UNIVERSITÉ DE MONTRÉAL

ÉVALUATION DE L'INTERFACE UTILISATEUR D'UN SYSTÈME D'AIDE À L'APPRENTISSAGE DE L'ÉCONOMIE ET CONCEPTION D'UNE NOUVELLE INTERFACE

AURÉLIEN SALOMON

DÉPARTEMENT DE MATHÉMATIQUES ET DE GÉNIE INDUSTRIEL

ÉCOLE POLYTECHNIQUE DE MONTRÉAL

MÉMOIRE PRÉSENTÉ EN VUE DE L'OBTENTION

DU DIPLÔME DE MAÎTRISE ÈS SCIENCES APPLIQUÉES

(GÉNIE INDUSTRIEL)

AOÛT 2014

© Aurélien Salomon, 2014.

UNIVERSITÉ DE MONTRÉAL

ÉCOLE POLYTECHNIQUE DE MONTRÉAL

Ce mémoire intitulé :

ÉVALUATION DE L'INTERFACE UTILISATEUR D'UN SYSTÈME D'AIDE À L'APPRENTISSAGE DE L'ÉCONOMIE ET CONCEPTION D'UNE NOUVELLE INTERFACE

présenté par : SALOMON Aurélien

en vue de l'obtention du diplôme de : Maîtrise ès sciences appliquées

a été dûment accepté par le jury d'examen constitué de :

Mme DE MARCELLIS-WARIN Nathalie, Doct., présidente

M. ROBERT Jean-Marc, Doct., membre et directeur de recherche

M. BOUDREAULT Yves, Ph.D., membre

REMERCIEMENTS

J'aimerais dans un premier temps remercier tout particulièrement M. Jean-Marc Robert, pour m'avoir aidé et guidé pendant mes études, et plus récemment pour ce mémoire. Sa disponibilité, sa bienveillance et son expertise m'ont permis de mener à bien ce projet.

Je voudrais également remercier l'équipe du CIRANO, et plus particulièrement Marine Demontaignac, Nathalie Viennot et Marco Lugo, pour leur coopération et leurs encouragements tout au long de ce projet.

Je souhaiterais aussi remercier toutes les personnes qui m'ont aidé en répondant aux questionnaires. Merci pour leur disponibilité et pour leur aide.

Enfin, je remercie ma famille et mes amis qui ont toujours été là pour m'apporter leur bonne humeur et leur soutien tout au long de ce projet.

RÉSUMÉ

Le système d'aide à l'apprentissage de notions d'économie dont il est question dans ce mémoire propose de nombreuses expériences pédagogiques aux apprenants. C'est un projet en cours de développement au CIRANO et les essais menés sur la première version avaient mis en évidence la nécessité d'améliorer celle-ci.

Ce mémoire a pour objectif d'améliorer la qualité de l'expérience de l'utilisateur avec ce système. Pour atteindre cet objectif, nous avons dans un premier temps évalué l'interface du système actuel. Cette évaluation a révélé que les attributs hédoniques de l'interface étaient limités. Nous avons aussi fait une évaluation heuristique des attributs pragmatiques de l'interface et plus particulièrement de l'utilisabilité, en ayant recours aux critères de Bastien & Scapin (1993). Cette évaluation a permis d'identifier 87 problèmes ergonomiques. De nombreux problèmes étaient récurrents dans les différentes expériences pédagogiques et ont été détaillés.

Nous avons ensuite procédé à la conception d'une nouvelle interface utilisateur visant à résoudre ces différents problèmes et à améliorer les attributs hédoniques de l'interface. Pour cela, nous avons effectué une conception à deux niveaux : au niveau supérieur, une structure globale commune à l'ensemble de l'outil a été conçue et développée afin de s'attaquer à la récurrence des problèmes décelés. Puis, à un niveau inférieur, chaque expérience a été spécifiquement conçue à nouveau, tout en respectant les contraintes établies par le CIRANO relatives au contenu économique à couvrir (celui-ci devait rester intact) et à l'implémentation rapide de la nouvelle interface. Nous avons choisi de présenter en détail la conception de l'une de ces expériences qui se nomme « Ressource commune ».

Pour évaluer et comparer l'EU vécue avec l'ancienne et la nouvelle interface de l'outil, nous avons utilisé l'outil « AttrakDiff » d'Hassenzahl en version réduite (2010). Cet outil est composé de 10 paires de mots opposés (échelle d'Osgood). L'échelle comporte sept points entre chaque mot et son opposé, permettant à l'utilisateur de choisir, dans chaque paire, le mot qui décrit le mieux le produit. Cet outil était intégré à un bref questionnaire que nous avons élaboré pour obtenir une mesure simple de l'apprentissage réalisé en économie ainsi qu'une mesure de l'appréciation du contenu économique de l'expérience pédagogique. Des questions ont été également ajoutées au sujet d'autres buts de l'entreprise tels que la satisfaction de l'utilisateur, l'envie de recommander le système, etc.

Vingt participants ont pris part à notre recherche. Ils ont été répartis en deux groupes. Le premier groupe essayait l'ancienne version, l'évaluait à l'aide du questionnaire, puis essayait la nouvelle version et l'évaluait à l'aide du même questionnaire. Le deuxième groupe effectuait un processus similaire, mais commençait d'abord par la nouvelle version pour contrebalancer les effets d'apprentissage. Les résultats montrent que la nouvelle version est mieux évaluée que l'ancienne version. Les qualités pragmatiques et hédoniques de la nouvelle version obtiennent des améliorations respectives de 17 % et de 33 %. L'évaluation de l'attractivité globale du système révèle que celle-ci s'améliore de 45 %. Ces résultats montrent une amélioration de l'EU vécue avec le système.

Lorsque les participants évaluent la nouvelle version de l'outil après avoir testé l'ancienne, leur évaluation de la nouvelle version est plus homogène que lorsqu'ils n'ont testé que la nouvelle version seulement. La qualité hédonique est encore mieux notée, par contre la qualité pragmatique semble légèrement moins bien notée, peut-être parce que le participant s'habitue à l'utilisabilité de l'ancienne version. Lorsque les participants évaluent l'ancienne version après avoir testé la nouvelle, leurs évaluations de l'ancienne version sont plus sévères, ce qui peut être expliqué par l'aversion à la perte.

Le niveau d'éducation élevé des participants à notre recherche ne nous a pas permis d'évaluer l'apprentissage réalisé en économie, ainsi que l'impact de l'appréciation du contenu économique traité par le système sur l'appréciation de l'interface ; tous les participants ont déclaré apprécier le domaine de l'économie.

La nouvelle version de l'interface du système affecte positivement les autres buts visés par le CIRANO avec ce système en améliorant la satisfaction de l'utilisateur de 13 %, l'envie de réutiliser et de recommander le système de 15 % et 18 %, l'envie que des expériences pédagogiques similaires soient développées de 20 %, et l'image du CIRANO que le système véhicule de 23 %. Seule la profitabilité, qui reflète le caractère avantageux de l'expérience pédagogique, ne semble pas s'être améliorée avec la nouvelle version, car elle serait vraisemblablement plus liée au contenu des expériences pédagogiques.

ABSTRACT

The economic learning support system discussed in this report offers numerous educational experiments to the learners. It is a project under development at the CIRANO. Tests conducted on the first version had highlighted the need to improve the user experience with the system.

The goal of this study is to enhance the quality of the user experience with this system. To reach this goal, we assessed the interface of the current system. This revealed that the hedonic attributes of the interface were limited. We also conducted a heuristic evaluation of the pragmatic attributes of the interface, and more particularly the usability, with the ergonomic criteria of Bastien & Scapin (1993). This evaluation helped to identify 87 ergonomic problems. Numerous problems were recurring in the various educational experiments and were detailed.

We then proceeded with the design of a new user interface to solve these problems and to improve the hedonic attributes of the interface. For that purpose, we focused on two levels: at the upper level, a global structure used by the whole system was designed and developed to deal with the recurrence of the revealed problems. Then, at a lower level, every experiment was specifically redesigned, while respecting the constraints established by the CIRANO, relative to the economic contents (which had to remain intact) and to the fast implementation of the new interface. We chose to present in detail the design of one of these experiments, which is called " Common Resource ".

To estimate and compare the user experience with the old and the new interface of the system, we used the reduced version of the tool " AttrakDiff " (Hassenzahl, 2010). This tool includes 10 pairs of opposite words (Osgood's scale). The scale contains seven points between every word and its opposite, enabling the user to choose, in every pair, the word that best describes the product. This tool was integrated into a short questionnaire we developed to obtain a simple measure of the learning realized as well as a measure of appreciation of the economic contents of the educational experiment. Questions were also added about other goals of the company, such as the user satisfaction, the desire to recommend the system, etc.

Twenty participants took part in our research. They were divided into two groups. The first group tried the old version, estimated it with the questionnaire, then tried the new version and estimated it with the same questionnaire. The second group followed a similar process, but began with the

new version to counterbalance the learning effects. Results show that the new version is better estimated than the old version. The pragmatic and hedonic qualities of the new version get respective improvements of 17% and 33%. The evaluation of the global attractiveness of the system reveals a 45% improvement with the new version. These results show an improvement of user experience with the system.

The evaluation of the new version is more homogeneous when participants already tested the old version, than when they only tested the new version; the hedonic quality gets an even better score, but the pragmatic quality gets a slightly worse one, maybe because the participant becomes familiar with the usability of the old version. When participants estimate the old version, having already tested the new one, their evaluations of the old version are more severe, which can be explained by the loss aversion.

The high educational level of the participants in this research did not allow us to estimate correctly the learning realized in economy, as well as the impact of the appreciation of the economic contents of the experiment on the appreciation of the interface. However, the new version of the system interface positively affects the other goals of the CIRANO with this system, by improving the satisfaction of the user by 13 %, the desire to reuse and to recommend the system by 15 % and 18 %, the desire that similar educational experiments be developed by 20 %, and the image of the CIRANO the system conveys by 23 %. Only the profitability, which reflects the advantageous character of the educational experiment, does not appear to have improved with the new version, because it would be more connected to the contents of the educational experiments.

TABLE DES MATIÈRES

REMERCIEMENTS III

RÉSUMÉ IV

ABSTRACT VI

TABLE DES MATIÈRES VIII

LISTE DES TABLEAUX XI

LISTE DES FIGURES XII

LISTE DES SIGLES ET ABRÉVIATIONS XIV

LISTE DES ANNEXES XV

INTRODUCTION 1

CHAPITRE 1 REVUE DE LITTÉRATURE SUR L'UTILISABILITÉ, L'EXPÉRIENCE UTILISATEUR 3

1.1 L'UTILISABILITÉ 3

1.1.1 Définitions de l'utilisabilité 3

1.1.2 Modèles de l'utilisabilité 4

1.2 L'EXPÉRIENCE UTILISATEUR 8

1.2.1 Définition de l'expérience utilisateur 8

1.2.2 Caractéristiques de l'expérience utilisateur 9

1.2.3 Modèles de l'expérience utilisateur 10

1.3 L'ÉVALUATION DE L'EXPÉRIENCE UTILISATEUR 16

1.3.1 Classification des méthodes d'évaluation de l'EU 16

1.3.2 Les évaluations d'experts pour l'utilisabilité 18

1.3.3 Les enquêtes 22

CHAPITRE 2 CONTEXTE ET OBJECTIFS 28

2.1 Contexte : Présentation de l'outil pédagogique du CIRANO 28

2.2 L'expérience « Ressource commune » 29

2.2.1 Ressource commune 30

2.3 Bilan du premier outil 32

2.4 Objectif 32

CHAPITRE 3 PROCESSUS D'ÉVALUATION ET DE CONCEPTION DE L'INTERFACE D'UN OUTIL D'AIDE À L'APPRENTISSAGE 33

3.1 Processus d'évaluation ergonomique de l'interface de l'outil actuel 33

3.2 Évaluation ergonomique de l'interface 34

3.2.1 Méthodologie d'évaluation ergonomique 34

3.2.2 Bilan des évaluations ergonomiques 37

3.2.3 Détails des problèmes ergonomiques communs à l'ensemble des expériences 39

3.3 Conception de la nouvelle interface utilisateur 41

3.3.1 Conception de la structure globale de l'interface 41

3.3.2 Conception de l'expérience « Ressource commune » 50

3.4 Développement de la nouvelle interface 54

3.5 Validation 55

CHAPITRE 4 MÉTHODOLOGIE D'ÉVALUATION DE L'EXPÉRIENCE UTILISATEUR VÉCUE AVEC LES DEUX VERSIONS DE L'INTERFACE 57

4.1 Méthodologie 57

4.1.1 Questionnaire d'évaluation 57

4.1.2 Distinction entre l'interface et le contenu 59

4.1.3 Participants 60

4.1.4 Procédure de test 61

4.2 Nettoyage des données 62

4.3 Traitements des données 62

4.3.1 Moyenne des scores de la qualité pragmatique et hédonique 64

4.3.2 Attractivité globale 65

4.3.3 Écart-type des qualités pragmatique et hédonique 65

CHAPITRE 5 RÉSULTATS ET ANALYSE 67

5.1 Retour sur le déroulement des essais 67

5.2 Analyse de l'évaluation de l'expérience utilisateur 67

5.2.1 Comparaison des résultats des versions A et B 67

5.2.2 Évolution vers le nouvel outil 69

5.2.3 Passer de la nouvelle à l'ancienne version 71

5.3 Impact sur les autres buts du CIRANO 73

5.3.1 Impact sur l'apprentissage réalisé 73

5.3.2 Impact sur les autres buts du CIRANO 74

5.3.3 Impact du contenu sur l'évaluation de l'interface 76

CONCLUSION 77

BIBLIOGRAPHIE 79

ANNEXES 98

LISTE DES TABLEAUX

Tableau 3-1 : Problèmes ergonomiques des interfaces utilisateur des six expériences pédagogiques évaluées classés par fréquence 37

Tableau 4-1 : Section 2 du questionnaire relatif au bilan de l'utilisation de l'outil 59

Tableau 4-2 : Essais des différents échantillons 61

LISTE DES FIGURES

Figure 1-1 : L'utilisabilité reliée aux autres concepts d'après Shackel 5

Figure 1-2 : L'utilisabilité reliée aux autres concepts de Nielsen..... 6

Figure 1-3 : Modèle d'Ahrippainen et de Tähti (2003) 11

Figure 1-4 : Éléments clés du modèle de l'EU selon Hassenzahl (2003)..... 12

Figure 1-5 : Le modèle CUE (component of user experience) de Mahlke et Thüring (2007) 14

Figure 1-6 : Les intrants et les extrants de l'EU selon Robert et Lesage (2011) 15

Figure 1-7 : Les neuf heuristiques de base de Nielsen et Molich 18

Figure 1-8 : Le modèle d'EU supportant Attrakdiff (extrait de http://attrakdiff.de) 24

Figure 1-10 : Les 28 paires de mots de AttrakDiff (Traduction libre)..... 25

Figure 1-11 : Version réduite de l'outil AttrakDiff (Traduction libre)..... 26

Figure 2-1 : Aperçu de l'expérience « Ressource commune » 31

Figure 3-1 : Boite de dialogue dans l'outil actuel..... 36

Figure 3-3 : Page de jeu de l'outil actuel - expérience pédagogique de la ressource commune ... 42

Figure 3-4 : Structure d'une page du nouvel outil 44

Figure 3-5 : L'application iBooks de Apple sur iPad en 2012 46

Figure 3-6 : Corbeille (design à plat)..... 46

Figure 3-7 : Écran d'accueil de Windows 8 (flat design) 47

Figure 3-8 : Maquette d'une expérience de l'outil..... 48

Figure 3-9 : Maquette d'une expérience de l'outil (version agrandie) 48

Figure 3-10 : Ancienne identification à droite et nouvelle à gauche 49

Figure 3-11 : Aperçu de l'expérience « Ressource commune » dans l'outil actuel 50

Figure 3-12 : Conteneur des informations récurrentes dans les expériences pédagogiques..... 52

Figure 3-13 : Nouvelle version de la page de jeu de l'expérience « Ressource commune »..... 52

Figure 3-14 : Tableau récapitulatif de l'ancienne version de l'expérience « Ressource commune » 53

Figure 3-15 : Tableau récapitulatif de la nouvelle version de l'expérience « Ressource commune » B 54

Figure 4-1 : Section 1 du questionnaire AttrakDiff en version réduite 58

Figure 4-2 : Distinction entre l'interface et le contenu 59

Figure 4-3 : Variation du graphique de l'outil AttrakDiff 63

Figure 4-4 : Version réduite de l'outil AttrakDiff 64

Figure 4-5 : Rectangle d'écart type 65

Figure 5-1 : Résultats d'évaluation des deux versions de l'outil d'aide à l'apprentissage 68

Figure 5-2 : Résultats d'évaluation des deux versions de l'outil et de la nouvelle version après que les participants aient essayé et évalué l'ancienne version 70

Figure 5-3 : Comparaison des versions et régression vers l'ancienne version A 72

Figure 5-4 : Impact de l'utilisation des versions A et B de l'outil sur différents buts de l'entreprise 74

LISTE DES SIGLES ET ABRÉVIATIONS

ATT	Attractivité
CIRANO	Centre interuniversitaire de recherche en analyse des organisations
CPL	Caractères par ligne
EU	Expérience utilisateur
IHM	Interface Humain-Machine
QH	Qualité hédonique
QP	Qualité pragmatique

LISTE DES ANNEXES

ANNEXE A : LES 10 HEURISTIQUES DE NIELSEN 86

ANNEXE B : LES SEPT PRINCIPES DE LA NORME ISO 9241 (110) 88

ANNEXE C : LES CRITÈRES DE BASTIEN & SCAPIN 89

ANNEXE D : CERTIFICAT D'ÉTHIQUE À LA RECHERCHE (CER) 93

ANNEXE E : QUESTIONNAIRE D'ÉVALUATION 94

ANNEXE F : CAPTURES D'ÉCRANS DE L'ANCIENNE ET LA NOUVELLE VERSION DE « RESSOURCE COMMUNE » PAR PAGE.....96

INTRODUCTION

Nous sommes aujourd'hui à l'ère de la virtualisation des produits et de l'essor des sites web, applications mobiles et autres interfaces humain-ordinateur (IHM). Ces interfaces occupent une place de plus en plus importante dans notre quotidien. En effet, nous les utilisons tous les jours pour, entre autres commander un produit sur internet, consulter ces courriels, ou tout simplement effectuer un appel téléphonique. Dans ce contexte, le nombre de sites internet, applications mobiles, et autres logiciels a considérablement augmenté. Le marché est donc devenu extrêmement concurrentiel, et les entreprises ont dû trouver un moyen pour convaincre les utilisateurs d'utiliser leurs produits plutôt que ceux des concurrents.

L'expérience utilisateur (EU) est un concept récent, apparu au milieu des années 90 chez Apple et héritant de l'utilisabilité. Il s'agit d'un concept holistique englobant "tous les aspects de l'interaction de l'utilisateur final avec l'entreprise, ses services et ses produits" (Nielsen Norman Group). L'EU est rapidement apparue comme un facteur clé permettant d'améliorer la satisfaction de l'utilisateur. Cette dernière est primordiale dans la mesure où elle permet d'augmenter le taux de recommandation d'une interface, l'envie de l'utiliser, et permet de transformer les utilisateurs en acheteurs.

Le CIRANO (centre interuniversitaire de recherche en analyse des organisations) développe actuellement un logiciel pédagogique d'économie. Il permet aux utilisateurs de prendre part de façon interactive et ludique à des expériences leur permettant ainsi d'être confrontés à des concepts économiques tels que les marchés de bien, la notion de risque, les impôts, etc. La première itération (version bêta) des expériences pédagogiques développées par le CIRANO a eu un succès convenable. Cependant, les expériences comportaient de nombreuses failles ergonomiques et des attributs hédoniques assez limités, impactant négativement sur l'EU. Ceci avait des répercussions aussi bien sur les demandes de financement que sur le taux d'utilisation des interfaces. Le but de ce mémoire est donc de mettre en avant le processus suivi dans la nouvelle itération de conception de l'interface de l'outil visant à améliorer l'EU vécue des expériences pédagogiques.

Ce projet comporte cinq chapitres. Le premier propose une revue de littérature sur l'utilisabilité et l'EU utilisateur afin de bien comprendre les concepts clés qui seront abordés dans ce mémoire. Le deuxième chapitre met l'accent sur le contexte dans lequel ce logiciel pédagogique intervient, et les objectifs spécifiques qui seront poursuivis. Dans le troisième chapitre, nous présenterons la démarche que nous avons suivie pour analyser l'anciennes interface et concevoir la nouvelle. Le chapitre 4 présentera les méthodes de recueil de données nous ayant permis de valider quantitativement l'amélioration de l'EU. Enfin, le dernier chapitre mettra en avant les résultats obtenus et l'interprétation que nous pouvons en faire. Les limites de notre travail ainsi que les futures améliorations à effectuer clôtureront ce mémoire.

CHAPITRE 1 REVUE DE LITTÉRATURE SUR L'UTILISABILITÉ, L'EXPÉRIENCE UTILISATEUR

Cette revue de littérature va s'organiser autour de deux concepts majeurs. Premièrement, nous présenterons le concept d'utilisabilité : les définitions, les modèles et les méthodes d'évaluation. Puis nous nous intéresserons au concept d'expérience utilisateur (EU) avec des produits interactifs ; ici aussi, nous allons présenter les définitions, les modèles et les méthodes d'évaluation.

1.1 L'UTILISABILITÉ

L'utilisabilité d'un système renvoie au sens large à la facilité d'apprentissage et d'utilisation de ce dernier. Ce concept a fortement gagné en popularité avec l'apparition des systèmes informatiques ou interactifs destinés à un large public, mais aussi à cause de l'augmentation de la concurrence entre les entreprises. Améliorer l'utilisabilité d'un système permet entre autres de réduire les frustrations de l'utilisateur et d'accomplir la tâche voulue avec plus d'aisance et de rapidité. Les responsables de marketing des entreprises ont donc rapidement compris qu'il s'agissait là d'un élément clé permettant de convaincre des clients d'acheter et d'utiliser leurs produits plutôt que ceux des concurrents. Le concept est large et beaucoup de définitions (Shackel 1991, Nielsen 1993, Preece et al. 1994, Schneiderman, 1998, ISO 9241-210 2000 et d'autres) ont été proposées afin de le définir. Nous allons présenter dans le paragraphe suivant quelques-unes des définitions les plus répandues.

1.1.1 Définitions de l'utilisabilité

Nous allons présenter ici trois définitions de l'utilisabilité. Les modèles associés à ces définitions seront décrits dans la section suivante.

- Shackel (1991) fait partie des premiers auteurs à s'intéresser à l'utilisabilité. Sa définition précise que « l'utilisabilité d'un système est la capacité, en terme de fonctionnalité humaine, d'être utilisé facilement et efficacement par un éventail d'utilisateurs, avec un certain niveau d'entrainement et de support, pour accomplir un éventail spécifique de

tâches, à l'intérieur d'un éventail spécifique de scénarios » (Traduction libre). On voit que l'utilisabilité est donc relative à la facilité et l'efficacité d'utilisation.

- Nielsen (1993), qui est aussi un auteur important du domaine, ne donne pas de définition explicite. Il précise que l'utilisabilité n'est pas une « propriété à une dimension » de l'interface utilisateur, mais qu'elle est composée des cinq attributs suivants : la facilité d'apprentissage, l'efficience, la mémorabilité, les erreurs, la satisfaction. Selon cet auteur, l'utilisabilité fait partie d'un concept plus grand qui est l'acceptabilité. Nous présenterons son modèle dans la section suivante.
- La définition la plus souvent citée est sans doute celle de la norme ISO 9241-210 (2010) (adapté de la 9241-11). Cette norme précise que l'utilisabilité est « le degré auquel un système, produit ou service peut être utilisé par des utilisateurs spécifiques pour atteindre des buts spécifiques avec efficacité, efficience, et satisfaction dans un contexte d'utilisation spécifique » (traduction libre). La distinction entre l'efficacité et l'efficience est importante : l'efficacité est liée à l'atteinte ou non des buts fixés, tandis que l'efficience est liée à la quantité de ressources (efforts) afin d'atteindre ces buts. On voit également que cette définition mentionne explicitement la satisfaction de l'utilisateur. L'utilisation redondante du mot spécifique exprime le fait que l'utilisabilité est relative et peut par exemple varier selon les utilisateurs.

1.1.2 Modèles de l'utilisabilité

Nous allons présenter ici les modèles associés aux définitions précédentes.

Modèle de Shackel (1991)

Shackel avait déjà bien compris que l'utilisabilité était un facteur important pour convaincre un utilisateur d'adopter un système. Mais il avait également déjà compris que ce n'était pas l'unique facteur et qu'il faisait partie d'un concept plus englobant. Il explique que *l'utilisabilité*, *l'utilité* et *la capacité à se faire apprécier* se combinent et doivent être considérées avec le *coût du produit* afin de savoir si l'utilisateur voudra acheter le produit (*acceptability*).

UTILITY	— will it do what is needed functionally?
+	
USABILITY	— will the users actually work it successfully?
+	
LIKEABILITY	— will the users *feel* it is suitable?
must be balanced in a trade-off against	
COST	— what are the capital and running costs? — what are the social and organisational consequences?
to arrive at a decision about	
ACCEPTABILITY	— on balance the best possible alternative for purchase.

Figure 1-1 : L'utilisabilité reliée aux autres concepts d'après Shackel (1991)

L'utilité exprime le fait que le produit répond aux besoins fonctionnels de l'utilisateur. *L'utilisabilité* porte sur le succès des utilisateurs lors de leur utilisation du produit, tandis que *la capacité de se faire apprécier* parle d'elle-même. Après avoir placé l'utilisabilité dans un cadre plus englobant, Shackel a utilisé sa définition de l'utilisabilité afin de déterminer les composantes de cette dernière. L'utilisabilité aurait alors une partie subjective et une partie objective. Il propose alors quatre critères de l'utilisabilité :

- *L'efficacité :* la capacité à atteindre le but voulu
- *La facilité d'apprentissage* : la capacité à apprendre rapidement à utiliser le système
- *La flexibilité* : l'adaptation aux différentes variantes des tâches
- *L'attitude :* la satisfaction de l'utilisateur avec le système.

Il donne pour ces quatre critères des valeurs très précises à atteindre en fonction de la tâche. Le modèle de Shackel a permis de poser les bases sur lesquelles s'appuieront les modèles suivants.

Modèle de Nielsen (1993)

Comme pour le modèle précédent, Nielsen précise que l'utilisabilité se situe dans un cadre plus large qu'il va nommer « *acceptabilité* du système ». Celle-ci est selon Nielsen la combinaison d'une acceptabilité *sociale* et d'une acceptabilité *pratique*.

L'acceptabilité pratique d'un produit prend en compte de nombreux facteurs relatifs à la volonté de l'utilisateur d'adopter un produit. On peut donc citer les fonctionnalités (qui déterminent

l'utilité du produit), les coûts, la fiabilité, le support, la compatibilité avec d'autres systèmes. Le caractère utilitaire (« usefulness ») correspond globalement au fait que le produit nous permet d'atteindre nos buts, et comprend deux facteurs que sont l'utilité et l'utilisabilité.

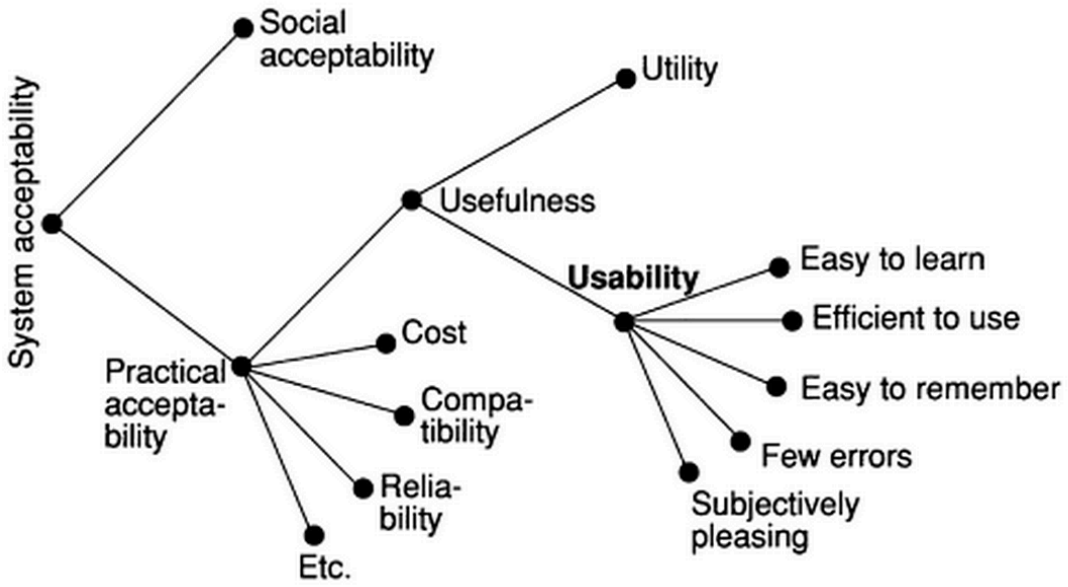


Figure 1-2 : L'utilisabilité reliée aux autres concepts de Nielsen (1993)

La figure 1-2 montre les cinq attributs qui définissent l'utilisabilité d'un système. Voici une brève description de chacun (traduction libre) :

- *Facilité d'apprentissage* : « Le système devrait être facile à apprendre afin que l'utilisateur puisse rapidement finir des tâches avec le système ».
- *Efficience* : « Le système devrait être efficient à utiliser afin que lorsque l'utilisateur a appris à l'utiliser, il puisse avoir un haut niveau de productivité ».
- *Mémorabilité* : « Le système devrait être facile à retenir, afin que l'utilisateur occasionnel soit capable de retourner au système, après un certain temps sans l'avoir utilisé, sans avoir à réapprendre depuis le début ».
- *Peu d'erreurs* : « Le système devrait occasionner un faible taux d'erreurs, de façon à ce que les utilisateurs fassent peu d'erreurs pendant l'utilisation du système et qu'ils soient capables de se rétablir facilement des erreurs qu'ils commettent. Également, les erreurs catastrophiques ne doivent pas arriver. »

- *Satisfaction* : « Le système devrait être plaisant à utiliser, afin que les utilisateurs soient subjectivement satisfaits en l'utilisant. »

On voit que globalement, les mêmes concepts reviennent d'un modèle à l'autre, bien que le modèle de Nielsen se révèle être plus complet. Sa vision de l'utilisabilité se voulait également plus holistique puisque « l'utilisabilité s'applique à tous les aspects du système avec lequel un humain peut interagir, en incluant l'installation et les opérations de maintenance ». L'utilisabilité ne se limite donc pas à l'interaction permettant d'accomplir les tâches principales.

Modèle de la norme ISO 9241-11

La norme ISO 9241-11 est un important standard de l'utilisabilité. Comme nous l'avons vu, selon la définition présentée dans cette norme, l'utilisabilité est pour un utilisateur spécifique, qui interagit avec un système dans un contexte d'utilisation spécifique afin de parvenir à ses buts. Elle se mesure selon trois principaux critères :

- *L'efficacité* : la capacité à atteindre les buts fixés ;
- *L'efficience* : la capacité à utiliser une faible quantité de ressources pour atteindre les buts ;
- *La satisfaction* : le confort et la satisfaction ressentie pour atteindre les buts.

Ce modèle se révèle beaucoup plus simple que les modèles précédents, mais il fait preuve d'un consensus important de la part de la communauté scientifique.

Tous ces modèles s'accordent sur un point : malgré le fait que l'utilisabilité représente une part importante de l'acceptabilité d'un produit, elle n'est pourtant pas suffisante. Elle permet de s'assurer que l'utilisateur arrive à bien effectuer ses tâches sans frustrations. Cependant, nos décisions d'adopter un produit ne se limitent pas à ce facteur. Lors de l'achat d'un véhicule par exemple, le design est un facteur de décision très important pour certains consommateurs[1], qui peuvent alors accorder une importance plus limitée aux autres facteurs tels que la consommation, l'habitabilité, la sécurité, etc.

---

[1] Un sondage en 2010, réalisé par Auto plus et aufeminin.com, auprès des femmes européennes montre que le design de la voiture est le premier critère de sélection (devant le prix et la consommation).

1.2 L'EXPÉRIENCE UTILISATEUR

L'utilisabilité n'est pas un concept suffisamment englobant pour décrire l'expérience vécue avec le produit. Limiter les frustrations de l'utilisateur est essentiel, mais certains produits peuvent également apporter du plaisir aux utilisateurs ainsi que leur faire ressentir différentes émotions. Le concept d'EU est justement voué à englober ces aspects.

À partir du milieu des années 1990, le terme d'EU fait son apparition chez Apple par le biais de Donald Norman (1995). En effet, ce dernier explique avoir proposé ce terme afin d'englober l'ensemble des éléments relatifs à l'expérience avec un produit, en incluant entres autres, des facteurs relatifs aux aspects affectifs et hédoniques.

De nos jours, l'EU est devenue un concept incontournable des IHM et occupe une place de plus en plus importante dans les entreprises qui s'intéressent aux interactions humain-machine. Nous présenterons dans la prochaine section une revue de littérature sur le concept de l'EU, en mettant en avant les différentes définitions, les caractéristiques et les modèles de l'EU.

1.2.1 Définition de l'expérience utilisateur

L'EU est un concept très vaste et assez complexe à délimiter. Plusieurs définitions ont été proposées ; en voici quatre qui montrent différents aspects de l'EU :

- Une des définitions les plus citées dans la littérature est celle d'Hassenzahl et Tractinsky (2006). Dans leur article, ces auteurs précisent que « l'EU est une conséquence de l'état interne de l'utilisateur (prédispositions, attentes, besoins, motivation, humeur, etc.), des caractéristiques du système conçu (ex. : complexité, but, utilisabilité, fonctionnalité, etc.) et du contexte (ou de l'environnement) dans lequel les interactions se déroulent (ex. : cadre organisationnel/social, le sens de l'activité, la velléité d'utilisation, etc.) » (Traduction libre). Ils précisent également que l'EU repose fortement sur l'idée de concevoir des systèmes dans le but de plaire aux utilisateurs plutôt que dans l'unique but de limiter les frustrations de ces derniers.

- La définition de la norme ISO 9241 (2010) stipule que l'EU correspond aux : « perceptions et réactions d'une personne qui résultent de l'utilisation effective ou/et anticipée d'un produit, système ou service ». Elle s'étire dans le temps puisqu'elle inclut aussi nos réactions qui surviennent quelque temps après avoir utilisé le produit. L'EU

dépend d'un nombre important de facteurs relatifs au produit, à son utilisation, à son contexte, ainsi qu'à l'utilisateur.

- La définition du groupe Nielsen-Norman (du site officiel : www.nngroup.com, 2014) précise que l'EU « englobe tous les aspects de l'interaction de l'utilisateur final avec l'entreprise, ses services et ses produits » (traduction libre). Cette définition montre bien l'étendue du concept de l'EU. Celle-ci ne dépend pas seulement du produit, mais des différents services de l'entreprise qui se rapportent au produit (ex., soutien, service après-vente).
- Enfin, la définition de Robert & Lesage (2011) précise que l'EU « est un construit multidimensionnel qui définit l'effet global à travers le temps de l'interaction d'un utilisateur avec un système ou un service dans un contexte spécifique » (traduction libre). Tous les aspects importants de l'EU sont réunis : un construit, le caractère multidimensionnel, un effet global dans le temps, et des interactions entre un utilisateur spécifique et un produit (implicitement : pour faire une tâche) dans un contexte spécifique.

1.2.2 Caractéristiques de l'expérience utilisateur

L'EU a de nombreuses caractéristiques (Robert & Larouche, 2011 ; Provost & Robert, 2012 ; Robert, 2014) :

- *Subjective* : étant donné que l'EU est perçue par un utilisateur, elle est subjective. Deux utilisateurs, utilisant le même système dans les mêmes conditions, pourront avoir eu des ressentis totalement différents. Elle dépend donc des caractéristiques de l'utilisateur, par exemple, sa motivation, ses intérêts, sa compétence, son humeur, ses expériences passées, etc.
- *Holistique* : comme le précise le groupe Nielsen Norman, elle « englobe tous les aspects de l'interaction de l'utilisateur final avec l'entreprise, ses services et ses produits »
- *Individuel* : un groupe peut vivre une EU, mais c'est parce que chacun des individus qui en fait partie vit une EU. Le groupe, dont fait partie un utilisateur, fait partie du contexte de sa propre EU (Law, Roto, Hassenzahl, Vermeeren, Kort, 2009).

- *Multidimensionnelle* : l'EU comprend plusieurs dimensions qui participent à la formation d'un ensemble. On peut ainsi vivre une expérience positive avec un site, car il nous aura permis de mener à bien notre tâche, sans aucune frustration et avec rapidité (dimension fonctionnelle), le tout avec une interface très agréable visuellement (psychologique). Provost (2012) a défini 12 dimensions de l'EU en se basant sur des travaux précédents (Robert & Lesage 2011) et sur les résultats d'une étude empirique qu'elle a réalisée.
- *Dynamique et cumulative* : l'EU évolue tout le temps et est le résultat des attentes envers le produit de même que des interactions avec celui-ci. Elle prend en compte toutes les attentes que l'utilisateur avait à l'égard du système avant de l'utiliser, toutes les interactions qu'il a pu avoir précédemment avec le système, ainsi que les différents contextes dans lesquels ces interactions ont eu lieu. Chaque nouvelle utilisation affecte donc notre EU.
- *Située dans un contexte* : chaque EU est vécue dans un certain contexte qui influe de façon plus ou moins importante sur cette dernière. Le contexte se définit par une multitude de facteurs qui peuvent bien sûr fortement varier d'une fois à l'autre : pression temporelle, lieu physique, présence de bruits, personnes présentes, conditions météo, etc.
- *Différents niveaux de granularité* : on peut, par exemple, considérer l'EU vécue avec une application mobile pendant une durée limitée, ou bien considérer l'EU vécue avec de nombreuses applications mobiles pendant une durée plus longue.

1.2.3 Modèles de l'expérience utilisateur

Il existe un grand nombre de modèles de l'EU dans la littérature. Ceux-ci possèdent de nombreux points en commun, mais ils ont aussi leurs particularités. Nous présenterons ici quelques-uns des principaux modèles de l'EU.

Modèle d'Arhippainen & Tähti (2003)

Le modèle d'Arhippainen et Tähti fut proposé en 2003. Ce modèle simple à comprendre, présente cinq catégories de facteurs qui interagissent les uns avec les autres afin de former l'EU.

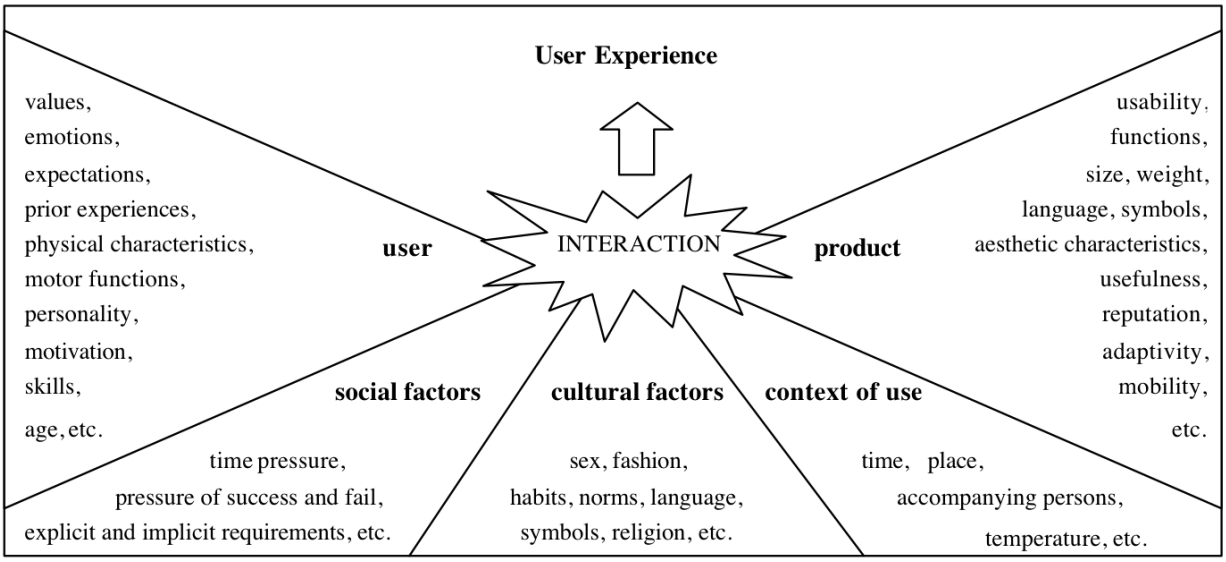


Figure 1-3 : Modèle d'Ahrippainen et de Tähti (2003)

À l'instar de la plupart des modèles de l'EU, ce dernier inclut bien évidemment *l'utilisateur*, qui *interagit* avec un *produit*, dans un *contexte d'utilisation*. Il ajoute deux catégories : les facteurs sociaux et les facteurs culturels. Cependant, on peut se demander si les facteurs culturels ne pourraient pas être inclus dans la catégorie des facteurs relevant de l'utilisateur. Roto (2006) propose quant à elle de déplacer les facteurs sociaux et culturels dans le contexte d'utilisation. Le modèle d'Ahrippainen et Tähti donne plusieurs exemples d'éléments faisant partie de chaque catégorie de facteurs.

Le modèle de Hassenzahl (2003)

Le modèle de Hassenzahl fut proposé en 2003. Cet auteur très prolifique continue encore aujourd'hui de publier de nombreux articles sur l'EU. Son modèle se révèle particulièrement intéressant puisqu'il comprend deux perspectives complémentaires : celle du designer et celle de l'utilisateur (figure 1-4).

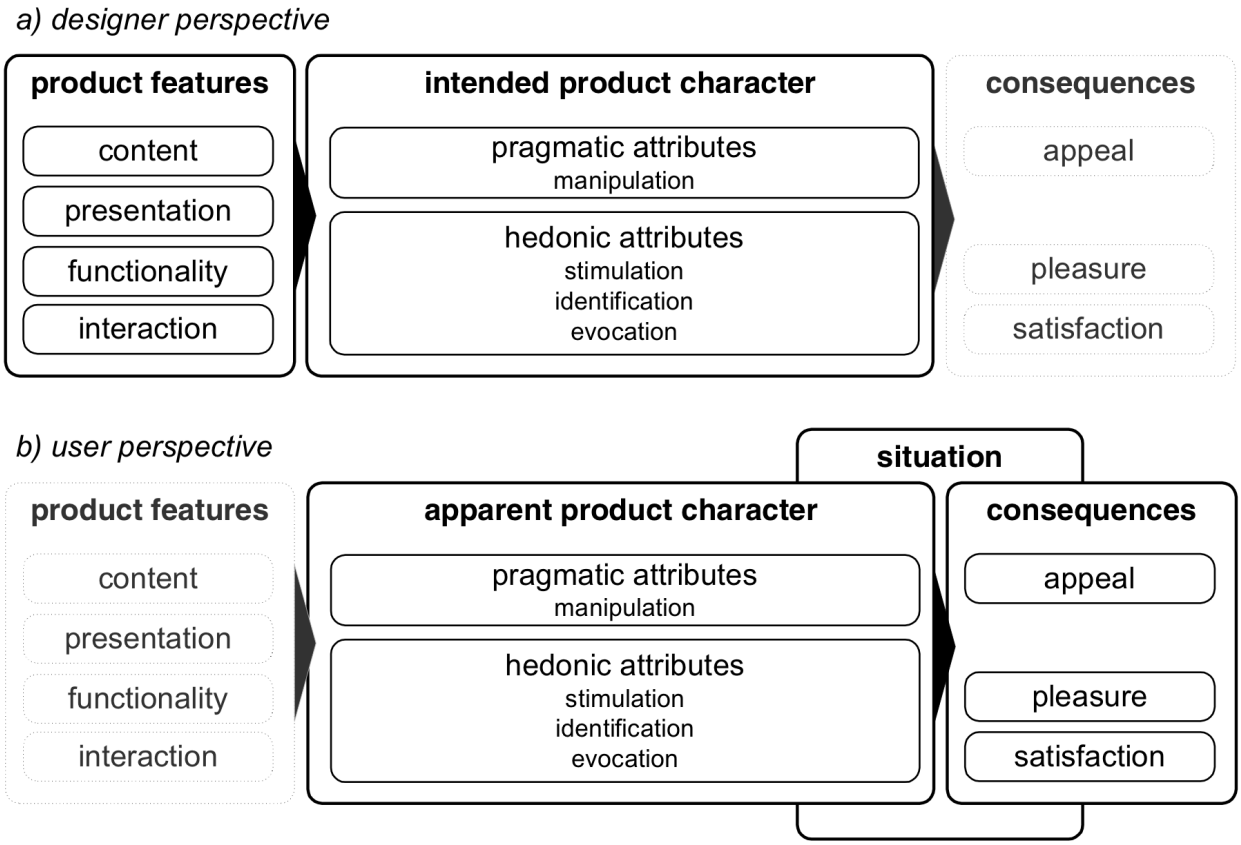


Figure 1-4 : Éléments clés du modèle de l'EU selon Hassenzahl (2003)

Dans chaque perspective, son modèle se présente sous la forme de trois éléments principaux. Premièrement, *les caractéristiques du produit* (contenu, présentation, fonctionnalité, et interaction), qui ont été conçues par le designer. Le second élément est le *caractère du produit*. Celui-ci est une description de haut niveau du produit qui résume ses attributs (par ex. : utile, prévisible, facile à utiliser, etc.). Il y a un caractère *voulu* par le designer (via les caractéristiques du produit), et un caractère *perçu* par l'utilisateur. Par exemple, un produit peut avoir comme caractéristique d'avoir une interface simple ; la perception de l'utilisateur sera qu'il est facile à utiliser. Enfin, le troisième élément porte sur *les conséquences* émotionnelles et comportementales que cela aura sur l'utilisateur.

Ce qui est critiquable dans la figure précédente est le fait que tous les éléments du modèle de l'EU ne sont pas présents. Par exemple, ce qui est perçu par l'utilisateur dépend des caractéristiques du produit, mais également des attentes et des connaissances de l'utilisateur.

Un élément très intéressant du modèle est de considérer deux catégories d'attributs du produit : *pragmatiques*, tels que l'utilité ou l'utilisabilité, et *hédoniques*, liés aux émotions, au plaisir. Les attributs pragmatiques servent à supporter les buts liés à l'action à accomplir (« do-goals », par exemple : faire un achat, envoyer un courriel, etc.), tandis que les attributs hédoniques supportent les buts liés à l'état de l'utilisateur (« be-goals », par exemple : être spécial, être connecté aux autres, etc.). Les attributs hédoniques jouent un rôle primordial dans l'expérience, et ce sont eux qui vont vraiment tirer l'expérience du côté positif. Les attributs pragmatiques ont plutôt un rôle de pilier : ils doivent être présents pour avoir une bonne expérience, mais ne sont pas suffisants. Il y a ici un lien à établir avec la théorie de la satisfaction des employés au travail de Herzberg (1976). Les attributs pragmatiques seraient alors des facteurs d'hygiène qui doivent être présents pour ne pas créer une expérience négative, mais qui n'assurent pas une expérience positive. Les facteurs de motivation seraient liés aux attributs hédoniques.

Le modèle de Mahlke & Thüring (2007)

Le modèle CUE (Component of User Experience) de Mahlke et Thüring comporte quelques éléments similaires à celui de Hassenzahl (voir figure 1-5). Ces auteurs définissent tout d'abord les caractéristiques de l'interaction, qui est formée par les propriétés du système, les attributs de l'utilisateur, la tâche ainsi que son contexte. Les caractéristiques de l'interaction vont faire percevoir à l'utilisateur deux types de qualités du système : instrumentales et non instrumentales. Les premières concernent le support, l'utilité et l'utilisabilité du système, tandis que les secondes réfèrent à l'esthétique du système et au ressenti lors de son utilisation. On voit alors que ceci rejoint les attributs pragmatiques et hédoniques d'Hassenzahl (2003).

Ces deux types de qualités perçues font partie de l'EU et vont provoquer des réactions émotionnelles chez l'utilisateur. Ainsi un système facile à utiliser et visuellement innovant peut par exemple créer des émotions telles que la surprise, ou le plaisir de l'utilisateur.

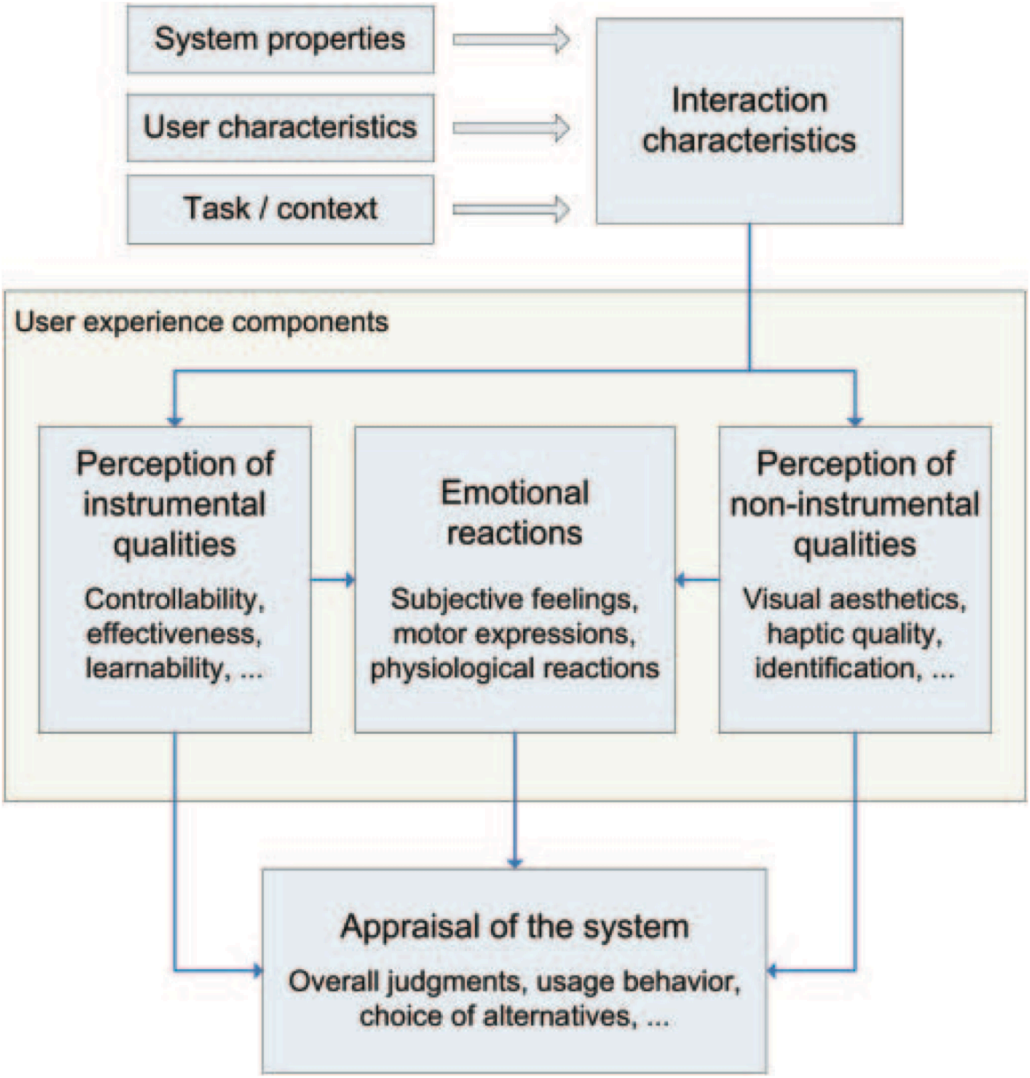


Figure 1-5 : Le modèle CUE (component of user experience) de Mahlke et Thüring (2007)

Une des critiques que l'on pourrait faire de leur modèle est le sens unidirectionnel des flèches. On peut se demander s'il n'y a pas une influence mutuelle par exemple entre les qualités instrumentales et les qualités non instrumentales, par ex., si une esthétique minimaliste d'un système peut influencer notre perception de la facilité d'utilisation.

On remarque également dans le modèle CUE que l'EU donne lieu à une évaluation globale du système par l'utilisateur. Cette évaluation ne fait pas partie de l'EU.

Le modèle de Robert et Lesage (2011)

Le modèle de Robert et Lesage a été proposé en 2010 et fut publié en 2011. Il présente les intrants et les extrants de l'EU. La figure présentée fait intervenir quatre éléments principaux : l'utilisateur, le produit, l'activité et le contexte.

L'activité est composée de buts intrinsèques et extrinsèques. Les buts intrinsèques peuvent être comparés aux « be-goals » d'Hassenzahl, relatifs à l'accomplissement de soi. Les buts extrinsèques (équivalent des do-goals) sont liés à l'action à accomplir.

Le produit possède des qualités instrumentales et non instrumentales, comme dans le modèle de Mahlke et Thüring (2007). L'ensemble des émotions de l'utilisateur, qui résultent de l'accomplissement de buts intrinsèques et extrinsèques et de la perception des qualités instrumentales et non instrumentales du produit, forment l'EU.

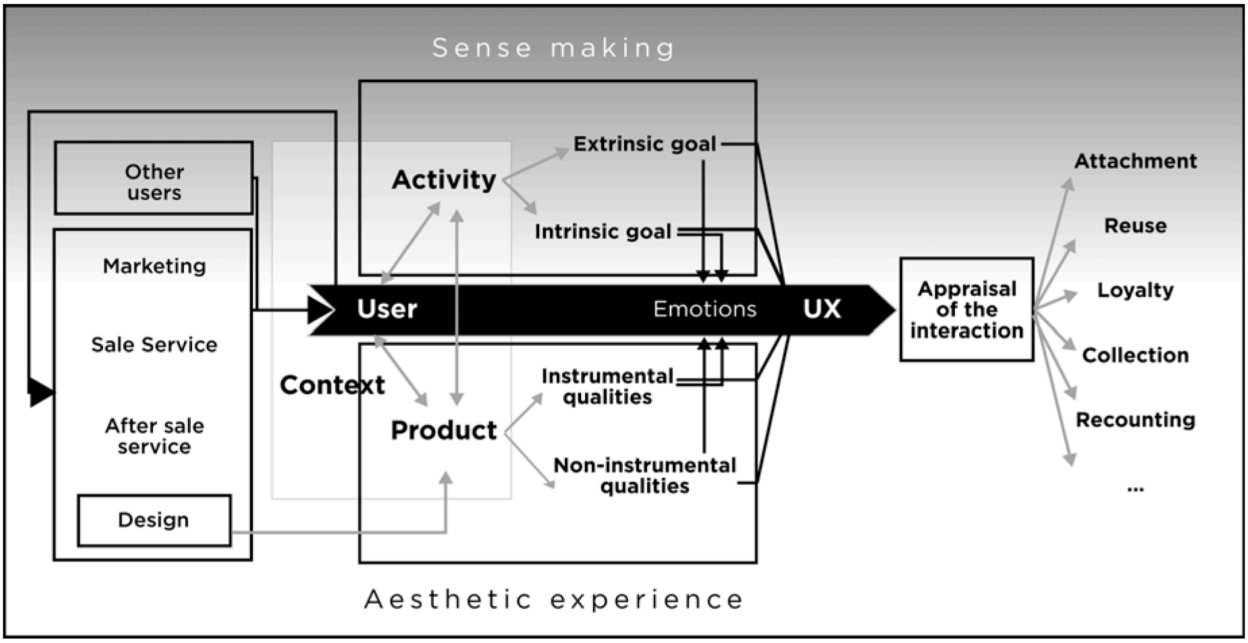


Figure 1-6 : Les intrants et les extrants de l'EU selon Robert et Lesage (2011)

Ce qui est intéressant dans ce modèle est que l'on voit le caractère holistique de l'EU, notamment le fait que l'utilisateur est soumis à des facteurs allant au-delà de l'utilisation stricte du produit, tels que le marketing, les tendances de design, les services gravitant autour du produit (ventes/après-vente). Le modèle montre aussi différents impacts possibles de l'appréciation du système par l'utilisateur : celui-ci pourra être attaché au produit, vouloir le réutiliser, collectionner des produits semblables, raconter ses expériences avec le produit, etc.

1.3 L'ÉVALUATION DE L'EXPÉRIENCE UTILISATEUR

Les différents produits ou systèmes que l'on utilise et qui sont à l'origine d'EU positives, neutres ou négatives peuvent être classés selon au moins trois catégories (Robert et Lesage, 2011) :

- *peu ou non ajustables, non interactifs*, que l'on utilise sans pouvoir en contrôler les paramètres : par ex., un télésiège, le métro, un autobus ;
- *ajustables*, qui peuvent être réglés et ajustés par l'utilisateur de façon numérique ou mécanique par ex., le siège d'une voiture, un vélo ;
- *interactifs*, avec lesquels l'utilisateur interagit via une série d'actions et de réactions ; par ex., les sites web, les applications mobiles, les assistants vocaux.

Nous nous concentrerons dans ce mémoire en particulier sur les systèmes interactifs avec utilisateur actif (et non pas passif), et nous présenterons les outils d'évaluation de l'EU portant sur les systèmes interactifs et les IHM en particulier.

1.3.1 Classification des méthodes d'évaluation de l'EU

Roto, Obrist et Väänänen-Vainio-Mattila (2009) ont proposé une classification des méthodes d'évaluation de l'EU selon cinq grandes catégories que nous présentons ci-dessous :

• Les études en laboratoire : les participants sont invités à effectuer une tâche avec un système pendant qu'un évaluateur les observe. Celui-ci peut également demander au participant de penser tout haut. On obtient alors de nombreuses informations sur l'EU (et plus particulièrement sur l'utilisabilité) directement pendant l'utilisation du produit, sans que le participant n'ait besoin de se remémorer son expérience. Cependant, ce genre d'études est relativement coûteux, car il est souvent nécessaire de rémunérer le participant qui va devoir se déplacer jusqu'au laboratoire.

• Les études sur le terrain : les participants sont observés, mais cette fois-ci, dans un contexte d'utilisation du système sur le terrain. Une des caractéristiques essentielles de l'EU que nous avons vues est qu'elle est dépendante du contexte, on comprend alors l'importance de ce type d'études.

• Les enquêtes : les enquêtes, notamment celles qui sont faites en ligne, permettent d'obtenir rapidement des informations à un coût plus faible et de rejoindre un grand nombre

d'utilisateurs. Elles s'appuient généralement sur des questionnaires contenant toutes les informations pertinentes pour que l'utilisateur puisse faire l'évaluation de façon autonome. Un des problèmes de cette méthode est qu'elle requiert un effort rétrospectif de la part du participant, ce qui peut entrainer la perte de certaines informations. On obtient des informations uniquement sur l'EU globale. Un autre problème de cette méthode d'évaluation est le faible taux de réponse.

• Les évaluations d'experts : elles sont couramment utilisées pour évaluer l'utilisabilité d'une interface. Elles s'appuient en général sur des critères (Bastien & Scapin 1993, Nielsen 1994, etc.) et permettent de dévoiler des problèmes de l'interface avant de faire (ou en remplacement) des tests utilisateurs. En ce qui concerne l'EU, le concept n'est pas assez mature pour qu'on aie des heuristiques d'évaluation.

• Les méthodes mixtes : combiner les méthodes d'évaluation (ex., une évaluation d'expert suivie de tests avec des utilisateurs) permet d'obtenir des informations beaucoup plus riches en profitant des avantages de chaque méthode. On profite également du croisement des informations pour valider les résultats. Des nouvelles méthodes mixtes apparaissent : l'outil « Peek » (2014), bien que encore en version beta, permet de faire évaluer son site Web par des utilisateurs à distance. Le responsable d'un site peut proposer des tâches à effectuer ainsi que des questions auxquelles des utilisateurs sont invités à répondre. Des participants peuvent, en tout lieu, tester le site. Leurs écrans sont enregistrés et ils doivent commenter en direct leurs impressions (comme pour le penser tout haut). Le concepteur du site a ensuite accès aux enregistrements de façon asynchrone. Il s'agit d'une méthode mixte, puisqu'elle peut s'apparenter à une enquête permettant ainsi d'atteindre un plus grand nombre d'utilisateurs, mais également à une étude sur le terrain, puisque le participant peut essayer le site chez lui, sur son lieu de travail etc.

Il est important de noter que les outils d'évaluation de l'utilisabilité peuvent être utilisés pour mesurer des aspects de l'EU d'après la norme ISO 9241 (2010). Cela est logique étant donné que l'utilisabilité et les aspects pragmatiques sont des composantes essentielles de l'EU. L'outil pédagogique de CIRANO comportait de nombreux problèmes d'utilisabilité. Nous allons donc présenter des outils spécifiques à l'utilisabilité dans un premier temps, notamment les évaluations d'experts que nous avons utilisées au début du projet. Ensuite, nous présenterons des outils mesurant de façon plus globale l'EU.

1.3.2 Les évaluations d'experts pour l'utilisabilité

L'utilisabilité des interfaces est un objet d'étude beaucoup plus ancien que l'EU. Ainsi, il existe de nombreuses heuristiques pour faire l'évaluation ergonomique des interfaces et en détecter les failles qui peuvent perturber les utilisateurs. Les heuristiques de Nielsen (1990), les principes de la norme ISO 9241-210 (2010), de même que les critères ergonomiques de Bastien et Scapin (1993) sont les plus communément cités. Nous les examinons en détails dans les sections qui suivent.

1.3.2.1 Les heuristiques de Nielsen et Molich (1990)

Nielsen et Molich ont proposé en 1990 neuf heuristiques de base pour évaluer l'utilisabilité d'un système interactif. Elles ont été créées à partir de leur expérience dans le domaine de l'utilisabilité et avaient pour but de guider l'évaluation d'une interface de façon simple. Comme les listes de lignes directrices pour la conception des interfaces utilisateurs étaient trop longues (944 dans le cas de Smith & Mosier 1986), Nielsen et Molich voulaient proposer une alternative efficace à ces dernières (voir Figure 1-7).

Simple and natural dialogue
Speak the user's language
Minimize user memory load
Be consistent
Provide feedback
Provide clearly marked exits
Provide shortcuts
Good error messages
Prevent errors

Figure 1-7 : Les neuf heuristiques de base de Nielsen et Molich

Les neuf heuristiques de Molich et Nielsen (1990) ont été améliorées en 1994 en les confrontant à des problèmes d'utilisabilité. Dix heuristiques, en grande partie basées sur les anciennes, ont alors été développées (traduction libre) :

1 - Visibilité de l'état du système

2 - Correspondance entre le système et le monde réel

3 - Contrôle de l'utilisateur et liberté

4 - Cohérence et standards

5 - Prévention des erreurs

6 - Reconnaissance plutôt que rappel

7 - Flexibilité et efficience d'utilisation

8 - Design esthétique et minimaliste

9 - Aider les utilisateurs à reconnaitre, diagnostiquer, et rattraper leurs erreurs.

10 - Aide et documentation

Ces heuristiques sont toujours très utilisées de nos jours. Elles constituent une bonne base pour évaluer une interface. Cependant, leur simplicité a un prix : leur faible niveau de spécification peut donner lieu à une bonne part d'interprétation de la part de l'évaluateur.

1.3.2.2 Norme ISO 9241-210 (2010)

La norme ISO 9241, qui est le fruit du travail de plusieurs experts et l'objet d'un consensus international, établit sept principes d'utilisabilité qui se présentent sous forme de questions (traduction libre) :

1 - Est-ce que le dialogue est approprié pour la tâche et le niveau de compétence de l'utilisateur ? (Pertinence pour la tâche)

2 - Est-ce que le dialogue est clair à propos de ce que devra faire l'utilisateur ensuite ? (Autodescription)

3 - Est-ce que le dialogue est cohérent ? (Conformité avec les attentes de l'utilisateur).

4 - Est-ce que le dialogue supporte l'apprentissage ? (Aptitude à l'apprentissage).

5 - Est-ce que le dialogue permet de contrôler la vitesse et la séquence de l'interaction ? (Contrôlabilité).

6 - Est-ce que le dialogue est indulgent ? (Tolérance aux erreurs).

7 - Est-ce que le dialogue peut être personnalisé pour convenir à l'utilisateur ? (Aptitude à l'individualisation).

Une explication détaillée de ces principes est disponible à l'Annexe 2.

Encore une fois, le degré de spécification reste encore assez faible. De plus, il y a des frais d'utilisation de la norme ISO, ce qui est sans doute un frein pour plusieurs utilisateurs potentiels.

1.3.2.3 Critères de Bastien & Scapin (1993)

Bastien et Scapin ont fait la synthèse de quelques 900 recommandations ergonomiques dans le domaine de l'ergonomie informatique. Ils ont abouti à l'élaboration de huit critères principaux composés de 18 sous-critères. Ces critères sont très largement utilisés en ergonomie des interfaces. De plus, des études (ex., Bastien, Scapin, & Leulier, 1999) ont montré que ces critères permettent de révéler davantage de défauts de l'interface en comparaison avec les principes de la norme ISO 9241-110, ou en comparaison avec une évaluation sans critères. Ils sont plus détaillés et plus spécifiques que les heuristiques précédentes, ce qui permet de mieux guider l'expert et de limiter la variabilité entre les experts. Le tableau 1-1 présente la liste des différents critères et sous-critères de Bastien et Scapin.

Tableau 1-1 : Les critères d'évaluation ergonomique de Bastien et Scapin (1993)

1. GUIDAGE	1.1. Incitation 1.2. Groupement/Distinction par le format ou le groupement 1.3. Feedback immédiat 1.4. Lisibilité
2. CHARGE DE TRAVAIL	2.1. Brièveté 2.2. Densité informationnelle
3. CONTRÔLE EXPLICITE	3.1. Actions explicites 3.2. Contrôle utilisateur
4. ADAPTABILITÉ	4.1. Flexibilité 4.2. Prise en compte de l'expérience
5. GESTION DES ERREURS	5.1. Protection contre les erreurs 5.2. Qualité des messages d'erreurs 5.3. Correction des erreurs
6. HOMOGÉNÉITÉ / COHÉRENCE	
7. SIGNIFIANCE CODES & DÉNOMINATIONS	
8. COMPATIBILITÉ	

L'utilisation de sous-critères permet de diminuer le niveau d'abstraction des critères. Une description détaillée de chacun des critères et des sous-critères est disponible à l'Annexe 3. Des justifications permettent de comprendre les conséquences positives du suivi de ces recommandations. Bastien & Scapin précisent qu'ils voulaient que ces recommandations puissent être utilisables également par des non spécialistes du domaine.

Certains critères sont généraux et assez peu spécifiques, par ex. « Compatibilité » ; d'autres sont très détaillés : par ex., le sous-critère « Brièveté » comporte lui-même deux sous-critères, « Concision » et « Actions minimales ».

Dans notre recherche, nous avons choisi d'utiliser ces critères pour évaluer l'utilisabilité de l'interface utilisateur du logiciel d'aide à l'apprentissage de notions d'économie qui est au centre de notre recherche. Nous avons également validé cette évaluation grâce à une enquête auprès des utilisateurs.

1.3.3 Les enquêtes

Bien que l'évaluation des interfaces à partir de critères ergonomiques présente certains avantages, il est toujours intéressant de la compléter à l'aide d'enquêtes ou de tests utilisateurs. Nous avons choisi d'utiliser les enquêtes car cela nous permettait d'obtenir des informations auprès d'un plus grand nombre d'utilisateurs. Il nous a également semblé important d'avoir des données quantitatives pour compléter notre analyse. De plus, comme nous l'avons vu, il n'existe pas de listes de critères reconnus pour évaluer l'EU comme c'est le cas pour l'utilisabilité. Nous présenterons brièvement les questionnaires d'évaluation de l'utilisabilité des interfaces les plus souvent cités, puis nous nous attarderons sur ceux permettant d'évaluer l'EU.

1.3.3.1 Les questionnaires d'évaluation de l'utilisabilité

Un grand nombre de questionnaires ont été développés afin d'évaluer l'utilisabilité des interfaces humain-machine. Voici un aperçu de ceux qui sont les plus souvent cités :

- QUIS : Questionnaire for user interface satisfaction (Chin et al., 1988).
- PUEU : Perceived usefulness and ease of use (Davis, 1989).
- NHE : Nielsen's heuristic evaluation (Nielsen, 1993).
- NAU : Nielsen's attributes of usability (Nielsen, 1993).
- CSUQ : Computer system usability questionnaire (Lewis, 1995).
- ASQ : After scenario questionnaire (Lewis, 1995).
- PUTQ : Purdue Usability Testing Questionnaire (Lin et al. 1997).
- USE : USE questionnaire (Lund, 2001).
- MUMMS : Measurement of usability of multimedia software (Human Factors Research Group, 2002)

Ces questionnaires présentent généralement des affirmations (ex., « Le système est facile à utiliser »), suivies de cases à cocher placées sur un continuum allant du désaccord à l'accord total et correspondant à une échelle de Likert en n points. Une alternative à ce type d'échelle est l'utilisation, aux extrémités du continuum, de paires de mots au sens opposé, tels que « Bon-Mauvais » (échelle d'Osgood). Certains problèmes bien connus dans la littérature sont pourtant récurrents dans ces questionnaires. C'est le cas par exemple du « biais de l'acquiescement » qui résulte du fait que les participants ont plus souvent tendance à être d'accord avec une question formulée positivement. En effet, les participants ayant un doute devant une question particulière ont plus tendance à se dire d'accord avec celle-ci (Watson, 1992). De plus, certains questionnaires ont des coûts d'utilisation (ex., MUMMS). La longueur de ces différents questionnaires varie de 3 à 100 questions. Les plus courts (ex., ASQ, PUEU) prétendent mesurer l'utilisabilité, mais ne permettent malheureusement pas de dévoiler les différents problèmes de façon détaillée; l'ASQ par exemple inclut uniquement trois questions sur la facilité, le temps et le support d'information pour mener à bien les tâches.

Certains questionnaires comme le NHE (Nielsen, 1993) sont les descendants des évaluations d'experts, cependant nous n'avons trouvé aucun questionnaire se basant sur les critères de Bastien & Scapin.

Peu de nouveaux questionnaires d'évaluation de l'utilisabilité voient le jour actuellement étant donné que l'évaluation s'est déplacée vers l'EU. Le support au questionnaire MUMMS est suspendu étant donné que les auteurs (Human Factors Research Group) travaillent sur une version visant à évaluer l'EU (WAMMI, 2013).

1.3.3.2 Les enquêtes pour l'expérience utilisateur

Bien que le concept de l'EU soit relativement nouveau, le nombre d'outils d'évaluation de l'EU est en constante augmentation. L'engouement des chercheurs pour ce thème a donné naissance à de nombreux questionnaires d'évaluation. Le site « All about UX », maintenu par des experts dans le domaine (tels que Roto), recense quelques-uns de ces questionnaires. Beaucoup d'entre eux sont à utiliser avec un système bien particulier, ou ne se focalisent que sur une partie de l'EU. C'est le cas de certains tests (ex., Emocard) permettant d'évaluer nos émotions (Desmet Overbeeke & Tax, 2001). La liste de questionnaires du site « All about UX » n'est pas exhaustive puisque certains questionnaires très cités dans la littérature ne sont pas présents, tels que le UEQ

(User Experience Questionnaire ; Laugwitz, 2008). Nous nous concentrerons en particulier sur un des questionnaires les plus cités dans la littérature : AttrakDiff.

1.3.3.3 AttrakDiff

AttrakDiff est un outil permettant d'évaluer rapidement l'EU vécue avec un système interactif. Il est gratuit et a été développé par Hassenzahl, Burmester et Koller (2003). Il permet d'évaluer les qualités pragmatiques et les qualités hédoniques d'un produit. Il est basé sur le modèle d'EU d'Hassenzahl que la figure 1-8 présente.

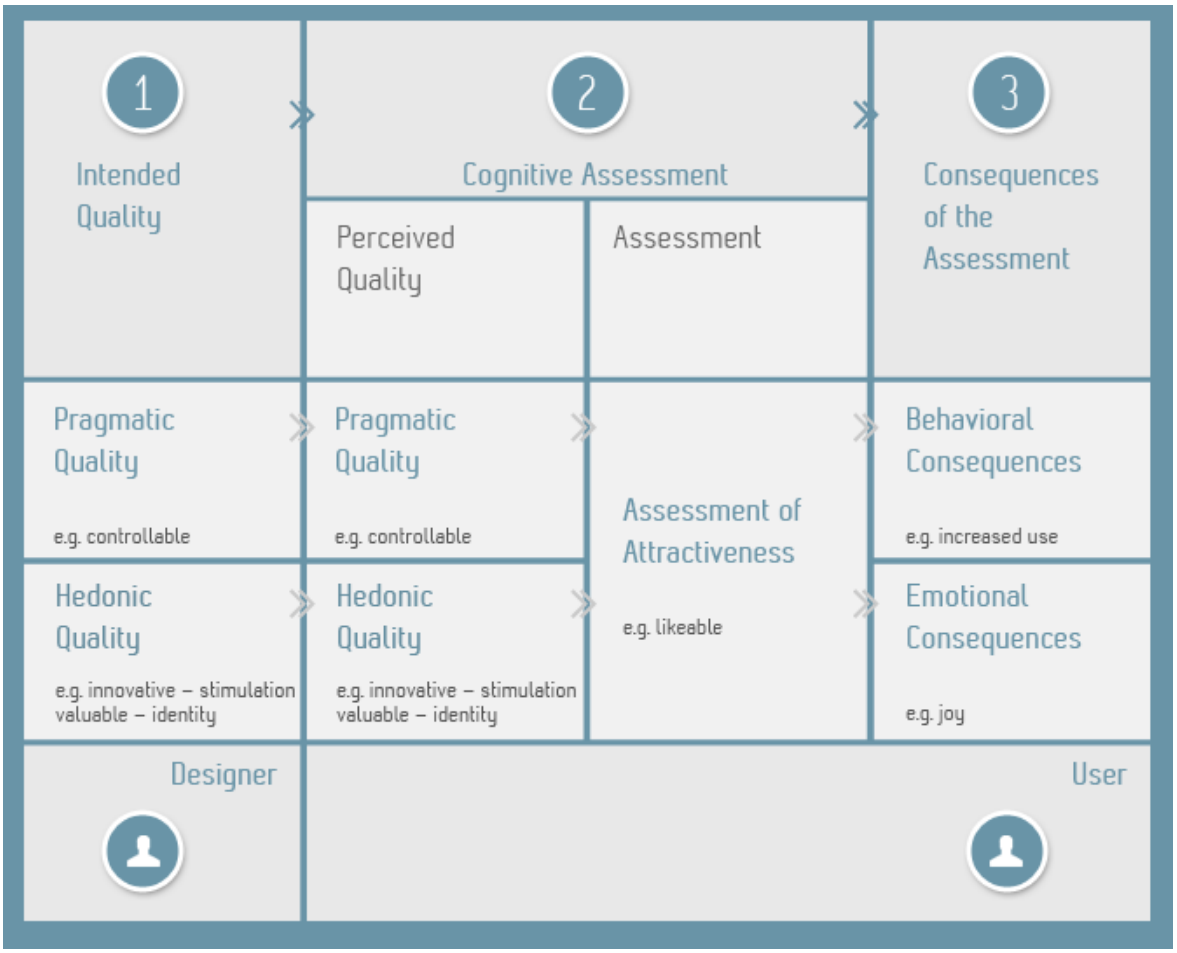


Figure 1-8 : Le modèle d'EU supportant Attrakdiff (extrait de http://attrakdiff.de)

Dans la colonne de gauche, on retrouve la qualité du produit voulue par le designer. La qualité du produit perçue et l'évaluation de celle-ci par l'utilisateur vont avoir des conséquences émotionnelles et comportementales.

L'outil permet donc d'évaluer les qualités pragmatiques du produit (QP : Pragmatic Qualities) (liées à l'utilisabilité) ainsi que ses qualités hédoniques (QH : Hedonic Quality). Une évaluation de l'attractivité globale (ATT), découlant des qualités QH et QP, a été ajoutée. Au cours des essais de la première version du produit, la nécessité de séparer les qualités hédoniques en deux aspects est apparue :

L'aspect stimulation des QH : il est lié aux attributs du produit permettant à l'utilisateur de se développer, d'évoluer, de voir des choses nouvelles et stimulantes.

L'aspect identité des QH : il est lié aux attributs du produit nous permettant de nous identifier à lui dans un contexte social.

Le questionnaire est composé de 28 paires de mots opposés (ce qui correspond à une échelle d'Osgood), afin que l'utilisateur choisisse dans chaque paire le mot qui décrit le mieux le produit. L'échelle de notation comporte sept points entre un mot et son opposé (ex. : Laid-Beau). Les 28 paires de mots sont réparties en quatre groupes (voir Figure 1-10).

Qualités hédoniques - Stimulation (QH-S)	Qualités pragmatiques (QP)
Conventionnel – Inventif Non imaginatif – Imaginatif Prudent – Audacieux Conservateur – Innovateur Ennuyant – Captivant Peu exigeant – Stimulant Ordinaire – Novateur	Technique – Humain Compliqué – Simple Non pratique – pratique Encombré – Direct Imprévisible – Prévisible Confus – Structuré Ingérable – Gérable
Qualités hédoniques - Identité (QH-I)	**Attractivité (ATT)**
Isolant – Connectant Non professionnel – Professionnel De mauvais goût – Élégant Bas de gamme – Haut de gamme Aliénant - Intégrant M'éloigne – M'approche	Déplaisant – Plaisant Laid – Beau Désagréable – Agréable Rejetant – Invitant Mauvais – Bon Repoussant – Attirant Décourageant – Motivant

Figure 1-9 : Les 28 paires de mots de AttrakDiff (Traduction libre)

Hassenzahl (2004) précise que les qualités hédoniques et les qualités pragmatiques sont indépendantes les unes des autres et contribuent de façon égale à l'attractivité globale du produit.

Hassenzahl et Monk (2010) proposent une version réduite qui conserve les éléments importants du modèle (voir Figure 1-11); cette réduction était nécessaire puisque les participants de leur étude devaient essayer plusieurs sites web. La distinction entre les qualités hédoniques de stimulation et d'identité a alors été retirée. Quatre paires de mots permettent de mesurer les qualités pragmatiques, et il en va de même pour les qualités hédoniques. L'attractivité globale se mesure en évaluant si le produit est *bon/mauvais* et *beau/laid.*

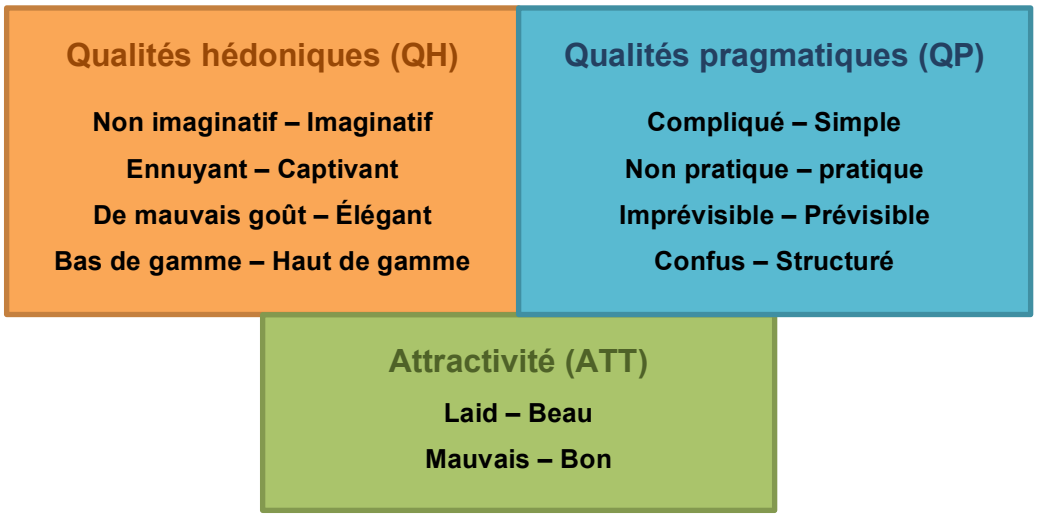


Figure 1-10 : Version réduite de l'outil AttrakDiff (Traduction libre)

Nous utiliserons cet outil pour évaluer l'EU vécue par les utilisateurs du logiciel d'aide à l'apprentissage des notions d'économie que nous avons contribué à développer. Il est important de noter qu'avec cet outil, l'utilisateur doit se remémorer son expérience, et il se peut donc que l'on manque certaines informations. Cependant, Hassenzahl précise dans un article de 2013 qu'il accorde plus d'importance à l'EU mémorisée qu'à l'EU vécue immédiatement. En effet, on peut penser que ce qui nous pousse à recommander un produit, à un instant donné, est le souvenir de l'expérience que l'on a eue avec celui-ci. Même si l'EU mémorisée évolue avec le temps (Karapanos 2013), c'est celle-ci qui va nous permettre de savoir si l'on a été satisfait avec un produit, si l'on veut le réutiliser ou même le recommander.

Ce chapitre nous a permis de présenter les principales définitions et les différents modèles de l'utilisabilité et de l'EU. Nous avons également mis en avant les méthodes d'évaluation de l'utilisabilité et de l'EU, ce qui nous a permis d'identifier les outils que nous utiliserons dans ce mémoire. Ceci nous permet maintenant d'exposer le contexte dans lequel ces différents éléments seront utilisés ainsi que les objectifs de ce mémoire.

CHAPITRE 2 CONTEXTE ET OBJECTIFS

Ce chapitre présente le contexte dans lequel ce mémoire a été effectué. Nous présenterons la première version de l'outil pédagogique du CIRANO, et plus particulièrement une expérience pédagogique nommée « Ressource commune ». Nous ferons un bilan de cette première version. Enfin, nous énoncerons les objectifs qui ont découlé de ce bilan et qui sont à la base de l'objectif général de notre recherche.

2.1 Contexte : Présentation de l'outil pédagogique du CIRANO

Le CIRANO est un centre universitaire de recherche du Québec regroupant plus de 180 chercheurs-professeurs. Il œuvre dans des domaines variés tels que l'économie, la finance, la sociologie, l'informatique, la médecine, etc. Il publie de nombreux travaux scientifiques chaque année. Il est également doté d'un laboratoire d'économie expérimentale. Celle-ci est une science récente ayant fait l'objet d'un prix Nobel en 2002 (Daniel Kahneman et Vernon Smith). Elle consiste à faire participer des individus ou des groupes à des expériences économiques afin d'observer leurs comportements et de les analyser par la suite. C'est dans ce laboratoire d'économie expérimentale que l'outil pédagogique du CIRANO a vu le jour.

Des études ont montré le faible niveau de connaissance en économie et en finance de la population québécoise par rapport à celle du reste du Canada (Lalime et Michaud, 2012). Pourtant, l'économie occupe une place importante dans la vie de chaque individu qui doit alors prendre de nombreuses décisions économiques et financières, sans forcément être correctement préparé pour celle-ci. La disparition des cours d'éducation économique des écoles secondaires du Québec en 2009 vient s'ajouter à cela. Ainsi, en 2012**,** le CIRANO a lancé un projet de développement d'expériences pédagogiques interactives sur l'économie et la finance, sous la direction de Claude Montmarquette et de Marine DeMontaignac.

L'approche théorique pour l'apprentissage de notions d'économie est parfois trop abstraite. Le lien entre la théorie et la réalité est parfois obscur. Ainsi, ce projet veut permettre aux étudiants de secondaire V et de CEGEP de se confronter directement à ces notions par le biais de l'économie expérimentale. Le but, à terme, est de leur permettre de faire des choix éclairés en matière

d'économie dans leur vie personnelle, et de mieux comprendre les problématiques économiques et financières auxquelles leur société est confrontée.

Ces expériences pédagogiques sont disponibles sur ordinateur et sur tablette interactive et permettent à l'étudiant d'expérimenter diverses notions d'économie et de finance : il peut prendre part à un marché de biens, découvrir son niveau de tolérance au risque, expérimenter le fonctionnement de l'épargne et de l'assurance, etc. Pour le moment, le CIRANO se concentre sur les étudiants de secondaire V et de CEGEP dans son développement d'outils adaptés ; à terme, il vise le grand public.

Certaines alternatives aux expériences du CIRANO telles que Moblab (www.moblab.com) existent sur le marché, mais ne sont pas disponibles en français et elles ciblent une population plus âgée et possédant de meilleures connaissances en économie et en finance.

L'outil du CIRANO, sans nom pour le moment, est actuellement composé d'une interface d'administration permettant de choisir, de configurer et de lancer les expériences pédagogiques. Même s'il s'agit d'une partie considérable de l'outil et qu'un travail important a été effectué sur cette dernière, nous ne la détaillerons pas dans ce mémoire. L'outil pédagogique comporte 12 expériences[2], mais par souci de concision, nous allons nous limiter à la présentation d'une seule expérience nommée « Ressource commune ». Nous avons choisi cette expérience car il s'agissait de la plus stable et la plus testée à l'interne. Les autres expériences nécessitaient encore des phases de tests à l'interne avant de pouvoir être évaluées à l'externe.

2.2 L'expérience « Ressource commune »

Les expériences pédagogiques sont généralement composées de trois pages. La page de consigne *présentant le jeu et les directives à suivre*, la page de jeu, *avec laquelle le participant interagit*, et enfin la page de résultat, *qui présente des résultats personnels et/ou sur le groupe.* Certaines expériences peuvent être réalisées si on est seul, d'autres nécessitent que plusieurs

---

[2] Les 12 expériences disponibles se nomment respectivement : Assurance, Bien public, Bulle spéculative, Double enchère, Épargne, Fraude fiscale, Loteries, Pollution, Ressource commune, Risque et rendements, Soldes et Ultimatum.

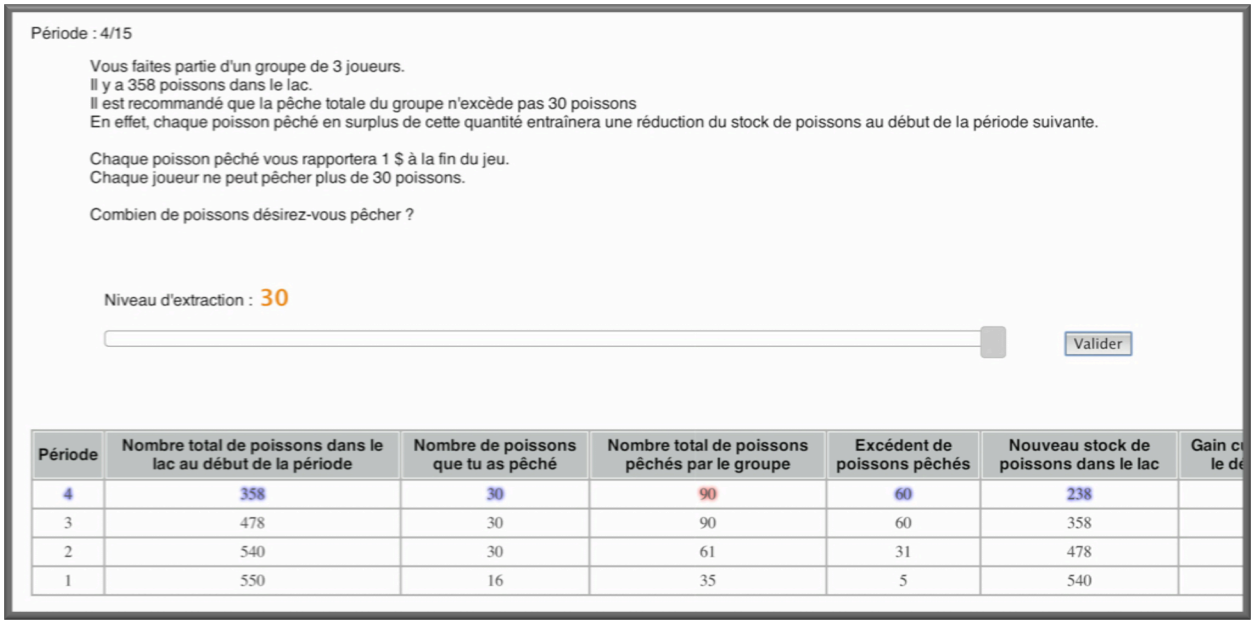
participants connectés sur le serveur. Nous présenterons ci-dessous la page de jeu de la version originale de l'expérience « Ressource commune ».

2.2.1 Ressource commune

L'expérience « Ressource commune » se joue à plusieurs. Elle permet d'appréhender le concept du « bien commun ». Dans cette expérience, chaque participant possède un bateau de pêche qui fait partie d'un groupe de bateaux de pêche dont il connait le nombre, mais avec lesquels il ne peut pas communiquer. La ressource commune est ici un lac comportant un stock initial de poissons que les participants connaissent. Chaque poisson pêché est bien sûr une source de revenus pour le pêcheur. Le fonctionnement est le suivant : il y a trois bateaux sur le lac. Chaque propriétaire de bateau peut pêcher jusqu'à 30 poissons. Cependant pour éviter d'épuiser trop rapidement le stock de poissons, les trois propriétaires de bateaux ne doivent pas pêcher plus de 30 poissons au total. Il s'agit d'une recommandation ; si le groupe pêche plus de 30 poissons, alors le stock de poissons baissera plus vite en fonction de l'excédent pêché. Le jeu se déroule sur plusieurs périodes, qui peuvent dans cette expérience être assimilées à des jours de pêche. Ainsi, les propriétaires de bateau doivent restreindre leur prise de poissons à chaque période (ou jour de pêche), quitte à diminuer leurs gains, pour ne pas épuiser le stock du lac trop rapidement.

Cette expérience permet d'expliquer la « tragédie des biens communs ». Il s'agit d'une théorie économique développée par Garrett Hardin en 1968. Celle-ci explique qu'un individu rationnel agissant pour son propre intérêt crée une relation *perdant-perdant* sur le long terme avec le groupe dont il fait partie. Cette relation *perdant-perdant* est ici illustrée par la surexploitation du lac, et donc la diminution du stock de poissons. Ce concept permet d'expliquer ensuite la nationalisation ou la privatisation de certaines ressources communes, ou permet également de présenter des problématiques plus générales telles que le réchauffement climatique.

La figure 2-1 présente un aperçu de l'interface de la page de jeu de l'expérience « Ressource commune ».



Période	Nombre total de poissons dans le lac au début de la période	Nombre de poissons que tu as pêché	Nombre total de poissons pêchés par le groupe	Excédent de poissons pêchés	Nouveau stock de poissons dans le lac	Gain cu le dé
4	358	30	90	60	238	
3	478	30	90	60	358	
2	540	30	61	31	478	
1	550	16	35	5	540	

Figure 2-1 : Aperçu de l'expérience « Ressource commune » dans l'ancienne version

En haut de la page, des informations telles que le stock de poissons du lac ou la période en cours sont présentées. Un rappel des consignes est également présent. Au milieu de l'écran, un bouton à glissière permet de choisir le nombre de poissons que le participant désire pêcher durant la période. Enfin, en bas de la page un tableau récapitulatif affiche un bilan des périodes précédentes. Il comporte un nombre important de colonnes et présente pour chaque période précédente le nombre de poissons dans le stock initial, le nombre de poissons pêchés par le participant et par le groupe, l'excédent de poissons pêchés, ainsi que les gains empochés.

Le participant est invité à lire la consigne puis est dirigé vers la page de jeu. Chaque jour, il choisit à l'aide du bouton à glissière le nombre de poissons à pêcher, puis confirme son choix (nombre de poissons pêchés). Une fois que tous les participants ont validé leur choix, les résultats de la journée de pêche apparaissent dans le tableau récapitulatif. À la fin des 15 journées de pêche, la page de résultat présente les résultats du groupe. Si celui-ci épuise le stock de poissons avant la fin du jeu, il ne peut plus pêcher jusqu'à la fin.

2.3 Bilan du premier outil

Le premier outil a été développé dans le but d'être d'abord fonctionnel. Il s'agissait d'une première version permettant de poser les bases de l'outil final. Il a permis aux utilisateurs d'essayer une dizaine d'expériences pédagogiques. Il a également servi à pointer les différentes corrections à apporter. Certains bogues mineurs étaient présents, le lancement des expériences était parfois fastidieux, mais l'outil était fonctionnel.

L'équipe était jusqu'à présent composée d'un directeur de projet, de plusieurs économistes et de développeurs. Le CIRANO souhaitait obtenir un engagement fort de la part des utilisateurs, et le besoin de faire appel à un spécialiste de l'expérience utilisateur est donc apparu assez tôt.

Les essais en interne avaient mis en avant la nécessité d'obtenir une interface plus intuitive, attirante et facile à utiliser. Les différents organismes finançant le projet soulignaient également le besoin de travailler sur ces éléments. D'où notre implication dans le projet depuis août 2013.

2.4 Objectif

L'objectif général de notre recherche est d'améliorer l'expérience des utilisateurs avec l'outil pédagogique du CIRANO; nous allons nous limiter à l'expérience pédagogique « Ressource commune » offerte par l'outil actuel en misant sur le fait que si on peut améliorer l'EU avec une partie de l'outil, on pourra alors le faire avec les autres parties en suivant les mêmes principes. Pour atteindre cet objectif, nous avons réalisé les trois activités suivantes :

- Évaluer l'interface de l'outil actuel;
- Concevoir et développer une nouvelle interface;
- Tester la nouvelle interface auprès des utilisateurs.

Les deux premières activités seront présentées dans le chapitre 3. L'essai et la validation de la nouvelle interface seront détaillés dans le chapitre 4. Enfin, les résultats obtenus et leur analyse seront présentés dans le chapitre 5.

CHAPITRE 3 PROCESSUS D'ÉVALUATION ET DE CONCEPTION DE L'INTERFACE D'UN OUTIL D'AIDE À L'APPRENTISSAGE

Dans ce chapitre, nous présentons le processus que nous avons suivi pour faire l'évaluation ergonomique de l'interface utilisateur de l'outil actuel et pour concevoir la nouvelle interface de l'outil, et plus particulièrement celle de l'expérience pédagogique « Ressource commune ». Enfin, nous détaillerons succinctement la phase de développement.

3.1 Processus d'évaluation ergonomique de l'interface de l'outil actuel

À notre arrivée sur le projet, nous avons assisté à de nombreuses présentations sur celui-ci. Nous avons utilisé l'outil d'administration et de configuration et fait quelques expériences pédagogiques supportées par l'outil. Nous avons identifié les utilisateurs principaux de l'outil qui sont les étudiants de secondaire V et de CEGEP, ainsi que les professeurs des cours d'économie. Nous avons également pu voir le contexte d'utilisation de l'outil en assistant à une session en classe avec une quinzaine d'élèves de CEGEP, essayant les expériences « Bien public » et « Ressource commune ». La session a duré une heure. Cela était essentiel pour bien évaluer l'outil actuel. Après cette phase, nous avons pu commencer notre évaluation de l'outil actuel.

L'essai de l'outil dans sa version originale nous a permis de déceler de nombreuses failles ergonomiques qui sont autant de freins à une bonne EU. Cette expérience personnelle nous a été très utile pour faire l'évaluation ergonomique de l'interface de l'outil d'administration et de configuration des expériences, ainsi que de l'interface des 12 expériences pédagogiques. Comme nous l'avons précisé plus haut, pour des raisons de concision, nous ne détaillons qu'une seule des expériences pédagogiques.

De plus, le CIRANO nous a fourni une synthèse des questionnaires qu'il avait administrés pour évaluer certaines expériences dans des classes de CEGEP d'une trentaine d'élèves environ (Pour un total de près de 600 élèves). Ces questionnaires demandaient surtout d'expliquer la stratégie qu'il avaient suivie durant le jeu, mais ils comportaient également des questions ouvertes demandant aux élèves de préciser ce qu'ils avaient aimé dans l'expérience, les difficultés auxquelles ils avaient été confrontés, ainsi que des suggestions d'améliorations. Ces retours d'information qualitatifs nous ont été fournis pour nous aider dans l'évaluation. À ce stade, nous

n'avions pas pu avoir accès directement aux utilisateurs (professeurs ou élèves). Ces retours d'informations ont été néanmoins été utiles pour faire l'évaluation ergonomique de l'interface de l'outil actuel.

3.2 Évaluation ergonomique de l'interface

L'utilisabilité joue un rôle de pilier dans l'EU : elle peut être considérée comme un facteur d'hygiène (Herzberg, 1976), c'est-à-dire pouvant causer de l'insatisfaction de l'utilisateur, mais non suffisante pour créer une EU positive. Il est donc primordial que les frustrations liées à l'utilisation du produit soient réduites considérablement et si possible inexistantes. L'évaluation heuristique de l'interface basée sur les critères de Bastien & Scapin nous a semblé être un bon moyen d'identifier et de classer de façon méthodique les différents problèmes ergonomiques de l'interface.

3.2.1 Méthodologie d'évaluation ergonomique

Les expériences pédagogiques comportent un faible nombre de pages (3 pages). Nous avons donc imprimé et assimilé la grille des critères de Bastien & Scapin puis chaque expérience était testée suivant différentes approches.

3.2.1.1 Approche d'évaluation

L'approche d'évaluation que nous avons suivie pour trouver le maximum de problèmes comprend quatre étapes d'analyses :

1) Analyse globale (vision extérieure au projet) : nous commencions par essayer l'expérience pédagogique au complet, sans nous être renseigné sur le concept économique sous-jacent, les limitations techniques et autres éléments liés à la conception de l'expérience. Ceci nous permettait d'être dans une situation semblable à un utilisateur découvrant l'expérience pédagogique et d'avoir une vision différente. Nous essayions donc l'expérience, et répertoriions les différents problèmes relevés une fois l'expérience terminée. Ceci nous permettait de nous faire une première idée des failles de l'expérience. Bien que cette méthode nous demandait un effort rétrospectif, nous avions une vision du fonctionnement « normal » de l'expérience sans interruption.

2) Analyse détaillée (vision extérieure au projet) : la deuxième analyse était plus détaillée. Il s'agissait alors d'essayer à nouveau la même expérience, mais de faire une interruption à chaque page et à chaque interaction afin de noter les différentes failles identifiables avec la grille de Bastien & Scapin. À chaque interruption, nous avons effectué une capture d'écran afin qu'elle soit accessible pour le reste de l'équipe.

3) Analyse détaillée (vision interne au projet) : un premier retour d'information était fourni à la directrice de projet. Nous présentions les différentes failles ergonomiques et autres problèmes que nous avions décelés. À partir de ce stade, nous nous empreignions du concept économique sous-jacent. Nous demandions également des informations sur le but de l'expérience, les limitations techniques et les causes des choix de conception qui ont été faits. Ceci nous permettait dans un premier temps d'affiner les problèmes trouvés, de découvrir de nouvelles failles, ou tout simplement de modérer certains problèmes liés à des contraintes économiques (certains éléments ne pouvaient pas être modifiés ou simplifiés d'après les économistes).

4) Analyse en condition inhabituelle (vision interne au projet) : nos connaissances en tant que développeur nous ont permis de nous rendre compte du nombre important de problèmes qui peuvent survenir lorsque le système n'est pas utilisé suivant une « utilisation normale ». En effet, de nombreuses exceptions doivent être prises en compte au niveau de l'interface : l'utilisateur peut décider de laisser un champ vide, de taper des caractères spéciaux qui n'étaient pas prévus, de cliquer de nombreuses fois sur un même bouton ou sur divers éléments, de ne pas suivre les instructions, etc. Cette dernière analyse consistait donc à tester ce type de comportement, ce qui permet dans un premier temps de voir la robustesse de l'interface face à une utilisation inhabituelle, et dans un second temps, d'observer la gestion des erreurs du système.

3.2.1.2 Présentation des problèmes ergonomiques

Après avoir terminé ces quatre évaluations, les problèmes étaient répertoriés et classés. Nous avons choisi de les classer par page de l'expérience pédagogique : page de consigne, page de jeu, page de résultats ; de plus nous classions les problèmes en leur attribuant un score de sévérité. Voici comment nous avons choisi de présenter les problèmes :

Titre : Une brève description du problème était offerte.

Élément de l'interface et critères non satisfaits : nous identifions l'élément de l'interface affecté et les critères de Bastien et Scapin qui n'étaient pas satisfaits pour ce problème. Ces critères étaient disponibles en annexe de chaque rapport ergonomique afin de permettre au lecteur d'accéder à leur définition.

Cause : la cause du problème et le contexte d'apparition étaient clairement expliqués.

Effet sur la tâche et l'utilisateur : les conséquences du problème étaient mises en avant afin de faire comprendre à l'équipe de développement pourquoi il était important d'y remédier.

Proposition de correction : une ou plusieurs propositions de corrections étaient faites afin d'amorcer la réflexion sur la résolution du problème.

Sévérité : enfin, afin de pouvoir cibler rapidement les problèmes les plus sévères, un score de sévérité sur 15 leur était attribué. Ce score prenait en compte trois éléments distincts notés respectivement de 1 à 5 : *le nombre d'utilisateurs* touchés par ce problème, *l'importance de la tâche* où le problème apparait, et *l'impact* du problème sur la tâche. Chaque score de 1 à 5 est un score heuristique attribué par l'évaluateur; par ex., pour le nombre d'utilisateurs, un score de 1 signifie que peu d'utilisateurs sont affectés. Un score de 5 signifie que la totalité des utilisateurs est affectée.

Voici un exemple de boite à dialogue utilisée dans l'interface de l'outil actuel (figure 3-1) :

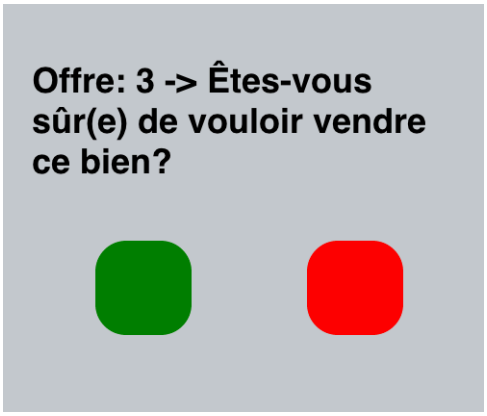


Figure 3-1 : Boite de dialogue utilisée dans l'outil actuel

La distinction entre l'action de valider et l'action d'annuler ne se fait qu'à l'aide d'un code couleur, ce qui est fortement déconseillé. Les utilisateurs daltoniens (plus de 8 % des hommes, et environ 0,5 % des femmes d'après Gegenfurtner & Sharpe, 2001) sont donc bloqués à cette étape. Il s'agit donc ici d'un problème de sévérité majeure, même si la proportion d'utilisateurs affectés est relativement faible:

Nombre d'utilisateurs : 1/5 (très peu d'utilisateurs), Importance de la tâche : 5/5, Impact : 5/5 (bloque la tâche). **Le score heuristique de sévérité est donc de 11/15.**

3.2.2 Bilan des évaluations ergonomiques

Parmi les 12 expériences pédagogiques, neuf étaient présentes à notre arrivée (trois ont été conçues par la suite). Sur ces neuf expériences, nous avons commencé la conception des nouvelles versions après avoir fait un bilan de six expériences principales (pour avoir une vue d'ensemble); les trois expériences restantes étaient très semblables au niveau des interactions et il était important de commencer la conception le plus tôt possible au vu des délais de développement. Nous allons donc présenter le bilan des rapports de six expériences principales qui nous ont guidés pour la conception du nouveau système.

Parmi ces six expériences principales, 87 problèmes ont été identifiés. Le tableau récapitulatif 3-1 présente les différents problèmes classés par fréquence d'apparition dans les expériences. Par exemple, une fréquence d'apparition de 3/6 signifie que le problèmes est apparu dans trois expériences, parmi les six expériences analysées. Il est intéressant de regarder particulièrement les fréquences d'apparition :

Tableau 3-1 : Problèmes ergonomiques des interfaces utilisateur des six expériences pédagogiques évaluées classés par fréquence

Fréquence	Problèmes	Sévérité (sur 15)
6/6	Utilisation répétée des boites de dialogue	10
6/6	Manque d'homogénéité/cohérence	9
6/6	Identification : dénomination non claire, titre non pertinent, reconnexion difficile	7
6/6	Manque d'adaptabilité à la taille de la fenêtre	6
6/6	Nombre de caractères par ligne (CPL) trop élevé	5
5/6	Densité d'information des tableaux trop élevée	9
5/6	Première période et rappel trop rapides	8

Tableau 3-1 : Problèmes ergonomiques des interfaces utilisateur des six expériences pédagogiques évaluées classés par fréquence (Suite et fin)

5/6	Apparition retardée du bouton « commencer »	7
5/6	Titres des messages d'erreurs non pertinents	5
5/6	Messages d'attente non pertinents (joueur seul). File d'attente non visible.	5
5/6	Confusion dans la capture d'écran de consigne	4
4/6	Bogues divers (déconnexion, informations erronées, etc.)	5
4/6	Incohérence mineure de dénominations	4
3/6	Curseur bloqué au début d'une nouvelle période	10
3/6	Nécessité de continuer toutes les périodes lors de l'élimination	9
2/6	Contrôle utilisateur faible lors du récapitulatif (trop rapide)	9
2/6	Changement d'état de l'expérience peu visible	6
2/6	Prévention des erreurs inexistantes (achat à la perte, validation des loteries)	5
1/6	Consigne trop longue et trop dense (double enchère)	11
1/6	Choix par couleurs non approprié (daltoniens)	11
1/6	Manque d'affordance (texte cliquable sans distinction)	11
1/6	Confusion du titre et axes du graphique	7
1/6	Appartenance des contrôles des graphiques non explicites	7
1/6	Résultats non explicites (Ambigüité du joueur et du groupe)	6
1/6	Informations liées présentées de façons dissociées (Numéro de période et temps restant)	3

On voit qu'un grand nombre de problèmes sont récurrents à plusieurs expériences. Au lieu de nous concentrer sur une expérience pédagogique en particulier, nous avons mené une réflexion de plus haut niveau : en proposant une structure globale qui résoudrait ces problèmes récurrents, nous cherchions à améliorer toutes les expériences à la fois. Puis dans un second temps, nous pourrions nous concentrer sur chacune des expériences.

Pour des raisons de longueur, nous détaillerons seulement les cinq problèmes communs à l'ensemble des expériences pédagogiques (fréquence de 6/6). Ces problèmes récurrents nous ont particulièrement guidés pour la conception d'une structure globale de plusieurs expériences, notamment dans « Ressource commune ». Il est donc important de les expliciter.

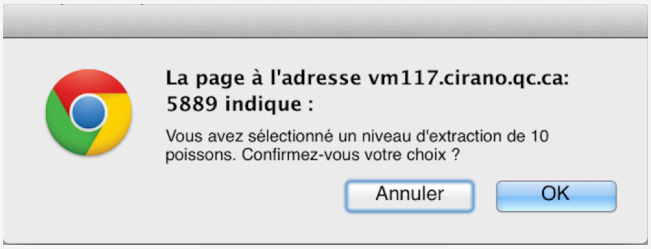
3.2.3 Détails des problèmes ergonomiques communs à l'ensemble des expériences

Ces cinq problèmes sont extraits des rapports d'évaluation ergonomique que nous avons produits pour le CIRANO. Ils sont communs à l'ensemble des expériences. Ils sont présentés sur la figure 3-2 (sur trois pages) de la même façon que dans nos rapports.

#1 : Boite de dialogue *(Charge de travail - Critère 2)*

L'utilisation répétée des boites de dialogue augmente la charge de travail.

Cause : Chaque fois que l'on choisit un nombre de poissons à pêcher (action principale), il faut cliquer sur valider, puis une boite de dialogue intervient pour confirmer notre choix. Les boites de confirmation sont recommandées pour les actions destructives ou lourdes en conséquence. Ici, cette boite apparait 15 fois, ce qui augmente la charge de travail et ne respecte pas le concept des actions minimales. D'autre part, le titre de la boite de dialogue est complexe et non pertinent pour l'utilisateur :

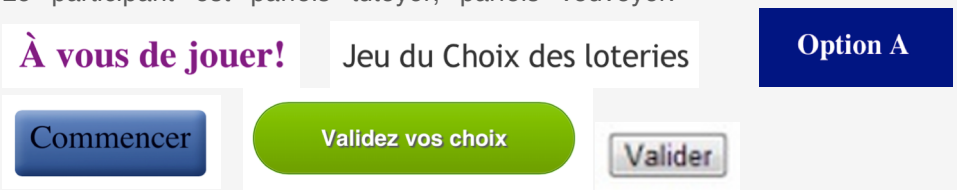


Effet sur l'utilisateur : Présentation d'une information non pertinente et augmentation de la charge de travail.
Effet sur la tâche : Perturbe la tâche et ralentit le processus.
Sévérité : 10/15 – Impact : 1/5 – Importance de la tâche : 4/5 – Nombre d'utilisateurs : 5/5
Proposition de correction : Retirer la boite de confirmation. Prévoir une action d'annulation..

#2 : Boutons, textes, tableaux *(Homogénéité/cohérence - Critère*

L'interface manque d'homogénéité et de cohérence

Cause : Les éléments de l'interface varient entre les expériences, et à travers une même expérience. La façon de présenter les éléments pour des contextes identiques varie (ex. gain final). Le participant est parfois tutoyer, parfois vouvoyer. .



Effet sur l'utilisateur : Le manque de cohérence est une raison importante du refus d'utilisation d'une interface. La prévisibilité du système et l'apprentissage sont diminués lorsque les éléments visuels varient trop d'une page à une autre ; l'utilisateur perd ses repères visuels.
Effet sur la tâche : Les éléments graphiques sont mieux reconnus et utilisés lorsqu'ils sont stables d'un écran à l'autre. Un manquement à ce critère augmente le nombre d'erreurs et réduit la compréhension.
Sévérité : 9/15 – Impact : 1/5 – Importance de la tâche : 3/5 – Nombre d'utilisateurs : 5/5
Proposition de correction : L'homogénéité et la cohérence de l'interface seront apportées par les maquettes graphiques.

#3 : Boite d'identification *(Guidage et dénomination - Critère 1 et 7)*

Les informations d'identification demandées ne sont pas explicites.

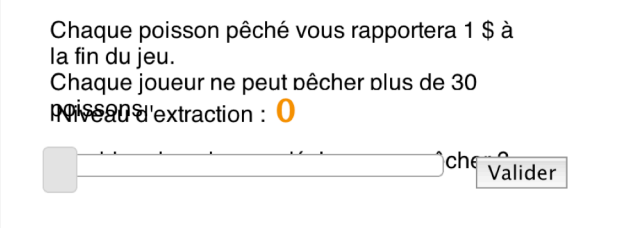
Cause : La boite d'identification avant chaque expérience, demande de remplir le champ "identification". Les utilisateurs novices (ou de la version grand public) ne sauront pas si il s'agit d'une adresse mail, d'un surnom, ou d'un nom. Le titre de la boite n'est pas pertinent. Si l'identifiant est oublié, le candidat ne peut se reconnecter.

Effet sur l'utilisateur : Présentation d'une information non pertinente, et ambigüité
Effet sur la tâche : Augmente le risque d'erreur et ralentit la tâche.
Sévérité : 7/15 – Impact : 1/5 – Importance de la tâche : 4/5 – Nombre d'utilisateurs : 2/5
Proposition de correction : Indiquer clairement l'information à fournir. Changer le titre de la boite. Préciser que cette information doit être retenue.

#4 : Zones de texte et tableaux *(Adaptabilité- Critère 4)*

L'interface ne s'adapte pas aux différentes tailles de la fenêtre.

Cause : Lors du redimensionnement de la fenêtre, certains éléments se chevauchent. Il est essentiel que l'interface s'adapte au support de l'utilisateur pour obtenir une meilleure expérience. Si à terme, les expériences sont proposées au grand public, certains utilisateurs ne pourront pas l'utiliser. (Voir capture suivante)



Sévérité : 6/15 – Impact : 4/5 – Importance de la tache : 1/5 – Nombre d'utilisateurs : 1/5
Proposition de correction : Prévoir l'adaptabilité de l'interface au niveau de l'implémentation.

#5 : Zones de texte *(Lisibilité - Critère 1.4)*

Le nombre de caractères par ligne est trop élevé.

Cause : La taille des caractères à l'écran est suffisante (il faut veiller à ce qu'elle soit d'au moins 16 pt). Cependant, le nombre de caractères par ligne (CPL) influe sur la lisibilité. Les standards de l'ergonomie web conseillent d'avoir 55 à 75 caractères par lignes. Cependant, sur la capture d'écran ci-dessous, le nombre de caractères par ligne est de 155 (bien au-dessus des critères) !

Exemple (non lisible, mais bon aperçu de la longueur) :

Si, au fur et à mesure des périodes de jeu, le nombre de poissons dans le lac devient plus petit que le niveau recommandé de pêche pour le groupe, alors les membres ne pourront plus pêcher.

Effet sur l'utilisateur : Si le CPL est trop faible, l'œil doit revenir à la ligne trop souvent, les recherches montrent que l'utilisateur manque des informations importantes. Si CPL est trop élevé, les recherches montrent que la rapidité de lecture diminue et que l'œil a du mal à visualiser la ligne suivante correctement.
Sévérité : 5/15 – Impact : 2/5 – Importance de la tâche : 1/5 – Nombre d'utilisateurs : 2/5
Proposition de correction : Il faudrait veiller à réduire considérablement le CPL actuel. La mise en page des nouvelles maquettes permettra de remédier à ce problème.

Figure 3-2 : Problèmes récurrents dans l'ensemble des expériences pédagogiques

Ces problèmes ont montré les lacunes de l'interface par rapport à de nombreux critères. Le nombre important de problèmes identifiés et leur récurrence justifient la conception d'une nouvelle interface. Nous allons détailler le processus de conception qui nous a permis de satisfaire ces critères.

3.3 Conception de la nouvelle interface utilisateur

Nous présenterons dans un premier temps la conception de la nouvelle structure globale, puisqu'elle sera utilisée dans toutes les pages de l'interface de l'outil du CIRANO. Puis nous détaillerons la conception de l'interface pour l'expérience pédagogique « Ressource commune ».

3.3.1 Conception de la structure globale de l'interface

La structure globale concerne l'organisation des différents éléments affichés à l'écran. Nous présentons dans un premier temps une critique de la structure actuelle.

3.3.1.1 Critique de la structure actuelle

En règle général, dans un site web, un menu en haut permet par exemple de savoir sur quelle page on se trouve et facilite la navigation. Ici, on ne trouve pas de tel menu. L'endroit où s'affichent des éléments identiques (tels que le gain) varie entre les expériences. Concernant

l'esthétique visuelle, il n'y a pas de style global et homogène s'appliquant à l'ensemble de l'outil. Les expériences sont différentes en fonction du développeur les ayant programmées. Certaines sont minimalistes avec par exemple un texte noir sur un simple fond blanc (voir figure 3-3). D'autres présentent des couleurs plus originales avec par exemple un texte violet pour présenter le même type de contenu.

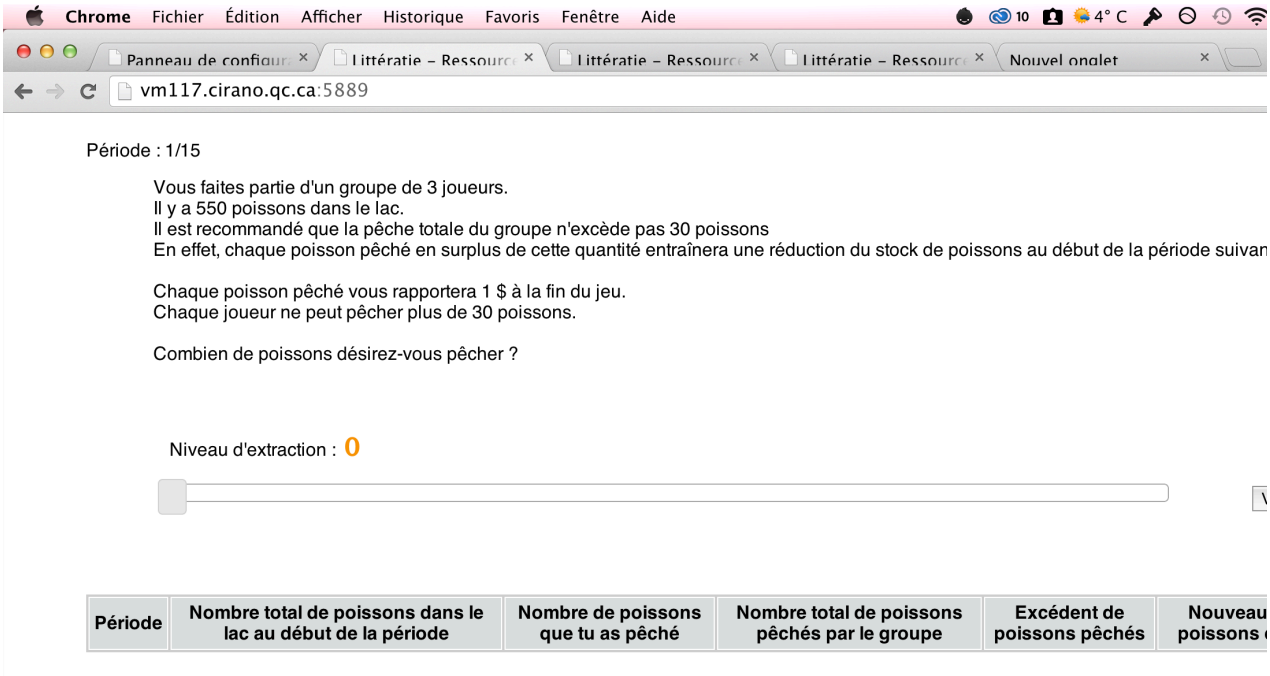


Figure 3-3 : Page de jeu de l'outil actuel - expérience pédagogique de la Ressource commune

Il y a un manque de cohérence entre les différentes expériences pour de nombreux éléments de l'interface (ex., zones de textes, tableaux, boutons, graphiques, etc.). L'identité visuelle d'un produit s'exprime par un style qui lui est propre et qui permet de reconnaitre le produit parmi d'autres. Cette identité est ici très limitée du fait du manque d'homogénéité décrite plus haut.

Pour ce qui est des aspects hédoniques, on peut voir que l'outil actuel ne s'apparente pas à première vue à un jeu. En effet, la mise en contexte visuel des expériences est très faible comme on peut le voir avec la figure précédente 3-3 présentant une session de pêche. Le caractère immersif est donc très limité.

3.3.1.2 Initialisation

Nous avons commencé par définir clairement les objectifs de la nouvelle interface, les utilisateurs cibles, ainsi que le contexte d'utilisation.

Objectifs de la nouvelle interface : améliorer l'EU des apprenants ; corriger l'ensemble des problèmes cités précédemment est une priorité. Les attributs esthétiques et hédoniques de l'outil seront également améliorés. Nous voulions donner au produit une image plus simple, mieux structurée et plus professionnelle.

Utilisateurs cibles : les utilisateurs principaux sont les étudiants de secondaire V et de CEGEP et les professeurs. De plus, le CIRANO voulait que ces expériences puissent être également utilisables par le grand public souhaitant en savoir plus sur l'économie.

Contexte d'utilisation : les expériences seront utilisées majoritairement en classe ou en salle de laboratoire, sur ordinateurs ou sur tablettes. Aucun effet sonore ne sera donc utilisé avec la nouvelle interface.

Nous avons donc ensuite porté notre attention sur la structure de l'information et le squelette d'écran que nous allions utiliser.

3.3.1.3 Schéma de l'interface

Nous avons proposé une nouvelle structure de l'information à l'écran qui pourra être utilisée dans toutes les pages du système. Pour cela, nous avons conçu des schémas de structure qui montrent le positionnement et le groupement des différents éléments qui supporteront la tâche de l'utilisateur. La présence de chaque élément dans cette nouvelle structure sera expliquée (figure 3-4) :

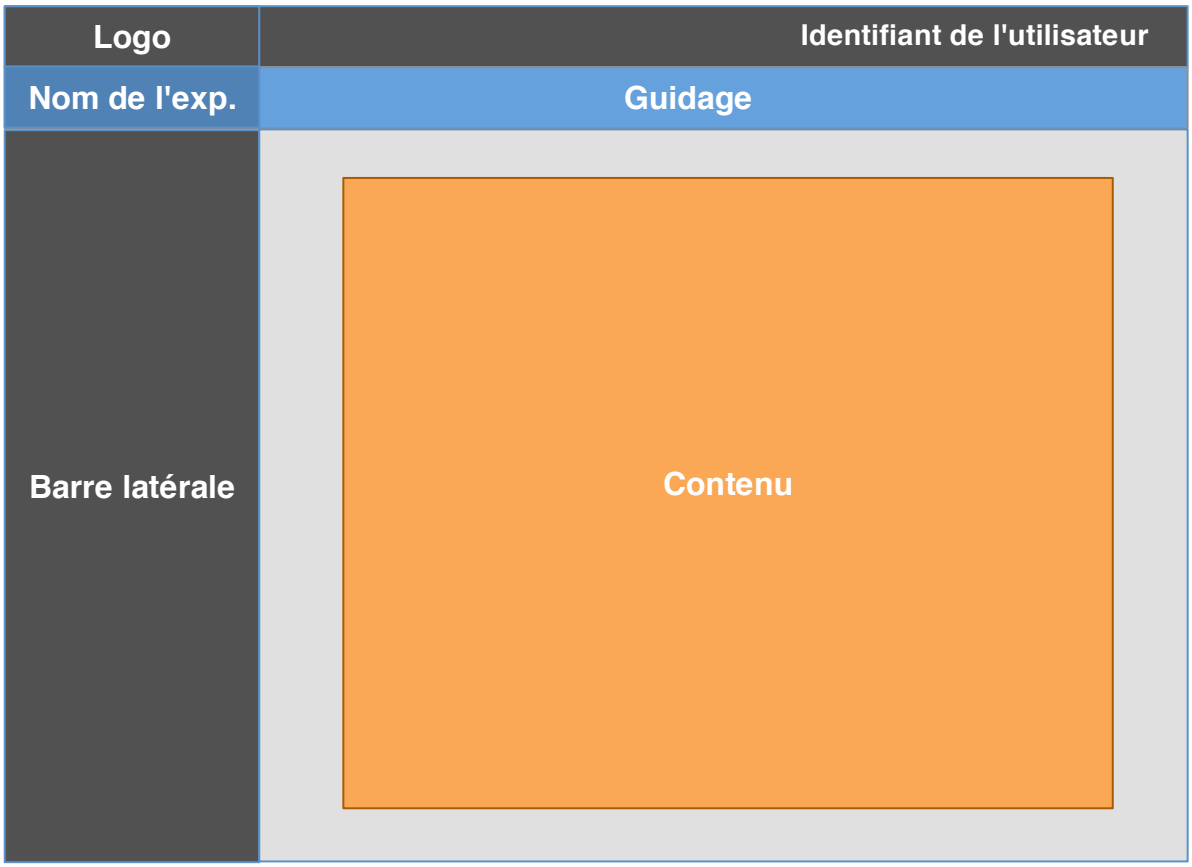


Figure 3-4 : Structure d'une page du nouvel outil

Logo : bien que le nom final de l'outil n'ait pas été encore choisi par le CIRANO, il est important que le nom du produit et le logo soient visibles afin que l'utilisateur s'en empreigne pour pouvoir le recommander.

Identifiant et état de connexion: il est important que l'utilisateur sache s'il est identifié ou déconnecté. Il est pertinent de lui rappeler son identifiant car il doit être réutilisé.

Titre de l'expérience : cela permet à l'utilisateur de se repérer et de savoir à tout moment dans quelle expérience pédagogique il se trouve.

Guidage : nous avons décidé d'ajouter une barre de guidage qui nous semble particulièrement utile pour les professeurs et les élèves. Cette barre sert à donner des instructions à l'utilisateur tout au long de son utilisation du produit afin de l'inciter à effectuer les bonnes actions (ex. « Veuillez lire attentivement la consigne », « Veuillez sélectionner le montant de votre offre », etc.). Cette barre présente également les erreurs à l'utilisateur (« Attention, votre choix précédent

est incorrect », etc.) ce qui permet de diminuer l'utilisation des boites de dialogue. Elle reste toujours visible.

Barre latérale de conseils : les écrans des ordinateurs et des iPad sont suffisamment larges pour qu'un texte en police 16, allant d'un bout à l'autre de l'écran, dépasse le nombre de caractères par ligne (CPL) recommandé. Dans l'ancien outil, les écritures allaient d'un bout à l'autre de l'écran, ce qui donnait un nombre de CPL beaucoup trop élevé, rendant la lecture difficile. La barre latérale permet donc de limiter cette largeur et est utilisée pour supporter la tâche de l'utilisateur. On affichera ici le rappel des consignes par exemple, et certaines astuces afin de mieux réussir la tâche en question.

Contenu : la zone de contenu sera utilisée pour afficher l'expérience en cours.

La nouvelle structure globale de la page a ensuite été présentée en interne à la directrice de projet, et aux deux développeurs. À ce stade il s'agissait de vérifier si la nouvelle structure serait programmable et si elle apportait un avantage vis-à-vis des problèmes ergonomiques par rapport à la précédente.

3.3.1.4 Design visuel

Nous nous sommes ensuite concentrés sur le design visuel de l'outil. Il joue un rôle important pour l'aspect esthétique de l'interface. Nous avons choisi d'utiliser le style « flat design » (design à plat) que nous décrirons plus bas.

Avant le design à plat, la tendance était celle du « Skeuomorphisme ». Le skeuomorphisme a été considérablement utilisé auparavant dans les IHM ; ce mouvement décrit les interfaces qui tentent de faire ressembler les éléments visuels à des objets physiques. On peut par exemple citer la métaphore du bureau, qui a été proposée commercialement en 1981 sur l'ordinateur Xerox Star; le contenu de l'ordinateur est représenté de façon similaire à celle d'un bureau. Les groupes de documents sont représentés sous forme de dossiers, une corbeille contient les éléments supprimés, etc. Pour soutenir ceci, de nombreux effets sont utilisés, tels que des ombres, des formes 3D, des textures, etc. La figure 3-5 présente un exemple d'une bibliothèque virtuelle skeuomorphe.

Figure 3-5 : L'application iBooks de Apple sur iPad en 2012

Cela permettait d'offrir un système plus facile à comprendre pour l'utilisateur qui n'était pas à l'aise avec les nouvelles interfaces. En effet, les actions demeurent explicites dans la mesure où les utilisateurs peuvent faire le lien entre les objets virtuels et les objets physiques réels. La suggestibilité (*affordance*) et la transparence des différents éléments sont donc augmentées, et les ambigüités sont réduites.

Cependant, ce style a été peu à peu remplacé par le design à plat au cours des dernières années dans les interfaces de géants tels que Microsoft (Windows 8), Apple (iOS7) et Google. Il s'agit donc de la tendance du moment. Cette approche se veut minimaliste sur le plan visuel comme on peut le voir avec la figure 3-6 représentant une corbeille.

Figure 3-6 : Corbeille (design à plat)

La plupart des effets donnant une illusion réelle aux éléments (ex., ombres, dégradés de couleurs, textures, ...) sont retirés. L'accent est mis sur un style beaucoup plus simple et minimaliste pour plusieurs éléments de l'interface tels que la typographie, les formes et les couleurs, comme on peut le voir sur la figure 3-7.

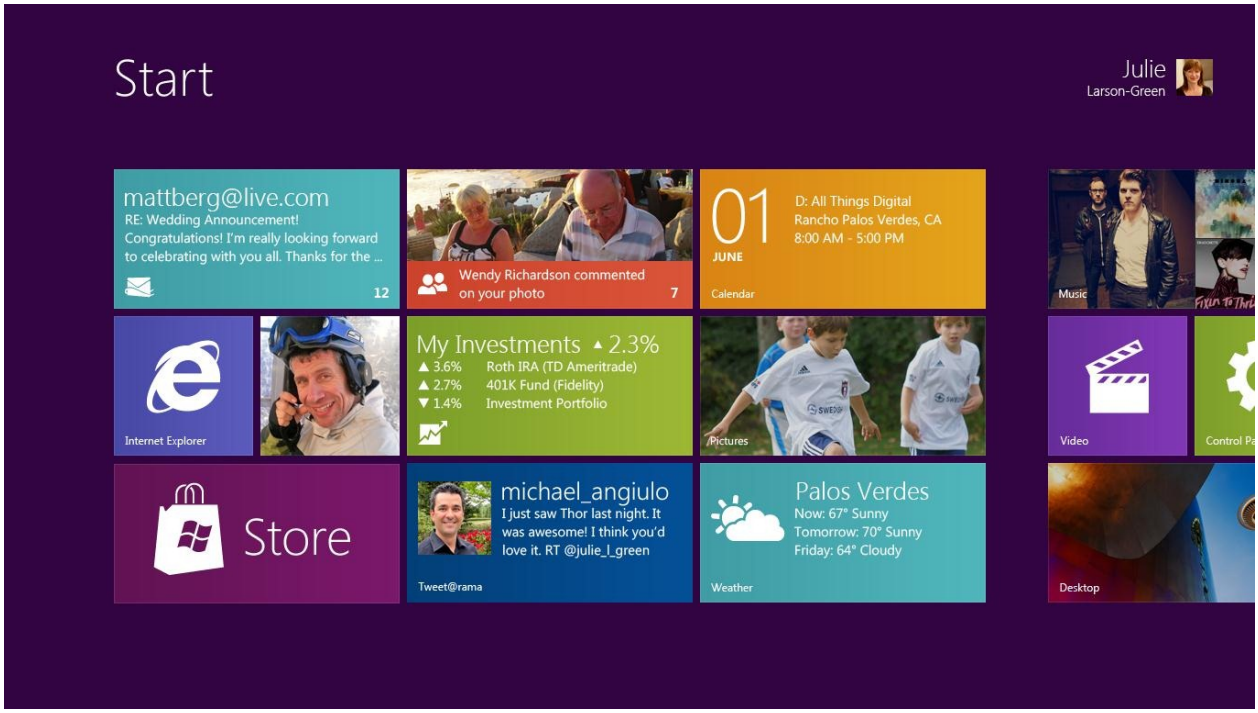


Figure 3-7 : Écran d'accueil de Windows 8 (flat design)

Son succès et les raisons qui font que nous allons l'utiliser s'expliquent principalement par les facteurs suivants : il permet d'apporter de la simplicité et de la clarté à l'interface, en diminuant la complexité visuelle avec les effets de textures. De plus, étant donné que les éléments visuels utilisent des couleurs unies (au lieu de textures plus complexes) et des formes de base très dépouillées, il est facile de les redimensionner aux différentes tailles des nombreux périphériques (téléphones intelligents, tablettes, ordinateurs), ce qui favorise l'adaptabilité. Enfin, la simplicité des éléments visuels permet des temps de chargement de la page plus rapide. En ce qui concerne la suggestibilité ou la transparence des éléments de l'interface, elle reste généralement plus faible que celle des interfaces skeuomorphes. Cependant, étant donné l'utilisation très répandue des interfaces de nos jours, les utilisateurs sont de plus en plus habitués aux formes et au style du design à plat.

Voici un aperçu d'une maquette de la structure globale de l'outil. La figure 3-8 présente une page de consignes :

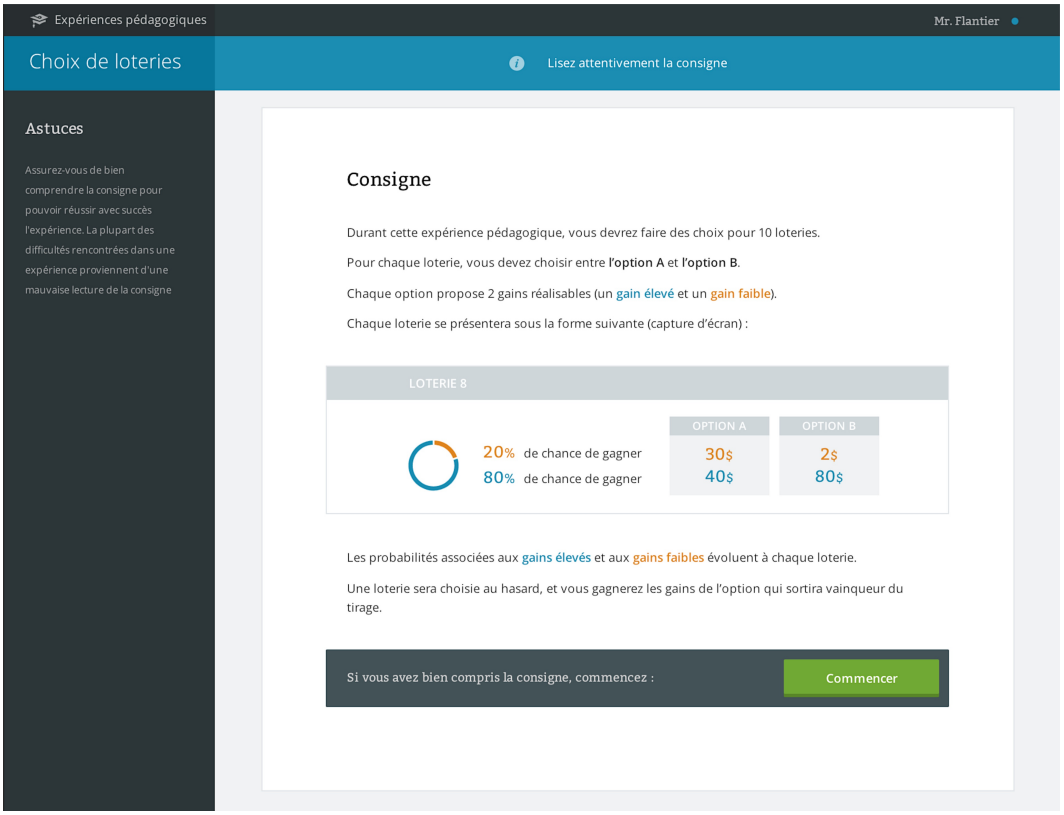


Figure 3-8 : Maquette de l'interface utilisateur de l'outil supportant une expérience pédagogique

La figure 3-8 présente une version agrandie de la figure précédente :

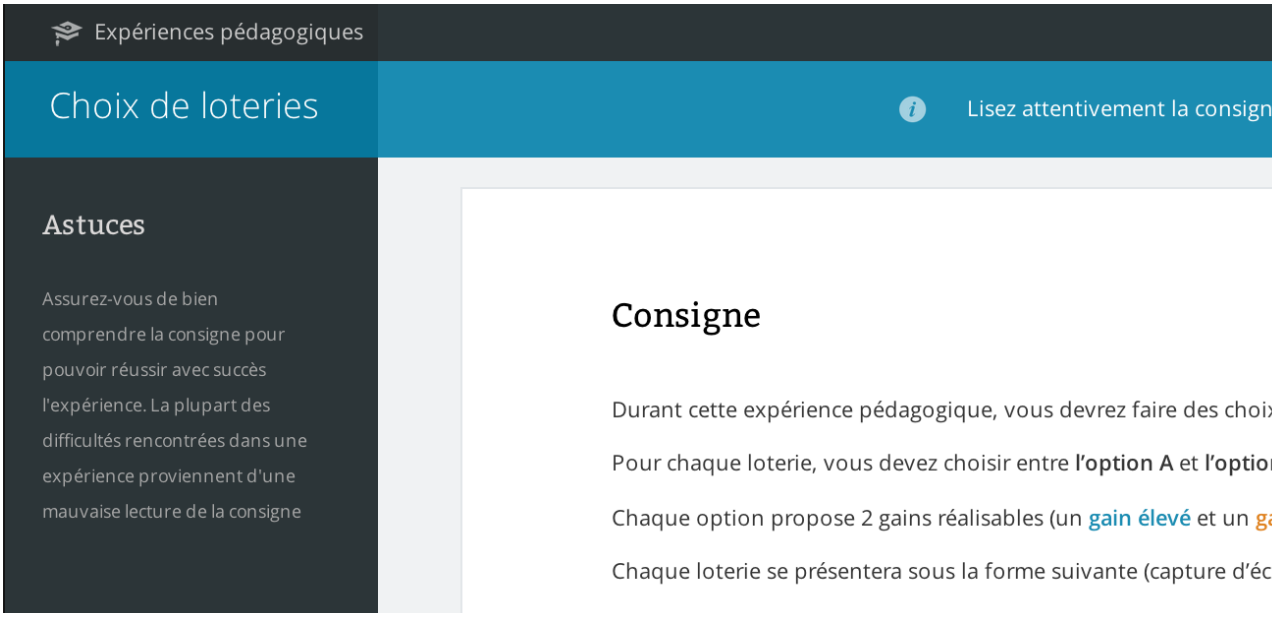


Figure 3-9 : Maquette de l'interface utilisateur de l'outil supportant une expérience pédagogique (version agrandie)

Cette nouvelle structure permet de régler de nombreux problèmes récurrents, tels que l'utilisation répétée des boites de dialogue, le CPL trop élevé, le manque d'adaptabilité, etc.

Les problèmes liés à l'identification de l'individu ont été également adressés (figure 3-10) en précisant l'information à fournir de façon explicite et en modifiant le titre non pertinent dans l'ancienne identification.

Figure 3-10 : Ancienne identification à droite et nouvelle à gauche

Enfin, afin de favoriser la cohérence et l'homogénéité à travers l'outil, nous avons produit un guide de style pour la conception des différents éléments de l'interface : polices, couleurs, styles et éléments à utiliser.

Au total, parmi les 87 problèmes ergonomiques identifiés dans les six expériences pédagogiques, la nouvelle structure permettait de résoudre ou d'entamer la résolution de 52 problèmes (certains problèmes nécessitaient l'apport de la nouvelle structure et un effort spécifique par expérience pour être totalement résolus).

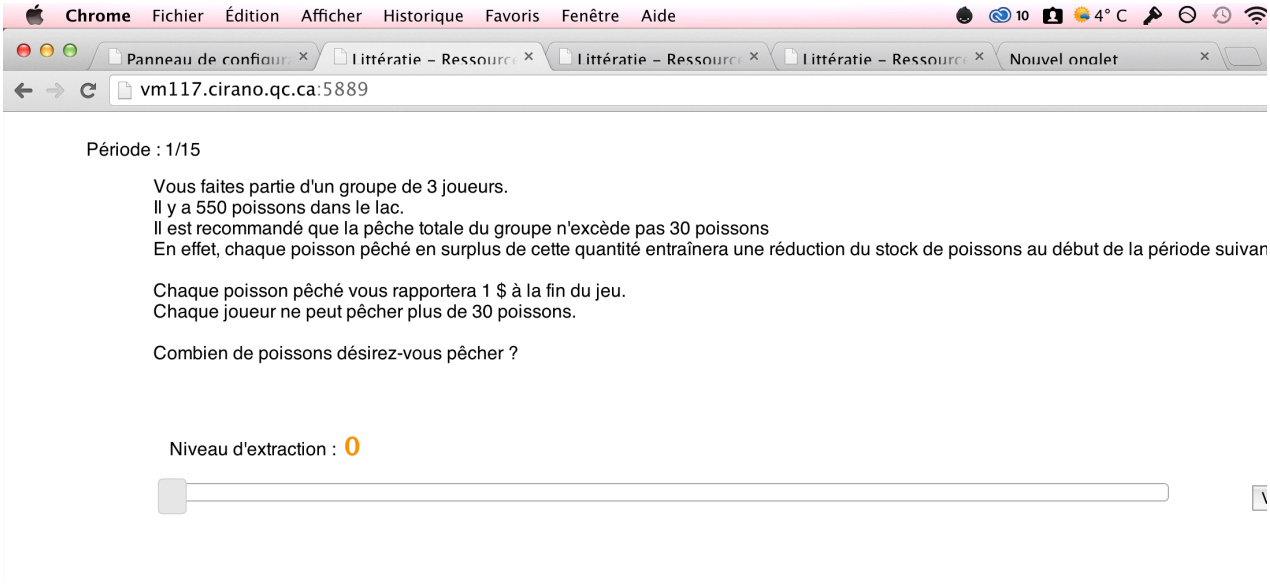
Une fois la structure de l'interface définie, nous avons pu nous concentrer sur les différentes expériences pédagogiques et proposer les premières maquettes afin de répondre à l'ensemble des problèmes restants.

Pour des raisons budgétaires et de délais, nous ne pouvions pas repenser les expériences pédagogiques depuis le début. Nous devions les améliorer tout en restant relativement proches de l'expérience originale, afin que l'implémentation ne soit pas trop longue. Les économistes avaient également des directives très claires sur le déroulement des actions. Le contenu économique des expériences était très rigide. Il fallait donc relever le défi de tenter d'améliorer les expériences pédagogiques tout en respectant les contraintes fournies.

3.3.2 Conception de l'expérience « Ressource commune »

Nous présentons ici les principaux axes d'amélioration que nous avons suivis pour la conception de la nouvelle version de l'expérience pédagogique « Ressource commune ».

Voici un aperçu de l'ancienne version de « Ressource commune ».



Période	Nombre total de poissons dans le lac au début de la période	Nombre de poissons que tu as pêché	Nombre total de poissons pêchés par le groupe	Excédent de poissons pêchés	Nouveau stock de poissons dans le lac	Gain cumulé depuis le début du jeu
6	118	30	90	60	-2	166
5	238	30	90	60	118	136
4	358	30	90	60	238	106
3	478	30	90	60	358	76

Figure 3-11 : Aperçu de l'expérience « Ressource commune » dans l'ancienne version de l'outil

Comme nous pouvons le voir, la page comporte un style qui s'apparente assez peu à un jeu. La page présente les consignes en haut, un bouton à glissière permettant de choisir le nombre de poissons à pêcher au milieu, puis en bas de la page, un tableau présente un récapitulatif des périodes (jours de pêche) précédentes. Ce tableau récapitulatif permet à l'utilisateur de visualiser ces choix précédents, ceux de son groupe, ainsi que l'évolution du stock du lac. On peut déjà remarquer que le tableau récapitulatif demande une charge de travail assez élevée à l'utilisateur puisqu'il comporte sept colonnes de chiffres à analyser.

Les questions qui nous ont guidés dans cette conception étaient les suivantes : quelles sont les informations les plus pertinentes pour l'utilisateur ? Quelles sont celles qui sont prioritaires ? Quelles sont les informations récurrentes à travers les expériences ?

Barre latérale

L'interface utilisateur de l'outil supportant l'expérience hérite de la nouvelle structure globale de l'interface présentée plus haut (cf section 3.3.1). Cela nous a permis de transférer le rappel de consigne dans la barre latérale prévue à cet effet (information récurrente) (section 3.3.1.b). Cela permet à l'utilisateur de localiser cette information de façon plus prévisible et nous offre plus d'espace pour le contenu principal. Dans l'ancienne version, un rappel de consigne conséquent était présent dans le contenu principal.

Barre d'état

De plus, le numéro de la période en cours et le gain cumulé sont des informations récurrentes dans toutes les expériences et sont prioritaires. Elles doivent donc ressortir pour l'utilisateur. Ceci peut être réalisé de plusieurs façons dans le design d'une interface. Nous avons principalement utilisé les deux façons suivantes :

- *En jouant sur la proximité des éléments*

- *En jouant sur le contraste avec les autres éléments (contraste de tailles, couleurs, épaisseurs, formes)*

Ceci nous a aidés à créer des différences entre les éléments. Nous avons donc choisi d'afficher ces informations dans un conteneur (nommé « barre récapitulative »), en mettant en avant le gain cumulé, qui est la variable que le joueur doit maximiser (voir figure 3-12). Cet élément a été retiré du tableau récapitulatif ce qui permet de retirer une colonne. Une icône pour le gain aide également à faire ressortir et à localiser cette information.

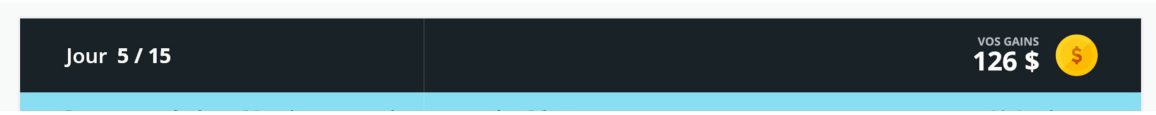


Figure 3-12 : Conteneur des informations récurrentes dans les expériences pédagogiques

Cette modification permet de réduire la taille du tableau récapitulatif et donc la charge de travail. Étant donné que cette barre est utilisée dans plusieurs expériences, on peut supposer que l'utilisateur saura où trouver ces informations dans chaque expérience et que ceci contribuera à l'amélioration du guidage. (Le joueur voit clairement à quelle période il se trouve et l'accent est mis sur son gain).

Contenu principal

Nous avons tenté d'améliorer les aspects hédoniques de l'expérience qui porte sur; celle-ci porte sur une pêche en groupe avec un stock de poissons limité. Ainsi nous avons cherché à rendre le contexte de l'expérience le plus réaliste et visuel possible afin d'augmenter l'immersion et l'engagement de l'utilisateur (voir figure 3-11).

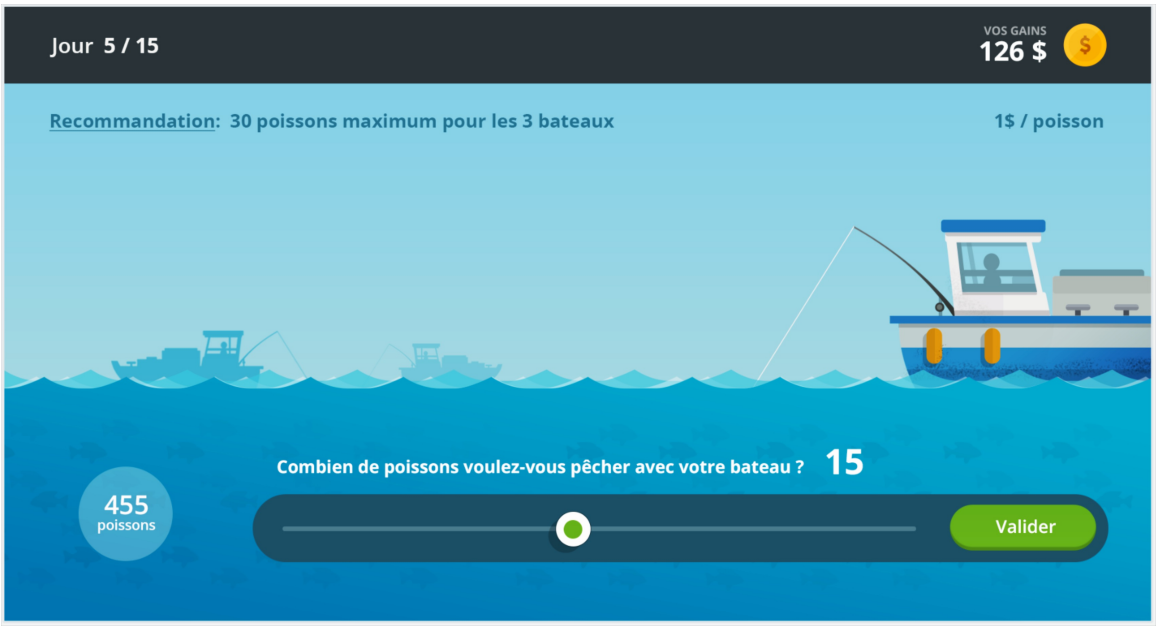


Figure 3-13 : Nouvelle version de la page de jeu de l'expérience « Ressource commune »

La figure 3-13 nous montre la représentation que nous avons faite de l'expérience de pêche. Étant donné que l'on voit plusieurs bateaux, le joueur prend conscience qu'il est en groupe et qu'il

partage une même ressource limitée. Certains éléments ont été replacés pour augmenter la prévisibilité. Par exemple, le nombre de poissons restant dans le lac (présent dans le tableau récapitulatif) a été logiquement déplacé au niveau du lac. Nous avons conservé l'utilisation du bouton à glissière pour choisir le nombre de poissons pêchés, car il s'agit d'un élément facile à utiliser et sa longueur limitée permet de prévenir certaines erreurs (le nombre minimum et maximum de poissons pouvant être pêchés). Enfin, dans l'ancienne version, le terme « période de pêche » était utilisé ce qui n'était pas très explicite pour l'utilisateur. Nous avons utilisé le terme « journée » afin de rendre le terme plus explicite.

Amélioration du tableau récapitulatif

Le tableau récapitulatif est présent sur la page de jeu et permet de voir un bilan des actions précédentes afin d'aider l'utilisateur à prendre une décision durant la période en cours. Le tableau récapitulatif de l'ancienne version de l'outil était chargé d'informations et son utilisation (lecture et comparaison des données) nous semblait donner lieu à une charge mentale de travail assez élevée. Le premier objectif a donc été de diminuer le nombre d'informations dans ce tableau.

Tableau 3-2 : Tableau récapitulatif de l'ancienne version de l'expérience « Ressource commune »

Période	Nombre total de poissons dans le lac au début de la période	Nombre de poissons que tu as pêché	Nombre total de poissons pêchés par le groupe	Excédent de poissons pêchés	Nouveau stock de poissons dans le lac	Gain cumulé depuis le début du jeu
6	118	30	90	60	-2	166
5	238	30	90	60	118	136
4	358	30	90	60	238	106
3	478	30	90	60	358	76

Cela a été possible grâce au changement de mode de présentation et au déplacement de certaines informations. Nous avons eu également de nombreuses discussions avec la directrice de projet et les économistes afin de trouver quelles sont les informations réellement prioritaires au moment de prendre une décision. Nous avons donc défini un tableau à cinq colonnes. Après avoir réduit la quantité d'informations, nous avons travaillé sur la façon de la présenter dans le but de diminuer la charge perceptive. À nouveau, ceci est possible grâce aux différents moyens à notre disposition pour distinguer les différents types d'informations : information relative à une période, information de la période précédente, les différentes quantités et excédents (figure 3-15).

Tableau 3-3 : Tableau récapitulatif de la nouvelle version de l'expérience « Ressource commune »

JOUR	POISSONS DANS LE LAC AU MATIN	POISSONS QUE TU AS PÊCHÉS	PÊCHE DE TOUS LES BATEAUX	EXCÉDENT
4 HIER	480	10	35	5
3	510	10	30	0
2	530	10	20	0
1	550	10	20	0

Réduction du nombre d'actions

Nous avons également enlevé les boites de confirmation qui demandaient à l'utilisateur de valider son choix à chaque période (15 périodes), ce qui augmentait considérablement le coût d'interaction et donc la charge de travail de l'utilisateur. Les boites de confirmation sont requises pour des actions destructives ou lourdes en conséquence, ce qui n'est pas le cas ici. Ceci ralentissait considérablement le déroulement du jeu.

Pour des raisons de longueur, nous ne présentons pas toutes les modifications apportées à l'ancienne version de l'interface utilisateur de l'expérience « Ressource commune »; le texte ci-dessus donne un bon aperçu de ce que nous avons fait pour améliorer l'ensemble des expériences, et plus particulièrement « Ressource commune ».

Suite à la conception de la nouvelle version des maquettes, nous avons procédé au développement du prototype que les utilisateurs pourraient essayer.

3.4 Développement de la nouvelle interface

Nous avons transformé les maquettes statiques en prototypes dynamiques utilisables. Dans le développement web, deux parties se distinguent : devant (Front-end) et arrière (Back-end). La partie devant concerne le code qui va être utilisé pour afficher l'interface du site (tout ce que l'on voit), alors que la partie arrière concerne le code permettant de gérer tout ce dont a besoin le site pour fonctionner : gestion des clients, gestion des bases de données, calculs sous-jacents, etc.

Ainsi, la plus grande différence entre l'ancienne et la nouvelle version de l'interface utilisateur porte sur la partie devant dont nous étions en charge. Le CIRANO nous avait précisé qu'il était important que la partie arrière ne soit pas trop affectée. Nous avons donc procédé au

développement de la nouvelle version pour le web en utilisant des langages tels que le HTML 5, le CSS3 et le JavaScript.

Une bonne partie de notre temps a donc été consacré à l'implémentation des prototypes. Dans la continuité du travail d'homogénéité et de cohérence, un effort a été fourni pour structurer et placer la structure globale ainsi que les éléments communs à un niveau plus haut, pouvant être réutilisés facilement dans tout l'outil. De nombreux tests ont également été effectués afin d'éliminer les bogues.

3.5 Validation

Durant tout le processus de développement présenté précédemment, des validations à l'interne étaient effectuées avec des membres du CIRANO pour des raisons budgétaires. Les prototypes étaient présentés à la directrice de projets, aux développeurs et aux économistes. Quelques professeurs ont pu également essayer la nouvelle version de l'outil d'administration et nous ont fourni des retours qualitatifs très positifs. Cependant, ceci n'est pas suffisant. Nous devions procéder à des essais nous permettant de mesurer quantitativement l'amélioration de l'EU, et surtout afin de déceler les prochaines améliorations nécessaires. La nouvelle version de « Ressource commune » était une des rares expériences (avec « choix de loteries ») à être suffisamment avancée et exempte de bogues pour être essayée en externe. Nous allons présenter dans le chapitre suivant les essais que nous effectuerons sur l'ancienne et la nouvelle version de « Ressource commune ».

L'évaluation heuristique à l'aide des critères de Bastien & Scapin a mis en avant de nombreux problèmes ergonomiques de l'interface. Ces problèmes affectent la qualité pragmatique de l'outil pédagogique, ce qui influe probablement négativement sur l'EU. Beaucoup de ces problèmes étaient récurrents. Le développement de la nouvelle version de l'outil a été effectué dans le but de régler ces problèmes. Nous avons remarqué que la qualité hédonique de l'interface utilisateur semblait également très limitée. L'interface ne s'apparentait pas directement à un jeu, la mise en contexte était faible et l'esthétique de l'interface était peu homogène. Le développement de la nouvelle structure globale et plus particulièrement de l'expérience « Ressource commune » vise à améliorer l'EU vécue avec l'outil. Cependant, l'implication des utilisateurs a été plutôt faible.

Beaucoup de retours ont été qualitatifs et il est maintenant nécessaire de tester la nouvelle version auprès des utilisateurs.

Le chapitre suivant présente la méthodologie que nous allons suivre pour évaluer la nouvelle interface et la comparer à l'ancienne.

CHAPITRE 4 MÉTHODOLOGIE D'ÉVALUATION DE L'EXPÉRIENCE UTILISATEUR VÉCUE AVEC LES DEUX VERSIONS DE L'INTERFACE

Ce chapitre présente la méthodologie que nous avons suivie pour évaluer la qualité de l'expérience que les utilisateurs ont vécue avec l'ancienne et la nouvelle interface de « Ressource commune ». Cette évaluation est importance pour connaitre les réactions des utilisateurs et les améliorations futures à apporter, s'il y a lieu.

4.1 Méthodologie

Cette section présente le questionnaire d'évaluation utilisé, les caractéristiques des participants auxquels nous avons fait appel et la procédure de tests.

4.1.1 Questionnaire d'évaluation

Tel que mentionné plus haut, nous avons choisi d'utiliser un questionnaire pour recueillir des données quantitatives sur l'EU. Un questionnaire a comme principaux avantages de permettre de rejoindre rapidement un grand nombre d'utilisateurs et de recueillir des données de manière complète et structurée.

Nous avions deux versions de notre expérience pédagogique : la version initiale, que nous nommerons *version A*, et la nouvelle version que nous avons développée, que nous nommerons *version B*. Il était donc important pour nous d'évaluer les versions A et B des interfaces utilisateur afin de les comparer.

Le questionnaire est composé de trois sections. La première est dédiée à évaluer l'EU : pour cela, nous avons traduit et utilisé la version réduite de l'outil AttrakDiff composée de 10 paires de mots opposés, sur une échelle d'Osgood à sept points. La figure 4-1 nous montre cette première section.

Confuse	☐-----☐-----☐-----☐-----☐-----☐-----☐	Structurée
Non pratique	☐-----☐-----☐-----☐-----☐-----☐-----☐	Pratique
Imprévisible	☐-----☐-----☐-----☐-----☐-----☐-----☐	Prévisible
Compliquée	☐-----☐-----☐-----☐-----☐-----☐-----☐	Simple
De mauvais goût	☐-----☐-----☐-----☐-----☐-----☐-----☐	Élégante
Bas de gamme	☐-----☐-----☐-----☐-----☐-----☐-----☐	Haut de gamme
Non imaginative	☐-----☐-----☐-----☐-----☐-----☐-----☐	Imaginative
Ennuyante	☐-----☐-----☐-----☐-----☐-----☐-----☐	Captivante
Mauvaise	☐-----☐-----☐-----☐-----☐-----☐-----☐	Bonne
Laide	☐-----☐-----☐-----☐-----☐-----☐-----☐	Belle

Figure 4-1 : Section 1 du questionnaire AttrakDiff en version réduite

Nous avons traduit les paires de mots en français à l'aide d'un traducteur.

La deuxième section comporte six questions composées par nous et portant sur le bilan de l'utilisation de l'outil pédagogique, en regard des buts de l'entreprise tels que la satisfaction de l'utilisateur, l'envie de recommander le système, l'image de marque de l'entreprise, etc. Le tableau 4-1 montre les six questions auxquelles on répond dans des cases à cocher placées sur un continuum allant du désaccord à l'accord total, sur une échelle de Likert à cinq points.

Tableau 4-1 : Section 2 du questionnaire relatif au bilan de l'utilisation de l'outil

	Pas du tout d'accord	Pas d'accord	Neutre	Plutôt d'accord	Tout à fait d'accord
Cette expérience m'a été profitable					
Je suis satisfait de l'expérience					
Je serais tenté de réutiliser le système					
Je recommanderais le système à mon entourage					
J'aimerais que d'autres expériences pédagogiques similaires					
Ces expériences m'ont donné une bonne image du CIRANO (centre de recherche ayant créé et développé le système)					

La troisième section du questionnaire est décrite dans la section suivante.

4.1.2 Distinction entre l'interface et le contenu

Il y a une distinction très importante à faire lors de l'évaluation de l'EU avec un outil tel que celui d'Hassenzahl entre l'interface utilisateur de l'outil et puis le contenu traité dans cette interface qui porte ici sur des notions d'économie et de finance (figure 4-2). La raison est simple : le participant peut apprécier l'un sans apprécier l'autre, et il ne faut pas que l'évaluation de l'un soit entachée par l'évaluation de l'autre.

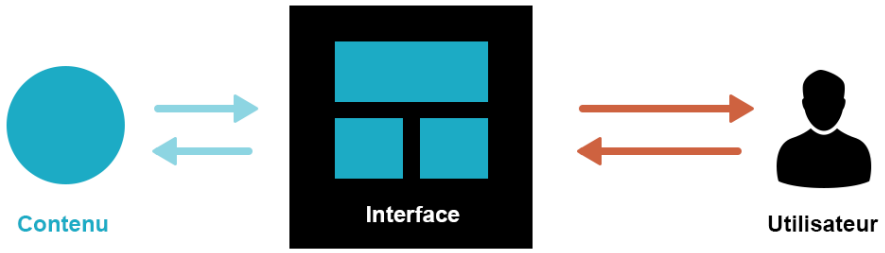


Figure 4-2 : Distinction entre l'interface et le contenu

Ainsi, lorsque les utilisateurs évaluent l'interface de l'outil pédagogique, ils risquent fort d'évaluer aussi le contenu traité, sans bien distinguer les deux. Si ce contenu est complexe et opaque et ne plait pas à l'utilisateur, cela aura vraisemblablement un effet négatif sur l'EU, même

si l'interface est bien conçue. Étant donné que nous ne pouvions pas changer le contenu présenté, il nous paraissait important de tenir compte de cette distinction dans les questionnaires. Nous avons donc demandé à l'utilisateur de tenir compte de cette distinction en évaluant son degré d'aisance avec les concepts d'économie et de finance et son appréciation de ceux-ci

Le questionnaire au complet a été revu par un expert du domaine de l'ergonomie, et par trois étudiants extérieurs au projet pour nous assurer de la compréhension et de la clarté des différentes questions avant d'être soumis aux participants.

4.1.3 Participants

Les participants ont été choisis parmi les contacts personnels du chercheur et ils constituent en ce sens un échantillon de convenance. Pour participer, il fallait être majeur, parler français et ne pas souffrir de surdité ou être malvoyant. Parmi les invitations envoyées, nous avons obtenu un petit nombre de répondants. Nous avons néanmoins réussi à obtenir 20 participants pour notre projet de recherche.

4.2.3.1 Représentativité de l'échantillon

Les expériences pédagogiques se destinent à plus long terme au grand public, mais ciblent dans un premier temps les étudiants de secondaire V et de CEGEP. Notre échantillon de 20 participants est composé de 70 % d'hommes et de 30 % de femmes. Les hommes sont donc un plus représentés que les femmes dans notre échantillon. Or, d'après le service régional d'admission du Montréal métropolitain (SRAM), il y avait en 2013, 58 % de femmes admises à un programme collégial contre 42 % d'hommes. Notre échantillon n'est donc pas représentatif des étudiants du CEGEP en qui concerne le sexe. Djamasbi et al. (2007) ont montré qu'il n'existait pas de différence significative entre les sexes pour ce qui est de l'évaluation de l'utilisabilité. Cependant, il se peut qu'il en existe pour la perception des attributs esthétiques et hédoniques des produits. 80 % des participants appartiennent à la tranche d'âge des 18-29 ans et 20 % à la tranche d'âge 45-59 ans. Notre échantillon comprend donc majoritairement des participants jeunes. L'âge médian des étudiants au collégial était de 22 ans (Statcan, 2006), tandis que l'âge moyen de notre échantillon est de 28 ans. L'échantillon est donc peu représentatif de la population visée pour ce qui est de l'âge. Enfin, concernant le niveau de scolarité, 5 % des participants ont complété l'équivalent d'un diplôme d'études collégiales, 65 % ont complété

l'équivalent d'une formation de premier cycle, 25 % ont complété l'équivalent d'une formation de second cycle, et 5 % ont complété l'équivalent d'une formation de 3e cycle (doctorat). L'échantillon que nous avons obtenu n'est pas représentatif de la première population cible du système. Cependant, les expériences visent aussi le grand public; celui-ci nous permettra d'avoir des indices sur l'amélioration de l'outil.

4.2.3.2 Définition des sous-échantillons

Nous avions besoin de deux sous échantillons (E-AB et E-BA). L'échantillon E-AB essayait la version A de l'outil puis la version B, tandis que l'échantillon E-BA le faisait dans l'ordre inverse; cela avait pour but de contrebalancer les biais causés par l'ordre d'utilisation des deux versions. Le tableau 4-2 présente un récapitulatif des essais des échantillons.

Tableau 4-2 : Essais des différents échantillons

E-AB	Ress-Comm A	Questionnaire	Ress-Comm B	Questionnaire
E-BA	Ress-Comm B	Questionnaire	Ress-Comm A	Questionnaire

Les participants ont été divisés en deux groupes égaux de 10 participants. À mesure que nous obtenions de nouveaux candidats, nous les avons répartis dans les deux groupes en tentant d'homogénéiser prioritairement l'âge.

La moyenne d'âge de l'échantillon E-AB est de 27 ans et celle de E-BA est de 29 ans, ce qui très comparable.

L'échantillon E-AB comporte 80 % d'hommes et 20 % de femmes, contre 60 % d'hommes et 40 % de femmes pour E-BA. Il nous a été impossible d'homogénéiser le niveau de scolarité.

4.1.4 Procédure de test

Les essais ont été effectués à distance. Les expériences pédagogiques ainsi que les questionnaires étaient accessibles en ligne. Le participant était invité à lire et signer le formulaire d'information et de consentement de participation à la recherche par courriel. Une fois le formulaire signé, il recevait la marche à suivre ainsi qu'un code identifiant. Il lui était demandé de se mettre dans un milieu calme afin d'éviter toute distraction.

La tâche à effectuer était la suivante. Pour le groupe E-AB, le participant devait tout d'abord essayer l'ancienne version de l'outil. Pour cela, il se connectait à l'adresse fournie et rentrait l'identifiant qui lui était distribué. Il devait lire la consigne de l'expérience « Ressource commune ». Puis il participait au jeu sur 15 périodes. Ensuite, il prenait connaissance de la page de résultats. Enfin, il devait répondre au questionnaire en ligne. Une fois que le questionnaire était rempli, un processus similaire était suivi avec la nouvelle version. Le participant répondait donc au même questionnaire pour cette nouvelle version. Pour le groupe E-BA, le participant commençait par la nouvelle version puis utilisait l'ancienne.

Les captures d'écran de l'ancienne et de la nouvelle version sont disponibles à l'Annexe 6.

Le chercheur était connecté sur Skype avec le participant afin de pouvoir répondre à ses questions à tout moment. Si ce n'était pas possible, le chercheur était disponible par le biais d'un système de message en temps réel. La durée des essais était d'environ 20 minutes. Notre rôle se limitait à présenter le projet et les tâches à effectuer, puis nous restions disponibles pour répondre à toutes questions.

4.2 Nettoyage des données

À la fin, nous avons passé en revue les différentes réponses afin de nous assurer que certains questionnaires ne présentaient pas des réponses douteuses ou incomplètes. Au final, sur nos 40 questionnaires (les 20 participants remplissaient chacun deux questionnaires), aucun n'a été retiré.

Nous avions quatre groupes de résultats pour l'expérience : version A (résultats avant et après combinés), version B (résultats avant et après combinés), version A en ayant testé la version B avant, version B en ayant testé la version A avant. Nous avons choisi de garder ces catégories.

4.3 Traitements des données

Un des graphes que nous allons présenter dans le chapitre des résultats doit être expliqué.

À l'instar de l'outil AttrakDiff qui est disponible en ligne, nous voulions pouvoir visualiser le positionnement des qualités pragmatiques et hédoniques du produit de façon graphique.

La figure 4-3 a été créée en se basant sur l'outil d'AttrakDiff. On pourra ainsi voir le positionnement des versions A et B de l'outil. Hassenzahl utilise des rectangles d'intervalle de

confiance, mais nous utiliserons des rectangles d'écart-type afin de montrer la variabilité des évaluations. Plus le rectangle est petit, plus le positionnement du produit est fiable, car cela reflète une plus grande unanimité de réponses des participants. Nous ne pouvons pas utiliser d'intervalles de confiance, car nos échantillons ne comportent pas un nombre suffisant de participants. À droite du graphe, les scores d'attractivité (ATT) des différentes versions sont également fournis. La façon dont nous avons obtenu ce graphe sera détaillée dans la section suivante.

Figure 4-3 : Exemple de graphique présenté par l'outil AttrakDiff

4.3.1 Moyenne des scores de la qualité pragmatique et hédonique

Pour réaliser ce graphe, nous allons calculer la moyenne QP_k des quatre dimensions pragmatiques $(\mathrm{QP}_i)_k$ pour l'utilisateur k. Ensuite, nous calculerons la moyenne globale des qualités pragmatiques QP de tous les utilisateurs.

Ainsi pour n participants : $QP_k = \frac{1}{4}\sum_{i=1}^{4}(\mathrm{QP}_i)_k$ et $QP = \frac{1}{n}\sum_{i=1}^{n} QP_k$

Nous procèderons de la même façon pour calculer la moyenne globale des qualités hédoniques QH.

Les dimensions pragmatiques et hédoniques sont notées sur une échelle d'Osgood allant de -3 à 3 (sept points). Donc les moyennes QP et QH seront également dans cet intervalle. QP sera reporté sur l'axe des abscisses, et QH sur celui des ordonnées, ce qui nous donnera le placement du produit. L'attractivité ATT n'intervient pas pour ce placement. Voici un rappel des dimensions d'AttrakDiff (figure 4-4) :

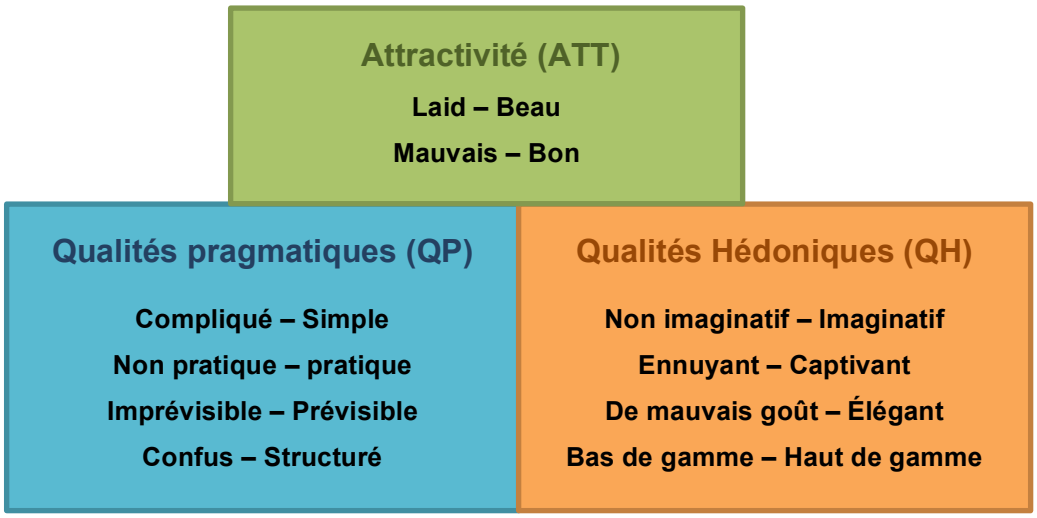


Figure 4-4 : Version réduite de l'outil AttrakDiff

Les pourcentages d'amélioration entre les différentes versions seront calculés sur une échelle fixe en considérant que le passage de -3 à 3 correspond à une amélioration de 100 %.

4.3.2 Attractivité globale

L'attractivité globale ATT est simplement la moyenne des attractivités ATT_k évaluées par chaque participant. L'attractivité globale est précisée à droite du graphe et donne un aperçu de l'impression globale que l'utilisateur a eue de l'expérience. Elle est une résultante de QP et QH. L'attractivité est elle aussi notée de -3 à 3.

4.3.3 Écart-type des qualités pragmatique et hédonique

Chaque version de l'outil est positionnée sur l'axe des abscisses et l'axe des ordonnées grâce au calcul des moyennes QP et QH. Ensuite, un rectangle d'écart-type permet de faire apparaitre la variabilité des évaluations (figure 4-5).

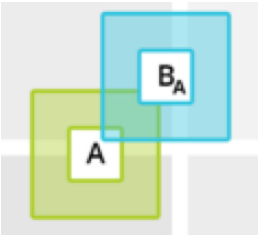


Figure 4-5 : Rectangle d'écart-type présenté dans nos résultats

La largeur et la longueur du rectangle seront respectivement calculées à l'aide de σ_{QP} et σ_{QH}. Ainsi, plus le rectangle est large, plus l'évaluation de la qualité pragmatique est variable. Plus il est long, plus l'évaluation de la qualité hédonique est variable.

L'écart-type de la qualité pragmatique σ_{QP} sera calculé à partir des évaluations QP_k de chaque participant en utilisant la formule standard :

$$\sigma_{QP} = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(QP_k - QP)^2}$$

Dans la formule, il est important de remarquer que, contrairement à la moyenne, nous divisons par (n-1) et non par (n) afin d'obtenir un estimateur non biaisé de l'écart-type (c.-à-d. tel que l'espérance de notre estimateur soit égale à l'écart-type de notre population).

De même, nous procéderons à un calcul semblable pour σ_{QH} avec les attributs hédoniques.

Ce graphique nous permettra ainsi de positionner et de comparer les différentes versions de l'outil selon les qualités pragmatiques et hédoniques de l'interface. Nous pourrons également connaître l'attractivité globale de l'interface de chaque outil. Les autres éléments mesurés dans nos évaluations seront présentés de façon détaillée dans le chapitre suivant consacré à la présentation et l'analyse des résultats.

CHAPITRE 5 RÉSULTATS ET ANALYSE

Ce chapitre est dédié à la présentation et l'analyse des résultats des évaluations de l'EU et d'autres facteurs. Nous ferons d'abord un retour sur le déroulement des essais, ensuite nous présenterons une analyse des résultats des évaluations de l'EU ainsi que l'impact sur les autres buts de l'entreprise.

5.1 Retour sur le déroulement des essais

Nous avons donc obtenu un total de 40 essais. Chacun de nos participants essayait chacune des deux versions des interfaces de l'expérience pédagogique. Les 40 essais des participants se sont bien déroulés et ont duré en moyenne 24 minutes chacun (avec un écart-type d'environ quatre minutes). Les participants semblaient généralement intéressés après avoir fini leurs essais puisqu'ils avaient de nombreuses questions à poser; ils demandaient si le système était accessible au grand public, si d'autres expériences existaient, etc.

Les essais se faisaient en ligne ainsi que le remplissage des questionnaires. Vingt essais ont été effectués en présence physique du candidat alors que vingt autres ont été effectués via Skype (système d'appel vidéo et vocal) afin d'éviter à certains participants de se déplacer. Parmi les vingt participants, nous en avions sept résidant dans différentes villes. Pour les essais qui ont été effectués en présence physique avec le participant, nous avons fait attention de ne pas perturber ce dernier.

Nous n'avons pas eu d'abandon pendant les essais.

5.2 Analyse de l'évaluation de l'expérience utilisateur

Cette section présente les résultats des évaluations des EU vécue avec chaque version de l'outil d'apprentissage et faites avec l'outil AttrakDiff. Nous allons présenter les résultats obtenus avec chaque version de l'outil.

5.2.1 Comparaison des résultats des versions A et B

La figure 5-1 présente les résultats des évaluations des versions A (ancienne) et B (nouvelle) par nos deux sous-échantillons. Nous pouvons voir le placement des différentes versions sur les

échelles d'évaluation des qualités pragmatiques et hédoniques. Les scores d'attractivité sont également disponibles à droite du graphique.

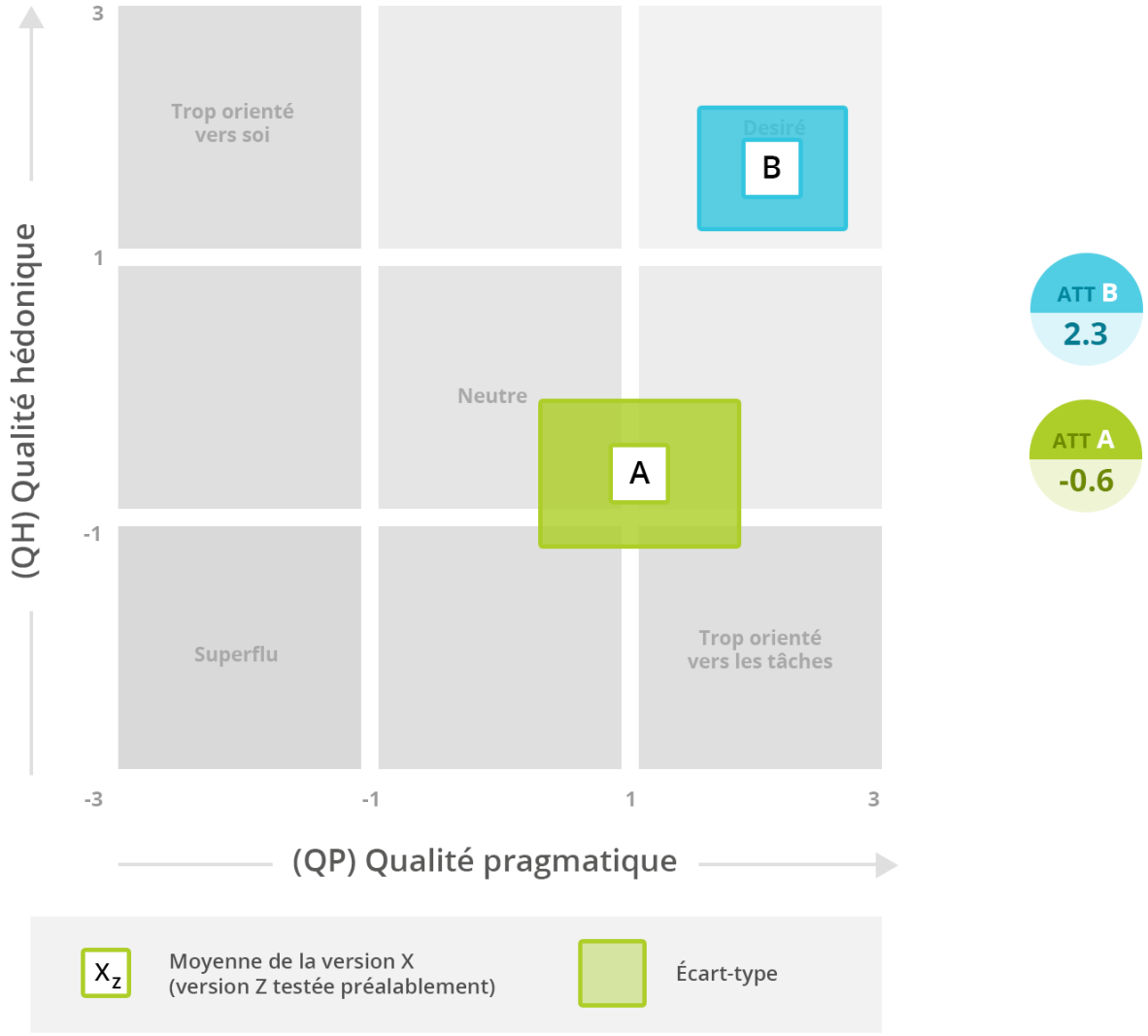


Figure 5-1 : Résultats d'évaluation des deux versions de l'outil d'aide à l'apprentissage

Tout d'abord, nous pouvons voir que la version A a un plus faible succès auprès des utilisateurs. Ceci confirme le fait que le développement d'une nouvelle version était fortement nécessaire. En effet, rappelons que la qualité pragmatique (QP), la qualité hédonique (QH) et l'attractivité globale (ATT) sont notées sur une échelle de -3 à 3, ce qui correspond à l'échelle d'Osgood à 7 points sur laquelle les candidats évaluaient l'EU. Ainsi la version A obtient une QP de 1.1 (écart-type de 1.5) ce qui est faible, mais passable. Nous avons déjà mentionné le fait que la section de l'outil supportant l'expérience « Ressource commune » ne comportait pas un nombre important de failles d'utilisabilité comparativement aux sections de l'outil supportant les autres expériences

pédagogiques. Cependant, nous savions que les attributs hédoniques de l'outil étaient très faibles et qu'il y avait place à amélioration : la version A obtient une QH de -0.7 (écart-type de 1.1). L'attractivité globale qui en résulte est ainsi tirée vers le bas. Le score d'attractivité globale est de -0.6 (écart-type 1.4) ce qui signifie que l'EU est négative.

La nouvelle version (B) de l'outil obtient heureusement des résultats plus satisfaisants. En effet, la QP, qui porte sur les attributs de l'utilisabilité, passe de 1.1 à 2.1 (écart-type 1.1), et connaît donc une amélioration de 17 %. La QH passe de -0.7 à 1.7 (écart-type 0.9) ce qui correspond à une amélioration de 33 %. Enfin, l'attractivité globale qui est la résultante de ces deux qualités passe de -0.6 à 2.3 (écart-type de 0.9). L'impression globale laissée par le produit s'est donc améliorée (de 45 %). L'écart-type de l'ATT est également plus faible, ce qui montre que les évaluations de l'EU sont plus homogènes.

Interprétations

Notre but principal a été atteint puisque l'EU vécue avec la nouvelle version de l'outil a été améliorée. Les utilisateurs évaluent mieux les qualités pragmatiques et hédoniques. De plus, l'ancienne version obtient une évaluation négative de l'attractivité (score de -0.6), alors que la nouvelle version obtient une bonne évaluation de l'attractivité (score de 2.3). On remarque également que les écarts-types des évaluations de la nouvelle version de l'outil sont plus faibles, ce qui dénote une plus grande homogénéité d'évaluation. La nouvelle version n'a pas obtenu un score de 3 sur 3 pour les QP et QH ce qui signifie qu'il y a encore une place à amélioration.

5.2.2 Évolution vers le nouvel outil

Un certain nombre de professeurs et d'étudiants de cours d'économie utilisent déjà les anciennes versions de la partie d'administration et des expériences pédagogiques. Ils vont bientôt devoir utiliser la nouvelle version de l'outil. Il nous a donc semblé pertinent d'examiner les résultats partiels portant sur l'essai et l'évaluation de la nouvelle version après l'essai et l'évaluation de l'ancienne version. Dans l'échantillon E-AB, les participants essayaient la version A (ancienne) et répondaient au questionnaire, puis essayaient la nouvelle version B (nouvelle) et répondaient au même questionnaire. La figure 5-2 présente les résultats d'évaluation des deux versions de l'outil ainsi que l'évaluation de la version B après que les participants aient essayé la version A.

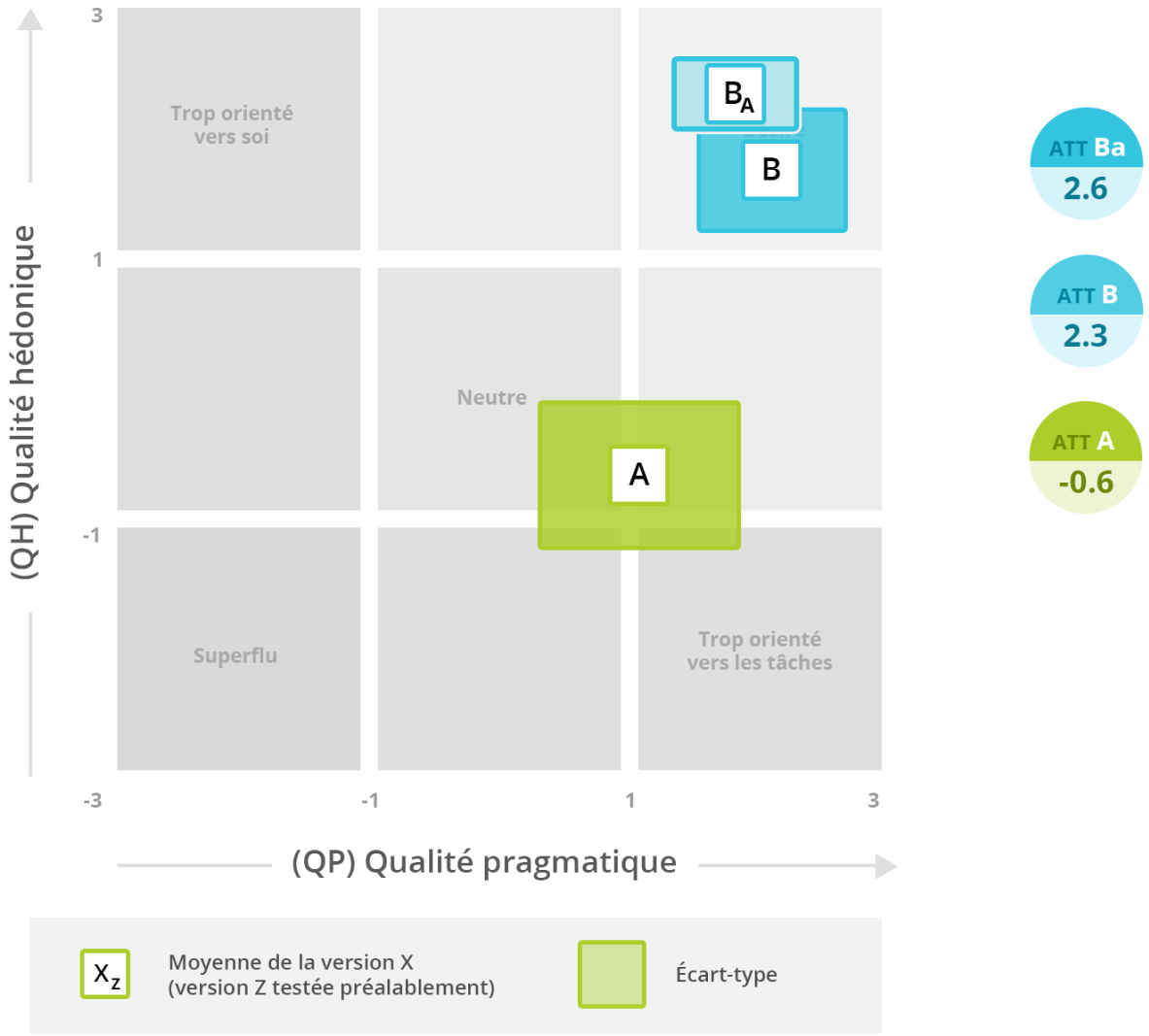


Figure 5-2 : Résultats d'évaluation des deux versions de l'outil et de la nouvelle version après que les participants aient essayé et évalué l'ancienne version

On obtient une attractivité globale encore plus élevée que lorsque la version B (nouvelle) est essayée en premier. En effet, l'attractivité globale passe de 2.3 à 2.6 (écart-type 0.5). La QH obtient également un score plus élevé en passant de 1.7 à 2.3 (écart-type de 0.5). Cependant, la QP passe de 2.1 à 1.7 (écart-type 0.9) ce qui est difficile à expliquer.

Interprétations

L'attractivité globale du produit s'est améliorée de 50 % plutôt que de 45 % (45% lorsque la nouvelle version B est testée en premier). De plus, la nouvelle ATT bénéficie d'un écart-type beaucoup plus faible (0.5) ce qui révèle une plus grande homogénéité d'évaluation. L'amélioration de la QH est également plus importante : 43 % au lieu des 33 % lorsque la version B est essayée en premier. Ceci met en avant quelques caractéristiques importantes de l'EU : elle est dynamique et cumulative, mais surtout elle est *comparative (particulièrement dans une même classe de produit)*. En effet, l'EU avec la nouvelle version (B) est affectée par le fait que le participant a essayé la version A au préalable. Ceci peut également intervenir avec deux produits différents (ex. : voitures). En effet, mon EU avec une voiture X va être affectée par l'ensemble de mes EU précédentes avec cette voiture. Cependant, elle pourra être davantage affectée par mon essai d'une voiture Y, ce qui aura pour effets de changer mes attentes, mes critères de qualité, etc. Ici, on suppose que l'ancienne version (A) crée chez l'utilisateur certaines attentes à l'égard d'une expérience pédagogique interactive. Il était donc prévu que la découverte de la nouvelle version produise un effet positif.

Cependant, pourquoi la QP n'a pas également bénéficié de cet effet? En effet, on obtient une amélioration plus faible (10 %) que lorsque la version B est testée en premier (17 %). Une des explications pourrait être que l'utilisateur essayant la version A s'habitue à l'interface et au placement des différentes informations et zones d'interaction. L'expérience dure une dizaine de minutes. Ainsi, à la découverte de la nouvelle version, l'utilisateur perd légèrement ses repères. Les endroits où il avait l'habitude de trouver ces informations ont changé, et ceci peut donc influencer la perception qu'il se fait de la simplicité de la nouvelle interface.

5.2.3 Passer de la nouvelle à l'ancienne version

L'échantillon E-BA a essayé la nouvelle version (B) puis l'a évaluée. Nous voulions également nous assurer que l'essai de l'ancienne version (A) par la suite montrerait également que la version B fournit une meilleure EU. Ainsi, nous avons estimé qu'il serait pertinent de comparer les résultats partiels de la régression vers l'ancienne version (A) après avoir testé la version B. Notre échantillon E-BA essayait donc la nouvelle version (B), puis répondait au questionnaire et essayait la version A avant de répondre au même questionnaire. Les résultats des ces évaluations sont disponibles dans la figure 5-3.

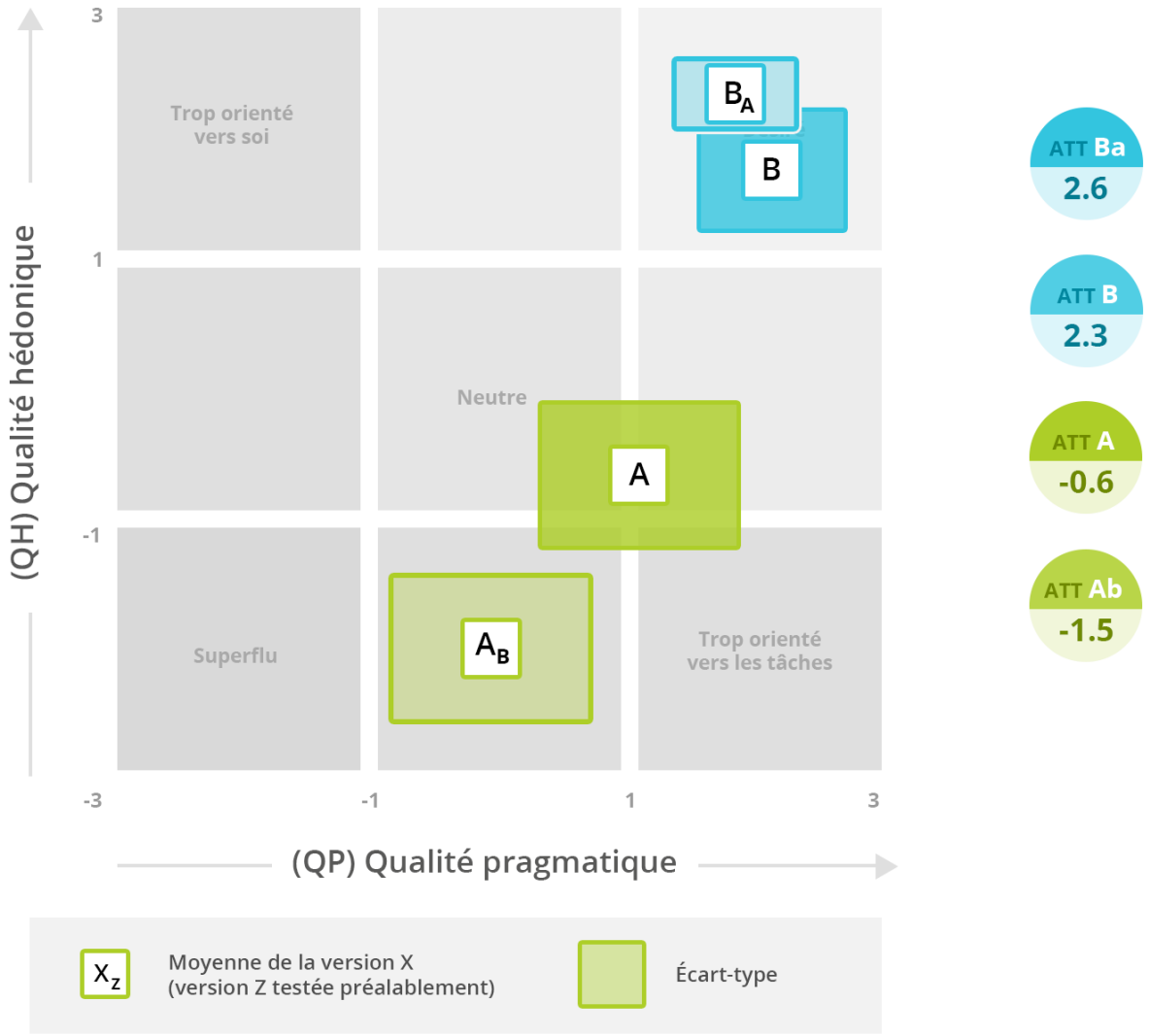


Figure 5-3 : Comparaison des versions et régression vers l'ancienne version A

Le passage de la version B à la version A donne lieu à des évaluations assez sévères. En effet, l'ATT passe de 2.3 à -1.5 (écart-type de 1.5), la QP passe de 2.1 à -0.3 (écart-type de 1.7) et la QH passe de 1.7 à -2.0 (écart-type de 1.1).

Interprétations

Les nouvelles évaluations sont particulièrement sévères et l'impression globale à l'égard de l'ancienne version A est négative (ATT de -1.5). L'EU est donc plutôt négative. La version A après avoir essayé la version B est notée beaucoup plus négativement que lorsque la version A est essayée seule. Ceci confirme à nouveau le caractère cumulatif et comparatif de l'EU. Mais ce qui est intéressant ici est de voir la différence entre l'évolution vers la nouvelle interface et la

régression vers l'ancienne. En effet, le passage de la version A à la version B apporte une amélioration de l'ATT de 50 %, mais le passage de la version B à A apporte une régression de 63 %. Il en va de même pour la QP et la QH dont les améliorations respectives de A vers B sont de 10 % et 43 %, et les régressions respectives (B vers A) sont de 40 % et 62 %. Les utilisateurs sont donc beaucoup plus affectés en valeur absolue par une régression que par une amélioration. Ceci est extrêmement intéressant puisqu'on peut comparer ce phénomène au concept de l'aversion à la perte. L'aversion à la perte a été démontrée par Kahneman et Tversky (1984) et réfère au fait que les individus préfèrent fortement éviter des pertes plutôt que d'acquérir de nouveaux gains. Ainsi, une personne perdant 10 $ perdra plus en satisfaction que ce qu'elle gagnerait en obtenant 10 $. Ce concept est utilisé en économie, mais également en marketing. Ici, on peut voir son application avec l'EU : la transition de la nouvelle version vers l'ancienne (possédant un ATT plus faible) influe négativement sur l'EU de façon plus forte que l'influence positive de la transition inverse.

5.3 Impact sur les autres buts du CIRANO

Nous avons également voulu mesurer l'influence de l'utilisation de la nouvelle version de l'outil sur des facteurs chers au CIRANO tels que l'apprentissage réalisé, la satisfaction de l'utilisateur, l'envie de recommander et de réutiliser l'outil ou encore l'envie de voir des produits similaires se développer.

5.3.1 Impact sur l'apprentissage réalisé

Le questionnaire comportait uniquement deux questions relatives à l'apprentissage. Pour mesurer l'apprentissage réalisé sur le concept de la tragédie des biens communs, nous avons demandé à l'utilisateur d'évaluer sa connaissance de ce concept avant et après avoir participé à l'expérience pédagogique sur une échelle de Likert à cinq points (de -2 à 2). Malheureusement, nous n'avons pas obtenu de résultats concluants.

En effet, seulement 20 % des participants (N=4) ont coché le point milieu « Neutre » lorsqu'on leur demandait s'il était à l'aise avec le concept de la tragédie des biens communs avant d'avoir essayé l'expérience pédagogique. Les autres candidats déclaraient être à l'aise ou très à l'aise. Ainsi, à la vue du faible nombre de participants ne connaissant pas ce concept, il ne nous a pas été possible de mesurer clairement l'impact sur l'apprentissage réalisé.

Parmi les quatre participants ayant coché le point milieu « neutre » (score de 0) pour l'évaluation des connaissances avant l'expérience, deux participants passent d'un score de 0 sur 2 à 1.0 sur 2 avec l'ancienne version, et deux participants passent d'un score de 0 sur 2 à 1.5 sur 2 avec la nouvelle version. Cependant, le nombre de réponses est trop faible pour interpréter ce résultat.

Il faudrait effectuer des évaluations à plus grande échelle auprès des étudiants afin de mesurer l'impact de l'amélioration de l'interface sur l'apprentissage réalisé à l'aide d'un questionnaire permettant d'évaluer les connaissances du candidat.

5.3.2 Impact sur les autres buts du CIRANO

Nous pouvons voir dans le modèle de l'EU supportant l'outil AttrakDiff (figure 1-8) que l'attractivité du produit va entrainer des conséquences comportementales et sentimentales chez l'utilisateur. Certaines de ces conséquences sont particulièrement importantes et constituent des buts pour le CIRANO dans ce projet. Voici les résultats des évaluations de six éléments relatifs au bilan de l'expérience que nous avons choisi d'évaluer (figure 5-4) :

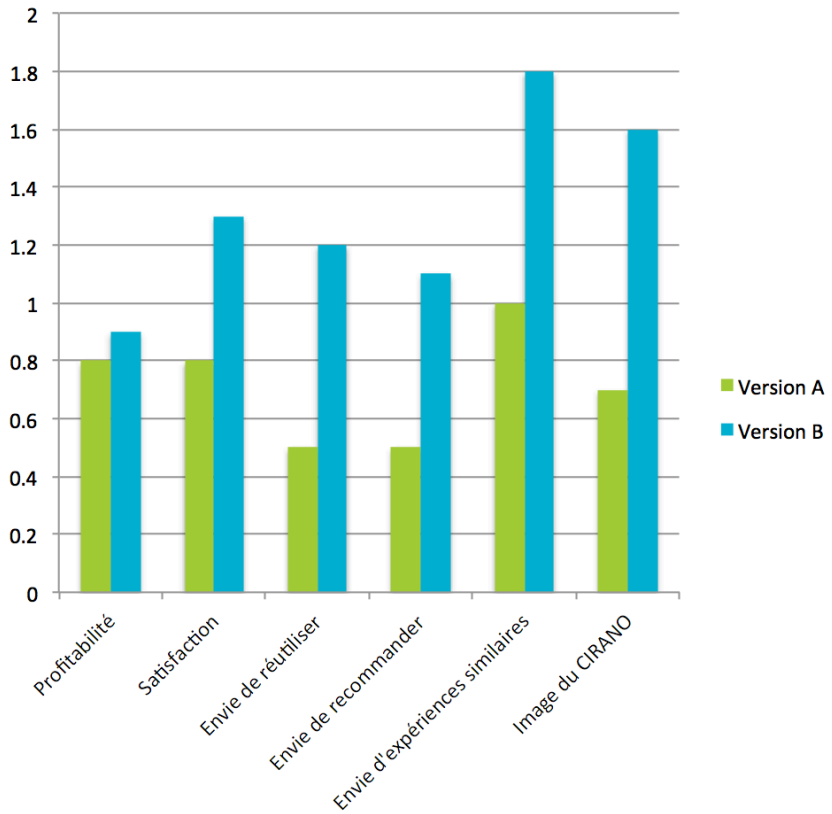


Figure 5-4 : Impact de l'utilisation des versions A et B de l'outil sur différents buts du CIRANO

Les six éléments étaient évalués sur une échelle de Likert à cinq points (de -2 à 2) présentant un continuum allant du désaccord total à l'accord total. Nous avons fait la moyenne des évaluations (0 correspondant à « Neutre », 1 correspondant à « d'accord », 2 à « tout à fait d'accord », et -2 à « Pas du tout d'accord »).

La profitabilité reflète le caractère avantageux de l'expérience pédagogique (ici Ressource commune), c'est-à-dire si le candidat a bénéficié de cette expérience. On observe qu'il n'y a pratiquement pas de différence de profitabilité entre les deux versions. Celle-ci passe de 0.8 (écart-type de 0.8) à 0.9 (écart-type de 0.7). Le niveau de satisfaction de l'utilisateur augmente avec la nouvelle version en passant de 0.8 (écart-type 0.8) à 1.3 (écart-type de 0.7). Les autres éléments évalués voient leur moyenne environ doubler lorsqu'on passe de l'ancienne à la nouvelle version : l'envie de réutiliser le système passe de 0.5 (écart-type de à 1.0) à 1.2 (écart-type de 0.6); l'envie de recommander le système passe de 0.5 (écart-type de 1.3) à 1.1 (écart-type de 0.9); l'envie que d'autres expériences similaires soient développées passe de 1.0 (écart-type de 0.8) à 1.8 (écart-type de 0.4); enfin, l'image du CIRANO passe de 0.7 (écart-type de 1.0) à 1.6 (écart-type de 0.7).

Interprétations

On observe que la profitabilité de l'expérience pédagogique n'est pas influencée par l'amélioration de l'EU avec l'outil. Cela peut provenir du fait que de nombreux participants ont répondu avoir certaines connaissances du concept de la tragédie des biens communs. La nouvelle version de l'outil influe très peu sur la profitabilité, vraisemblablement parce que cette dernière serait plus liée au contenu des expériences pédagogiques.

De plus, l'utilisateur sera satisfait si ses attentes vis-à-vis du système sont comblées. Sa satisfaction augmente de 13 % avec la nouvelle version. Avec l'ancienne version, les participants n'étaient que faiblement d'accord en moyenne pour réutiliser et recommander le système (score de 0.5). Avec la nouvelle version, les envies de recommander et de réutiliser le système sont plus élevées avec des taux d'amélioration respective de 18 % (score de 1.2) et 15 % (score de 1.1). Les participants semblent avoir particulièrement apprécié les expériences pédagogiques puisqu'ils sont en moyenne d'accord pour que des expériences similaires soient développées : score de 1.0 sur 2 avec l'ancienne version et de 1.9 sur 2 avec la nouvelle version (amélioration

de 20 %). Enfin, l'image du CIRANO est meilleure avec la nouvelle version qu'avec l'ancienne puisqu'on passe d'un score de 0.7 à un score de 1.6 (amélioration de 23 % en moyenne). On voit ainsi que l'amélioration de l'EU vécue par le participant à travers les expériences pédagogiques proposées dans ce projet a permis au CIRANO d'atteindre ses buts. Seule la profitabilité ne semble pas avoir bénéficié de l'amélioration de l'EU.

5.3.3 Impact du contenu sur l'évaluation de l'interface

L'impact du contenu correspond à l'impact de l'appréciation par l'utilisateur du contenu traité dans le système. Il était donc intéressant de mesurer si les candidats apprécient ou non le contenu économique présenté et si cela influe sur l'appréciation de l'interface.

Malheureusement, l'ensemble de nos participants a répondu être d'accord ou tout à fait d'accord lorsqu'on leur demande s'ils ont aimé les concepts économiques présentés. Nous n'avons donc pas pu déterminer si les participants n'ayant pas apprécié le contenu auraient pu tout de même apprécier l'interface. Parmi ceux ayant déclaré avoir apprécié le contenu, les scores d'attractivité ou des autres buts de du CIRANO étaient très variables.

Dans ce chapitre, nous avons fait un rappel sur le déroulement des essais et nous avons analysé les résultats d'évaluation de l'EU vécue avec l'ancienne et la nouvelle version par notre échantillon. Nous avons également analysé l'impact de l'amélioration de l'interface sur les différents buts du CIRANO. L'évaluation de l'apprentissage réalisé et de l'impact de l'appréciation du contenu économique sur l'évaluation de l'interface n'a pas pu être réalisée.

CONCLUSION

L'évaluation de l'ancienne interface de l'outil a permis d'identifier de nombreux problèmes ergonomiques qui influent négativement sur l'EU. Plusieurs problèmes étaient récurrents et ont justifié la conception et le développement d'une nouvelle interface visant à améliorer les qualités pragmatiques et hédoniques de l'outil.

Le principal objectif visé par notre recherche, qui était d'améliorer l'EU avec le système, a été atteint. Ceux qui apprennent l'économie auront vraisemblablement une meilleure expérience avec l'outil, et ceci aura des conséquences positives pour le CIRANO telles qu'une plus grande satisfaction de l'utilisateur, des envies plus fortes chez les utilisateurs de réutiliser et de recommander le produit, d'où l'acquisition de meilleures connaissances en économie, et une meilleure image véhiculée. L'amélioration de l'attractivité du produit a également aidé à l'obtention de nouveau financement pour le projet. Ces améliorations ont également eu un impact sur l'interface qu'utilise le professeur, afin que la mise en place et le lancement des expériences pédagogiques soient faciles et agréables. Ceci est un point important pour aider le CIRANO à convaincre ces professeurs d'utiliser l'outil pédagogique.

Plusieurs limitations de notre recherche sont à prendre en compte. Le nombre de participants (N=20) ayant pris part aux évaluations est faible de sorte que les résultats ne sont pas robustes. De plus, notre échantillon est peu représentatif des utilisateurs finaux du système (étudiants du Secondaire V et de CEGEP) à cause de l'âge et du niveau d'éducation de nos participants.

Le contexte dans lequel l'expérience du participant a lieu joue un rôle important dans l'évaluation de celle-ci. Le contrôle du contexte était ici limité étant donné que certains essais ont été effectués à distance. Il aurait été préférable de pouvoir les effectuer en classe.

D'autres itérations devront être faites afin de poursuivre l'amélioration des expériences pédagogiques. L'accès aux utilisateurs finaux a été restreint de sorte qu'il a été difficile d'impliquer les étudiants durant le processus de conception de la nouvelle interface. De plus, les prochaines itérations devraient impliquer plus fortement les étudiants en utilisant des méthodes de recueil de données plus riches telles que le penser tout haut.

Concernant la nouvelle version du système, il faudra évaluer et valider le nouvel outil d'administration ainsi que 12 expériences pédagogiques (présentement en version alpha) sur

lesquelles nous avons travaillé. Nous avons présenté ici la seule expérience pédagogique qui était suffisamment avancée et exempte de bogues pour faire l'objet de tests avec des utilisateurs.

Un très faible nombre de participants de notre échantillon n'avait pas de connaissance du concept de la tragédie des biens communs qui était au centre de l'expérience pédagogique que nous avons améliorée, de sorte qu'il ne nous a pas été possible d'évaluer l'apprentissage de l'économie. Il est clair qu'il faudra aussi évaluer l'impact de l'amélioration de l'EU sur l'apprentissage de l'économie.

Enfin, nous avons également mentionné que l'appréciation par l'utilisateur du contenu ou du domaine qui est présenté dans une interface devrait être prise en compte lors de l'évaluation de l'EU. Si le contenu ou le domaine présenté est complexe et ne plait pas à l'utilisateur, cela pourrait avoir un effet négatif sur l'EU, même si l'interface est bien conçue. La pertinence et l'appréciation du contenu du domaine par l'utilisateur influent vraisemblablement sur l'évaluation de l'EU. Il pourrait donc être intéressant d'inclure cette distinction dans des questionnaires tels que AttrakDiff.

BIBLIOGRAPHIE

Abras, C., Maloney-Krichmar, D., Preece, J. (2004). User-Centered Design. In Bainbridge, W. (Ed.). Encyclopedia of Human-Computer Interaction, 445-456.

Arhippainen, L., Tähti, M. (2003). Empirical Evaluation of User Experience in Two Adaptive Mobile Application Prototypes. Communication at the 2nd International Conference on Mobile and Ubiquitous Multimedia, Norrköping, Sweden, 27-34.

Baek, E., Cagiltay, K., Boling, E., Frick, T. (2007). User-centered design and development. In J. Spector, M. Merrill, J. van Merrienboer, & M. Driscoll (Eds.), Handbook of research on educational communications and technology, New York: Routledge/Taylor & Francis Group, 659–670.

Ball, S., Eckel, C., Rojas, C. (2006). Technology Improves Learning in Large Principles of Economics Classes: Using Our WITS, The American Economic Review, 96(2) (May), 442-446.

Bastien, J.M.C., Scapin, D.L. (1993). Critères Ergonomiques pour l'Évaluation d'Interfaces Utilisateurs (version 2.1). Technical report May 1993. INRIA. Programme 3 Artificial intelligence, cognitive systems, and man-machine interaction, Rocquencourt, France.

Bastien, J.M.C., Scapin, D.L., Leulier, C. (1999). The ergonomic criteria and the ISO/DIS 9241-10 dialogue principles: a pilot comparison in an evaluation task. Interacting with computers, 11(3), 299-322.

Chin, J.P., Diehl, V.A., Norman, K.L. (1988). Development of an instrument measuring user satisfaction of the human–computer interface. In: Proceedings of the CHI 88 Conference on Human Factors in Computing Systems. ACM Press, New York, 213– 218.

Csikszentmihalyi, M. (1975). Beyond boredom and anxiety: The experience of play in work and games. San Francisco, CA, U.S.A.: Jossey-Bass.

Davis, F.D. (1989). Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology. MIS Quarterly 13 (3), 319–340.

Desmet, P.M.A., Overbeeke, C.J., Tax, S.J.E.T. (2001). Designing Products with Added Emotional Value: Development and Application of an Approach for Research through Design. The Design Journal, 4 (1), 32-47.

Desmet, P.M.A., Hekkert, P. (2007). Framework of product experience. International Journal of Design. 1(1), 57-66.

Dickie, M. (2006). Do Classroom Experiments Increase Learning in Introductory Microeconomics? The Journal of Economic Education, 37(3), 267-288, DOI: 10.3200/JECE.37.3.

Djamasbi, S., Tullis, T., Hsu, J., Mazuera, E., Osberg, K., Bosch, J. (2007). Gender Preferences in Web Design: Usability Testing through Eye Tracking. Proceedings of the Thirteenth Americas Conference on Information Systems, 1-8.

Friess, E. (2008). The User-centered Design Process: Novice Designers' Use of Evidence in designing from Data, Carnegie Mellon University.

Gegenfurtner, K., Sharpe, L. (2001). Color vision: From genes to perception. Cambridge university press, New York.

Gould, J.D., Boies, S.,J., Ukelson, J. (1997). How to Design Usable Systems, How to design usable systems. Handbook of human-computer interaction, 2, 231-54.

Gould, J.D., Lewis, C. (1983). Designing for usability—key principles and what designers think. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '83), in Janda (Ed.). ACM, New York, NY, USA, 50-53. DOI=10.1145/800045.801579 http://doi.acm.org/10.1145/800045.801579

Gould, J.D., Lewis, C. (1985). Designing for usability: key principles and what designers think. Communications of the ACM 28, 3 (March 1985), 300-311, DOI:10.1145/3166.3170 http://doi.acm.org/10.1145/3166.3170

Hardin, G. (1968). The tragedy of the commons. Science, 162, 1243–1248

Hartwell L.H., Hopfield J.J, Leibler S., Murray A.W. (1999). From molecular to modular cell biology. Nature, 402, 47-52.

Hansen, W. J. (1971). User engineering principles for interactive systems. Communication at the fall joint computer conference, Argonne, Illinois, États-Unis, 523-532.

Hassenzahl, M., Beu, A., Burmester, M. (2001). Engineering Joy. IEEE Software, 70-76.

Hassenzahl, M. (2003). The Thing and I: Understanding the Relationship between User and Product. In Funology. Springer Netherlands, 2005, 31-42.

Hassenzahl, M., Burmester, M., Koller, F. (2003). AttrakDiff: Ein Fragebogen zur Messung

wahrgenommener hedonischer und pragmatischer Qualität. [AttrakDiff: A questionnaire to measure perceived hedonic and pragmatic quality.] In J. Ziegler & G. Szwillus (Eds.), Mensch & Computer 2003. Interaktion in Bewegung, Stuttgart, Germany, 187–196.

Hassenzahl, M. (2004). The interplay of beauty, goodness, and usability in interactive products. Human-Computer Interaction, 19, 319-349.

Hassenzahl, M., Tractinsky, N. (2006). User Experience - a Research Agenda. Behaviour and Information Technology, 25(2), 91-97.

Hassenzahl, M. (2008). User experience (UX): towards an experiential perspective on product

quality. Proceedings of IHM'08, New York, NY, USA, 11-15.

Hassenzahl, M. (2007). AttrakDiff. http://www.attrakdiff.de (acces on March 29th)

Hassenzahl, M., Diefenbach, S., Goritz, A. (2010). Needs, affect, and interactive products – facets of user experience. Interacting with Computers, 22(5), 353–362.

Hassenzahl, M., Monk, A. (2010). The inference of perceived usability from beauty. Human–Computer Interaction, 25(3), 235-260.

Hassenzahl, M. (2013). User Experience and Experience Design. In Soegaard, M., Dam R.F., (Eds.). "The Encyclopedia of Human-Computer Interaction, 2nd Ed.". Aarhus, Denmark.

Herzberg, F. (1966). Work and the Nature of Man (Vol. 3). Cleveland, Ohio, USA: World Publishing Company.

Human Factors Research Group (2002). Available from: http://www.ucc.ie/hfrg/.

Holt, C., Laury, S. (1997). Classroom Games: Voluntary Provision of a Public Good.

Journal of Economic Perspectives, 11, 209-215.

International Organization for Standardization (2010). (Created on 1994). Ergonomics of human system interaction - Part 210: Human-centred design for interactive systems. ISO 9241-210. (Acces on Iso.org on March 27th)

International Organization for Standardization (1991). Software Product Evaluations—Quality Characteristics and Guidelines for their Use, ISO DIS 9126.

International Organization for Standardization (1999). Human-centred design processes for interactive systems, ISO 13407.

Kahneman, D., Tversky, A. (1984). Choices, Values, and Frames. American Psychologist, 39 (4), 341–350. doi:10.1037/0003-066x.39.4.341

Karapanos, E. (2013). User Experience Over Time. In Modeling Users' Experiences with Interactive Systems, Springer Berlin Heidelberg, 57-83.

Kerkow, D. (2007). Don't have to know what it is like to be a bat to build a radar reflector– Functionalism in UX. International Conference HCI 2007, Lancaster, UK. 19-25.

Kort, J., Vermeeren, A., Fokker, J.E. (2007). Conceptualizing and Measuring UX. Communication at the International Conference HCI 2007, Lancaster, UK. 57-64.

Lalime, T., Michaud, P.-C. (2012). Littératie financière et préparation à la retraite au Québec et dans le reste du Canada, Série Scientifique, Montréal Décembre 2012.

Larouche, A. (2011). Enquête visant à déterminer les dimensions de l'expérience utilisateur. Rapport de projet de maîtrise, Polytechnique Montréal, non publié.

Laugwitz, B.; Held, T. & Schrepp, M. (2008). Construction and evaluation of a user experience questionnaire. In Holzinger, A. (Ed.): USAB 2008, LNCS 5298, 63-76.

Law, E., Roto, V., Vermeeren, A., Kort, J., & Hassenzahl, M. (2008). Towards a Shared Definition for User Experience. Special Interest Group in CHI'08 Proceedings, Florence, Italy, 2395-2398.

Law, E., Roto, V., Hassenzahl, M., Vermeeren, A., Kort, J. (2009). Understanding, Scoping and Defining User experience: A Survey Approach. Proceedings of CHI'09, Boston, MA, USA, 719-728.

Lewis, J.R. (1995). IBM Computer Usability Satisfaction Question- naires: Psychometric Evaluation and Instructions for Use/James R. Lewis. International Journal of Human–Computer Interaction, 7 (1), 57–78.

Lin, H.X., Choong, Y.Y., & Salvendy, G. (1997). A proposed index of usability: a method for comparing the relative usability of different software systems. Behaviour & Information Technology, Chicago, 16(4-5), 267-277.

Lindgaard, G., Fernandes, G., Dudek, C., Brown, J. (2006). Attention web designers: You have 50 milliseconds to make a good first impression! Behavious & Information Technology, 25(2), 115-126.

Lund, A.M. (2001). Measuring usability with the USE questionnaire. Usability Interface, 8(2), 3-6.

Mahlke, S., Thüring, M. (2007). Usability, aesthetics and emotions in human-technology interaction. International Journal of Psychology, 42, 253–264.

Mao, J.-Y., Vredenburg, K., Smith, P., Carey, T. (2005). The state of user-centered design practice. Commun. ACM 48, 3 (March 2005), 105-109. DOI=10.1145/1047671.1047677 http://doi.acm.org/10.1145/1047671.1047677

Molich, R., Nielsen, J. (1990). Improving a human-computer dialogue, Communications of the ACM 33 , 3 (March), 338-348.

Nagamachi, M. (2002). Kansei engineering as a powerful consumer-oriented technology for product development. Applied Ergonomics, 33(3), 289-294.

Nielsen, J., Molich, R. (1990). Heuristic evaluation of user interfaces, Proceedings ACM CHI'90 Conf. (Seattle, WA, 1-5 April), 249-256.

Nielsen, J. (1993). Usability Engineering, Academic press, San Diego, CA

Nielsen, J. (1994a). Enhancing the explanatory power of usability heuristics. Proceedings ACM CHI'94 Conf. (Boston, MA, April 24-28), 152-158.

Nielsen, J. (1994b). Heuristic evaluation. In Nielsen, J., and Mack, R.L. (Eds.), Usability Inspection Methods, John Wiley & Sons, New York, NY.

Norman, D., Draper, S. (1986). User Centered System Design; New Perspectives on Human-Computer Interaction. L. Erlbaum Assoc. Inc., Hillsdale, N.J., USA.

Norman, D., Miller J., Henderson, A. (1995). What You See, Some of What's in the Future, And How We Go About Doing It. HI at Apple Computer. Proceedings of CHI 1995, Denver, Colorado, USA.

Norman, D. (1988). The Design of Everyday Things. New York: Basic Books. ISBN 978-0-465-06710-7.

Norman, D., (2003). 3 ways good design makes you happy , Conference at TED 2003. Available on http://www.ted.com/talks/don_norman_on_design_and_emotion#t-652569

Norman, D., Nielsen J. The definition of user experience. Nielsen Norman group website: http://www.nngroup.com/articles/definition-user-experience (Accès le 24 Mars 2014)

Preece, J., Rogers, Y., Sharp, H., Benyon, D., Holland, S., Carey, S. (1994). Human–Computer Interaction, Addison-Wesley, Reading, MA

Provost, G. (2012). Étude des expériences des utilisateurs avec des produits interactifs. Rapport de projet de maitrise, Polytechnique Montréal, non publié.

Robert, J.-M. (2003). Que faut-il savoir sur les utilisateurs pour réaliser des interfaces de qualité? In G.A. Boy (Ed.). Ingénierie cognitive : IHM et Cognition, Hermès, France, 249-284.

Robert, J.-M., Lesage, A. (2011). From usability to user experience with user interfaces (chap. 14). G.A. Boy (Ed.). Handbook of Human-Computer Interaction. A human-centered design approach. Ashgate, U.K., 303-320

Robert, J.-M., Lesage, A. (2011). Designing and evaluating user experience (chap. 15). In G.A. Boy (Ed.). Handbook of Human-Computer Interaction. A human-centered design approach. Ashgate, U.K., 321-338.

Robert, J.-M. (2014). Defining and Structuring the Dimensions of User Experience with Interactive Products. Proceedings of HCII 2014 (Human Computer Interaction International), 22-27 June, Crête, Grèce. 10 pages.

Rochefeuille, M. (2013). Analyse des dimensions de l'expérience utilisateur avec des produits intéractifs. Mémoire de maitrise, Polytechnique Montréal, non publié.

Roto, V. (2006). User experience building blocks. In The 2nd COST294-MAUSE International Open Workshop, (Vol. 14).

Roto, V. (2007). User Experience from Product Creation Perspective. Towards a UX Manifesto: COST294-MAUSE, Lancaster, UK, 31-34.

Roto, V., Obrist, M., Väänänen-Vainio-Mattila, K. (2009). User Experience Evaluation Methods in Academic and Industrial Contexts. Proceedings of UXEM'09, Uppsala, Sweden.

Scapin, D.L., Bastien, J.M.C. (1997). Ergonomic criteria for evaluating the ergonomic quality of interactive systems. Behaviour and Information Technology, 6 (4-5), 220-231.

Scapin, D., Leulier, C., Vanderdonckt, J., Mariage, C., Bastien, C., Farenc, C., Bastide, R. (2000). A framework for organizing web usability guidelines. In Proceedings of the Sixth Conference on Human Factors & the Web.

Shackel, B., Richardson, S. (1991). Human Factors for Informatics Usability, Cambridge University Press, Cambridge, 21-38.

Shneiderman, B. (1998). Designing the user interface : Strategies for effective human-computer interaction. 2nd edition Addison Wesley, Reading.

Smith, S. L., Mosier, J. N. (1986). Guidelines for designing user interface software. Report ESD-TR-86-278. Bedford, MA: The MITRE Corporation.

Statcan (2006). Les tendances de la composition selon l'âge des étudiants et des diplômés collégiaux et universitaires : http://www.statcan.gc.ca/pub/81-004-x/2010005/article/11386-fra.htm (Accès le 2 Juillet 2014).

Watson, D. (1992). Correcting for Acquiescent Response Bias in the Absence of a Balanced Scale: An Application to Class Consciousness. Sociological Methods Research, August, 21(1) 52-88. doi:10.1177/0049124192021001003.

ANNEXE A – LES 10 HEURISTIQUES DE NIELSEN

Tiré su site du Nielsen Norman group : http://www.nngroup.com/articles/ten-usability-heuristics/

Visibility of system status

The system should always keep users informed about what is going on, through appropriate feedback within reasonable time.

Match between system and the real world

The system should speak the users' language, with words, phrases and concepts familiar to the user, rather than system-oriented terms. Follow real-world conventions, making information appear in a natural and logical order.

User control and freedom

Users often choose system functions by mistake and will need a clearly marked "emergency exit" to leave the unwanted state without having to go through an extended dialogue. Support undo and redo.

Consistency and standards

Users should not have to wonder whether different words, situations, or actions mean the same thing. Follow platform conventions.

Error prevention

Even better than good error messages is a careful design which prevents a problem from occurring in the first place. Either eliminate error-prone conditions or check for them and present users with a confirmation option before they commit to the action.

Recognition rather than recall

Minimize the user's memory load by making objects, actions, and options visible. The user should not have to remember information from one part of the dialogue to another. Instructions for use of the system should be visible or easily retrievable whenever appropriate.

Flexibility and efficiency of use

Accelerators -- unseen by the novice user -- may often speed up the interaction for the expert user such that the system can cater to both inexperienced and experienced users. Allow users to tailor frequent actions.

Aesthetic and minimalist design

Dialogues should not contain information which is irrelevant or rarely needed. Every extra unit of information in a dialogue competes with the relevant units of information and diminishes their relative visibility.

Help users recognize, diagnose, and recover from errors

Error messages should be expressed in plain language (no codes), precisely indicate the problem, and constructively suggest a solution.

Help and documentation

Even though it is better if the system can be used without documentation, it may be necessary to provide help and documentation. Any such information should be easy to search, focused on the user's task, list concrete steps to be carried out, and not be too large.

ANNEXE B – LES SEPT PRINCIPES DE LA NORME ISO 9241 (110)

1 - Is the dialogue suitable for the user's task and skill level? (Suitability for the task)

"A dialogue is suitable for a task when it supports the user in the effective and efficient completion of the task. In a dialogue which is suitable for the task, the user is enabled to focus on the task itself rather than the technology chosen to perform that task."

2 - Does the dialogue make it clear what the user should do next? (Self-descriptiveness)

"A dialogue is self-descriptive to the extent that at any time it is obvious to the users which dialogue they are in, where they are within the dialogue, which actions can be taken and how they can be performed."

3 - Is the dialogue consistent? (Conformity with user expectations)

"A dialogue conforms with user expectations if it corresponds to predictable contextual needs of the user and to commonly accepted conventions."

4 - Does the dialogue support learning? (Suitability for learning)

"A dialogue is suitable for learning when it supports and guides the user in learning to use the system."

5 - Can the user control the pace and sequence of the interaction? (Controllability)

"A dialogue is controllable when the user is able to initiate and control the direction and pace of the interaction until the point at which the goal has been met."

6 - Is the dialogue forgiving? (Error tolerance)

"A dialogue is error-tolerant if, despite evident errors in input, the intended result may be achieved with either no or minimal corrective action by the user. Error tolerance is achieved by means of damage control, error correction, or error management to cope with errors that occur."

7 - Can the dialogue be customised to suit the user? (Suitability for individualisation)

"A dialogue is capable of individualization when users can modify interaction and presentation of information to suit their individual capabilities and needs."

ANNEXE C – LES CRITÈRES DE BASTIEN & SCAPIN

1. GUIDAGE Le guidage est l'ensemble des moyens mis en œuvre pour conseiller, orienter, informer, et conduire l'utilisateur lors de ses interactions avec l'ordinateur (messages, alarmes, labels, etc.). • Un bon guidage facilite l'apprentissage et l'utilisation du système en permettant à l'utilisateur de savoir, à tout moment où il se trouve dans une séquence d'interactions, ou dans l'accomplissement d'une tâche, de connaître les actions permises ainsi que leurs conséquences, d'obtenir de l'information supplémentaire.	**1.1. Incitation** Moyens pour l'utilisateur de connaître les actions disponibles, les alternatives, l'état du contexte où il se trouve. **1.2. Groupement/Distinction par le format ou le groupement** Il prend en compte la localisation et le format pour indiquer les relations entre les divers items affichés, et leur appartenance ou non à une même classe d'items. **1.3. Feedback Immédiat** Une réponse doit être fournie à l'utilisateur le renseignant sur l'action accomplie et sur son résultat, ceci, avec un délai de réponse approprié et homogène selon les types de transactions (souvent, délai immédiat). **1.4. Lisibilité.** Caractéristiques de présentation des informations qui doivent en faciliter la lecture (typographie, espacement, ...)
2. CHARGE DE TRAVAIL La charge de travail concerne l'ensemble des éléments de l'interface qui ont un rôle dans la réduction de la charge perceptive ou mnésique des utilisateurs et dans l'augmentation de l'efficacité du dialogue. • Plus la charge de travail est élevé, plus grands sont les risques d'erreurs. De même, moins l'utilisateur sera distrait par des informations non pertinentes, plus il pourra effectuer sa tâche	**2.1. Brièveté** Limiter le travail de lecture et d'entrée d'information : • Concision : pour les éléments individuels d'entrée ou de sortie. • Actions Minimales : limiter le nombre d'actions successives nécessaires pour atteindre un but. **2.2. Densité informationnelle** Réduire au maximum la charge informationnelle

efficacement. Par ailleurs, plus les actions requises seront courtes, plus rapides seront les interactions.	
3. CONTROLE EXPLICITE Le contrôle explicite concerne la transparence entre les actions de l'utilisateur et les réponses du système, le contrôle qu'ont les utilisateurs sur le traitement de leurs actions. • Quand les entrées des utilisateurs sont explicitement définies par eux-mêmes et sous leur contrôle, les ambiguïtés et les erreurs sont limitées. De plus, le contrôle qu'ont les utilisateurs sur le dialogue est un facteur d'acceptation du système.	**3.1. Actions Explicites** La relation entre les actions de l'utilisateur et les réponses du système doit être explicite, c'est-à-dire que le système doit exécuter seulement les actions demandées par l'utilisateur et au moment où il(elle) les demande. **3.2. Contrôle Utilisateur** L'utilisateur doit toujours pouvoir contrôler le déroulement des traitements informatiques en cours.
4. ADAPTABILITE L'adaptabilité d'un système concerne sa capacité à réagir selon le contexte, et selon les besoins et préférences des utilisateurs. • Plus les façons d'effectuer une même tâche sont diverses, plus les chances que l'utilisateur puisse choisir et maîtriser l'une d'entre elles, au cours de ses apprentissages, sont importantes. Il faut donc fournir à l'utilisateur des procédures, options, et commandes différentes lui permettant d'atteindre un même objectif. Par ailleurs, une interface ne peut convenir à la fois à tous ses utilisateurs potentiels. Pour	**4.1. Flexibilité** Réfère aux moyens disponibles à l'utilisateur pour personnaliser son interface de façon à prendre en compte ses stratégies de travail et/ou ses habitudes et les exigences de ses tâches. **4.2. Prise en compte de l'expérience** Les différents moyens disponibles pour prendre en compte le niveau d'expérience de l'utilisateur (expérimenté, débutant).

qu'elle n'ait pas d'effets négatifs sur l'utilisateur, cette interface doit, selon les contextes, s'adapter à l'utilisateur.	
5. GESTION DES ERREURS Le critère Gestion des erreurs concerne tous les moyens permettant d'éviter ou de réduire les erreurs, et de les corriger lorsqu'elles surviennent. Les erreurs sont ici considérées comme des saisies de données incorrectes, des saisies dans des formats inadéquats, des saisies de commandes avec une syntaxe incorrecte. • Les interruptions provoquées par les erreurs ont des conséquences négatives sur l'activité des utilisateurs. De manière générale, elles rallongent les transactions et perturbent la planification et diminuent la performance.	**5.1. Protection contre les erreurs** Réfère aux moyens disponibles pour détecter et prévenir (avant validation) : les erreurs d'entrée de données, les erreurs de commandes, les actions à conséquences destructrices **5.2. Qualité des messages d'erreurs** Réfère à l'expression et au contenu des messages d'erreur : leur pertinence, leur facilité de lecture, leur précision quant à la nature des erreurs, l'indication des actions de correction **5.3. Correction des erreurs** Réfère aux moyens disponibles pour l'utilisateur pour corriger immédiatement ses erreurs.
6. HOMOGÉNÉITÉ / COHÉRENCE Le critère Homogénéité / Cohérence ou « consistence », se réfère à la façon avec laquelle les choix de conception de l'interface (codes, dénominations, formats, procédures, etc.) sont conservés pour des contextes identiques, et sont différents pour des contextes différents.	Les procédures, labels, commandes, etc., sont d'autant mieux reconnus, localisés et utilisés, que leur format, localisation, ou syntaxe sont stables d'un écran à l'autre, d'une session à l'autre. Dans ces conditions le système est davantage prévisible et les apprentissages plus généralisables ; les erreurs sont réduites. Le manque d'homogénéité peut augmenter considérablement le temps de recherche. Le manque d'homogénéité est aussi une raison importante de refus d'utilisation.

7. SIGNIFIANCE CODES & DENOMINATIONS Le critère Signifiance des Codes et Dénominations concerne l'adéquation entre l'objet ou l'information affichée ou entrée, et son référent. Des codes et dénominations " signifiants " disposent d'une relation sémantique forte avec leur référent.	Lorsque le codage est signifiant, le rappel et la reconnaissance sont meilleurs. De plus, les codes et dénominations non significatifs pour les utilisateurs peuvent leur suggérer des opérations inappropriées et ainsi conduire à des erreurs.
8. COMPATIBILITE Le critère Compatibilité se réfère à l'accord pouvant exister entre les caractéristiques des utilisateurs (mémoire, perceptions, habitudes, compétences, âge, attentes, etc.) et des tâches, d'une part, et l'organisation des sorties, des entrées et du dialogue d'une application donnée, d'autre part. De plus, la Compatibilité concerne également le degré de similitude entre divers environnements ou applications.	Le transfert d'information d'un contexte à un autre est d'autant plus rapide et efficace que le volume d'information à recoder par l'utilisateur est réduit. L'efficacité est accrue lorsque les procédures nécessaires à l'accomplissement de la tâche sont compatibles avec les caractéristiques psychologiques des utilisateurs, les procédures et les tâches sont organisées de manière à respecter les attentes, ou habitudes des utilisateurs ; les traductions, les transpositions, les interprétations, ou références à la documentation sont minimisées. Les performances sont meilleures lorsque l'information est présentée sous une forme directement utilisable.

ANNEXE D – CERTIFICAT D'ÉTHIQUE À LA RECHERCHE (CER)

POLYTECHNIQUE MONTRÉAL
LE GÉNIE EN PREMIÈRE CLASSE

CERTIFICAT D'ACCEPTATION D'UN PROJET DE RECHERCHE
PAR LE COMITÉ D'ÉTHIQUE DE LA RECHERCHE
AVEC DES ÊTRES HUMAINS DE POLYTECHNIQUE MONTRÉAL

Le 29 juillet 2014

M. Aurélien Salomon
M. Jean-Marc Robert
Département de mathématiques et génie industriel
Polytechnique Montréal

N/Réf : Dossier CÉR-13/14-25

Messieurs,

J'ai le plaisir de vous informer que les membres du Comité d'éthique de la recherche ont procédé à l'évaluation en comité restreint de votre projet de recherche intitulé « *Processus d'amélioration de l'expérience utilisateur avec un système d'aide à l'apprentissage de notions d'économie* » et en ont recommandé l'approbation sur la base des précisions/modifications apportées à votre documentation en date du 13 et du 28 juillet 2014.

Veuillez noter que le présent certificat est valable pour une durée d'un an, soit du **29 juillet 2014 au 28 juillet 2015**, pour le projet tel que soumis au Comité d'éthique de la recherche avec des sujets humains.

Nous vous prions de nous faire parvenir un bref **rapport annuel** (http://www.polymtl.ca/recherche/document/deonto.php) afin de renouveler votre certificat au moins un mois avant l'expiration du présent certificat. La secrétaire du Comité d'éthique de la recherche avec des sujets humains devra également être informée de toute modification qui pourrait être apportée ultérieurement au protocole expérimental, de même que de tout problème imprévu pouvant avoir une incidence sur la santé et la sécurité des personnes impliquées dans le présent projet de recherche (sujets, professionnels de recherche ou chercheurs).

Je vous souhaite bonne chance dans vos travaux de recherche,

Farida Cheriet, présidente
Comité d'éthique de la recherche avec des êtres humains

c.c.: Céline Roehrig, DRI

ANNEXE E – QUESTIONNAIRE D'ÉVALUATION

Ce questionnaire a pour objectif de connaître votre opinion sur l'apprentissage des notions d'économie que vous pensez avoir réalisé s'il y a lieu en utilisant le système, et sur la qualité de l'interface utilisateur de ce système.
Vous avez fait une expérience nommée "ressource commune" qui porte sur le concept économique de la "tragédie des biens communs". Cette expérience a été développée par le CIRANO (centre de recherche).
Il n'y a pas de bonnes ou de mauvaises réponses.

Votre identifiant (ex., P1, P2, ...) :.....

Sexe : ☐ Masculin ☐ Féminin **Âge :** ..... ans **Profession :** .....

Dernier niveau de scolarité complété :
☐ Avant le secondaire ☐ Secondaire ☐ CEGEP (ou collège)
☐ Baccalauréat ☐ Maîtrise ☐ Doctorat

SVP prêter attention aux définitions des 2 termes suivants:

1) Le contenu: il englobe les concepts économiques sur lesquels l'expérience s'appuie, dans la Tragédie des biens communs.

2) L'interface utilisateur: elle comprend l'ensemble des éléments que vous avez vus à l'écran et avec lesquels vous avez interagi : par ex., visuel de la page web, positionnement des éléments, zone de textes, boutons, images, animations, etc..

1) Le contenu *(Cochez la réponse qui vous convient) :*

	Pas du tout d'accord	Pas d'accord	Neutre	Plutôt d'accord	Tout à fait d'accord
1.1 - Avant de commencer l'expérience, j'étais à l'aise avec les concepts économiques présentés dans la Tragédie des biens communs.					
1.2 - Après l'expérience, j'étais à l'aise avec les concepts économiques présentés dans la Tragédie des biens communs.					
1.3 - J'ai aimé les concepts économiques présentés					

2) L'interface utilisateur était :

Confuse	☐-----☐-----☐-----☐-----☐-----☐-----☐	Structurée
Non pratique	☐-----☐-----☐-----☐-----☐-----☐-----☐	Pratique
Imprévisible	☐-----☐-----☐-----☐-----☐-----☐-----☐	Prévisible
Compliquée	☐-----☐-----☐-----☐-----☐-----☐-----☐	Simple
De mauvais goût	☐-----☐-----☐-----☐-----☐-----☐-----☐	Élégante
Bas de gamme	☐-----☐-----☐-----☐-----☐-----☐-----☐	Haut de gamme
Non imaginative	☐-----☐-----☐-----☐-----☐-----☐-----☐	Créative
Ennuyante	☐-----☐-----☐-----☐-----☐-----☐-----☐	Captivante
Mauvaise	☐-----☐-----☐-----☐-----☐-----☐-----☐	Bonne
Laide	☐-----☐-----☐-----☐-----☐-----☐-----☐	Belle

	Pas du tout d'accord	Pas d'accord	Neutre	Plutôt d'accord	Tout à fait d'accord
3) Bilan					
3.1 - Cette expérience m'a été profitable					
3.2 - Je suis satisfait de l'expérience					
3.3 - Je serais tenté de réutiliser le système					
3.4 - Je recommanderais le système à mon entourage					
3.5 - J'aimerais que d'autres expériences pédagogiques similaires soient développées					
3.6 - Ces expériences m'ont donné une bonne image du CIRANO (centre de recherche ayant créé et développé le système)					

4) Questions facultatives:

4.1 - Qu'avez vous aimé dans cette expérience ?
.....
.....
.....

4.2 - Avez-vous eu des difficultés avec le système? Si oui, lesquelles ?
.....
.....
.....
.....

4.3 - Comment pourrions-nous améliorer le système ? .....
.....
.....
.....
.....

ANNEXE F – CAPTURES D'ÉCRANS DE L'ANCIENNE ET LA NOUVELLE VERSION DE « RESSOURCE COMMUNE » PAR PAGE

Page de consigne (ancienne puis nouvelle)

vm117.cirano.qc.ca:5889

Instructions

Tu vas participer à une expérience pédagogique, qui dure plusieurs périodes.

Dans cette expérience, tu vas faire partie d'un groupe de **3** joueurs, qui va rester le même durant toute la durée du jeu.

Ce groupe se retrouve chaque année au lac Piri-Piri pour une journée de pêche.

Au début du jeu, le lac contient **550** poissons.

Chacun d'entre vous est autorisé à pêcher jusqu'à **30** poissons.

Cependant, vous êtes informé que, pour assurer le renouvellement des poissons dans le lac, l'ensemble du groupe ne devrait pas en pêcher plus de **30** durant la période.

Chaque poisson pêché en plus de cette quantité entraîne une baisse du nombre de poissons dans le lac à la période suivante.

Par exemple :
- il y a 550 poissons dans le lac;
- il est recommandé de ne pas pêcher plus de 30 poissons dans le lac;
- tu as décidé de pêcher 5 poissons;
- à la fin de la période, 35 poissons ont été pêchés par ton groupe.

5 poissons de plus que ce qui était recommandé ont été pêchés. Au début de la période suivante, il y aura : 550 - (5x2) = 540 poissons dans le lac.

Si, au fur et à mesure des périodes de jeu, le nombre de poissons dans le lac devient plus petit que le niveau recommandé de pêche pour le groupe, alors les membres ne pourront plus pêcher.

Voici en quoi consiste le jeu :
Tu vas devoir décider du nombre de poissons que tu veux pêcher dans le lac à chaque période.

GAIN
Chaque poisson pêché te rapportera 1 $.

ÉCRAN DU JEU
Voici à quoi ressemble ton écran :

Période : 1/15

Vous faites partie d'un groupe de 3 joueurs.

Expérience prédagogiques

Mr. Leblanc

Ressource commune

Lisez attentivement la consigne

Astuces

Assurez-vous de bien comprendre la consigne pour pouvoir réussir avec succès l'expérience. La plupart des difficultés rencontrées dans une expérience proviennent d'une mauvaise lecture de la consigne

Consigne

Tu vas faire partie d'un groupe de 3 pêcheurs ayant chacun leur bateau.

Ce groupe se rejoint chaque année au lac Piri-piri pour une journée de pêche. Le lac contient au début 550 poissons.

Chaque jour, chaque pêcheur peut attraper jusqu'à 30 poissons. Cependant, pour assurer le renouvellement du lac, l'ensemble du groupe ne devrait pas pêcher plus de 30 poissons par jour. Chaque poisson pêché en surplus entraine une diminution du nombre de poisson le lendemain.

Exemple

- *Il y a 550 poissons dans le lac*
- *Il est recommandé de ne pas pêcher plus de 30 poissons*
- *Tu décides de pêcher 5 poissons.*
- *A la fin de la journée, 35 poissons ont été pêchés par ton groupe.*

Page de jeu (ancienne puis nouvelle)

vm117.cirano.qc.ca:5889

Période : 4/15

Vous faites partie d'un groupe de 3 joueurs.
Il y a 358 poissons dans le lac.
Il est recommandé que la pêche totale du groupe n'excède pas 30 poissons
En effet, chaque poisson pêché en surplus de cette quantité entraînera une réduction du stock de poissons au début de la période suivante.

Chaque poisson pêché vous rapportera 1 $ à la fin du jeu.
Chaque joueur ne peut pêcher plus de 30 poissons.

Combien de poissons désirez-vous pêcher ?

Niveau d'extraction : 30

Valider

Période	Nombre total de poissons dans le lac au début de la période	Nombre de poissons que tu as pêché	Nombre total de poissons pêchés par le groupe	Excédent de poissons pêchés	Nouveau stock de poissons dans le lac	Gain cumulé depuis le début du jeu
4	358	30	90	60	238	106
3	478	30	90	60	358	76
2	540	30	61	31	478	46
1	550	16	35	5	540	16

Expérience prédagogiques

Mr. Leblanc

Ressource commune

Utiliser le slider pour choisir le nombre de poissons

Rappel de consigne

Vous êtes 3 bateaux à partager le lac.

Il est recommandé que la pêche totale des bateaux ne dépasse pas 30 poissons par jour, sinon le nombre de poisson du lac diminuera le lendemain.

Vos gains augmentent de 1$ par poisson pêché !

Jour 5 / 15

VOS GAINS 126 $

Recommandation: 30 poissons maximum pour les 3 bateaux

1$ / poisson

455 poissons

Combien de poissons voulez-vous pêcher avec votre bateau ? 15

Valider

JOUR	POISSONS DANS LE LAC AU MATIN	POISSONS QUE TU AS PÊCHÉS	PÊCHE DE TOUS LES BATEAUX	EXCÉDENT
4 HIER	480	10	35	5
3	510	10	30	0
2	530	10	20	0
1	550	10	20	0

Page de résultats (ancienne puis nouvelle)

vm117.cirano.qc.ca:5889

Mon gain cumulé : 166

Groupe	**1**
Stock de poissons à la fin du jeu	**-2**
Période à laquelle il n'était plus possible de pêcher	**14**

Période	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Joueurs qui ont pêché 10 poissons ou moins	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Joueurs qui ont resquillé en pêchant plus que 10 poissons	1	3	3	3	3	3	0	0	0	0	0	0	0	0	0

RÉCAPITULATIF DU JEU

Période	Nombre total de poissons dans le lac au début de la période	Nombre de poissons que tu as pêché	Nombre total de poissons pêchés par le groupe	Excédent de poissons pêchés	Nouveau stock de poissons dans le lac	Gain cumulé depuis le début du jeu
1	550	16	35	5	540	16
2	540	30	61	31	478	46
3	478	30	90	60	358	76
4	358	30	90	60	238	106
5	238	30	90	60	118	136
6	118	30	90	60	-2	166

Expériences pédagogiques

Mr. Leblanc

Ressource commune

Félicitations !

Résultat personnels

Résultat du groupe

Récapitulatif

Résultats personnels

Expérience réussie
Félicitations

Profil économique
Coopératif

Gain final
total : 80$

Résultats du groupe

BILAN	
Groupe	1
Stock de poissons à la fin du jeu	150
Période à laquelle il n'était plus possible de pêcher	14

REPARTITION DES JOUEURS PAR JOUR

Jour	1	2	3	4	5	6	7	8	9	10
Joueurs ayant pêcher 10 poissons ou moins	0	3	2	1	0	1	3	2	0	0
Joueurs ayant pêcher plus de 10 poissons !	3	0	1	2	3	2	0	1	3	3