec le problème 25 et l'affirmation que les hyperliens ne sont pas très visibles.

Tableau 14 : L'essentiel en bref : exploitation de la question ouverte

Test automatisé	Test modéré
<ul style="list-style-type: none"> + Repérer des problèmes (de tout type) ou en confirmer l'existence (en particulier ceux qui ont échappé à l'analyse des tâches) + Repérer des problèmes rares. + Information abondante récoltée (en raison du grand nombre de participants). ✗ Commentaires non spontanés liés à une action, risque de se perdre dans de faux problèmes. 	<ul style="list-style-type: none"> • Possibilité modérée de trouver de nouveaux problèmes d'utilisabilité. • Peu d'informations pertinentes récoltées (en partie à cause du nombre réduit de participants). ✗ Commentaires non spontanés liés à une action, risque de se perdre dans de faux problèmes.

5.6 Fiabilité des tests

Le participant 11 germanophone a écrit ces mots parmi d'autres considérations dans la question ouverte : « Bei der Frage nach dem Faulcon sort habe ich eine Weile lang nach dem besten Weg suchen müssen (der beste wäre dann der kürzeste) und bin dabei die Inhaltsverzeichnisse aller Bände durchgegangen. »

Quand on observe le chemin de clics de ce participant, on constate qu'il ne contient que le chemin le plus court avec le passage par le menu « contenu ». Voici encore un autre exemple : le participant 56 du test francophone explique ses difficultés pour retrouver Erasme, mais lorsqu'on consulte les pages visitées, il a systématiquement validé la tâche dès la première page. Ces deux participants décrivent leurs actions et les difficultés rencontrées, mais qu'en est-il des autres ? Combien d'entre eux ont-ils fait des essais avant de valider le test ? Combien ont en réalité effectué plusieurs fois le test ? Qu'en est-il des problèmes techniques ?

On peut supposer que l'effort pour falsifier délibérément les résultats est suffisamment grand pour décourager la plupart de ceux qui seraient pour une raison ou une autre tentés de le faire, toutefois les raisons pouvant amener des données parasites ne manquent pas et celles-ci sont difficiles à détecter.

Naturellement ce problème se pose beaucoup moins dans les tests modérés puisqu'il est possible d'observer directement les actions des participants.

5.7 Capacité à comparer différentes versions d'un site

Pour simuler la comparaison de plusieurs versions d'un site, nous allons prendre en considération la différence linguistique. La troisième tâche francophone demandait de retrouver le plus ancien livre présent sur le site et dont Erasme est l'auteur. La difficulté consistait dans le fait que le système ne retourne aucun résultat pertinent si on effectue une recherche avec « Erasme » : il fallait indiquer le nom latin « Erasmus », nom que nous avons indiqué dans la description de la tâche dans sa version alémanique.

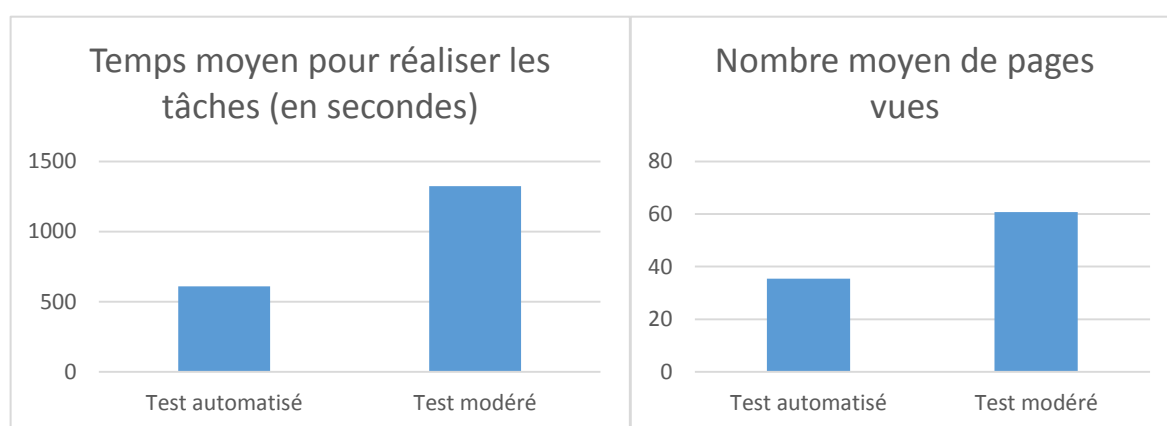
A travers le test automatisé, il est ensuite facile de comparer le taux de succès de cette tâche en fonction de la langue : 61% de réussite parmi les germanophones et 22% parmi les francophones. On peut donc constater que la différence est de près de 40%. Plus d'un tiers des participants au test français ont commis l'erreur de chercher « Erasme ». Ces chiffres permettent d'évaluer assez précisément l'impact d'une différence précise entre deux versions d'un site.

Un test modéré parviendrait beaucoup moins efficacement à donner une telle estimation. Il permet au mieux au cas par cas de constater que telle difficulté est rencontrée ou non. Dans ce contexte, les analyses quantitatives présentent un avantage indéniable.

5.8 Attitude des participants

Pour clore ce chapitre, nous allons nous poser la question de savoir si une différence significative dans les interactions des participants existe selon que le test est modéré ou automatisé. Afin de tenter d'apporter une réponse, nous allons comparer deux critères relativement simples : le temps moyen employé à réaliser les tâches et le nombre de pages vues.

Figure 11 : Comparaison du temps et du nombre moyen de pages vues entre les participants du test automatisé et ceux du test modéré



Ces chiffres ne prennent en compte que les participants qui ont terminé l'ensemble des tâches et répondu à toutes les questions, ce qui est le cas pour tous ceux qui ont pris part au test modéré. Même si pour ce dernier, les résultats se fondent uniquement sur 7 personnes, le constat est sans appel. Les participants au test modéré y passent deux fois plus de temps et visitent deux fois plus de pages. Au cours des différentes passations, plusieurs utilisateurs nous ont paru faire preuve d'acharnement pour trouver une solution. La présence d'un

observateur n'est donc pas sans conséquence et exerce probablement une forme de pression sur les participants qui ne veulent pas se retrouver en situation d'échec, même si on leur explique bien qu'on évalue un site et non leur prestation. D'une certaine manière, ce facteur peut compenser en partie le fait que les tâches ne répondent pas à des besoins spécifiques des participants. Quant au test automatisé, les éléments de motivation sont relativement faibles hormis la volonté de contribuer à l'amélioration d'un site.

Il est difficile d'évaluer l'impact réel de ces différences sur la capacité des tests à détecter des problèmes d'utilisabilité. Quoi qu'il en soit, il faut avoir conscience que le comportement des participants est différent en fonction de la méthode de test utilisée.

6. Conclusion

L'ensemble de nos résultats est décrit en détail dans le chapitre précédent. Dans la conclusion nous reviendrons uniquement sur les objectifs de ce travail de recherche. Le premier consistait à comparer l'efficacité et l'efficacité des deux méthodes de test utilisées.

Commençons par la question de l'efficacité. Nous avons pu montrer que l'analyse des tâches du test automatisé était largement moins performante que celle du test modéré, du moins en quantité de problèmes détectés. Toutefois, les éléments qui entravent le plus l'utilisabilité sont détectées dans les deux cas.

Avec un grand nombre de réponses, la question ouverte devient pour le test automatisé un outil essentiel pour rétablir une certaine parité dans le nombre de faits problématiques constatés.

La différence entre les deux formes de test, plutôt que dans le nombre de problèmes détectés, consiste dans la qualité des données. Avec le test automatisé, on peut assez souvent (et presque toujours pour les plus sérieux) évaluer leur fréquence et leur impact en terme de taux d'échec aux tâches et d'effort nécessaire pour les dépasser. Dans le cas du test modéré, il n'est en général pas envisageable d'établir de telles statistiques, par contre on observe mieux au cas par cas l'impact réel des problèmes sur les participants : agacement, sentiment d'être perdu, etc.

Concernant l'efficacité, nous avons tenté d'estimer l'effort nécessaire pour réaliser le test avec l'une et l'autre méthode, mais comme les deux approches sont assez radicalement différentes au niveau des tâches qu'elles impliquent, un comparatif était très compromis. En effet, comment comparer le temps nécessaire pour organiser un test modéré (déplacements, enregistrement, exploitation, etc.) et celui pour analyser les résultats du test automatisé ? Dans ces conditions, plutôt que de comparer l'efficacité du point de vue des organisateurs du test, pourquoi ne pas aborder cette question en considérant la perspective des participants ? Si on ne considère que les tâches, les 160 personnes qui ont pris part au test automatisé y ont passé en tout près de 19 heures. Les sept participants cumulés des tests avec modérateur y ont passé un peu plus de trois heures au total. A ces chiffres, il faut encore ajouter le temps passé à répondre aux questions et à lire les énoncés, ce qui, même s'il n'est pas possible avec nos données d'en fournir une estimation exacte, revient environ à doubler ces temps.

Dans un contexte économique, il faudrait probablement rémunérer les participants et, même si on peut supposer que recruter des personnes pour un test à distance soit meilleur marché qu'en présentiel, il reste que l'énergie humaine dépensée pour le test automatisé est largement supérieure à celle du test modéré pour des résultats pas fondamentalement meilleurs, mais qualitativement différents.

Au final, quelle est la meilleure méthode ? Le plus efficace évidemment serait de recourir aux deux, ce qui garantit le repérage d'un maximum de problèmes, mais cela n'est pas forcément pertinent (à moins de disposer de gros moyens) vu que l'une et l'autre méthode permettent de repérer les points les plus importants. Pour se déterminer entre elles, il y a plusieurs critères à prendre en compte comme les contraintes organisationnelles (a-t-on accès facilement ou non à une population de participants potentiels ?) ou le type de données dont on a besoin. Par exemple, s'il s'agit de convaincre une autorité d'investir de l'argent

pour corriger des problèmes d'utilisabilité, si ces montants sont importants, il vaut mieux disposer de données quantitatives. Il est bien plus percutant de pouvoir affirmer que 20% des visiteurs qui souhaitent acheter un article sur le site Internet interrompent le processus à tel endroit en raison d'un formulaire problématique plutôt que d'expliquer que lors des tests, on a pu observer qu'un des utilisateurs a eu un problème avec un formulaire.

Dans la majorité des cas, s'il s'agit uniquement de repérer les problèmes d'utilisabilité, un test modéré nous semble malgré tout très efficace et demander un investissement moindre en ressources.

Concernant le second objectif de ce travail, à savoir les recommandations d'amélioration pour le site e-rara.ch, celles-ci ont été consignées dans l'annexe 3.

Bibliographie

ALBERT, Bill, TULLIS, Tom, et TEDESCO, Donna, 2010. *Beyond the usability lab : conducting large-scale online user experiences studies*. San Fransisco : Morgan Kaufmann. ISBN 978-0-12-374892-8

BANGOR, Aaron, KORTUM, Philip T., et MILLER, James T, 2008. An empirical evaluation of the system usability scale. *International journal of human-computer interaction* [en ligne]. 2008. Vol. 24. No. 6. P. 574-594. [Consulté le 16 octobre 2013]. Disponible à l'adresse : <http://content.ebscohost.com/pdf9/pdf/2008/7M5/01Aug08/33402155.pdf?T=P&P=AN&K=33402155&S=R&D=buh&EbscoContent=dGJyMNxb4kSep7Q4zdnyOLCmr0ueprVSsq64SrKWxWXS&ContentCustomer=dGJyMOzprk20prJKuePfgeyx44Dt6fIA>

BANGOR, Aaron, KORTUM, Philip T., et MILLER, James T, 2009. Determining what individual SUS scores mean : adding an adjective rating scale. *Journal of usability studies* [en ligne]. 2009. Vol. 4. No. 3. P. 114-123. [Consulté le 17 décembre 2013]. Disponible à l'adresse : http://66.39.39.113/upa_publications/jus/2009may/JUS_Bangor_May2009.pdf

BARNUM, Carol M., 2011. *Usability testing essentials : ready, set... test!* [en ligne]. Amsterdam : Elsevier, 2011. [Consulté le 17 décembre 2013]. Disponible à l'adresse : <http://hesge.scholarvox.com/reader/index/docid/88805879/page/1> [accès par abonnement]

BEVAN, Nigel, 2003. *Usabilitynet* [en ligne]. 2003. 2006. [Consulté le 18 juin 2013]. Disponible à l'adresse : <http://www.usabilitynet.org/home.htm>

BLUMER, Eliane, 2011. *Benutzerorientierte Evaluation der Webseite e-rara.ch anhand von Usability- und Blickmessungstests* [en ligne]. Genève : Haute école de gestion de Genève. Travail de Bachelor. [Consulté le 18 juin 2013]. Disponible à l'adresse : http://doc.rero.ch/record/28011/files/TDB_2139.pdf

BOLT, Nate, 2007. *Remote research* [en ligne]. 2007. 2013. [Consulté le 18 juin 2013]. Disponible à l'adresse : <http://remotereseaer.ch/>

BOLT, Nate, TULATHIMUTTE, Tony, 2009. *Remote research : real users, real time, real research*. Brooklyn : Rosenfeld Media. ISBN 978-1-933820-77-4

BOREN, Ted et RAMEY, Judith, 2000. Thinking aloud : reconciling theory and practice. *IEEE Transactions on professional communication* [en ligne]. Septembre 2000. Vol. 43. No. 3. P. 261-278. [Consulté le 20 décembre 2013]. Disponible à l'adresse : <http://www.pippinbarr.com/MAHS-F2010/Boren2000.pdf>

BORSCI, Simone, FEDERICI, Stefano, et LAURIOLA, Marco, 2009. On the dimensionality of the System Usability Scale : a test of alternative measurement models. *Cognitive processing*, [en ligne], 2009. Vol. 10. P. 193-197. [Consulté le 17 décembre 2013]. Disponible à l'adresse : <http://link.springer.com/article/10.1007/s10339-009-0268-9#page-1>

BOUTIN, Mario, et MARTIAL, O., 2001. *Évaluation de l'utilisabilité d'un site Web : tests d'utilisabilité versus évaluation heuristique*. Montréal : Centre de recherche informatique de Montréal. Rapport technique, [CRIM-01/01-08]

BROOKE, John, 2011. SUS : a quick and dirty usability scale. In : *ICLE Wiki* [en ligne]. 2011. 10.12.2013. [Consulté le 19 décembre 2013]. Disponible à l'adresse : http://cui.unige.ch/isi/icle-wiki/_media/ipm:test-suschapt.pdf

CHAPMAN, Cameron, 2011. Comprehensive Review Of Usability And User Experience Testing Tools. *Smashing magazine* [en ligne]. 20 octobre 2011. [Consulté le 2 janvier 2014]. Disponible à l'adresse : <http://uxdesign.smashingmagazine.com/2011/10/20/comprehensive-review-usability-user-experience-testing-tools/>

D'APPILI, Florian, 2011. 5 choses à savoir sur les problèmes d'utilisabilité pour ne pas mourir idiot. *Appili – Le Blog de l'Ergonomie Web* [en ligne]. 08 août 2011. [Consulté le 18 décembre 2013]. Disponible à l'adresse : <http://blog.appili.com/2011/08/08/5-choses-savoir-problemes-utilisabilite/>

- DUMAS, Joseph S., 2007. The great leap forward : the birth of the usability profession (1988-1993). *Journal of usability studies* [en ligne], 2007. Vol. 2. No. 2. p. 54-60. [Consulté le 18 décembre 2013]. Disponible à l'adresse : http://www.upassoc.org/upa_publications/jus/2007_february/dumas_birth_of_usability_profession.pdf
- DUMAS, Joseph S. et LORING, Beth A., 2008. *Moderating usability tests : principles and practices for interacting*. Amsterdam : Elsevier. ISBN 978-0-12-373933-9
- DUMAS, Joseph S. et REDISH, Janice C., 1999. *A practical guide to usability testing*. Revised ed. Exeter : Intellect. ISBN 978-1-84150-020-8
- ERICSSON, K. A. et SIMON, H. A., 1993. *Protocol analysis : verbal report as data*. Rev. Ed. Cambridge : MIT Press. ISBN 0-262-55023-7
- GUO, Frank, 2012. More Than Usability : The Four Elements of User Experience, Part I. In : *UXmatters* [en ligne]. 24.04.2012. [Consulté le 17 décembre 2013]. Disponible à l'adresse : <http://www.uxmatters.com/mt/archives/2012/04/more-than-usability-the-four-elements-of-user-experience-part-i.php>
- HERTZUM, M., HANSEN, K. et ANDERSON, H., 2009. Scrutinising usability evaluation: does thinking aloud affect behavior and mental workload ? *Behavioral & Information Technology* [en ligne]. Mars-avril 2009. Vol. 28. No. 2. P. 165-181. [Consulté le 20 décembre]. Disponible à l'adresse : http://www.akira.ruc.dk/~mhaz/Research/Publ/BIT2009_preprint.pdf
- JENG, Judy, 2005. What is usability in the context of the digital library and how can it be measured ? *Information technology and libraries*, 2005. Vol. 24. No.2. P. 47-56. [Consulté le 02 avril 2013]. Disponible à l'adresse : https://inventory.pbworks.com/f/Usability_DL.pdf
- KRAHMER, E., UMMELEN, N., 2004. Thinking about thinking aloud : a comparison of two verbal protocols for usability testing. *IEEE Transactions on professional communication* [en ligne]. Janvier 2004. Vol. 47. No. 2. P. 105-117. [Consulté le 20 décembre 2013]. Disponible à l'adresse : <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.103.34&rep=rep1&type=pdf>
- KRUG, Steve, 2006. *Don't make me think ! A common sense approach to web usability*. 2nd ed. Berkeley : New Riders. ISBN 0-321-34475-8
- LOOP11, 2013. *Loop11 : online usability testing* [en ligne]. 2009-. 2014. [Consulté le 5 janvier 2014]. Disponible à l'adresse : http://www.loop11.com/?utm_expid=13390663-6.iWJa0ABOThaJKcZb3XWfww.0&utm_referrer=http%3A%2F%2Fwww.loop11.com%2Ffeatures%2F
- MORVILLE, Peter, 2004. User Experience Design. In : *Semantics studios* [en ligne.] 21 juin 2004. [Consulté le 17 décembre 2013]. Disponible à l'adresse : <http://semanticstudios.com/publications/semantics/000029.php>
- NIELSEN, Jakob, 1993. *Usability engineering*. Boston : AP Professionnal. ISBN 0-12-518406-9
- NIELSEN, Jakob, 2000. Why you only need to test with five users. *Alertbox* [en ligne]. 19 mars 2000. [Consulté le 13 mai 2013]. Disponible à l'adresse : <http://www.nngroup.com/articles/why-you-only-need-to-test-with-5-users/>
- NIELSEN, Jakob, 2012a. Usability 101 : introduction to usability. *Alertbox* [en ligne]. 04 janvier 2012. [Consulté le 13 mai 2013]. Disponible à l'adresse : <http://www.nngroup.com/articles/usability-101-introduction-to-usability/>
- NIELSEN, JAKOB, 2012b. Thinking Aloud : The #1 Usability Tool. *Nielsen Norman Group* [en ligne]. 16 janvier 2012. [Consulté le 17 décembre 2013]. Disponible à l'adresse : <http://www.nngroup.com/articles/thinking-aloud-the-1-usability-tool/>

- NIELSEN, Jakob, 2012c. How Many Test Users in a Usability Study? *Nielsen Norman Group* [en ligne]. 04 juin 2012. [Consulté le 17 décembre 2013]. Disponible à l'adresse : <http://www.nngroup.com/articles/how-many-test-users/>
- NIELSEN, Jakob, 2012d. Traveling Usability Lab. *Nielsen Norman Group* [en ligne]. 10 septembre 2012. [Consulté le 18 décembre 2013]. Disponible à l'adresse : <http://www.nngroup.com/articles/traveling-usability-lab/>
- NIELSEN, Jakob, LORANGER, Hoa, 2007. *Site web : priorité à la simplicité*. Paris : CampusPress. ISBN 978-2744021527
- OLMSTED-HAWALA, Erica L. et al., 2010. Think-aloud protocols : a comparison of three think-aloud protocols for use in testing data-dissemination web sites for usability. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems CHI '10, 5-10 avril 2010* [en ligne]. New York : ACM, 2010. [Consulté le 20 décembre 2013]. Disponible à l'adresse : <http://dmrussell.net/CHI2010/docs/p2381.pdf>
- ORGANISATION INTERNATIONALE DE NORMALISATION, 1998. *Exigences ergonomiques pour travail de bureau avec terminaux à écrans de visualisations (TEV) : partie 11 : lignes directrices concernant l'utilisabilité*. Paris : AFNOR, 1998. ISO 9241-11.
- PRONGUÉ, Nicolas, 2012. *Évaluation de l'utilisabilité de RODIN au moyen d'un test utilisateur asynchrone* [en ligne]. Genève : Haute école de gestion de Genève. Travail de Bachelor. [Consulté le 16 avril 2013] Disponible à l'adresse : <http://doc.rero.ch/record/30353/files/TDB2184.pdf>
- ROBILLARD-BASTIEN, Alain, 2001. Ergonomie du web : test d'utilisabilité. *Ergoweb.ca* [en ligne]. 2001. Octobre 2013. [Consulté le 19 décembre 2013]. Disponible à l'adresse : <http://www.ergoweb.ca/utilisabilite.html>
- RUBIN, Jeffrey, et CHRISNELL, Dana, 2008. *Handbook of usability testing : how to plan, design and conduct effective tests*. 2nd ed. Indianapolis : Wiley Publ. ISBN 978-0-470-18548-3
- SAURO, Jeff, 2011. Measuring Usability with the System Usability Scale (SUS) : Measuring Usability. In : *Measuring usability : quantitative usability, statistics & six sigma* [en ligne]. Mis en ligne le 02 février 2011. [Consulté le 17 décembre 2013]. Disponible à l'adresse : <https://www.measuringusability.com/sus.php>
- SCHMIDT, Eveline, 2013. *Remote oder In-person Usability-test? Methodenvergleich am Beispiel des Metakatalogs Swisslib*. Genève : Haute école de gestion de Genève. Travail de Bachelor. [Consulté le 19 décembre 2013]. Disponible à l'adresse : <http://doc.rero.ch/record/208871?ln=fr>
- SPOOL, Jared, SCHROEDER, Will, 2001. Testing Web sites: five users is nowhere near enough. In : *Proceedings of the Conference extended abstracts on Human Factors in Computing Systems, CHI'2001*. New York : ACM Press, 2001. P. 285-286.
- TOMLIN, Craig, W., 2009. My big list of 24 web site usability testing tools. *Useful Usability* [en ligne]. 2 décembre 2009, 12 :21. [Consulté le 2 janvier 2014]. Disponible à l'adresse : <http://www.usefulusability.com/24-usability-testing-tools/>
- TSAKONAS, Giannis, et PAPATHEODOROU, Christos, 2008. Exploring usefulness and usability in the evaluation of open access digital libraries. *Information processing and management*. 2008. Vol. 44. No 3, p. 1234-1250
- TULLIS, Tom et al., 2002. An empirical comparison of lab and Remote usability testing of web sites [en ligne]. In : *Usability Professionals Association Conference*. Boston : Fidelity Investments. [Consulté le 16 décembre 2013]. Disponible à l'adresse : <http://home.comcast.net/~tomtullis/publications/RemoteVsLab.pdf>
- TULLIS, Thomas S. et STETSON, Jacqueline N. 2004. A comparison of questionnaires for assessing website usability. In : *Usability Professional Association Conference* [en ligne].

2004. [Consulté le 17 décembre 2013]. Disponible à l'adresse : <http://home.comcast.net/~tomtullis/publications/UPA2004TullisStetson.pdf>

USABILIS, sans date. Test d'utilisabilité - Test utilisateur. *Usabilis.com* [en ligne]. Sans date. [Consulté 19 décembre 2013]. Disponible à l'adresse : <http://www.usabilis.com/methode/test-utilisateur.htm>

VAN DEN HAAK, Maaïke J., DE JONG, Menno, D. T. et SCHELLENS, Peter Jan, 2003. Retrospective vs. concurrent think-aloud protocols : testing the usability of an online library catalogue. *Behavior & Information technology* [en ligne]. Septembre-octobre 2003. Vol. 22. No 5. p. 339-351. [Consulté le 20 décembre 2013]. Disponible à l'adresse : http://www.ecs.csun.edu/~rlingard/COMP595VAV/thinking_aloud_testing.pdf

VAN DEN HAAK, Maaïke J., DE JONG, Menno, D. T. et SCHELLENS, Peter Jan, 2009. Evaluating municipal websites: a methodological comparison of three think-aloud variants. *Government Information Quarterly* [en ligne]. Janvier 2009. Vol. 26. No 1. P. 193-202. [Consulté le 20 décembre 2013]. Disponible à l'adresse : http://repository.ubn.ru.nl/bitstream/2066/78902/1/M_287664.pdf

WHATUSERDO, 2013. *Whatuserdo : online usability testing and user experience research* [en ligne]. 2009-. 2013. [Consulté le 9 janvier 2014]. Disponible à l'adresse : <http://www.whatusersdo.com/>

Annexe 1 : Analyse du test automatisé à distance

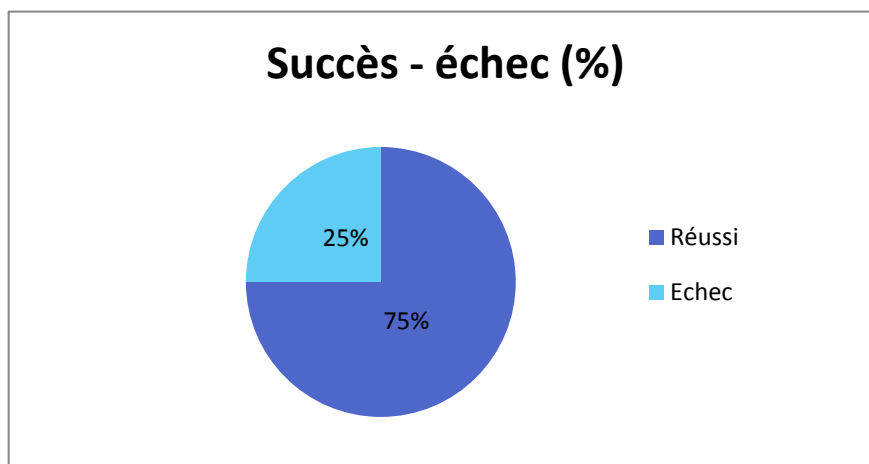
Nous présentons ici les taux de succès, pourcentages de premiers clics, stratégies suivies et erreurs commises pour les différentes tâches des tests en français et en allemand. Seuls les participants ayant achevé les tâches en question ont été pris en compte.

Tâche 1 : participation de la bibliothèque des Pasteurs (Neuchâtel) :

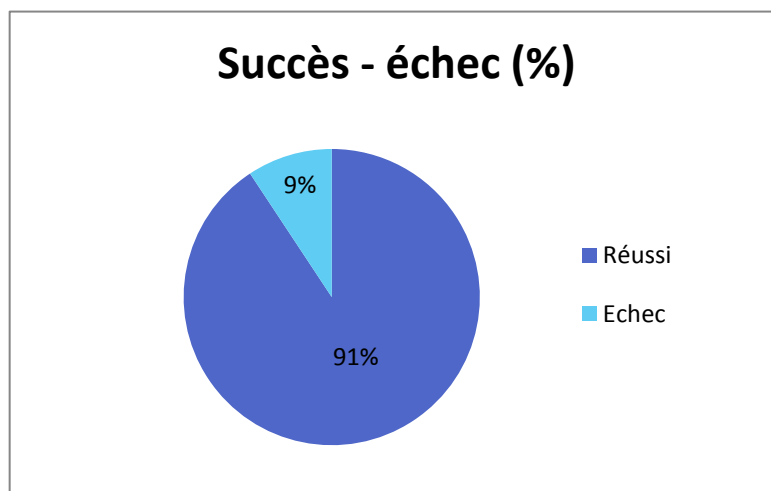
38 participants francophones et 43 germanophones ont réalisé cette tâche.

Plusieurs chemins étaient possibles pour parvenir à la réponse, principalement via l'index des « collections thématiques » (la collection Bibleikon est associée à la bibliothèque des Pasteurs), la recherche avancée (la Bibliothèque des pasteurs est dans la liste des bibliothèques), la recherche avec les termes « Bibliothèque des pasteurs », ou pasteur (et utiliser facette « bibliothèque » pour filtrer les résultats) ou tout simplement en utilisant la facette « Bibliothèques ».

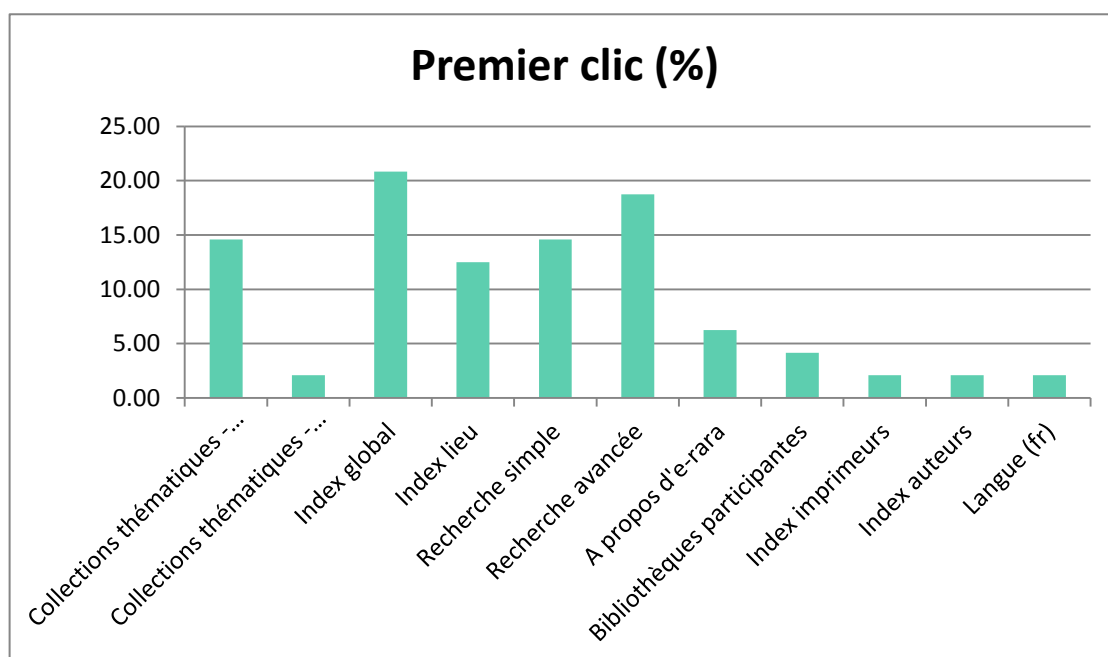
Pourcentage de réussite et d'échec à la tâche 1 du test en français



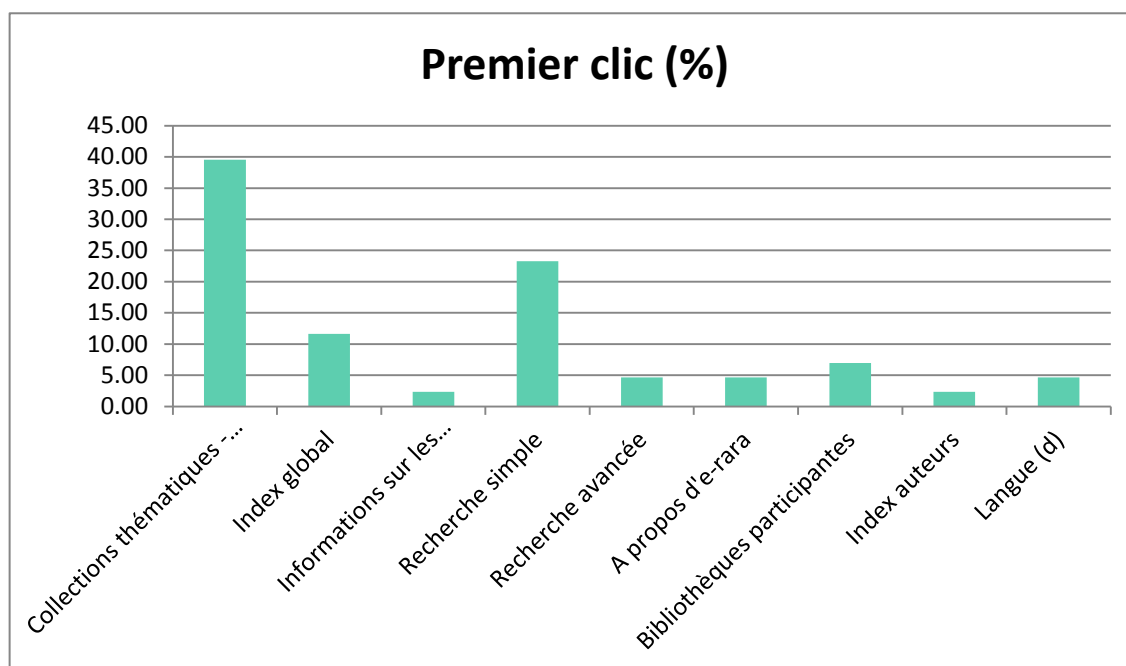
Pourcentage de réussite et d'échec à la tâche 1 du test en allemand



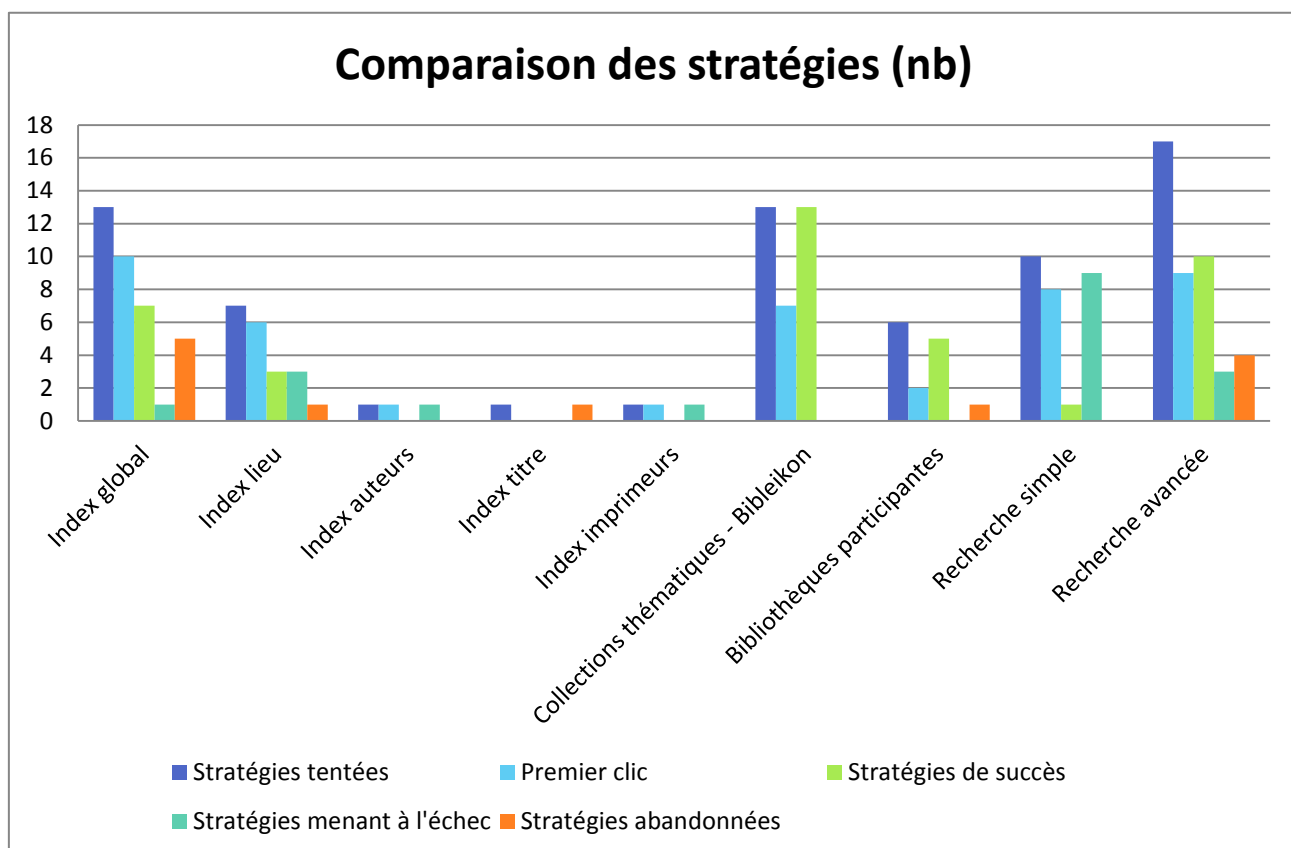
Pourcentage des premières pages cliquées pour le test en français (tâche 1)



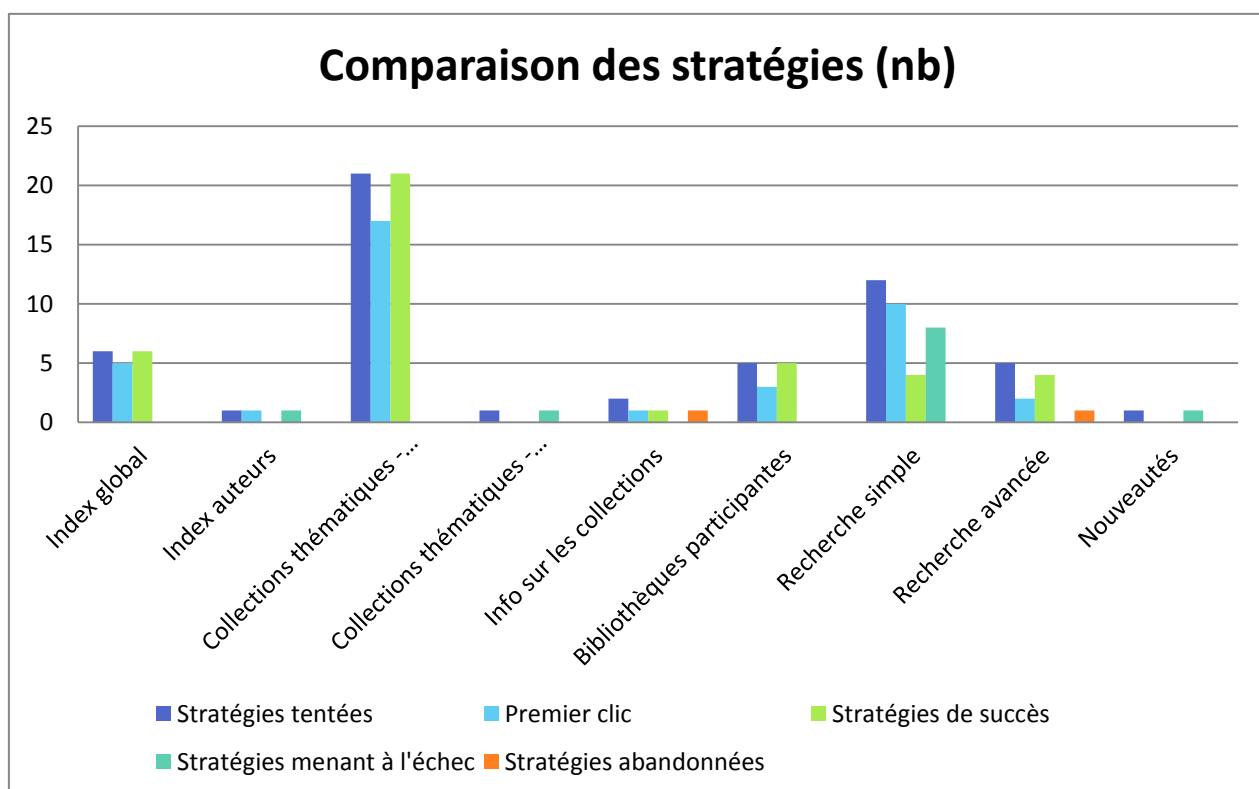
Pourcentage des premières pages cliquées pour le test en allemand (tâche 1)



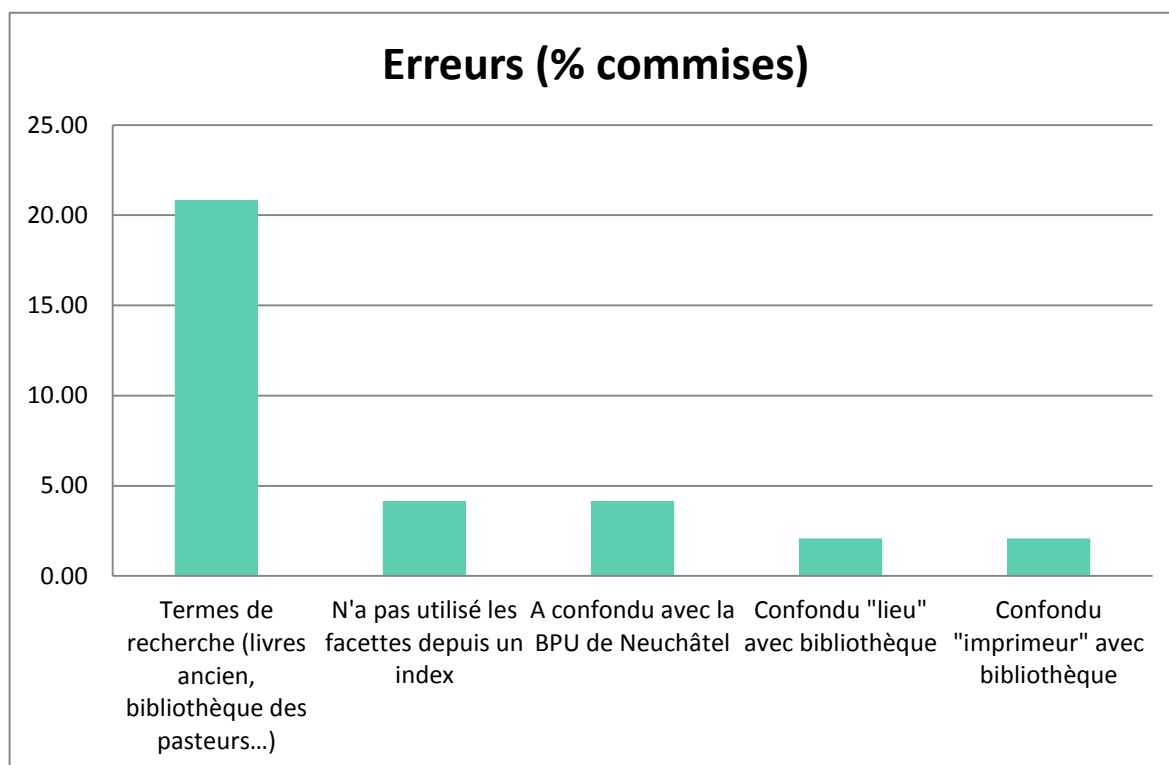
Comparaison des stratégies employées par les participants du test en français (tâche 1)



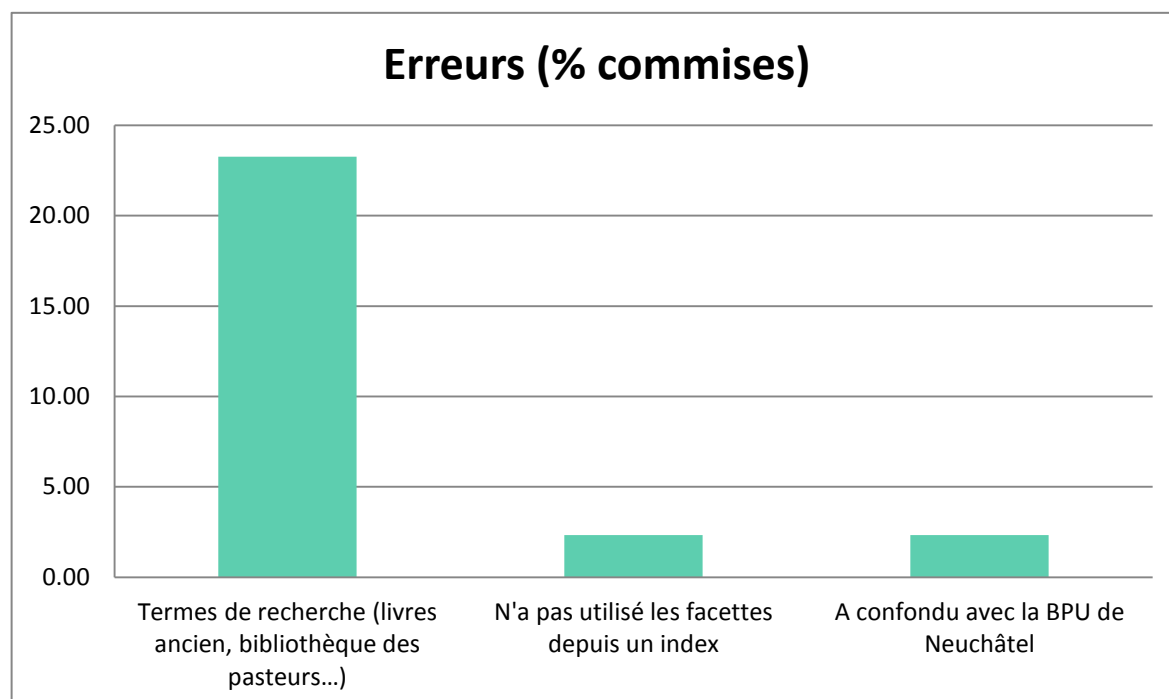
Comparaison des stratégies employées par les participants du test en allemand (tâche 1)



Types d'erreurs commises par les participants du test en français (tâche 1)



Types d'erreurs commises par les participants du test en allemand (tâche 1)



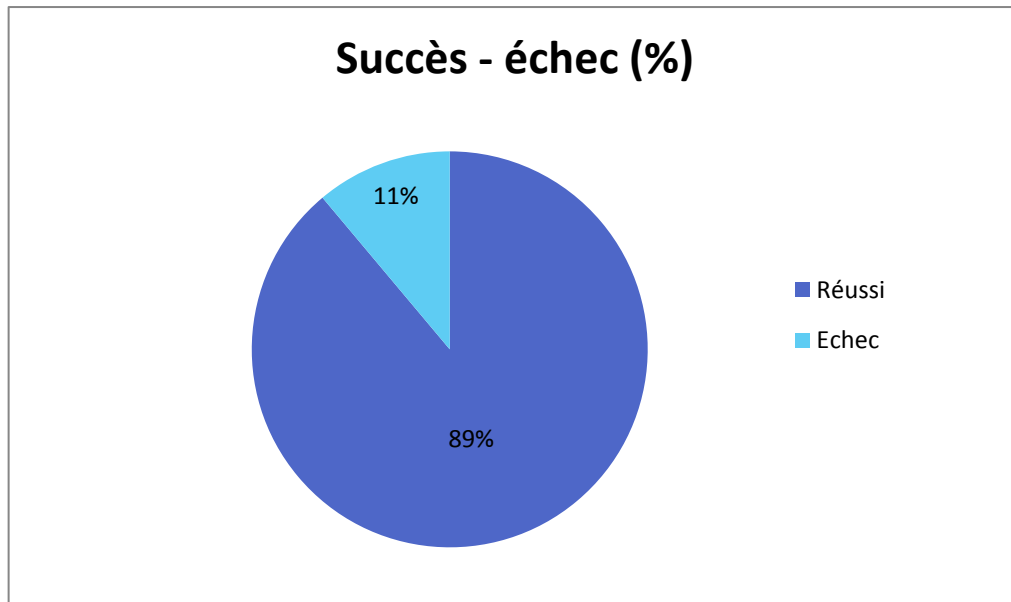
Vu les nombreux chemins possibles pour accéder à l'information demandée, cette tâche n'a finalement révélé que peu d'erreurs. Elle nous a toutefois permis de suspecter un problème :

- Visibilité des bibliothèques participantes insuffisante (problème 24)

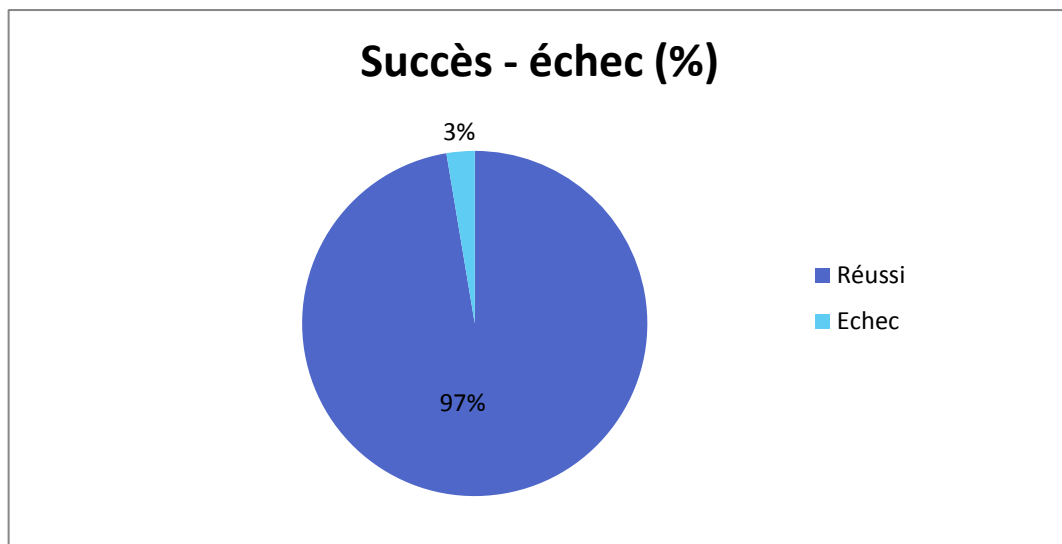
Tâche 2 : atteindre la page des flux rss

80 participants (42 francophones et 38 germanophones) ont réalisé cette tâche. Elle comporte un grand taux de réussite, et ce dans les deux langues. La façon dont Loop11 collecte les données ne nous permet pas de chiffrer l'utilisation respective de l'icône et du lien dans le texte de la page d'accueil, mais la heat map nous montre que les deux ont été cliqués.

Pourcentage de réussite et d'échec à la tâche 2 du test en français



Pourcentage de réussite et d'échec à la tâche 2 du test en allemand

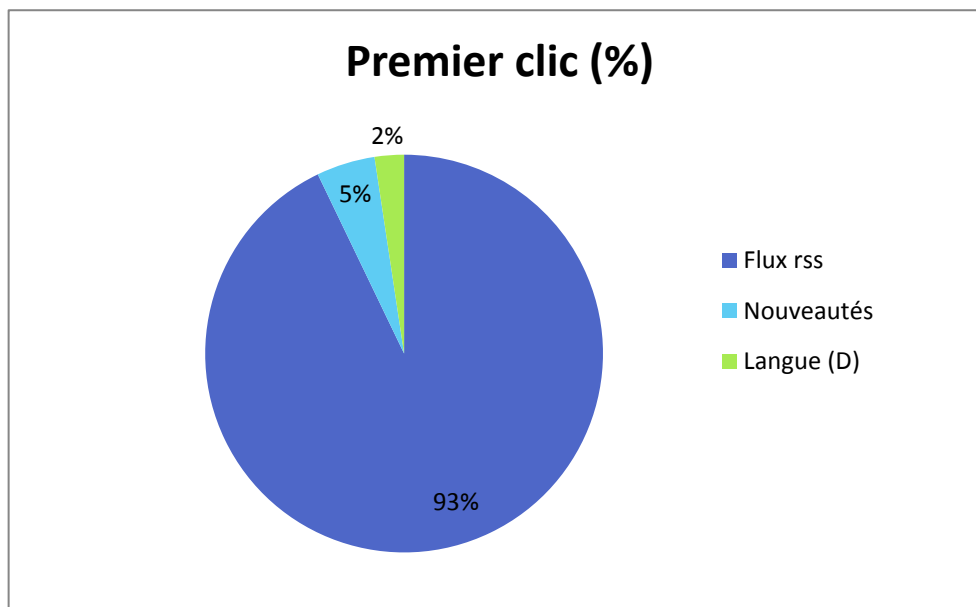


Cela représente au final 2 participants francophones et 1 germanophone. A noter que 2 des participants indiquant qu'ils ne savaient pas ce qu'étaient des flux rss ont tout de même réalisé et réussi la tâche. Au final, seul 1 participant savant ce qu'était un flux rss n'a pas

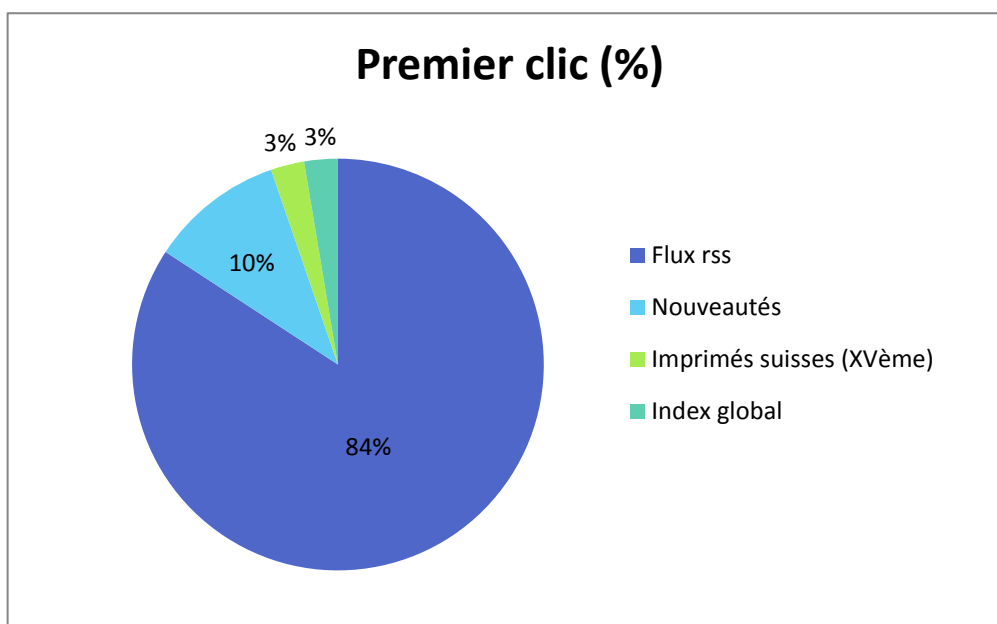
trouvé la réponse. Les modifications suivant le travail d'Eliane Blumer (2011) ont donc porté leurs fruits.

Cette tâche étant relativement courte et réalisable en un clic, les erreurs commises se réduisent aux erreurs de premier clic, à l'exception d'un participant du test allemand, qui a également cliqué sur les nouveautés et l'index global. Nous n'avons donc pas fait d'autres analyses pour cette tâche.

Pourcentage des premières pages cliquées pour le test en français (tâche 2)



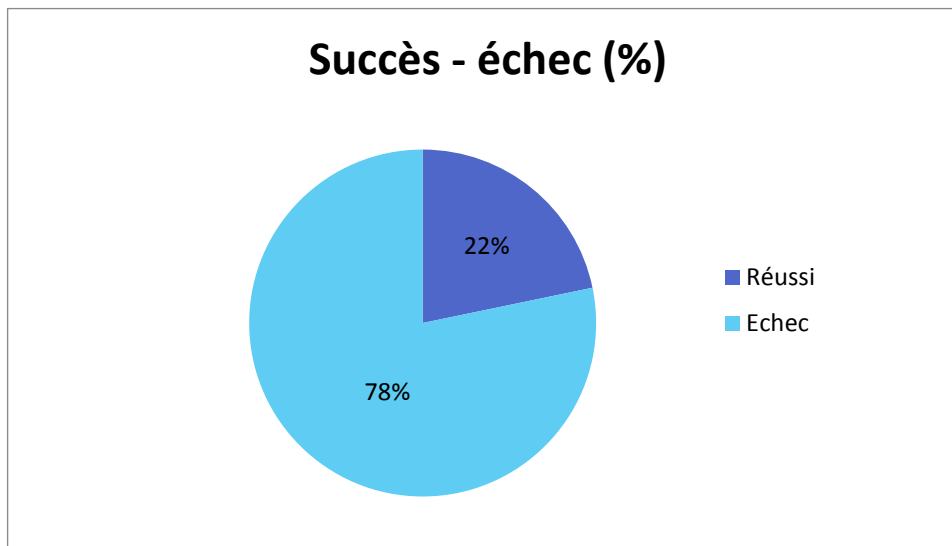
Pourcentage des premières pages cliquées pour le test en allemand (tâche 2)



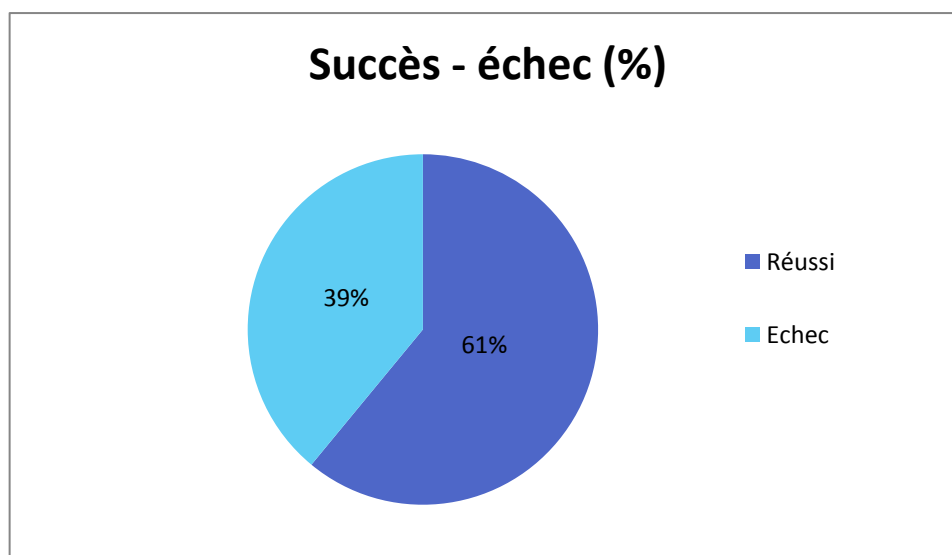
Tâche 3 : trouver le plus ancien ouvrage d'Erasme disponible sur le site

87 participants (46 en français et 41 en allemand) ont rempli cette tâche.

Pourcentage de réussite et d'échec à la tâche 3 du test en français

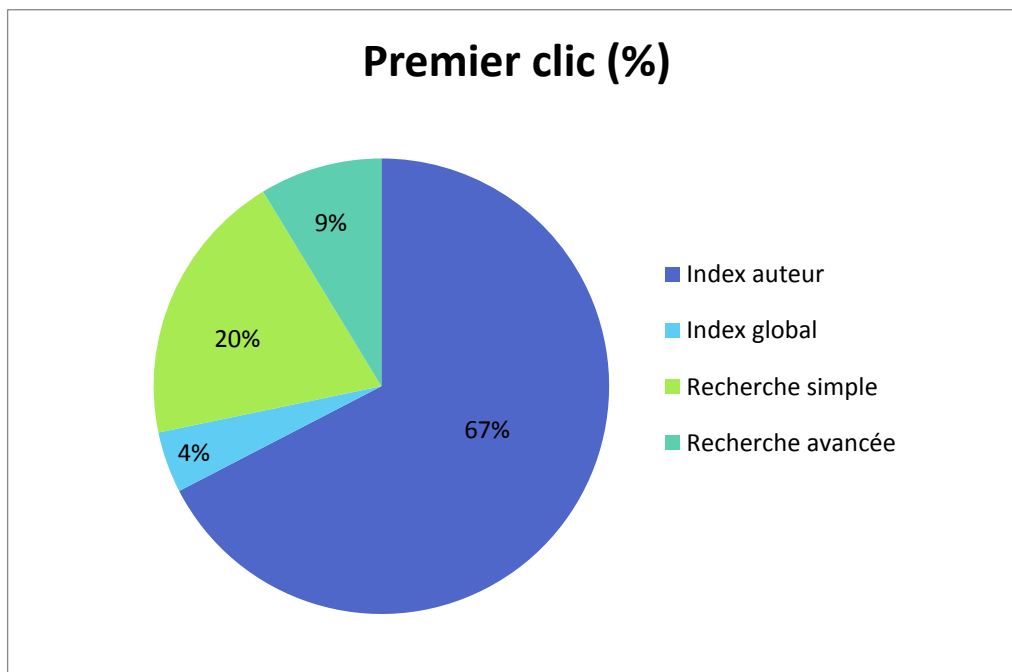


Pourcentage de réussite et d'échec à la tâche 3 du test en allemand

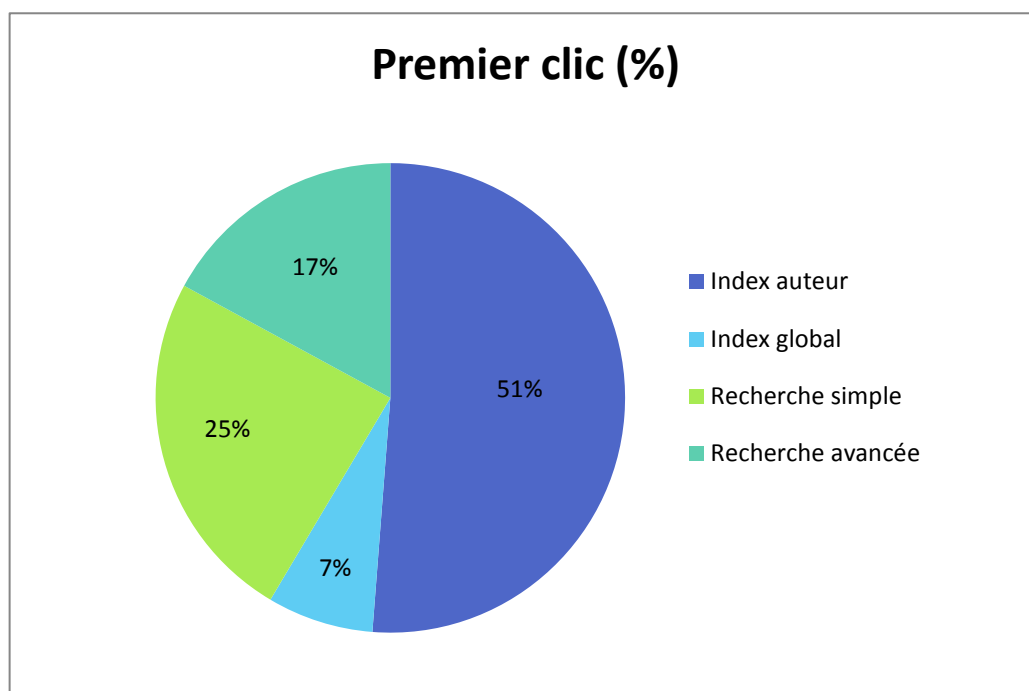


La différence de difficulté en fonction de la langue ressort nettement dans le pourcentage de succès et d'échec.

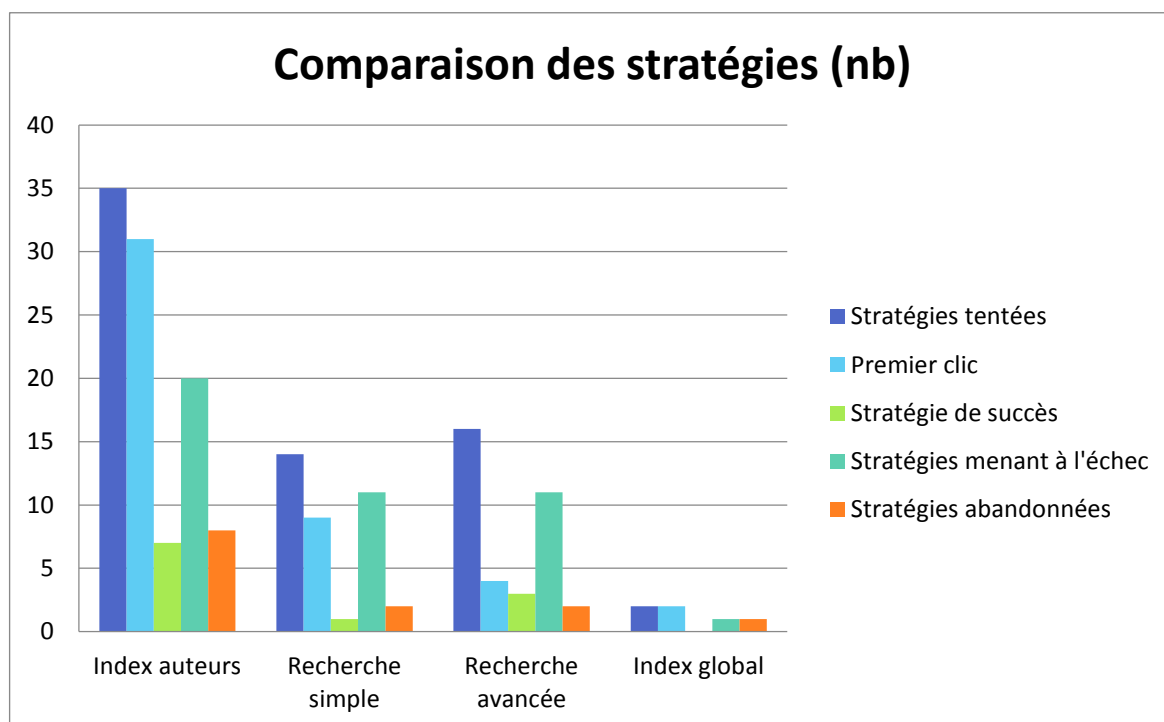
Pourcentage des premières pages cliquées pour le test en français (tâche 3)



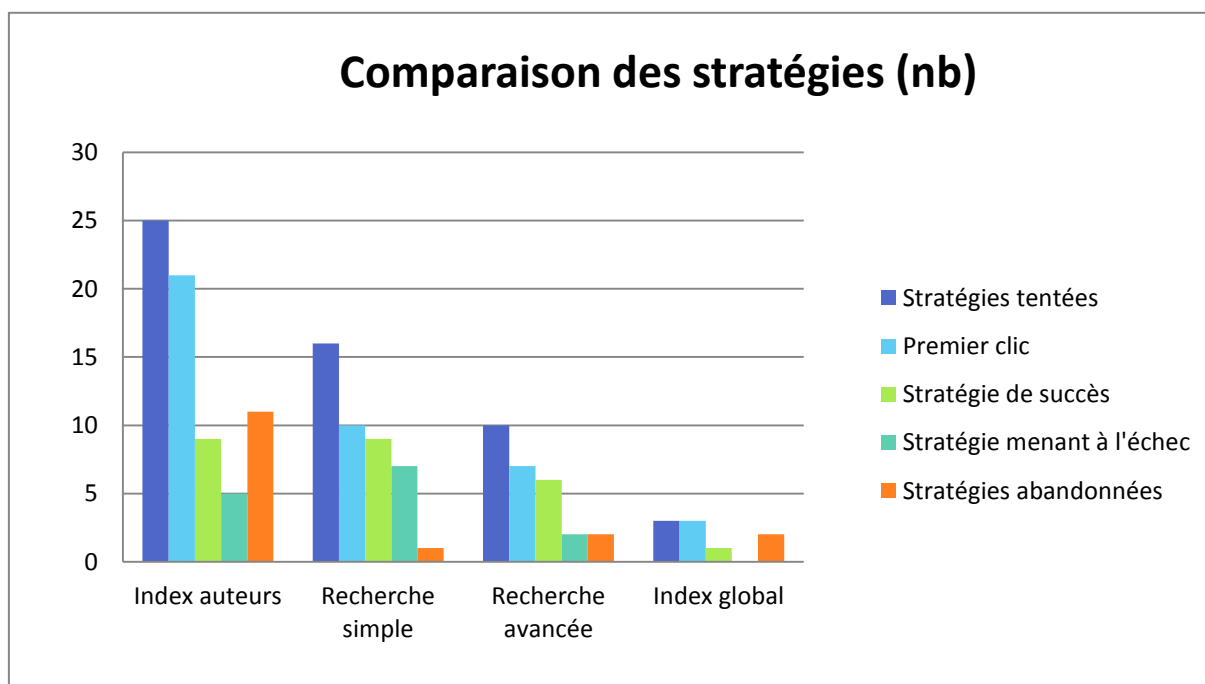
Pourcentage des premières pages cliquées pour le test en allemand (tâche 3)



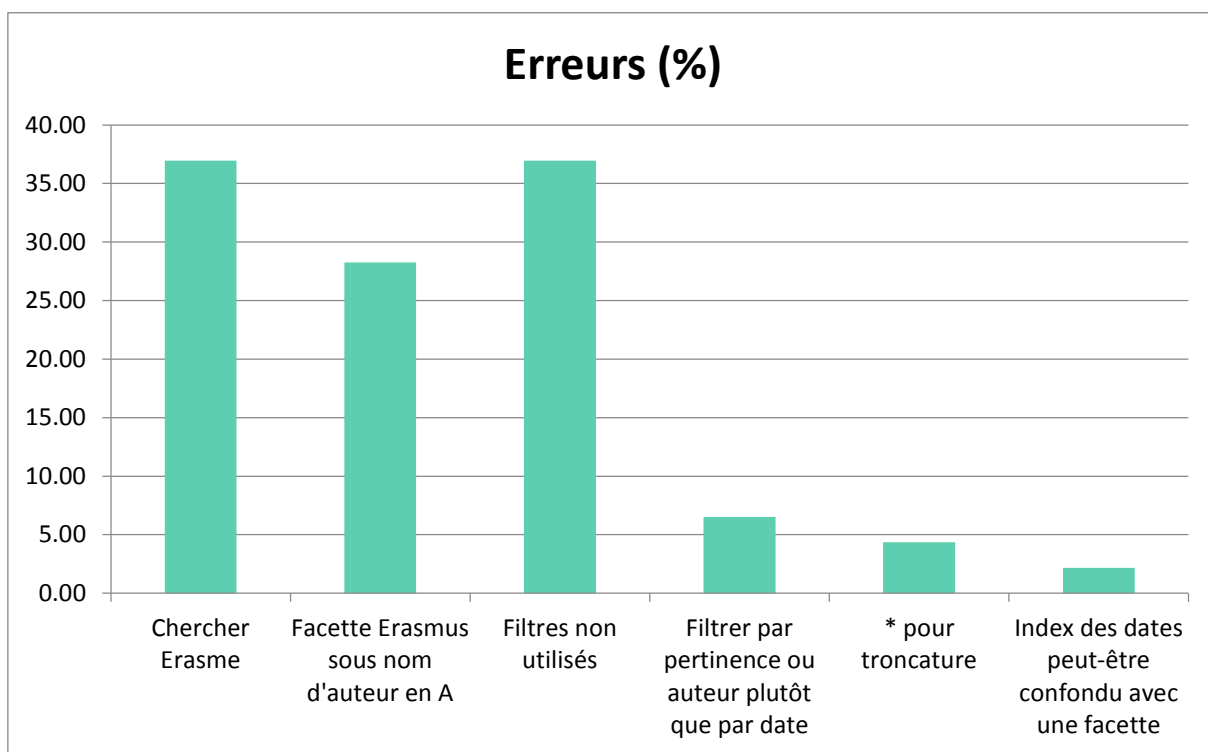
Comparaison des stratégies employées par les participants du test en français (tâche 3)



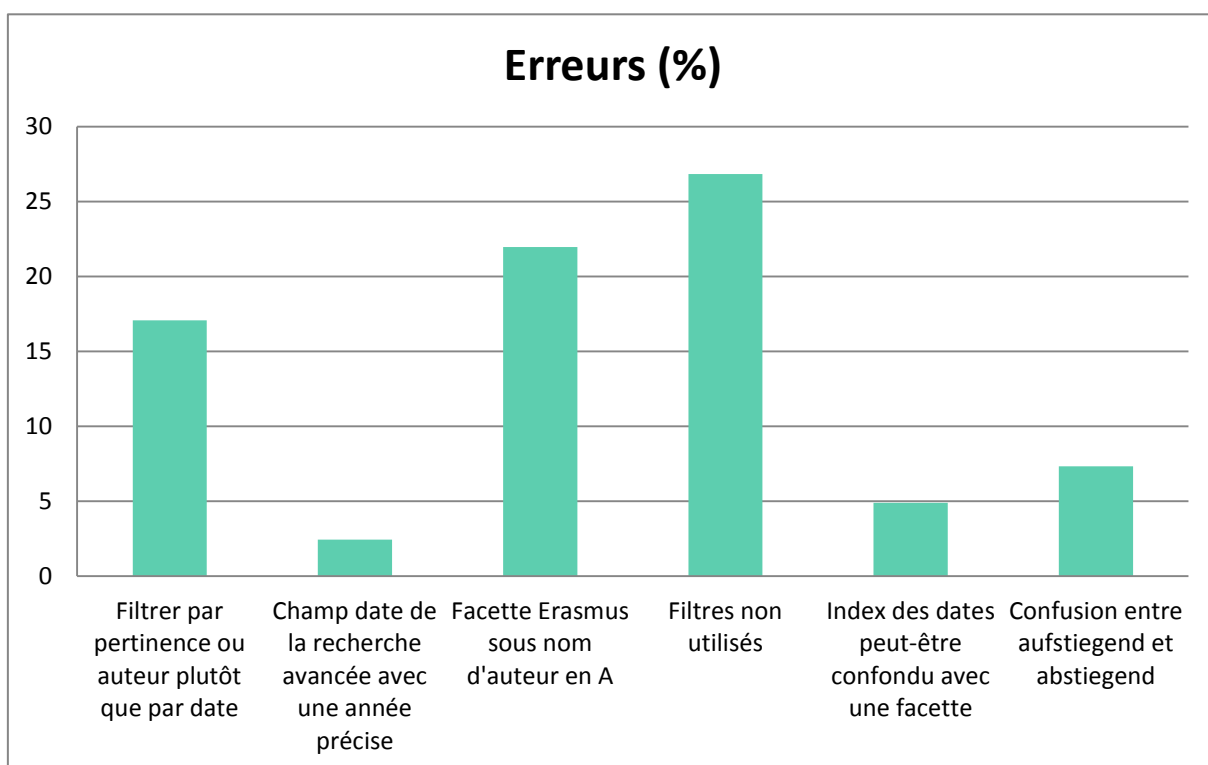
Comparaison des stratégies employées par les participants du test en allemand (tâche 3)



Types d'erreurs commises par les participants du test en français (tâche 3)



Types d'erreurs commises par les participants du test en allemand (tâche 3)



Problèmes repérés :

- Absence de renvois entre les différentes dénominations des auteurs (problème 1)
- Fonctionnalité « parcourir » peu claire avec l'affichage des notices de documents (problème 2)
- Troncature "*" non acceptée (problème 23)

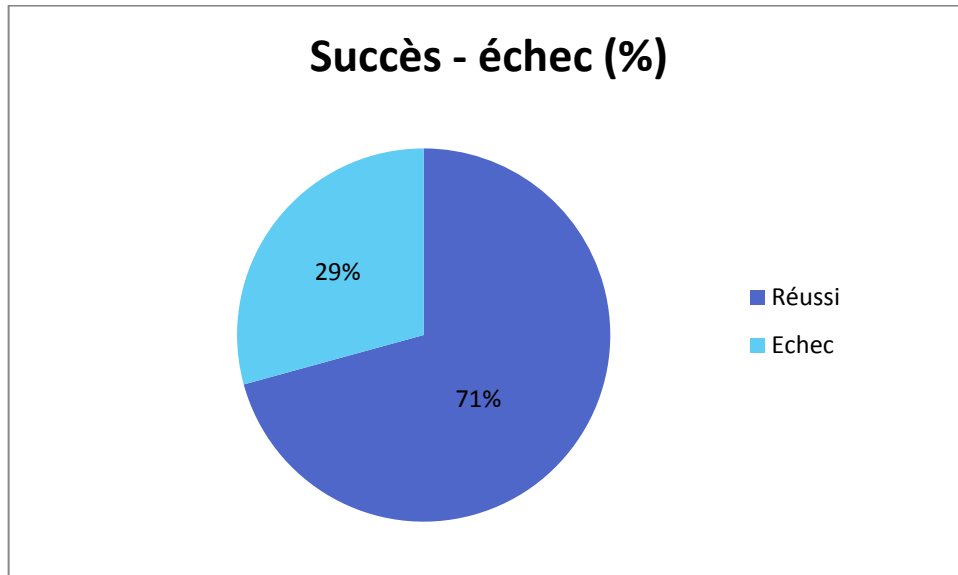
Problèmes suspectés :

- Fenêtre de recherche peu visible (problème 9)
- Affichage des résultats trop peu clair (pas d'étiquettes comme titre, auteur, etc.) (problème 6)
- Interaction entre les différents menus des outils de tri peu claire (problème 10)

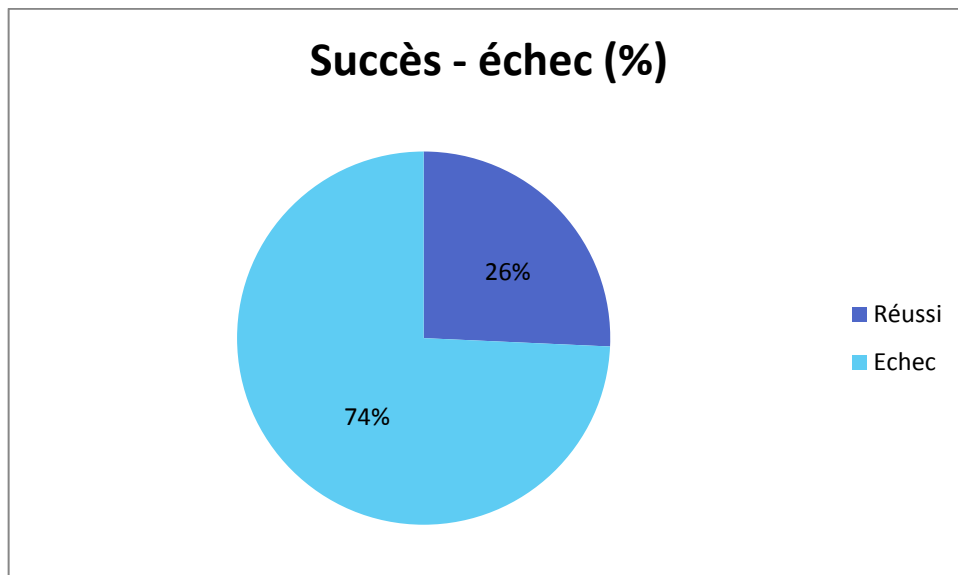
Tâche 4 : trouver quel type d'oiseau était un griffon

41 participants francophones et 35 germanophones ont réalisé la tâche.

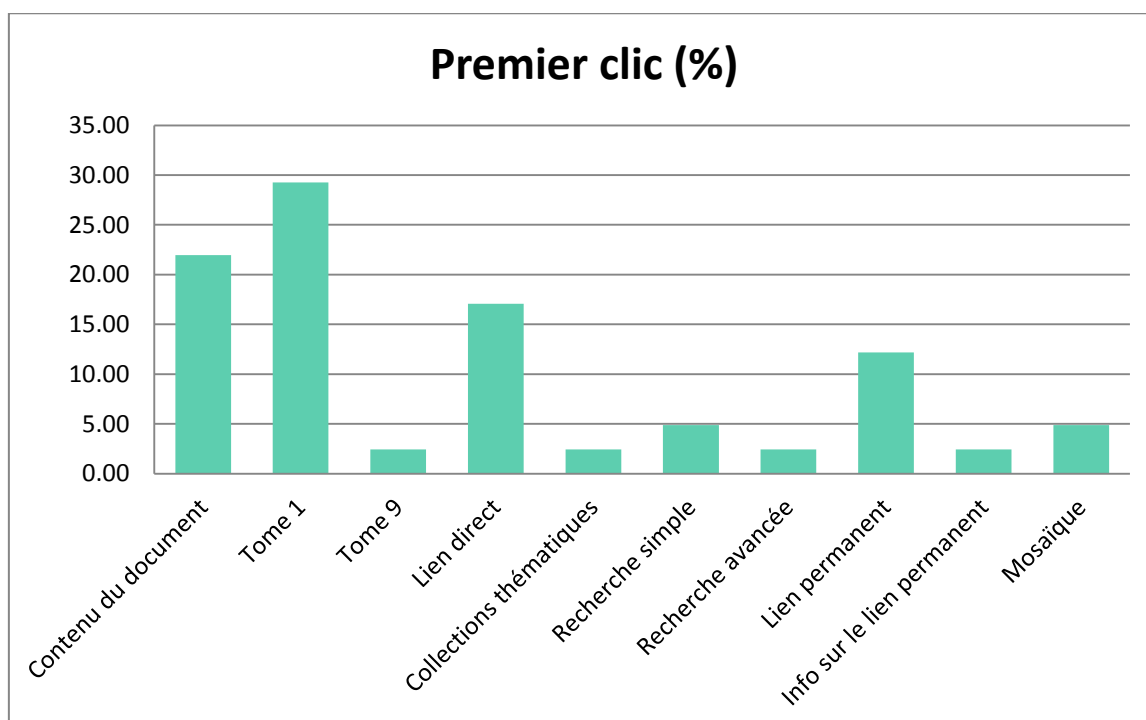
Pourcentage de réussite et d'échec à la tâche 4 du test en français



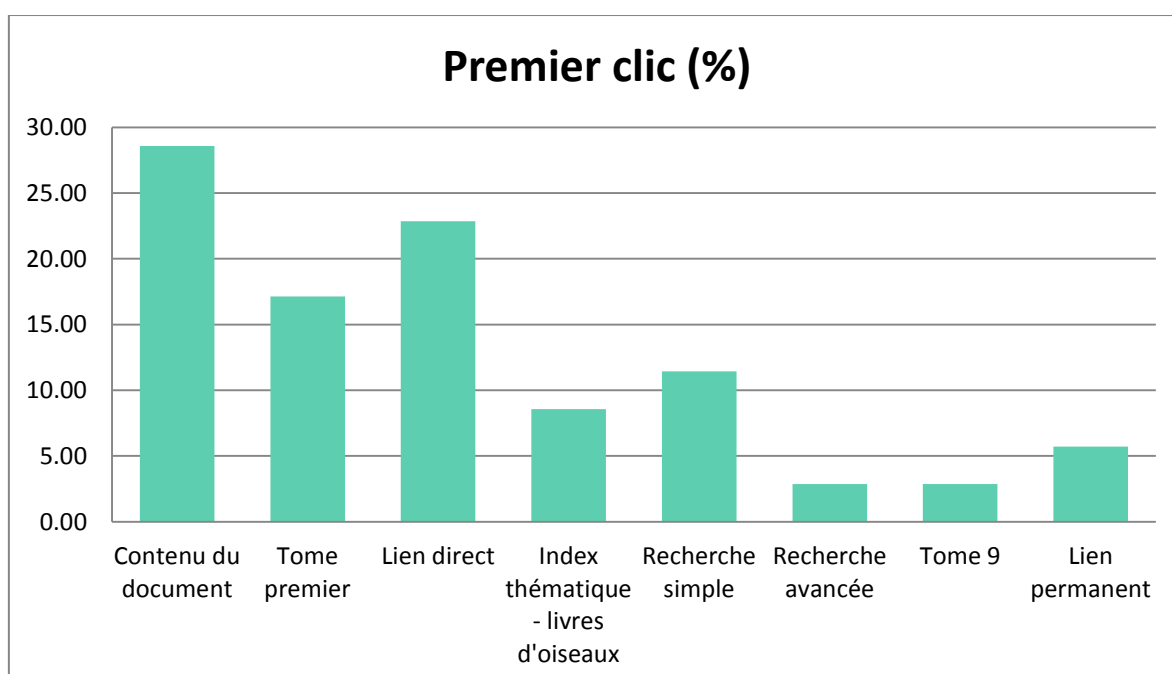
Pourcentage de réussite et d'échec à la tâche 4 du test en allemand



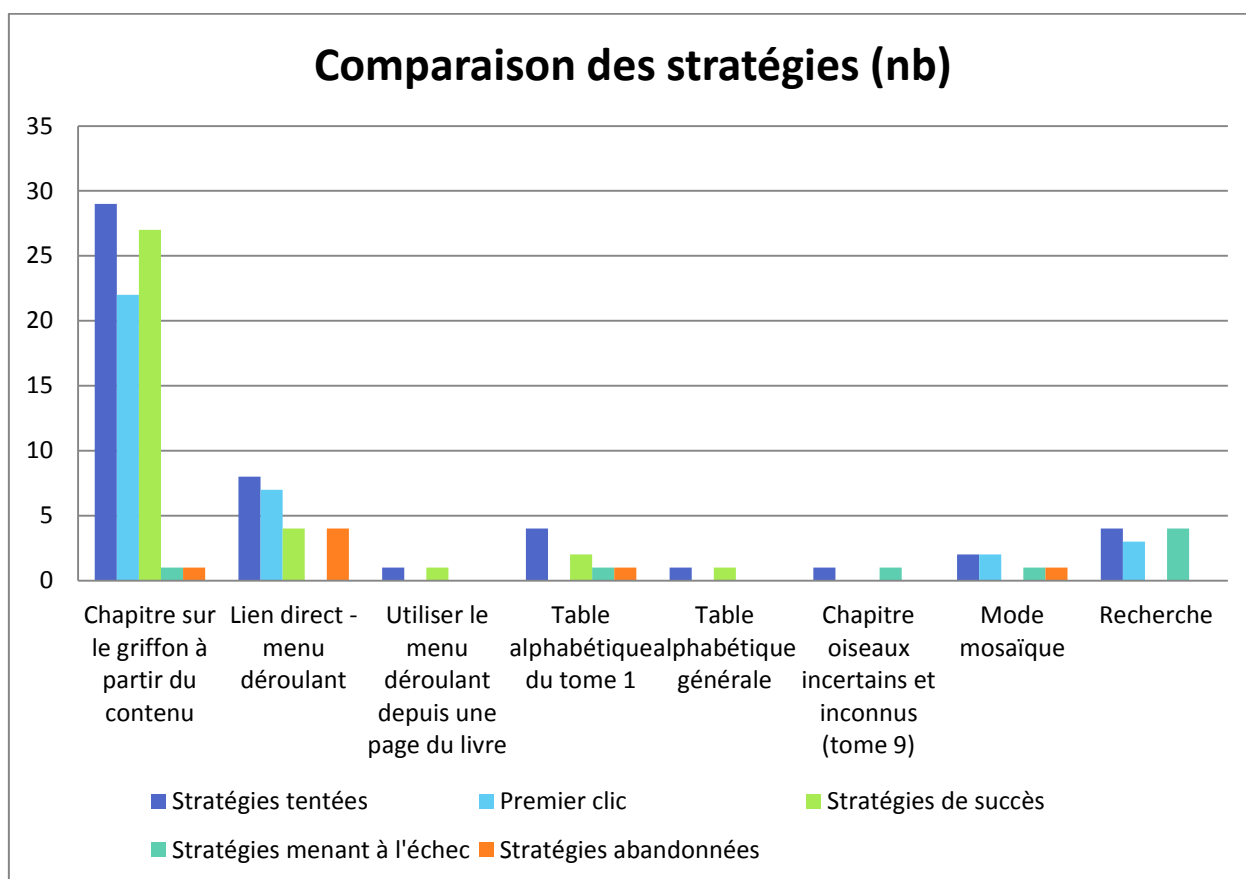
Pourcentage des premières pages cliquées pour le test en français (tâche 4)



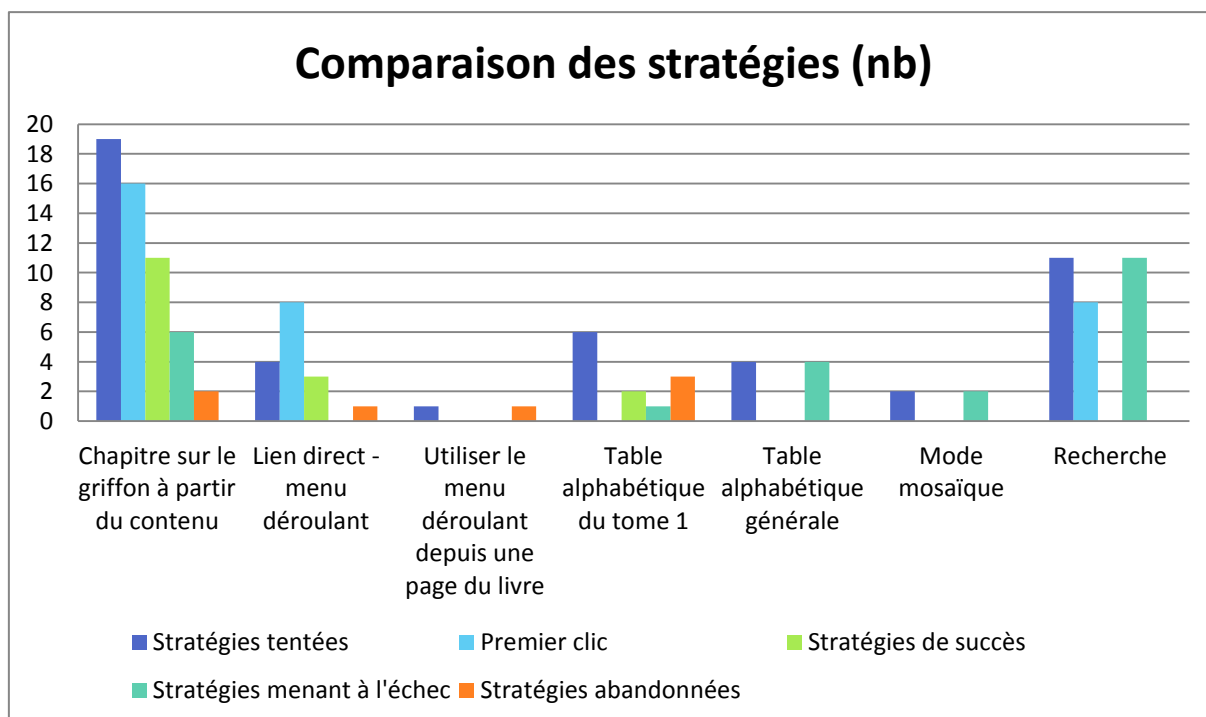
Pourcentage des premières pages cliquées pour le test en allemand (tâche 4)



Comparaison des stratégies employées par les participants du test en français (tâche 4)

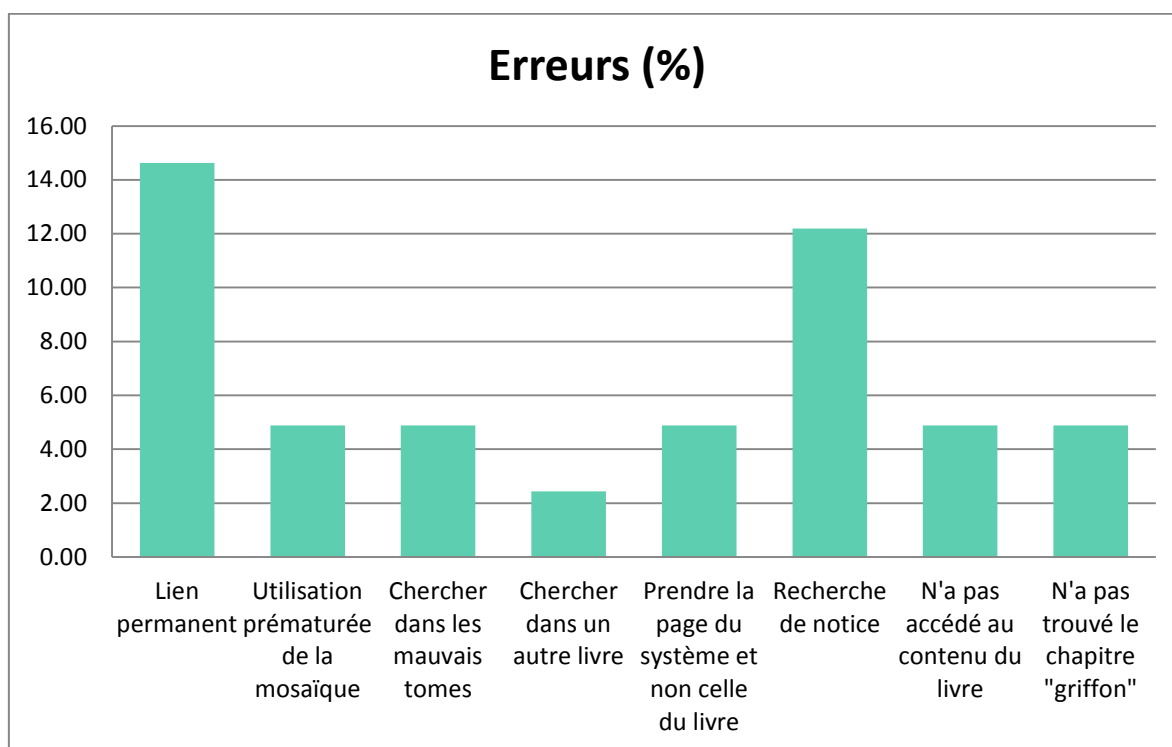


Comparaison des stratégies employées par les participants du test en allemand (tâche 4)

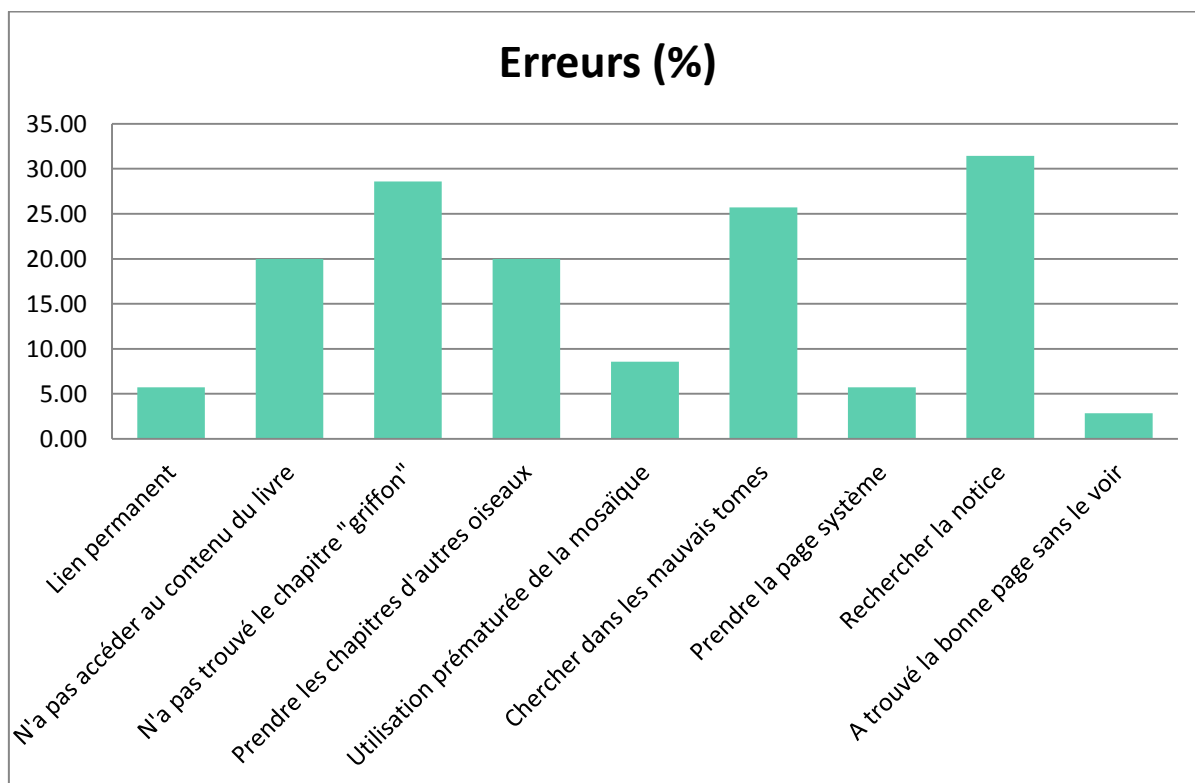


Par « chapitre sur le griffon à partir du contenu », nous entendons que les participants ont accédé au chapitre depuis le sommaire du lien « contenu » ou « tome premier ».

Types d'erreurs commises par les participants du test en français (tâche 4)



Types d'erreurs commises par les participants du test en allemand (tâche 4)



Problèmes repérés :

- Accès difficile à l'information à l'intérieur des documents (problème 5)
- Notion de lien permanent peu claire (problème 17)

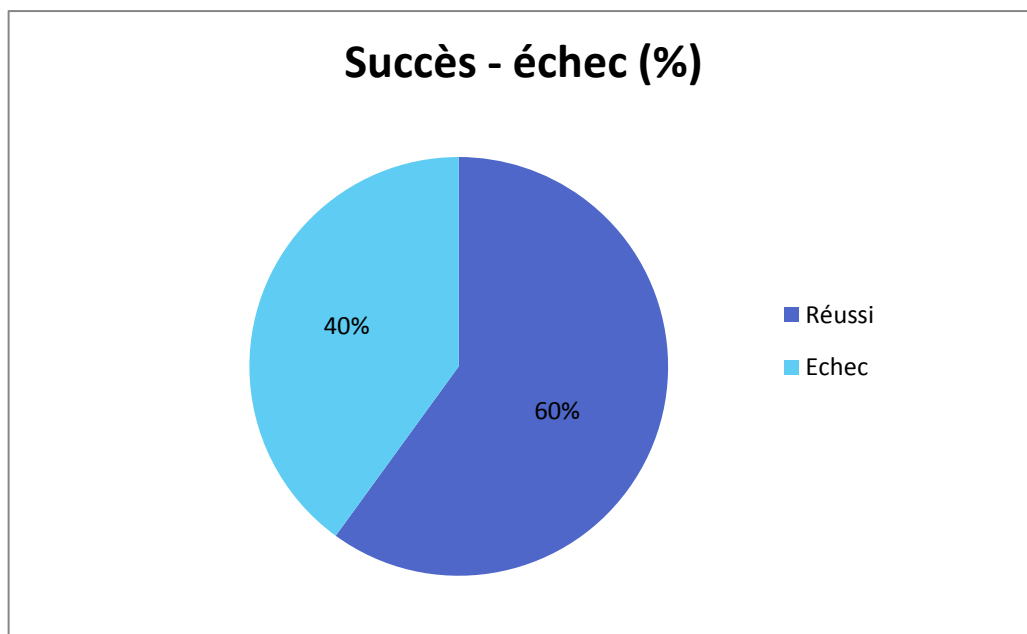
Problèmes suspectés :

- Manque de visibilité du caractère cliquable de l'image du livre des notices (problème 12)

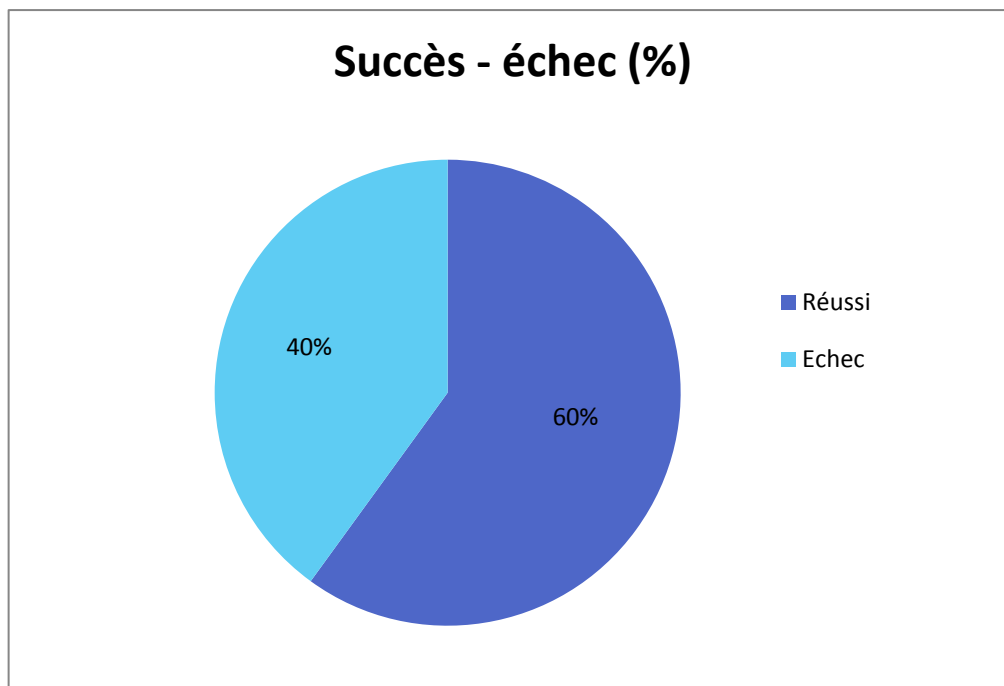
Tâche 5 : trouver la page de l'illustration du faucon sort

Cette tâche a été effectuée par 40 participants francophones et 35 participants germanophones.

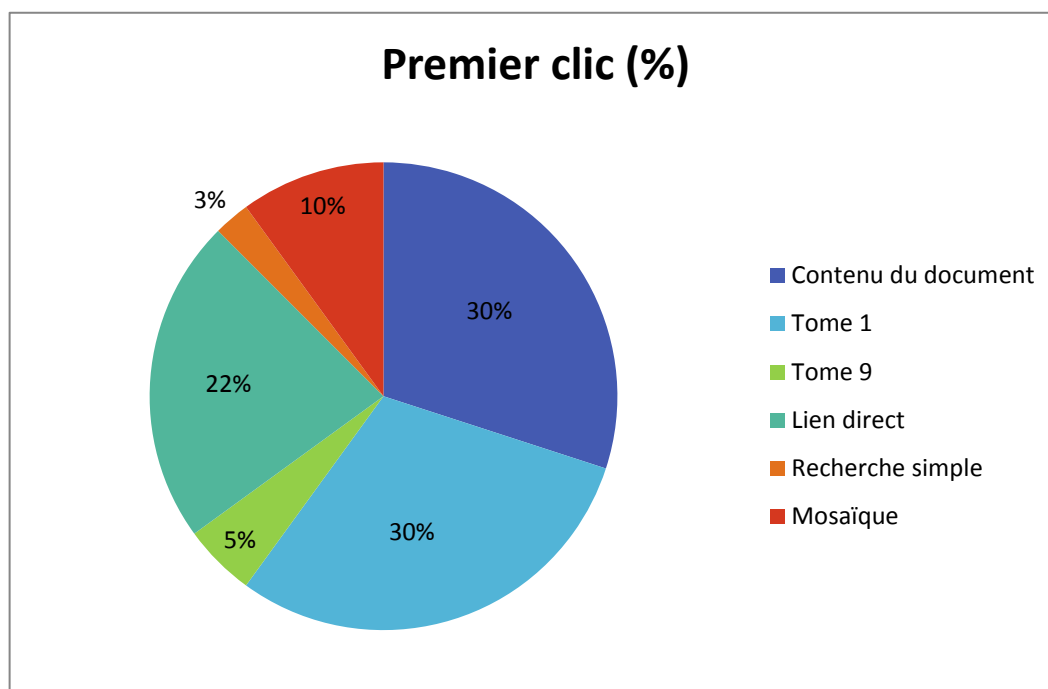
Pourcentage de réussite et d'échec à la tâche 5 du test en français



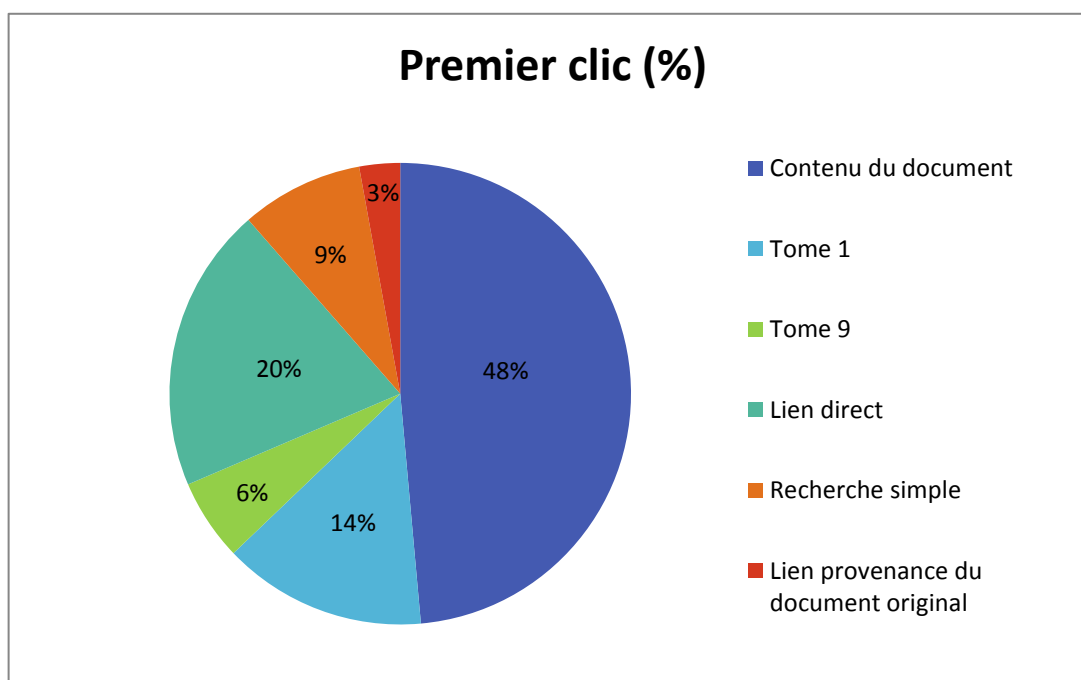
Pourcentage de réussite et d'échec à la tâche 5 du test en allemand



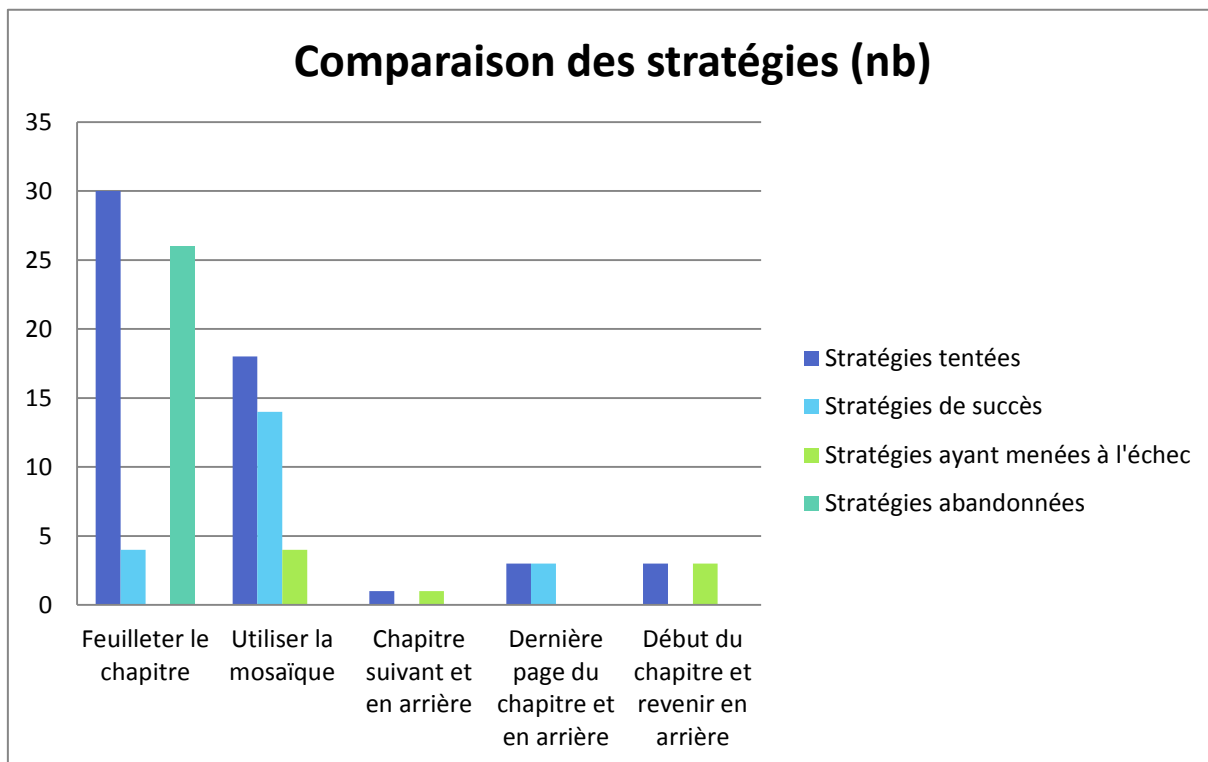
Pourcentage des premières pages cliquées pour le test en français (tâche 5)



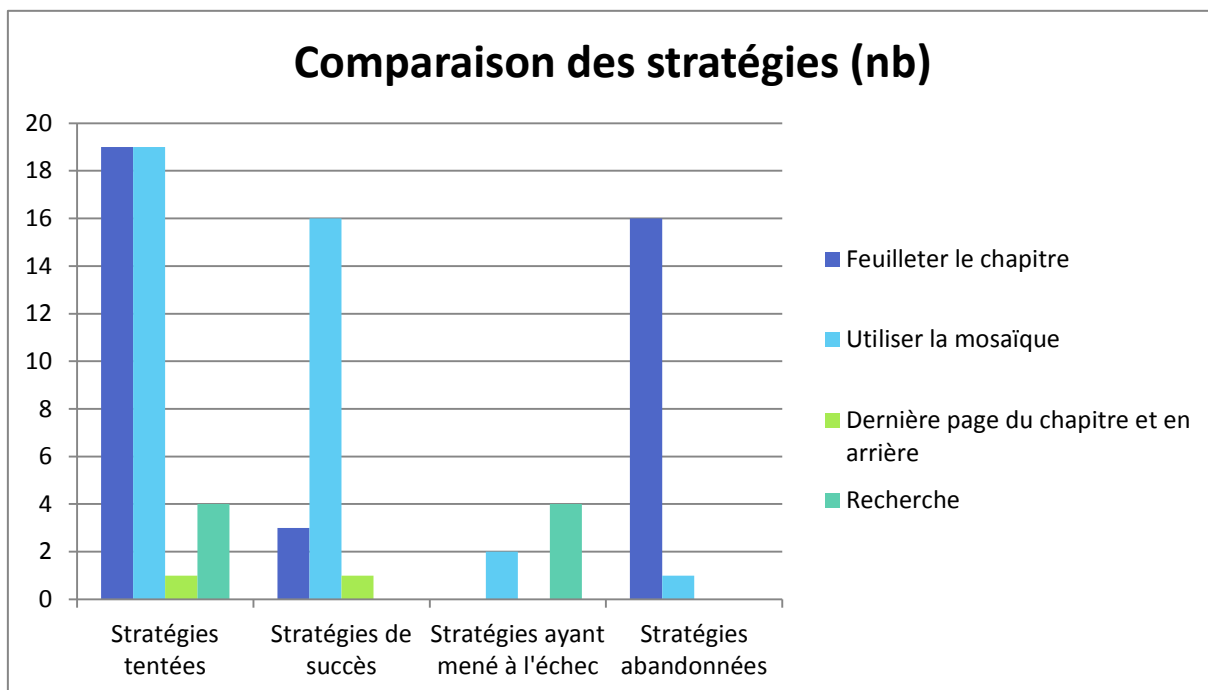
Pourcentage des premières pages cliquées pour le test en allemand (tâche 5)



Comparaison des stratégies employées par les participants du test en français (tâche 5)

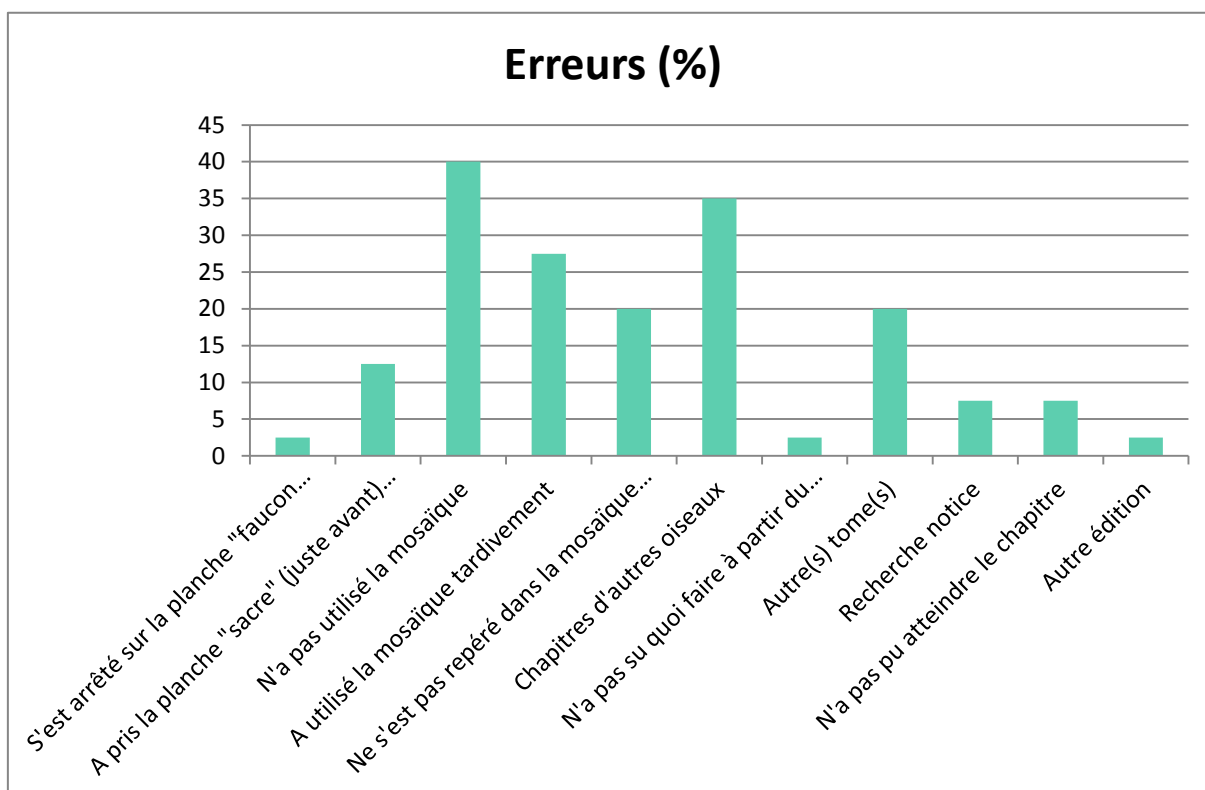


Comparaison des stratégies employées par les participants du test en allemand (tâche 5)

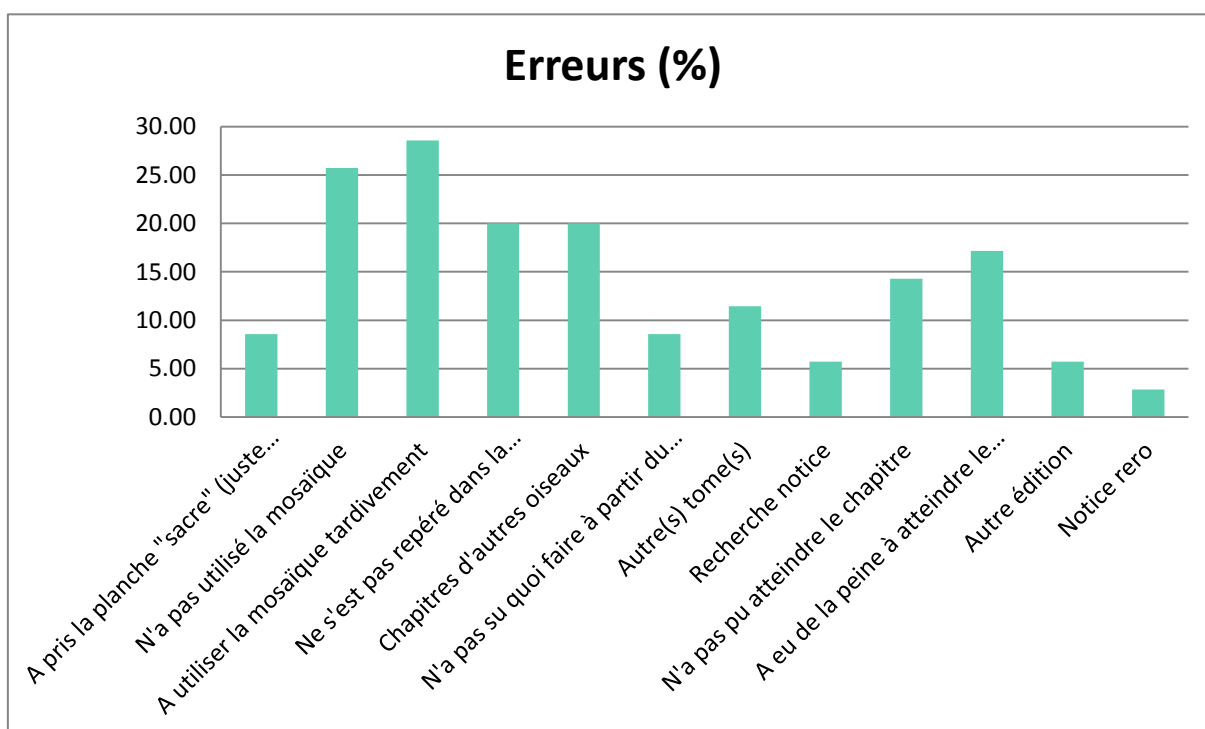


Pour cette tâche, nous nous sommes concentrés sur les stratégies mises en œuvre pour atteindre l'illustration, la manière d'atteindre le chapitre concerné étant la même que pour la tâche 4.

Types d'erreurs commises par les participants du test en français (tâche 5)



Types d'erreurs commises par les participants du test en allemand (tâche 5)



Problèmes repérés :

- Accès difficile à l'information à l'intérieur des documents (problème 5)
- Confusion liée à la double pagination (problème 8)
- Navigation et repérage peu aisé dans les miniatures (problème 3)

Problèmes suspectés :

- Intitulé « provenance du document original » trop peu clair (problème 14)
- Manque de visibilité du caractère cliquable de l'image du livre des notices (problème 12)

Annexe 2 : Comptes rendus des tests modérés

Participant francophone 1

Problèmes d'utilisabilité détectés : 8 et 1 à confirmer

Profil du participant	Résultats des tâches
Profil général : spécialiste de l'information notamment dans le domaine audio-visuel.	Tâche 1 : succès
Tranche d'âge : 26-35 ans	Tâche 2 : succès
Expertise dans l'utilisation d'Internet : 4	Tâche 3 : abandonnée
Utilisation bibliothèque numérique : 3	Tâche 4 : succès
Expertise travail avec livres anciens : 3	Tâche 5 : échec
Fréquentation d'e-rara.ch au préalable : jamais	Temps total 11 min. 18
Impliqué dans le projet e-rara.ch : non	des tâches : sec.
	Pages vues : 26

Tâche 1 (Bibliothèque des Pasteurs) : succès

Stratégie(s) : utilisation des facettes.

Premier clic : index global.

Temps : 133 secondes (3 pages vues).

Index global

Après avoir lu en diagonal ce qui se trouvait sur la page

Facette
« bibliothèques »

A affiché toutes les bibliothèques et a sélectionné la Bibliothèque des Pasteurs

Tâche 2 (flux RSS) : succès

Stratégie(s) : icône.

Premier clic : icône permettant d'accéder à la liste des flux.

Temps : 47 secondes (2 pages vues).

Tâche 3 (livre le plus ancien d'Erasmus) : abandonnée

Stratégie(s) : parcourir auteurs.

Premier clic : parcourir auteurs.

Erreurs : facette « Erasmus » sur la page des auteurs en « A », échec dans le repérage de la barre de recherche.

Temps : 196 secondes (7 pages vues).

Parcourir auteur

La première page comporte des notices avec des auteurs commençant par A. Les facettes ne concernent que les notices affichées.

Clique sur la facette « Erasmus » et obtient la liste des ouvrages dont Erasmus est un des auteurs, mais dont un autre commence par A.

A dû scroller latéralement pour parvenir aux facettes, la page ne

	se redimensionnant pas en fonction du petit écran de l'ordinateur utilisé pour le test.
	S'est interrogée sur l'équation Erasmus-Erasme (et est partie du principe que les deux renvoyaient à la même personne).
Tri par date	A trié ces résultats selon la date, par ordre croissant.
Parcourir auteur	A répété la manœuvre. S'est interrogée plusieurs fois concernant la fonction recherche : « mais elle est où, la recherche ? ». A eu l'impression de parvenir à un résultat de la tâche par hasard.

Problèmes repérés :

- Croisement entre les différents points d'accès aux documents (menus « collections » et « parcourir », facettes, classement alphabétique).
- Scroll latéral fréquent sur des écrans étroits.
- Fenêtre de recherche peu visible.
- Fonctionnalité « parcourir auteur » peu claire avec l'affichage des notices de documents (on attendrait une liste d'auteurs, cela permettrait de répondre à des questions comme celle de savoir si Erasme est équivalent d'Erasmus).

Tâche 4 (nature du griffon chez Buffon) : succès

Stratégie(s) : parcourir les tables des contenus des différents volumes

Premier clic : lien permanent

Erreurs : lien permanent

Temps : 93 secondes (4 pages vues)

Lien permanent	A pensé qu'il conduisait au contenu.
Contenu	A passé en revue les contenus des 9 tomes.
Chapitre griffon	Trouve la réponse après lecture de la page.

Problème repéré :

- Notion de « lien permanent » peu claire.
- Recherche difficile dans la liste des contenus.

Tâche 5 (trouver une illustration de faucon) : échec

Stratégie(s) : contenu – chapitre.

Premier clic : contenu.

Erreurs : pas su gérer les positions du cadre rouge (prendre images après, à la fin du chapitre).

Temps : 209 secondes (10 pages vues).

Contenu	
Chapitre faucon	
Mosaïque	A mis du temps avant de trouver cette fonction. Clique sur la page avec gravure du sacre (fin du chapitre)

	précédent). En fait, un cadre rouge entoure la page au point de lecture où on se trouve l'internaute, ici la première page du chapitre consacrée faucon.
Première de couverture	Mauvaise manipulation des outils de navigation au sein de l'ouvrage.
Mosaïque	
Chapitre faucon	Toujours en mode mosaïque
Image du sacre	Clique sur la page avec gravure du sacre (fin du chapitre précédent).

Problèmes repérés :

- Navigation et repérage peu aisés dans les pages des miniatures.
- Fonctionnalité « mosaïque » difficile à repérer ou peu explicite. ?

Questionnaire posttest :

	Complètement en désaccord		Complètement d'accord	
J'ai réussi à obtenir des résultats pertinents aux tâches proposées	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>
Ce site est facile d'utilisation.	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>
Ce site est inutilement compliqué.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>
Je recommanderai ce site, si une de mes connaissances s'intéresse aux livres anciens.	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>
J' imagine que, pour la plupart des gens, apprendre à utiliser ce site ne demande que peu de temps.	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
L'utilisation de ce site m'a paru très agréable.	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>

Recommandations et impressions du participant :

- Petit tutoriel vidéo (préfère quelque chose à regarder plutôt qu'un texte à lire).
- Plus de liens entre les différents chemins ? (impression de cloisonnement, n'a pas su donner davantage de précisions).

On peut en déduire le problème suivant :

- Manque d'aide générale.

Commentaires finaux :

L'ordinateur utilisé par le participant avait un petit écran et la taille de la page du site ne s'y adaptait pas. Il a donc fallu scroller latéralement, ce qui était ennuyeux et ne permettait pas au participant de bien voir les facettes disposées à droite.

Le participant a rencontré de nombreuses difficultés dans la réalisation des tâches. La lecture des scénarios des tâches était souvent trop rapide et demandait ensuite des retours. Une des raisons est probablement à chercher dans le capital de patience émoussé par les problèmes techniques que nous avons eus pour mettre en place le dispositif (notamment trouver une connexion Internet, car le wifi n'offrait pas suffisamment de stabilité).

Un point particulièrement remarquable de ce test consiste dans l'échec du repérage de la fenêtre de recherche alors que le participant a affirmé plusieurs fois l'avoir cherchée.

Participant francophone 2

Problèmes d'utilisabilité détectés : 5 et 2 à confirmer

Profil du participant	Résultats des tâches
Profil général : formation universitaire, spécialiste de l'information et de la mise en valeur du patrimoine Tranche d'âge : 36-50 ans Expertise dans l'utilisation d'Internet : 3 Utilisation bibliothèque numérique : 3 Expertise travail avec livres anciens : 4 Fréquentation d'e-rara.ch au préalable : une fois ou l'autre Impliqué dans le projet e-rara.ch : non	Tâche 1 : succès Tâche 2 : succès Tâche 3 : échec Tâche 4 : échec Tâche 5 : succès Temps total des tâches : 22 min. 42 sec. Pages vues : 48

Tâche 1 (Bibliothèque des Pasteurs) : succès

Stratégie(s) : utilisation des facettes.

Premier clic : parcourir lieu.

Temps : 553 secondes (12 pages vues).

Parcourir lieu

Sous la lettre A, a utilisé la facette bibliothèque pour sélectionner l'ouvrage de la bibliothèque des Pasteurs.

A désactivé la facette et choisi la lettre B.

Consulte une notice d'un des livres de la Bibliothèque des Pasteurs et accède à la liste des contenus puis à un chapitre.

Collections thématiques

A sélectionné la collection Bibleikon de la Bibliothèque des Pasteurs, a cliqué sur la première notice et est entrée dans le livre (lien image de l'ouvrage)

Le participant déclare la tâche « moins claire que prévue », car il est parvenu à deux listes comprenant des ouvrages de la Bibliothèque des Pasteurs avec un nombre d'ouvrages différent et ne comprend pas la raison de cette différence.

Parcourir lieu

Accède à la lettre B et décide de mettre un terme à la tâche

A profité de la tâche 1 pour explorer un peu le site. Il a d'ailleurs fallu lui rappeler plusieurs fois que, si elle estimait avoir achevé la tâche, il lui fallait cliquer sur fin de tâche et passer à la suivante.

Problèmes repérés :

- Croisement entre les différents points d'accès aux documents (menus « collections » et « parcourir », facettes, classement alphabétique).

Tâche 2 (flux RSS) : succès

Stratégie(s) : icône

Premier clic : icône permettant d'accéder à la liste des flux

Temps : 47 secondes (2 pages vues)

Tâche 3 (livre le plus ancien d'Erasmus) : échec

Stratégie(s) : parcourir auteur.

Premier clic : parcourir auteur.

Erreurs : facette « Erasmus » sur la page des auteurs en A, échec dans le repérage de la barre de recherche.

Temps : 196 secondes (7 pages vues).

Parcourir auteur	Choisit la lettre E, mais clique ensuite par mégarde sur R. Revient sur E.
Facette « Erasmus »	Espère que les facettes proposées lui permettent de retrouver Erasme. Et en effet, le participant constate qu'Erasmus s'y trouve.
Facette « 1501-1600 »	Tentative pour repérer le plus ancien, mais finalement pas de tri.

Problèmes repérés :

- Fonctionnalité « parcourir auteur » peu claire avec l'affichage des notices de documents. Dans ce cas, le participant en est réduit à « espérer » de l'aide du côté des facettes.

Tâche 4 (nature du griffon chez Buffon) : échec

Stratégie(s) : rechercher le livre, parcourir les tables des contenus des différents volumes.

Premier clic : parcourir auteur.

Erreurs : rechercher l'ouvrage alors que la notice était la page de départ, confusion au niveau des systèmes de pagination, n'a pas su trouver l'information recherchée sur la page qui la contenait pourtant.

Temps : 447 secondes (17 pages vues).

Parcourir auteur	N'a pas remarqué que la notice était la page de départ.
Table des matières du livre numérisé	Retrouve la notice et accède au contenu en cliquant sur l'image du livre. Parcourt le livre de page de page et consulte la table des matières du livre numérisé. On peut se demander pourquoi le participant n'utilise pas les outils de navigation offerts par le système. Trouve que le griffon se situe à la page 151. A l'aide du menu se rend à la page [151] (pagination du système) et aboutit à l'illustration du balbuzard. Retourne à la table et accède à la page [197] 151 qui cette fois porte sur le griffon.
Feuilleter le chapitre	Parcourt les six premières pages sans trouver l'information demandée.

Problème repéré :

- Confusion liée à la double pagination.
- Menu horizontal (titre, contenu, etc.) peu visible. ?

Tâche 5 (trouver une illustration de faucon) : succès

Stratégie(s) : contenu.

Premier clic : contenu.

Erreurs : pas su gérer les positions du cadre rouge (prendre images après, à la fin du chapitre).

Temps : 159 secondes (10 pages vues).

Contenu

Chapitre faucon

Mosaïque

Clique sur la page avec gravure du sacre (fin du chapitre précédent). En fait, un cadre rouge entoure la page au point de lecture où on se trouve l'internaute, ici la première page du chapitre consacrée faucon.

Constate l'erreur et retourne à la mosaïque.

Passe à la série suivante qui comprend le chapitre du faucon.

Clique sur l'image du faucon.

Problèmes repérés :

- Navigation et repérage peu aisés dans les pages des miniatures

Questionnaire posttest :

	Complètement en désaccord			Complètement d'accord		
J'ai réussi à obtenir des résultats pertinents aux tâches proposées	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	
Ce site est facile d'utilisation.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	
Ce site est inutilement compliqué.	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	
Je recommanderai ce site, si une de mes connaissances s'intéresse aux livres anciens.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	
J'imagine que, pour la plupart des gens, apprendre à utiliser ce site ne demande que peu de temps.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	
L'utilisation de ce site m'a paru très agréable.	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	

Recommandations et impressions du participant :

- Au niveau du design, manque de contrastes et de couleurs.
- Le site « pourrait être plus intuitif, plus simple, mais ce n'est pas catastrophique ».
- Page un peu surchargées : beaucoup d'options sur l'écran, certaines peut-être superflues selon les choix précédents de l'utilisateur et pourraient s'effacer si on ne les a pas choisies, donc plus personnalisable. Elle a donné en exemple le site d'une bibliothèque de Pennsylvanie.
- Certains outils et certaines fonctionnalités pourraient être mieux mis en relief.
- Le chemin le plus probable pour parvenir à l'information qu'on cherche pourrait être mieux mis en avant, l'utilisateur plus guidé.

On peut en déduire les problèmes suivants :

- Aide insuffisante.
- Contrastes trop faibles, manque de couleur. ?

Pour les autres points, il est difficile d'être suffisamment spécifique pour en déduire un problème donné.

Commentaires finaux :

L'ordinateur utilisé par le participant avait un petit écran et la taille de la page du site ne s'y adaptait pas. Il a donc fallu scroller latéralement, ce qui était ennuyeux et ne permettait pas au participant de bien voir les facettes disposées à droite.

Le participant semblait passablement gêné par la méthode du think-aloud. En effet, il était difficile de l'amener à verbaliser ce qu'il faisait. Il semblait même parfois complètement concentré sur sa tâche et ignorer complètement nos injonctions.

Comme pour le premier participant francophone, la lecture de certains scénarios de tâche lui a posé problème (peut-être la langue maternelle a-t-elle un rôle dans ce cas).

Participant francophone 3

Problèmes d'utilisabilité détectés : 8 et 3 à confirmer

Profil du participant	Résultats des tâches
Profil général : historienne de l'art, spécialiste d'art moderne Tranche d'âge : 36-50 ans Expertise dans l'utilisation d'Internet : 3 Utilisation bibliothèque numérique : 3 Expertise travail avec livres anciens : 3 Fréquentation d'e-rara.ch au préalable : jamais Impliqué dans le projet e-rara.ch : non	Tâche 1 : succès Tâche 2 : ignorée Tâche 3 : succès Tâche 4 : échec Tâche 5 : succès Temps total des tâches : 42 min. 1 sec. Pages vues : 101

Tâche 1 (Bibliothèque des Pasteurs) : succès

Stratégie(s) : parcourir les collections.

Premier clic : collections thématiques.

Temps : 553 secondes (12 pages vues)

Intention de téléphoner	Spontanément après lecture de la tâche, le participant s'exclame « c'est facile, on peut téléphoner », mais il comprend ensuite qu'on attend de lui qu'il trouve l'information sur le site.
Collections thématiques	« C'est bizarre », liste qui ne parle pas immédiatement au participant. Après avoir déroulé la liste des « collections thématiques », regarde assez longuement la liste pour y repérer une mention de la Bibliothèque des Pasteurs. Trouve la collection « Bibleikon », clique dessus.
Consultation d'un ouvrage	Consulte une bible de la collection et tente même de télécharger l'ouvrage. Le modérateur intervient afin que le test se poursuive sur le navigateur et que Loop11 puisse tracer les actions. Le participant affirme ne pas être sûr que ce livre est mis à disposition. Il a expliqué après la tâche qu'il entendait « mettre à disposition » dans un sens physique. C'est pourquoi, il a ensuite accédé au catalogue RERO pour vérifier si on pouvait réellement accéder au livre dans cette bibliothèque.

Problèmes repérés :

- Liste des collections thématiques peu compréhensible (sentiment de manque de clarté chez le participant). ?

Tâche 2 (flux RSS) : ignorée

Ne sait pas ce qu'est un flux RSS.

Tâche 3 (livre le plus ancien d'Erasmus) : succès

Stratégie(s) : parcourir auteur, parcourir date, facette « Erasmus », facettes de période.

Premier clic : parcourir auteur.

Erreurs : facette « Erasmus » sur la page des auteurs en A, tri par auteur au lieu de par année.

Temps : 1038 secondes (38 pages vues).

Parcourir auteur

La première page comporte des notices avec des auteurs commençant par A. Les facettes ne concernent que les notices affichées.

Clique sur la facette « Erasmus » et obtient la liste des ouvrages dont Erasmus est un des auteurs mais dont un autre commence par A.

« quelle drôle manière de faire », le participant explique : « j'ai déjà demandé de trier par auteur et après il te redemande de trier par auteur, et avec tout ça tu n'as toujours pas le nom de l'auteur à mettre quelque part. »

Manipule les filtres et croit que « croissant » et « décroissant » concerne forcément la date et ne saisit pas l'interaction avec le menu « trier par ».

Tente d'afficher plus de résultats par page (100).

Supprime la facette « Erasmus » et se retrouve à rechercher dans l'ensemble du catalogue. Sans vraiment en prendre conscience.

Trouve le tri par date et le livre le plus ancien (avec un auteur commençant par A) et constate qu'Erasmus n'est pas l'auteur.

Recommence en cliquant sur « parcourir auteur ».

Choisit à nouveau la facette « Erasmus », mais affirme « je ne comprends pas la logique ».

Parcourir date

Renonce immédiatement à cette option vu qu'elle ne favorise pas la recherche d'auteur. retourne aux auteurs, puis repasse par les dates.

Parcourir auteur

Retourne sur « parcourir auteur ».

Le participant s'étonne de ne pas pouvoir chercher à la fois par date et auteur : « soit on met auteur soit on met titre soit on met lieu, pourquoi on n'arrive pas à faire plusieurs recherches en même temps ».

Recherche sur Wikipedia les dates d'activité d'Erasmus.

« Bon, il est nul ce site ! »

« Avec ça je ne sais toujours pas ce que signifie croissant et décroissant. Il y a plein de mots qui ne sont pas définis, il y a plein de sous-entendus, il y a plein d'implicite, quand on est sensé savoir alors que le contexte ne donne pas la définition, alors ça ça m'énerve. »

Comprend enfin le système de tri. Il s'agit d'un classement alphabétique (estime que croissant et décroissant ne s'applique qu'à des chiffres) de la page « parcourir auteur ».

clique sur la lettre E.

Facette « Erasmus »

Facettes périodes

Grâce une sélection de facettes successives, trouve le bon livre.

Problèmes repérés :

- Croisement entre les différents points d'accès aux documents (menus « collections » et « parcourir », facettes, classement alphabétique).
- Fonctionnalité « Parcourir auteur » peu claire avec l'affichage des notices de documents. Dans ce cas, le participant en est réduit à « espérer » de l'aide du côté des facettes.
- Fenêtre de recherche peu visible.
- Page parcourir auteur pas assez documentée par de l'aide.
- Termes « croissant » et « décroissant » indifféremment appliqué à des chiffres qu'à des lettres. ?
- Interaction entre les différents menus des outils de tri peu claire.

Tâche 4 (nature du griffon chez Buffon) : **échec**

Stratégie(s) : consulter la notice RERO, parcourir les tables des contenus des différents volumes.

Premier clic : consulter la notice RERO.

Erreurs : rechercher l'ouvrage alors que la notice était la page de départ, confusion au niveau des systèmes de pagination, a oublié de rechercher l'information sur la page après l'avoir trouvée.

Temps : 348 secondes (12 pages vues).

Recherche d'un accès au contenu du livre

Au début de la tâche, le participant souhaite ouvrir un fichier pdf, mais le modérateur explique que cette action n'est pas autorisée.

Le participant relit la tâche afin de mieux la comprendre.

Il met beaucoup de temps jusqu'à trouver une solution (plus d'une minute)

Mode page

Finalement il clique sur le lien « mode page ».

Il parcourt les premières pages jusqu'à la table des matières.

Trouve le numéro de page du chapitre du griffon (p. 151).

Accède à la page 151 du système et tombe sur une illustration d'un balbuzard.

Constata son erreur et atteint cette fois-ci le chapitre du griffon.

Le participant est heureux du résultat et clique sur achever la tâche sans lire (« voilà c'est bon...Ah merde j'aurais dû lire ! »).

Problème repéré :

- Confusion liée à la double pagination.
- Menu horizontal (titre, contenu, etc.) peu visible. ?

Tâche 5 (trouver une illustration de faucon) : succès

Stratégie(s) : après un temps de confusion, menu contenu, utilisation de la mosaïque.

Premier clic : contenu.

Erreurs : chercher une illustration de griffon, pas su gérer les positions du cadre rouge (sélection d'une image incorrecte).

Temps : 923 secondes (46 pages vues).

Confusion avec la tâche précédente

Au début de la tâche, le participant croit devoir rechercher une illustration de griffon. Cela l'amène à poursuivre une série de mauvaises pistes.

Accède au chapitre du griffon.

Parcourt cinq pages.

Essaie le chapitre précédent et le suivant via le menu.

Atteint la page précédente (dernière du chapitre du griffon) pour voir s'il y a pas une illustration du griffon à la fin du chapitre.

Recourt à la fonctionnalité « mosaïque ».

Atteint la dernière page du volume (affichage mosaïque).

Atteint la première page du livre (affichage mosaïque).

Parcourt le livre via les pages de mosaïques.

Clique sur les images quand il en voit et si ce n'est pas la bonne retourne sur l'affichage mosaïque.

Décide finalement de retourner au chapitre du griffon via le menu.

Choisit l'affichage mosaïque.

Constata qu'il n'y a pas d'illustration.

Comprend son erreur et décide de chercher un faucon sort.

Chapitre faucon

Atteint le chapitre faucon.

Enclenche l'affichage mosaïque.

Clique sur l'image du sacre (chapitre précédent).

Constata l'erreur et retourne à l'affichage des mosaïques sous « faucon ».

Tome 2 et 3

Un peu perdu, tente de regarder si dans les autres tomes, il n'y a pas aussi des faucons. Affichage sous forme de mosaïque du début des tomes 2 et 3.

Page « contenu »

Trouve un autre chapitre en rapport avec le faucon : Oiseaux étrangers, qui ont rapport au gerfaut & aux faucons (en fait le suivant de celui consacré exclusivement au faucon).

Choisit ce chapitre.

Affichage sous forme de mosaïque.

Clique sur une image du chapitre précédent qui se trouve être celle qui est recherchée.

Faire exécuter deux tâches à l'intérieur du même livre sur des sujets similaires pouvait en effet engendrer de la confusion. Le participant prend énormément de temps avant de constater son erreur.

Problèmes repérés :

- Navigation et repérage peu aisés dans les pages des miniatures

Questionnaire posttest :

	Complètement en désaccord			Complètement d'accord		
J'ai réussi à obtenir des résultats pertinents aux tâches proposées	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	
Ce site est facile d'utilisation.	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	
Ce site est inutilement compliqué.	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	
Je recommanderai ce site, si une de mes connaissances s'intéresse aux livres anciens.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	
J'imagine que, pour la plupart des gens, apprendre à utiliser ce site ne demande que peu de temps.	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	
L'utilisation de ce site m'a paru très agréable.	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	

Recommandations et impressions du participant :

- Plus d'aide, fonctions plus explicites.
- Image du livre cliquable pas visible.
- Estime inadmissible de mettre la fenêtre de recherche dans l'en-tête, on ne la voit pas.

On peut en déduire le problème suivant :

- Aide insuffisante.
- Manque de visibilité du caractère cliquable de l'image du livre des notices.
- Fenêtre de recherche peu visible.

Commentaires finaux :

La méthode du think-aloud a été très efficace au début, mais le test a été tellement long (plus de 45 minutes) que cette contrainte devenait de plus en plus difficile à tenir sur la durée. Les commentaires liés à la tâche 5 sont très pauvres.

Comme pour les deux premiers participants francophones, la lecture de certains scénarios de tâche lui a posé problème (peut-être la langue maternelle a-t-elle un rôle dans ce cas).

On remarquera que l'utilisateur conserve des habitudes de lecteur de livres physiques comme lorsqu'il consulte la table du livre lui-même et non celle qui est proposée par le système.

Participant francophone 4

Problèmes d'utilisabilité détectés : 6 et 1 à confirmer

Profil du participant	Résultats des tâches
Profil général : historien	Tâche 1 : succès
Tranche d'âge : 26-35 ans	Tâche 2 : succès
Expertise dans l'utilisation d'Internet : 4	Tâche 3 : succès
Utilisation bibliothèque numérique : 2	Tâche 4 : succès
Expertise travail avec livres anciens : 3	Tâche 5 : échec
Fréquentation d'e-rara.ch au préalable : jamais	Temps total 15 min. 6
Impliqué dans le projet e-rara.ch : non	des tâches : sec.
	Pages vues : 76

Tâche 1 (Bibliothèque des Pasteurs) : succès

Stratégie(s) : parcourir les collections, lien « à propos ».

Premier clic : collections thématiques.

Temps : 553 secondes (12 pages vues).

Collections
thématiques

A cliqué dans les « divers » puis est retourné à la page d'accueil.

Page « à propos d'e-
rara.ch »

A cliqué sur les « bibliothèques participantes ».

Clique « Bibliothèque des Pasteurs », puis la liste des ouvrages et la notice du premier d'entre eux en s'exclamant : « Attends, j'aimerais juste voir... ah oui, là on a les textes et après on peut les chercher. (...) Ah y a même les pdf ! C'est magique ! »

Tâche 2 (flux RSS) : succès

Stratégie(s) : icône flux RSS.

Premier clic : icône flux RSS.

Temps : 25 secondes (3 pages vues)

Tâche 3 (livre le plus ancien d'Erasmus) : succès

Stratégie(s) : recherche simple, index auteur, recherche avancée, recherche simple (toujours avec Erasmus), index auteurs, recherche simple (avec Erasmus), recherche avancée avec Erasmus.

Premier clic : recherche simple

Erreurs : rechercher « Erasmus »

Temps : 230 secondes (18 pages vues).

Recherche simple

A tapé Erasmus (et a testé l'ordre croissant / décroissant des résultats).

A vu que les résultats ne concordaient pas (« ben là, c'est Eratosthène, c'est pas Erasmus »).

Parcourir auteur	N'a pas poursuivi cette stratégie
Recherche avancée	Erasme dans le champ auteur et a sélectionné toutes les autres possibilités (bibliothèques, époques, collections, etc.) : cela n'aboutit à aucun résultat.
Recherche simple	Erasme : même résultats qu'au début, et a cliqué sur la notice du livre d'Eratosthène pour s'assurer que ce n'était pas ça. « Bon, c'est marrant, mais c'est marqué nulle part que c'est Erasme, c'est un peu bizarre ».
Parcourir auteur	A pris la lettre E « C'est nul, on a trois trucs pis on sait pas à quoi ça sert... » A vu Erasmus dans les facettes, du coup a voulu essayer avec cette dénomination, mais sans cliquer sur la facette elle-même.
Recherche simple	Entre « Erasmus ». « Ah ben voilà ! Il s'appelle pas Erasme, il s'appelle Erasmus, déjà, un ».
Recherche avancée	Entre Erasmus dans le champ auteur, pour vérifier que ce qu'il a eu en recherche simple était bien les bonnes données. « Pis je devrais avoir que lui. Je ne suis pas vraiment sûr d'avoir trouvé, mais allons-y comme ça »
Tri par date	A trié les résultats obtenus et est parvenu à la bonne réponse.

Problèmes repérés :

- Absence de renvois entre les différentes dénominations des auteurs.
- Page parcourir auteur pas assez documentée par de l'aide.

Tâche 4 (nature du griffon chez Buffon) : succès

Stratégie(s) : consulter la notice RERO, parcourir les tables des contenus des différents volumes.

Premier clic : lien permanent

Erreurs : lien permanent

Temps : 86 secondes (7 pages vues)

Lien permanent	A cliqué sur l'icône info du lien permanent, pour voir ce que c'était, et s'en est aussitôt rendu compte. « Non, mais c'est le lien pour toujours arriver dessus en fait ». Reprend la notice de départ (lien « titre »).
Tome premier	« Alors, si je fais ça, il se passe quoi ? Ah, pas mal ! Bon, qu'est-ce que j'ai fait pour arriver à ça ? » A refait la manœuvre (passage de la notice à la liste des chapitres du tome 1).
Chapitre griffon	« Je vois qu'il y a un truc qui s'appelle le griffon dans le contenu de ce bouquin ». Clique sur le lien du chapitre et lit la réponse sur la première page.

Problème repéré :

- Notion de « lien permanent » peu claire. ?

Tâche 5 (trouver une illustration de faucon) : échec

Stratégie(s) : contenu, menu déroulant au-dessus du livre,

Premier clic : contenu du tome 1.

Erreurs : utiliser un chapitre du livre appelé « plan de l'ouvrage », feuilleter au lieu d'utiliser la mosaïque.

Temps : 424 secondes (39 pages vues).

Contenu tome 1	A cliqué sur le lien tome premier et a pris le chapitre « plan de l'ouvrage » en pensant que c'était une sorte de sommaire : « Ah, mais je voulais le sommaire en fait (...) c'est pas ce que je veux »
Chapitre faucon	<p>A utilisé le menu déroulant en-dessus de l'ouvrage pour parvenir au chapitre concernant le faucon. Il feuillette les 7 premières pages.</p> <p>« Bon, je pense qu'en fait il n'y aura pas d'illustrations dans ce bouquin »</p> <p>Veut tester une autre stratégie, mais note le numéro de la page où il s'est arrêté, au cas où, et fait « page précédente » pour revenir à la notice de départ.</p>
Recherche d'une liste des figures	<p>Cherche s'il n'y aurait pas une liste des figures dans la liste des chapitres du tome 1, puis du tome 2, la table de ce qui est contenu dans le volume 1.</p> <p>« Bon, ok, ça ne me dit pas s'il y a des images »</p>
Chapitre faucon	<p>Retourne à la page du chapitre faucon où il s'était arrêté grâce aux menus déroulants et feuillette les 4 pages suivantes.</p> <p>« Bon, j'espère qu'il n'y a pas mille pages sur ce faucon ! »</p>
Chapitre suivant	<p>« J'ai vu un truc. Peut-être que c'est à la fin du chapitre, donc je prends le chapitre suivant pour juste devoir reculer et pas me taper tout le chapitre »</p> <p>Il s'exécute et parvient au faucon hagard, puis au faucon sort.</p>
Trouver le numéro de page	<p>« Je dois noter le numéro de la page. 333. C'est noté deux cents... Oh les salauds, y a deux numéros de page ! »</p> <p>A finalement regardé le no de la page précédente (la dernière du texte) et fait plus un (267 : or, les illustrations ont une pagination spéciale.</p>

Problèmes repérés :

- Navigation et repérage peu aisés dans les pages des miniatures.
- Confusion liée à la double pagination.
- Manque un index d'images.

Questionnaire posttest :

	Complètement en désaccord			Complètement d'accord	
J'ai réussi à obtenir des résultats pertinents aux tâches proposées	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>
Ce site est facile d'utilisation.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>
Ce site est inutilement compliqué.	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Je recommanderai ce site, si une de mes connaissances s'intéresse aux livres anciens.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>
J'imagine que, pour la plupart des gens, apprendre à utiliser ce site ne demande que peu de temps.	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
L'utilisation de ce site m'a paru très agréable.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>

Recommandations et impressions du participant :

- Guide d'utilisation.

On peut en déduire le problème suivant :

- Aide insuffisante.

Commentaires finaux :

Cet utilisateur anticipe certains problèmes d'utilisabilité. Par exemple, il sélectionne toutes les sources (alors que si le système fonctionne bien, c'est en principe le choix par défaut). Il anticipe aussi les « pièges » potentiels puisqu'il fait une recherche avancée pour être assuré qu'Erasme soit bien auteur pour les documents donnés en résultat.

Participant francophone 5

Problèmes d'utilisabilité détectés : 5 et 1 à confirmer

Profil du participant	Résultats des tâches
Profil général : historien et expérience dans l'enseignement de l'histoire. Tranche d'âge : 51-65 ans Expertise dans l'utilisation d'Internet : 3 Utilisation bibliothèque numérique : 3 Expertise travail avec livres anciens : 5 Fréquentation d'e-rara.ch au préalable : jamais Impliqué dans le projet e-rara.ch : non	Tâche 1 : échec Tâche 2 : ignorée Tâche 3 : échec Tâche 4 : succès Tâche 5 : échec Temps total des tâches : 11 min. 20 sec. Pages vues : 36

Tâche 1 (Bibliothèque des Pasteurs) : **échec**

Stratégie(s) : aucune.

Premier clic : aucun.

Temps : 28 secondes (1 pages vue).

Le participant a cliqué sur « achever la tâche » sans avoir compris qu'on attendait une action de sa part.

Tâche 2 (flux RSS) : **ignorée**

Le participant ne savait pas ce qu'est un flux RSS. Il en a profité pour parcourir le site et on a pu constater le problème suivant :

- Position du lien vers le pdf sur la droite trop éloignée du texte si l'écran est grand. Le participant s'est plaint de n'avoir pas vu le fichier dans un premier temps.

Tâche 3 (livre le plus ancien d'Erasme) : **échec**

Stratégie(s) : recherche simple

Premier clic : recherche simple

Erreurs : rechercher « Erasme »

Temps : 45 secondes (non enregistrés par Loop11 suite à une mauvaise manipulation : cliquer terminer la tâche sans le vouloir).

A tapé Erasme et a trouvé les deux résultats sortant dont Erasme n'est pas l'auteur. Il perd assez rapidement espoir « Bon, alors là, je sais pas du tout quoi faire » et a terminé la tâche.

Problèmes repérés :

- absence de renvois entre les différentes dénominations des auteurs.

Tâche 4 (nature du griffon chez Buffon) : succès

Stratégie(s) : consulter la notice RERO, parcourir les tables des contenus des différents volumes.

Premier clic : menu contenu (tous les tomes)

Erreurs : lien permanent

Temps : 117 secondes (4 pages vues)

Menu contenu	A cliqué sur « contenu » sans trop espérer y trouver une liste de ce que contient l'ouvrage : « Oh, mais ça ne me donnera pas l'intérieur... Oh, ben oui ». A mis un moment avant de trouvé le bon chapitre, car n'avait pas vu qu'il fallait scroller pour voir tous les chapitres.
Chapitre griffon	A choisi le chapitre consacré au griffon dans la liste des chapitres et a parcouru les deux premières pages.

Problèmes repérés :

- Recherche difficile dans la liste des contenus.
- Menu horizontal (titre, contenu, etc.) peu visible.

Tâche 5 (trouver une illustration de faucon) : échec

Stratégie(s) : contenu, menu déroulant au-dessus du livre,

Premier clic : lien « provenance du document original », menant vers la notice rero du livre à la BPU de Neuchâtel.

Erreurs : absence d'utilisation de la mosaïque et de menus en général au-dessus du document, tentative de suivre des hyperliens dans les pages du livre.

Temps : 490 secondes (30 pages vues).

Notice rero	Clique sur le lien « provenance du document original » et aboutit à une notice RERO. Le participant commente sa déconvenue : « Ah, ben je me suis fait avoir ».
Mode page	A cliqué sur l'image du livre, pour entrer dans le contenu. A feuilleté 2 pages et est revenu à la notice.
Contenu	A passé un long moment avant de retrouver où il avait cliqué pour parvenir à la liste des chapitres, se focalisant sur la notice, mais pas sur ce qui se trouvait plus haut.
Chapitre faucon	A passé un long moment à aller et revenir entre les trois premières pages du chapitre : le texte mentionnait des planches et il essayait d'y accéder en cliquant dessus (lien hypertexte), alors qu'en fait cliquer sur la page mène à la suivante (donc retour ensuite sur les premières pages). Sait d'après les livres anciens qu'il a chez lui que « à certaines places dans le livre, il y a des animaux, mais alors comment y arriver... ». Ne savait pas comment procéder à partir de là et a cliqué sur terminer la tâche. « Alors là, je ne sais pas ».

Problèmes repérés :

- Intitulé « provenance du document original » trop peu clair. ?
- Menu horizontal (titre, contenu, etc.) peu visible.

Questionnaire posttest :

	Complètement en désaccord			Complètement d'accord		
J'ai réussi à obtenir des résultats pertinents aux tâches proposées	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>
Ce site est facile d'utilisation.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>
Ce site est inutilement compliqué.	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Je recommanderai ce site, si une de mes connaissances s'intéresse aux livres anciens.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>
J'imagine que, pour la plupart des gens, apprendre à utiliser ce site ne demande que peu de temps.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>
L'utilisation de ce site m'a paru très agréable.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>

Recommandations et impressions du participant :

- Aide plus visible.

On peut en déduire le problème suivant :

- Aide insuffisante.

Commentaires finaux :

Cet utilisateur a exécuté souvent rapidement les tâches sans bien comprendre le fonctionnement du site. Il a toutefois permis de trouver des problèmes encore inconnus et pour cette raison se révèle malgré tout intéressant. La diversité des profils semble un avantage.

Participant germanophone 1

Problèmes d'utilisabilité détectés : 3 et 2 à confirmer

Profil du participant	Résultats des tâches
Profil général : écrivain, accomplit parfois des recherches historiques pour ses œuvres. Tranche d'âge : 51-65 ans Expertise dans l'utilisation d'Internet : 2 Utilisation bibliothèque numérique : 1 Expertise travail avec livres anciens : 1 Fréquentation d'e-rara.ch au préalable : jamais Impliqué dans le projet e-rara.ch : non	Tâche 1 : échec Tâche 2 : succès Tâche 3 : échec Tâche 4 : succès Tâche 5 : échec Temps total des tâches : 22 min. 30 sec. Pages vues : 30

Tâche 1 (Bibliothèque des Pasteurs) : **échec**

Stratégie(s) : parcourir les collections, lien « à propos ».

Premier clic : collections thématiques.

Temps : 212 secondes (2 pages vues).

Accueil

Lecture du texte de la page d'accueil

Index des titres

Index des titres. Ne sait pas comment retrouver un livre d'une bibliothèque spécifique s'il ne connaît pas au préalable au moins un livre possédé par cette bibliothèque. « Ich muss ein Titel eines Buches haben »

Commentaires : ne concevait pas qu'on pouvait trouver une telle information sans connaître de livre appartenant à cette bibliothèque. Le problème relevait assez clairement d'une faible pratique de ce type de bases de données.

Problème repéré :

- Aide insuffisante. ?
- Visibilité des bibliothèques participantes insuffisante. ?

Tâche 2 (flux RSS) : **succès**

Stratégie(s) : lecture du texte et repérage du lien.

Premier clic : lien textuel vers la page des flux rss.

Erreurs : -

Temps : 64 secondes (2 pages vues).

Tâche 3 (livre le plus ancien d'Erasmus) : **échec**

Stratégie(s) : recherche simple « Erasmus ».

Premier clic : recherche simple « Erasmus ».

Erreurs : choisir un livre dont Erasmus n'est pas l'auteur (1513).

Temps : 127 secondes (3 pages vues).

Recherche simple	Lecture du texte de la page d'accueil
Utilisation de filtres	Utilisation de filtres « Aufsteigend ist gut », confirme que l'option par défaut est la bonne. Ne voit pas que le livre de 1513 n'a pas Erasme pour auteur.

Problèmes repérés :

- Affichage des résultats trop peu clair (pas d'étiquettes comme titre, auteur, etc.).

Tâche 4 (nature du griffon chez Buffon) : succès

Stratégie(s) : recherche rapide, parcourir les contenus.

Premier clic : recherche rapide.

Erreurs : recherche griffon dans la recherche rapide en croyant éventuellement chercher dans le livre.

Temps : 479 secondes (5 pages vues).

Traduire « Greif » en français	Via Wikipedia trouve griffon
Recherche simple	« da musst es irgendwo ein Suchdings haben ». Tente d'utiliser la fenêtre de recherche générale à cet effet. Constate qu'il aboutit à une liste d'ouvrage en lien avec l'animal et en conclut que cette piste n'est pas la bonne.
Recherche dans les contenus	Retour à la notice, clique sur « Inhaltsansicht » et parcours les contenus. Choisit le chapitre des griffons et lit : « distinguer des autres vautours... vautour ».
Traduire « vautour » en allemand	Via Wikipedia, le participant trouve la traduction : « Geier ».

Problèmes repérés :

- Accès difficile à l'information à l'intérieur des documents.

La possibilité d'une recherche plein texte devient de plus en plus un standard. Comme on s'attend à la trouver, le participant est tenté d'utiliser la fenêtre de recherche standard en espérant qu'elle remplira cette fonction également.

Tâche 5 (trouver une illustration de faucon) : échec

Stratégie(s) : index des contenus, miniatures.

Premier clic : index des contenus.

Erreurs : clic sur le sacre, puis sur le gerfaut.

Temps : 468 secondes (18 pages vues).

Index des contenus	Trouve faucon mais ne clique pas
Seitenansicht	Affiche les premières pages du livre.
Recherche dans le menu des titres de chapitres	Trouve le faucon

Parcourir la partie de livre autour du faucon	<p>Clique sur le sacre puis sur le gerfaut.</p> <p>Ne comprend pas pourquoi il ne s'agit pas de faucons.</p> <p>Retour au chapitre du faucon.</p>
Page de titre du chapitre faucon	<p>Clique sur la page de titre du chapitre. Il est bien question de faucon, mais il n'y a pas d'image.</p>
Repère le cadre rouge rouge	<p>Repère le cadre rouge rouge (« oder ist es das da rot (impression d'avoir compris) – was bedeutet rot - dass das eine Seite wo der faucon erwähnt ist aber ich habe kein Bild »)</p>
Atteint le chapitre suivant mentionnant également un faucon	<p>Trouve plusieurs illustrations et clique dessus. Il parcourt même l'image recherchée, mais sans le remarquer.</p>

Problèmes repérés :

- Navigation et repérage peu aisés dans les pages des miniatures (notamment le cadre rouge semble peu visible).

Questionnaire posttest :

	Complètement en désaccord			Complètement d'accord		
Ich habe relevante Ergebnisse erhalten, indem ich die Aufgaben löste.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	
E-rara.ch ist leicht zu benutzen.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	
Diese Website ist unnützlich kompliziert.	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	
Ich werde e-rara.ch Bekannten empfehlen, die sich für alte Bücher interessieren.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	
Ich denke, dass die meisten Leute ganz schnell lernen, e-rara.ch zu benutzen.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	
Es hat mir gefallen, e-rara.ch zu benutzen.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	

Recommandations et impressions du participant :

- aucune suggestion (le participant ne se sentait pas assez expert pour donner un avis).

Commentaires finaux :

Cet utilisateur n'a pas une grande pratique des banques de données. Par contre, il pratique Google Books et s'attendait probablement à y retrouver des fonctionnalités similaires comme la recherche plein texte. Il lui faut beaucoup de temps pour lire les pages. Il ne fait que peu de tentatives « pour voir ». Il souhaite comprendre véritablement avant de cliquer.

Participant germanophone 2

Problèmes d'utilisabilité détectés : 4 et 1 à confirmer.

Profil du participant	Résultats des tâches
Profil général : spécialiste de langues de l'Antiquité et de littérature comparée.	Tâche 1 : succès
Tranche d'âge : 26-35 ans	Tâche 2 : succès
Expertise dans l'utilisation d'Internet : 3	Tâche 3 : succès
Utilisation bibliothèque numérique : 3	Tâche 4 : succès
Expertise travail avec livres anciens : 4	Tâche 5 : échec
Fréquentation d'e-rara.ch au préalable : jamais	Temps total 25 min. 5
Impliqué dans le projet e-rara.ch : non	des tâches : sec.
	Pages vues : 94

Tâche 1 (Bibliothèque des Pasteurs) : succès

Stratégie(s) : parcourir les collections, lien « à propos ».

Premier clic : collections thématiques.

Temps : 173 secondes (5 pages vues).

Accueil

Lecture du texte de la page d'accueil

Page « über e-rara.ch »

Lecture, ne sait pas où trouver.

Clique sur « Infos zu den Kollektionen », et trouve la réponse sans en être convaincu.

Problème repéré :

- Visibilité des bibliothèques participantes insuffisante. ?

Tâche 2 (flux RSS) : succès

Stratégie(s) : icône flux RSS.

Premier clic : icône flux RSS.

Erreurs : -

Temps : 87 secondes (2 pages vues).

Tâche 3 (livre le plus ancien d'Erasmus) : succès

Stratégie(s) : recherche avancée.

Premier clic : recherche avancée.

Erreurs : utilisation de l'astérisque pour la troncature

Temps : 256 secondes (11 pages vues).

Recherche avancée

Recherche avancée Cherche Erasmus comme auteur et ajoute une indication de date avec troncature « * ». « Jetzt weiss ich nicht leider nicht ob dieses Datenbank Trunkatur annimmt ».

Refait la recherche sans indication de date.

Menu de tri	Trie par date, prend la peine de regarder les derniers résultats pour s'assurer que le tri est bon.
-------------	---

Problème repéré :

- utilisation de la troncature « * » non acceptée.

Tâche 4 (nature du griffon chez Buffon) : succès

Stratégie(s) : recherche via l'index du dernier volume

Premier clic : tome 8 qui conduit à la liste de son contenu

Erreurs : ne pas utiliser les outils de repérage mis à disposition pour

Temps : 521 secondes (40 pages vues)

Observation de la notice	Remarque que le livre est cliquable.
Traduire Greif en français	Trouve qu'il s'agit d'un griffon
Recherche dans l'index	Voudrait faire une recherche plein texte (ich weiss nicht, ob ich das Wort eingeben kann).
général du dernier volume	Trouve la page et le volume (vol. I p. 151).
Utilisation des menus	Pour trouver le volume et la page, trouve le chapitre et la réponse.

Problèmes repérés :

- Accès difficile à l'information à l'intérieur des documents.

Ce participant utilise exactement la méthode la plus pertinente lorsqu'on consulte le document sous forme papier. Il possède d'ailleurs une grande expérience à ce niveau.

A noter que le participant n'a pas été gêné par la double pagination. Il connaît les conventions qui veulent que les textes entre crochets sont des ajouts des éditeurs et donc que la pagination réelle n'en comporte pas.

Tâche 5 (trouver une illustration de faucon) : échec

Stratégie(s) : index des contenus, miniatures.

Premier clic : index des contenus.

Erreurs : clic sur le sacre, puis sur le gerfaut.

Temps : 468 secondes (36 pages vues).

Clic sur le livre	Réflexe du lecteur de livres papier qui commence par ouvrir.
Utilisation du menu	Parvient au chapitre du griffon
Parcourir les pages	« Ich möchte irgendwo das Wort eingeben können. Jetzt blättere ich mal. ». Parcourt 6 pages.
Fenêtre de recherche	Tente une recherche faucon sort, mais n'aboutit à rien.

Consultation de la page « Inhaltsansicht »	<p>Pour trouver le volume et la page, trouve la réponse. Tente de télécharger le fichier pdf, mais le modérateur prévient la manipulation.</p> <p>Se rend au chapitre suivant : <i>Oiseaux étrangers, qui ont rapport au gerfaut & aux faucons</i>.</p> <p>Parcourt sans succès 5 pages du chapitre.</p>
Va à la fin du livre	<p>Parcourt en reculant, ne sert à rien.</p> <p>Sélectionne la vue des miniatures.</p>
Miniaturansicht	<p>Retourne via le menu au faucon et se trompe en cliquant sur les illustrations du gerfaut puis du sacre. « Ich bin eigentlich beim faucon und hat zwei bilder aber kein Bild entspricht dem... das faucon sort ist vielleicht kein Faucon. ».</p>
Chapitre <i>Oiseaux étrangers, qui ont rapport au gerfaut & aux faucons</i> .	<p>Clique sur les images et trouve plus ou moins par hasard le faucon sort à la fin du chapitre précédent celui du faucon.</p>

Problèmes repérés :

- Navigation et repérage peu aisés dans les pages des miniatures (notamment le cadre rouge semble peu visible).
- Accès difficile à l'information à l'intérieur des documents.

Questionnaire posttest :

	Complètement en désaccord			Complètement d'accord		
Ich habe relevante Ergebnisse erhalten, indem ich die Aufgaben löste.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>
E-rara.ch ist leicht zu benutzen.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>
Diese Website ist unnützlich kompliziert.	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Ich werde e-rara.ch Bekannten empfehlen, die sich für alte Bücher interessieren.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>
Ich denke, dass die meisten Leute ganz schnell lernen, e-rara.ch zu benutzen.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>
Es hat mir gefallen, e-rara.ch zu benutzen.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>

Recommandations et impressions du participant :

- souhait d'avoir une recherche plein texte.

Problème déduit :

- Accès difficile à l'information à l'intérieur des documents.

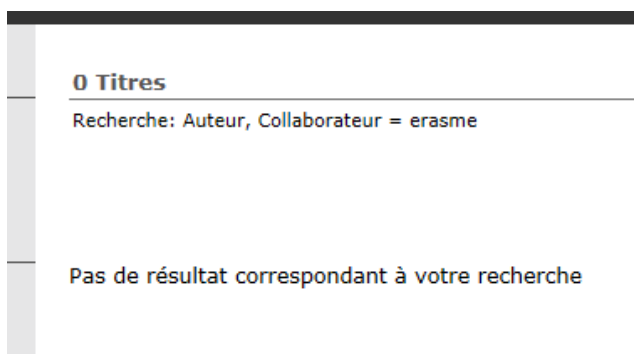
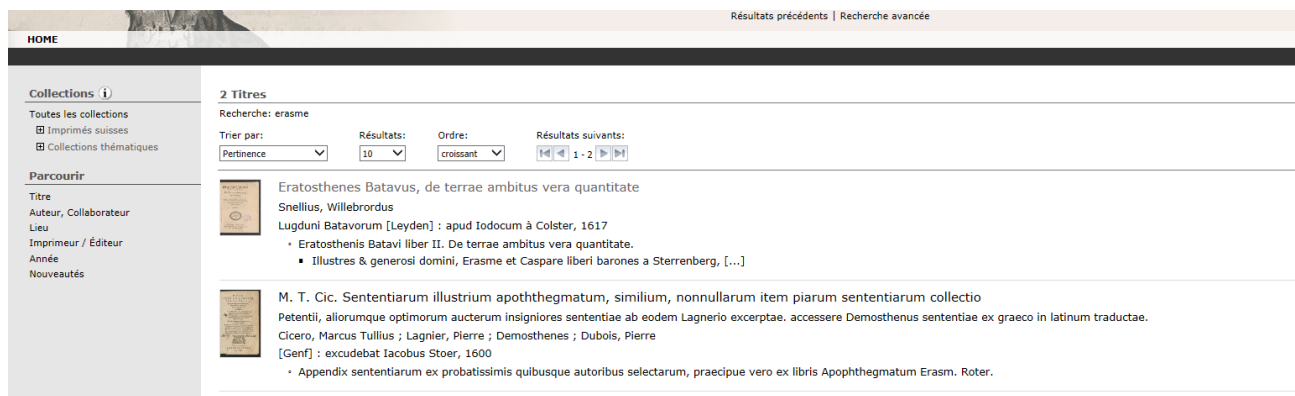
Commentaires finaux :

Cet utilisateur a une très grande pratique de la consultation de livres anciens. Ce fait est visible dans sa manière d'aborder certains problèmes. Il s'oriente facilement dans les tables et les textes des livres eux-mêmes et ne pense pas forcément à recourir à l'aide que met à sa disposition la bibliothèque numérique. Il consulte les ouvrages comme s'il s'agissait de papier.

Annexe 3 : Problèmes d'utilisabilité et recommandations

Les différents problèmes sont classés en fonction de trois niveaux. Le niveau 3, le plus sérieux, est réservé aux éléments fréquents qui entraînent un fort taux d'échec à la tâche. Le niveau 2 à ceux qui occasionnent une forte insatisfaction ou des erreurs qui s'accompagnent d'une perte de temps. Les problèmes classés au niveau 1 sont mineurs par leurs conséquences et en général se manifestent peu fréquemment.

Problème 1 : Absence de renvois entre les différentes dénominations des auteurs (niveau 3)



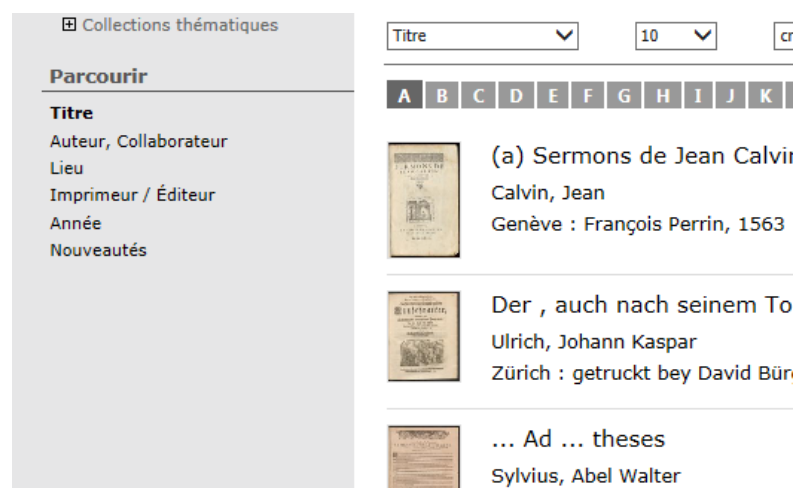
Problème :

Les noms de certains auteurs varient selon les langues. Nous avons vu l'exemple d'Erasmus, plus couramment connu en français sous la forme d'Erasme. Or, lorsque les participants francophones tapent « Erasme » dans la recherche, ils obtiennent deux notices n'ayant rien à voir avec l'auteur Erasme (pour la recherche simple), voire aucun résultat en utilisant la recherche avancée. Il faut donc connaître au préalable la forme sous laquelle l'auteur est indexé dans le site pour avoir une chance de le trouver avec la fonction « recherche », ce qui n'est pas nécessairement le cas de tous les utilisateurs.

Recommandation :

Nous proposons d'instaurer des notices d'autorités, avec des renvois aux formes utilisées.

Problème 2 : Fonctionnalité « parcourir » peu claire avec l’affichage des notices de documents (niveau 3)



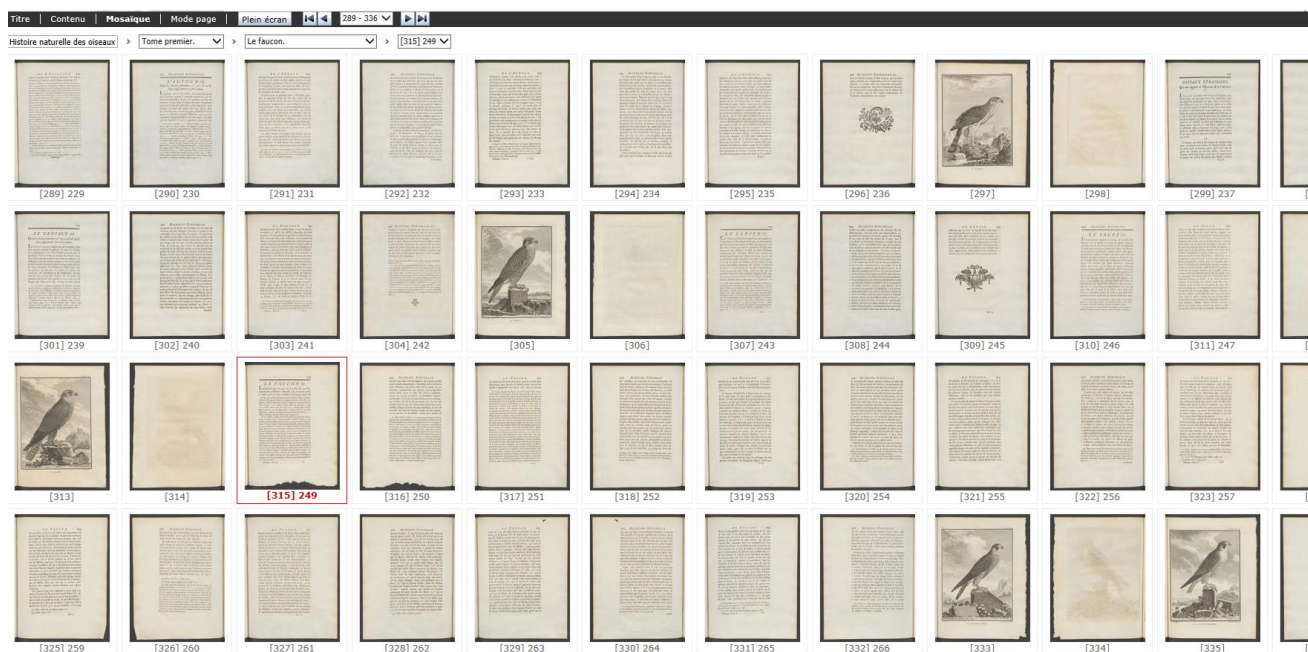
Problème :

Les menus « parcourir » donne accès à toutes les notices listée dans l’ordre du critère de recherche. Cela fournit une masse de notices importante et peu pratique à parcourir lorsque l’on souhaiterait atteindre un titre en particulier. Comme le mentionne le participant 22 du test francophone à distance, « c’est illisible, et très fatigant de repérer dans la notice où se trouve ce qu’on cherche ».

Recommandation :

Nous proposons de présenter à la place une liste alphabétique des titres, des noms d’auteurs, des lieux d’édition, etc.

Problème 3 : Navigation et repérage peu aisés dans les miniatures (niveau 3)



Problème :

Il est difficile de bien se rendre compte d'où l'on est parti dans la fonction mosaïque. Les séparations entre chapitres ne sont pas bien visibles et l'on se retrouve avec plusieurs illustrations sans savoir à quoi elles se rattachent. De plus, lorsque l'on en choisit une et qu'on reprend ensuite la fonction mosaïque, les pages montrées ne sont plus les mêmes. S'il l'on ne s'y attend pas, cela provoque de la confusion supplémentaire et on risque de se perdre.

Recommandation :

Nous proposons d'organiser la fonction mosaïque par chapitre et d'augmenter l'épaisseur du cadre rouge, de manière à ce qu'il soit mieux visible.

Problème 4 : Absence d'outil de recherche pour les images (niveau 2)

The screenshot displays the e-rara.ch search interface. On the left is a sidebar with a 'Collections' section containing 'Toutes les collections', 'Imprimés suisses', and 'Collections thématiques'. Below this is a 'Parcourir' (Browse) section with a list of fields: 'Titre', 'Auteur, Collaborateur', 'Lieu', 'Imprimeur / Éditeur', 'Année', and 'Nouveautés'. The main area on the right contains several search filters. The 'Métadonnées bibliographiques' (Bibliographic metadata) section has a table with columns 'Relation', 'Chercher dans' (Search in), and 'Terme(s)' (Term(s)). It includes four rows for searching by 'et' (and) in 'Tous les champs' (All fields), 'Titre' (Title), 'Auteur, Collaborateur' (Author, Collaborator), and 'Imprimeur / Éditeur' (Printer / Editor). Below this is the 'Recherche dans le contenu' (Search in content) section with a text input field. The 'Troncature' (Truncation) section has radio buttons for 'non' (no) and 'oui' (yes), with 'oui' selected. The 'Filtre' (Filter) section includes three lists: 'Bibliothèques' (Libraries) with 'BPU Neuchâtel', 'Biblioteca Salita dei Frati, Lugano', 'Bibliothèque de Genève', and 'Bibliothèque des Pasteurs, Neuchâtel'; 'Langues' (Languages) with 'Allemand', 'Anglais', 'Français', and 'Latin'; and 'Collections' with 'Imprimés suisses du XVe siècle', 'Imprimés suisses du XVIe siècle', 'Imprimés suisses du XVIIe siècle', and 'Imprimés suisses du XVIIIe siècle'.

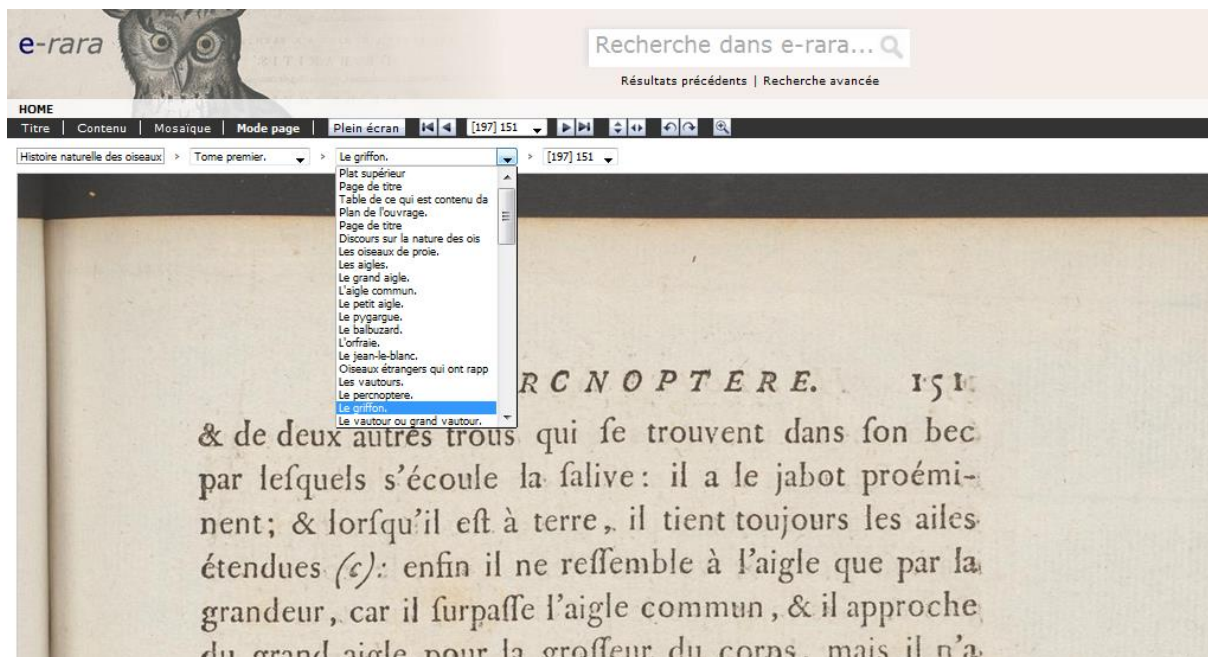
Problème :

Le seul moyen de visionner les images d'un ouvrage sans devoir en parcourir toutes les pages est de passer par la fonction « mosaïque ». Or, cela présuppose d'avoir au préalable une idée des chapitres ou parties du livre où elles se trouvent, ce qui n'est pas nécessairement toujours le cas. De plus, il est impossible de demander à rechercher spécifiquement des images dans l'ensemble de la bibliothèque numérique. Cela peut se révéler problématique pour les utilisateurs.

Recommandation :

Nous proposons de créer une indexation des images et d'offrir la possibilité de les rechercher à partir de la recherche avancée. Comme le mentionne le participant 32 du test francophone, cela « serait un plus pour certains ouvrages, compte tenu de la popularité des planches qu'ils contiennent, comme Buffon ou l'Encyclopédie ».

Problème 5 : Accès difficile à l'information à l'intérieur des documents (niveau 2)



Problème :

Plusieurs de nos participants ont éprouvé de la difficulté à accéder au contenu des documents que nous leur demandions d'explorer, à y naviguer et surtout à s'y repérer. Les menus et index mis en place sont d'une grande aide, mais pas toujours suffisants.

Recommandation :

Nous proposons la mise en place d'une possibilité de recherche plein texte. Cette mesure serait certes coûteuse, mais simplifierait grandement la recherche d'éléments particuliers à l'intérieur des livres.

Problème 6 : Affichage des résultats trop peu clair (pas d'étiquettes comme titre, auteur, etc.) (niveau 2)

206 Titres

Recherche: erasmus

Trier par:

Résultats:

Ordre:


Résultats suivants:

Pertinence


10

croissant

1 - 10



Das regiment der gesuntheit
In dysem biechlin findestu wie sich ein jeglich mensch halten sol mit essen unc
[Basel] : [Pamphilus Gengenbach], 1513



Novum Instrumentum omne
diligenter ab Erasmo Roterodamo recognitum & emendatum, ... una cum Anno
Erasmus, Desiderius
Apud inclytam Germaniae Basileam : [Johannes Froben], [1516]
• Leoni decimo, pontifici modis omnibus summo, Erasmus Roterodamus theo

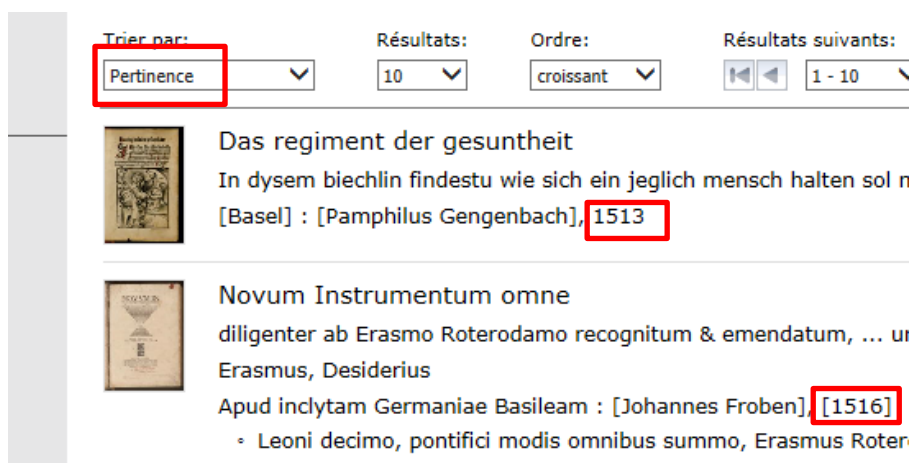
Problème :

Les informations des notices sont présentées sans étiquette spécifiant qu'il s'agit d'un titre, d'un auteur ou d'un éditeur. Nous avons ici l'exemple de la recherche d'Erasmus : une fois trié, le premier résultat de la liste est un ouvrage n'ayant pas été écrit par Erasme, mais en l'absence d'une étiquette indiquant clairement un champ destiné pour le nom d'auteur, ce fait est peu visible. La référence bibliographique en question comporte l'indication « Erasmushaus ». Plusieurs participants ont relevé la date de ce livre comme ouvrage d'Erasme.

Recommandation :

Nous proposons d'ajouter des étiquettes (titre, auteur, etc.) de manière à clarifier la nature des informations des notices.

Problème 7 : Après recherche, option de tri "relevance" (pertinence) toujours affichée automatiquement quand on modifie l'ordre (niveau 2)



Problème :

Lorsque l'on modifie les filtres d'une liste de résultats, par exemple pour trier les livres d'Erasme par année de manière à avoir le plus ancien en tête de liste, le filtre « trier par » se remet automatiquement sous le label « pertinence » (le premier de la liste du menu déroulant). Cela peut entraîner de la confusion chez les utilisateurs, qui ne savent pas si l'opération s'est effectuée correctement ou non.

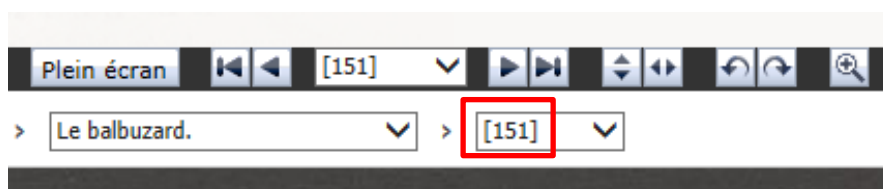
A noter que ce problème ne se présente pas lorsque l'on modifie les filtres des index (index auteurs par exemple).

Recommandation :

Nous recommandons de garder le filtre utilisé dans la fenêtre, de manière à toujours avoir trace de ce qu'on a fait, comme c'est déjà le cas pour les filtres des index.

Problème 8 : Confusion liée à la double pagination (niveau 2)

<i>Le Pygargue.....</i>	99
<i>Le Balbuzard.....</i>	103
<i>L'Orfraie.....</i>	112
<i>Le Jean-le-blanc.....</i>	124
<i>Oiseaux étrangers qui ont rapport aux Aigles & Balbuzards.....</i>	136
<i>Les Vautours.....</i>	146
<i>Le Percnoptère.....</i>	149
<i>Le Griffon.....</i>	151
<i>Le Vautour ou grand Vautour.....</i>	158
<i>Le Vautour à aigrettes.....</i>	159
<i>Le petit Vautour.....</i>	164



Problème :

Le fait qu'il y ait deux types de numéros de page peut induire en erreur, notamment lorsque l'on recherche un chapitre via l'index alphabétique de l'ouvrage. Celui-ci nous donne un numéro de page (151 dans le cas du griffon de notre tâche 4). Il faut alors avoir constaté que cette page correspond au deuxième nombre du menu permettant l'accès aux différentes pages, pour ne pas se retrouver sur une autre sans lien avec ce que l'on avait cherché (l'illustration du balbuzard dans le cas mentionné plus haut, ce qui est arrivé à 4 de nos participants).

Recommandation :

Nous proposons de placer la page utilisée par le système (entre crochets) en deuxième position, de manière à ce que les utilisateurs puissent se référer au premier numéro de page depuis la table alphabétique. Une icône d'information à côté du menu pourrait contribuer à lever l'ambiguïté.

Problème 9 : Fenêtre de recherche peu visible (niveau 2)



Problème :

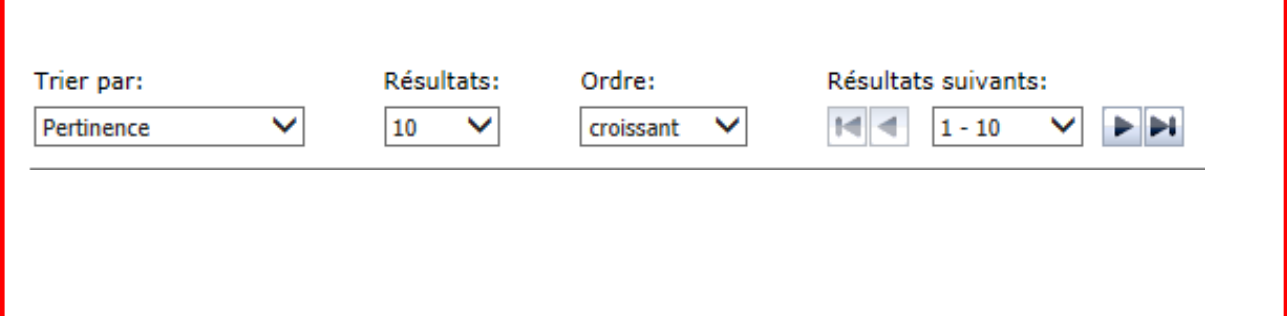
La barre de recherche ne ressort pas assez du bandeau dans lequel elle se trouve. Ainsi, plusieurs de nos participants l'ont cherchée sans la trouver immédiatement, et parfois même sans la trouver du tout. Par exemple, le participant 6 du test germanophone à distance : « Die Suchbox "Suche in e-rara" auf der Startseite ist zwar einerseits gross und prominent, andererseits aber farblich sehr dezent in den Seitenkopf eingebettet und befindet sich an einer Position, an der man auf Websites selten eine Suchbox antrifft (eher rechts oben oder am linken Rand), was dazu führte, dass ich sie zunächst ganz übersehen hatte. Erst nach mehreren Nutzungen der Website habe ich gemerkt "oh, da oben ist ja eine Suchbox" ».

Cela peut donc se révéler très problématique.

Recommandation :

Nous recommandons d'augmenter le contraste de la barre de recherche, par exemple en ajoutant un cadre noir autour, de manière à ce qu'elle ressorte mieux.

Problème 10 : Interaction entre les différents menus des outils de tri peu clair (niveau 2)



Problème :

Il n'est pas évident au premier coup d'œil de comprendre que ces trois fenêtres de filtres sont en fait liées et que le critère « croissant » (et « décroissant ») opérera en fonction du critère choisi dans le filtre « trier par ». Ainsi, l'outil de navigation « résultats suivants » n'est pas directement dépendant de ces trois filtres, mais les écarts entre ces différentes fenêtres n'est pas significativement plus grand.

Recommandation :

Nous proposons de rajouter un cadre englobant les trois menus interdépendants et d'éloigner l'outil « résultats suivants » en le décalant un peu vers la droite.

Problème 11 : Manque d'aide générale (niveau 2)

The screenshot shows the e-rara.ch website interface. At the top, there's a navigation bar with 'HOME' and links for 'Titre', 'Contenu', 'Mosaïque', 'Mode page', and 'Plein écran'. Below this, a breadcrumb trail reads 'Histoire naturelle des oiseaux > Tome premier. > Le faucon.' with a page indicator '[315] 249'. The main content area shows '920 titres' and search filters: 'Trier par: Auteur, Collaborateur', 'Résultats: 10', 'Ordre: croissant', and 'Résultats suivants: 1-10'. A alphabetical index bar is visible. A small thumbnail of a book cover is shown next to the text: 'Iacobi Acontii Tridentini de methodo, hoc est, de recta investigandarum, tradendarum nunc iterum editus liber.' On the left, a sidebar titled 'Collections' lists 'Toutes les collections', 'Imprimés suisses', and 'Collections thématiques'. Below it, 'Parcourir' lists 'Titre', 'Auteur, Collaborateur', 'Lieu', 'Imprimeur / Éditeur', 'Année', and 'Nouveautés'. On the right, a 'Affiner la recherche' sidebar shows filters for 'Auteurs, Collaborateurs', 'Périodes', 'Langues', 'Lieux', 'Imprimeurs / Éditeurs', and 'Bibliothèques' with corresponding counts.

Auteurs, Collaborateurs	Count
Erasmus, Desiderius	153
Estienne, Henri	103
Euler, Leonhard	57
Egli, Raphael	31
Erni, Heinrich	26

Afficher 45 autres résultats

Périodes	Count
1401-1500	5
1501-1600	459
1601-1700	201
1701-1800	142
1801-1900	110
1901-2000	3

Afficher 6 autres résultats

Langues	Count
Latin	581
Allemand	234
Grec ancien	103
Français	76
Anglais	14

Afficher 6 autres résultats

Lieux	Count
Basel	283
Zürich	207
Genève	153
Paris	40
Berlin	25

Afficher 45 autres résultats

Imprimeurs / Éditeurs	Count
Estienne, Henri II	62
[s.n.]	44
Froben, Johannes (Offizin, Basel)	42
Froschauer (Offizin, Zürich)	41
Wolf (Offizin, Zürich)	29

Afficher 45 autres résultats

Bibliothèques	Count
UB Basel	267
ETH-Bibliothek Zürich	256
ZB Zürich	202
Bibliothèque de Genève	145
UB Bern	17

Afficher 7 autres résultats

Problème :

Le site est doté de nombreuses fonctionnalités, tant au niveau des possibilités de recherche (fonction « recherche », index, facettes, filtres, etc.) que de navigation dans les ouvrages mis à disposition (accès direct au contenu, accès à un sommaire, fonction « mosaïque », menus déroulants pour accéder à un chapitre ou une page particulière, etc.). Elles ne sont toutefois répertoriées nulle part. Beaucoup de nos participants auraient voulu utiliser une fonction qui existait sur le site, mais qu'ils ne sont pas parvenu à trouver.

Ainsi, plusieurs de nos participants ont feuilleté l'entier du chapitre du faucon pour arriver aux illustrations de fin de chapitre, tout en étant convaincu qu'il n'y avait pas d'autre moyen d'y parvenir, alors qu'il existe une fonction « mosaïque » permettant d'avoir une vue d'ensemble.

Recommandation :

Nous proposons la création d'une rubrique « aide », répertoriant les fonctionnalités proposées, des marches à suivre pour la réalisation de certaines actions clés de la recherche, peut-être sous forme de tutoriels vidéo.

Problème 12 : Manque de visibilité du caractère cliquable de l'image du livre des notices (niveau 2)



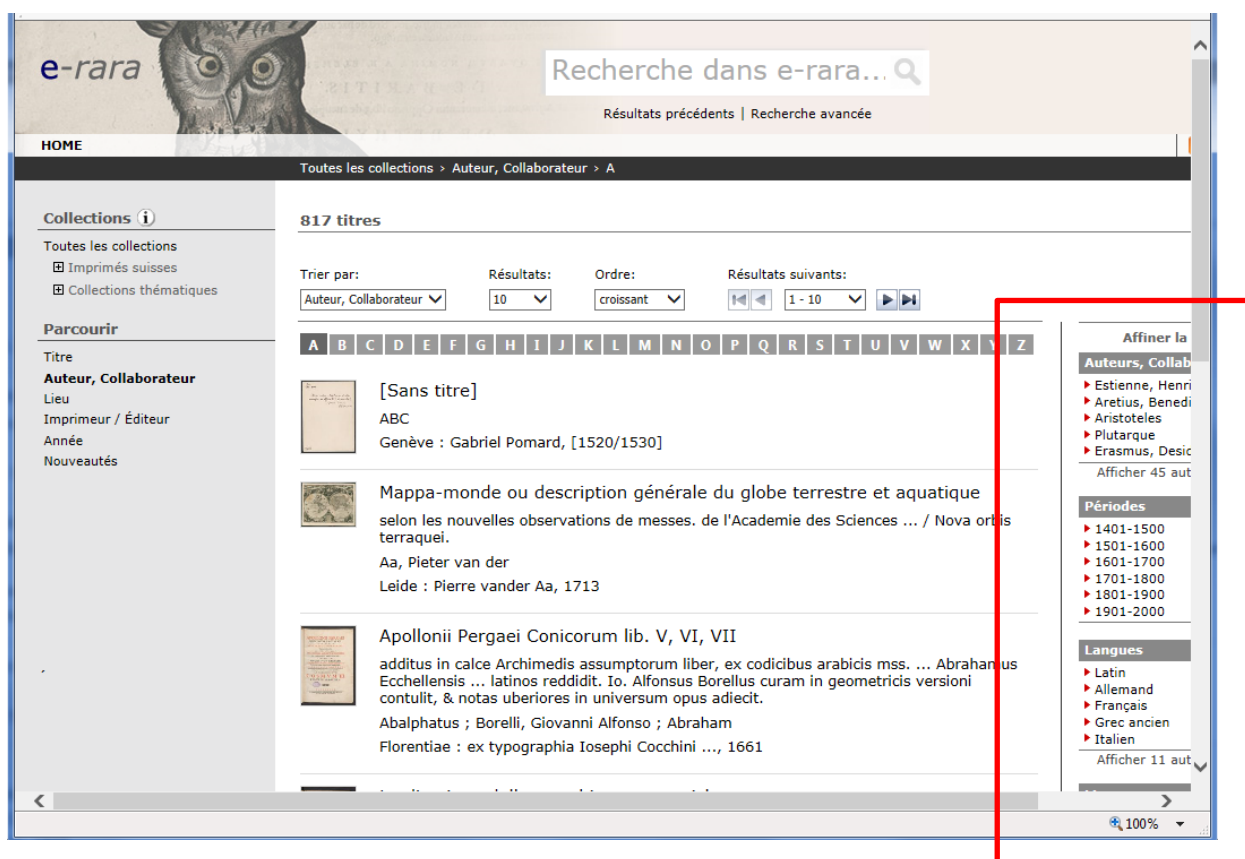
Problème :

Il n'est pas à première vue évident que l'image de l'ouvrage permet d'accéder à son contenu. La majorité des participants ont utilisé un autre moyen d'accès pour parcourir le livre. Une de nos participantes au test modéré a beaucoup hésité avant de tester cette fonction.

Recommandation :

Nous proposons de rajouter une icône visible sur le livre, peut-être dans un coin avec une infobulle invitant à cliquer pour accéder au contenu.

Problème 13 : Scroll latéral fréquent sur des écrans étroits (niveau 2)



Problème :

Sur de petits écrans d'ordinateurs, les facettes sont coupées, voire carrément invisibles. Il faut donc scroller latéralement pour les obtenir. Cela se révèle très gênant à l'usage. De plus, pour un utilisateur effectuant sa première visite, il risque de manquer complètement cette fonctionnalité si elle est invisible. « Da ich auf einem kleinen Bildschirm suchte, hatte zum Beispiel den rss-feed nicht im Blickfeld » (participant 12 du test à distance en allemand).

Recommandation :

Nous recommandons d'améliorer le dynamisme des pages de manière à pouvoir visionner les facettes et autres fonctionnalités même avec de petits écrans.

Problème 14 : Intitulé « provenance du document original » trop peu clair (niveau 1)



Notice détaillée

Titre	Histoire naturelle des oiseaux / [par M. de Buffon et M. Gueneau de Montbeillard]
Auteur, Collaborateur	Buffon, Georges Louis Leclerc de ; Montbeillard, Gueneau de ; Sève, Jacques Eustache de ; Bex
Collectivité	Imprimerie royale (Paris)
Adresse bibliographique	A Paris : de l'imprimerie royale, 1770-1783
Collation	9 vol. : ill. ; 4° (27 cm)
Langue	Français
Référence bibliographique	Ronsil, R.: Bibliogr. ornithol. franç., no 413, p. 76
Provenance du document original	BPU Neuchâtel, ZR 200/16-24
Série	Histoire naturelle, générale et particulière, avec la description du cabinet du Roi 16-24
Lien permanent	http://dx.doi.org/10.3931/e-rara-7296

Volumes

[Tome premier.](#)

[Tome second.](#)

[Tome troisième.](#)

Auteur	Buffon, Georges Louis Leclerc de
Titre	Histoire naturelle des oiseaux / [par M. de Buffon et M. Gueneau de Montbeillard]
Lieu / Date	A Paris : de l'imprimerie royale, 1770-1783
Collection	Histoire naturelle, générale et particulière, avec la description du cabinet du Roi / [Buffon] ; t. 24
Accès	Texte intégral: http://www.e-rara.ch/evn_r/content/titlenfo/1894161
Collation	9 vol. : ill. ; 4° (27 cm)

Demandes en cours 0 Demandes totales 0

Exemplaires 1 - 1 / 1

1	<input type="checkbox"/> Sélectionner Cet exemplaire ne peut pas être demandé à distance. Voir le statut ci-dessous. Cote BPUN ZR 200/16-24 Dépôt NE BPUN Réserve Statut A consulter sur place Détails de l'exemplaire 031056768
---	---

[Ajouter sélection au panier](#)

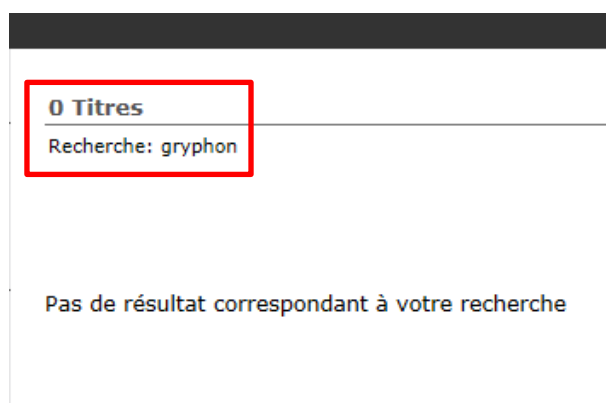
Problème :

Un de nos participants a voulu feuilleter l'ouvrage via ce lien, pensant qu'il le mènerait au document original. Le fait d'utiliser ces termes dans l'intitulé peut porter à confusion.

Recommandation :

Nous proposons de remplacer l'intitulé par « Provenance du document physique ».

Problème 15 : Limitation dans la recherche des variantes orthographiques au sein des équations de recherche (niveau 1)



Problème :

Le moteur de recherche ne permet pas facilement la prise en charge d'une équation parcourant un terme sous ses différentes variantes orthographiques via la troncature interne ou des opérateurs booléens.

« Es ist unbefriedigend, dass Platzhalter nicht verwendet und dabei an beliebigen Orten gesetzt werden können; Beispiel e?dgenossenschaft für eidgenossenschaft bzw. eydgenossenschaft » (participant 31 du test en allemand à distance).

Recommandation :

Si le problème 2 est résolu et qu'il devient possible d'afficher par exemple une liste d'auteurs classé alphabétiquement, il sera déjà plus facile de retrouver une information même si un nom se présente sous plusieurs formes. Nous recommandons sinon de permettre la troncature interne et d'autoriser l'usage de booléens accompagnés de parenthèses.

Problème 16 : Menu horizontal (titre, contenu, etc.) peu visible (niveau 1)



Problème :

Les menus, fonctionnalités et filtres situés au-dessus des notices et des ouvrages ne sont pas vus par tous les participants. Certains partent du principe que tout ce qui est au-dessus fait partie de la zone du bandeau et ne comporte donc pas de fonctions. D'autres ne se sont pas aperçus qu'il y avait des boutons cliquables à ces endroits-là.

Recommandation :

Nous proposons de les agrandir, d'augmenter le contraste au niveau des couleurs et peut-être de les espacer un peu plus, de manière à ce qu'ils ressortent mieux.

Problème 17 : Notion de «lien permanent» peu claire (niveau 1)

Titre | Contenu | Mosaïque | Mode page | **Plein écran**



Notice détaillée

Titre	Histoire naturelle des oiseaux / [par M. de Buffon et M. Gueneau de Montbeillard]
Auteur, Collaborateur	Buffon, Georges Louis Leclerc de  ; Montbeillard, Gueneau de  ; Sève, Jacques Eustache de  ; Bexon, Gab
Collectivité	Imprimerie royale (Paris) 
Adresse bibliographique	A Paris : de l'Imprimerie royale. 1770-1783
Collation	9 vol. : ill. ; 4° (27 cm)
Langue	Français
Référence bibliographique	Ronsil, R.: Bibliogr. ornithol. franç., no 413, p. 76
Provenance du document original	BPU Neuchâtel, ZR 200/16-24 
Série	Histoire naturelle, générale et particulière, avec la description du cabinet du Roi 16-24
Lien permanent	http://dx.doi.org/10.3931/e-rara-7296 

Volumes

[Tome premier.](#)

[Tome second.](#)



Problème :

Plusieurs participants ont pensé que le « lien permanent » menait au contenu de l'ouvrage et ont tenté ce moyen pour accéder au contenu. D'autres l'ont utilisé en s'interrogeant sur sa nature.

Recommandation :

Nous proposons de reformuler ce lien en « Lien permanent pour accéder à cette page ».

Problème 18 : Pas de possibilité de classer alphabétiquement les facettes disponibles (niveau 1)

Affiner la recherche

Auteurs, Collaborateurs	
▶ Erasmus, Desiderius	153
▶ Jud, Leo	10
▶ Estienne, Henri	9
▶ Gwalther, Rudolf	6
▶ Reinhold, Erasmus	6
▶ Cicero, Marcus Tullius	5
▶ Isocrates	4
▶ Johannes	4
▶ Rhenanus, Beatus	4
▶ Terentius Afer, Publius	4
▶ Agricola, Georgius	3
▶ Alcman	3
▶ Alcée	3
▶ Anacréon	3
▶ Bacchylide	3
▶ Bibliander, Theodor	3
▶ Buchanan, George	3
▶ Castellio, Sebastian	3
▶ Cato, Dionysius	3
▶ Cato, Marcus Porcius Censorius	3
▶ Céos, Simonide de	3
▶ Eberlin von Günzburg, Johann	3
▶ Gesner, Conrad	3
▶ Ibycos	3
▶ Kolin, Peter	3
▶ Linacre, Thomas	3
▶ Negri, Stefano	3
▶ Pellikan, Konrad	3
▶ Pindare	3
▶ Pirckheimer, Willibald	3
▶ Poliziano, Angelo	3
▶ Publilius Syrus	3
▶ Rivius, Johannes	3
▶ Sappho	3
▶ Spauter, Johannes de	3
▶ Stésichore	3
▶ Valla, Laurentius	3
▶ Valla, Lorenzo	3
▶ Vives, Juan Luis	3
▶ Agricola, Rudolf	2
▶ Alphée de Mytilène	2
▶ André, Élie	2
▶ Anysius, Janus	2
▶ Archiloque de Paros	2
▶ Aristoteles	2
▶ Ausone	2
▶ Ausonius, Decimus Magnus	2
▶ Bertulphus, Hilarius	2
▶ Caesar, Carl	2
▶ Camerarius, Joachim	2

Afficher les 5 premiers résultats

Affiner la recherche

Auteurs, Collaborateurs	
▶ Erasmus, Desiderius	153
▶ Jud, Leo	10
▶ Estienne, Henri	9
▶ Gwalther, Rudolf	6
▶ Reinhold, Erasmus	6

Afficher 45 autres résultats

Problème :

Les facettes donnent leurs énoncés en fonction du nombre d'occurrence dans la liste des résultats (de la recherche ou de l'index). Or, l'on ne cherche pas nécessairement le plus fréquemment cité et, lorsqu'il y a 45 termes dans la liste, il est parfois laborieux de tous les parcourir. Ce problème se poserait moins si le problème 2 était résolu et qu'il devenait possible d'obtenir une liste d'auteurs classés alphabétiquement par exemple.

Recommandation :

Nous proposons d'ajouter la possibilité de classer les termes de la facette par ordre alphabétique ou du moins de résoudre le problème 2.

Problème 19 : Position du lien vers les fichiers pdf trop à droite (niveau 1)

Recherche dans e-rara... Q

Résultats précédents | Recherche avancée

HOME

Titre | Contenu | Mosaïque | Mode page | Plein écran

Histoire naturelle des oiseaux / [par M. de Buffon et M. Gueneau de Montbeillard]. A Paris : de l'Imprimerie royale, 1770-1783. Tome premier.

Contenu du document

Plat supérieur	PDF
Page de titre	PDF
Table de ce qui est contenu dans ce volume.	PDF
Plan de l'ouvrage.	1 PDF
Page de titre	PDF
Discours sur la nature des oiseaux.	PDF
Les oiseaux de proie.	61 PDF
Les aigles.	71 PDF
Le grand aigle.	76 PDF
L'aigle commun.	86 PDF
Le petit aigle.	91 PDF
Le pygargue.	99 PDF
Le balbuzard.	103 PDF
L'orfraie.	112 PDF
Le Jean-le-blanc.	124 PDF
Oiseaux étrangers qui ont rapport aux aigles & balbuzards.	136 PDF
Les vautours.	146 PDF
Le percnoptère.	149 PDF
Le griffon.	151 PDF
Le vautour ou grand vautour.	158 PDF
Le vautour à aigrettes.	159 PDF
Le petit vautour.	164 PDF
Oiseaux étrangers qui ont rapport aux vautours.	167 PDF
Le condor.	184 PDF
Le milan et les buses.	197 PDF
La buse.	206 PDF
La bondrée.	208 PDF
L'oiseau saint-martin.	212 PDF
La soubuse.	215 PDF
La harpye.	217 PDF
Le busard.	218 PDF
Oiseaux étrangers, qui ont rapport au milan, aux buses & soubuses.	221 PDF

Problème :

Lorsque l'on travaille sur un écran relativement grand, les pdf (ainsi que les numéros de pages des différents chapitres) alignés à droite se retrouve loin de la liste des chapitres, et donc de la zone où se porte principalement le regard. Ils risquent ainsi de passer inaperçus.

Recommandation :

Nous recommandons de ne pas les aligner à droite, mais de les rapprocher sans les coller. Cela éviterait également le risque qu'ils n'apparaissent pas sur écran trop petit.

Problème 20 : Recherche avancée impossible sans entrer un terme (niveau 1)

Métadonnées bibliographiques

Relation	Chercher dans	Terme(s) recherché(s)
-	Tous les champs	
et	Titre	
et	Auteur, Collaborateur	
et	Imprimeur / Éditeur	

Recherche dans le contenu

Troncature

☐ non ☒ oui

Filtre

Bibliothèques

BPU Neuchâtel
Biblioteca Salita dei Frati, Lugano
Bibliothèque de Genève
Bibliothèque des Pasteurs, Neuchâtel

Langues

...

Sélecti

Message de la page Web

⚠ S'il vous plaît entrer un mot.

OK

Problème :

Il n'est pas possible de lancer une recherche avancée en sélectionnant une bibliothèque ou un autre filtre sans entrer de terme de recherche.

Recommandation :

Nous proposons de permettre ce type de recherche.

Problème 21 : Recherche difficile dans la liste des contenus (niveau 1)

Contenu du document	
✿ Tome premier.	✿ Le merle de roche.
✿ Plat supérieur	✿ Le merle bleu.
✿ Page de titre	✿ Le merle solitaire.
✿ Table de ce qui est contenu dans ce volume.	✿ Oiseaux étrangers qui ont rapport au merle solitaire.
✿ Plan de l'ouvrage.	✿ Oiseaux étrangers qui ont rapport aux merles d'Europe.
✿ Page de titre	✿ Le griséin de cayenne.
✿ Discours sur la nature des oiseaux.	✿ Le verdin de la Cochinchine.
✿ Les oiseaux de proie.	✿ L'azurin.
✿ Les aigles.	✿ Les breves.
✿ Le grand aigle.	✿ Le mainate des Indes orientales.
✿ L'aigle commun.	✿ Variétés du mainate.
✿ Le petit aigle.	✿ Le goulin.
✿ Le pygargue.	✿ Le martin.
✿ Le balbuzard.	✿ Le jaseur.
✿ L'orfraie.	✿ Variété du jaseur.
✿ Le jean-le-blanc.	✿ Le gros-bec.
✿ Oiseaux étrangers qui ont rapport aux aigles & balbuzards.	✿ Le bec-croisé.
✿ Les vautours.	✿ Oiseaux étrangers qui ont rapport au gros-bec.
✿ Le percnoptère.	✿ Le moineau.
✿ Le griffon.	✿ Oiseaux étrangers qui ont rapport au moineau.
✿ Le vautour ou grand vautour.	✿ Le friquet.
✿ Le vautour à aigrettes.	✿ Oiseaux étrangers qui ont rapport au friquet.
✿ Le petit vautour.	✿ La soulcie.
✿ Oiseaux étrangers qui ont rapport aux vautours.	✿ Oiseaux étrangers qui ont rapport à la soulcie.
✿ Le condor.	✿ Table des matières.
✿ Le milan et les buses.	✿ Fautes à corriger dans ce volume.
✿ La buse.	✿ Plat inférieur
✿ La bondrée.	✿ Tome quatrième.
✿ L'oiseau saint-martin.	✿ Plat supérieur
✿ La soubuse.	✿ Page de titre
✿ La harpaye.	✿ Table de ce qui est contenu dans ce volume.
✿ Le busard.	✿ Le serin des Canaries.
✿ Oiseaux étrangers, qui ont rapport au milan, aux buses & soubuses.	✿ Oiseaux étrangers qui ont rapport aux serins.
✿ L'épervier.	✿ L'habesch de Syrie.
✿ L'autour.	✿ La linotte.
✿ Oiseaux étrangers, qui ont rapport à l'épervier & l'autour.	✿ Variétés de la linotte.
✿ Le gerfaut.	✿ Oiseaux étrangers qui ont rapport à la linotte.
✿ Le lanier.	✿ Le ministre.
✿ Le sacre.	✿ Les bengalis et les sénégalis, &c.
✿ Le faucon.	✿ Le maia.
✿ Oiseaux étrangers, qui ont rapport au gerfaut & aux faucons.	✿ Le maian.
	✿ Le pinson.
	✿ Variétés du pinson.

Problème :

La liste des chapitres d'un ouvrage peut se révéler très longue, particulièrement lorsqu'il s'agit d'un ouvrage en plusieurs volumes. Dans ce cas de figure, il peut être difficile de repérer le chapitre qui nous intéresse.

Recommandation :

Nous recommandons de permettre une recherche plein texte sinon dans l'ouvrage du moins dans la page de ses contenus en aménageant une fenêtre de recherche équivalente à la fonction « ctrl + f ».

Problème 22 : Termes « croissant » et « décroissant » indifféremment appliqué à des chiffres qu'à des lettres (niveau 1)

The image shows two identical sorting controls. Each control has four main sections: 'Trier par:', 'Résultats:', 'Ordre:', and 'Résultats suivants:'. In the top control, 'Trier par:' is set to 'Auteur, Collaborateur' and 'Ordre:' is set to 'croissant'. In the bottom control, 'Trier par:' is set to 'Année' and 'Ordre:' is set to 'croissant'. Red boxes highlight the 'Trier par:' and 'Ordre:' dropdowns in both controls. The 'Résultats:' dropdown is set to '10' and 'Résultats suivants:' is set to '1 - 10' in both. Below each control is a horizontal bar with letters A through Y.

Problème :

La notion d'ordre croissant / décroissant s'applique généralement à des chiffres, aussi l'appliquer à des lettres (en triant par auteur, titre ou encore lieu) a surpris un de nos participants. Il n'a donc pas compris que le filtre ordre était lié à celui du triage par auteur et à quel critère l'ordre croissant se référait. Ce problème serait moins sensible si le lien entre les différents outils de tri était davantage visible (problème 10).

Recommandation :

Nous proposons de remplacer les termes « croissant » et « décroissant » par « ascendant » et « descendant » qui peuvent s'appliquer autant aux lettres qu'aux chiffres. A noter que la version allemande utilise justement les termes « aufsteigend » et « absteigend ».

Problème 23 : Troncature « * » non acceptée (niveau 1)

The screenshot shows the e-rara search interface. At the top, there is a search bar with the text "Recherche dans e-rara..." and a magnifying glass icon. Below the search bar, there are links for "Résultats précédents" and "Recherche avancée". On the right side, there is a red box containing an information icon and the text "INFORMATION [AFFICHER]". Below this, there is a section for "Métadonnées bibliographiques" with a table for searching by relation, field, and term. The table has columns for "Relation", "Chercher dans", and "Terme(s) recherché(s)". The first row shows "et" for "Titre" with the term "erasm*". The second row shows "et" for "Auteur, Collaborateur" with the term "erasm*". The third row shows "et" for "Imprimeur / Éditeur" with an empty term. Below the table, there is a "Recherche dans le contenu" section with an empty search bar. At the bottom, there is a "Troncature" section with a checkbox. The search results section shows "0 Titres" and "Recherche: Auteur, Collaborateur = erasm*". Below this, it says "Pas de résultat correspondant à votre recherche". On the right side of the results, there is a red box containing three rows of a dropdown menu with a chevron icon and the text "erasm*".

Problème :

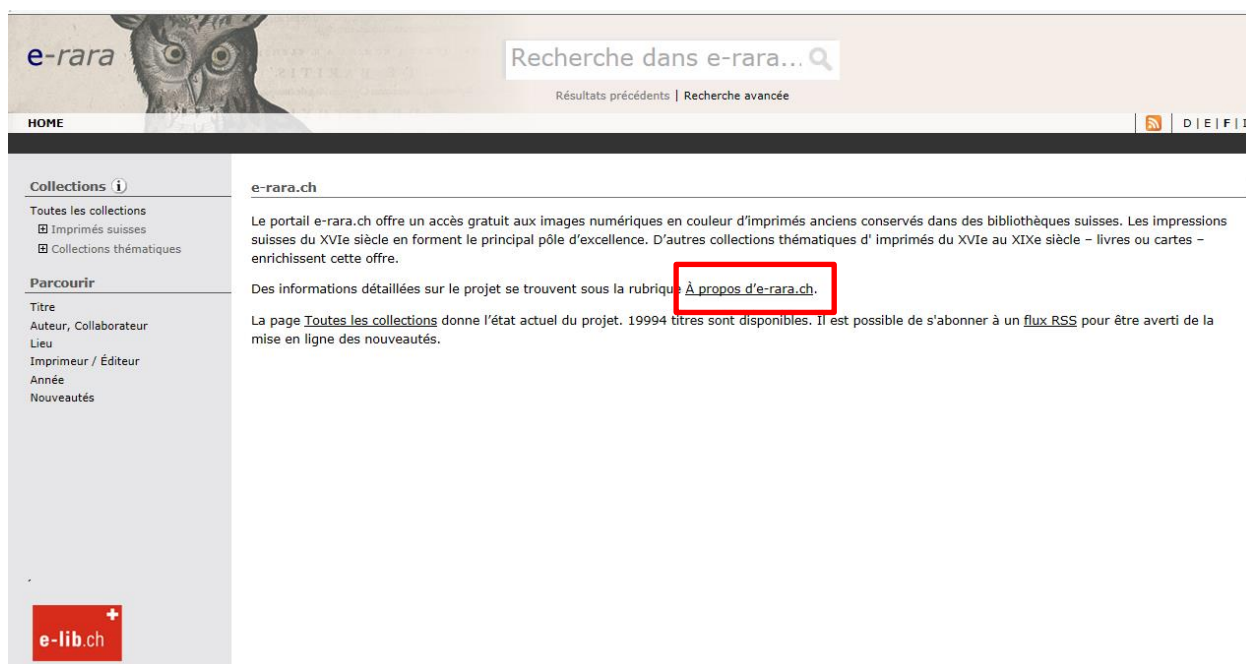
Dans de nombreux sites et moteurs de recherche, la troncature se symbolise par l'astérisque « * ». Or, ce n'est pas le cas ici. Cela a posé problème à quelques-uns de nos participants, qui n'ont ainsi obtenus aucun résultat à leur recherche.

Recommandation :

Nous proposons d'agrandir et de rapprocher l'icône « information » placée à droite du formulaire de recherche avancée, et d'afficher ces informations par défaut. L'utilisateur peut cliquer pour les fermer s'il le souhaite. Cela l'aiderait à se rendre compte que cette aide est placée là.

En outre, il serait judicieux de faire en sorte que la recherche ignore purement et simplement le symbole, puisqu'il effectue automatiquement la troncature à droite sauf si on l'en empêche via une option.

Problème 24 : Visibilité des bibliothèques participantes insuffisante (niveau 1)



Problème :

Pour découvrir les bibliothèques participant au projet (et donc mettant leurs documents à disposition), la page d'accueil offre plusieurs moyens. Un lien « à propos d'e-rara.ch » permet d'accéder à une présentation du site, mais en bas à droite il y a également un lien « bibliothèques participantes ». Ces deux moyens ne sont pas extrêmement visibles.

Recommandation :

Nous proposons d'ajouter la liste bien en vue sur la page d'accueil avec éventuellement des illustrations. La page d'accueil actuelle propose surtout du texte et n'est pas très intéressante pour le regard. Cet espace pourrait être exploité pour mettre en évidence davantage les parties prenantes du projet.

Problème 25 : Visibilité des liens hypertextes (niveau 1)

Kollektionen ⓘ

Alle Kollektionen

- ☑ Schweizer Drucke
- ☑ Thematische Kollektionen


Browsing

- Titel
- Autor, Beteiligte
- Druckort
- Drucker / Verlag
- Jahr

206 Titel

Suche nach: Erasmus

Sortieren nach: Relevanz Max. Trefferanzahl: 10 Reihenfolge: aufsteigend Blät

 Tomus secundus continens Paraphrasim D. Erasmi ipso autore recognitam, emendatanque

Erasmus, Desiderius

[Basileae] : [in officina Frobeniana], [1532]

7 Treffer im Inhaltsverzeichnis

Non cliquable Cliquable

Problème :

Les liens hypertextes ne sont marqués que par le fait qu'ils sont soulignés quand on passe la souris dessus. Cela ne les fait pas suffisamment ressortir et ne renseigne pas immédiatement sur leur nature de lien.

Recommandation :

Nous recommandons de signaler les liens hypertextes non seulement par un souligné, mais également par une couleur différente. Cela permettrait de voir plus aisément ce qui est cliquable ou non sur la page.

Annexe 4 : Comparaison des méthodes en fonction des problèmes d'utilisabilité

Problème 1 : absence de renvois entre les différentes dénominations des auteurs (niveau 3)

Méthode de test automatisée	Méthode de test modérée
Capacité au repérage <ol style="list-style-type: none"> 1. Problème passé inaperçu 2. Problème soupçonné 3. Problème identifié ✓ 	<ol style="list-style-type: none"> 1. Problème passé inaperçu 2. Problème soupçonné 3. Problème identifié ✓
Probabilité de repérage <ol style="list-style-type: none"> 1. Repérage ou calcul problématique 2. Repérage (peu) probable 3. - 4. Repérage très probable ✓ 	<ol style="list-style-type: none"> 1. Repérage ou calcul problématique 2. Repérage peu probable 3. Repérage probable 4. Repérage très probable ✓
Efficience du repérage <ol style="list-style-type: none"> 1. Repérage requérant une manipulation particulière des données 2. Repérage ne requérant pas de manipulation particulière des données ✓ 	<ol style="list-style-type: none"> 1. Repérage requérant une manipulation particulière des données 2. Repérage ne requérant pas de manipulation particulière des données ✓
Evaluation de la gravité du problème <ul style="list-style-type: none"> • Fréquence du problème ✓ • Niveau d'insatisfaction soulevé • Taux d'échec de la tâche causé ✓ • Temps perdu ou effort nécessaire pour dépasser le problème ✓ 	<ul style="list-style-type: none"> • Fréquence du problème • Niveau d'insatisfaction soulevé ✓ (part. fr. 4 a par exemple directement stoppé la tâche) • Taux d'échec de la tâche causé • Temps perdu ou effort nécessaire pour dépasser le problème ✓
Facilitation du design de la solution <ol style="list-style-type: none"> 1. Le problème est repéré mais avec peu d'indice sur la manière de le résoudre ✓ 2. Une solution est plus ou moins induite 	<ol style="list-style-type: none"> 1. Le problème est repéré mais avec peu d'indice sur la manière de le résoudre ✓ 2. Une solution est plus ou moins induite
Score : 13	Score : 12

Commentaires : pour ce problème, la méthode automatisée est très avantageuse parce que la mesure à prendre pour résoudre ce problème est relativement complexe : l'ajout de notices d'autorité. Il convient donc d'évaluer en détail l'impact négatif d'un tel problème afin de décider en bonne connaissance de cause de le corriger ou non.

Problème 2 : Fonctionnalité « parcourir » peu claire avec l’affichage des notices de documents (niveau 3)

Méthode de test automatisée	Méthode de test modérée
Capacité au repérage <ol style="list-style-type: none"> 1. Problème passé inaperçu 2. Problème soupçonné 3. Problème identifié ✓ 	<ol style="list-style-type: none"> 1. Problème passé inaperçu 2. Problème soupçonné 3. Problème identifié ✓
Probabilité de repérage <ol style="list-style-type: none"> 1. Repérage ou calcul problématique 2. Repérage (peu) probable 3. - 4. Repérage très probable ✓ <p><i>La première piste pour ce problème a été donnée par un commentaire. Sans cela, cerner précisément le problème aurait demandé un effort supplémentaire.</i></p>	<ol style="list-style-type: none"> 1. Repérage ou calcul problématique 2. Repérage peu probable 3. Repérage probable ✓ 4. Repérage très probable <p><i>Dans le cas présent, la probabilité dépend du soin avec lequel l’analyse est menée. Le risque de passer outre est assez élevé.</i></p>
Efficience du repérage <ol style="list-style-type: none"> 1. Repérage requérant une manipulation particulière des données ✓ 2. Repérage ne requérant pas de manipulation particulière des données <p><i>Ce problème est à l’origine de plusieurs autres (impression de manque de clarté, demande d’aide, etc.) et ramener toutes ces difficultés à celle-ci qui en constitue la source a été complexe.</i></p>	<ol style="list-style-type: none"> 1. Repérage requérant une manipulation particulière des données ✓ 2. Repérage ne requérant pas de manipulation particulière des données <p><i>Ce n’est qu’à la fin de l’analyse des données en constatant que plusieurs problèmes se recoupaient que nous avons pu identifier celui-ci.</i></p>
Evaluation de la gravité du problème <ul style="list-style-type: none"> • Fréquence du problème ✓ • Niveau d’insatisfaction soulevé • Taux d’échec de la tâche causé ✓ • Temps perdu ou effort nécessaire pour dépasser le problème ✓ 	<ul style="list-style-type: none"> • Fréquence du problème • Niveau d’insatisfaction soulevé ✓ <i>(part. fr. 3 a par exemple fortement exprimé son agacement de ne pas comprendre la logique)</i> • Taux d’échec de la tâche causé • Temps perdu ou effort nécessaire pour dépasser le problème ✓
Facilitation du design de la solution <ol style="list-style-type: none"> 1. Le problème est repéré mais avec peu d’indice sur la manière de le résoudre 2. Une solution est plus ou moins induite ✓ 	<ol style="list-style-type: none"> 1. Le problème est repéré mais avec peu d’indice sur la manière de le résoudre 2. Une solution est plus ou moins induite ✓
Score : 13	Score : 11

Commentaires : ce problème a été très complexe à détecter. Une des réponses à la question ouverte du test automatisé le signalait tel quel, mais il s’est avéré que ce point était à l’origine de plusieurs autres difficultés, comme le fait de ne pas toujours comprendre la logique des outils de tri. Pour le test automatisé, aucune mention directe de ce problème n’a

été faite par les participants. Ce n'est qu'en observant que plusieurs problèmes se recoupaient en réalité, qu'il a été possible de remonter à celui-ci.

Problème 3 : Navigation peu aisée dans les miniatures (niveau 3)

Méthode de test automatisée	Méthode de test modérée
Capacité au repérage 1. Problème passé inaperçu 2. Problème soupçonné 3. Problème identifié ✓	1. Problème passé inaperçu 2. Problème soupçonné 3. Problème identifié ✓
Probabilité de repérage 1. Repérage ou calcul problématique 2. Repérage (peu) probable 3. - 4. Repérage très probable ✓	1. Repérage ou calcul problématique 2. Repérage peu probable 3. Repérage probable 4. Repérage très probable ✓
Efficience du repérage 1. Repérage requérant une manipulation particulière des données 2. Repérage ne requérant pas de manipulation particulière des données ✓	1. Repérage requérant une manipulation particulière des données 2. Repérage ne requérant pas de manipulation particulière des données ✓
Evaluation de la gravité du problème <ul style="list-style-type: none"> • Fréquence du problème ✓ • Niveau d'insatisfaction soulevé • Taux d'échec de la tâche causé ✓ • Temps perdu ou effort nécessaire pour dépasser le problème ✓ 	<ul style="list-style-type: none"> • Fréquence du problème • Niveau d'insatisfaction soulevé ✓ (en général, les participants ne voyaient pas leurs difficultés comme un problème du système) • Taux d'échec de la tâche causé • Temps perdu ou effort nécessaire pour dépasser le problème ✓
Facilitation du design de la solution 1. Le problème est repéré mais avec peu d'indice sur la manière de le résoudre ✓ 2. Une solution est plus ou moins induite	1. Le problème est repéré mais avec peu d'indice sur la manière de le résoudre ✓ 2. Une solution est plus ou moins induite
Score : 13	Score : 13

Commentaires : le problème est très fréquent et ne pose pas de problème à être découvert. A noter qu'il comporte plusieurs parties : visibilité du cadre rouge et délimitation du chapitre ou de la section en cours de lecture. Les deux éléments sont assez facilement décelables à l'aide de l'un et l'autre test.

Problème 4 : Absence d'outil de recherche pour les images (niveau 2)

Méthode de test automatisée	Méthode de test modérée
Capacité au repérage 1. Problème passé inaperçu 2. Problème soupçonné 3. Problème identifié ✓	1. Problème passé inaperçu 2. Problème soupçonné 3. Problème identifié ✓
Probabilité de repérage 1. Repérage ou calcul problématique 2. Repérage (peu) probable 3. - 4. Repérage très probable ✓ <i>Cette haute probabilité est peut-être due au public qui a effectué le test (spécialiste de l'information pour la plupart)</i>	1. Repérage ou calcul problématique 2. Repérage peu probable ✓ 3. Repérage probable 4. Repérage très probable
Efficience du repérage 1. Repérage requérant une manipulation particulière des données 2. Repérage ne requérant pas de manipulation particulière des données ✓ <i>Via les commentaires.</i>	1. Repérage requérant une manipulation particulière des données 2. Repérage ne requérant pas de manipulation particulière des données ✓
Evaluation de la gravité du problème <ul style="list-style-type: none"> Fréquence du problème Niveau d'insatisfaction soulevé Taux d'échec de la tâche causé Temps perdu ou effort nécessaire pour dépasser le problème ✓ 	<ul style="list-style-type: none"> Fréquence du problème Niveau d'insatisfaction soulevé Taux d'échec de la tâche causé Temps perdu ou effort nécessaire pour dépasser le problème
Facilitation du design de la solution 1. Le problème est repéré mais avec peu d'indice sur la manière de le résoudre 2. Une solution est plus ou moins induite ✓	1. Le problème est repéré mais avec peu d'indice sur la manière de le résoudre 2. Une solution est plus ou moins induite ✓
Score : 12	Score : 10

Commentaires : vu la difficulté souvent rencontrée pour trouver une image, la mise à disposition d'un index permettrait de gagner à la fois du temps et de les retrouver plus vite. L'inconvénient est évidemment le coût d'une telle mesure.

Problème 5 : Accès difficile à l'information à l'intérieur des documents (niveau 2)

Méthode de test automatisée	Méthode de test modérée
Capacité au repérage <ol style="list-style-type: none"> 1. Problème passé inaperçu 2. Problème soupçonné 3. Problème identifié ✓ 	<ol style="list-style-type: none"> 1. Problème passé inaperçu 2. Problème soupçonné 3. Problème identifié ✓
Probabilité de repérage <ol style="list-style-type: none"> 1. Repérage ou calcul problématique 2. Repérage (peu) probable 3. - 4. Repérage très probable ✓ 	<ol style="list-style-type: none"> 1. Repérage ou calcul problématique 2. Repérage peu probable 3. Repérage probable ✓ 4. Repérage très probable
Efficience du repérage <ol style="list-style-type: none"> 1. Repérage requérant une manipulation particulière des données 2. Repérage ne requérant pas de manipulation particulière des données ✓ 	<ol style="list-style-type: none"> 1. Repérage requérant une manipulation particulière des données 2. Repérage ne requérant pas de manipulation particulière des données ✓
Evaluation de la gravité du problème <ul style="list-style-type: none"> • Fréquence du problème • Niveau d'insatisfaction soulevé • Taux d'échec de la tâche causé • Temps perdu ou effort nécessaire pour dépasser le problème 	<ul style="list-style-type: none"> • Fréquence du problème • Niveau d'insatisfaction soulevé ✓ • Taux d'échec de la tâche causé • Temps perdu ou effort nécessaire pour dépasser le problème
Facilitation du design de la solution <ol style="list-style-type: none"> 1. Le problème est repéré mais avec peu d'indice sur la manière de le résoudre 2. Une solution est plus ou moins induite ✓ 	<ol style="list-style-type: none"> 1. Le problème est repéré mais avec peu d'indice sur la manière de le résoudre 2. Une solution est plus ou moins induite ✓
Score : 11	Score : 11

Commentaires : vue la difficulté souvent rencontrée pour trouver un élément spécifique à l'intérieur du texte des documents, la mise à disposition d'un outil de recherche plein texte serait un réel avantage, surtout pour une population composée en bonne partie de chercheurs qui ne lisent que rarement les ouvrages in extenso. L'inconvénient est évidemment le coût d'une telle mesure.

Problème 6 : Affichage des résultats trop peu clair (pas d'étiquette comme titre, auteur, etc.) (niveau 2)

Méthode de test automatisée	Méthode de test modérée
Capacité au repérage 1. Problème passé inaperçu 2. Problème soupçonné 3. Problème identifié ✓	1. Problème passé inaperçu 2. Problème soupçonné 3. Problème identifié ✓
Probabilité de repérage 1. Repérage ou calcul problématique ✓ 2. Repérage (peu) probable 3. - 4. Repérage très probable <i>Repérage via un unique commentaire. Le comportement donne peu d'indications sur la facilité de lecture des notices.</i>	1. Repérage ou calcul problématique 2. Repérage peu probable ✓ 3. Repérage probable 4. Repérage très probable <i>Repérage via une erreur spécifique, une réponse erronée précise s'explique parce que la personne n'a pas vu la mention d'auteur.</i>
Efficience du repérage 1. Repérage requérant une manipulation particulière des données 2. Repérage ne requérant pas de manipulation particulière des données ✓	1. Repérage requérant une manipulation particulière des données 2. Repérage ne requérant pas de manipulation particulière des données ✓
Evaluation de la gravité du problème <ul style="list-style-type: none"> Fréquence du problème Niveau d'insatisfaction soulevé Taux d'échec de la tâche causé Temps perdu ou effort nécessaire pour dépasser le problème 	<ul style="list-style-type: none"> Fréquence du problème Niveau d'insatisfaction soulevé ✓ Taux d'échec de la tâche causé Temps perdu ou effort nécessaire pour dépasser le problème
Facilitation du design de la solution 1. Le problème est repéré mais avec peu d'indice sur la manière de le résoudre 2. Une solution est plus ou moins induite ✓	1. Le problème est repéré mais avec peu d'indice sur la manière de le résoudre 2. Une solution est plus ou moins induite ✓
Score : 8	Score : 10

Commentaires : a posteriori, on pourrait montrer que ce problème explique éventuellement plusieurs difficultés dans la lecture et l'utilisation du site. Par exemple, tous ceux qui avaient une impression de confusion et de ne pas comprendre ce à quoi ils aboutissent quand ils cliquent sur « parcourir auteur ». Une analyse plus approfondie de toutes les réponses du test automatisé nous aurait toutefois permis de repérer le problème avec plus de probabilité, vu qu'une des erreurs peut s'expliquer à travers cet élément (répondre 1513 à la tâche 3, année que donne quatre participants du test germanophone).

Problème 7 : Après recherche, option de tri « relevance » toujours affichée automatiquement quand on modifie l'ordre (niveau 2)

Méthode de test automatisée	Méthode de test modérée
Capacité au repérage <ol style="list-style-type: none"> 1. Problème passé inaperçu ✓ 2. Problème soupçonné 3. Problème identifié 	<ol style="list-style-type: none"> 1. Problème passé inaperçu ✓ 2. Problème soupçonné 3. Problème identifié <p><i>A noter que ce problème était apparu lors des prétests.</i></p>
Probabilité de repérage <ol style="list-style-type: none"> 1. Repérage ou calcul problématique ✓ 2. Repérage (peu) probable 3. - 4. Repérage très probable <p><i>Comme ce problème est de l'ordre du bug, on peut s'attendre à ce que les utilisateurs qui l'observent en fassent mention dans la question ouverte, mais cela n'a pas été le cas.</i></p>	<ol style="list-style-type: none"> 1. Repérage ou calcul problématique 2. Repérage peu probable ✓ 3. Repérage probable 4. Repérage très probable <p><i>En réalité, la faible probabilité est surtout due au design de la tâche qui ne nécessite pas, en principe, l'usage de cette fonctionnalité.</i></p>
Efficience du repérage <ol style="list-style-type: none"> 1. Repérage requérant une manipulation particulière des données 2. Repérage ne requérant pas de manipulation particulière des données <p><i>Le repérage est difficile en analysant le flux de clics. En effet, il n'y a pas concordance entre l'affichage et l'url. Sans commentaire direct d'utilisateur, le repérage est très compromis.</i></p>	<ol style="list-style-type: none"> 1. Repérage requérant une manipulation particulière des données 2. Repérage ne requérant pas de manipulation particulière des données <p><i>Si le problème est rencontré, le repérage est ici aisé, puisqu'il est possible d'observer ce qui se passe à l'écran de l'utilisateur.</i></p>
Evaluation de la gravité du problème <ul style="list-style-type: none"> • Fréquence du problème • Niveau d'insatisfaction soulevé • Taux d'échec de la tâche causé • Temps perdu ou effort nécessaire pour dépasser le problème 	<ul style="list-style-type: none"> • Fréquence du problème • Niveau d'insatisfaction soulevé • Taux d'échec de la tâche causé • Temps perdu ou effort nécessaire pour dépasser le problème
Facilitation du design de la solution <ol style="list-style-type: none"> 1. Le problème est repéré mais avec peu d'indice sur la manière de le résoudre 2. Une solution est plus ou moins induite 	<ol style="list-style-type: none"> 1. Le problème est repéré mais avec peu d'indice sur la manière de le résoudre 2. Une solution est plus ou moins induite
Score : 3	Score : 7

Commentaires : ce bug est apparu lors des prétests, il n'a ensuite plus été détecté. A posteriori, on pourrait, en parcourant les listes de pages parcourues par certains participants à distance, supposer qu'ils ont eu un problème, mais rien de manifeste. Ce cas met en évidence un des problèmes des tests à distance où aucun observateur ne regarde directement l'écran du participant. En effet, dans les url, l'utilisation de l'option de tri est bien indiquée, mais l'affichage replace le menu sous relevance. Lors des prétests, une observation directe avait permis de repérer facilement le problème. Malheureusement, cela

ne s'est pas reproduit dans la série de tests successive. En effet, la tâche ne nécessitait pas qu'on modifie l'ordre et n'était pas très favorable pour la découverte de cet élément.

Problème 8 : Confusion liée à la double pagination (niveau 2)

Méthode de test automatisée	Méthode de test modérée
Capacité au repérage <ol style="list-style-type: none"> 1. Problème passé inaperçu 2. Problème soupçonné 3. Problème identifié ✓ 	<ol style="list-style-type: none"> 1. Problème passé inaperçu 2. Problème soupçonné 3. Problème identifié ✓
Probabilité de repérage <ol style="list-style-type: none"> 1. Repérage ou calcul problématique 2. Repérage (peu) probable 3. - 4. Repérage très probable ✓ 	<ol style="list-style-type: none"> 1. Repérage ou calcul problématique 2. Repérage peu probable 3. Repérage probable 4. Repérage très probable ✓
Efficience du repérage <ol style="list-style-type: none"> 1. Repérage requérant une manipulation particulière des données 2. Repérage ne requérant pas de manipulation particulière des données ✓ <p><i>Le problème est assez manifeste et plusieurs l'ont signalé dans la question ouverte. Par contre, repérer le problème dans les données demande d'analyser les réponses et peut présenter davantage de complexité.</i></p>	<ol style="list-style-type: none"> 1. Repérage requérant une manipulation particulière des données 2. Repérage ne requérant pas de manipulation particulière des données ✓ <p><i>L'observation des choix en direct rend assez manifeste si un utilisateur s'appuie sur le mauvais numéro de page sans compter l'aide du think-aloud.</i></p>
Evaluation de la gravité du problème <ul style="list-style-type: none"> • Fréquence du problème • Niveau d'insatisfaction soulevé • Taux d'échec de la tâche causé • Temps perdu ou effort nécessaire pour dépasser le problème 	<ul style="list-style-type: none"> • Fréquence du problème • Niveau d'insatisfaction soulevé ✓ • Taux d'échec de la tâche causé • Temps perdu ou effort nécessaire pour dépasser le problème ✓
Facilitation du design de la solution <ol style="list-style-type: none"> 1. Le problème est repéré mais avec peu d'indice sur la manière de le résoudre ✓ 2. Une solution est plus ou moins induite 	<ol style="list-style-type: none"> 1. Le problème est repéré mais avec peu d'indice sur la manière de le résoudre ✓ 2. Une solution est plus ou moins induite
Score : 11	Score : 12

Commentaires : En général, les utilisateurs peuvent comparer les deux systèmes de pagination proposés dans le menu déroulant et les comparer avec ce qu'on observe sur la page même de l'ouvrage consulté. Cette démarche ne laisse toutefois pas de trace dans le flux de clics. Par contre, si un utilisateur consulte l'index du dernier tome et ensuite accède à la page selon la numérotation propre à e-rara.ch, on pourra en conclure que l'utilisateur a confondu cette dernière pagination avec celle de l'ouvrage lui-même. Le test modéré est donc ici assez largement avantage.

Problème 9 : Fenêtre de recherche peu visible (niveau 2)

Méthode de test automatisée	Méthode de test modérée
Capacité au repérage 1. Problème passé inaperçu 2. Problème soupçonné 3. Problème identifié	1. Problème passé inaperçu 2. Problème soupçonné 3. Problème identifié ✓
Probabilité de repérage 1. Repérage ou calcul problématique ✓ 2. Repérage (peu) probable 3. - 4. Repérage très probable <i>Uniquement de vagues indices, à l'exception d'un commentaire qui souligne le problème.</i>	1. Repérage ou calcul problématique 2. Repérage peu probable 3. Repérage probable ✓ 4. Repérage très probable
Efficience du repérage 1. Repérage requérant une manipulation particulière des données 2. Repérage ne requérant pas de manipulation particulière des données ✓ <i>C'est surtout un commentaire qui a permis de détecter d'identifier avec sûreté le problème. Sinon, la faible utilisation nous a amené à le soupçonner.</i>	1. Repérage requérant une manipulation particulière des données 2. Repérage ne requérant pas de manipulation particulière des données ✓
Evaluation de la gravité du problème <ul style="list-style-type: none"> Fréquence du problème Niveau d'insatisfaction soulevé Taux d'échec de la tâche causé Temps perdu ou effort nécessaire pour dépasser le problème 	<ul style="list-style-type: none"> Fréquence du problème Niveau d'insatisfaction soulevé ✓ Taux d'échec de la tâche causé Temps perdu ou effort nécessaire pour dépasser le problème ✓
Facilitation du design de la solution 1. Le problème est repéré mais avec peu d'indice sur la manière de le résoudre ✓ 2. Une solution est plus ou moins induite	1. Le problème est repéré mais avec peu d'indice sur la manière de le résoudre ✓ 2. Une solution est plus ou moins induite
Score : 7	Score : 10

Commentaires : Pour les questions de repérage des éléments sur la page, nous sommes très dépendants des commentaires. En leur absence, on peut uniquement constater la faible utilisation d'une fonctionnalité, mais cela ne nous en donne pas la raison. Quoi qu'il en soit, le test modéré est ici avantagé, puisque sur sept, deux participants ont signalé spontanément le problème, parfois même avec virulence.

Problème 10 : Interaction entre les différents menus des outils de tri peu claire (niveau 2)

Méthode de test automatisée	Méthode de test modérée
Capacité au repérage <ol style="list-style-type: none"> 1. Problème passé inaperçu 2. Problème soupçonné ✓ 3. Problème identifié 	<ol style="list-style-type: none"> 1. Problème passé inaperçu 2. Problème soupçonné 3. Problème identifié ✓
Probabilité de repérage <ol style="list-style-type: none"> 1. Repérage ou calcul problématique ✓ 2. Repérage (peu) probable 3. - 4. Repérage très probable <p><i>Il y a 22 participants qui utilisent le tri, mais ne modifient pas le type de tri en laissant l'option par défaut.</i></p>	<ol style="list-style-type: none"> 1. Repérage ou calcul problématique 2. Repérage peu probable 3. Repérage probable 4. Repérage très probable ✓ <p><i>Un seul participant a permis de repérer le problème dans les tests modérés, mais celui-ci est souvent observé dans le test automatisé (oubli de modifier le critère de tri). On peut supposer qu'on a joué de malchance.</i></p>
Efficience du repérage <ol style="list-style-type: none"> 1. Repérage requérant une manipulation particulière des données ✓ 2. Repérage ne requérant pas de manipulation particulière des données <p><i>Avec autant de non-utilisations, on peut en déduire qu'il y a un problème, mais cela nécessite malgré tout un effort de déduction et de prendre en compte plusieurs données. Le risque serait de considérer cela comme des oublis.</i></p>	<ol style="list-style-type: none"> 1. Repérage requérant une manipulation particulière des données ✓ 2. Repérage ne requérant pas de manipulation particulière des données <p><i>Le participant en question n'a pas directement mentionné le problème, mais plutôt fait état de plusieurs incompréhensions liées à ces fonctions. On en déduit que le participant les croyait indépendantes.</i></p>
Evaluation de la gravité du problème <ul style="list-style-type: none"> • Fréquence du problème ✓ • Niveau d'insatisfaction soulevé • Taux d'échec de la tâche causé ✓ • Temps perdu ou effort nécessaire pour dépasser le problème 	<ul style="list-style-type: none"> • Fréquence du problème • Niveau d'insatisfaction soulevé ✓ • Taux d'échec de la tâche causé • Temps perdu ou effort nécessaire pour dépasser le problème
Facilitation du design de la solution <ol style="list-style-type: none"> 1. Le problème est repéré mais avec peu d'indice sur la manière de le résoudre ✓ 2. Une solution est plus ou moins induite 	<ol style="list-style-type: none"> 1. Le problème est repéré mais avec peu d'indice sur la manière de le résoudre ✓ 2. Une solution est plus ou moins induite
Score : 7	Score : 10

Commentaires : Le repérage via la méthode automatisée est délicat. On constate que 22 participants ne modifient pas le menu « trier par ». On en conclut soit à un oubli soit à une incompréhension du fonctionnement du système. Le repérage est donc délicat. L'avantage du test modéré ici est de comprendre pourquoi un utilisateur accomplit telle ou telle action. C'est moins clair quand il ne fait pas quelque chose, mais dans le cas présent suffisamment de « symptômes » indiquaient cette incompréhension.

Problème 11 : Manque d'aide générale (niveau 2)

Méthode de test automatisée	Méthode de test modérée
Capacité au repérage 1. Problème passé inaperçu 2. Problème soupçonné 3. Problème identifié ✓	1. Problème passé inaperçu 2. Problème soupçonné 3. Problème identifié ✓
Probabilité de repérage 1. Repérage ou calcul problématique 2. Repérage (peu) probable 3. - 4. Repérage très probable ✓ <i>Uniquement dans les commentaires, mais ceux-ci sont suffisamment abondants.</i>	1. Repérage ou calcul problématique 2. Repérage peu probable 3. Repérage probable 4. Repérage très probable ✓
Efficience du repérage 1. Repérage requérant une manipulation particulière des données 2. Repérage ne requérant pas de manipulation particulière des données ✓	1. Repérage requérant une manipulation particulière des données 2. Repérage ne requérant pas de manipulation particulière des données ✓
Evaluation de la gravité du problème <ul style="list-style-type: none"> Fréquence du problème Niveau d'insatisfaction soulevé Taux d'échec de la tâche causé Temps perdu ou effort nécessaire pour dépasser le problème 	<ul style="list-style-type: none"> Fréquence du problème Niveau d'insatisfaction soulevé ✓ Taux d'échec de la tâche causé Temps perdu ou effort nécessaire pour dépasser le problème
Facilitation du design de la solution 1. Le problème est repéré mais avec peu d'indice sur la manière de le résoudre ✓ 2. Une solution est plus ou moins induite	1. Le problème est repéré mais avec peu d'indice sur la manière de le résoudre ✓ 2. Une solution est plus ou moins induite
Score : 10	Score : 11

Commentaires : Plusieurs commentaires indiquent que le site manque d'aide. La difficulté est ici de déterminer dans quelle mesure cette aide est nécessaire ou si ce type de demande ne provient pas en réalité d'autres problèmes d'utilisabilité. Comme le site ne propose pas de rubrique « aide », on peut supposer que cet élément constitue un manque pour les utilisateurs ou du moins est manifestement perçu de cette manière. Les deux méthodes sont à peu près équivalentes pour l'analyse de ce problème.

Problème 12 : Manque de visibilité du caractère cliquable de l'image du livre des notices (niveau 2)

Méthode de test automatisée	Méthode de test modérée
Capacité au repérage 1. Problème passé inaperçu 2. Problème soupçonné ✓ 3. Problème identifié	1. Problème passé inaperçu 2. Problème soupçonné 3. Problème identifié ✓
Probabilité de repérage 1. Repérage ou calcul problématique ✓ 2. Repérage (peu) probable 3. - 4. Repérage très probable <i>Peu d'utilisateurs ont cliqué sur le livre, mais il n'est pas évident de conclure à un manque de visibilité.</i>	1. Repérage ou calcul problématique 2. Repérage peu probable ✓ 3. Repérage probable 4. Repérage très probable
Efficience du repérage 1. Repérage requérant une manipulation particulière des données ✓ 2. Repérage ne requérant pas de manipulation particulière des données <i>Comme le repérage du problème repose sur une conjecture, les vérifications à faire engendrent un travail supplémentaire (quelles solutions alternatives ? Avec quelle efficacité ?).</i>	1. Repérage requérant une manipulation particulière des données 2. Repérage ne requérant pas de manipulation particulière des données ✓
Evaluation de la gravité du problème <ul style="list-style-type: none"> Fréquence du problème Niveau d'insatisfaction soulevé Taux d'échec de la tâche causé Temps perdu ou effort nécessaire pour dépasser le problème 	<ul style="list-style-type: none"> Fréquence du problème Niveau d'insatisfaction soulevé ✓ Taux d'échec de la tâche causé Temps perdu ou effort nécessaire pour dépasser le problème
Facilitation du design de la solution 1. Le problème est repéré mais avec peu d'indice sur la manière de le résoudre ✓ 2. Une solution est plus ou moins induite	1. Le problème est repéré mais avec peu d'indice sur la manière de le résoudre ✓ 2. Une solution est plus ou moins induite
Score : 5	Score : 9

Commentaires : Lorsqu'il s'agit d'évaluer la visibilité d'un élément, le test automatisé semble ici encore défavorisé. A noter toutefois que seul un participant du test modéré a permis de relever ce problème. Comme un chemin alternatif existe dans le menu horizontal pour la même fonctionnalité, difficile d'interpréter la préférence pour l'un ou l'autre outil. Une différence statistique même significative n'est en elle-même pas problématique si les deux options apparaissent clairement au visiteur. Ceci peut surtout être vérifié via des commentaires ou le think-aloud.

Problème 13 : Scroll latéral fréquent sur des écrans étroits (niveau 2)

Méthode de test automatisée	Méthode de test modérée
Capacité au repérage 1. Problème passé inaperçu 2. Problème soupçonné 3. Problème identifié ✓	1. Problème passé inaperçu 2. Problème soupçonné 3. Problème identifié ✓
Probabilité de repérage 1. Repérage ou calcul problématique 2. Repérage (peu) probable ✓ 3. - 4. Repérage très probable	1. Repérage ou calcul problématique 2. Repérage peu probable ✓ 3. Repérage probable 4. Repérage très probable
Efficience du repérage 1. Repérage requérant une manipulation particulière des données ✓ 2. Repérage ne requérant pas de manipulation particulière des données <i>Deux participants ont signalé ce problème dans la question ouverte. Il était donc pertinent de tester nous-même avec un écran étroit.</i>	1. Repérage requérant une manipulation particulière des données 2. Repérage ne requérant pas de manipulation particulière des données ✓
Evaluation de la gravité du problème <ul style="list-style-type: none"> Fréquence du problème Niveau d'insatisfaction soulevé Taux d'échec de la tâche causé Temps perdu ou effort nécessaire pour dépasser le problème 	<ul style="list-style-type: none"> Fréquence du problème Niveau d'insatisfaction soulevé ✓ Taux d'échec de la tâche causé Temps perdu ou effort nécessaire pour dépasser le problème
Facilitation du design de la solution 1. Le problème est repéré mais avec peu d'indice sur la manière de le résoudre ✓ 2. Une solution est plus ou moins induite	1. Le problème est repéré mais avec peu d'indice sur la manière de le résoudre ✓ 2. Une solution est plus ou moins induite
Score : 7	Score : 9

Commentaires : Le problème est difficile à détecter via le test automatisé sauf par la question ouverte. En effet, il n'est pas possible avec Loop11 de vérifier l'interface et le matériel de l'utilisateur (sauf éventuellement en posant des questions à cet effet). Pour le test modéré, celui qui mène les tests voit les conditions concrètes dans lesquelles se déroulent le test et peut donc facilement constater ce type de problème.

Problème 14 : Intitulé « provenance du document » peu clair (niveau 1)

Méthode de test automatisée	Méthode de test modérée
Capacité au repérage <ol style="list-style-type: none"> 1. Problème passé inaperçu ✓ 2. Problème soupçonné 3. Problème identifié 	<ol style="list-style-type: none"> 1. Problème passé inaperçu 2. Problème soupçonné 3. Problème identifié ✓
Probabilité de repérage <ol style="list-style-type: none"> 1. Repérage ou calcul problématique 2. Repérage (peu) probable ✓ 3. - 4. Repérage très probable <p><i>Une seule personne a cliqué sur ce lien, il n'est pas possible d'en conclure quelque chose.</i></p>	<ol style="list-style-type: none"> 1. Repérage ou calcul problématique 2. Repérage peu probable ✓ 3. Repérage probable 4. Repérage très probable
Efficience du repérage <ol style="list-style-type: none"> 1. Repérage requérant une manipulation particulière des données 2. Repérage ne requérant pas de manipulation particulière des données <p><i>Les tâches ne permettent pas vraiment de tester la compréhension de cette rubrique. Comme celle-ci a été presque toujours ignorée, il n'est pas possible d'en dire beaucoup.</i></p>	<ol style="list-style-type: none"> 1. Repérage requérant une manipulation particulière des données 2. Repérage ne requérant pas de manipulation particulière des données ✓
Evaluation de la gravité du problème <ul style="list-style-type: none"> • Fréquence du problème • Niveau d'insatisfaction soulevé • Taux d'échec de la tâche causé • Temps perdu ou effort nécessaire pour dépasser le problème 	<ul style="list-style-type: none"> • Fréquence du problème • Niveau d'insatisfaction soulevé ✓ • Taux d'échec de la tâche causé • Temps perdu ou effort nécessaire pour dépasser le problème
Facilitation du design de la solution <ol style="list-style-type: none"> 1. Le problème est repéré mais avec peu d'indice sur la manière de le résoudre 2. Une solution est plus ou moins induite 	<ol style="list-style-type: none"> 1. Le problème est repéré mais avec peu d'indice sur la manière de le résoudre ✓ 2. Une solution est plus ou moins induite
Score : 3	Score : 9

Commentaires : Ce problème n'a été détecté que par un utilisateur du test modéré. Sur la quasi centaine de participants au test automatisé, une seule personne a cliqué sur ce lien. Il s'agit donc d'un fait très marginal. S'agit-il véritablement d'un problème d'utilisabilité ? Comme l'enjeu est de modifier un intitulé, l'investissement n'est pas très important si on trouve une meilleure formulation (sans l'expression « document original » qui laisse peut-être entendre qu'on peut y consulter directement le texte original).

Problème 15 : Limitation dans la recherche des variantes orthographiques au sein des équations de recherche (niveau 1)

Méthode de test automatisée	Méthode de test modérée
Capacité au repérage 1. Problème passé inaperçu ✓ 2. Problème soupçonné 3. Problème identifié ✓	1. Problème passé inaperçu ✓ 2. Problème soupçonné 3. Problème identifié
Probabilité de repérage 1. Repérage ou calcul problématique 2. Repérage (peu) probable ✓ 3. - 4. Repérage très probable <i>Deux commentaires mentionnent de telles difficultés.</i>	1. Repérage ou calcul problématique 2. Repérage peu probable ✓ 3. Repérage probable 4. Repérage très probable
Efficience du repérage 1. Repérage requérant une manipulation particulière des données ✓ 2. Repérage ne requérant pas de manipulation particulière des données <i>Deux problèmes différents ont été signalés, l'un concernait la possibilité d'une troncature interne et l'autre l'usage de booléens+parenthèses. Ces demandes doivent servir à répondre à un problème : difficulté pour rechercher des variantes notamment orthographiques.</i>	1. Repérage requérant une manipulation particulière des données 2. Repérage ne requérant pas de manipulation particulière des données
Evaluation de la gravité du problème <ul style="list-style-type: none"> Fréquence du problème Niveau d'insatisfaction soulevé Taux d'échec de la tâche causé Temps perdu ou effort nécessaire pour dépasser le problème 	<ul style="list-style-type: none"> Fréquence du problème Niveau d'insatisfaction soulevé Taux d'échec de la tâche causé Temps perdu ou effort nécessaire pour dépasser le problème
Facilitation du design de la solution 1. Le problème est repéré mais avec peu d'indice sur la manière de le résoudre 2. Une solution est plus ou moins induite ✓	1. Le problème est repéré mais avec peu d'indice sur la manière de le résoudre 2. Une solution est plus ou moins induite
Score : 8	Score : 3

Commentaires : Ce problème n'a été détecté que par les commentaires du test automatisé. En effet, dans les livres anciens l'orthographe est moins bien fixée et les variantes orthographiques assez nombreuses. Ce problème est d'autant plus présent en l'absence d'un système de renvoi pour les noms d'auteur. Il n'y a guère que les commentaires qui permettent de repérer ce type de problème qui ne concerne qu'une minorité d'utilisateurs (peu d'entre eux vont de toute façon utiliser troncature interne ou équations complexes avec booléens).

Problème 16 : Menu horizontal (titre, contenu, etc.) peu visible (niveau 1)

Méthode de test automatisée	Méthode de test modérée
Capacité au repérage <ol style="list-style-type: none"> 1. Problème passé inaperçu ✓ 2. Problème soupçonné 3. Problème identifié 	<ol style="list-style-type: none"> 1. Problème passé inaperçu 2. Problème soupçonné 3. Problème identifié ✓
Probabilité de repérage <ol style="list-style-type: none"> 1. Repérage ou calcul problématique ✓ 2. Repérage (peu) probable 3. - 4. Repérage très probable <i>Pas d'indication.</i>	<ol style="list-style-type: none"> 1. Repérage ou calcul problématique 2. Repérage peu probable ✓ 3. Repérage probable 4. Repérage très probable <i>Un seul participant a permis de détecter clairement ce problème.</i>
Efficience du repérage <ol style="list-style-type: none"> 1. Repérage requérant une manipulation particulière des données 2. Repérage ne requérant pas de manipulation particulière des données <i>Une absence de clics sur ce menu pourrait l'indiquer, mais cela se chiffre plutôt en perte de temps pour se repérer.</i>	<ol style="list-style-type: none"> 1. Repérage requérant une manipulation particulière des données 2. Repérage ne requérant pas de manipulation particulière des données ✓
Evaluation de la gravité du problème <ul style="list-style-type: none"> • Fréquence du problème • Niveau d'insatisfaction soulevé • Taux d'échec de la tâche causé • Temps perdu ou effort nécessaire pour dépasser le problème 	<ul style="list-style-type: none"> • Fréquence du problème • Niveau d'insatisfaction soulevé ✓ • Taux d'échec de la tâche causé • Temps perdu ou effort nécessaire pour dépasser le problème ✓
Facilitation du design de la solution <ol style="list-style-type: none"> 1. Le problème est repéré mais avec peu d'indice sur la manière de le résoudre 2. Une solution est plus ou moins induite 	<ol style="list-style-type: none"> 1. Le problème est repéré mais avec peu d'indice sur la manière de le résoudre ✓ 2. Une solution est plus ou moins induite
Score : 2	Score : 10

Commentaires : Une première question que soulèvent ces résultats est celle de savoir si en effet il y a problème ou s'il s'agit seulement de la bizarrerie d'un des participants. Une variable supplémentaire pourrait être la certitude de la « réalité » d'un problème. A nouveau, pour une question de visuel, le test automatisé n'est pas avantage.

Problème 17 : Notion de lien permanent peu claire (niveau 1)

Méthode de test automatisée	Méthode de test modérée
Capacité au repérage 1. Problème passé inaperçu 2. Problème soupçonné 3. Problème identifié ✓	1. Problème passé inaperçu 2. Problème soupçonné 3. Problème identifié ✓
Probabilité de repérage 1. Repérage ou calcul problématique 2. Repérage (peu) probable 3. - 4. Repérage très probable ✓	1. Repérage ou calcul problématique 2. Repérage peu probable ✓ 3. Repérage probable ✓ 4. Repérage très probable
Efficiency du repérage 1. Repérage requérant une manipulation particulière des données ✓ 2. Repérage ne requérant pas de manipulation particulière des données <i>Une petite dizaine de participants aux tests ont cliqué sur ce lien non pertinent pour leur recherche. On peut en conclure qu'il y a là un problème de clarté.</i>	1. Repérage requérant une manipulation particulière des données 2. Repérage ne requérant pas de manipulation particulière des données ✓
Evaluation de la gravité du problème <ul style="list-style-type: none"> Fréquence du problème ✓ Niveau d'insatisfaction soulevé Taux d'échec de la tâche causé ✓ Temps perdu ou effort nécessaire pour dépasser le problème 	<ul style="list-style-type: none"> Fréquence du problème Niveau d'insatisfaction soulevé ✓ Taux d'échec de la tâche causé Temps perdu ou effort nécessaire pour dépasser le problème ✓
Facilitation du design de la solution 1. Le problème est repéré mais avec peu d'indice sur la manière de le résoudre ✓ 2. Une solution est plus ou moins induite	1. Le problème est repéré mais avec peu d'indice sur la manière de le résoudre ✓ 2. Une solution est plus ou moins induite
Score : 12	Score : 11

Commentaires : Ce problème est plus visible dans le test modéré où il est possible de constater pourquoi les participants cliquent sur ce lien. Dans le cas du test automatisé, on peut seulement constater qu'environ 10% des participants cliquent dessus. Ce lien n'a d'ailleurs pas d'effet concret si ce n'est de modifier l'url en en lien permanent de type DOI. Ce problème pourrait paraître important parce qu'il touche trois participants sur sept du test modéré, mais le test automatisé montre que cette proportion n'est pas si élevée au final, puisque dans ce contexte, il n'y a que huit utilisateurs qui ont cliqué sur ce lien.

On voit donc l'importance d'un échantillon suffisamment grand pour donner une représentation plus pertinente de la fréquence d'un phénomène.

Problème 18 : Pas de possibilité de classer alphabétiquement les facettes (niveau 1)

Méthode de test automatisée	Méthode de test modérée
Capacité au repérage <ol style="list-style-type: none"> 1. Problème passé inaperçu 2. Problème soupçonné 3. Problème identifié ✓ 	<ol style="list-style-type: none"> 1. Problème passé inaperçu ✓ 2. Problème soupçonné 3. Problème identifié
Probabilité de repérage <ol style="list-style-type: none"> 1. Repérage ou calcul problématique 2. Repérage (peu) probable ✓ 3. - 4. Repérage très probable 	<ol style="list-style-type: none"> 1. Repérage ou calcul problématique 2. Repérage peu probable ✓ 3. Repérage probable 4. Repérage très probable
Efficiency du repérage <ol style="list-style-type: none"> 1. Repérage requérant une manipulation particulière des données ✓ 2. Repérage ne requérant pas de manipulation particulière des données <p><i>Un seul commentaire signale ce problème, de plus il parle d'un mauvais classement. Or ces facettes sont classées en fonction du nombre de résultats qu'elles renvoient puis en fonction de l'alphabet. On peut supposer que c'est l'ordre alphabétique seul qui intéresse la personne.</i></p>	<ol style="list-style-type: none"> 1. Repérage requérant une manipulation particulière des données 2. Repérage ne requérant pas de manipulation particulière des données <p><i>A posteriori, peut-être que nous aurions pu détecter ce problème dans les tests modérés en analysant plus en détail les interactions avec les facettes.</i></p>
Evaluation de la gravité du problème <ul style="list-style-type: none"> • Fréquence du problème • Niveau d'insatisfaction soulevé • Taux d'échec de la tâche causé • Temps perdu ou effort nécessaire pour dépasser le problème 	<ul style="list-style-type: none"> • Fréquence du problème • Niveau d'insatisfaction soulevé • Taux d'échec de la tâche causé • Temps perdu ou effort nécessaire pour dépasser le problème
Facilitation du design de la solution <ol style="list-style-type: none"> 1. Le problème est repéré mais avec peu d'indice sur la manière de le résoudre ✓ 2. Une solution est plus ou moins induite 	<ol style="list-style-type: none"> 1. Le problème est repéré mais avec peu d'indice sur la manière de le résoudre 2. Une solution est plus ou moins induite
Score : 7	Score : 3

Commentaires : On peut se demander si ce problème serait encore perçu si e-rara.ch proposait une liste d'auteurs lorsqu'on clique sur « parcourir auteur ». En effet, ce que recherche probablement la personne qui a fait ce commentaire, c'est probablement une telle liste. Quoi qu'il en soit, ce problème ne semble pas avoir concerné beaucoup de personnes et semble assez marginal.

Problème 19 : Position du lien vers les fichiers pdf trop à droite (niveau 1)

Méthode de test automatisée	Méthode de test modérée
Capacité au repérage <ol style="list-style-type: none"> 1. Problème passé inaperçu ✓ 2. Problème soupçonné 3. Problème identifié 	<ol style="list-style-type: none"> 1. Problème passé inaperçu 2. Problème soupçonné 3. Problème identifié ✓
Probabilité de repérage <ol style="list-style-type: none"> 1. Repérage ou calcul problématique 2. Repérage (peu) probable ✓ 3. - 4. Repérage très probable <p><i>Les tâches ne sont pas du tout centrées autour de cette fonctionnalité, au contraire, il est interdit explicitement de télécharger des pdf.</i></p>	<ol style="list-style-type: none"> 1. Repérage ou calcul problématique 2. Repérage peu probable ✓ 3. Repérage probable 4. Repérage très probable <p><i>Repérage via observation des actions d'un participant.</i></p>
Efficience du repérage <ol style="list-style-type: none"> 1. Repérage requérant une manipulation particulière des données 2. Repérage ne requérant pas de manipulation particulière des données <p><i>Il n'y a guère qu'un commentaire qui puisse efficacement faire noter ce problème. Mais aucun n'a été donné dans ce sens.</i></p>	<ol style="list-style-type: none"> 1. Repérage requérant une manipulation particulière des données 2. Repérage ne requérant pas de manipulation particulière des données ✓
Evaluation de la gravité du problème <ul style="list-style-type: none"> • Fréquence du problème • Niveau d'insatisfaction soulevé • Taux d'échec de la tâche causé • Temps perdu ou effort nécessaire pour dépasser le problème 	<ul style="list-style-type: none"> • Fréquence du problème • Niveau d'insatisfaction soulevé ✓ • Taux d'échec de la tâche causé • Temps perdu ou effort nécessaire pour dépasser le problème
Facilitation du design de la solution <ol style="list-style-type: none"> 1. Le problème est repéré mais avec peu d'indice sur la manière de le résoudre 2. Une solution est plus ou moins induite 	<ol style="list-style-type: none"> 1. Le problème est repéré mais avec peu d'indice sur la manière de le résoudre ✓ 2. Une solution est plus ou moins induite
Score : 3	Score : 9

Commentaires : Contrairement au scroll, qui touche les petits écrans (problème 16), celui-ci touche plutôt les grands écrans. Le test automatisé ne permet pas de détecter aisément ce type de problèmes sauf via des commentaires. Si on avait conçu une tâche spécifique sur ce bouton, on aurait pu voir le nombre d'erreur de ligne et demander quelque part la taille de l'écran pour vérifier s'il y a corrélation.

Les tâches ne sont pas pensées pour évaluer cette partie du site, si bien que les informations sont très pauvres ce qui rend la méthode automatisée assez peu efficace, et ce d'autant plus pour un problème de disposition des éléments sur l'écran et de forme dynamique de la page.

Problème 20 : Recherche avancée impossible sans entrer un terme (niveau 1)

Méthode de test automatisée	Méthode de test modérée
Capacité au repérage 1. Problème passé inaperçu 2. Problème soupçonné 3. Problème identifié ✓	1. Problème passé inaperçu ✓ 2. Problème soupçonné 3. Problème identifié
Probabilité de repérage 1. Repérage ou calcul problématique 2. Repérage (peu) probable ✓ 3. - 4. Repérage très probable	1. Repérage ou calcul problématique 2. Repérage peu probable ✓ 3. Repérage probable 4. Repérage très probable
Efficience du repérage 1. Repérage requérant une manipulation particulière des données 2. Repérage ne requérant pas de manipulation particulière des données ✓ <i>Un seul commentaire signale ce problème.</i>	1. Repérage requérant une manipulation particulière des données 2. Repérage ne requérant pas de manipulation particulière des données <i>Un repérage via observation directe aurait tout à fait été possible.</i>
Evaluation de la gravité du problème <ul style="list-style-type: none"> • Fréquence du problème • Niveau d'insatisfaction soulevé • Taux d'échec de la tâche causé • Temps perdu ou effort nécessaire pour dépasser le problème 	<ul style="list-style-type: none"> • Fréquence du problème • Niveau d'insatisfaction soulevé • Taux d'échec de la tâche causé • Temps perdu ou effort nécessaire pour dépasser le problème
Facilitation du design de la solution 1. Le problème est repéré mais avec peu d'indice sur la manière de le résoudre 2. Une solution est plus ou moins induite ✓	1. Le problème est repéré mais avec peu d'indice sur la manière de le résoudre 2. Une solution est plus ou moins induite
Score : 12	Score : 3

Commentaires : Bien que marginal, ce problème en est manifestement un. Pourquoi exiger un terme absolument ? Il est possible ailleurs sur le site d'accéder à la liste des livres d'une bibliothèque en particulier. Pour ce genre de problème qui ne touche pas directement les tâches, plus de personnes parcourent le site, plus il y a de chance de les faire émerger.

Problème 21 : Recherche difficile dans la liste des contenus (niveau 1)

Méthode de test automatisée	Méthode de test modérée
Capacité au repérage 1. Problème passé inaperçu 2. Problème soupçonné 3. Problème identifié ✓	1. Problème passé inaperçu ✓ 2. Problème soupçonné 3. Problème identifié
Probabilité de repérage 1. Repérage ou calcul problématique 2. Repérage (peu) probable 3. - 4. Repérage très probable ✓	1. Repérage ou calcul problématique 2. Repérage peu probable ✓ 3. Repérage probable 4. Repérage très probable
Efficience du repérage 1. Repérage requérant une manipulation particulière des données 2. Repérage ne requérant pas de manipulation particulière des données ✓ <i>Quatre commentaires en relation avec ce problème.</i>	1. Repérage requérant une manipulation particulière des données 2. Repérage ne requérant pas de manipulation particulière des données <i>Un repérage via observation directe ou commentaire aurait tout à fait été possible.</i>
Evaluation de la gravité du problème <ul style="list-style-type: none"> • Fréquence du problème • Niveau d'insatisfaction soulevé • Taux d'échec de la tâche causé • Temps perdu ou effort nécessaire pour dépasser le problème 	<ul style="list-style-type: none"> • Fréquence du problème • Niveau d'insatisfaction soulevé • Taux d'échec de la tâche causé • Temps perdu ou effort nécessaire pour dépasser le problème
Facilitation du design de la solution 1. Le problème est repéré mais avec peu d'indice sur la manière de le résoudre ✓ 2. Une solution est plus ou moins induite	1. Le problème est repéré mais avec peu d'indice sur la manière de le résoudre 2. Une solution est plus ou moins induite
Score : 10	Score : 3

Commentaires : Les participants demandaient un moyen de faire comme « ctrl+f » pour faire une recherche dans la table des matières afin de gagner du temps. Il est vrai que parcourir la table peut demander énormément de temps et qu'un outil à cet effet pourrait en faire gagner. Pour ce type de fonctionnalité de confort qui fait gagner du temps, mais dont l'absence n'engendre pas d'erreur, il n'y a guère que les commentaires ou l'observation directe qui permette d'en déceler le manque.

Problème 22 : Termes « croissant » et « décroissant » indifféremment appliqué à des chiffres et à des lettres (niveau 1)

Méthode de test automatisée	Méthode de test modérée
Capacité au repérage 1. Problème passé inaperçu ✓ 2. Problème soupçonné 3. Problème identifié	1. Problème passé inaperçu 2. Problème soupçonné 3. Problème identifié ✓
Probabilité de repérage 1. Repérage ou calcul problématique ✓ 2. Repérage (peu) probable 3. - 4. Repérage très probable <i>Il n'y a guère que via commentaire qu'un tel problème peut être repéré.</i>	1. Repérage ou calcul problématique 2. Repérage peu probable ✓ 3. Repérage probable 4. Repérage très probable
Efficience du repérage 1. Repérage requérant une manipulation particulière des données 2. Repérage ne requérant pas de manipulation particulière des données <i>Un repérage via commentaire serait possible.</i>	1. Repérage requérant une manipulation particulière des données 2. Repérage ne requérant pas de manipulation particulière des données ✓
Evaluation de la gravité du problème <ul style="list-style-type: none"> Fréquence du problème Niveau d'insatisfaction soulevé Taux d'échec de la tâche causé Temps perdu ou effort nécessaire pour dépasser le problème 	<ul style="list-style-type: none"> Fréquence du problème Niveau d'insatisfaction soulevé Taux d'échec de la tâche causé Temps perdu ou effort nécessaire pour dépasser le problème
Facilitation du design de la solution 1. Le problème est repéré mais avec peu d'indice sur la manière de le résoudre 2. Une solution est plus ou moins induite	1. Le problème est repéré mais avec peu d'indice sur la manière de le résoudre ✓ 2. Une solution est plus ou moins induite
Score : 2	Score : 9

Commentaires : Il est vrai qu'on attendrait plutôt « ascendant » et « descendant ». Le participant au test modéré qui nous a permis de découvrir ce problème n'avait en outre pas compris le lien entre le menu « ordre » et le menu « trier par ». Cette imprécision au niveau du vocabulaire a encouragé la confusion. Concernant le repérage spécifique de ce problème, il n'y a guère qu'un commentaire qui permette de le découvrir, que ce soit avec l'une ou l'autre méthode.

Problème 23 : Troncature « * » non acceptée (niveau 1)

Méthode de test automatisée	Méthode de test modérée
Capacité au repérage 1. Problème passé inaperçu 2. Problème soupçonné 3. Problème identifié ✓	1. Problème passé inaperçu 2. Problème soupçonné 3. Problème identifié ✓
Probabilité de repérage 1. Repérage ou calcul problématique 2. Repérage (peu) probable ✓ 3. - 4. Repérage très probable	1. Repérage ou calcul problématique 2. Repérage peu probable ✓ 3. Repérage probable 4. Repérage très probable
Efficience du repérage 1. Repérage requérant une manipulation particulière des données 2. Repérage ne requérant pas de manipulation particulière des données ✓	1. Repérage requérant une manipulation particulière des données 2. Repérage ne requérant pas de manipulation particulière des données ✓
Evaluation de la gravité du problème <ul style="list-style-type: none"> • Fréquence du problème ✓ • Niveau d'insatisfaction soulevé • Taux d'échec de la tâche causé ✓ • Temps perdu ou effort nécessaire pour dépasser le problème ✓ 	<ul style="list-style-type: none"> • Fréquence du problème • Niveau d'insatisfaction soulevé ✓ • Taux d'échec de la tâche causé • Temps perdu ou effort nécessaire pour dépasser le problème ✓
Facilitation du design de la solution 1. Le problème est repéré mais avec peu d'indice sur la manière de le résoudre ✓ 2. Une solution est plus ou moins induite	1. Le problème est repéré mais avec peu d'indice sur la manière de le résoudre ✓ 2. Une solution est plus ou moins induite
Score : 11	Score : 10

Commentaires : Le problème est marginal et ne touche que très peu d'utilisateurs. Néanmoins, il occasionne une perte de temps chez ceux qui tentent cette troncature. Les deux formes de tests permettent d'évaluer ce phénomène avec des nuances quelque peu différentes surtout dans le domaine de l'évaluation de la gravité du problème.

Problème 24 : Visibilité des bibliothèques participantes insuffisantes (niveau 1)

Méthode de test automatisée	Méthode de test modérée
Capacité au repérage 1. Problème passé inaperçu 2. Problème soupçonné ✓ 3. Problème identifié	1. Problème passé inaperçu 2. Problème soupçonné 3. Problème identifié ✓
Probabilité de repérage 1. Repérage ou calcul problématique 2. Repérage (peu) probable 3. - 4. Repérage très probable	1. Repérage ou calcul problématique 2. Repérage peu probable 3. Repérage probable ✓ 4. Repérage très probable
Efficience du repérage 1. Repérage requérant une manipulation particulière des données ✓ 2. Repérage ne requérant pas de manipulation particulière des données <i>Des commentaires auraient pu aider à repérer ce problème, mais le nombre de gens qui échouent à la tâche 1 (déterminer si une certaine bibliothèque met ses livres à disposition sur e-rara.ch) constitue un indice également.</i>	1. Repérage requérant une manipulation particulière des données 2. Repérage ne requérant pas de manipulation particulière des données ✓ <i>Problème mentionné dans la discussion après les tâches.</i>
Evaluation de la gravité du problème <ul style="list-style-type: none"> Fréquence du problème ✓ Niveau d'insatisfaction soulevé Taux d'échec de la tâche causé ✓ Temps perdu ou effort nécessaire pour dépasser le problème 	<ul style="list-style-type: none"> Fréquence du problème Niveau d'insatisfaction soulevé ✓ Taux d'échec de la tâche causé Temps perdu ou effort nécessaire pour dépasser le problème
Facilitation du design de la solution 1. Le problème est repéré mais avec peu d'indice sur la manière de le résoudre ✓ 2. Une solution est plus ou moins induite	1. Le problème est repéré mais avec peu d'indice sur la manière de le résoudre ✓ 2. Une solution est plus ou moins induite
Score : 8	Score : 10

Commentaires : Ce problème est formulé de manière assez large et ne peut difficilement être repéré par le test automatisé que par le taux d'erreur à la première tâche ou des commentaires. Ce sont surtout les discussions sur l'organisation du site en général après les tâches qui ont permis de mettre en évidence cette faiblesse assez réelle. En effet, la page d'accueil offre un texte assez peu... accueillant. Pourquoi ne pas y mettre en évidence les bibliothèques participantes ? La méthode modérée est ici avantagée.

Problème 25 : Visibilité des liens hypertextes (niveau 1)

Méthode de test automatisée	Méthode de test modérée
Capacité au repérage 1. Problème passé inaperçu 2. Problème soupçonné 3. Problème identifié ✓	1. Problème passé inaperçu ✓ 2. Problème soupçonné 3. Problème identifié
Probabilité de repérage 1. Repérage ou calcul problématique 2. Repérage (peu) probable ✓ 3. - 4. Repérage très probable	1. Repérage ou calcul problématique 2. Repérage peu probable ✓ 3. Repérage probable 4. Repérage très probable
Efficience du repérage 1. Repérage requérant une manipulation particulière des données 2. Repérage ne requérant pas de manipulation particulière des données ✓ <i>Un seul commentaire mentionne le fait.</i>	1. Repérage requérant une manipulation particulière des données 2. Repérage ne requérant pas de manipulation particulière des données
Evaluation de la gravité du problème <ul style="list-style-type: none"> • Fréquence du problème ✓ • Niveau d'insatisfaction soulevé • Taux d'échec de la tâche causé ✓ • Temps perdu ou effort nécessaire pour dépasser le problème 	<ul style="list-style-type: none"> • Fréquence du problème • Niveau d'insatisfaction soulevé ✓ • Taux d'échec de la tâche causé • Temps perdu ou effort nécessaire pour dépasser le problème
Facilitation du design de la solution 1. Le problème est repéré mais avec peu d'indice sur la manière de le résoudre ✓ 2. Une solution est plus ou moins induite	1. Le problème est repéré mais avec peu d'indice sur la manière de le résoudre 2. Une solution est plus ou moins induite
Score : 8	Score : 3

Commentaires : différencier les liens hypertextes par un soulignement et un changement de couleur fait partie de conventions bien établies sur le web qu'il n'est pas absolument nécessaire de suivre. On peut se demander dans quelle mesure ce participant a réellement été gêné par ce problème ou s'il ne fait que rapporter des conventions. S'il y avait des liens réellement peu visibles, on peut supposer que la méthode modérée permet plus facilement de repérer des absences de comportement ou qui sont différés parce qu'il a fallu chercher.

Annexe 5 : poster de présentation du projet

