



HAL
open science

Evaluation comparative des interfaces homme-machine : convergence entre plasticité et créativité pour la qualité de l'interaction homme-machine

Hayet Hammami

► **To cite this version:**

Hayet Hammami. Evaluation comparative des interfaces homme-machine : convergence entre plasticité et créativité pour la qualité de l'interaction homme-machine. Interface homme-machine [cs.HC]. Université Grenoble Alpes [2020-..]; Université de Tunis El-Manar. Faculté des Sciences de Tunis (Tunisie), 2021. Français. NNT : 2021GRALM074 . tel-03684167

HAL Id: tel-03684167

<https://theses.hal.science/tel-03684167>

Submitted on 1 Jun 2022

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

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



THÈSE

Présentée en vue de l'obtention du

DIPLÔME DE DOCTORAT

Préparée dans le cadre d'une cotutelle entre l'*Université Grenoble Alpes* et l'*Université de Tunis El Manar*

Spécialité : **Informatique**

Présentée par

Hayet HAMMAMI

Thèse dirigée par **Gaëlle CALVARY** et **Faouzi MOUSSA**

Codirigée par **Meriem RIAHI** et **Fatoumata G. CAMARA**

Préparée au sein du **Laboratoire d'Informatique de Grenoble (LIG)** et du **Laboratoire d'Informatique en Programmation Algorithmique et Heuristique (LIPAH)**

Dans les Écoles Doctorales **Mathématiques, Sciences et Technologies de l'Information, Informatique et Mathématiques, Informatique, Sciences et Technologies de la Matière**

Évaluation comparative des Interfaces Homme-Machine : convergence entre Plasticité et Créativité pour la Qualité de l'Interaction Homme-Machine

Thèse soutenue publiquement le **20 Décembre 2021**
devant le jury composé de :

Madame Leila BEN AYED

Professeur, École Nationale des Sciences de l'Informatique, Université de Manouba, Rapporteur

Madame Gaëlle CALVARY

Professeur, Institut polytechnique de Grenoble, Directeur de thèse

Madame Fatoumata CAMARA

Docteur en sciences, Hochschule für Wirtschaft und Recht Berlin, Co-encadrante (invitée)

Monsieur Patrick GIRARD

Professeur, Université de Poitiers, Examineur

Monsieur Faouzi MOUSSA

Professeur, Université de Tunis El Manar, Codirecteur de thèse

Madame Meriem RIAHI

Maître de conférences, École nationale supérieure d'ingénieurs de Tunis, Co-encadrante (invitée)

Monsieur Jean VANDERDONCKT

Professeur, Université catholique de Louvain, Rapporteur

Madame Christine VERDIER

Professeur, Université Grenoble Alpes, Présidente



Résumé

Cette thèse s'inscrit dans le domaine de l'ingénierie de l'Interaction Homme-Machine (IHM). Elle traite de l'évaluation des IHM. Plus précisément, elle contribue à la classe d'approches d'évaluations comparatives. La littérature leur voit comme atouts d'augmenter la quantité et la qualité des critiques lors d'évaluations expérimentales, permettant ainsi d'améliorer la qualité de l'interaction. Cependant, la production d'IHM alternatives pour l'évaluation est coûteuse en l'absence d'outils de génération automatique de ces alternatives.

L'objectif de la thèse est d'automatiser la production d'IHM alternatives pour accélérer la convergence vers une IHM de qualité. Il s'agit de définir les transformations à appliquer à une IHM initiale pour générer automatiquement la meilleure IHM alternative, c'est-à-dire celle qui, vis-à-vis de l'IHM initiale, permettra de recueillir le plus efficacement possible des retours utilisateurs enclins à améliorer la qualité de l'interaction. L'IHM alternative pourrait devenir l'IHM de référence pour itérativement converger vers l'IHM finale.

Pour répondre à cet objectif, nous nous sommes premièrement interrogés sur la nécessité ou non de transformer l'IHM : un changement de matérialité pourrait-il suffire ? Nous nous sommes placés dans le cadre des IHM graphiques et avons exploré la voie d'alternatives vidéo. En effet, la production d'un prototype vidéo est rapide, peu coûteuse, et permet de raisonner sur l'interaction, à la différence des évaluations comparatives de l'état de l'art qui portent sur l'interface. Nous avons mené une étude expérimentale consistant à présenter successivement aux participants une IHM initiale en version papier puis la même sous la forme d'un film déroulant un scénario d'interaction. Cette étude montre que, dans la version papier, les utilisateurs se concentrent sur les fonctionnalités offertes par le système, c'est-à-dire principalement son utilité, tandis qu'avec un film ils s'expriment davantage sur l'interaction, c'est-à-dire l'utilisabilité. Aussi, l'utilisation combinée des deux supports, papier et film, enrichit les résultats. Cependant les critiques se sont limitées à l'IHM proposée sans suggestions alternatives.

Nous avons alors étudié la transformation de l'IHM et, inspirés par DesignScape, nous nous sommes interrogés sur le caractère mineur ou majeur que ces changements devaient avoir pour donner une évaluation comparative efficace. Sur la base du cadre de référence CAMELEON, nous avons identifié trois niveaux d'abstraction possibles pour ancrer les transformations de l'IHM initiale : les tâches utilisateur, l'interface abstraite et l'interface concrète. Les changements aux deux premiers niveaux d'abstraction (tâche et interface abstraite) sont considérés comme des changements majeurs. Contrairement aux changements au niveau de l'interface concrète qui sont considérés comme mineurs au regard de leur profondeur dans CAMELEON. Nous avons mené une étude expérimentale dans laquelle trois groupes de participants ont évalué comparativement une IHM initiale et une IHM alternative de transformation respectivement au niveau du modèle de la tâche, de l'interface abstraite et de l'interface concrète. L'étude montre que les retours utilisateur ont été les plus nourris en qualité et en quantité sur la version majeure au niveau de l'interface abstraite.

Dans l'optique d'une génération automatique d'alternatives, nous avons alors défini des règles de transformation de l'interface abstraite consistant à modifier la navigation sans changer la présentation des espaces de travail élémentaires. Nous avons implémenté ces règles dans GRA:FITI (Great Reshuffling of Abstractions: From Interface To Interaction), notre démonstrateur de concept d'un possible outil de génération automatique d'IHM alternatives. GRA:FITI génère automatiquement trois IHM à partir d'une IHM initiale qui réduisent la navigation globale, c'est-à-dire, la navigation entre les grands espaces de travail de l'IHM, et/ou la navigation locale, c'est-à-dire la navigation au sein des espaces de travail élémentaires de l'IHM. L'évaluation expérimentale montre que : (1) GRA:FITI aide le concepteur à obtenir des critiques pertinentes sur son IHM initiale par des évaluations comparatives, (2) GRA:FITI peut servir de source d'inspiration en phase de conception.

En synthèse, la thèse apporte deux contributions : (1) elle définit le type de transformation à appliquer à une IHM pour des évaluations comparatives efficaces et en propose un ensemble de règles, (2) elle propose GRA:FITI un outil de génération automatique d'alternatives comme démonstrateur de concept.

Abstract

This thesis contributes to the field of Human-Computer Interaction (HCI) engineering. It deals with the evaluation of User Interfaces (UI). More precisely, it contributes to the class of comparative evaluation approaches. They are known as having the advantage of increasing the quantity and quality of criticism in experimental evaluations, thus improving the quality of interaction. However, the production of alternative UIs to support evaluation is costly in the absence of tools that automatically generate these alternatives.

The objective of this thesis is to automate the production of alternative HCIs to accelerate the convergence towards high quality UIs. The aim is to define the transformations to be applied to an initial UI in order to automatically generate the best alternative UI, i.e. the one which, shown alongside the initial UI, will allow collecting as efficiently as possible user feedback inclined to improve the quality of interaction. The alternative UI could become the reference UI to iteratively converge towards the final UI.

To meet this objective, we first explored whether it was necessary to transform the UI: would a change in materiality be sufficient? We placed our study within graphical UIs and explored the path of video alternatives. Indeed, the production of a video prototype is fast, inexpensive, and allows to reason about interaction, unlike the comparative evaluations of the state of the art which focus on the interface. We conducted an experimental study consisting in presenting successively to the participants an initial UI in paper version and then the same one in the form of a video demonstrating an interaction scenario. This study shows that, in the paper version, users focus on the features offered by the system, i.e. mainly its utility, whereas with a video they express themselves more on the interaction, i.e. the usability. Also, the combined use of the two prototypes, paper and video, enriches the results. However, the criticisms were limited to the explored design space without alternative suggestions. We then looked at the transformation of the UI and, inspired by DesignScape, questioned whether these changes should be minor or major to give an effective comparative evaluation. Based on the CAMELEON framework, we identified three possible levels of abstraction to anchor the transformations of the initial UI: user tasks, abstract User Interface, and concrete UI. Changes at the first two levels of abstraction (task and abstract UI) are considered major changes, in contrast to the changes at the concrete UI level which are considered minor with respect to their depth in CAMELEON. We conducted an experimental study in which three groups of participants comparatively evaluated an initial and an alternative UI presenting transformations respectively at the task model, abstract UI and concrete UI. The study shows that the user feedback was the richest in quality and quantity on the major version of transformation at the level of the abstract UI.

In the perspective of an automatic generation of alternatives, we then defined transformation rules for the abstract UI consisting in modifying the navigation without changing the presentation of the elementary workspaces. We implemented these rules in GRA:FITI (Great Reshuffling of Abstractions: From Interface To Interaction), our proof of concept of a

possible tool for automatic generation of alternative UIs. GRA:FITI automatically generates three UIs from an initial UI that reduce global navigation, i.e., navigation between the main workspaces of the UI, and/or local navigation, i.e., navigation within the elementary workspaces of the UI. The experimental evaluation shows that: (1) GRA:FITI helps the designer to obtain relevant feedback on the initial UI through comparative evaluations, (2) GRA:FITI can serve as a source of inspiration in the exploration of the design process.

In summary, the thesis presents two contributions: (1) it defines the type of transformation to be applied to an UI for efficient comparative evaluations and proposes a set of rules for it, (2) it proposes GRA:FITI a tool for automatic generation of alternatives as a proof of concept.

Dédicace

Je dédie ce travail à :

Mes chers parents, Ayachi et Wahida, sans qui je ne serai pas arrivée là,

Ma sœur Fida qui a été toujours là quand j'ai besoin,

Mon frère Ahmed, qui m'a toujours soutenu,

Mon amie Soumaya, qui a toujours été à mes côtés,

Mon mari Mehdi, qui m'a toujours poussé vers l'avant, et

Tous ceux que j'aime,

Remerciements

Je remercie M. Jean Vanderdonckt et Mme. Hela Ben Ayed pour m'avoir fait l'honneur d'évaluer ce travail de thèse.

Je remercie également M. Patrick Girard et Mme. Christine Verdier pour avoir accepté d'examiner mon travail.

Je remercie Gaëlle Calvary et Faouzi Moussa pour leur encadrement, toute leur aide et leurs précieux conseils.

Je remercie Meriem Riahi sans qui cette thèse aurait été une tâche plus difficile.

Je remercie Fatoumata Camara pour son aide.

Je tiens aussi à remercier les membres des équipes IIHM du LIG et IHM du LIPAH pour leur bon humeur et qualités humaines.

Je remercie également ma famille pour leur soutien inconditionnel, leur amour et encouragement.

Je remercie mes amis pour m'avoir aidé à garder le moral pendant ces années de thèse.

Je tiens à remercier Mehdi, mon mari qui m'a soutenu tout au long de cette thèse et qui ne m'a jamais laissé abandonner.

Finalement, je tiens à remercier tous ceux qui m'ont aidé et soutenu tout au long de ce travail.

Table des matières

Introduction générale	1
1	1
2	2
3	2
4	3
I État de l’art	4
1 Évaluation comparative des IHM	5
1	6
2	6
2.1	6
2.2	7
3	9
3.1	10
3.2	13
3.2.1	13
3.3	17
3.4	19
4	20
2 Exploration des alternatives	23
1	24
2	24
2.1	25

2.1.1	Le prototypage	25
2.1.2	Différents types et formes des prototypes	26
2.2	Explorer des alternatives à l'aide des galeries web	28
2.2.1	Outil Adaptive Ideas Web Design	28
2.2.2	Outil d.tour	30
3	Générer automatiquement des alternatives à partir d'une IHM	30
3.1	DesignScape	31
3.2	Sketchplore	32
4	Produire des IHM par transformation de modèles	34
4.1	Cadre de référence CAMELEON pour la génération des IHM	35
4.2	Magellan	37
5	Synthèse et conclusion	39
	Discussion sur l'état de l'art	40

II Contributions **42**

3	Quelle alternative pour les évaluations comparatives ?	43
1	Introduction	44
2	Démarche de travail	44
2.1	Changement de support	44
2.2	Changement d'IHM	44
3	Étude #1 : changement de support	45
3.1	Cas d'étude	45
3.2	Élaboration des prototypes	45
3.3	Participants	46
3.4	Protocole expérimental	46
3.5	Collecte des données	48
3.6	Résultats	49
3.6.1	Quantité des retours utilisateur	49
3.6.2	Qualité des retours utilisateur	50
3.6.3	Temps d'évaluation	54
3.7	Synthèse de l'étude #1	54
4	Étude #2 : changement d'IHM	55
4.1	Cas d'étude	56
4.2	Élaboration des prototypes	56
4.3	Participants	58
4.4	Protocole expérimental	59
4.5	Résultats	60

4.5.1	Notes obtenues pour l'IHM de référence	60
4.5.2	Quantité des retours utilisateur	62
4.5.3	Type de retours utilisateur	63
4.6	Synthèse de l'étude #2	63
5	Étude #3 : Transformation au niveau de l'IUA	65
5.1	Élaboration des prototypes	65
5.2	Participants	68
5.3	Protocole expérimental	68
5.4	Résultats	69
5.4.1	Notes obtenues par l'IHM de références	69
5.4.2	Quantité des retours utilisateur	69
5.4.3	Qualité des retours utilisateur	70
6	Conclusion	72
4	GRA:FITI, un environnement de génération automatique d'alternatives	74
1	Introduction	75
2	Principes de génération	75
3	Démonstrateur de concept	78
3.1	Choix techniques	78
3.2	GRA:FITI en action	82
4	Processus d'évaluation de GRA:FITI	83
4.1	Protocole	83
4.1.1	Cas d'étude	83
4.1.2	Participants	86
4.1.3	Protocole expérimental	87
4.1.4	IHM générées	87
4.1.5	Scénario d'interaction	89
4.2	Résultats	89
4.2.1	Score SUS	89
4.2.2	Notes obtenues pour l'utilité des IHM alternatives proposées par GRA:FITI	91
4.2.3	Retours des utilisateurs par rapport à GRA:FITI	92
4.2.4	Retours des utilisateurs par rapport aux différentes IHM	94
5	Discussion	94
	Conclusion générale et perspectives	96
1	Rappel des contributions	96
1.1	Définir la meilleure IHM alternative	97
1.2	Production de l'IHM alternative	97

1.3	L'obtenir par transformation de l'IHM de référence	98
1.4	Démontrer la faisabilité technique	98
1.5	Synthèse	99
2	Perspectives	101
A	Annexe A - Questionnaire utilisé dans l'étude expérimentale #2	111
B	Annexe B - Questionnaire SUS	113
C	Annexe C - Concepts fondamentaux en architecture PAC	115

Table des figures

1.1	Principe d'évaluation comparative des IHM	9
1.2	Prototypes utilisés (Tohidi et al. 2006b)	10
1.3	Catégorisation des retours utilisateurs (Tohidi et al. 2006b)	12
1.4	Crowdsourced Feedback With Imagery (Robb et al. 2015)	13
1.5	Exemples d'interfaces présentées aux utilisateurs (Kang et al. 2018)	15
1.6	IHM de <i>Paragon</i> (Kang et al. 2018)	16
1.7	Test A/B	18
1.8	Évaluation de la performance du design dans (Dow et al. 2011)	20
2.1	Utilisation des exemples dans le processus de conception des IHM (Buxton 2010)	25
2.2	Exemple de création de prototype de basse fidélité avec un outil digital (Mock-flow (<i>Super-easy tool to brainstorm UI ideas</i> Retrieved : December, 2019)) .	27
2.3	À gauche, une vidéo basse-fidélité et, à droite, une vidéo haute-fidélité (Dhillon et al. 2011)	28
2.4	Outil <i>Adaptive Ideas Web Design</i> (Lee et al. 2010)	29
2.5	Détection d'alignement (gauche) et segmentation du design en zones de texte et graphiques (droite) (O'Donovan et al. 2015)	31
2.6	IHM de l'outil DesignScape (O'Donovan et al. 2015)	32
2.7	Outil Sketchplore - Utilisation (Todi et al. 2016)	33
2.8	Idées de conception dans Microsoft Office PowerPoint	34
2.9	Les quatre niveaux d'abstraction du cadre de référence CAMELEON (Calvary et al. 2003)	35
2.10	Exemple de modèle de tâche	36
2.11	Exemple d'espaces de travail sur le modèle de tâche	37
2.12	Exemples d'IHM générées par Magellan (Masson et al. 2010)	39

3.1	Prototype statique du site e-commerce	47
3.2	Prototype dynamique du site e-commerce	47
3.3	Situation d'évaluation 1	48
3.4	Un participant qui saisit ses retours	49
3.5	Nombre de retours utilisateur par rapport à leur type	51
3.6	Nombre de retours utilisateur par rapport à leur qualité	52
3.7	Résumé des résultats de la première étude - Situation d'évaluation 1	53
3.8	Résumé des résultats de la première étude - Situation d'évaluation 2	53
3.9	Taux de retours utilisateur par catégorie pour chaque situation d'évaluation .	54
3.10	Élaboration des prototypes	56
3.11	Modèle de tâche 1	57
3.12	IU Abstraite 1	57
3.13	IU Concrète 1	58
3.14	IU Concrète 2	59
3.15	IU Abstraite 2	60
3.16	IU Concrète 3	61
3.17	Modèle de tâche 2	61
3.18	Score accordée à IU Concrète 1 pour chaque situation d'évaluation	62
3.19	Nombre de retours utilisateur par rapport à leurs types pour chaque situation d'évaluation	64
3.20	Impact du choix de l'IHM alternative sur les retours utilisateur dans l'étude #2.	64
3.21	Élaboration des IHM - Étude #3	65
3.22	Interface Utilisateur Abstraite 0	66
3.23	IUA 1 (à gauche) et IUA 2 (à droite)	66
3.24	IU Concrète 0	67
3.25	IU Concrète 1-B	67
3.26	IU Concrète 2-B	68
3.27	Score de IUC0 pour chaque situation d'évaluation - Étude #3.	69
3.28	Nombre de retour utilisateur en fonction de leur type pour chaque situation d'évaluation - Étude #3.	70
3.29	Nombre de retours utilisateur pour chaque situation d'évaluation en fonction de leur qualité - Étude #3.	71
4.1	Produire des IHM par transformation de modèles	76
4.2	Exemple de tâche articulatoire	77
4.3	Exemple d'agent articulatoire (agent C)	78
4.4	Agent ciment dans (Calvary et al. 1997)	78
4.5	GRA:FITI : Principe	79
4.6	Vue globale de notre boîte à outil	80

4.7	Exemple d'une arborescence d'une IHM implémentée suivant l'architecture PAC	81
4.8	Exemple de deux alternatives d'une IHM présentant des systèmes de navigation différents	82
4.9	Vue globale	83
4.10	Exemple de trois alternatives (à droite) d'une IHM de référence (à gauche) présentant des systèmes de navigation différents	84
4.11	Page d'accueil de GRA:FITI : téléchargement de l'IHM de référence	85
4.12	GRA:FITI : Comparaison des IHM	85
4.13	IHM source	86
4.14	IHM alternative 1 - Page d'accueil	88
4.15	IHM alternative 2 - Page parcours	89
4.16	Création d'un parcours - IHM de référence	90
4.17	Création d'un parcours - IHM alternative N°2	90
4.18	Score SUS par participant	91
4.19	Score SUS pour la première question	91
4.20	Notes accordées pour l'utilité des IHM alternatives proposées par GRA:FITI	92
4.21	Processus de développement d'IHM (Gulliksen et al. 2003).	101
C.1	Hiérarchie d'agents PAC (Coutaz 1987)	116

Liste des tableaux

1.1	Exemples de méthodes classiques en évaluation des IHM	8
1.2	Description des critères utilisés dans (Tohidi et al. 2006b) pour l'évaluation des IHM.	11
1.3	Tableau comparatif.	21
2.1	Classes de variations dans les trois modèles : tâches, IUA et IUC	38
3.1	Nombre de retours utilisateur pour chaque situation d'évaluation- Étude #1.	50
3.2	Nombre de retours utilisateur pour chaque situation d'évaluation - Étude #2	62
3.3	Nombre total de retours utilisateur pour chaque situation d'évaluation - Étude #3.	70
4.1	Nombre des participants par profil	86
4.2	Tableau comparatif.	100

Introduction générale

1 Domaine de recherche

Cette thèse s’inscrit dans le domaine de l’ingénierie de l’Interaction Homme-Machine (IHM). Ce domaine couvre l’étude des concepts, modèles et outils nécessaires à la conception et à l’évaluation de l’interaction au service des personnes. Le processus de conception d’une IHM est une tâche difficile. La sous-spécification des exigences est une difficulté majeure accroissant d’autant l’espace de conception à explorer. Aussi, l’évaluation des IHM est primordiale.

Cette thèse traite de l’évaluation des IHM. Elle contribue à la classe d’approches d’évaluations dites comparatives consistant à comparer différents exemples d’alternatives d’IHM. La littérature montre que travailler en s’appuyant sur des exemples présente plusieurs avantages lors du processus de conception des IHM. En effet, tel que le montre B. Buxton, l’exploration d’alternatives améliore à la fois le processus d’apprentissage et le résultat (IHM finale) (Buxton 2010). Ainsi, les travaux de la littérature recommandent aux concepteurs de rechercher des exemples d’alternatives lors de la conception d’une IHM. La littérature indique que les alternatives explorées favorisent leur créativité et qu’utiliser simultanément une référence et des alternatives peut être bénéfique pour améliorer la qualité de l’IHM. En psychologie cognitive, il est indiqué que l’analogie imprègne la pensée et que *«l’analogie est présentée à la fois comme outil et comme méthode utilisée pour produire de la connaissance et du savoir»* (Zouhairi 2010).

Par ailleurs, il a été montré que l’approche d’évaluation comparative améliore les résultats d’évaluations expérimentales des IHM en augmentant la quantité et la qualité des critiques fournies par les utilisateurs (Tohidi et al. 2006b ; Dow et al. 2011 ; Kang et al. 2018).

Pourtant, ce champ de recherche est aujourd’hui sous-exploré, n’expliquant notamment pas comment produire et choisir les alternatives.

2 Problématique et questions de recherche

Différents travaux dans la littérature ont étudié les vertus de l'évaluation comparative des IHM (Tohidi et al. 2006b ; Dow et al. 2011 ; Kang et al. 2018). En effet, il a été trouvé que la présentation de plusieurs alternatives d'une IHM aux testeurs permet de mieux critiquer ces dernières, augmente les chances de trouver la bonne IHM et de converger vers une IHM satisfaisante.

La littérature documente la performance des évaluations comparatives mais n'en développe pas les principes méthodologiques : ni la méthode, ni les critères de choix de conception des IHM alternatives ne sont expliqués.

Cette thèse vise à automatiser la production d'IHM alternatives pour optimiser l'évaluation comparative et ainsi accélérer la convergence vers une IHM de qualité. De ce fait, notre question de recherche est double :

- En termes de produit, quelle est l'IHM à présenter pour optimiser les résultats des évaluations comparatives ?
- En termes de processus, comment générer cette IHM ?

3 Démarche

Cette thèse se situe à l'intersection de trois domaines de l'ingénierie des IHM : l'évaluation comparative, l'exploration des alternatives et la génération des IHM. Nous avons commencé par étudier l'état de l'art couvrant ces trois domaines. Cet état de l'art s'organise en deux pôles :

- Une première étude sur l'évaluation comparative des IHM. Cette étude nous a permis de mettre en évidence le bénéfice de l'utilisation des alternatives pour les évaluations des IHM.
- Une deuxième étude sur l'exploration et la génération des alternatives dans le processus de conception des IHM. Cette étude nous a fourni des lignes directrices permettant de générer des exemples d'alternatives d'une IHM.

Dans un deuxième temps, nous présentons nos contributions. La première contribution répond à notre première question de recherche : « *quelle est l'IHM à présenter pour optimiser les résultats des évaluations comparatives ?* ». Nous définissons donc le type de transformation à appliquer à une IHM initiale pour générer la meilleure IHM alternative. Sur la base de ces résultats nous répondons à notre deuxième question de recherche « *comment générer l'IHM alternative ?* » et nous définissons un ensemble de règles de transformation à appliquer pour la génération automatique des IHM. Enfin, dans l'optique d'automatiser cette transformation,

nous présentons GRA:FITI (*Great Reshuffling of Abstractions : From Interface To Interaction*) un démonstrateur démontrant la faisabilité technique de la génération automatique des alternatives.

4 Organisation du mémoire

Le plan du rapport s'articule en deux parties : l'état de l'art et les contributions.

La première partie est composée de deux chapitres qui synthétisent l'état de l'art. Le premier chapitre (chapitre 1) présente les notions théoriques et fondamentales de l'évaluation comparative des IHM et expose les travaux présentés de la littérature. À la fin de ce chapitre, nous présentons une discussion critique soulevant les questions fondamentales de nos travaux de recherche. Le second chapitre (chapitre 2) se focalise sur l'exploration et la génération des alternatives dans le processus de conception des IHM. Cette première partie se conclut par une discussion sur les différents travaux présentés.

La deuxième partie est consacrée aux contributions. Elle est organisée en deux chapitres. Le premier (chapitre 3) présente la démarche de travail et expose les différentes études et expérimentations menées pour la définition des transformations à appliquer à une IHM initiale pour générer la meilleure IHM alternative. Le deuxième (chapitre 4) détaille les règles de transformations élaborées et présente l'outil GRA:FITI développé pour la génération automatique des alternatives. À la fin de ce chapitre nous présentons une implémentation illustrée sur un cas d'étude dans le but de démontrer la faisabilité technique de notre proposition.

Enfin ces contributions ouvrent la voie à de nombreuses perspectives présentées à la fin de ce mémoire.

Première partie
État de l'art

1

Évaluation comparative des IHM

1 Introduction

Ce premier chapitre expose et analyse l'approche d'évaluation comparative en IHM. Il s'organise en trois sections. Tout d'abord, il présente les notions fondamentales en évaluation des IHM et plus spécialement en évaluation comparative. Ensuite, il présente un panorama de travaux inspirants sur le sujet. Enfin, il propose une discussion critique soulevant les questions fondamentales de notre recherche.

2 Notions fondamentales en évaluation des IHM

Le processus de développement des IHM se structure en trois phases principales : la phase d'analyse et de spécification des besoins, la phase d'implémentation et la phase d'évaluation. Le cadre général du cycle de vie des IHM met l'accent sur la phase d'évaluation (Millot 1987 ; Hartson et Hix 1989). Elle est considérée comme une discipline à part entière ayant ses méthodes et ses outils. Nous présentons dans cette section le principe d'évaluation des IHM ainsi qu'un panorama des techniques d'évaluation existantes. Nous nous focalisons ensuite sur l'évaluation comparative.

2.1 Principe d'évaluation des IHM

L'évaluation des IHM consiste à tester et à valider un système interactif en se basant sur différents critères pour améliorer et affiner sa qualité et assurer son bon fonctionnement. Ces critères tiennent compte des performances globales du système et sont exprimés en termes d'écart entre la production réelle et les objectifs définis (Millot 1987 ; Bastien et Scapin 1992 ; Vanderdonck 1994). Cet écart peut être mesuré en termes d'utilisabilité du système, de son utilité et/ou accessibilité. L'utilisabilité est considérée comme un attribut de qualité. Elle évalue la facilité d'utilisation des IHM (Nielsen et Faber 1996). Elle pose des requis sur la facilité d'apprentissage, l'efficacité, l'efficience et la satisfaction, dans un contexte d'utilisation spécifique. L'utilité est relative aux services utilisateur attendus. Elle représente selon C. MacDonald et M. Atwood "le degré selon lequel les fonctions d'un système permettent aux utilisateurs de compléter un ensemble de tâches et de répondre aux objectifs spécifiques dans un contexte d'utilisation particulier" (MacDonald et Atwood 2014). L'accessibilité d'une IHM définit sa capacité à être utilisée par tout type d'utilisateur, surmontant les diversités culturelles, techniques ou sociales (Rutter et al. 2007).

Au cours des vingt dernières années, plusieurs travaux portant sur les méthodes, outils et techniques d'évaluation des IHM ont été publiés. Ces travaux proposent des techniques d'aide à l'évaluation qui diffèrent sur un certain nombre de points tels que les facteurs inspectés (utilisabilité, utilité, accessibilité) ou la nature des données pour l'évaluation (quantitatives ou qualitatives). Ainsi, différentes classifications de ces techniques ont été définies permettant de

les répertoire. Par exemple Ivory et Hearst classifient ces techniques autour de cinq catégories (Ivory et Hearst 2001) : l’inspection de l’utilisabilité, les tests utilisateur, les enquêtes, les techniques analytiques et prédictives et la simulation.

Par ailleurs, différents types d’évaluation d’IHM ont été définis. Nous présentons dans ce qui suit quatre types d’évaluation parmi les plus cités dans la littérature :

- **Évaluation subjective** : ce type d’évaluation a pour but d’identifier des problèmes d’utilisabilité d’un point de vue subjectif. Elle est réalisée par des utilisateurs experts ou non experts. Les utilisateurs expriment leurs opinions sur la qualité de l’IHM étudiée (Nishiuchi et Takahashi 2015).
- **Évaluation objective** : l’évaluation objective vise à obtenir des résultats indépendants des opinions personnelles subjectives des utilisateurs (Albert et Tullis 2013).
- **Évaluation formative** : ce type d’évaluation est réalisé en cours de conception. Son but est d’identifier les problèmes d’utilisabilité avant la finalisation et la livraison du produit (Albert et Tullis 2013).
- **Évaluation sommative** : ce type d’évaluation se réalise à la fin de la conception. Son but est d’évaluer la qualité du produit final. Elle détermine si ce dernier répond ou non aux objectifs fixés (Albert et Tullis 2013). Selon Harton et al. elle peut être utilisée pour comparer des alternatives de conception (Hartson et al. 2001).

Différentes méthodes et techniques ont été proposées pour guider et assister les équipes projets pendant le processus de développement (Assila 2016 ; Balbo 1994 ; Kolski et GRISLIN 1995 ; Bastien et Scapin 1992 ; Nielsen 1994). Chaque méthode a ses propres exigences et peut être appropriée pour une ou plusieurs étapes du cycle de vie de conception des IHM. Ces méthodes permettent de recueillir un ensemble de données quantitatives et/ou qualitatives. Ces données servent ensuite à identifier les problèmes à traiter. Leur gravité peut varier de critique à cosmétique, définissant ainsi la priorité à porter aux problèmes dans la correction du système (Nielsen 1994). Le tableau 1.1 présente un panorama non-exhaustif des méthodes les plus utilisées et référencées dans la littérature pour l’évaluation des IHM.

2.2 Principe d’évaluation comparative des IHM

Les publications en évaluation d’IHM sont nombreuses (Kurosu 2015), mais peu d’entre elles traitent de l’évaluation comparative. Cette approche repose sur la présentation simultanée d’une IHM de référence et d’une ou plusieurs alternatives de cette IHM au cours d’une session d’évaluation. Ces alternatives représentent des variations de conception de l’IHM de référence à plusieurs niveaux d’abstraction tels que les niveaux lexicaux, syntaxiques, sémantiques ou pragmatiques (Figure 1.1). Elles sont conçues par la même personne ou non.

Au fil des années, l’évaluation comparative s’est développée (Dicks 2002 ; Tohidi et al. 2006b ; Dow et al. 2010 ; Dow et al. 2011 ; Mora 2015 ; Kang et al. 2018). Par exemple, Tohidi

Méthode d'évaluation	Principe d'évaluation
Évaluation heuristique	C'est une méthode d'inspection basée sur un ensemble d'heuristiques qui permettent d'identifier des problèmes potentiels d'utilisabilité (Nielsen et Landauer 1993). Cette méthode d'évaluation est rapide et facile à mettre en place par les concepteurs. Le principe est simple : différentes personnes (expertes en utilisabilité) évaluent une IHM en s'appuyant sur une liste donnée d'heuristiques.
Guides ergonomiques	L'objectif de cette méthode est d'évaluer la qualité ergonomique d'une IHM et de détecter les anomalies en s'appuyant sur des règles ergonomiques. Le principe de l'évaluation consiste à inspecter et à juger la conformité des éléments de l'IHM par rapport à un ensemble de recommandations ergonomiques. Différents guides ont été proposés dans la littérature tels que les critères ergonomiques de (Bastien et Scapin 1992) ou (Shneiderman et al. 2016).
Pensée à voix haute (Think-Aloud Protocol)	Cette méthode est utilisée pour collecter des données pendant les tests d'utilisabilité. Il est demandé aux utilisateurs de verbaliser leurs pensées et raisonnements à voix haute lors de l'interaction avec le système. Ces données sont enregistrées afin d'être analysées et étudiées ultérieurement. Cette méthode d'évaluation facilite la compréhension de ce que l'utilisateur pense lors de l'utilisation du système en question.
Mesure des performances	Cette méthode d'évaluation permet aux concepteurs de mesurer la performance des utilisateurs lors de la réalisation de tâches. Le principe consiste à définir les mesures de performances souhaitées (par exemple le temps de réalisation d'une tâche) et de les capturer en cours de tests utilisateurs. Ces mesures peuvent être réalisées manuellement par les concepteurs ou automatiquement avec des outils dédiés. Les résultats de cette évaluation permettent de juger si les objectifs d'utilisabilité sont atteints.
Questionnaires	La méthode des questionnaires est couramment utilisée dans les évaluations des IHM. Elle consiste à questionner les utilisateurs sur leur perception de différents critères.

TABLE 1.1 – Exemples de méthodes classiques en évaluation des IHM

a démontré qu'exposer plusieurs alternatives d'IHM à un utilisateur augmente le nombre de critiques et de suggestions d'amélioration (Tohidi et al. 2006b). De son côté, S. Dicks montre que présenter différentes alternatives d'IHM à un utilisateur permet une évaluation

plus juste. Selon lui, la pression sur l'utilisateur est moindre car il se sent déchargé de devoir juger une IHM particulière (Dicks 2002). Wiklund avance d'autres arguments. En particulier, il défend l'idée que les utilisateurs ne souhaitent pas être assimilés à des personnes négatives ou offensantes vis-à-vis des concepteurs ou à des personnes ayant une dextérité insuffisante vis-à-vis de la technologie (Wiklund et al. 2016). Enfin, Mora met l'accent sur l'aide qu'offrent les alternatives pour fournir un vocabulaire permettant de discuter l'espace de conception (Mora 2015).

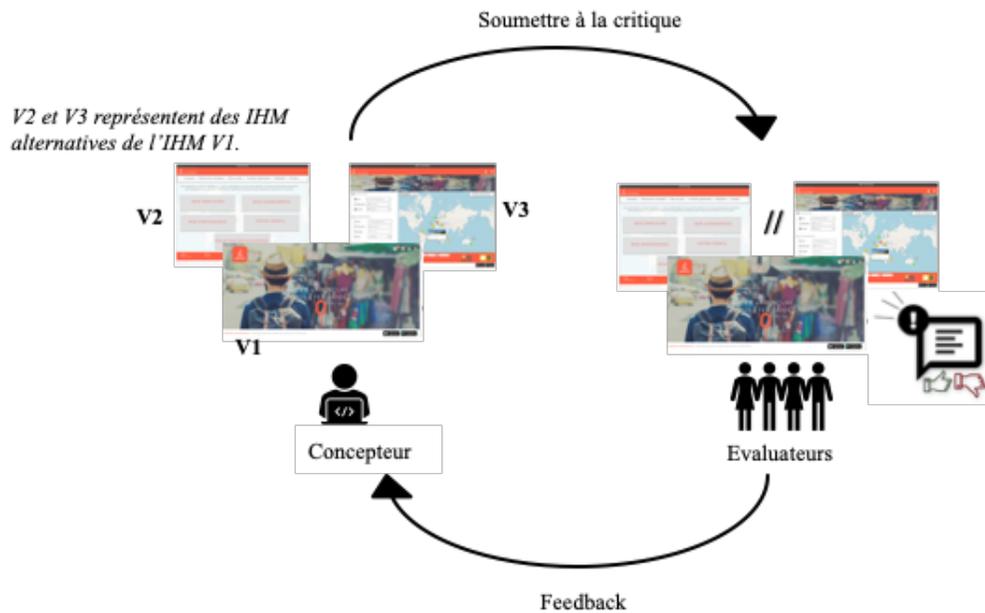


FIGURE 1.1 – Principe d'évaluation comparative des IHM

Dans la section suivante, nous détaillons les travaux clés traitant de l'évaluation comparative et démontrant les bénéfices de cette approche.

3 Revue critique de l'état de l'art en évaluation comparative des IHM

Peu de travaux traitent de l'évaluation comparative des IHM. Les résultats de Tohidi et al. (Tohidi et al. 2006b) constituent l'un des travaux fondateurs de notre recherche. Nous commençons par présenter ce travail dans la première partie de cette section. Ensuite, nous exposons d'autres travaux de la littérature qui abordent des approches utiles au sujet ici traité.

3.1 Comparer des IHM plutôt que n'en évaluer qu'une !

Tohidi et al. étudient le bénéfice de l'évaluation simultanée de trois alternatives d'IHM comparativement à une seule (Tohidi et al. 2006b). L'IHM étudiée est graphique, monopage, relative à un système de contrôle de climatisation domestique. Les fonctionnalités offertes concernent la création, le changement ou l'activation des paramètres ou des programmes du système de climatisation. La couverture fonctionnelle des trois IHM est la même. Seule l'IHM change dans les couleurs et widgets utilisés : liste déroulante, bouton, barre coulissante, horloge analogique ou numérique (Figure 1.2).

Les chercheurs ont conçu les trois IHM en basse-fidélité et support papier (Figure 1.2), en utilisant scotch, marqueurs et feutres dans le but de pouvoir simuler un scénario pour des tests utilisateur. Le principal avantage de ce prototypage est son faible coût.

Quarante-huit participants, répartis en deux groupes, ont évalué soit une IHM seule (groupe 1), soit trois IHM simultanément (groupe 2). Les participants du premier groupe ont été répartis en trois sous-groupes, traitant chacun une des trois IHM.

Les participants ont réalisé différentes tâches telles que la réinitialisation de l'heure et la date du thermostat, ou la modification de la température affichée. Les tâches ont été réalisées par une tierce personne dont le rôle était de simuler le comportement du système. Cette approche est économique, permettant de réduire le coût des évaluations.

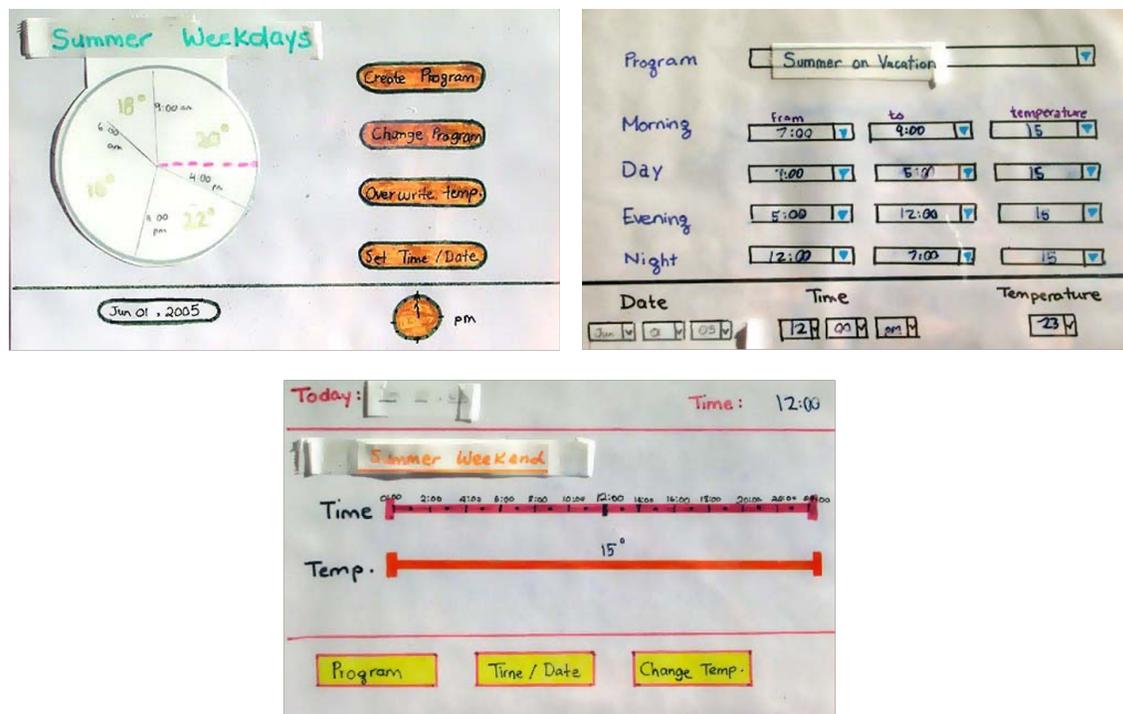


FIGURE 1.2 – Prototypes utilisés (Tohidi et al. 2006b)

Chaque prototype était évalué sur une échelle de Likert de 10 points selon douze critères : quatre critères sur la facilité d'utilisation pour chacune des tâches réalisées, et sept critères concernant l'esthétique, la lisibilité, la cohérence, la correspondance avec les métaphores du monde réel, la navigation, l'exploitabilité et l'apprentissage. Nous présentons une définition de ces critères dans le tableau 1.2.

Critère	Définition
Esthétique	Ce critère concerne le choix des visuels de l'IHM. Par exemple, le choix des couleurs, des images, etc.
Lisibilité	Ce critère fait partie des sous-critères ergonomiques proposés par (Bastien et Scapin 1992). Il concerne les caractéristiques lexicales de présentation des informations sur l'écran pouvant entraver ou faciliter la lecture de ces dernières.
Cohérence	Ce critère fait partie des sous-critères ergonomiques proposés par (Bastien et Scapin 1992). Il concerne la façon selon laquelle les choix de conception de l'interface sont conservés pour des contextes identiques et sont différents pour des contextes différents.
Correspondance avec les métaphores du monde réel	Ce critère mesure la correspondance du contenu de l'interface avec le monde réel.
Navigation	Ce critère mesure la facilité de navigation entre les différents espaces de travail de l'IHM.
Exploitabilité	Ce critère mesure le degré d'exploitation possible du système étudié.
Apprentissage	Ce critère mesure la facilité d'apprentissage du système à évaluer.

TABLE 1.2 – Description des critères utilisés dans (Tohidi et al. 2006b) pour l'évaluation des IHM.

Ensuite, il était demandé aux participants d'accorder une note globale au(x) prototype(s) évalué(s). Une moyenne a été calculée à partir des notes de chaque participant pour les onze premières questions, excluant la note globale donnée par le participant.

Enfin, les participants étaient conviés à un entretien semi-structuré afin de commenter ce qu'ils apprécient ou non dans les prototypes présentés et suggérer des améliorations.

Les retours utilisateurs collectés ont été classés en commentaires et suggestions. Les commentaires représentent des opinions personnelles sur les IHM et sont classés selon leur polarité positive (ce que les participants apprécient) et négative (ce que les utilisateurs n'apprécient pas). Les suggestions sont nuancées en superficielles versus substantielles. Les substantielles

sont à leur tour distinguées selon qu'elles sont nouvelles ou empruntées. Les suggestions dites nouvelles représentent des idées d'amélioration originales, c'est-à-dire inexistantes dans les trois IHM présentées, comme, par exemple, ajouter deux options d'unité Celsius ou Fahrenheit. Les suggestions dites empruntées sont des idées présentes dans l'une des trois IHM présentées, par exemple, afficher la température actuelle sur l'écran principal, idée utilisée dans l'une des trois IHM (Figure 1.3).

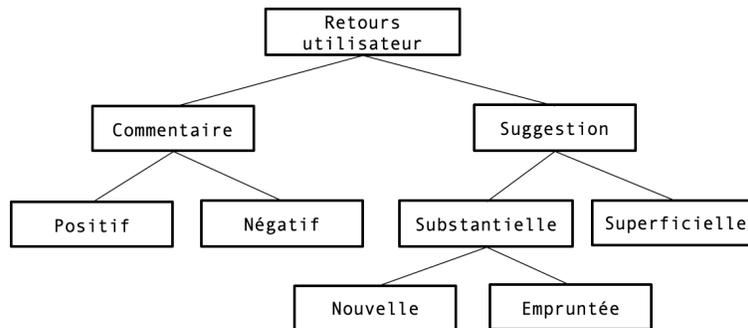


FIGURE 1.3 – Catégorisation des retours utilisateurs (Tohidi et al. 2006b)

Le score donné pour chaque prototype, le nombre de commentaires positifs et négatifs ainsi que le nombre de suggestions ont été analysés pour chaque condition d'évaluation. Les résultats de l'étude montrent que, lorsque les participants évaluent simultanément des alternatives d'IHM, ils fournissent un plus grand nombre de retours que quand ils évaluent une seule IHM. Cette situation d'évaluation favorise donc l'esprit critique chez les participants. Ces retours se constituent essentiellement de commentaires négatifs et de suggestions. En outre, il est montré que les participants notent plus généreusement les IHM quand elles sont évaluées seules et que les commentaires positifs sont alors également plus nombreux et qu'inversement cela diminue lorsque plusieurs alternatives sont présentées pour l'évaluation. Les participants se focalisent alors sur les problèmes de l'IHM et multiplient les commentaires négatifs et suggestions constructives. Une autre différence constatée entre les groupes est le rejet d'IHM lors d'évaluations comparatives, contrairement aux évaluations mono-IHM où aucun rejet n'a été exprimé.

Cette étude porte les vertus de l'évaluation comparative des IHM. En revanche, le choix des caractéristiques des différentes IHM n'est pas discuté. Les alternatives ont été conçues à la main par la même équipe de concepteurs, mais la méthode de génération et le choix des variantes ne sont pas décrits. Ces constats nous ont incités à nous demander dans quelle mesure le choix des alternatives affecte les retours utilisateurs.

3.2 Utiliser des galeries d'exemples pour évaluer !

Dans les évaluations d'IHM basées sur du *Crowdsourcing*¹ (Zhao et Zhu 2014; Luther et al. 2015), il a été montré que se baser sur des exemples visuels dans l'évaluation améliore les retours utilisateurs, conduisant à une réflexion plus approfondie (Robb et al. 2015; Robb et al. 2017). Cette méthode consiste en la sélection intuitive et rapide dans une banque d'exemples visuels présentés aux utilisateurs pour exprimer une idée (figure 1.4).

Cependant, chercher et trouver des exemples pertinents par rapport à un domaine donné n'est pas facile. Dans ce contexte, Kang et al. présentent *Paragon*, un système permettant de parcourir efficacement un ensemble d'exemples à l'aide de méta-données et de fournir des retours en ligne (Kang et al. 2018).

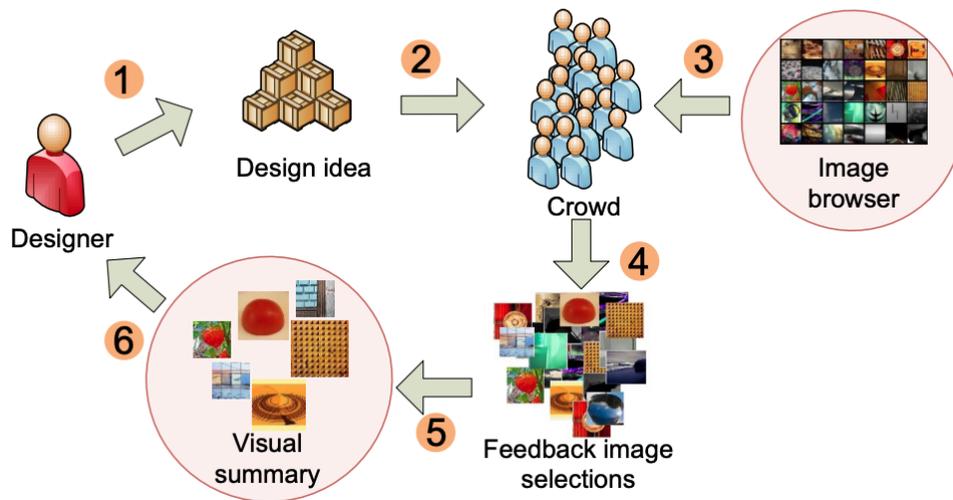


FIGURE 1.4 – Crowdsourced Feedback With Imagery (Robb et al. 2015)

3.2.1 Travaux de Kang et al. (Kang et al. 2018)

Kang et al. étudient comment les évaluateurs exploitent des exemples d'IHM pour fournir des retours sur une IHM de référence et comment ces exemples impactent la qualité de leurs retours (Kang et al. 2018). Dans ce cadre, ils présentent deux études expérimentales au cours desquelles des participants évaluent différentes IHM en s'aidant de galeries d'exemples. Dans chaque étude, il est demandé aux participants d'évaluer une IHM sous trois conditions : une première condition où l'interface est présentée seule et deux autres conditions où l'interface est présentée avec une galerie d'exemples.

1. Recueillir des retours utilisateurs d'un grand nombre de personnes via Internet ou une communauté en ligne.

Dans la première étude, l'évaluation porte sur une affiche de conférence, dans trois conditions :

1. **Contrôle** : l'affiche est présentée seule.
2. **Contenu similaire**: l'affiche est évaluée avec l'aide d'une galerie de 24 exemples similaires en termes de contenu.
3. **Organisée** : l'affiche est évaluée en s'appuyant sur 24 exemples collectés sur le web via une recherche d'affiches présentant des contenus très variés (figure 1.5).

Les participants à l'évaluation ont été recrutés via Amazon Mechanical Turk (MTurk)¹. Il s'agit de trente-deux personnes dont des designers professionnels et des étudiants avec une expérience de design graphique. Les participants ont accédé via le web au système d'évaluation Paragon (Figure 1.6) qui contient l'interface à évaluer ainsi qu'une galerie navigable d'exemples, des champs de texte vides pour soumettre des retours, et enfin huit rubriques cliquables créées par des experts :

- Besoin de considérer le public.
- Fournir une meilleure concentration visuelle.
- Beaucoup d'informations.
- Créer une mise en page plus sensible.
- Besoin d'utiliser des visuels et du texte complémentaires.
- Besoin d'une hiérarchie visuelle claire.
- Choisir judicieusement la police et couleurs des caractères.
- Autre.

À la fin de chaque évaluation, les participants devaient remplir un questionnaire portant sur leurs compréhension des rubriques, efficacité pour exprimer leurs intentions dans leurs retours et facilité à restituer leurs retours. Ils répondaient sur une échelle Likert de 7 points (allant de "pas du tout d'accord" à "tout à fait d'accord") sur les deux questions suivantes : "*j'aimerais choisir un exemple quand je donne mon retour*" et "*j'aimerais choisir des exemples similaires plutôt que des exemples différents lorsque je donne mon retour.*" Enfin, trois questions ouvertes étaient posées sur les préférences des participants par rapport à la qualité des exemples fournis, la manière selon laquelle ils avaient parcouru les exemples, et les fonctionnalités qu'ils souhaitaient avoir à leur disposition en parcourant les exemples.

Les retours utilisateurs recueillis ont été analysés et notés par un designer professionnel. Ils indiquent que les participants préfèrent s'appuyer sur des exemples pour fournir leurs retours. L'utilisation des exemples les a aidés à utiliser les rubriques créées par les experts plus efficacement. Par ailleurs, les participants indiquent que s'appuyer sur des exemples les

1. MTurk est une plateforme web de crowdsourcing qui vise à faire effectuer par des humains, contre rémunération, des tâches plus ou moins complexes.

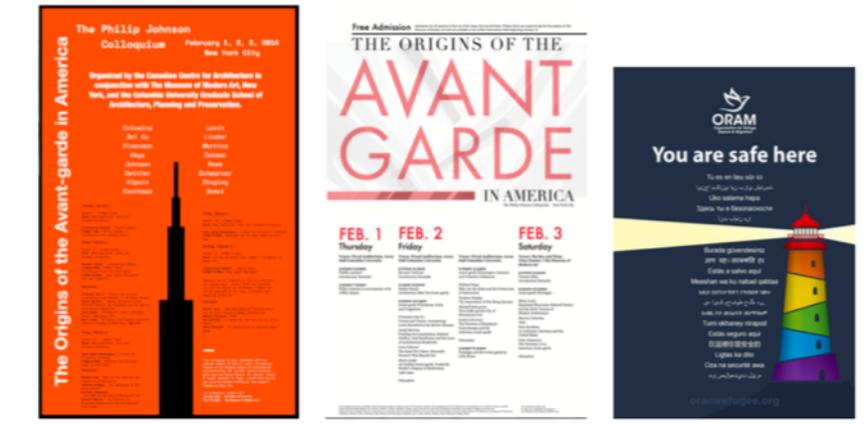


FIGURE 1.5 – Exemples d’interfaces présentées aux utilisateurs (Kang et al. 2018)

aide à mieux développer leurs idées pour la critique. Nous présentons ci-dessous des exemples de retours utilisateurs fournis par quelques participants illustrant ces points :

« *Le fait d’avoir des exemples m’a aidé à réfléchir à d’autres critiques et à citer des exemples contenant les bons éléments que je recherchais ou que j’essayais de décrire* » (Kang et al. 2018).

« *Il était parfois difficile de mettre des mots sur ce que je cherchais ..., les exemples m’ont beaucoup aidé car ils m’ont permis de montrer ce à quoi je pensais ...* » (Kang et al. 2018).

Cependant les chercheurs n’ont pas constaté de différence significative dans la qualité des retours utilisateurs entre les deux conditions contenu similaire et organisé.

Dans la deuxième étude, les participants ont évalué l’IHM dans l’une des trois conditions suivantes :

1. **Contrôle** : l’interface est présentée seule.
2. **Exemples** : l’interface est évaluée avec des exemples librement choisis : les participants sélectionnent eux même les exemples qui leurs semblent les plus pertinents depuis une galerie d’exemples présentés.
3. **Exemples + méta-données** : l’interface est évaluée avec des exemples avec méta-données permettant aux participants, par des filtres, de choisir les exemples dans une galerie. Les filtres utilisés consistent en quatre attributs concernant la quantité de texte, l’espace blanc dans l’interface, les couleurs primaires utilisées et l’alignement du contenu, ainsi que trois attributs présentant trois qualités d’interprétation pour les principes de conception graphique inclus dans la rubrique.

De la même manière que pour la première étude, les participants ont été recrutés via Amazon Mechanical Turk (MTurk). Ils étaient 124 personnes dont des designers professionnels et

des étudiants avec une expérience de design graphique. L'évaluation s'est déroulée à l'aide de l'outil Paragon. Chaque participant était affecté à l'une des trois conditions d'évaluation au hasard. L'interface à évaluer a également été choisie au hasard parmi des posters d'un festival de musique.

La figure 1.6 représente l'interface de Paragon. Elle contient sept zones :

1. L'interface à évaluer et sa description.
2. Les rubriques proposées : les participants passent en revue la rubrique et saisissent leurs retours.
3. Exemple de rubrique sélectionnée.
4. Exemples sélectionnés par les participants dans la condition méta-données.
5. Galerie d'exemples mise à disposition des participants.
6. Le nombre d'exemples disponibles.
7. Les filtres à appliquer lors de la sélection des exemples pour la condition d'évaluation exemple avec méta-données.

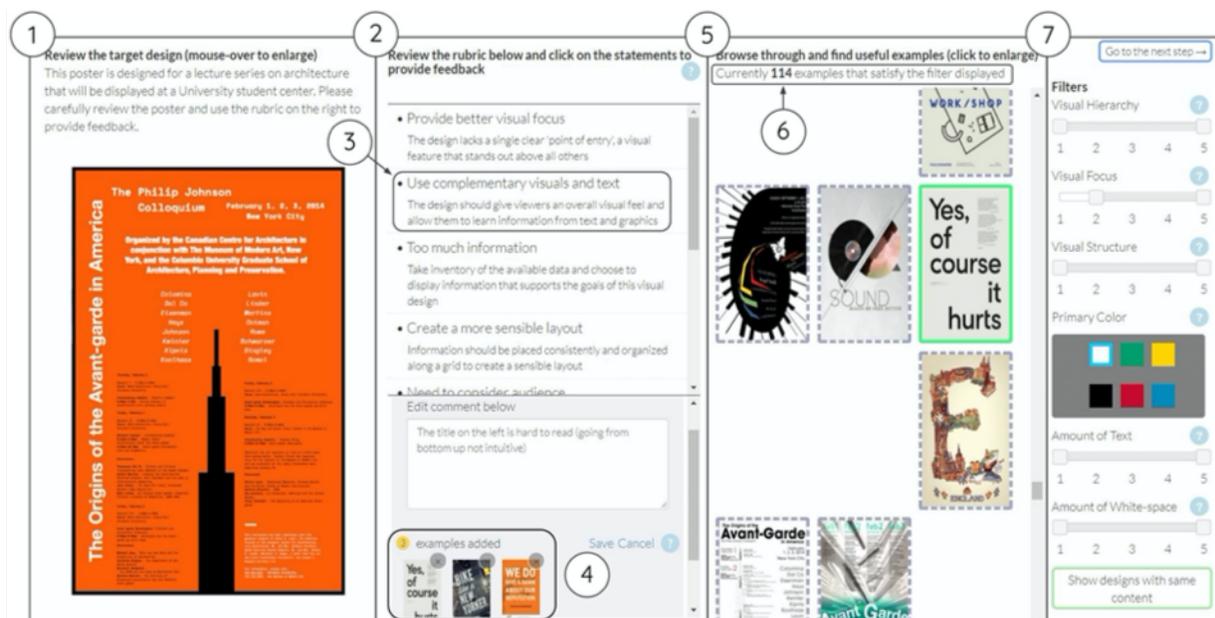


FIGURE 1.6 – IHM de *Paragon* (Kang et al. 2018)

Dans la première condition d'évaluation (contrôle), seuls les blocs 1 et 2 sont affichés aux participants. Dans la deuxième condition (exemples), le septième panel n'est pas affiché. L'ordre de présentation des exemples a été randomisé pour la deuxième et la troisième condition d'évaluation. Enfin, pour la troisième condition, ce sont 287 exemples d'interfaces qui

ont été fournis. Les filtres permettaient de réduire le nombre d'interfaces et de rechercher les exemples les plus adaptés pour l'évaluation.

Les retours utilisateurs recueillis dans cette deuxième étude ont été analysés et jugés par des experts. Leur qualité a été appréciée selon trois dimensions de qualité : spécifique, actionnable et novateur (Sadler 1989).

Dans les deux conditions menées avec des d'exemples (Exemples et Exemples + métadonnées), la qualité des retours utilisateurs a été jugée meilleure : ils sont plus spécifiques, actionnables et innovants. En effet, la présentation des exemples a amélioré la description des problèmes de l'interface ainsi que les étapes à suivre pour l'améliorer.

Concernant l'utilisation des filtres dans la troisième condition d'évaluation, la majorité des participants ont estimé que cette option était utile pour faire correspondre des attributs tels que les couleurs, la quantité de texte, etc. ou simplement pour faciliter la navigation. Toutefois, les auteurs indiquent que certains participants ont exprimé de la frustration soulignant une difficulté à trouver un équilibre entre la charge cognitive supplémentaire et l'utilité d'un tel mécanisme. Il est aussi important de noter que les participants qui ont utilisé des exemples ont passé beaucoup plus de temps pour fournir leurs retours et ont rédigé des commentaires plus longs.

Les deux études présentées dans ce travail confirment les bénéfices de l'utilisation d'exemples pour l'évaluation des IHM (Kang et al. 2018). Les retours fournis par les participants en situation d'évaluation comparative sont plus circonstanciés, innovants et actionnables. Cependant, nous constatons qu'il n'a pas été défini ce qui constitue un "bon" ou "mauvais" exemple à présenter pour maximiser les bénéfices de la comparaison et fournir des retours utilisateurs pertinents.

3.3 Comparer des paires pour choisir !

Le test A/B (*A/B testing* ou *split testing* en anglais) est une méthode d'évaluation permettant de comparer des alternatives d'une certaine IHM afin de déterminer celle qui obtient les meilleurs résultats (Kohavi et Longbotham 2017).

Le test A/B couvre de multiples champs d'application et est très utilisé surtout en marketing de sites e-commerce, tel qu'*Amazon*. Cette méthode permet de mesurer l'impact d'un changement d'une variable sur l'atteinte d'un objectif (clic, validation, remplissage d'un formulaire, etc.). Les IHM utilisées peuvent être statiques sans aucune interaction ou interactives. Le test peut s'effectuer sur des plateformes web et/ou mobiles. « *Un test A/B permet d'évaluer les performances respectives d'une ou plusieurs versions partiellement ou totalement différentes d'une même page web (A et B) en les comparant à la version originale. Le test consiste à créer des versions modifiées de la page originale (A appelée aussi version contrôle) en modifiant autant d'éléments qu'on le souhaite ou que le trafic le permet (contenu, aspect général, présentation...)* » ((*Guide complet de l'A/B testing* 2019)). La figure 1.7 présente un

exemple de deux interfaces A et B qui présentent un même contenu, avec une différence au niveau de sa présentation.

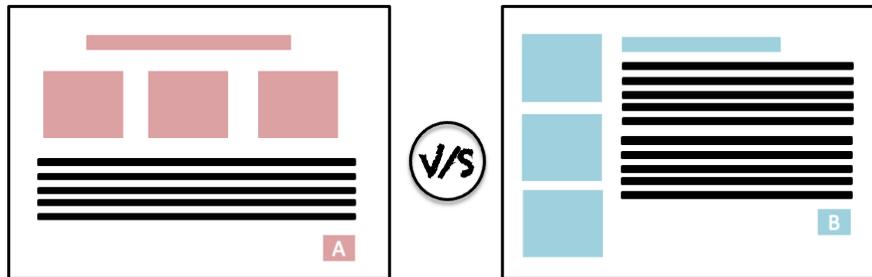


FIGURE 1.7 – Test A/B

Le test A/B fonctionne selon un principe simple : la comparaison des statistiques entre deux versions d'une même page. La démarche est la suivante (*Qu'est-ce que l'A/B testing, Guide complet pour tout savoir sur l'A/B testing. 2019*) :

1. Déterminer l'objectif : identifier le problème à résoudre. Cela peut concerner un élément de l'interface ou une fonction.
2. Déterminer les variantes : créer une version B alternative de A. Cette étape consiste à changer un élément qui peut être l'emplacement ou la couleur du bouton d'ajout au panier. Les deux interfaces doivent être similaires, avec la seule différence de l'élément à tester.
3. Faire le test : l'évaluation est menée dans des conditions réelles en soumettant les options potentielles à un ensemble d'utilisateurs d'une façon équitable : les utilisateurs sont répartis de manière égale entre les deux versions A et B.
4. Analyser les données et examiner les résultats : il existe deux critères incontournables pour s'assurer que les résultats a priori positifs d'un test A/B soient fiables et reproductibles : la taille de l'échantillon testé et l'indice de confiance statistique (*AB Tasty 2020*).
5. Apporter des modifications en fonction des résultats.
6. Préparer des tests de suivi.

Le test A/B permet d'analyser le comportement d'un utilisateur par rapport aux différentes variantes présentées (A ou B) afin de faciliter la prise de décision et atteindre plus rapidement un objectif donné. Il est généralement utilisé pour améliorer le taux de conversion d'un site web. Le taux de conversion est le ratio entre le nombre de visiteurs ayant effectué une action (achat par exemple) sur un site web par rapport au nombre total de visiteurs

sur une période donnée. Les concepteurs cherchent alors à optimiser leurs IHM pour engager l'utilisateur à passer à l'action.

Il existe plusieurs outils pour effectuer les tests A/B, la plupart étant commercialisés. Par exemple, AB Tasty ([AB Tasty 2020](#)), Kameleoon ([Une plateforme d'A/B testing et de personnalisation, web et full stack 2020](#)) ou Google Analytics ([Google Analytics 2020](#)) offrent une évaluation du taux de conversion. Dans la littérature, différents chercheurs ont présenté d'autres outils qui permettent d'évaluer d'autres critères tels que l'utilisabilité (Speicher et al. [2014](#)), la préférence de l'utilisateur (Vanderdonckt et al. [2019](#)) ou l'expérience utilisateur (UX) (Gardey et Garrido [2020](#)).

3.4 Concevoir et partager plusieurs IHM plutôt qu'une !

S.Dow et al. ont étudié les avantages que présentent la conception et le partage de plusieurs IHM sur les performances des concepteurs (Dow et al. [2011](#)). Ils présentent une étude dans laquelle il s'agit de concevoir et publier sur le web des IHM de publicité, sous une des trois conditions suivantes :

- Partager plusieurs : concevoir et partager trois propositions d'interfaces.
- Partager la meilleure : concevoir trois propositions d'interfaces mais ne partager qu'une seule.
- Partager une seule : ne concevoir et partager qu'une seule interface.

Quatre-vingt-quatre participants ont été divisés en trois groupes et ont travaillé par paires pendant l'étude. L'étude comportait trois étapes. Tout d'abord, il était demandé aux participants de préparer des propositions d'interfaces individuellement : une seule pour la troisième condition et trois pour les deux premières. La conception était réalisée via un outil graphique basé sur le web. Ensuite, il leur était demandé de mutuellement critiquer leurs interfaces. Puis, chacun devait produire une interface finale. Le même temps était alloué à chaque participant quelle que soit condition. Le seul paramètre qui changeait était le nombre d'interfaces à créer et à partager. Les différentes interfaces ont été jugées essentiellement par rapport à leur qualité en terme de performance. Elles ont été publiées sur une plateforme de publicité et jugées sur le nombre de sélections. Elles ont en plus été notées et classées par rapport à la préférence des différents utilisateurs et à leur qualité. Pour ce faire, il a été demandé de noter les interfaces sur leurs performances dans une campagne publicitaire sur une échelle de Likert de 7 points (Figure [1.8](#)).

Les interfaces du premier groupe (partager plusieurs) ont été jugées plus divergentes, performantes, de meilleure qualité, au plus grand nombre de sélections et aux notes les plus élevées.

Ainsi, cette étude montre que, lorsque les concepteurs élaborent et partagent plusieurs IHM dans des groupes de discussions, ils (1) explorent des idées plus diverses, (2) s'inspirent

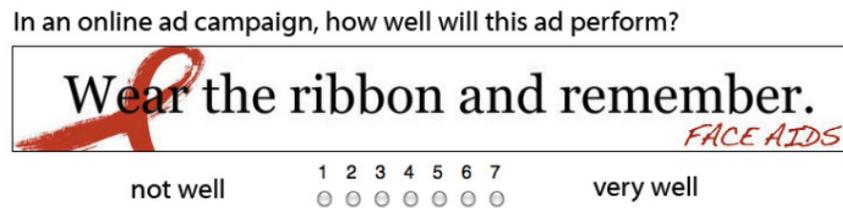


FIGURE 1.8 – Évaluation de la performance du design dans (Dow et al. 2011)

plus des travaux de leurs partenaires, (3) s’engagent dans des conversations plus productives et finalement, (4) produisent un travail de meilleure qualité. En outre, lorsque plusieurs alternatives sont présentées, les utilisateurs fournissent une critique plus approfondie et sont susceptibles de mieux explorer l’espace de conception. Ainsi, créer et partager rapidement plusieurs alternatives est bénéfique. Un outil logiciel permettant une exploration plus large est dès lors nécessaire.

4 Synthèse et conclusion

En IHM, l’ingénierie d’évaluation soulève plusieurs questions, notamment celle relative au choix de la méthode. Ce choix dépend de plusieurs facteurs comme la phase au cours de laquelle s’effectue l’évaluation : par exemple, en début de conception ou en test final après production. D’autres paramètres entrent en jeu comme, pour les évaluations comparatives, le nombre et la nature des alternatives présentées. Ces paramètres ont une incidence sur les résultats obtenus.

Plusieurs travaux ont étudié l’intérêt de l’évaluation comparative et confirment son bénéfice. Cependant, le nombre et le type d’alternatives à utiliser varient d’un travail à l’autre. Nous en présentons dans le tableau 1.3 une revue selon les critères suivants :

- **Phase** : phase du processus de développement pendant laquelle l’évaluation de l’interface est réalisée.
- **Méthode** : méthode d’évaluation utilisée.
- **IHM de référence** : caractéristiques de l’IHM de référence et propriétés du prototype utilisé.
- **IHM alternatives** : caractéristiques des alternatives de l’IHM utilisées.
- **Nombre d’alternatives** : nombre de variantes présentées aux utilisateurs.
- **Chronologie** : ordre de présentation des alternatives, successif ou simultané avec l’IHM de référence.
- **Génération** : méthode automatique ou manuelle de production des alternatives d’IHM.

	(Tohidi et al. 2006b)	(Kang et al. 2018)	Test A/B
Phase	Conception précoce	Produit final	Produit final
Méthode	Évaluation subjective. Test utilisateurs	Évaluation heuristique réalisée en ligne	Évaluation subjective réalisée en ligne
IHM de référence	Prototype interactif de basse-fidélité : maquette papier d'une seule page	Poster de haute-fidélité	IHM interactive (ou pas) de haute-fidélité
IHM alternatives	Prototype interactif de basse-fidélité qui représente une variation au niveau du style de présentation	Poster de haute-fidélité qui représente : - Une variation au niveau visuel - Une variation au niveau de contenu	Dépend du besoin de l'utilisateur ; Ajouter ou supprimer des fonctionnalités ou changer de style
Nombre d'alternatives	3	24 (étude 1) et 268 (étude 2)	2 (IHM A et IHM B)
Chronologie	Simultanée (3 interfaces ensembles à chaque fois)	L'interface de référence est présentée simultanément avec une galerie exemples	-
Génération	Manuelle	Automatique : Interfaces collectées, créées par d'autres designers (variation au niveau visuel). Interfaces collectées via une recherche sur le web (variation au niveau de contenu)	Manuelle

TABLE 1.3 – Tableau comparatif.

Cet état de l'art met l'accent sur le bénéfice de l'utilisation de différentes alternatives d'IHM pour son évaluation. Néanmoins, nous avons distingué quelques limites :

- Il n'existe pas de justification des alternatives à utiliser, ni de méthode de génération. Aujourd'hui, les variantes sont, soit conçues manuellement, soit collectées. Le premier cas nécessite des règles de transformation d'IHM et induit un temps de réflexion substantiel. Le deuxième cas suppose des banques d'IHM et un temps de filtrage et de sélection dans différents domaines.
- Il n'a pas été défini ce qu'est un bon ou un mauvais exemple à présenter aux utilisateurs. En effet, la qualité des retours des utilisateurs n'a pas été évaluée en fonction des alternatives présentées.
- Enfin, les travaux traitent d'IHM graphiques mono-page, sans considérer véritablement l'interaction.

Une approche systématique pour soutenir la génération d'IHM alternatives est donc une piste prometteuse. Nous explorons au chapitre suivant les différentes méthodes et outils d'exploration et de génération d'IHM disponibles. Cette étude nous aidera pour la génération automatique de la meilleure alternative en évaluation comparative.

2

Exploration des alternatives

1 Introduction

Nous étudions ici les techniques d’exploration de l’espace de conception pour la génération des IHM alternatives. Nous commençons par étudier les techniques de prototypage qui aident les concepteurs à explorer rapidement un maximum d’alternatives lors des premières phases de conception. Nous nous intéressons ensuite aux outils de génération automatique d’IHM. Puis, nous présentons les approches dirigées par les modèles pour la génération d’IHM. Enfin, nous concluons le chapitre par une synthèse.

2 Explorer des IHM alternatives

En phase amont de conception, les concepteurs esquissent plusieurs variantes d’IHM (phase d’élaboration). Le processus est itératif, aboutissant à des IHM soit rejetées, soit sélectionnées pour amélioration. Les IHM conservées sont ensuite affinées au fur et à mesure jusqu’à parvenir à la bonne IHM (phase de réduction) (Buxton 2010). Afin d’augmenter les chances de trouver cette bonne IHM, les premières phases de conception nécessitent l’exploration de nombreuses solutions et la proposition de plusieurs variantes de l’IHM (Figure 2.1). Afin de favoriser la créativité des concepteurs et augmenter le nombre de pistes explorées, plusieurs techniques sont utilisées, telles que le *brainstorming* (Taylor et al. 1958), le *brain-writing* (Heslin 2009; Linsey et Becker 2011), les activités de prototypage (Schütze et al. 2003; Tohidi et al. 2006a) ou l’exploration des alternatives. Par exemple, afin de produire un maximum de variantes d’une IHM au début de la conception, les concepteurs utilisent des prototypes de faible coût tels que des croquis, storyboards ou vidéos. Le but ici est d’explorer rapidement plusieurs idées et solutions à un problème de conception afin de les tester et de les discuter avec les membres de l’équipe. En outre, l’exposition à des exemples ou à des suggestions d’alternatives augmente le nombre d’idées générées et favorise la créativité (Smith et al. 1993; Bonnardel 1999; Masson et al. 2011). En effet, l’exploration des alternatives est considérée avoir un rôle critique lors du processus créatif de conception et est couramment utilisée par les concepteurs (Herring et al. 2009; Masson et al. 2010). Marsh et al. (Marsh et al. 1996) suggèrent que l’exposition aux exemples impacte la conformité et l’originalité dans les tâches créatives et qu’un transfert positif d’éléments existe depuis les exemples vers les futures propositions. Cela permet de prendre conscience de l’espace de conception et des options potentielles (Kim et Yoon 2005).

Diverses approches et différents outils ont été proposés dans la littérature servant comme assistants à l’inspiration et à l’exploration. Nous explorons dans ce qui suit les travaux les plus pertinents par rapport à notre travail de recherche. Nous commençons par présenter deux techniques de prototypage parmi les plus utilisées dans les premières phases de conception pour la génération et l’exploration rapide des alternatives. Ensuite, nous présentons des travaux permettant l’exploration d’alternatives d’IHM et nous finissons par la présentation

de travaux qui permettent une génération (automatique) de différentes alternatives d'une IHM donnée.

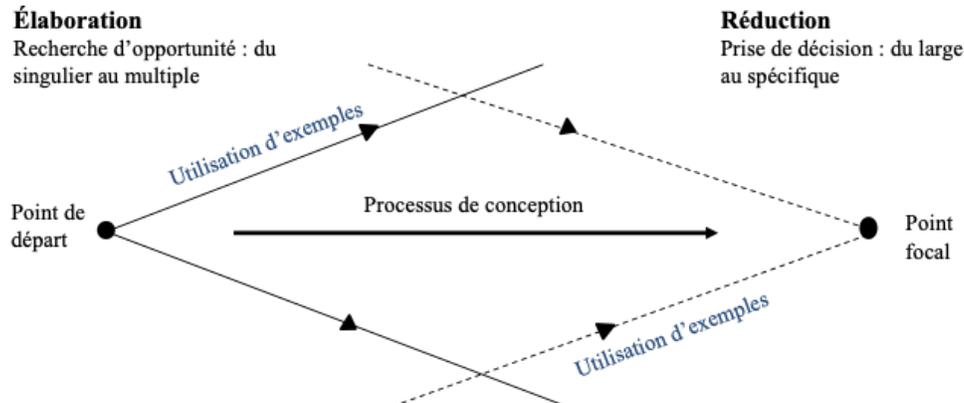


FIGURE 2.1 – Utilisation des exemples dans le processus de conception des IHM (Buxton 2010)

2.1 Explorer des alternatives en esquissant

Une façon d'explorer rapidement différentes alternatives dans les premières phases de conception est de créer des esquisses des différentes conceptions possibles de l'IHM étudiée. Ces esquisses, élaborées dans la plupart du temps de manière aléatoire, sont ensuite développées au fil des itérations. Nous présentons ici les différentes techniques de prototypage permettant une exploration rapide d'alternatives.

2.1.1 Le prototypage

Le prototypage sert à explorer et à tester différentes alternatives de conception, à obtenir des informations sur l'interaction des utilisateurs avec le futur produit, à aider l'utilisateur à imaginer le système et à se concentrer sur les parties problématiques d'une IHM. Un prototype peut être une entité aussi simple qu'un croquis ou aussi complexe qu'un système interactif entièrement fonctionnel. Ainsi, il existe différents types de prototypes qui peuvent être utilisés pour des finalités différentes. En effet, un prototype peut être représenté avec différents outils, sur différents supports (papier ou écran) et avec différents niveaux de détails et de fidélité par rapport à l'IHM finale (Coyette et al. 2007). Le choix du niveau de fidélité du prototype est généralement basé sur les objectifs de prototypage, de la complétude de la conception et des ressources disponibles (Walker et al. 2002). Le niveau de fidélité est considéré haut si le prototype a le même style de présentation que l'IHM finale en termes de mise en page ou d'éléments choisis pour la navigation (globale et locale) par exemple. Au contraire, le

niveau de fidélité est considéré bas si le prototype ne représente que partiellement l'IHM finale (Coyette et al. 2007). D'après la littérature, le niveau de fidélité d'un prototype à évaluer peut affecter les retours des participants pendant l'évaluation (Coyette et al. 2007 ; Tohidi et al. 2006a ; Wiklund et al. 2016). En effet, lorsque l'on évalue un prototype de basse ou de moyenne fidélité, les évaluateurs se sentent davantage impliqués dans le processus de conception et se sentent plus libres de critiquer le prototype. Cependant, il a été constaté que les retours utilisateurs collectés portant sur l'utilisabilité d'une IHM en utilisant des prototypes de basse-fidélité ou de moyenne-fidélité sont comparables à ceux collectés en utilisant un prototype de haute-fidélité (Wiklund et al. 2016 ; Walker et al. 2002 ; Dhillon et al. 2011). L'usage de différents prototypes peut également varier selon ce que nous avons besoin de présenter : le rôle, l'aspect ou le comportement (*look and feel*) (Houde et Hill 1997).

2.1.2 Différents types et formes des prototypes

Il existe différentes approches et méthodes de prototypage. Le but de cette partie n'est pas de présenter une étude exhaustive sur les différentes caractéristiques des prototypes mais d'étudier les techniques de prototypage les plus utilisées pour la génération et l'exploration rapide des alternatives. Nous présentons ici deux types de prototypage : papier et vidéo.

Le prototypage papier est une des techniques les plus utilisées par les concepteurs grâce à son faible coût et son efficacité. Le prototypage papier peut jouer le rôle d'esquisse (croquis/sketch) afin de développer, d'explorer, de communiquer et d'évaluer les idées du concepteur. Il offre un moyen rapide et facile de communiquer et de tester les idées initiales, par exemple dans des séances de brainstorming. Cette technique de prototypage aide à obtenir un retour substantiel des utilisateurs, favorisant ainsi le développement itératif rapide (Walker et al. 2002) et permettant d'appliquer des modifications faciles et peu coûteuses (en termes de ressources matérielles) sur les IHM proposées (Tohidi et al. 2006a).

Le prototypage papier consiste à produire des artefacts qui mettent en avant les éléments essentiels d'une idée de conception. Ils sont rapides à produire et facilitent l'exploration en largeur de l'espace de conception. La production de ces derniers offre la possibilité aux concepteurs "*de détecter des problèmes qu'ils n'avaient pas encore envisagés, voire découvrir de nouvelles caractéristiques et relations entre les concepts qu'ils n'avaient pas intentionnellement inclus dans leurs croquis*" (Masson 2014).

Ce type de prototype peut être dessiné à la main (Tohidi et al. 2006a) ou produit à l'aide d'outils de prototypage tel que *Balsamiq Mockups* ou *Mockflow* qui offrent un moyen de création de prototypes de basse et de moyenne fidélité (Figure 2.2).

Un autre type de prototype rapide et économique est le prototype vidéo. Le prototypage vidéo est une technique généralement utilisée dans les premières étapes de conception pour permettre aux concepteurs de systèmes interactifs d'évaluer l'interaction avant la mise en œuvre réelle de ces systèmes. Les prototypes vidéo présentent de grands avantages (Dhillon

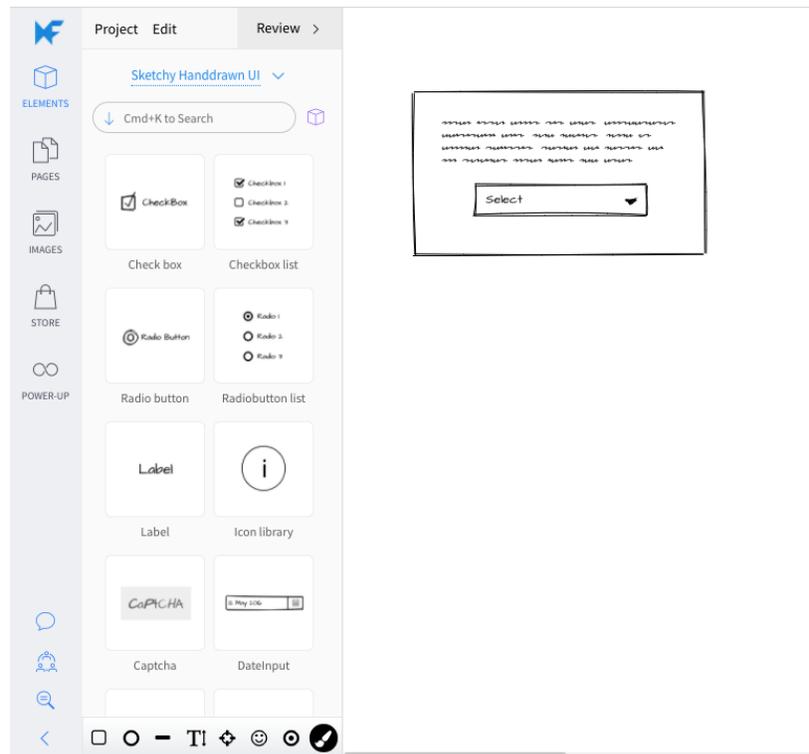


FIGURE 2.2 – Exemple de création de prototype de basse fidélité avec un outil digital (Mock-flow (*Super-easy tool to brainstorm UI ideas* Retrieved : December, 2019))

et al. 2011 ; Leiva et Beaudouin-Lafon 2018 ; Mackay 1988 ; Mackay 2002 ; Mackay et al. 2000). Ils favorisent l’exploration de différentes conceptions (Mackay et al. 2000 ; Mackay 1988) et aident à capturer et à communiquer les détails sur la façon selon laquelle les utilisateurs interagissent avec un système interactif (Leiva et Beaudouin-Lafon 2018). Le gros avantage de cette technique est que la production d’un prototype vidéo coûte moins cher et prend moins de temps à préparer que la construction d’un prototype entièrement fonctionnel.

Les prototypes vidéo représentent un outil puissant pour les évaluations des IHM car ils aident à collecter des retours utilisateur qui peuvent être comparés aux retours collectés avec d’autres techniques d’évaluation plus coûteuses telles que des tests d’utilisabilité. En effet, Zwinderman et al. (Zwinderman et al. 2013) ont comparé les résultats d’une évaluation utilisant des prototypes vidéo avec les résultats d’une évaluation s’appuyant sur des tests d’utilisabilité. Les auteurs confirment que l’utilisation des prototypes vidéo aide à obtenir des retours utilisateur similaires à ceux recueillis lorsque les utilisateurs testent un produit final.

Une vidéo peut présenter un utilisateur en train d’utiliser un système dans la vie réelle (Leiva et Beaudouin-Lafon 2018 ; Dhillon et al. 2011) ou exposer un scénario d’utilisation du système. Il existe différents moyens de les concevoir, soit en filmant une scène d’utilisation

à l'aide de caméras, soit en faisant une capture vidéo ou en filmant un utilisateur en train de simuler l'interaction sur un prototype papier. Une vidéo peut donc présenter différents niveaux de fidélité. Une étude élaborée par Dhilton et al. (Dhillon et al. 2011) indique que le niveau de fidélité visuel du prototype vidéo n'affecte pas la quantité et la qualité des retours utilisateurs pendant les sessions d'évaluation. La figure 2.3 présente un exemple d'une vidéo de haute-fidélité et une de basse-fidélité utilisées dans l'étude de Dhillon et al. (Dhillon et al. 2011)



FIGURE 2.3 – À gauche, une vidéo basse-fidélité et, à droite, une vidéo haute-fidélité (Dhillon et al. 2011)

Les prototypes papier et vidéo sont des moyens faciles et rapides pour explorer de nombreuses pistes en termes d'interface et d'interaction en début de conception. Ils sont peu coûteux à produire (en termes de ressources matérielles et de moyens humains) et se sont avérés bénéfiques pour l'exploration et l'évaluation des IHM. Cependant, le concepteur explore ici différentes alternatives produites par sa propre imagination, mais n'explore pas ce qui a été fait par d'autres concepteurs, ce qui limite l'espace de conception exploré.

Nous nous intéressons dans ce qui suit aux différentes techniques facilitant l'exploration et la génération des alternatives à finalité d'inspiration.

2.2 Explorer des alternatives à l'aide des galeries web

Les galeries web sont des sites web qui collectent, catégorisent et présentent différents exemples de sites web dans leur contenu principal. Nous présentons ici deux outils existants parmi les plus cités dans la littérature.

2.2.1 Outil Adaptive Ideas Web Design

Lee et al. présentent un outil de conception de sites web nommé *Adaptive Ideas Web Design* qui se base sur les galeries d'exemples (Lee et al. 2010). Dans ce travail, la galerie des exemples exposée est un ensemble de sites web qui sont organisés selon quatre attributs :

- couleur d'arrière-plan
- police principale
- nombre de colonnes
- densité visuelle.

Les utilisateurs peuvent donc chercher des exemples similaires à un exemple sélectionné par rapport à un des attributs définis, pour s'inspirer de variations mineures. Ou bien, ils peuvent choisir des exemples qui représentent une variété d'options le long d'un attribut particulier, pour s'inspirer de variations majeures. Par exemple, si l'utilisateur choisit de chercher des exemples par rapport à la couleur d'arrière-plan, la galerie d'exemples affiche les sites web qui présentent une similarité par rapport à la couleur de l'exemple sélectionné.

L'outil utilise une approche de conception par la modification d'exemples. Cette approche consiste à créer des IHM en réutilisant et en modifiant des IHM existantes. Adaptive Ideas se compose d'une galerie d'exemples de sites web, d'un volet d'édition et d'un volet d'aperçu. Il permet de parcourir de manière paramétrique l'ensemble des exemples, ensuite les utilisateurs peuvent en récupérer des éléments et les utiliser dans leurs propres travaux. Ils peuvent ainsi modifier et combiner des éléments de plusieurs exemples 2.4. L'aperçu permet de voir un exemple en détail et d'en sélectionner un élément à utiliser. En outre, l'exemple sélectionné peut servir de pivot pour mettre à jour la galerie proposée.

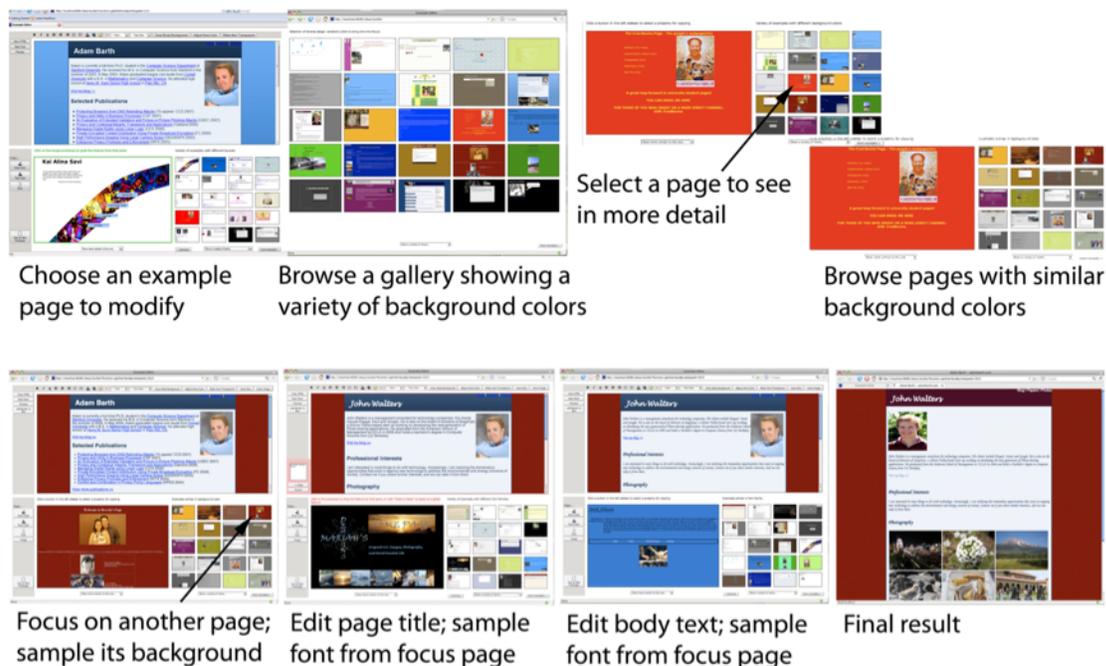


FIGURE 2.4 – Outil *Adaptive Ideas Web Design* (Lee et al. 2010)

Cet outil propose une recherche par propriété de style et présente un mécanisme pour copier directement des éléments intéressants d'une IHM existante. Cependant, du point de vue de l'interactivité d'un site web, cet outil ne permet ni la recherche par type d'interaction, ni l'emprunt de fonctionnalités interactives.

2.2.2 Outil d.tour

Ritchie et al. présentent d.tour, un outil proposant des galeries d'exemples de sites web basées sur des techniques de recherche par similitude stylistique (Ritchie et al. 2011). Contrairement à l'outil *Adaptive Ideas* proposé par Lee et al. (Lee et al. 2010), la recherche des exemples dans ce travail se fait de manière automatique. Le style est automatiquement identifié à partir de la structure d'une page HTML de référence. Le style est ici défini par quatre aspects majeurs :

- la complexité de la page, calculée à partir du nombre d'éléments et de leur positionnement dans l'arbre HTML
- la densité visuelle, basée sur la quantité d'images et de texte dans le document
- la couleur exprimée en température, en luminosité et en saturation
- la typographie limitée à la taille de la police.

Cependant, comme *Adaptive Ideas Web Design*, cet outil ne traite pas l'aspect interactif de l'IHM. Comme l'indiquent les auteurs : "d.tour. ne tente pas de capturer des aspects non statiques tels que l'animation et l'interaction" (Ritchie et al. 2011).

Les galeries web permettent aux concepteurs d'avoir une idée sur la structure de sites web dans différents domaines (par exemple, sites de e-commerce, sites vitrine, blogs, etc.), d'être à jour par rapport aux nouvelles tendances en développement web et d'évaluer leurs idées par rapport à des sites web similaires. Cependant ces exemples sont les résultats d'autres concepteurs, ce qui limite l'espace de conception exploré à une partie déjà explorée par d'autres concepteurs humains. Dans d'autres travaux ((O'Donovan et al. 2015; Todi et al. 2016)), les alternatives sont générées (semi-)automatiquement en faisant évoluer un ensemble de variables d'une IHM de référence (celle du concepteur) de sorte que les résultats produits présentent une grande variété de possibilités.

3 Générer automatiquement des alternatives à partir d'une IHM

Nous présentons dans cette section deux outils parmi les plus connus de la littérature permettant la génération automatique d'IHM alternatives.

3.1 DesignScape

O’Donovan et al. présentent DesignScape (O’Donovan et al. 2015), un éditeur visuel d’interfaces graphiques statiques telles que des affiches, dépliants ou posters. La particularité de cet outil est qu’il fournit une assistance pendant le processus de conception en proposant des suggestions interactives à une interface de référence en temps réel. Ces dernières exposent des changements de mises en pages (layout) dans une variété de styles. Les suggestions proposées portent sur des changements au niveau de la disposition des éléments qui composent l’interface, leur position et leur taille. Elles sont divisées en deux types :

- Des suggestions de raffinement : ce type de suggestion vise à améliorer une mise en page courante (créée par l’utilisateur). Dans ce cas, des améliorations légères dans la disposition des composants et l’alignement sont proposées.
- Des suggestions de plus grande échelle appelées suggestions de brainstorming, où chaque mise en page est distincte des autres. Dans ce cas, les suggestions varient en termes de fonction de la symétrie, de la taille du texte et des graphiques ou des préférences d’alignement. Ce type de suggestion aide à explorer les mises en pages possibles.

Les suggestions de raffinement sont générées en se basant sur un modèle de calcul de la disposition des différentes composantes. Ce modèle se base sur une formule d’énergie qui évalue une mise en page X par rapport à différents principes de design tels que l’alignement et la symétrie (figure 2.5) Les suggestions de brainstorming sont générées en se basant sur une approche d’apprentissage automatique (*Machine Learning*).



FIGURE 2.5 – Détection d’alignement (gauche) et segmentation du design en zones de texte et graphiques (droite) (O’Donovan et al. 2015)

La figure 2.6 présente l’IHM de DesignScape. Le canevas central permet à l’utilisateur de créer des mises en page dans un éditeur simple. Sur la gauche, le système propose des suggestions de raffinement (mineures). Sur la droite, le système fournit des suggestions de brainstorming (majeures).

Le système fournit des suggestions à l’utilisateur tout au long du processus de conception. À chaque modification de l’utilisateur, de nouvelles suggestions lui sont proposées. En outre,

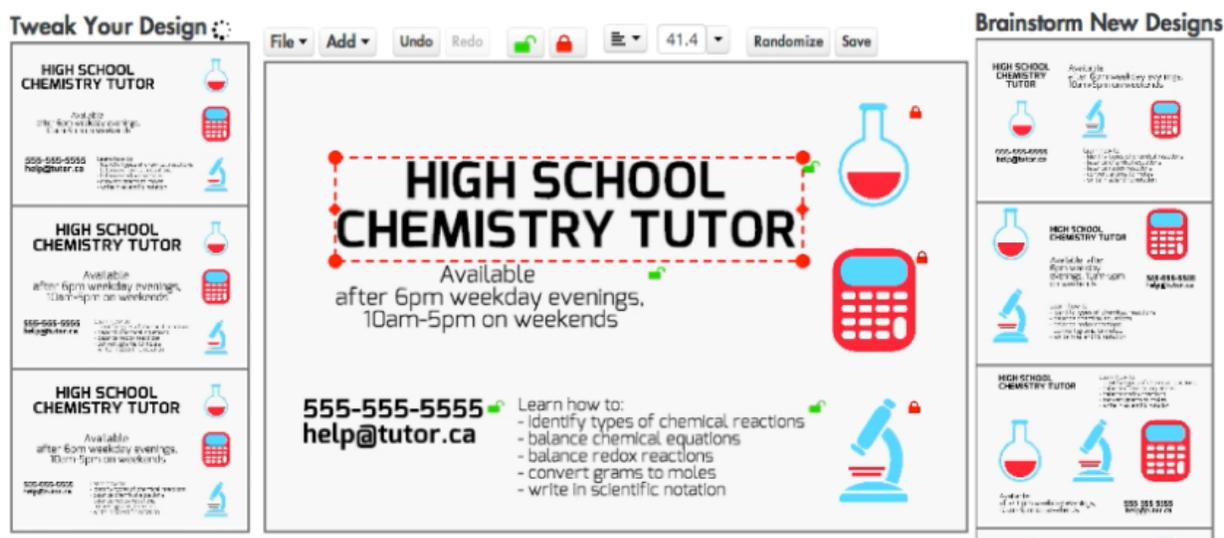


FIGURE 2.6 – IHM de l’outil DesignScape (O’Donovan et al. 2015)

l’utilisateur peut fixer la taille d’un élément ou contrôler les suggestions en verrouillant un élément, ce qui évite qu’il soit repositionné automatiquement.

Cet outil présente un moyen permettant d’améliorer automatiquement la mise en page d’une interface de référence en temps réel et de manière fluide. L’évaluation de DesignScape a montré que les utilisateurs créent de meilleures interfaces avec les suggestions proposées par l’outil que sans s’appuyer sur ces suggestions.

3.2 Sketchplore

Dans le même esprit que DesignScape, Todi et al. ont développé un outil de prototypage interactif qui combine esquisse et exploration par optimisation de la mise en page (layout), d’où son nom Sketchplore (pour *Sketch and explore*). L’outil permet l’optimisation en temps réel, c’est-à-dire pendant que le concepteur fait son croquis, rendant possible l’exploration et l’exploitation de différentes alternatives basées sur ce dernier.

Les alternatives proposées par le système portent sur des changements locaux et/ou globaux. Ceci permet au concepteur de rechercher à la fois des améliorations locales à son esquisse actuelle (explorer) et des alternatives plus radicales (exploiter). Les changements locaux proposent des améliorations au niveau des deux aspects spatiaux et colorimétriques de l’esquisse initiale (réparer et recolorer). Le but ici est d’améliorer l’esquisse du concepteur en minimisant l’encombrement visuel par exemple ou en proposant des couleurs plus harmonieuses. Cette optimisation est basée sur des modèles permettant de couvrir des caractéristiques esthétiques et des mesures de performance sensorimotrice comme le modèle Kieras-Hornof proposé par D.Kieras et al. (Kieras et Hornof 2014). L’outil dans ce cas cherche des

alternatives qui sont meilleures en termes de qualité esthétique, ainsi que des performances des utilisateurs. Les alternatives présentant des changements globaux exposent, quant à elles, des modifications sur une plus grande échelle spatiale. Ces suggestions orientent le concepteur vers des IHM qui sont esthétiquement meilleures et plus utilisables sans le surcharger avec le processus d'optimisation.

L'outil Sketchplore se concrétise dans un large espace de travail qui permet à un utilisateur de (1) dessiner son IHM, (2) de modifier/raffiner cette dernière et d'ajouter des détails aux éléments utilisés. Ces détails concernent le type de l'élément utilisé ou son importance qui est présenté par une valeur de probabilité d'utilisation. Ensuite (3) l'outil propose des alternatives d'amélioration dans le cas où l'utilisateur choisit l'option " réparer et recolorer " ou bien (4) d'explorer des alternatives de conception qui soient optimisées de façon plus globale (Figure 2.7).

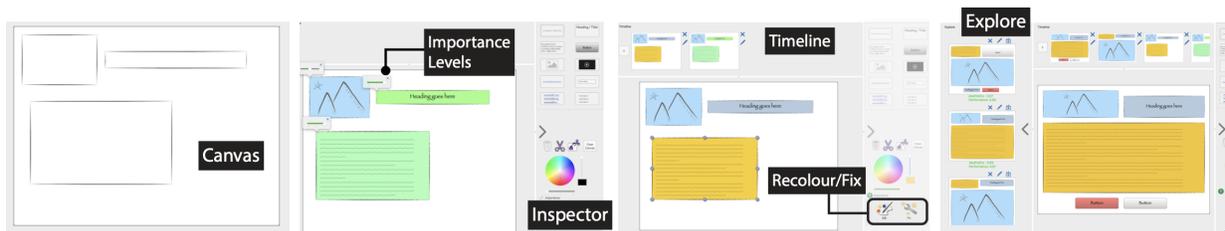


FIGURE 2.7 – Outil Sketchplore - Utilisation (Todi et al. 2016)

Cet outil permet aux utilisateurs de mieux explorer l'espace de conception. Cependant une limite importante de cet outil est qu'il ne traite pas les aspects sémantiques de l'IHM, comme le regroupement des éléments par exemple.

Par ailleurs, ces travaux de recherche se sont concrétisés dans différents outils commercialisés tels que *PowerPoint* de *Microsoft Office*, qui a intégré depuis sa seizième version (2016) une fonctionnalité appelée " idées de conception (*design ideas* en anglais) ". Cette fonctionnalité propose automatiquement et en temps réel des alternatives de mise en page des diapositives (Figure 2.8). Le logiciel détecte les images, graphiques, tableaux et textes saisis par l'utilisateur dans sa diapositive et propose soit des images qui reflètent le texte de la diapositive, soit des suggestions d'organisation de la diapositive. Le logiciel propose aussi un modèle de conception avec des couleurs assorties à une image choisie et associe visuellement toutes les diapositives de la présentation.

Le processus de conception d'une IHM consiste en une succession d'étapes de génération et de sélections d'idées. Une façon de faire est de (1) commencer directement par esquisser une idée et la concrétiser dans des prototypes, (2) améliorer les prototypes jusqu'à arriver au résultat souhaité. Cette méthode définit l'essentiel, mais pas forcément la totalité, faute de

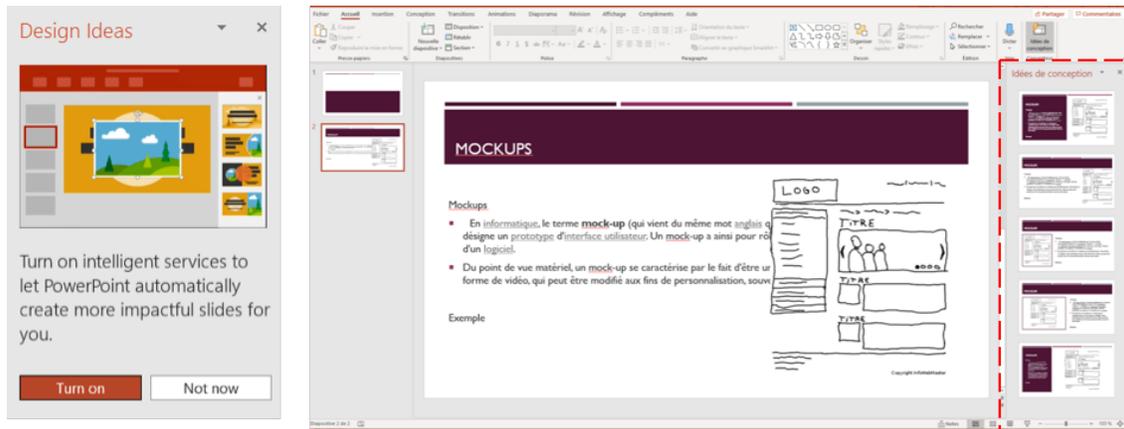


FIGURE 2.8 – Idées de conception dans Microsoft Office PowerPoint

temps ou de spécifications initiales exhaustives, de l'IHM.

Une autre méthode de travail consiste à utiliser des patrons pour la conception des IHM. "Un patron décrit un problème à résoudre, une solution, et le contexte dans lequel cette solution se situe" (Johnson 1997). Un point fort de cette approche est qu'elle offre la possibilité d'explorer différentes alternatives tout au long du processus de conception. En effet, des transformations peuvent être appliquées aux différents modèles de l'IHM tels que le modèle de concepts et de tâches afin de produire différentes IHM (Calvary et al. 2003). Ces transformations reposent sur une capitalisation de la manière dont les concepts, tâches et opérateurs entre tâches peuvent être représentés concrètement (Calvary et al. 2003). Dans ce contexte, différents travaux se sont intéressés à cette approche pour la génération d'alternatives d'IHM (Calvary et al. 2002 ; Masson et al. 2010 ; Limbourg et al. 2004).

Nous présentons dans la section suivante le principe de l'approche par modèles en présentant les travaux les plus pertinents par rapport à notre recherche.

4 Produire des IHM par transformation de modèles

L'Ingénierie Dirigée par les Modèles (IDM) a pour point de départ la séparation des modèles indépendants des plates-formes (Platform Independent Model ou PIM) et des modèles spécifiques aux plates-formes (Platform Specific Model ou PSM) (Sottet et al. 2005).

L'IDM pour la conception des IHM consiste à concevoir et à développer une ou plusieurs IHM d'un système interactif en faisant intervenir divers modèles tels que le modèle de tâche, le modèle de domaine, le modèle d'interface utilisateur, le modèle de plateforme, le modèle utilisateur, etc. La conception de l'IHM, s'articule alors autour de différents niveaux d'abstraction allant de la spécification du domaine (concept et tâches) à l'IHM finale (Calvary

et al. 2002).

4.1 Cadre de référence CAMELEON pour la génération des IHM

Le projet CAMELEON (Calvary et al. 2003) propose un cadre conceptuel adapté à la spécification d'IHM qui s'adaptent à une plateforme d'accueil : " *Le cadre de conception CAMELEON explicite et formalise les différents aspects qui concourent à la production d'IHM plastiques en termes de modèles et de relations entre modèles. Il est en cela conforme aux principes de l'Ingénierie Dirigée par les Modèles (IDM) du Génie Logiciel* " (Coutaz et al. 2012). Ce cadre de conception a identifié différents niveaux d'abstraction clés en conception d'IHM. Les auteurs distinguent quatre niveaux principaux : le niveau domaine (modèle de tâches utilisateur et concepts), l'Interface Utilisateur Abstraite (IUA), l'Interface Utilisateur concrète (IUC) et l'Interface Utilisateur Finale (IUF) (Figure 2.9).

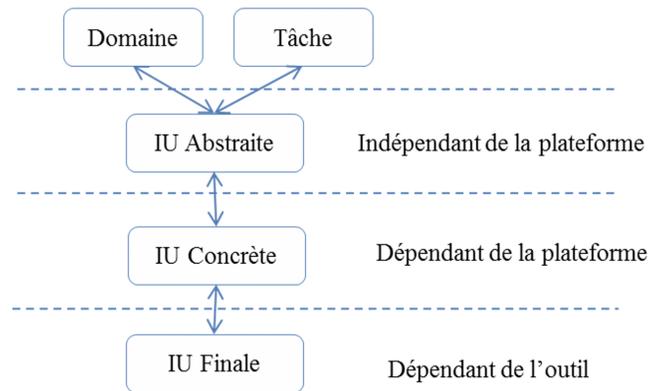


FIGURE 2.9 – Les quatre niveaux d'abstraction du cadre de référence CAMELEON (Calvary et al. 2003)

- **Le modèle de tâche et le modèle des concepts** décrivent respectivement les tâches utilisateur en termes d'objectifs et de procédures, et les concepts du domaine en termes d'entités et de relations. Les procédures explicitent comment atteindre les buts et sous-butts que l'utilisateur se fixe avec le système. Les tâches sont décorées des concepts du domaine qu'elles manipulent pour l'accomplissement de leurs buts (Coutaz et al. 2012). À ce niveau, la description est indépendante de toute modalité d'interaction.

La figure 2.10 présente une modélisation pour la tâche de réservation d'une chambre d'hôtel formalisée avec la notation ConcurTaskTrees (CTT) (Paternò 2004). Cette dernière est une notation visuelle pour décrire les modèles de tâches.

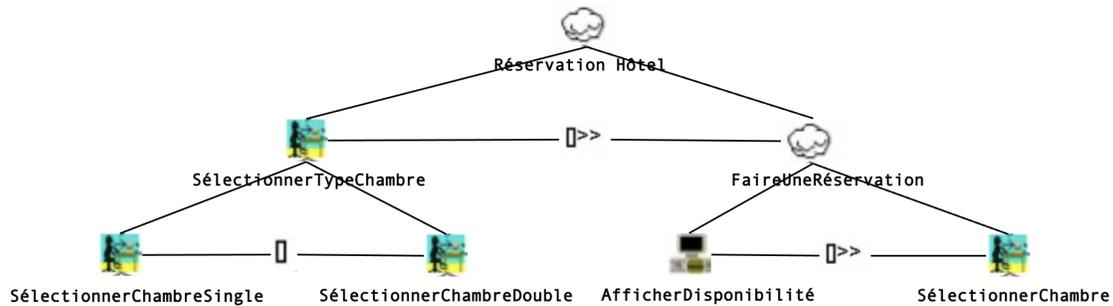


FIGURE 2.10 – Exemple de modèle de tâche

Dans CTT, il existe quatre types de tâches : les tâches abstraites, interactives, système et utilisateur. L'icône en forme de nuage représente une tâche abstraite ; les icônes avec les personnes en face d'écrans représentent une tâche interactive, enfin l'icône en forme de machine représente une tâche système. Les deux opérateurs utilisés dans ce modèle sont l'opérateur de choix ([]). Cet opérateur active deux (ou plus) tâches possibles. L'opérateur d'activation avec un passage d'information ([]>) signifie qu'une deuxième tâche (T2) ne peut s'effectuer qu'après l'exécution d'une première tâche (T1) : quand T1 se termine elle fournit des informations à T2 et l'active.

Les tâches peuvent être décorées de propriétés pour préciser, par exemple, que la tâche est itérative (c'est-à-dire réalisable plusieurs fois) ou optionnelle (c'est-à-dire non obligatoire). Les tâches sont également décorées de concepts pour spécifier les entités du domaine manipulées dans la tâche. Les concepts sont généralement modélisés dans un diagramme de classe UML.

- **L'Interface Utilisateur Abstraite (IUA)** décrit la structuration de l'IHM en termes d'espaces de travail (appelés aussi espaces de dialogue) et de navigation entre ces espaces, aussi indépendamment que possible de la plate-forme cible (Coutaz et al. 2012). Un espace de travail est un groupement d'informations nécessaires et suffisantes pour l'accomplissement d'une tâche. Un espace de travail peut être composé de sous espaces de travail. L'ensemble des espaces de travail d'un système interactif et leurs enchaînements constituent son IHM abstraite. À ce niveau, l'apparence de l'IHM est encore indéterminée.

La figure 2.11 présente un exemple d'identification d'espaces de travail à partir du modèle de tâches présenté dans la figure 2.10.

- **L'Interface Utilisateur Concrète (IUC)** représente l'IHM en termes d'interacteurs (par exemple, des boutons ou des menus). Ici, l'apparence de l'IHM est fixée (Coutaz et al. 2012).
- **L'Interface Utilisateur Finale (IUF)** désigne l'IHM telle que l'utilisateur la perçoit

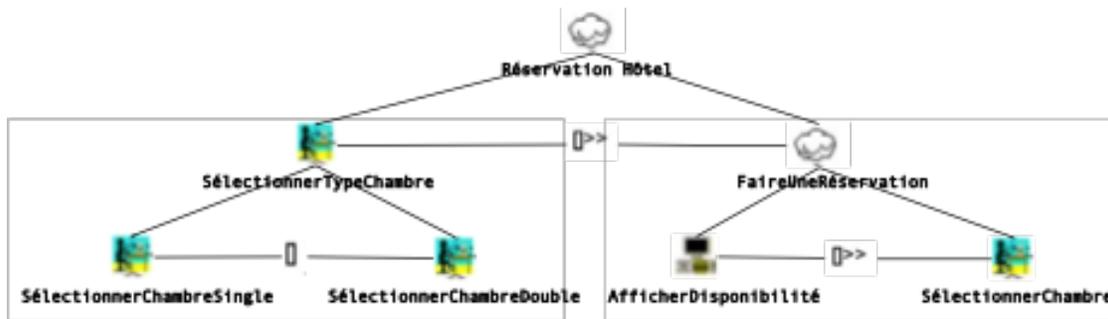


FIGURE 2.11 – Exemple d’espaces de travail sur le modèle de tâche

et la manipule à l’exécution. L’IUF fait donc un choix d’espace technologique (JAVA-SWING, PHP-HTML, etc.).

Chaque niveau d’abstraction possède son vocabulaire et fixe, en conséquence, le type de choix de conception possibles à ce niveau. Les caractéristiques sont présentées dans le tableau 2.1.

Ces quatre niveaux (tâche et concept, IUA, IUC et IUF) sont en correspondance par une relation de réification (passant d’un niveau abstrait à un niveau concret) et d’abstraction (d’un niveau concret à un niveau abstrait) (Calvary et al. 2003). Dans une approche descendante, le code est généré par transformation de ces modèles (*forward engineering*). Inversement, les modèles peuvent être découverts à partir du code (*reverse engineering* et *models recovery*). Différents travaux ont été proposés dans la littérature proposant des processus de transformation des différents modèles (Tran et al. 2012; Ammar 2019).

4.2 Magellan

D. Masson propose un outil interactif pour générer semi-automatiquement des alternatives d’IHM correspondant à un modèle de tâches donné (Masson et al. 2010). L’outil, nommé Magellan, cible les concepteurs en leur proposant des exemples d’IHM adaptées au problème de conception traité dans le but de favoriser leur créativité. Magellan combine l’utilisation des Algorithmes Génétiques Interactifs (AGI) et les approches basées sur les modèles pour la génération de ces alternatives. Ces dernières consistent en des combinaisons des éléments d’une IHM et de leurs paramètres (couleur, taille, positionnement, etc.) en respectant un périmètre fonctionnel fixé défini par le modèle de tâches donné par l’utilisateur.

L’outil prend en entrée un modèle de tâches et crée puis fait évoluer une population d’IHM et de transformations en s’appuyant sur une base de connaissances structurées en un réseau sémantique. Cette base permet d’associer à chaque tâche ou opérateur entre tâches les IHM possibles ainsi que les paramètres de personnalisation de ces dernières. Ces transformations sont ensuite appliquées au modèle de tâches afin de générer des interfaces utilisateur

Modèle de tâches	Modèle de l'IUA	Modèle de l'IUC
<ul style="list-style-type: none"> - Objectifs de l'utilisateur : une transformation possible est d'ajouter ou de supprimer une fonctionnalité interactive - Opérateurs entre tâches : une transformation possible est de remplacer un choix par une séquence. - Décorations de tâches : une transformation possible est de déclarer une tâche fréquente ou itérative. 	<ul style="list-style-type: none"> - Groupement des tâches : une transformation possible est de grouper les tâches fréquentes pour les séparer des non fréquentes. Une autre transformation possible est de factoriser un concept du domaine partagé par plusieurs tâches - Navigation : une transformation possible est de lancer l'application sur l'espace des tâches fréquentes et de forcer les tâches non fréquentes à repasser par les tâches fréquentes 	<ul style="list-style-type: none"> - Interacteurs : une transformation possible est de remplacer un calendrier par un champ texte pour la saisie d'une date, ou une liste de boutons radio par un menu déroulant - Paramètres interacteurs : une transformation possible est de changer la couleur, la taille, ou la position d'un champ texte

TABLE 2.1 – Classes de variations dans les trois modèles : tâches, IUA et IUC

concrètes. Les COMETs (COntext Mouldable widgET) ont été utilisés afin d'aider à générer des transformations qui s'appuient sur les concepts associés aux tâches pour configurer les éléments d'IHM dans un but esthétique et/ou de guidage. Une COMET est un widget, capable de s'adapter à son contexte d'usage : "*une COMET est polymorphe en termes d'incarnation. Le polymorphisme peut s'exercer à différents niveaux d'abstraction : concepts et tâches, interfaces abstraite, concrète et finale. Chaque niveau se décrit d'un point de vue fonctionnel et non fonctionnel*" (Calvary et al. 2004). Enfin, Magellan permet à l'utilisateur de définir son propre AGI. Ce dernier est configuré afin d'avoir une population initiale de seize IHM générées aléatoirement. L'utilisateur sélectionne ensuite les IHM qui l'intéressent. Ces dernières participent ensuite à la reproduction d'autres exemples d'IHM.

La figure 2.12 montre un exemple de deux IHM générées à partir d'un modèle de tâche décrivant un système de messagerie instantanée. L'IHM à gauche présente un exemple où les lignes de discussion sont différenciées par couleur selon l'intervenant. L'IHM à droite présente un exemple où la partie supérieure de l'IHM contient un panneau à onglets.

L'évaluation de cet outil s'est effectuée à l'aide de onze participants présentant des profils différents : des programmeurs, des designers, des experts en IHM et des concepteurs débu-

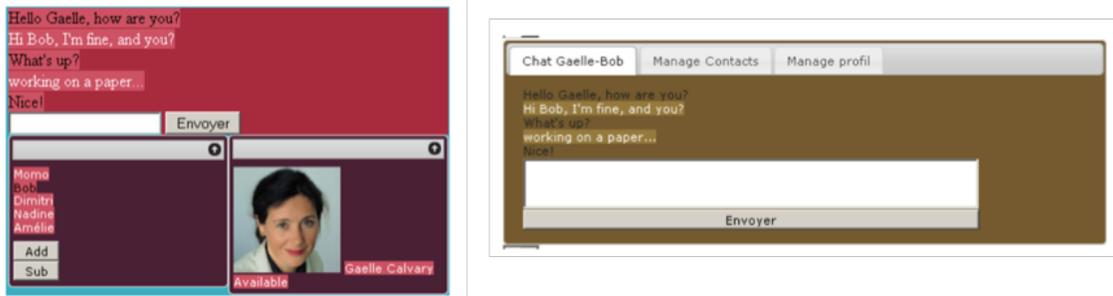


FIGURE 2.12 – Exemples d’IHM générées par Magellan (Masson et al. 2010)

tants. Elle a montré que les galeries d’exemples présentées inspirent les concepteurs et les aident à générer de nouvelles IHM. Les IHM proposées par Magellan présentent un référentiel de solutions possibles à un problème donné. En outre, les utilisateurs peuvent discuter de leurs productions dans le but de sélectionner les pistes les plus prometteuses ou d’engager un travail collaboratif.

Magellan permet la génération de différentes IHM. Cependant il ne traite pas l’aspect interactif. En effet, dans ce travail, le modèle de tâches utilisé n’est pas relié à un noyau fonctionnel (Masson et al. 2010).

5 Synthèse et conclusion

Plusieurs approches et outils ont été proposés dans la littérature pour soutenir le processus d’exploration d’alternatives. Cependant, les paramètres de génération et de sélection de ces alternatives varient d’un travail à l’autre. Ils peuvent se baser sur différents critères tels que des changements au niveau de la mise en page, des changements au niveau du style de présentation ou des présentations différentes d’un modèle de tâches. Par exemple, Lee et al. proposent un outil qui permet aux utilisateurs d’explorer des alternatives en se basant sur des critères de recherches stylistiques. DesignScape fournit aux utilisateurs différentes suggestions d’alternatives qui se basent sur des transformations de mise en page (layout) d’une IHM fournie par le concepteur. Masson et al. proposent un outil qui génère une multitude d’exemples d’IHM en se basant sur un modèle de tâche donné. Enfin, il est important de mentionner que la génération des alternatives, en plus du fait qu’elle aide les concepteurs à s’inspirer et favorise leur créativité, fournit également la possibilité de rejeter des idées en les voyant se concrétiser dans les exemples.

Une limite que nous avons observée est que les travaux présentés ne traitent pas l’aspect interactif de l’IHM (la sémantique ou la dynamique des interactions) : ils se limitent à l’interface.

Discussion sur l'état de l'art

Trouver la bonne interface et bien réussir l'interface (*getting the right design and getting the design right*) (Buxton 2010) sont deux défis complémentaires dans le domaine de l'interaction homme-machine.

La littérature montre qu'exposer simultanément des utilisateurs à des alternatives d'IHM lors d'évaluations expérimentales est bénéfique pour répondre au premier défi, c'est-à-dire trouver les bons principes de conception. En effet, la comparaison d'alternatives augmente la quantité et la qualité des critiques et suggestions d'amélioration. De plus, en présentant des alternatives, on permet aux utilisateurs d'exprimer quelle est leur IHM favorite et de rejeter celles qu'ils n'apprécieraient pas. Ces retours utilisateur sont donc cruciaux dans une démarche d'efficacité : ils permettent de diminuer le nombre d'itérations en conception et de converger plus rapidement vers la bonne IHM. Cependant, la production des alternatives est coûteuse (en termes de ressources matérielles, de moyens humains etc.) en l'absence de méthodes et d'outils pour leur choix et/ou génération. En particulier, rien ne cadre le choix d'une *bonne* ou *mauvaise* alternative pour optimiser les bienfaits de la comparaison.

Notre objectif dans cette thèse consiste à relever cette première limite de l'état de l'art : trouver le bon compromis entre les résultats bénéfiques des évaluations comparatives et le temps requis à la production des alternatives. Au regard de cet objectif, nous étudions dans le chapitre suivant (Chapitre 3) le(s) type(s) de transformation à appliquer à une IHM de référence pour pouvoir produire la meilleure IHM alternative possible.

Notre analyse de l'état de l'art montre aussi que les alternatives, aussi variées soient-elles, représentent des IHM mono-pages et sont toujours présentées par un même type de prototype. Aussi, nous relevons cette deuxième limite de l'état de l'art et étudions, dans un premier temps, si l'utilisation de différents types de prototypes (papier et vidéos) peut présenter une alternative intéressante à la génération d'IHM pour l'évaluation comparative.

Dans un deuxième temps, nous relevons une troisième limite de l'état de l'art : une focalisation sur l'interface sans considération de l'interaction. Nous nous intéressons alors à

l’IDM pour étudier globalement les différentes transformations possibles d’un modèle initial. En effet, nous avons constaté que différents travaux se sont intéressés à l’IDM pour la génération des alternatives dans une finalité d’aider les concepteurs à mieux explorer l’espace de conception. Toutefois, l’utilisation de ce type de transformations à finalité d’évaluation comparative n’a pas été étudié. L’avantage de cette approche est qu’elle permet une génération de différentes alternatives d’une IHM à différents niveaux d’abstraction, en couvrant la préoccupation de l’interaction.

Dans la première étude (Chapitre 3), nous examinons l’impact des différents types de transformations sur les résultats des évaluations des IHM dans le but de déterminer quel serait le plus bénéfique pour les évaluations comparatives. L’étude oriente nos principes de génération automatique des alternatives, méthode que nous détaillons au dernier chapitre de ce manuscrit (Chapitre 4).

Deuxième partie

Contributions

3

Quelle alternative pour les évaluations comparatives ?

1 Introduction

Définir les caractéristiques de la propriété « Meilleure IHM alternative à présenter lors d'évaluations comparatives » est notre première contribution. L'identification des caractéristiques de cette IHM permettra ensuite de définir les transformations à appliquer à une IHM initiale et de générer automatiquement cette « meilleure » alternative. Nous présentons dans ce chapitre notre démarche de travail pour arriver à ce but. Nous présentons ensuite trois études expérimentales et concluons.

2 Démarche de travail

Nous défendons dans cette thèse que présenter la bonne IHM alternative en face d'une IHM de référence optimise le processus de conception. Dans ce contexte, nous nous intéressons aux transformations à appliquer à l'IHM de référence pour la production d'alternatives. Nous identifions deux classes :

- Un changement de support présentant l'IHM : par exemple, passer d'un prototype papier à une vidéo.
- Un changement d'IHM nécessitant dès lors une transformation de l'IHM.

2.1 Changement de support

Dans notre étude de l'état de l'art, nous avons constaté que les travaux sur les évaluations comparatives étaient principalement basés sur des prototypes statiques. L'interaction, qui représente pourtant un aspect essentiel de l'IHM, a été peu explorée. Ce constat nous a motivées à étudier en premier lieu le changement de support de l'IHM de référence sans la modifier elle-même. Nous étudions donc si un tel changement pourrait suffire pour optimiser les résultats des évaluations.

Nous avons mené une étude expérimentale (étude 3) dans laquelle nous examinons ce type de transformation. Nous nous intéressons à l'utilisation simultanée d'un prototype papier et d'un prototype vidéo présentant la même IHM respectivement de manière statique (interface) et dynamique (interaction).

2.2 Changement d'IHM

Dans notre étude de l'état de l'art, nous avons constaté que les transformations d'IHM peuvent s'opérer à différents niveaux d'abstraction : les tâches utilisateur, l'IUA et l'IUC. Nous cherchons ici à découvrir le niveau le plus judicieux pour optimiser la génération de l'IHM alternative.

Pour y répondre, nous avons conduit une étude expérimentale (étude 4) dans laquelle nous examinons l'impact de la transformation sur les évaluations comparatives. Nous montrons que l'IUA est le niveau le plus bénéfique.

Ensuite, en nous basant sur les résultats de cette étude, nous avons réalisé une troisième étude expérimentale (étude 5) dans laquelle nous nous intéressons au type de transformation à appliquer au niveau de l'IUA. Cette étude a pour objectif de préciser le type de transformation à appliquer à ce niveau d'abstraction afin de générer automatiquement la meilleure IHM alternative.

Nous détaillons dans les sections suivantes les protocoles expérimentaux de chaque étude ainsi que leurs résultats. Nous concluons par une synthèse et discussion.

3 Étude #1 : changement de support

Cette première expérimentation a été menée dans le but d'étudier l'impact du changement de support sur l'évaluation des IHM. Nous avons considéré deux supports différents : d'une part, un prototype papier statique et, d'autre part, un prototype vidéo dynamique. À travers cette étude, nous explorons les résultats de l'évaluation en utilisant ces deux supports simultanément, mais dans un ordre de présentation différent dans deux groupes de testeurs. Ainsi, nous avons émis les deux hypothèses suivantes :

- **H1** : voir une IHM sous différents supports peut impacter la quantité et la qualité des retours utilisateur.
- **H2** : l'ordre de présentation des supports lors de l'évaluation a une influence en termes de la qualité des retours utilisateurs.

3.1 Cas d'étude

Pour cette étude, nous avons choisi de travailler sur une IHM d'un site web de commerce électronique spécialisé dans la vente de produits high-tech, tels que des smartphones, des télévisions, des ordinateurs, etc. La raison derrière ce choix est que la plateforme de e-commerce est très répandue de nos jours. Il s'agit alors d'un domaine facile à appréhender par le plus grand nombre de participants, sans nécessité de mise en contexte particulière.

3.2 Élaboration des prototypes

Pour l'élaboration des prototypes, nous avons utilisé l'outil de prototypage *Balsamiq Wireframes* (*Balsamiq Wireframes, Quick and Easy Wireframing Tool Retrieved : December, 2019*). Nous avons dans un premier temps conçu le prototype statique avec un niveau de fidélité moyen. Il couvre les différentes pages et fonctionnalités du site web : page d'accueil,

menus et sous-menus pour présenter les différents produits, en obtenir des détails, les ajouter au panier, gérer les produits sélectionnés, etc. (Figure 3.1). Ce premier prototype a été imprimé sur papier : une feuille papier représentant chaque page du site web.

Le prototype dynamique a été réalisé en enregistrant des interactions à travers le même prototype statique via Balsamiq. En effet, l'outil permet une navigation entre les différentes pages conçues, nous avons donc affecté des liens aux différents éléments de l'interface pour assurer la navigation entre les pages. La vidéo a été enregistrée en utilisant un outil de capture d'écran vidéo. Elle dure 42 secondes et illustre le scénario suivant : l'utilisateur commence par naviguer sur le site web et se concentre en particulier sur le menu permettant d'explorer les différentes catégories de produits disponibles sur le site. Ensuite, il sélectionne un article (un smartphone) et l'ajoute au panier. Puis, l'utilisateur change d'avis après avoir consulté le panier. Il supprime le smartphone sélectionné du panier, recherche un autre article et l'ajoute au panier. L'ensemble du processus est filmé dans la vidéo et présenté aux utilisateurs (Figure 3.2).

3.3 Participants

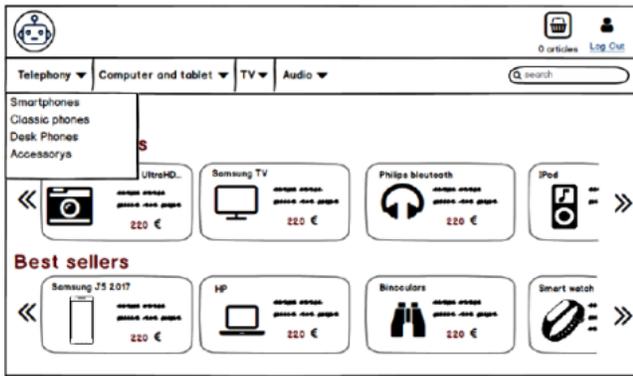
Au total, 22 personnes ont participé à cette étude. Les participants étaient composés de 11 femmes et de 11 hommes âgés de 22 à 58 ans. L'étude a impliqué aussi bien des personnes familiarisées avec le domaine de l'IHM que des personnes n'ayant aucune connaissance dans le domaine. Par conséquent, le panel de participants était varié. Il comprenait : des chercheurs, des ingénieurs et des étudiants (master et doctorat) dans différents domaines ainsi que des membres du personnel administratif de l'université.

3.4 Protocole expérimental

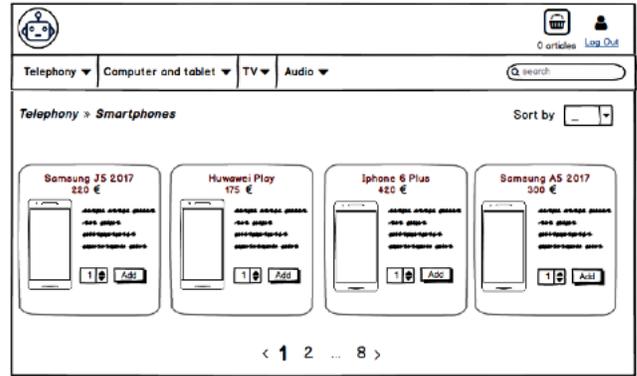
Afin de vérifier les hypothèses formulées ci-dessus, nous avons élaboré deux situations d'évaluation. Chaque situation se déroulait en deux étapes. Dans la première situation d'évaluation, les participants commencent par évaluer le prototype statique puis évaluent le prototype dynamique. Dans la deuxième situation d'évaluation, les participants commencent par évaluer le prototype dynamique puis évaluent le prototype statique. Les participants ont été divisés en deux groupes, chaque groupe ayant testé une situation d'évaluation. Chaque session commençait par présenter brièvement l'étude au participant et le prototype à évaluer.

— Situation d'évaluation 1 : statique puis dynamique :

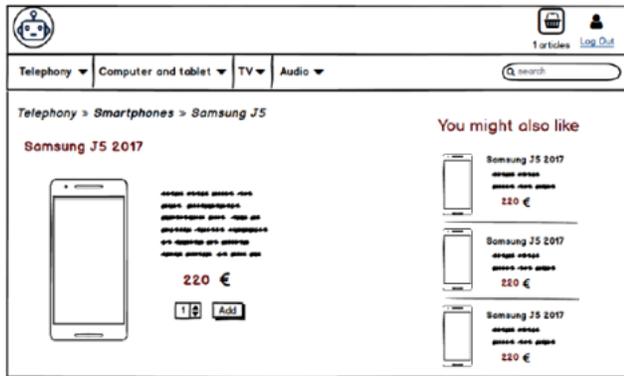
Dans cette situation d'évaluation, nous commençons par présenter le prototype statique au participant en première étape. Le participant observe le prototype sans que nous lui imposions de limite de temps. Une fois terminé, nous lui demandons son avis sur



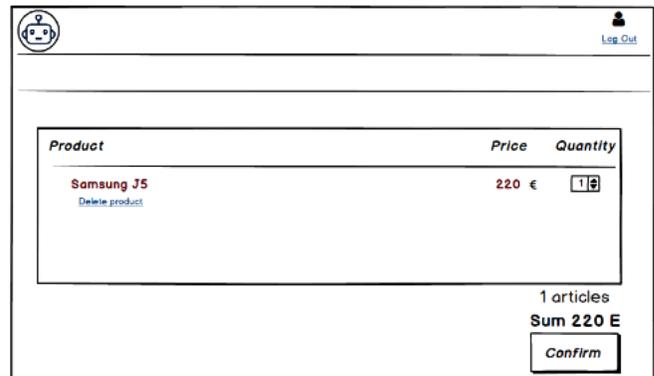
Page d'accueil



Page catégorie smartphone



Page d'un produit sélectionné



Page panier

FIGURE 3.1 – Prototype statique du site e-commerce



FIGURE 3.2 – Prototype dynamique du site e-commerce

l'interface : ce qu'il/elle aime, n'aime pas, changerait/suggère pour améliorer l'IHM. Cette catégorisation est inspirée de la taxonomie de Tohidi et al. (Tohidi et al. 2006b).

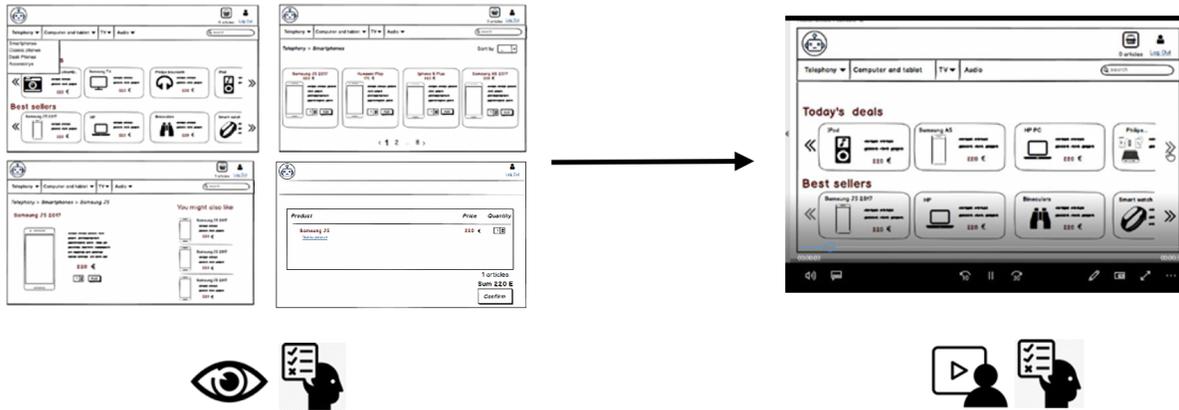


FIGURE 3.3 – Situation d'évaluation 1

Une fois la première étape terminée, nous remplaçons le prototype statique par le prototype dynamique. La vidéo est jouée sur un ordinateur portable. S'il le souhaite, le participant peut la regarder jusqu'à trois fois au maximum. Nous avons ensuite demandé aux participants s'ils avaient quelque chose à ajouter ou à modifier par rapport à leur première évaluation de l'IHM (dans la première étape).

— Situation d'évaluation 2 : Dynamique puis statique

Cette situation d'évaluation est similaire à la première avec une seule différence principale : le prototype dynamique, c'est-à-dire ici la vidéo, est présenté au participant en première étape, puis le prototype statique.

3.5 Collecte des données

Lors de l'évaluation, chaque participant a été invité à présenter ses retours en les inscrivant sur un tableau blanc. Sur ce tableau nous avons placé trois cercles représentant chacun une des catégories suivantes : ce que les participants aiment ; ce que les participants n'aiment pas et ce qu'ils proposent d'améliorer (suggestions d'amélioration). Ensuite, des post-it ont été placés sur le tableau grâce à un aimant pour que les participants écrivent leurs retours dessus. Les participants placent alors les post-it dans le cercle approprié. Enfin, deux couleurs de post-it ont été utilisées pour distinguer chaque étape de l'évaluation.

Cette technique de travail nous a permis de collecter des données qualitatives et quantitatives. En outre, elle a été opérationnalisée conformément à la taxonomie proposée par (Tohidi et al. 2006b).

Notons que la majorité des participants ont expliqué ce qu'ils écrivaient et exprimaient leurs pensées oralement. Les observations ont été enregistrées en prenant des notes tout au long de la session de l'évaluation. Ainsi, nous avons enregistré les sessions d'évaluation via un dictaphone.



FIGURE 3.4 – Un participant qui saisit ses retours

3.6 Résultats

Afin d'évaluer l'impact de la présentation de l'IHM avec différents supports (prototype papier et prototype vidéo) sur les résultats de l'évaluation, nous avons analysé la quantité et la qualité des retours collectés pour chaque étape de chaque situation d'évaluation. Cette analyse nous permet d'évaluer la volonté critique des utilisateurs et la pertinence de leurs retours par rapport à chaque situation d'évaluation afin de valider les hypothèses posées.

3.6.1 Quantité des retours utilisateur

Comme indiqué au tableau 3.1, le nombre total des retours utilisateur était de 211 retours qui se composent de 106 retours dans la première situation d'évaluation, et 105 retours dans la deuxième. Nous remarquons que le nombre de retours collectés lors de chaque situation d'évaluation est similaire.

Nous avons par la suite analysé le nombre de retours utilisateur pour chaque étape dans chaque situation d'évaluation. Cette analyse indique que pour les deux situations d'évaluation, le nombre de retours utilisateur obtenus lors de la première étape est nettement supérieur au nombre de retours utilisateur obtenus lors de la deuxième étape : 72,64% pour le prototype statique dans la première situation et 75,96% pour le prototype dynamique dans la deuxième situation.

Ces premiers résultats indiquent que (1) le nombre total de retours collecté est indépendant de l'ordre de présentation des prototypes, et (2) quelque soit le type de prototype présenté en premier, celui-ci générera un plus grand nombre de retours utilisateur (environ les 3/4 des commentaires sont donnés après le premier prototype, contre seulement 1/4 pour le deuxième).

	Étape 1	Étape 2	Nombre total des retours utilisateur
Situation d'évaluation 1	77	29	106
Situation d'évaluation 2	79	26	105

TABLE 3.1. Nombre de retours utilisateur pour chaque situation d'évaluation- Étude #1.

3.6.2 Qualité des retours utilisateur

Dans un premier temps, nous avons commencé par étudier l'impact de chaque situation d'évaluation sur le type des retours utilisateur fournis. Le type des retours se rapporte aux retours positifs (ce que les testeurs aiment), aux retours négatifs (ce que les testeurs n'aiment pas) et aux suggestions d'amélioration. Nous avons ensuite procédé à une classification des retours utilisateurs selon leur qualité, c'est-à-dire selon les trois aspects suivants : utilité, utilisabilité et esthétique. Les retours sur l'utilité concernent les fonctionnalités offertes par le site web. Les retours sur l'utilisabilité englobent la facilité de la navigation dans le site web, les modalités d'interaction, le guidage et la charge de travail. Le guidage couvre les quatre sous-critères : incitation, groupement/distinction entre items, feedback immédiat et lisibilité, tel que défini par C. Bastien et D. Scapin (Bastien et Scapin 1992). La charge de travail couvre les deux sous-critères brièveté et densité informationnelle. Enfin les retours sur l'esthétique de l'IHM concernent l'aspect visuel et esthétique générale de l'interface tel que les couleurs utilisées, le choix des différentes icônes ou le logo du site web.

Type des retours utilisateurs

La figure 3.5 présente le nombre de retours utilisateur par rapport à leur type (positif, négatif, suggestion). Nous exposons dans ce qui suit quelques déclarations de participants illustrant ces types de retours :

- *Exemples de retours positifs* : la possibilité d'accéder aux meilleures ventes sur la page d'accueil et la simplicité de l'interface utilisateur.
- *Exemples de retours négatifs* : manque de couleurs, peu de produits affichés, pas de menu dans le panier, on ne peut pas ajouter un article dans le panier depuis la page d'accueil, le logo n'est pas joli.
- *Exemples de suggestions* : ajout d'une fonction de recherche avancée (plus détaillée), changer le défilement horizontal en vertical.

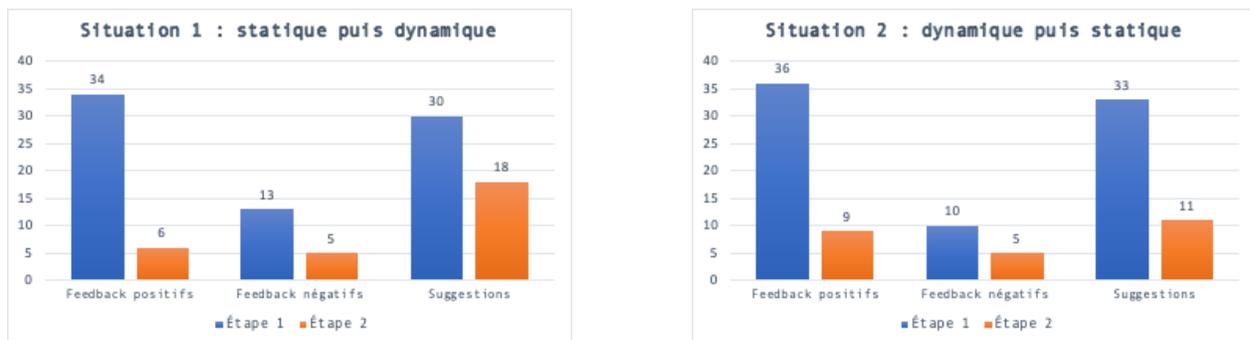


FIGURE 3.5 – Nombre de retours utilisateur par rapport à leur type

Nous pouvons effectuer les observations suivantes :

- Le nombre de commentaires négatifs ajouté au nombre de suggestions est nettement supérieur au nombre de commentaires positifs dans les deux situations d'évaluation. En conséquence, si les résultats de cette étude ont mis en évidence les aspects positifs, ils ont encore plus permis de révéler des éléments d'amélioration. Nous pouvons également en déduire que la réévaluation de l'IHM, sous une autre forme, pousse le testeur à être plus critique et à fournir plus de suggestions d'amélioration.
- Le nombre de suggestions collectés dans la deuxième étape de la première situation d'évaluation est plus important que dans la deuxième situation.
- Enfin, nous avons remarqué qu'il y a plus de suggestions superficielles que de suggestions substantielles et que les suggestions proposées sont limitées à l'IHM proposée sans suggestions alternatives. Dans ce cadre, une suggestion substantielle représente des nouvelles idées et des suggestions qui portent essentiellement sur l'amélioration de l'interaction et de l'utilisabilité du site (par exemple, ajouter un message de confirmation pour l'ajout d'un produit dans le panier), tandis que les suggestions superficielles représentent des commentaires peu profonds et frivoles (par exemple, mettre des prix plus réalistes ou ajouter plus de couleurs).

Qualité des retours utilisateur

Nous analysons dans cette section les retours des utilisateurs par rapport à l'utilité, l'utilisabilité et l'esthétique de l'IHM.

Comme indiqué dans la figure 3.6, il y a une différence significative entre les deux situations d'évaluation par rapport aux aspects définis ci-dessus. En effet, lors de la première étape de l'évaluation de la première situation (statique puis dynamique) le nombre de retours par rapport à l'utilité est le plus grand par rapport à l'utilisabilité et l'esthétique. Tandis que dans la deuxième situation (dynamique puis statique), c'est l'utilisabilité qui a régné en nombre de retours par rapport à l'utilité et l'esthétique.

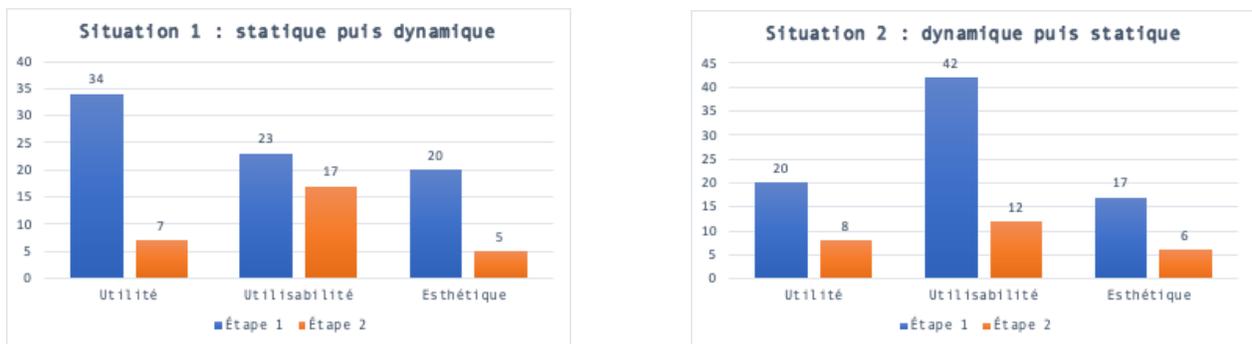


FIGURE 3.6 – Nombre de retours utilisateur par rapport à leur qualité

Nous concluons que :

- L'utilisation du prototype statique a permis aux utilisateurs de se concentrer sur les fonctionnalités offertes par le système : ils ont pris le temps de les explorer, ce qui a permis de développer une compréhension du noyau fonctionnel. Par conséquent, ils ont critiqué principalement l'utilité.
- L'utilisation du prototype vidéo a permis aux utilisateurs de se concentrer davantage sur l'interaction avec le système, c'est-à-dire sur la manière dont un utilisateur peut effectuer une tâche spécifique. En observant le système en action, les utilisateurs ont développé une compréhension des aspects interactifs du système et se sont concentrés sur l'utilisabilité.
- Enfin, en ce qui concerne l'esthétique, le nombre de retours utilisateur est à peu près le même dans la première étape des deux situations d'évaluation. Il convient de noter que le nombre de retours est légèrement plus élevé avec le prototype statique lors de la première situation d'évaluation que le nombre de retours collectés avec la vidéo lors de la deuxième situation d'évaluation.

Les figures 3.7 et 3.8 résument les résultats de l'évaluation dans les deux situations d'évaluation par rapport aux deux classifications : type et qualité des retours utilisateur collectés.

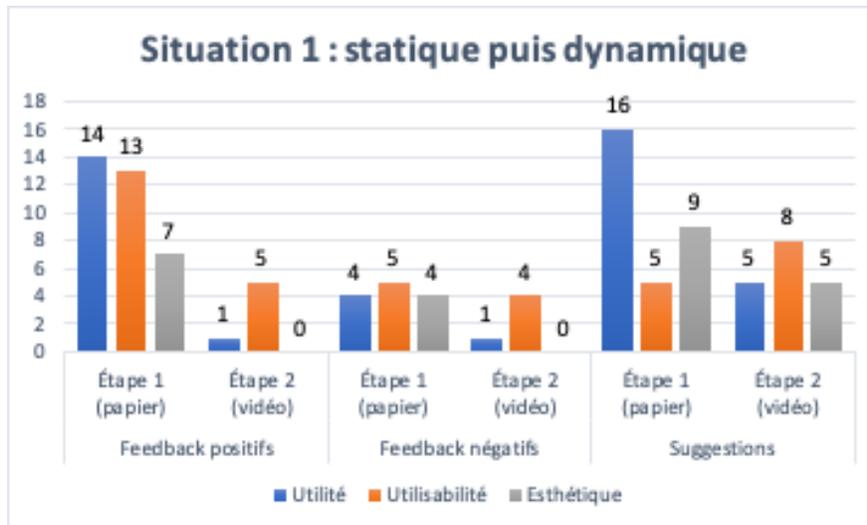


FIGURE 3.7 – Résumé des résultats de la première étude - Situation d'évaluation 1

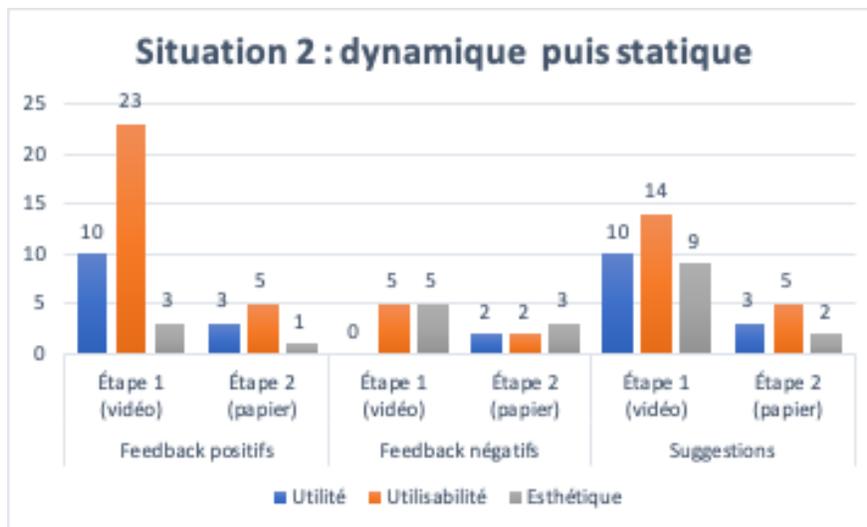


FIGURE 3.8 – Résumé des résultats de la première étude - Situation d'évaluation 2

Nous remarquons que le nombre de retours utilisateur par rapport à chaque classification est différent dans les deux situations d'évaluation. Dans la première situation, nous remarquons que le nombre de retours par rapport à la qualité est plus homogène que dans la deuxième situation d'évaluation, où le nombre de retours utilisateur par rapport à l'utilisabilité est plus élevé (51%) (figure 3.9) que le nombre de retours par rapport à l'utilité et l'esthétique (Hammami et al. 2020).

Nous concluons que la présentation de la vidéo en premier lieu lors de l'évaluation pousse

les utilisateurs à plus se concentrer sur l'aspect interactif du système que sur les autres aspects, contrairement à la première situation d'évaluation où les utilisateurs ont exploité tous les aspects tout au long de l'évaluation.

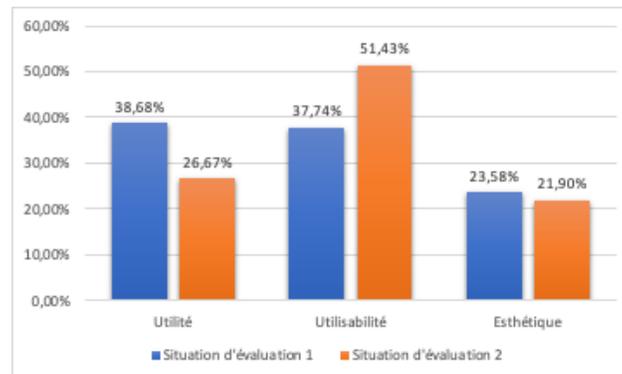


FIGURE 3.9 – Taux de retours utilisateur par catégorie pour chaque situation d'évaluation

3.6.3 Temps d'évaluation

Pour finir, nous avons examiné les résultats par rapport au temps mis pour chaque situation d'évaluation. Nous avons constaté que la durée de l'évaluation était similaire quelle que soit la situation d'évaluation. Présentés à la première ou à la deuxième étape, les participants ont passé le même temps à regarder les prototypes statiques et ont visionné la vidéo deux fois en moyenne dans les deux situations d'évaluation. Il n'y a donc pas un gain significatif en termes de temps d'évaluation entre les deux situations d'évaluation.

3.7 Synthèse de l'étude #1

Le but de cette étude était de répondre à la question suivante : est-ce que nous pouvons nous contenter d'un changement de support pour l'évaluation comparative, c'est-à-dire utiliser la même IHM mais représentée par deux supports différents ? Nous avons présenté, lors d'une évaluation expérimentale, un prototype papier et un autre sous forme d'une vidéo successivement aux participants, dans deux ordres différents. Les résultats de cette étude montrent que l'utilisation combinée des deux prototypes, papier et vidéo, enrichit les résultats et est bien adaptée pour couvrir tous les aspects de l'IHM (utilité, utilisabilité et esthétique). Par ailleurs, nous avons observé que le type de prototype présenté impacte significativement les résultats de l'évaluation d'une IHM. Cette différence a été manifestée en termes de la qualité des retours utilisateur, mais pas en terme de la quantité des retours. Ces résultats confirment en partie notre première hypothèse : "voir une IHM sous différents supports peut

impacter la quantité et la qualité des retours utilisateur." En effet, le prototype statique produit plus de commentaires liés à l'utilité, alors que la vidéo produit plutôt des commentaires relatifs à l'utilisabilité. Toutefois, le type de prototype n'a pas d'impact significatif sur le nombre de commentaires liés à l'esthétique du design. Enfin, en ce qui concerne l'ordre de présentation des prototypes, nous avons constaté que le fait de présenter le prototype papier en premier suivi de la vidéo peut être plus bénéfique en termes de la qualité des retours utilisateurs : dans cette situation d'évaluation les retours ne sont pas concentrés sur un seul aspect de l'IHM mais couvrent des commentaires par rapport à son utilité, son utilisabilité ainsi que son esthétique de manière équitable. Ce résultat confirme notre deuxième hypothèse : "l'ordre de présentation des prototypes lors de l'évaluation a une influence en termes de la qualité des retours utilisateurs".

— *Les résultats de cette étude ont été présentés dans un papier lors de la conférence Advances in Computer-Human Interactions en 2020 (Hammami et al. 2020).*

Le changement de support n'étant pas suffisant, nous étudions ci-après les changements d'IHM.

4 Étude #2 : changement d'IHM

Nous nous intéressons dans cette étude expérimentale à l'impact de la transformation d'IHM sur l'évaluation comparative. Nous cherchons à déterminer quel type de transformation est le plus adapté pour la génération d'IHM alternatives afin de maximiser l'efficacité des évaluations comparatives.

Nous nous sommes basés sur le cadre de référence CAMELEON pour structurer la production des alternatives. Ainsi, nous avons classé les types de transformations par rapport au (1) niveau tâches et concepts qui décrit la sémantique de l'application du point de vue de l'utilisateur, (2) le niveau de l'Interface Utilisateur Abstraite (IUA) qui modélise l'IHM d'une manière abstraite et précise le dialogue entre l'utilisateur et le système, et (3) le niveau de l'Interface Utilisateur Concrète (IUC) qui modélise l'IHM en termes d'interacteurs concrets (boutons, listes, etc.).

Les transformations appliquées représentent des changements majeurs et mineurs au regard de leur profondeur dans le cadre de référence CAMELEON : les transformations aux niveaux tâches-concepts et IUA sont dites majeures ; les transformations au niveau de l'IUC sont qualifiées de mineures.

Notre démarche pour cette étude se présente comme suit : nous commençons par créer des IHM alternatives en partant d'une IHM de référence selon les trois classes de transformations définies précédemment. Ensuite nous utilisons ces alternatives en évaluation comparative. Notre point de départ est un modèle de tâches que nous avons élaboré comme cas d'étude.

4.1 Cas d'étude

Pour cette étude, nous avons choisi un cas d'étude d'un système de gestion de sécurité d'une maison à distance permettant à un utilisateur peut gérer : l'accès à sa maison, les caméras de surveillance et le système d'alarme. L'utilisateur peut verrouiller ou déverrouiller les entrées/sorties de la maison. Il/elle peut accéder aux vidéos de surveillance, les regarder en direct, supprimer des vidéos d'une date précise ou regarder des anciennes vidéos en précisant la date et l'heure de la séquence à voir. Enfin, l'utilisateur peut activer, désactiver le système de d'alarme, ou l'arrêter en cas de déclenchement.

Ce cas d'étude est un exemple pédagogique simplifié que nous avons élaboré afin d'étudier l'impact des transformations au niveau des différents modèles décrits.

4.2 Élaboration des prototypes

Nous avons commencé par créer le modèle de tâches correspondant à notre cas d'étude, puis généré des IUA et IUC correspondantes. Par la suite, nous avons généré des alternatives de ces modèles comme indiqué à la figure 3.10. Au final, nous avons quatre alternatives d'une IHM, que nous avons dessinées à l'aide de l'outil Balsamiq pour produire les prototypes à évaluer. Nous avons défini l'IHM concrète 1 comme IHM de référence.

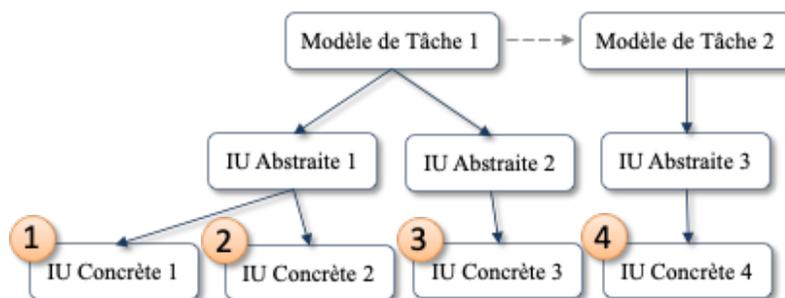


FIGURE 3.10 – Élaboration des prototypes

Le modèle de tâche est présenté dans la figure 3.11). À partir de ce modèle, nous avons généré manuellement deux modèles différents de l'interface utilisateur abstraite. La différence entre les deux se manifeste en termes de navigation et de groupements des tâches (figures 3.12 et 3.15). Enfin, nous avons créé manuellement des IUC correspondant à ces IUA.

Les deux figures 3.12 et 3.13 présentent respectivement les IUA1 et IUC1.

La figure 3.14 présente IUC2. L'IUC1 et l'IUC2 diffèrent en termes de types et de paramètres de widgets utilisés. Par exemple, les boutons radio permettant la gestion des accès ont été remplacés par des boutons de type *switch*. Dans ce cas le modèle de tâche ainsi que l'IUA ne sont pas modifiés.

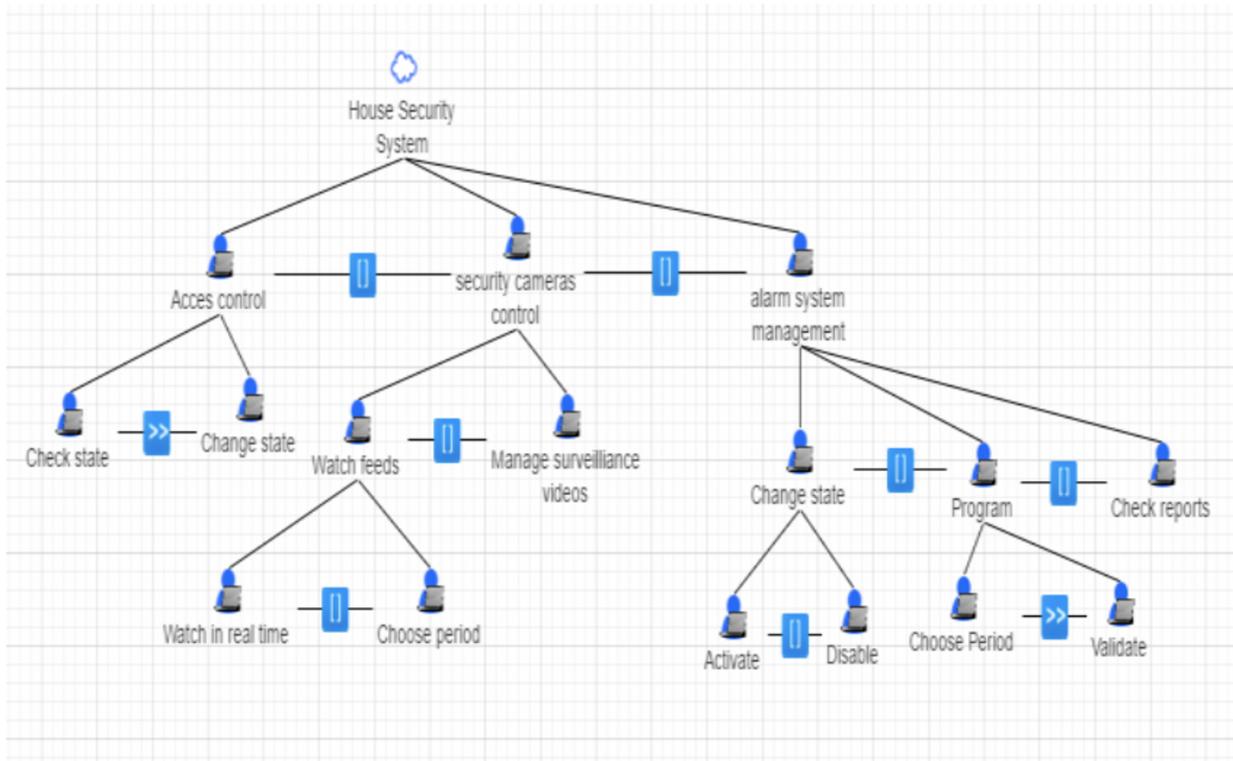


FIGURE 3.11 – Modèle de tâche 1

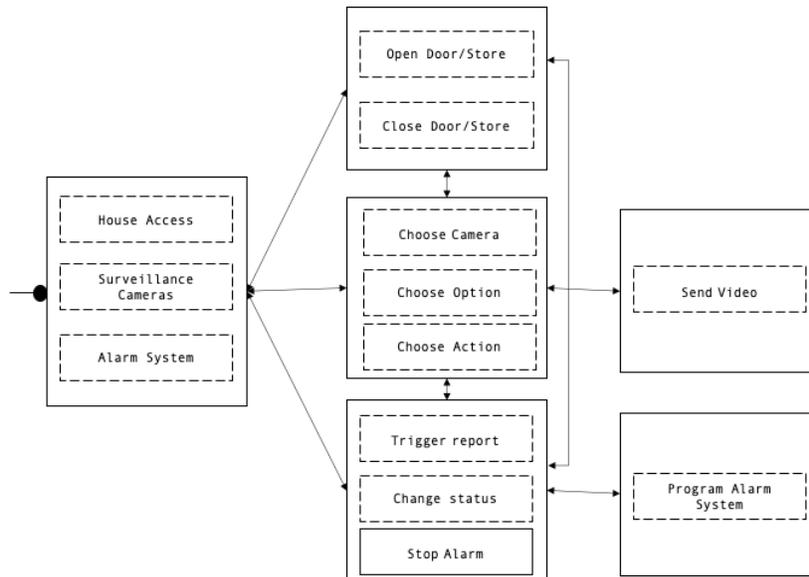


FIGURE 3.12 – IU Abstraite 1

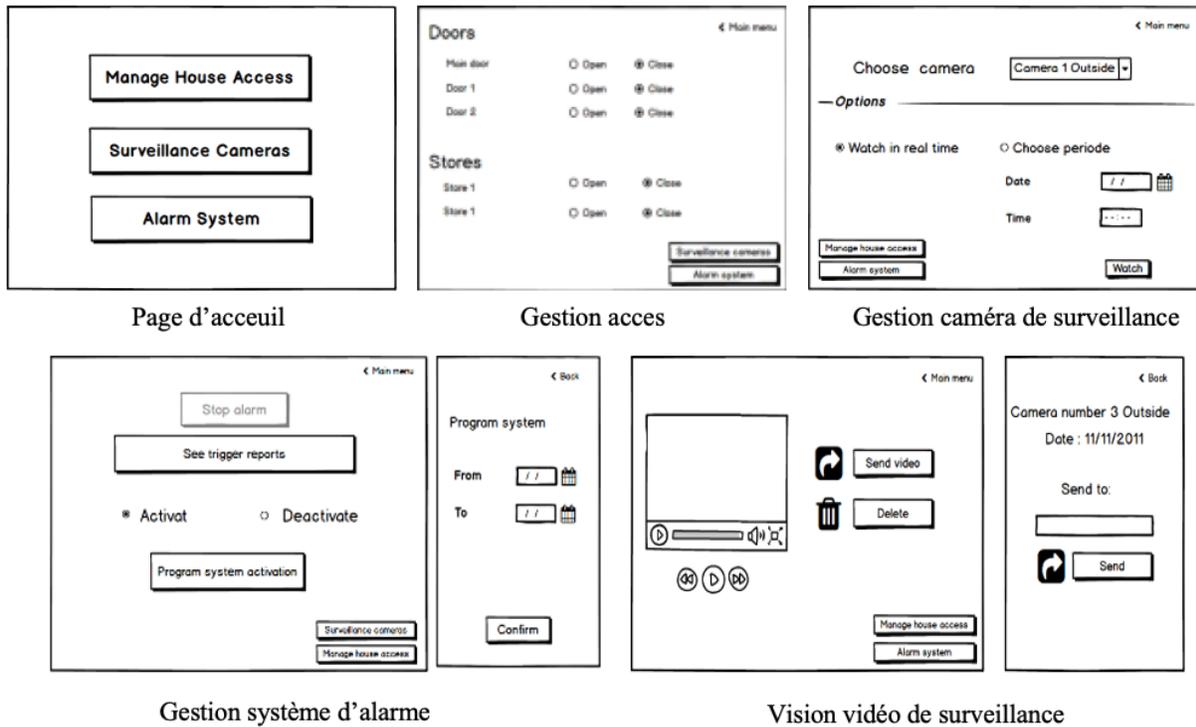


FIGURE 3.13 – IU Concrète 1

La troisième IUC (figure 3.16) a été conçue en utilisant le même style de présentation que l’IUC1 afin d’évaluer l’impact de la transformation au niveau de l’IUA, indépendamment de l’aspect concret de l’interface.

Enfin, nous avons généré une variation du premier modèle de tâches afin de produire un deuxième modèle de tâches (Figure 3.17). Les différences entre les deux sont des transformations de structuration du modèle et d’opérateurs utilisés entre les tâches. Par exemple, pour la gestion de la caméra de surveillance, l’utilisateur doit choisir la caméra, puis choisir une option., alors que dans le premier modèle, l’utilisateur peut directement choisir l’option souhaitée avant de choisir la caméra.

4.3 Participants

Nous avons recruté un total de 28 personnes pour cette étude. L’étude a impliqué à la fois des personnes familiarisées avec le domaine de l’IHM ainsi que des personnes n’ayant aucune connaissance du domaine, comme par exemple des étudiants en master et doctorat de différents domaines.

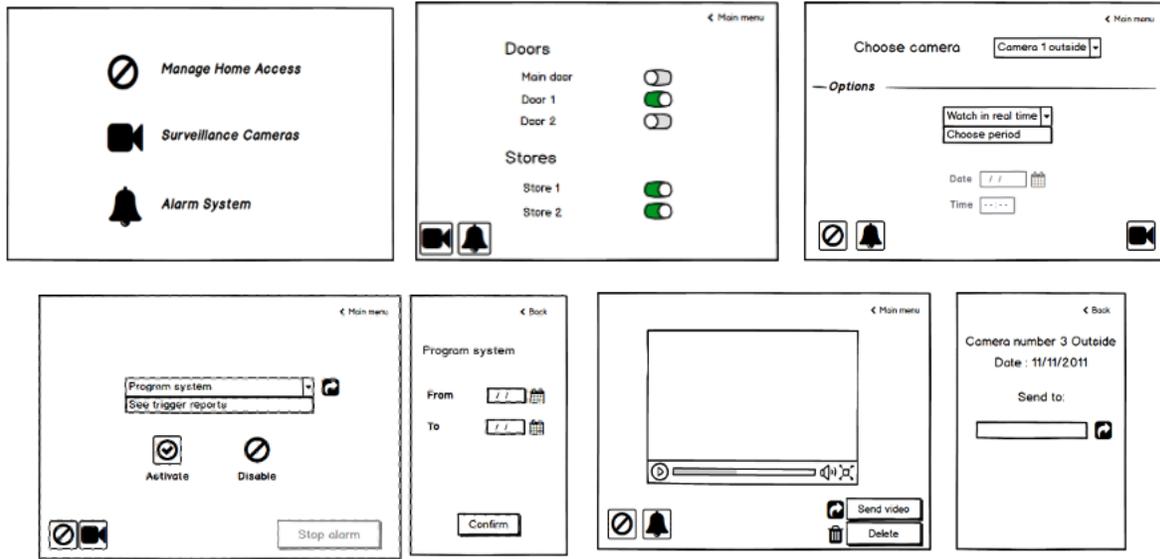


FIGURE 3.14 – IU Concrète 2

4.4 Protocole expérimental

Les participants ont été divisés en trois groupes. Nous leur avons présenté à chaque fois deux IHM simultanément : l'IHM de référence (IU Concrète 1) et une des alternatives. Un premier groupe (situation d'évaluation 1) a évalué IU Concrète 1 avec IU Concrète 2, un deuxième groupe (situation d'évaluation 2) a évalué IU Concrète 1 avec IU Concrète 3 et un troisième groupe (situation d'évaluation 3) a évalué IU Concrète 1 avec IU Concrète 4.

Nous avons commencé par présenter brièvement l'étude aux participants et l'application de gestion de sécurité. Ensuite nous avons donné les prototypes des deux IHM à évaluer. Les prototypes sont présentés simultanément en format papier : une feuille papier pour chaque IHM, représentant chacune l'ensemble des pages de l'application. Les participants les observent sans que nous leur imposions une limite de temps.

Les participants ont été invités à répondre à un questionnaire pour l'évaluation de l'IHM de référence (IU Concrète 1). Ce dernier est divisé en deux parties. La première partie, destinée à noter l'IHM, est composée de neuf questions basées sur une échelle de Likert à 5 points (allant de "pas du tout d'accord" à "tout à fait d'accord"). Les questions de cette première partie portent sur quatre aspects principaux : le contenu, la structure, la navigation, et le visuel (esthétique). Dans la deuxième partie du questionnaire, les participants ont été invités à exprimer leurs avis (ce qu'ils aiment, ce qu'ils n'aiment pas et s'ils avaient des suggestions d'amélioration), au travers de questions ouvertes. Enfin, ils ont été invités à

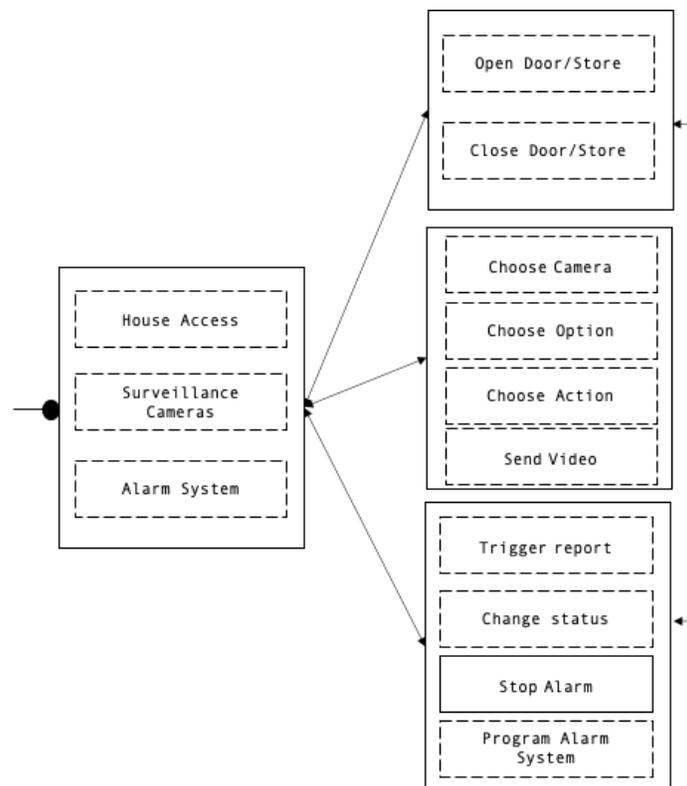


FIGURE 3.15 – IU Abstraite 2

exprimer leurs préférences entre les deux IHM présentées (l’IHM de référence et l’alternative). Ce questionnaire est présenté dans l’annexe A.

4.5 Résultats

Dans le but d’évaluer l’impact de la présentation de différents types d’alternatives en évaluation comparative sur les retours des utilisateurs, nous avons commencé par calculer le score donné à IUC 1 en nous basant sur les notes accordées par les participants dans la première partie du questionnaire. Par la suite, nous avons analysé les retours fournis lors de la deuxième partie du questionnaire, pour chaque groupe, en termes de quantité et de type.

4.5.1 Notes obtenues pour l’IHM de référence

Dans la première partie du questionnaire les participants ont répondu à neuf questions permettant de noter l’IHM de référence sur une échelle Likert de 5 points. Nous avons calculé une moyenne en s’appuyant sur ces évaluations unitaires pour calculer le score final de l’IHM de référence. Ce score définit l’appréciation globale des participants en fonction des 4 mesures : contenu, structure, navigation et esthétique.

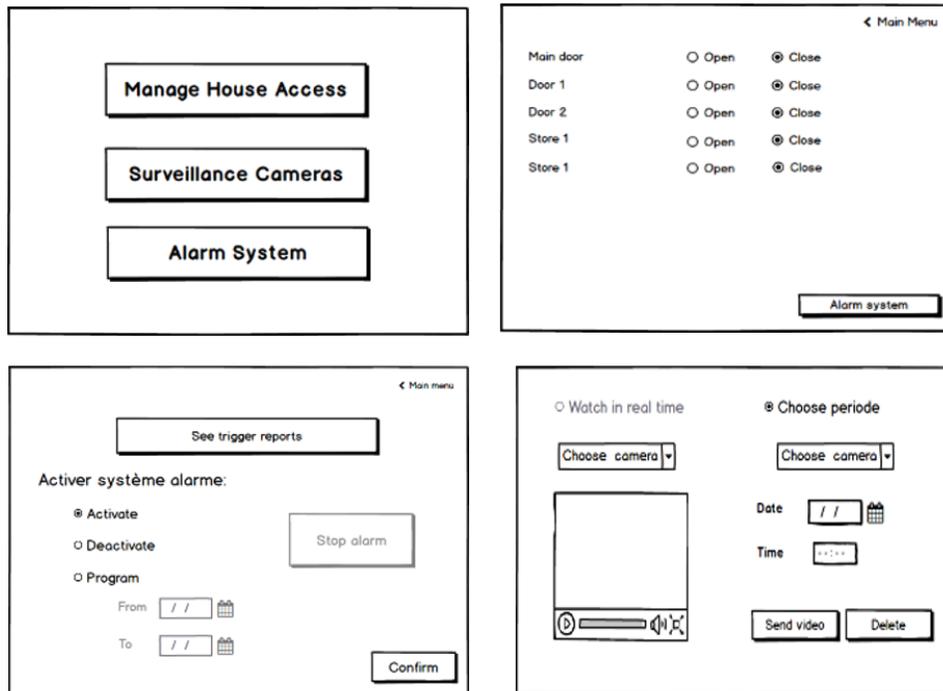


FIGURE 3.16 – IU Concrète 3

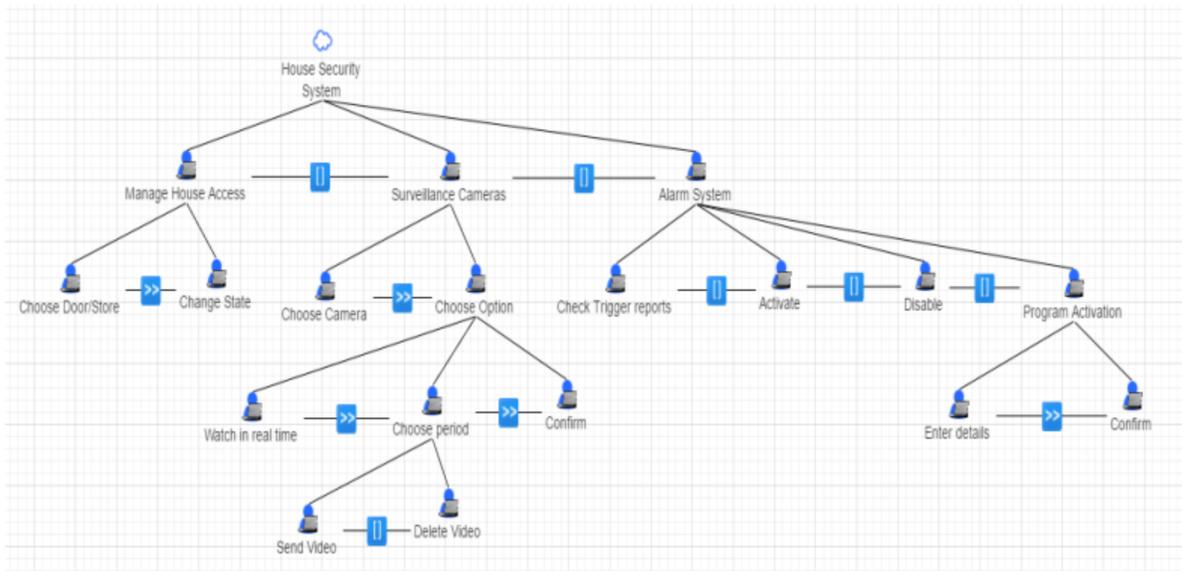


FIGURE 3.17 – Modèle de tâche 2

Comme indiqué dans la figure 3.18 le score accordé à l'IUC1 dans la deuxième situation d'évaluation (évaluer deux IHM présentant une variation au niveau du modèle de l'IUA) est le plus bas. Nous constatons qu'en comparant deux IHM ayant des IUA différentes les participants sont plus conscients de sa structure et sont plus exigeants par rapport à l'organisation des différents espaces de travail de l'IHM et le lien entre ces espaces, ce qui se traduit par des scores plus bas dans cette situation d'évaluation.

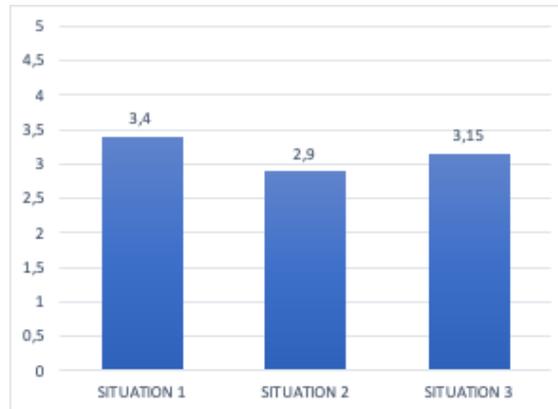


FIGURE 3.18 – Score accordée à IU Concrète 1 pour chaque situation d'évaluation

4.5.2 Quantité des retours utilisateur

Dans une deuxième étape, nous avons calculé le nombre de retours utilisateur fournis. En effet, en répondant aux questions de la deuxième partie du questionnaire, les participants ont écrit leurs critiques et avis par rapport à l'IHM de référence. Ils ont listé ce qu'ils appréciaient, ce qu'ils n'aimaient pas ou ce qui leur semblait incorrect, et enfin leurs idées et suggestions d'amélioration. Le tableau 3.2 présente le nombre de retours utilisateur obtenus pour chacune des trois situations d'évaluation.

<i>Situation d'évaluation</i>	<i>Nombre total de retours utilisateur</i>
Situation d'évaluation 1	18
Situation d'évaluation 2	29
Situation d'évaluation 3	17

TABLE 3.2 – Nombre de retours utilisateur pour chaque situation d'évaluation - Étude #2

Nous constatons que le nombre de retours utilisateur fournis lors de la deuxième situation d'évaluation est significativement plus important que le nombre de retours utilisateur fournis lors des deux autres situations d'évaluation. En effet, les participants ont été plus critiques

quand ils ont évalué deux IHM présentant des IUA différentes. Les retours ont été plus élaborés, et les commentaires négatifs ont été appuyés par des suggestions d'amélioration dans la majorité des cas. Nous concluons que fournir une IHM présentant une organisation différente des espaces de travail et de la navigation entre ces espaces peut fournir un bon support pour favoriser la créativité des évaluateurs afin qu'ils produisent un maximum de retours pertinents.

4.5.3 Type de retours utilisateur

La troisième étape de notre analyse des résultats consiste à étudier les retours utilisateurs en termes de type, et cela afin d'étudier l'impact de chaque situation d'évaluation sur la volonté de critique des participants. Pour cela, nous avons classé les retours en :

- Commentaires négatifs, par exemple : *"Il y a trop d'informations dans l'interface."*, *"Je n'aime pas les icônes choisis"*, *"Il y a beaucoup d'espace vide non exploité."* .
- Commentaires positifs, par exemple : *"Navigation facile et pratique."* , *"J'aime bien ce type de bouton"*.
- Suggestions d'amélioration, par exemple : *"Remplacer les boutons radios par une liste déroulante pour le choix de visualisation des vidéos de surveillance"*, *"Remplacer le texte par une icône"*, *"Ajouter des boutons de navigation dans toutes les pages"*, *"Ajouter une barre de menu, verticale par exemple"*.

Une première observation était que le type des retours utilisateur a été significativement impacté par le choix de l'IHM alternative présentée. Nous avons constaté que les transformations appliquées au niveau du modèle de l'IUA impactent le plus ces retours en termes de nombre et de type (Figure 3.19). En effet dans cette situation d'évaluation, le nombre de commentaires négatifs et de suggestions d'amélioration est le plus important. Dans ce contexte, les commentaires négatifs sont considérés comme des retours pertinents car ils peuvent révéler des erreurs et problèmes de conception de l'IHM.

4.6 Synthèse de l'étude #2

Dans cette étude expérimentale, nous avons étudié l'impact du choix de l'IHM alternative à présenter dans un contexte d'évaluation comparative. Pour cela, nous avons présenté à trois groupes de participants différentes alternatives d'une IHM qui varient en fonction du modèle de la tâche et concepts, du modèle abstrait de l'interface et du modèle concret de l'interface.

Nous avons observé que le choix de l'IHM alternative impactait la quantité et la qualité des retours utilisateur. En effet, le nombre de suggestions d'amélioration de l'IHM, ainsi que le niveau de précision des retours utilisateurs fournis n'est pas le même dans les différentes situations d'évaluation. Les utilisateurs ne présentent donc pas la même volonté critique

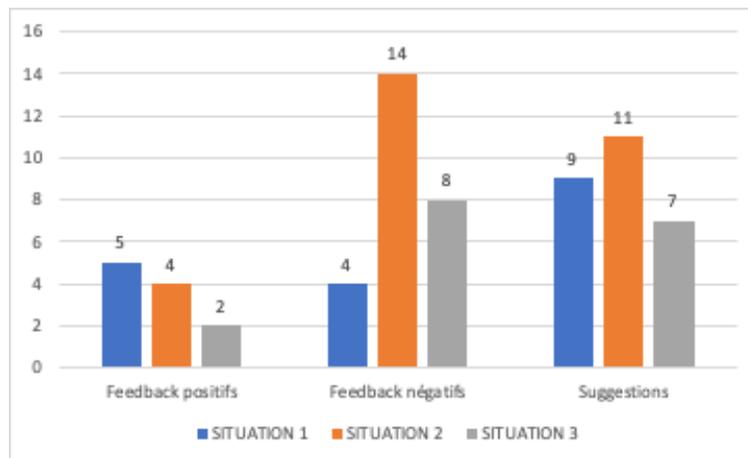


FIGURE 3.19 – Nombre de retours utilisateur par rapport à leurs types pour chaque situation d'évaluation

ni le même intérêt à tous les aspects de l'IHM quand ils regardent deux IHM ayant des changements au niveau concret, abstrait ou au niveau de la tâche et concepts.

Enfin, nous avons trouvé que présenter deux IHM avec des IUA différentes améliore les résultats de la comparaison : les utilisateurs fournissent plus de critiques (plus de commentaires négatifs) et aussi plus de suggestions d'amélioration. Ces retours sont précieux au concepteur pour améliorer la qualité de l'IHM.

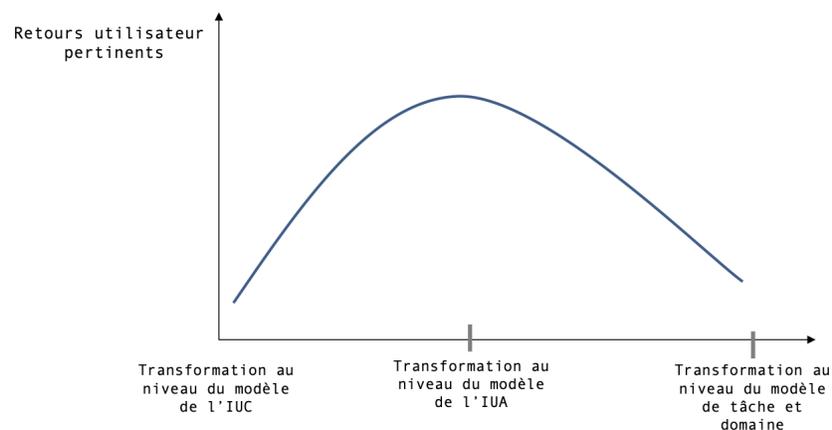


FIGURE 3.20 – Impact du choix de l'IHM alternative sur les retours utilisateur dans l'étude #2.

— *Les résultats de cette étude ont été présentés dans un papier lors de la conférence British HCI en 2017 (Hammami et al. 2017).*

Cette étude nous a orientées pour définir la meilleure alternative à générer pour les évaluation comparative. Cependant, le modèle de l'IUA peut varier selon deux paramètres principaux : le groupement entre tâches et la navigation entre tâches. Nous étudions dans la section suivante quelle transformation est la plus bénéfique, entre navigation et groupement, pour générer la meilleure alternative.

5 Étude #3 : Transformation au niveau de l'IUA

Cette étude est réalisée dans le but d'examiner plus en détail le type de transformations à appliquer au modèle de l'IUA pour la génération automatique des alternatives. Pour cela, différents participants ont évalué un ensemble d'IHM présentant des transformations différentes respectivement au niveau de la navigation et des groupements des tâches.

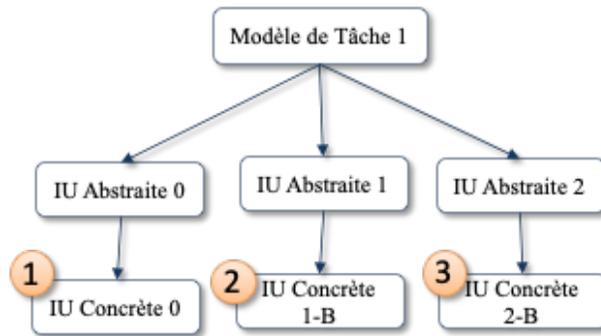


FIGURE 3.21 – Élaboration des IHM - Étude #3

5.1 Élaboration des prototypes

En partant du modèle de tâche élaboré dans la deuxième étude (Figure 3.11), nous avons généré trois IUA :

- IUA0 : représente l'IHM de référence. Cette IUA a été conçue manuellement en partant du modèle de tâche présenté dans la figure 3.11.
- IUA1 : une variation de IUA0 représentant des transformations au niveau de la navigation entre les tâches.
- IUA2 : une variation de IUA0 représentant des transformations au niveau des groupement des tâches.

Les figures 3.22 et 3.23 illustrent les trois IUA conçues pour cette étude.

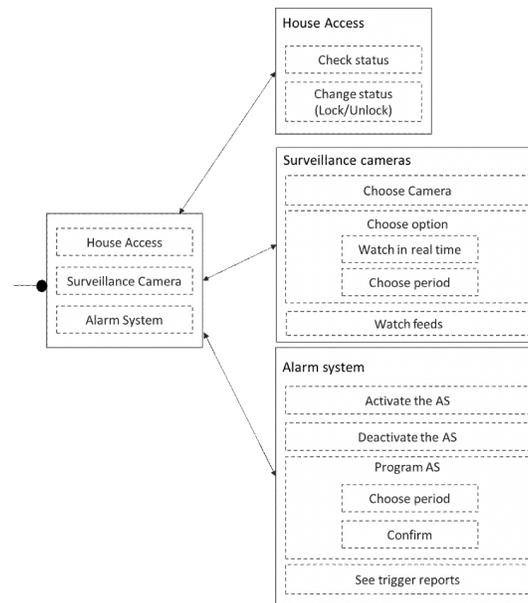


FIGURE 3.22 – Interface Utilisateur Abstraite 0

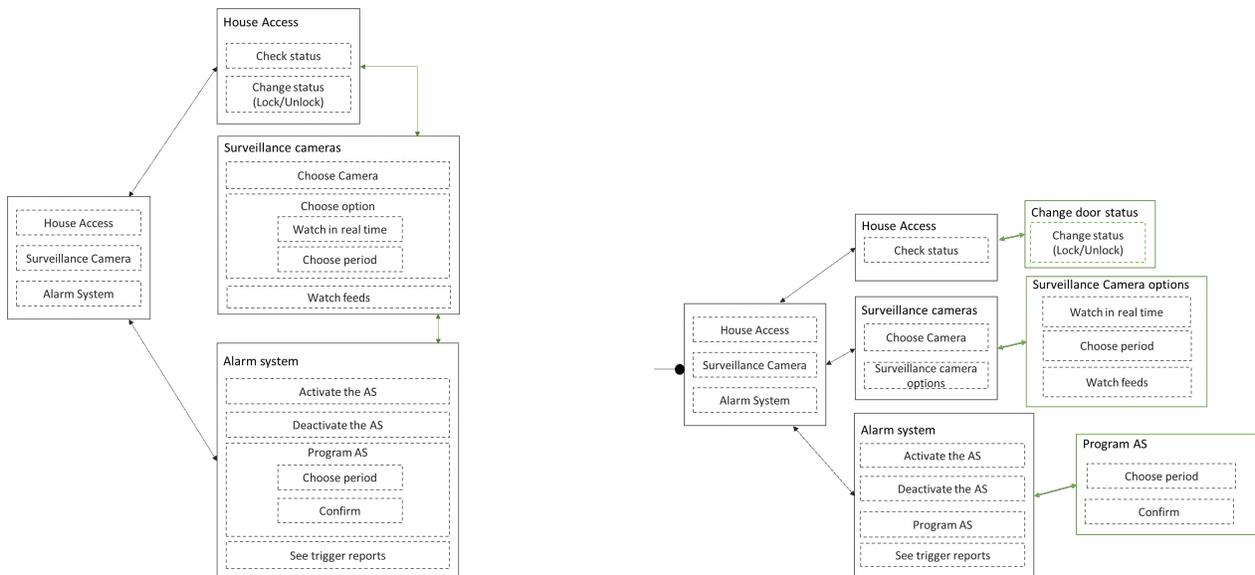


FIGURE 3.23 – IUA 1 (à gauche) et IUA 2 (à droite)

Nous avons ensuite conçu trois IUC correspondantes à ces IUA, ces dernières ont été conçues à l'aide de *Balsamiq Mockups* et présentent un niveau de fidélité moyen. Les figures 3.24, 3.25, et 3.26 illustrent les trois IUC produites.

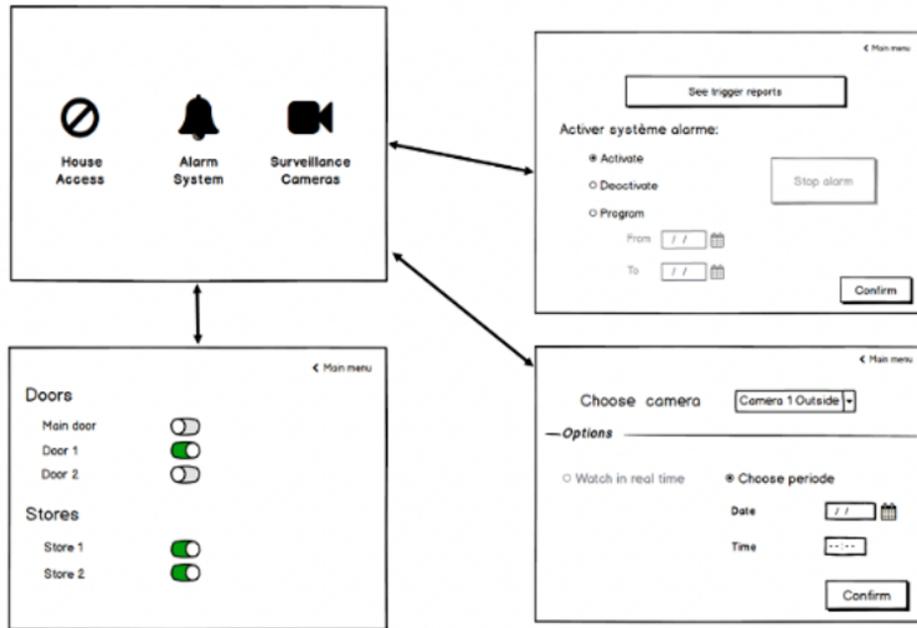


FIGURE 3.24 – IU Concrète 0

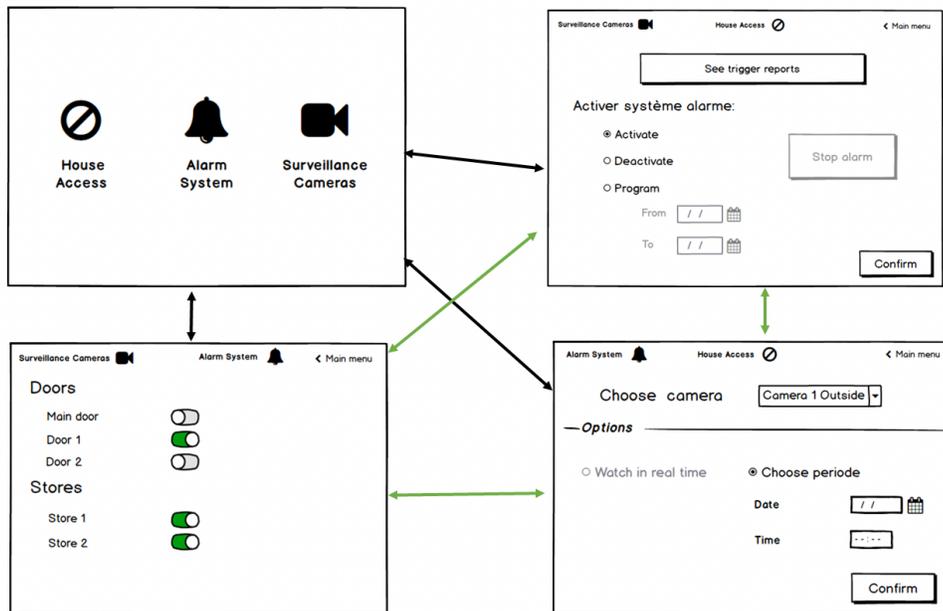


FIGURE 3.25 – IU Concrète 1-B

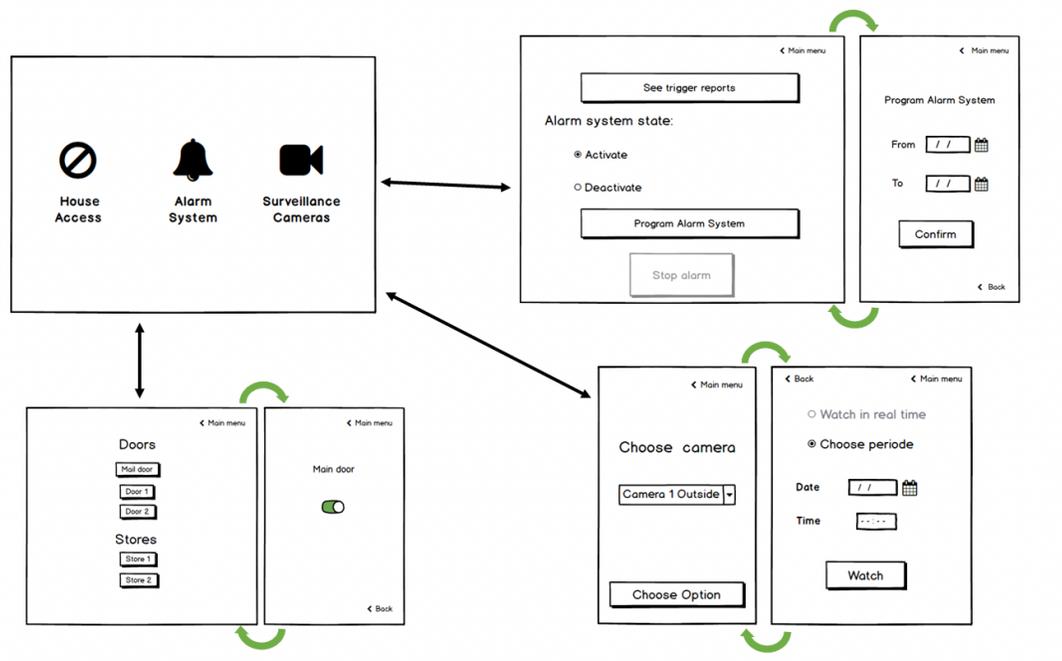


FIGURE 3.26 – IU Concrète 2-B

5.2 Participants

Nous avons recruté un total de 32 personnes pour cette étude. L'étude a impliqué à la fois des personnes familiarisées avec le domaine de l'IHM ainsi que des personnes n'ayant aucune connaissance du domaine comme par exemple des étudiants en ingénierie, en master et en Doctorat venant d'autres domaines.

5.3 Protocole expérimental

Le protocole expérimental appliqué dans cette étude est le même que celui de la deuxième étude (section 4.4).

Les participants ont été répartis en deux groupes. Nous leur avons présenté à chaque fois deux IHM simultanément : l'IHM de référence (IU concrète 0) et l'une des alternatives. Le premier groupe (situation d'évaluation 1) a évalué IUC 0 avec IUC 1-B, et le deuxième groupe (situation d'évaluation 2) a évalué IUC 0 avec IUC 2-B.

Chaque groupe a été invité à répondre à un questionnaire structuré en deux parties. La première partie était composée de neuf questions basées sur une échelle de Likert à 5 points (allant de "Pas du tout d'accord" à "tout à fait d'accord") qui portent sur le guidage, la charge de travail, la navigation et l'esthétique.

Dans la deuxième partie du questionnaire, les participants ont été invités à exprimer leurs

avis (ce qu'ils aimaient, ce qu'ils n'aimaient pas et s'ils avaient des suggestions d'amélioration), à travers de questions ouvertes. Enfin, ils ont exprimé leurs préférences entre les deux IHM présentées.

Nous n'avons imposé aucune limite de temps : les participants ont pris le temps nécessaire pour observer et évaluer les prototypes présentés.

5.4 Résultats

Les résultats de cette étude ont été analysés comme suit : nous avons calculé le score donné à IUC0 en nous basant sur les notes données par les différents participants dans la première partie du questionnaire. Ensuite, nous avons calculé le nombre total des retours fournis dans la deuxième partie du questionnaire, pour chaque situation d'évaluation. Enfin, nous avons analysé la qualité et le type de ces retours utilisateur.

5.4.1 Notes obtenues par l'IHM de références

Nous avons calculé le score accordé à IUC0 à partir des évaluations unitaires effectuées dans la première partie du questionnaire. La figure 3.27 présente ce score dans les deux situations d'évaluation.

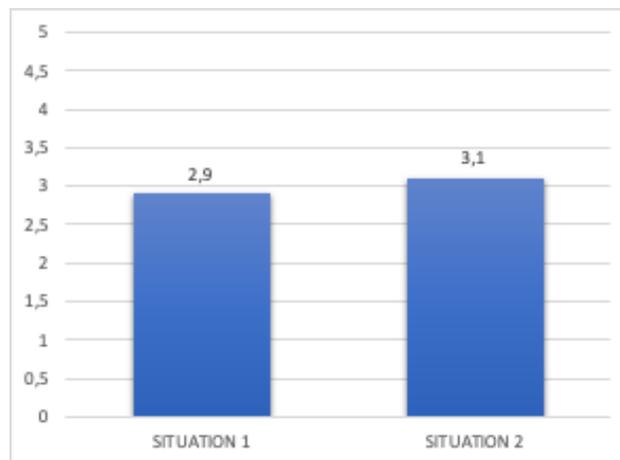


FIGURE 3.27 – Score de IUC0 pour chaque situation d'évaluation - Étude #3.

Nous constatons que les scores calculés montrent une légère différence entre les deux situations d'évaluation.

5.4.2 Quantité des retours utilisateur

Le tableau 3.3 indique le nombre total de retours utilisateur collectés. Nous avons obtenu 99 retours : 54 dans la première situation d'évaluation et 45 dans la seconde.

<i>Situation d'évaluation</i>	<i>Nombre de retours utilisateur</i>
Situation d'évaluation 1	54
Situation d'évaluation 2	45

TABLE 3.3 – Nombre total de retours utilisateur pour chaque situation d'évaluation - Étude #3.

Nous observons que les participants fournissent plus de retours en comparant deux IHM ayant une différence de navigation. Nous analysons dans les sections qui suivent ces retours en termes de leur qualité afin de déterminer quelle transformation est plus bénéfique pour les évaluations comparatives.

5.4.3 Qualité des retours utilisateur

Dans cette section, nous analysons la qualité des retours utilisateur. Nous commençons par définir leur type et nous les classons en commentaires négatifs, commentaires positifs et suggestions d'amélioration. La figure 3.28 montre les résultats pour chaque situation d'évaluation par rapport à cette classification. Par la suite, nous les analysons par rapport à leur qualité et nous les classons selon s'ils portent sur l'utilisabilité, l'utilité et l'esthétique. Ces deux classifications nous permettent d'étudier l'impact du choix de l'alternative sur la volonté critique des participants et la pertinence de leurs retours pour améliorer l'IHM de référence.

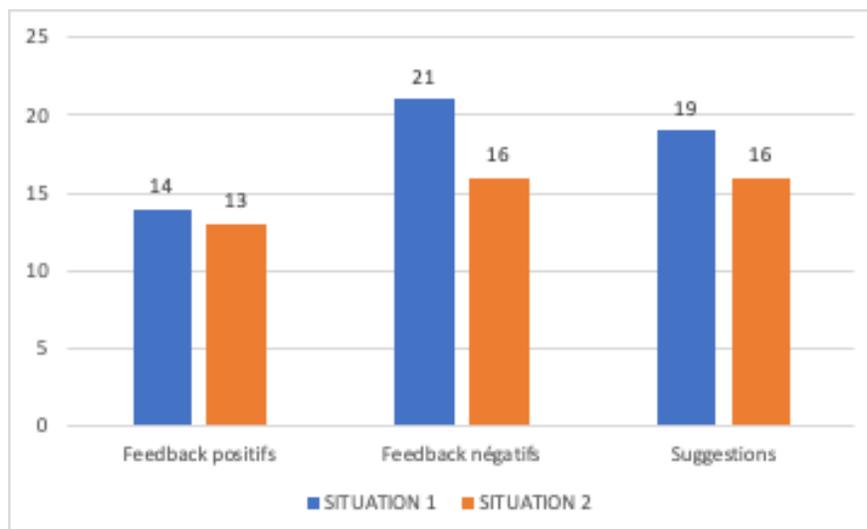


FIGURE 3.28 – Nombre de retour utilisateur en fonction de leur type pour chaque situation d'évaluation - Étude #3.

- Exemples de commentaires positifs fournis par les participants : "*J'aime les widgets de la page du menu principal*", "*Simple et clair*", "*Regroupement de toutes les options dans une seule fenêtre, facile à manipuler*".
- Exemples de commentaires négatifs fournis par les participants : " *Icône d'accès à la maison: mauvais choix.*", "*Il n'existe pas un bouton d'annulation d'action de plus absence de l'ergonomie de l'interface*".
- Exemples de suggestions fournies par les participants : «*J'ajouterais un bouton de confirmation dans la page d'accès à la maison*», «*Vous deviez ajouter une description pour chaque action*», «*Je changerais la position du bouton d'arrêt d'alarme*».

Les résultats soulignent que le choix de l'IHM alternative pour l'évaluation influence le nombre et le type de retours fournis par les participants. Par exemple, 26% des retours fournis dans la première situation d'évaluation sont des commentaires positifs contre 37% dans la deuxième situation. En outre, le nombre de commentaires négatifs ainsi que le nombre de suggestions sont plus importants dans la première situation d'évaluation que dans la deuxième, ce qui montre une volonté critique plus importante dans cette situation d'évaluation.

Tel que le montre la figure 3.29, nous n'avons pas observé de différence significative du nombre de retours utilisateur entre les deux situations d'évaluation en fonction de leur type. Cependant une observation importante est que la présentation de différentes IHM avec des variations au niveau du modèle de l'IUA produit essentiellement des retours portant sur l'utilisabilité de l'IHM.

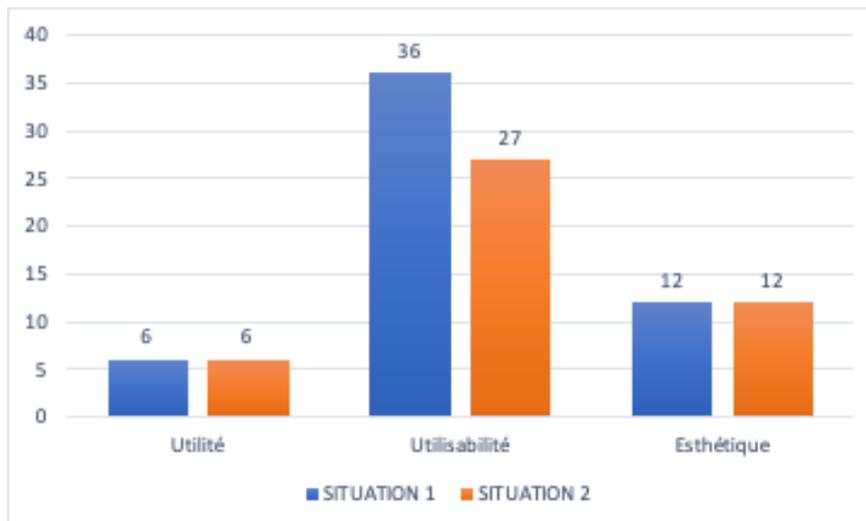


FIGURE 3.29 – Nombre de retours utilisateur pour chaque situation d'évaluation en fonction de leur qualité - Étude #3.

Exemples de retours utilisateur :

- "J'ai eu un peu de mal à faire le lien entre le menu principal et les 3 autres pages. Du coup, je suggère que chaque interface soit intitulée avec son icône (accès, alarme, caméra)."
- "Il faut toujours retourner à l'interface principale pour changer de service, je n'aime pas ça."
- "Manque de facilité de navigation entre les pages. Ce n'est pas possible de naviguer d'une page à une autre sans repasser par la page du menu principal : lent."
- "L'interface est simple mais en l'utilisant il faut prendre un moment pour comprendre ce que fait exactement chaque action."
- "Je n'aime pas qu'il faille ouvrir les pages pour choisir les options."
- "L'icône House Access n'est pas bien choisi."
- "Le regroupement de toutes les options dans une seule fenêtre facilite la manipulation."

Finalement, cette étude montre que la présentation d'une alternative d'une IHM avec des transformations de la navigation au niveau de l'IUA est mieux adapté pour recueillir des retours utilisateur servant à améliorer la qualité de l'interaction, que la présentation de deux IHM présentant des transformations au niveau des groupements des tâches.

6 Conclusion

Dans ce chapitre, nous avons présenté les différentes études expérimentales que nous avons menées afin de déterminer les critères de génération de la meilleure IHM alternative pour les évaluations comparatives. Ces études ont permis de mesurer l'impact et le bénéfice de l'utilisation de différents types d'alternatives dans un contexte d'évaluation expérimentale d'IHM. Les alternatives utilisées présentaient des changements de support et d'IHM. Les résultats de ces études ont révélé que :

1. La diversité des supports est bénéfique, ici un prototype statique et un prototype dynamique, en termes de retours utilisateurs : les utilisateurs fournissent des retours portant sur l'interaction et l'interface, permettant ainsi d'améliorer la qualité de l'IHM.
2. La transformation d'IHM est bénéfique et maximale lorsqu'elle est opérée au niveau de l'IUA.
3. Les transformations de navigation sont les plus performants pour les évaluations comparatives.

Ce travail répond à notre première question de recherche : en termes de produit, quelle est l'IHM à présenter pour optimiser les résultats des évaluations comparatives ? La réponse est une interface abstraite alternative dans laquelle la navigation a été changée.

Le chapitre suivant complète les contributions sur l'aspect processus : comment générer une telle alternative à moindre coût ?

4

**GRA:FITl, un environnement de génération
automatique d'alternatives**

1 Introduction

Ce chapitre propose un environnement logiciel de génération automatique d'alternatives d'IHM. Les variations se situent au niveau de l'interface abstraite, consistant en un changement de navigation. Nous présentons ici les principes de génération puis décrivons GRA:FITI, notre démonstrateur de concept. Nous présentons ensuite son évaluation et concluons sur la valeur de la proposition.

2 Principes de génération

Au chapitre 3, nous avons montré que l'évaluation comparative fournit de meilleurs résultats lorsque les utilisateurs comparent des IHM présentant un schéma de navigation différent. La question est donc comment produire une telle variante et à partir de quelle IHM.

L'espace d'exploration est le suivant :

- En entrée de l'outil, peuvent être fournis une IHM finale et/ou une IHM concrète et/ou une IHM abstraite et/ou un modèle de tâche et de concepts de référence,
- En sortie, doit être produite une IHM concrète ou finale alternative.

Pour ce faire, la transformation de l'entrée peut être une séquence de réifications et/ou abstractions et/ou translations et/ou reciblages (Cf. chapitre 2, avec des illustrations en figure 4.1).

Nous présentons ci-dessous les principes retenus.

Principe #1. Laisser les praticiens s'exprimer en terme d'interface concrète

En pratique, les concepteurs se placent naturellement au niveau de l'interface concrète pour esquisser leurs IHM. Ils sont rarement familiers des approches basées modèles et, même formés à ces abstractions, l'exercice leur reste difficile. Ce constat nous amène au principe pragmatique de laisser les praticiens esquisser des IHM concrètes.

Principe #2. Aller au-delà de l'interface et démontrer l'interaction

Au chapitre précédent, nous avons montré le bénéfice de jouer l'interaction en évaluation comparative. Aussi notre objectif est de générer des alternatives interactives, permettant par exemple la production automatique de vidéos et la comparaison de scénarios d'interaction.

Principe #3. Utiliser l'architecture logicielle comme langage pivot

Des travaux historiques ont montré la possibilité de générer l'architecture logicielle PAC (Annexe C) d'un système interactif à partir de ses spécifications externes (Degriigny et Duval 2004). Cet apport permet, à partir d'une spécification externe, de générer une IHM finale, certes incomplète, mais interactive.

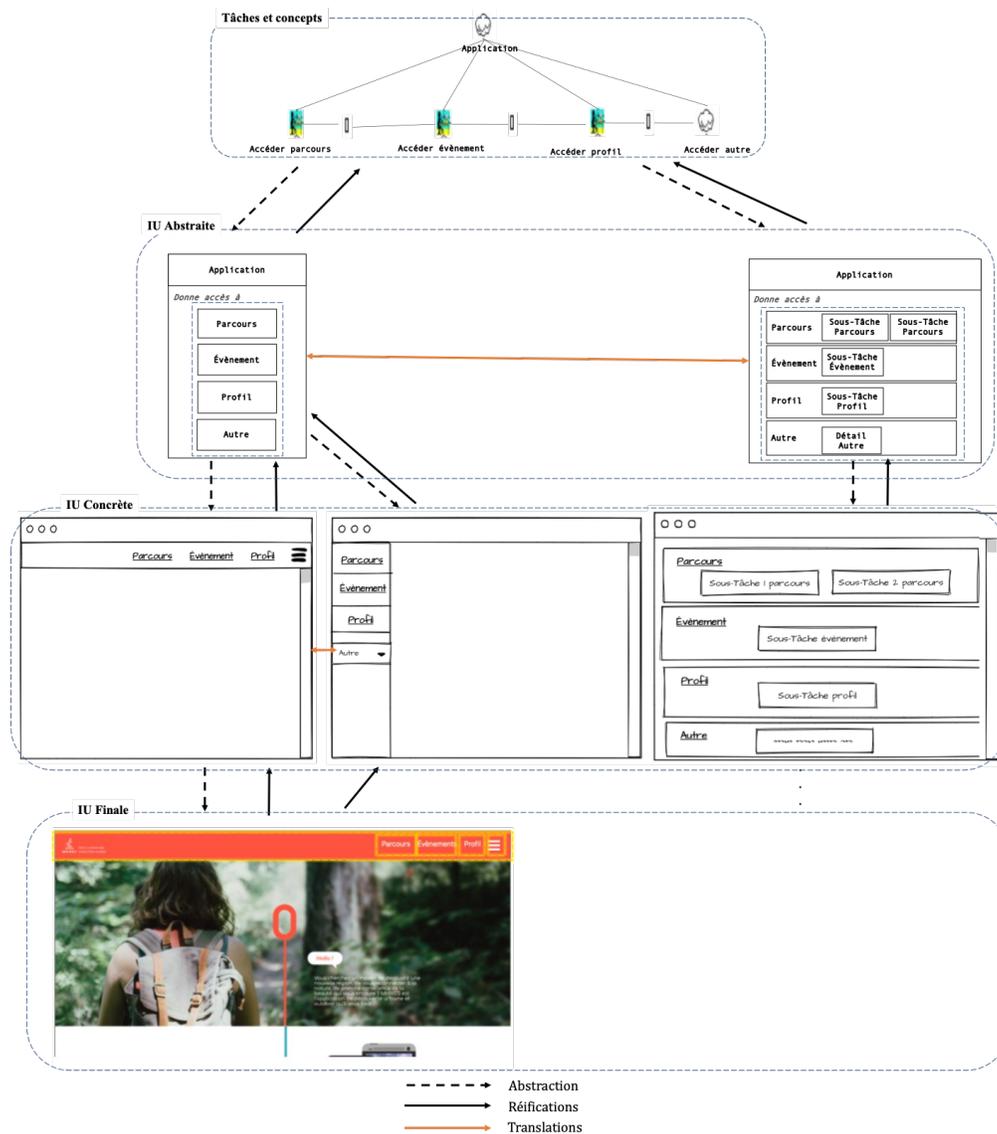


FIGURE 4.1 – Produire des IHM par transformation de modèles

Il s'agit donc ici de produire le modèle PAC de l'IHM de référence et de le transformer pour générer une IHM alternative. Nous optons pour des transformations de type *Translation*, par modification des facettes Présentation des agents PAC. Ce choix se justifie, d'une part, par l'impossibilité de raisonner sur les facettes Abstraction faute de sémantique et, d'autre part, par la faisabilité technique de changer de présentations, comme démontré dans des approches à COMET par exemple (Daassi 2007). Les COMET sont des widgets polymorphes, c'est-à-dire à présentations multiples, offrant donc, par construction, une capacité de transformation.

Le chapitre précédent ayant conclu que les changements devaient être focalisés sur la

navigation, les translations cibleront les agents PAC de navigation. Par exemple, un agent entrelacement, incarné par un séparateur commutera vers une présentation multi-fenêtres. Autrement dit, il s'agit de modifier les tâches articulatoires de l'IHM de référence (Thevenin et al. 1999) : "*des tâches utilisateurs qui ne modifient pas le noyau fonctionnel d'un système, elles permettent d'accéder à un élément de deuxième niveau dans un modèle de tâche*". Par exemple, dans la figure 4.2-A à gauche, l'utilisateur a besoin de cliquer sur le bouton Main door pour consulter son état et le changer via une autre page. Dans l'exemple à droite (Figure 4.2-B), l'utilisateur dispose d'un accès direct à toutes les informations et tâches sans avoir besoin de naviguer entre différentes pages et peut donc changer l'état de la porte sans être dirigé vers une page différente.

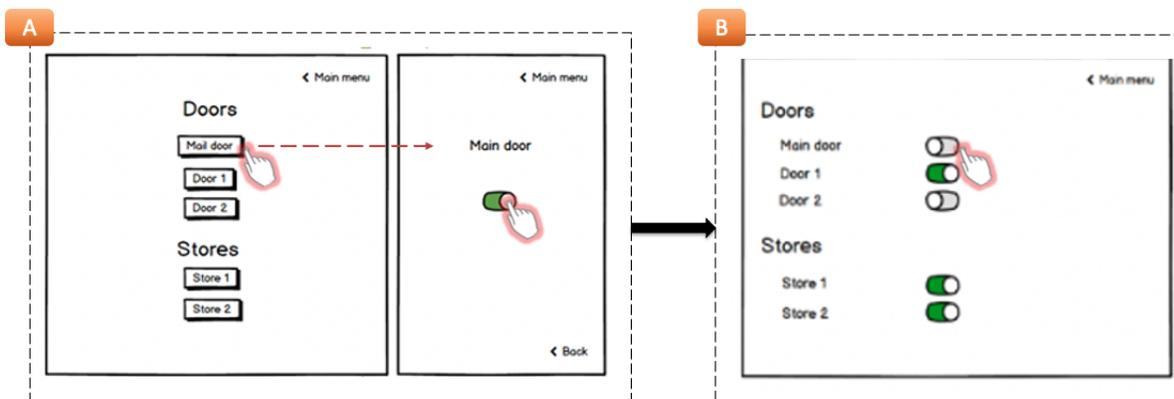


FIGURE 4.2 – Exemple de tâche articulatoire

Dans la hiérarchie d'agents PAC, nous ciblons donc les agents qui ne disposent pas de facette Abstraction. Ces agents correspondent à des tâches articulatoires ou à des agents Ciment (par exemple, l'agent C dans la figure 4.3) : "*Si les actions de l'utilisateur sont réparties sur plusieurs agents et si ces actions sont significatives pour le système, alors un agent ciment doit être introduit pour synthétiser ces actions réparties sur plusieurs agents.*" (Calvary et al. 1997).

Principe #4. Générer une alternative de moindre coût articulatoire

Pour l'adoption de l'outil par les praticiens, l'IHM générée doit être de meilleure qualité, sur certains points au moins, que l'IHM de référence. Par ce progrès, c'est dans un processus de convergence progressive vers l'IHM finale que nous inscrivons l'outil : le praticien pourrait même, sans mener d'évaluations comparatives, décider de prendre l'IHM générée comme IHM de référence. Dans notre approche, c'est le coût articulatoire que nous retenons comme critère qualité, avec la nécessité de réduire la trajectoire d'interaction.

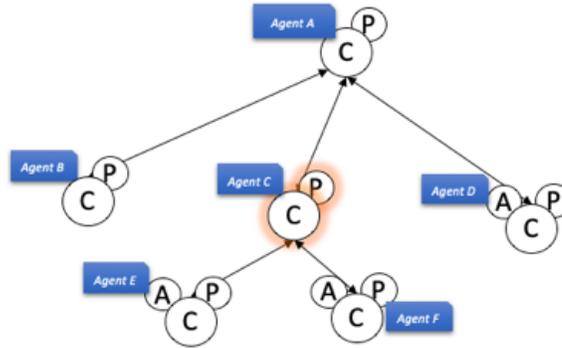


FIGURE 4.3 – Exemple d’agent articulatoire (agent C)

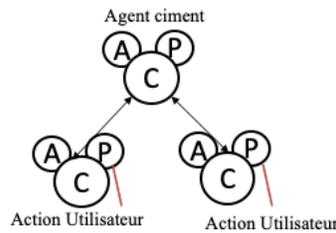


FIGURE 4.4 – Agent ciment dans (Calvary et al. 1997)

Principe #5. Produire automatiquement les vidéos de scénarios d’interaction sur les IHM initiales et générées

Grâce aux modèles PAC d’entrée et de sortie et aux correspondances entre ces deux modèles, l’outil peut automatiquement jouer l’interaction à partir d’un scénario des actions accomplies par un utilisateur sur l’IHM de référence. Ce principe est intéressant pour, à faible coût, produire les vidéos, en support à l’évaluation comparative, conformément au chapitre précédent (chapitre 3).

3 Démonstrateur de concept

Nous présentons ici nos choix techniques puis décrivons l’outil en action.

3.1 Choix techniques

GRA:FITI est un outil web programmé en PHP. Il prend en entrée un modèle PAC de l’IHM de référence et génère des IHM alternatives en terme de navigation. Chaque modèle

PAC généré peut, à son tour, conformément au principe #4, être utilisé comme référence pour itérativement converger vers la bonne IHM finale (figure 4.5).

La génération des alternatives est basée sur une substitution des éléments de présentation de la facette Présentation (P) des agents articulatoires du modèle de référence de l'IHM par d'autres éléments de présentation. Pour cela, nous avons développé une boîte à outils qui contient un ensemble de scripts permettant d'identifier les agents articulatoires dans une hiérarchie d'agents PAC et de changer leurs facettes Présentation. Cette boîte à outil permet de charger des fichiers d'entrée, de les modifier et de les sauvegarder.

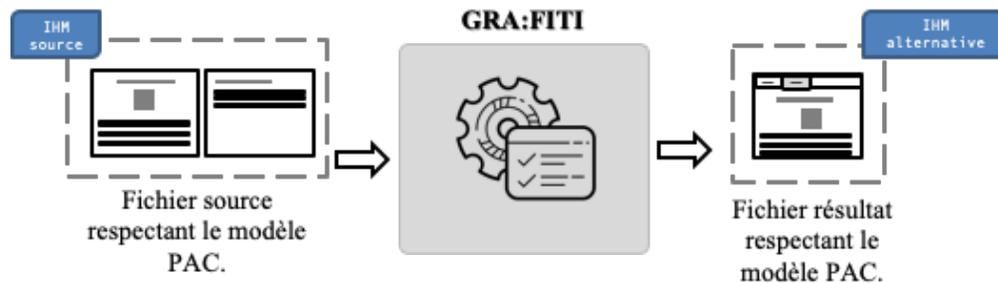


FIGURE 4.5 – GRA:FITI : Principe

Le processus de génération de l'IHM alternative se structure en deux étapes. La première consiste à analyser le modèle PAC de référence afin d'identifier les agents articulatoires (Script 1). La deuxième étape consiste à modifier la facette Présentation (P) de ces agents (Script 2). Cette transformation est basée sur un ensemble de paramètres définis dans un fichier structuré JSON dans la boîte à outils (figure 4.6). La présentation des agents non articulatoires de l'IHM est conservée telle que définie par le concepteur.

Étape 1 : Identification des agents articulatoires de l'IHM d'entrée

PAC modélise une IHM en une hiérarchie d'agents, chacun étant composé de trois facettes : Abstraction, Présentation et Contrôle. Chaque agent correspond à une partie spécifique de l'IHM. Nous présentons dans la figure 4.7 l'arborescence de fichiers, selon ce modèle, pour un site web. Chaque élément de l'IHM est représenté par un agent et est enregistré dans un dossier à son nom qui contient son contrôleur (fichier commençant par "C_ "), son abstraction (fichier commençant par "A_ ") et sa présentation (fichier commençant par "P_ "). Cette organisation permet de repérer facilement les différentes parties de chaque agent dans l'arborescence.

Le premier script parcourt l'arborescence à la recherche d'agents n'ayant pas de facette Abstraction. Ces agents sont déterminés par le fait qu'ils n'ont que deux parties : une partie contrôle (C) et une partie présentation (P). La partie abstraction (A) dans ce cas n'est pas définie car ce type d'agents ne communique pas avec le modèle de données.

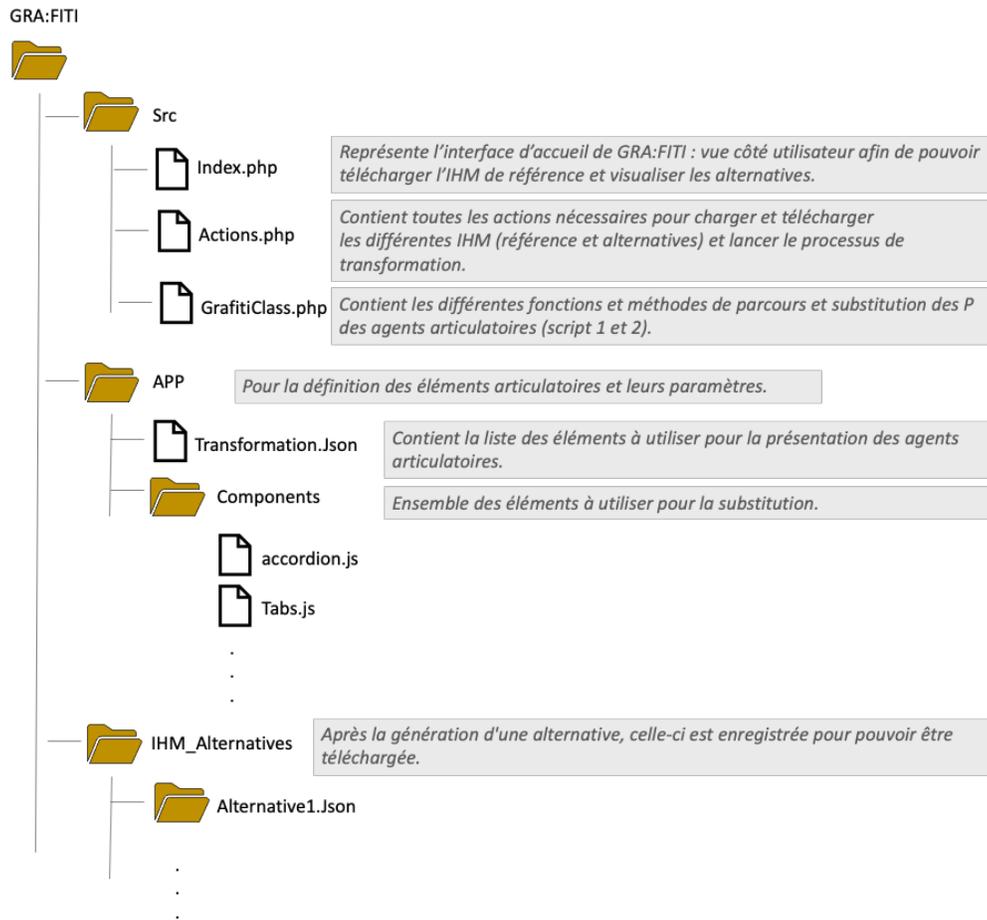


FIGURE 4.6 – Vue globale de notre boîte à outil

Étape 2 : Modification de la facette Présentation (P) des agents articulatoires

La deuxième étape consiste à parcourir la liste des agents articulatoires identifiés en étape 1 et de modifier leurs facettes présentation. Pour cela, un deuxième script est activé, dédié à la substitution des éléments utilisés pour la présentation de ces agents par des éléments ayant les mêmes fonctionnalités mais qui présentent un moindre coût articulatoire.

Le coût articulatoire est défini comme étant la somme des coûts de navigation (NC : *Navigation Cost*) et de trajectoire d'interaction (ITC : *Interaction Trajectory Cost*).

Le coût de navigation est modélisé comme le nombre de clics qu'un utilisateur doit effectuer pour accomplir une tâche en manipulant un élément d'interaction (un widget). Par exemple, pour choisir une option en utilisant un widget de type liste déroulante, un utilisateur (1) ouvre la liste déroulante pour voir les différentes options et (2) choisit une option. Cependant, dans le cas de l'utilisation d'un widget de type bouton radio l'utilisateur n'a qu'à choisir l'option souhaitée, il gagne donc en nombre de clics.

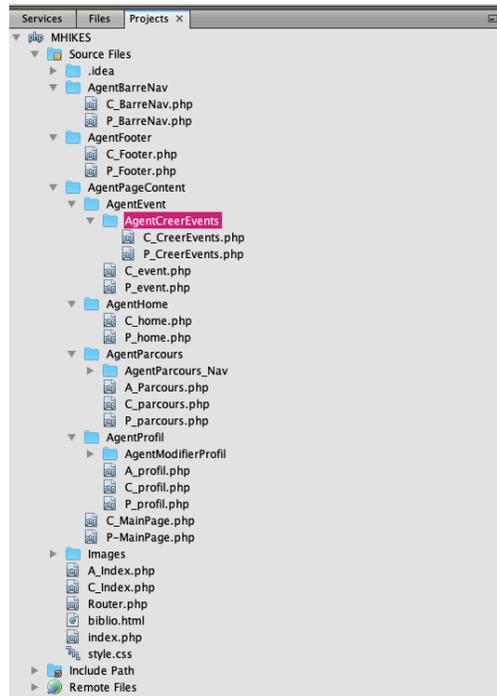


FIGURE 4.7 – Exemple d’une arborescence d’une IHM implémentée suivant l’architecture PAC

Ce coût est prédéfini dans le fichier structuré JSON (*Transformation.Json*) de notre boîte à outils à la granularité des widgets élémentaires. Dans la version actuelle, ces widgets sont :

- Les menus de navigation : navigation (navbar), accordéon, onglet (tabulation) et menu de recherche.
- Les sections : blocs séparés par un simple séparateur qui segmentent des portions de l’IHM (regrouper d’autres éléments HTML).

La trajectoire d’interaction (ITC) est calculée en fonction du nombre de requêtes passées entre les différents agents afin d’afficher un résultat demandé selon la hiérarchie du modèle de l’IHM.

Pour chaque widget élémentaire, une liste de présentations possibles est définie dans le fichier *Transformation.Json*. Par exemple, un menu de navigation (NavBar) peut être remplacé par un menu à onglet ou un menu accordéon. En outre, chaque widget est défini par les propriétés suivantes :

- NC : une valeur qui représente le coût de navigation.
- Balise : définition de la balise HTML qui correspond à l’élément.
- Description : description de l’élément.

Les widgets disponibles sont en outre définis dans un dossier "*components*" sous forme de fichiers *JavaScript* afin de pouvoir les intégrer dans le code HTML pour la substitution.

Dans l'exemple de la figure 4.8, nous avons à gauche une IHM dans laquelle un utilisateur a la possibilité de rechercher ou de créer un itinéraire. Ces deux tâches sont séparées par un menu à onglet : l'utilisateur a donc besoin de cliquer sur l'onglet souhaité pour changer de tâche. Le menu à onglet est un agent articulatoire qui peut être substitué par un élément de type section afin de minimiser le coût articulatoire (IHM à droite).

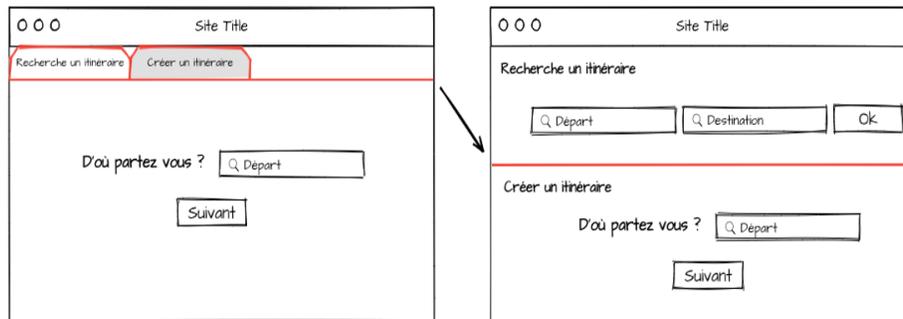


FIGURE 4.8 – Exemple de deux alternatives d'une IHM présentant des systèmes de navigation différents

Concernant le style de présentation des widgets substitués, notre outil, dans sa version actuelle, ne traite pas cet aspect. Les widgets insérés héritent donc du style défini dans le fichier CSS de référence. En outre, si le style du widget utilisé n'est pas défini, celui-ci garde un style standard HTML.

Notre outil génère trois alternatives présentant des transformations : de navigation globale, de navigation locale et une combinaison de navigations globales et locales. La définition du type de navigation se fait par rapport au positionnement des agents concernés dans l'arbre HTML du modèle de l'IHM. La navigation globale, ou navigation principale, est responsable d'afficher les différentes pages de niveau supérieur du modèle. Elle est disponible sur chaque page et répertorie les principales sections et/ou pages de contenu du système. La navigation locale, ou navigation au niveau des pages, permet d'accéder aux niveaux inférieurs d'une structure. La navigation locale montre généralement d'autres options au même niveau d'une hiérarchie, ainsi que les options sous une page en cours. Nous présentons dans la figure 4.10 un exemple de trois alternatives représentant ces transformations.

3.2 GRA:FITI en action

L'utilisateur charge son IHM de référence via un bouton de téléchargement placé dans la page d'accueil (*index.php*) (Figure 4.11). L'outil génère ensuite automatiquement les IHM

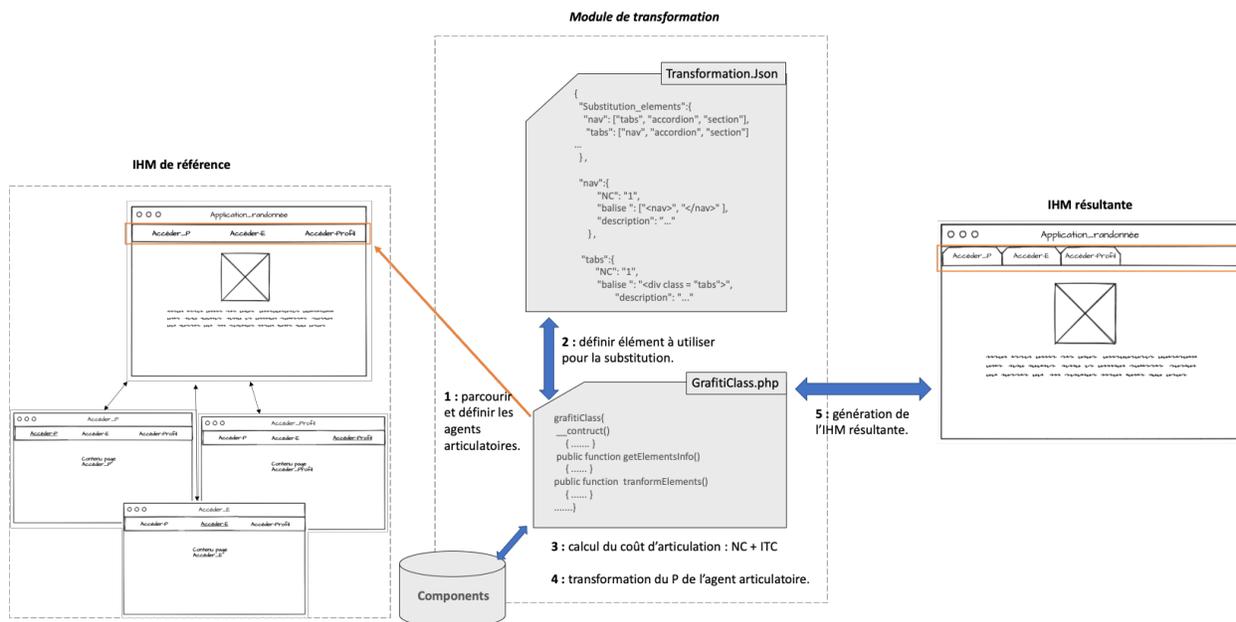


FIGURE 4.9 – Vue globale

alternatives et les affiche à côté de l'IHM de référence, toutes intégrées dans des éléments de type IFrame, placées en juxtaposition (Figure 4.12). En d'autres termes, les IHM sont contenues dans des fenêtres indépendantes les unes des autres. Cela permet à l'utilisateur de naviguer et d'exécuter les différentes tâches dans les IHM respectives.

L'utilisateur peut indiquer s'il apprécie ou non telle alternative. Il peut aussi télécharger les alternatives et les utiliser comme nouveau point d'entrée pour générer d'autres alternatives en retour et ainsi explorer plus largement l'espace de conception.

L'utilisateur peut aussi sélectionner deux alternatives et les comparer au lieu de conserver son IHM source (affichée à gauche).

Nous présentons dans la section suivante une première évaluation expérimentale de GRA:FITI.

4 Processus d'évaluation de GRA:FITI

Nous présentons ici le protocole ainsi que les résultats d'évaluation.

4.1 Protocole

4.1.1 Cas d'étude

Le cas d'étude utilisé est un site web de randonnées. Le site offre essentiellement la possibilité de trouver et de créer différents parcours et événements pour des randonnées à

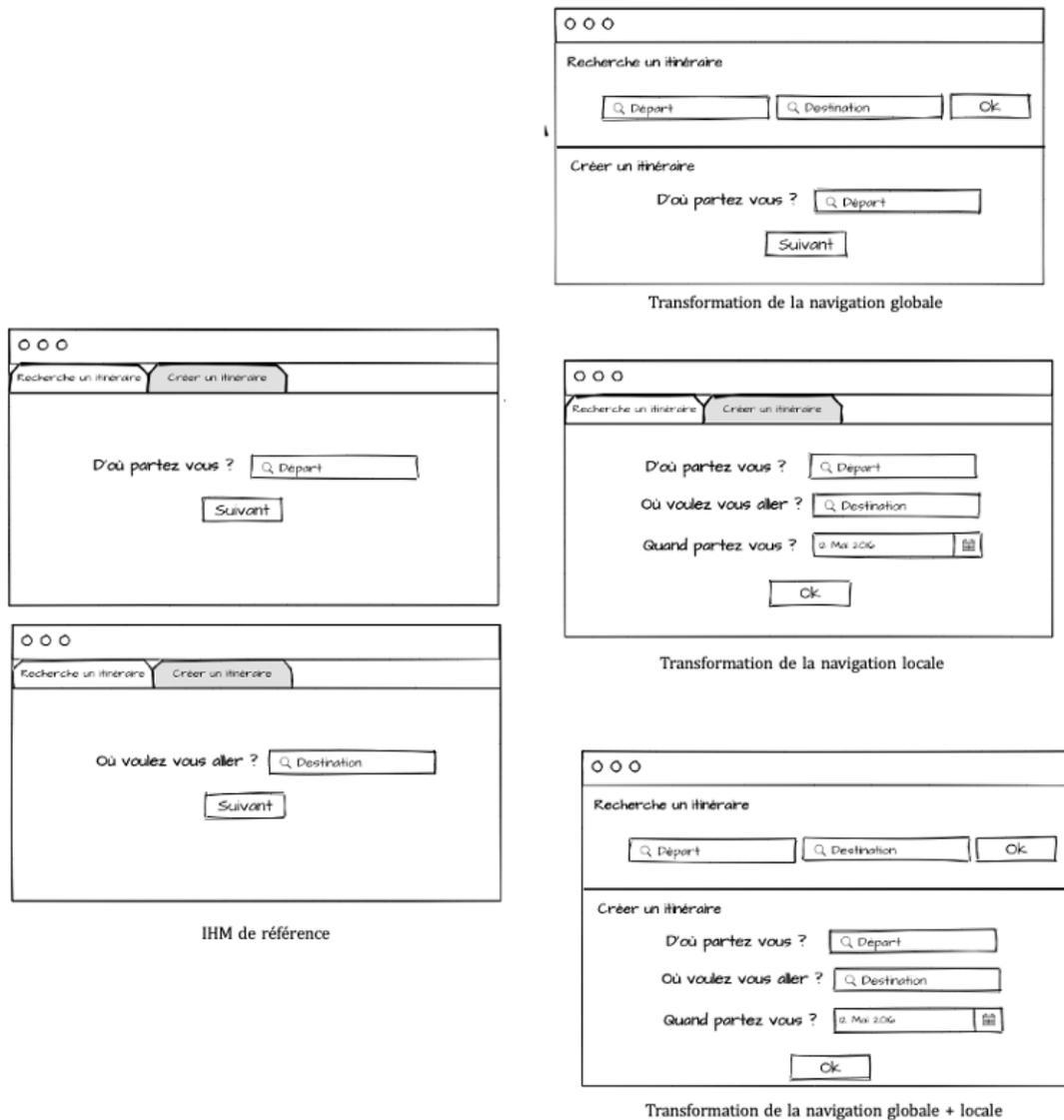


FIGURE 4.10 – Exemple de trois alternatives (à droite) d'une IHM de référence (à gauche) présentant des systèmes de navigation différents

une échelle internationale. Les parcours et les événements sont caractérisés par un nom, un type d'activité et un emplacement. Ces caractéristiques servent de filtres de recherche.

Via ce site web, un utilisateur peut consulter la liste de tous les parcours disponibles, rechercher ou créer un parcours. Il peut aussi consulter la liste de tous les événements, rechercher ou créer un événement. Un utilisateur peut créer un compte (son profil) qui contiendra la liste des événements qu'il a créés ainsi que ses informations personnelles qu'il peut modifier. L'utilisateur a aussi la possibilité de partager son profil s'il le souhaite. Le site Web

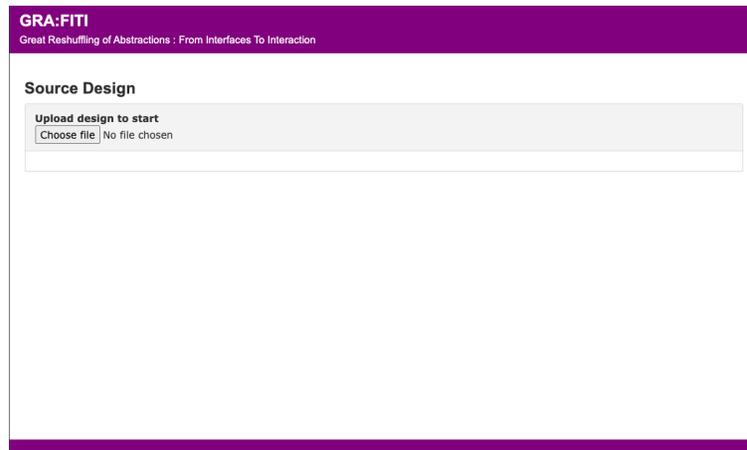


FIGURE 4.11 – Page d'accueil de GRA:FITI : téléchargement de l'IHM de référence

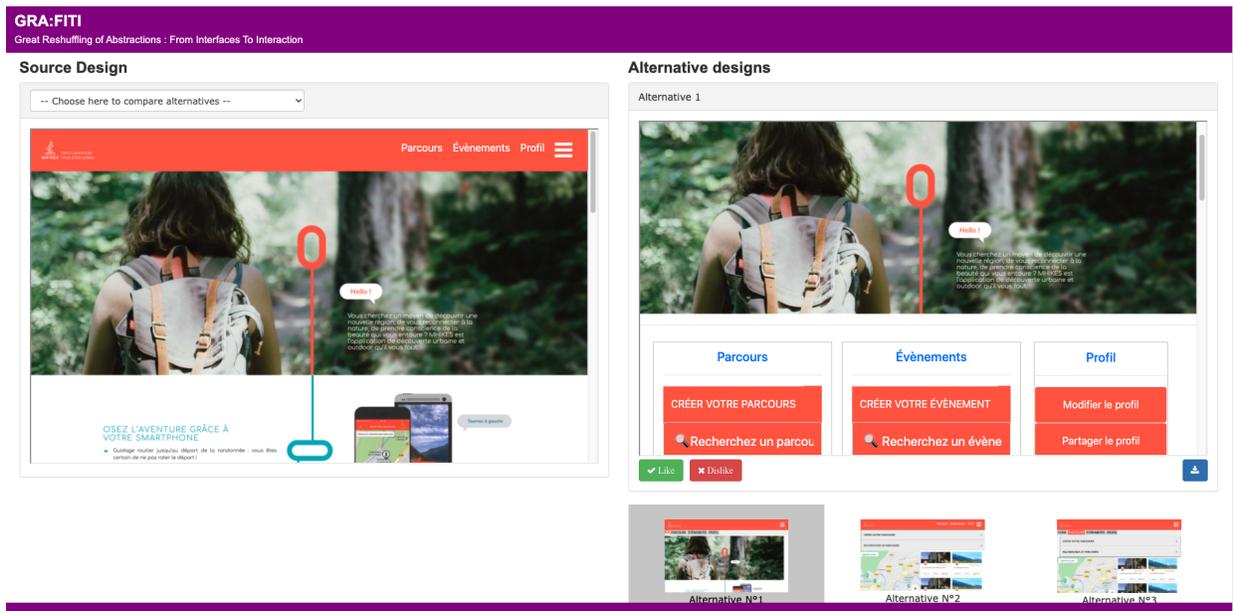


FIGURE 4.12 – GRA:FITI : Comparaison des IHM

contient différentes autres pages comme par exemple un formulaire de contact ou la liste des partenaires.

Nous avons choisi ce site car il contient plusieurs pages et offre différentes fonctionnalités. En plus, ce type de sites web répandu est familier à de nombreuses personnes, ce qui simplifie les sessions d'évaluation.

La figure 4.13 présente l'IHM source du cas d'étude. La figure couvre les différentes pages et fonctionnalités présentées : page d'accueil, page des parcours, page des événements et page

du profil utilisateur. La navigation entre les différentes pages se fait via la barre de navigation présentée en haut de chaque page.

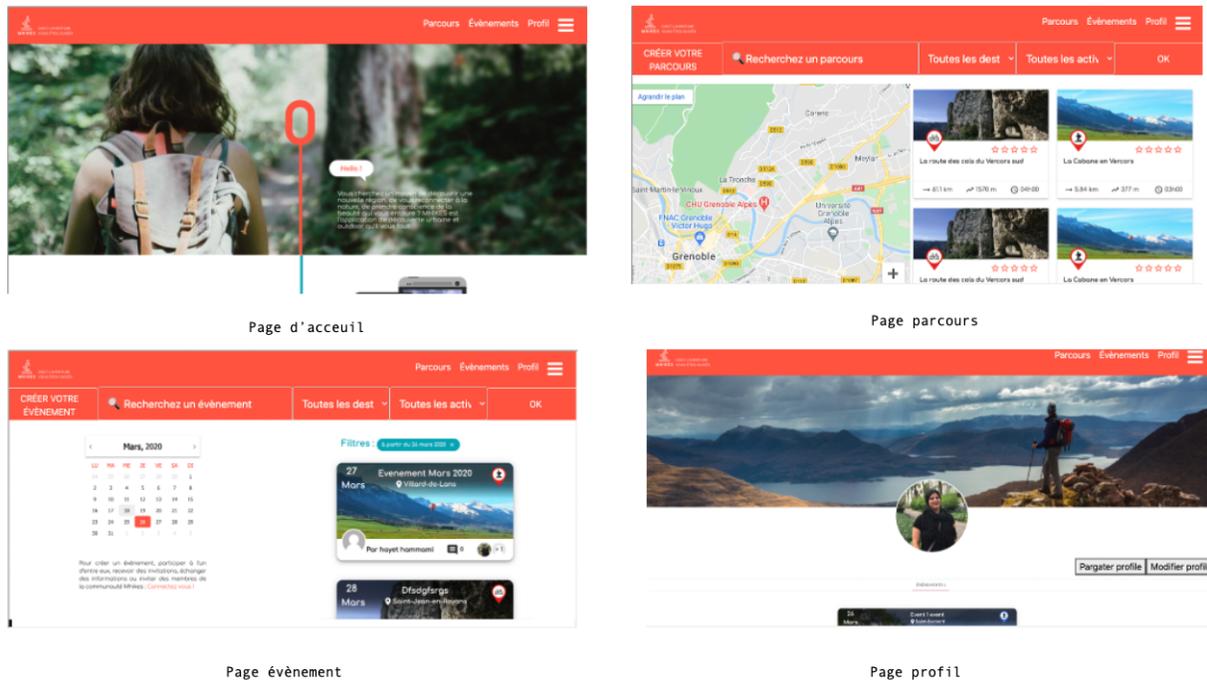


FIGURE 4.13 – IHM source

4.1.2 Participants

Nous avons recruté un total de 16 personnes, panel composé de 6 femmes et 10 hommes, âgés de 24 à 34 ans. Tous avaient tous des connaissances en conception des IHM mais n'étaient pas tous professionnels du domaine. Nous les classons en trois catégories : développeurs (*front-end* ou *back-end*), concepteurs (parfois en *free-lance*) et chercheurs (dont doctorants).

Le tableau 4.1 présente le nombre de participants par rapport à leur profil.

Profil du participant	Nombre
Développeur	5
Concepteur	5
Chercheur	6

TABLE 4.1 – Nombre des participants par profil

4.1.3 Protocole expérimental

Nous avons commencé par présenter l'objectif général de l'étude à chaque participant ainsi que les principales fonctionnalités du site web de référence (IHM source). Nous avons demandé aux participants de tester et d'explorer l'IHM afin de s'y familiariser avant d'afficher les alternatives. Ensuite, nous avons donné la liberté aux participants de tester GRA:FITI et de générer les IHM alternatives. Aucune contrainte de temps n'a été imposée, les participants prenaient leurs temps pendant toute la séance d'évaluation. Une fois que le participant indique avoir terminé son évaluation (avoir exploré les différentes alternatives et les avoir testées), nous lui montrons trois vidéos qui exposent l'exécution d'un scénario de création de parcours sur les trois alternatives présentées en plus de l'IHM de référence.

Nous avons adopté la méthode de la pensée à voix haute (Think-Aloud Protocol) et avons donc demandé aux participants d'exprimer leurs avis à voix haute pendant l'évaluation. Nous leur avons également demandé de répondre à un questionnaire qui se divise en deux parties. La première partie correspond à un questionnaire de type SUS (*System Usability Scale*) (Brooke 2013). Il se compose de dix questions associées toutes à un choix de réponse sous forme d'échelle de Likert en cinq points (allant de "pas du tout d'accord" à "tout à fait d'accord"). Ce questionnaire a pour objectif de déterminer le niveau de satisfaction des utilisateurs vis-à-vis de GRA:FITI, d'évaluer sa facilité d'utilisation et les incohérences détectées (le questionnaire SUS est fourni en annexe B). Le résultat final du questionnaire SUS est un score de satisfaction compris entre zéro et cent. Un score supérieur à 70 représente un bon résultat, un score compris entre 50 et 70 est considéré comme passable, et un score inférieur à 50 révèle de gros problèmes en termes de satisfaction de l'utilisateur.

La deuxième partie du questionnaire comporte trois questions. Dans la première question, les participants devaient répondre, sur une échelle de Likert en cinq points (allant de "pas du tout d'accord" à "tout à fait d'accord") si les alternatives proposées par l'outil pouvaient leur donner des idées auxquelles ils n'avaient pas pensé lors de la conception de l'IHM source. Les deux autres questions sont des questions ouvertes où les participants devaient donner leurs avis par rapport à l'outil et aux alternatives proposées.

4.1.4 IHM générées

La figure 4.14 présente la page d'accueil de la première IHM alternative. La transformation appliquée concerne la navigation globale au sein du site web. Cette transformation permet à l'utilisateur d'apercevoir toutes les actions possibles dans les différentes pages exposées dans la barre de navigation. Ces pages se chargent (s'affichent) à la demande de l'utilisateur. Par cette transformation, l'utilisateur peut directement rechercher un parcours ou demander la création d'un parcours sans avoir à naviguer vers la page parcours par exemple. Le coût articulatoire est moins important : l'utilisateur peut accéder directement à la création d'un parcours sans naviguer vers par la page parcours (moins de clics à effectuer pour savoir ce

qui existe dans la page parcours ou pour exécuter une sous-tâche).

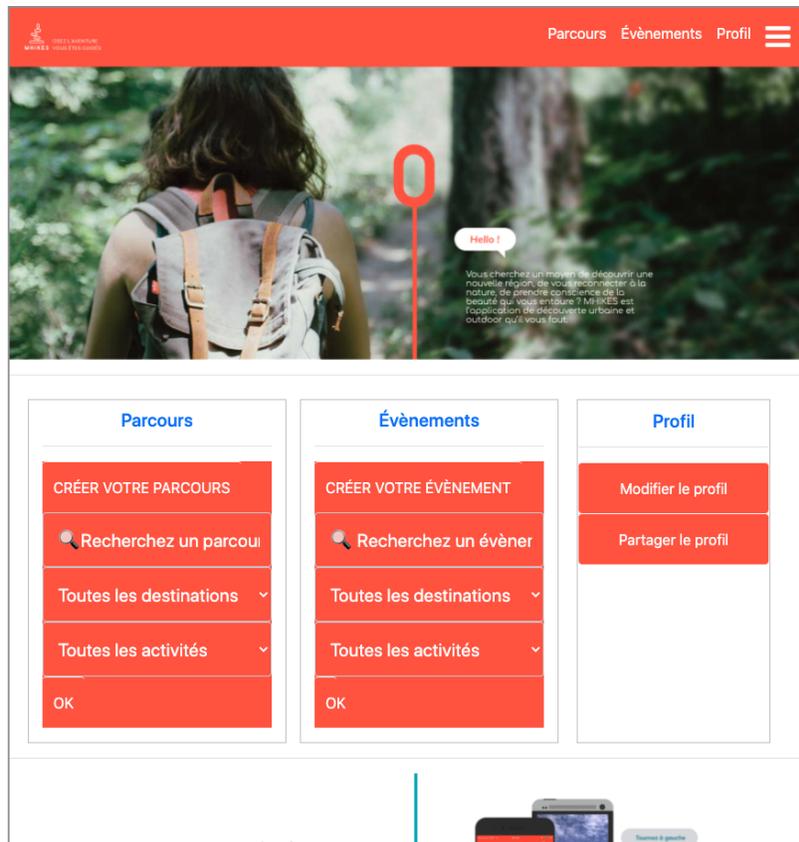


FIGURE 4.14 – IHM alternative 1 - Page d'accueil

La figure 4.15 présente la page parcours de la deuxième IHM alternative qui expose un changement au niveau de la navigation locale. Dans l'IHM de référence, le contenu des deux agents "créer parcours" et "rechercher parcours" est présenté dans deux pages différentes. Dans cette alternative, ces deux agents sont présentés dans la même page via un élément de type accordéon. L'avantage de ce type de transformation est qu'il permet à l'utilisateur de rester sur la même page lors de la création d'un parcours, de la recherche d'un parcours existant ou de la consultation de la liste des parcours.

L'avantage de l'élément de type accordéon par rapport à la navigation entre différentes pages est que tous les éléments sont affichés et traités dans un même espace de travail dès le début (chargement de la page), ce qui évite de naviguer entre les différentes pages en cas de besoin : la navigation se réalise dans un seul espace de travail et par conséquent le coût de la trajectoire d'interaction est moindre.

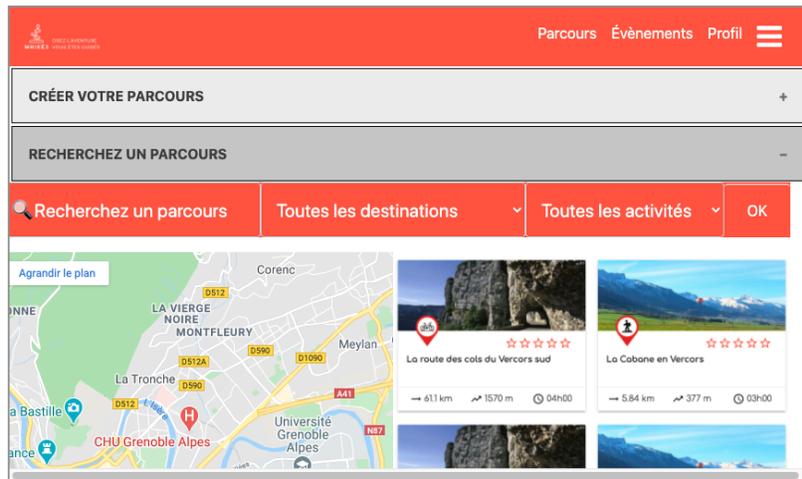


FIGURE 4.15 – IHM alternative 2 - Page parcours

4.1.5 Scénario d'interaction

La génération automatique de vidéos (principe #5) n'étant pas implémentée aujourd'hui, nous avons exposé aux participants en fin d'évaluation différentes vidéos (créées manuellement) présentant un scénario de création d'un parcours sur les différentes IHM (référence et alternatives). Nous avons présenté en premier la vidéo illustrant ce scénario sur l'IHM de référence, ensuite sur les trois IHM alternatives générées. Le scénario décrit les différentes étapes pour la création d'un parcours.

Nous présentons dans les figures 4.16 et 4.17 les différentes étapes de création d'un parcours respectivement sur l'IHM de référence et sur la deuxième IHM alternative (transformation de la navigation locale).

4.2 Résultats

Nous avons commencé par calculer le score SUS afin de mesurer l'expérience utilisateur en utilisant notre outil. Ensuite nous avons analysé les retours fournis par les utilisateurs. Ils consistent en des retours exprimés lors de la phase de test et aux retours rédigés dans la deuxième partie du questionnaire.

4.2.1 Score SUS

Nous avons calculé le score SUS à partir des réponses des participants dans le questionnaire. Ce score permet de déterminer le niveau de satisfaction des utilisateurs vis-à-vis de notre outil.

GRA:FITI a obtenu un score SUS de 81,25/100. Nous présentons les notes individuelles

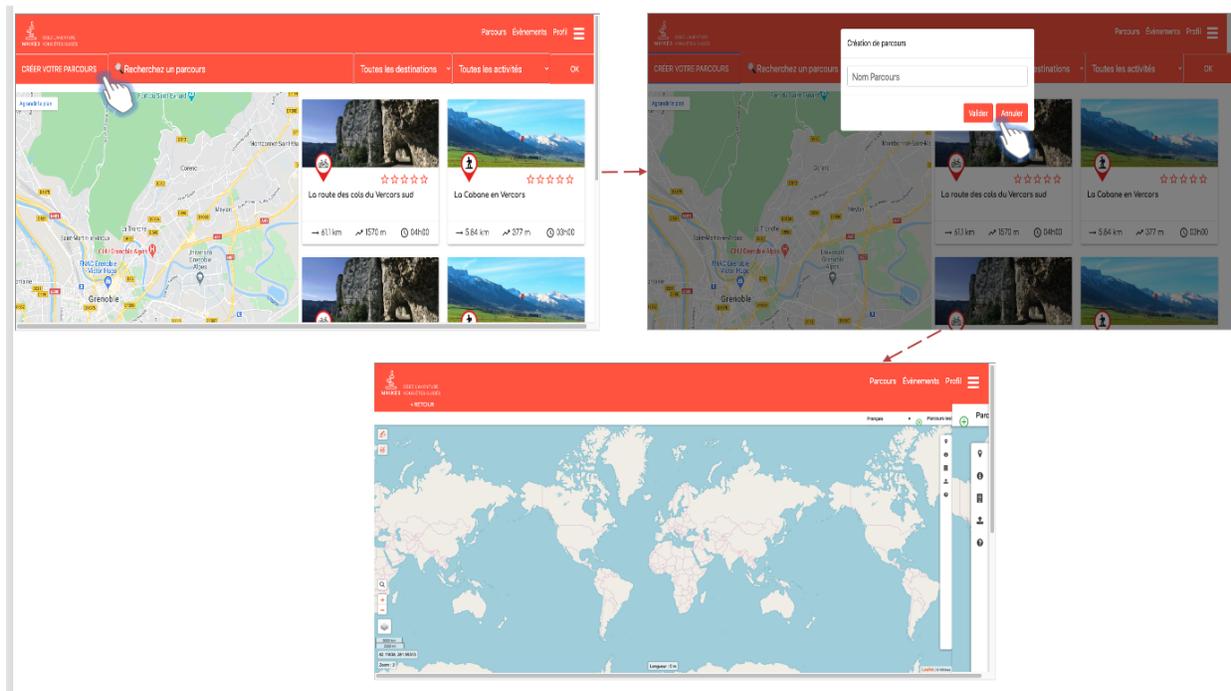


FIGURE 4.16 – Création d'un parcours - IHM de référence

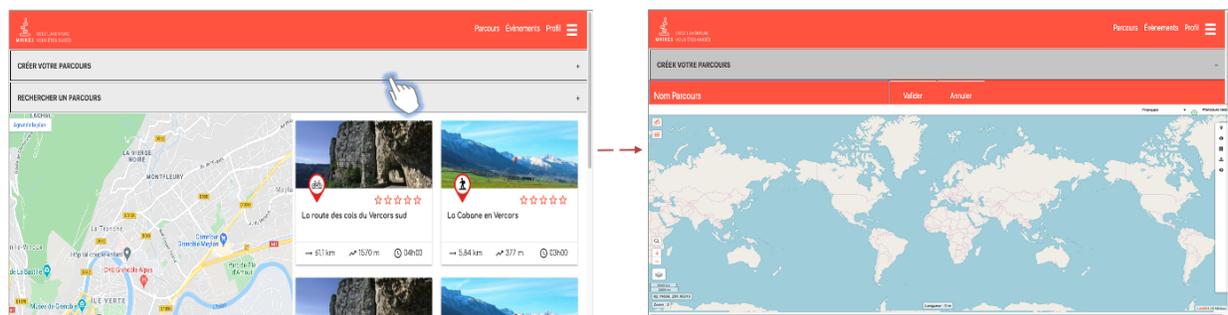


FIGURE 4.17 – Création d'un parcours - IHM alternative N°2

attribuées par chaque participant dans la figure 4.18. La méthode de calcul de ce score est présenté dans l'annexe B.

Ces résultats indiquent que les utilisateurs étaient satisfaits de l'utilisation de GRA:FITI et qu'il n'y a pas de problème majeur d'utilisation rencontré. Par exemple, les réponses à la première question "Je pense que je vais utiliser ce système fréquemment" sont plutôt positives avec 12 personnes qui ont répondu entre 4 et 5 (figure 4.19). Ces résultats sont confortés par plusieurs commentaires des participants exprimés lors de l'évaluation :

— "Utilisation facile, je n'ai seulement mon interface à mettre."

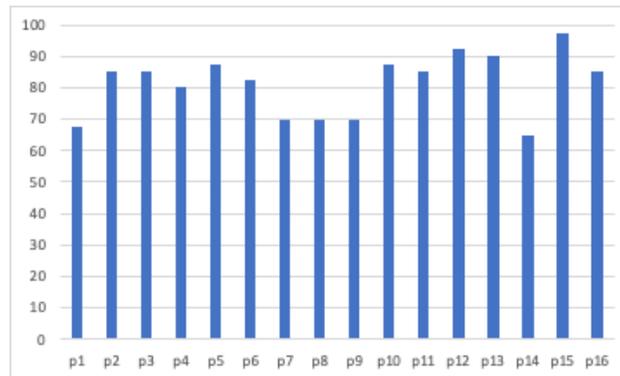


FIGURE 4.18 – Score SUS par participant

- "Il est facile à manipuler et utile dans son utilisation."
- "Tout est très clair, je n'ai pas besoin qu'on m'explique ce que je peux faire ou comment le faire."

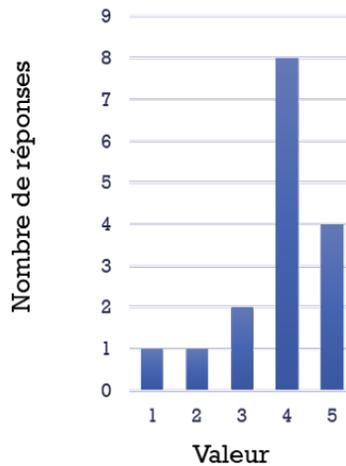


FIGURE 4.19 – Score SUS pour la première question

4.2.2 Notes obtenues pour l'utilité des IHM alternatives proposées par GRA:FITI

La deuxième partie du questionnaire comportait trois questions qualitatives portant sur l'utilité de GRA:FITI. Dans la première question, les participants devaient noter de un à cinq (1="Pas du tout d'accord" et 5="Tout à fait d'accord") si les IHM proposées par GRA:FITI pourraient les inspirer. Comme le montre la figure 4.20, 37,5% des participants ont accordé une note de 5/5 comme réponse à cette question, 50% des participants ont accordé une note

de 4/5 et 12,5% des participants ont accordé une note de 3/5. Aucun des participants n'a accordé une note de 1/5 ou de 2/5 à cette question.

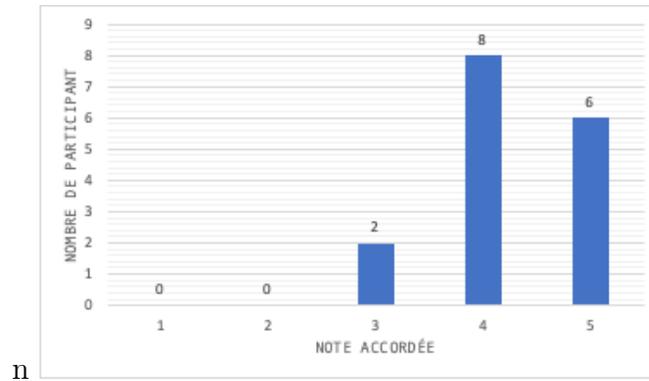


FIGURE 4.20 – Notes accordées pour l'utilité des IHM alternatives proposées par GRA:FITI

Les participants ont indiqué que les alternatives générées les ont aidés à réfléchir à différentes façons de présenter le contenu de l'IHM, autres que celle proposée dans l'IHM de référence. Ils ont indiqué que de telles propositions peuvent "*aider un concepteur à tester plusieurs façons de concevoir son IHM*" et que la comparaison de la référence avec les alternatives proposées les aide à mieux critiquer l'IHM de référence. Nous analysons par la suite ces retours.

Les deux dernières questions de notre questionnaire étaient des questions ouvertes. Les participants avaient la possibilité de donner leurs avis sur (1) notre outil GRA:FITI en général et (2) les alternatives proposées. Nous analysons leurs retours dans les sections 4.2.3 et 4.2.4

4.2.3 Retours des utilisateurs par rapport à GRA:FITI

Nous analysons dans cette section les réponses à la question : "Que pensez-vous globalement de l'outil?".

Globalement, les participants ont exprimé des retours favorables par rapport à l'outil. Ils apprécient le concept et encouragent son évolution et sa commercialisation. Par exemple, un des participants a indiqué que "*l'utilisation d'un tel outil pourrait minimiser le nombre d'itérations de conception en améliorant le processus d'évaluation*". Selon lui, montrer un ensemble d'alternatives avec une référence aide un client à fournir assez de retours permettant au concepteur de mieux comprendre ses besoins. Un autre participant a mentionné que l'utilisation d'un tel outil peut permettre à un concepteur débutant ou à un étudiant de concevoir une bonne IHM et que cet outil "*peut être un co-designer pour aider quelqu'un qui ne peut pas payer un professionnel*". Les participants trouvent que GRA:FITI peut être pratique pour explorer rapidement différentes propositions possibles de conception et pour

aider le concepteur à converger rapidement vers la bonne IHM. Nous rapportons dans ce qui suit d'autres retours positifs exprimés par les participants :

- *Cet outil permet de proposer facilement différentes alternatives au concepteur d'une IHM, notamment en terme de structure, ce qui permet d'apporter plus d'accessibilité au besoin.*
- *Il n'y a rien à perdre à l'utiliser, au contraire il peut nous donner des idées auxquelles nous nous n'avons pas pensé.*
- *Cet outil fournit des propositions auxquelles nous n'avons pas réfléchi, permet de mieux exploiter les parties de l'IHM et de mieux organiser ses composants.*

En résumé, les participants indiquent que GRA:FITI :

1. Peut être pratique pour explorer rapidement différentes propositions possibles de conception.
2. Permet de gagner du temps lors de la conception et aide à converger plus rapidement vers une IHM satisfaisante.
3. Peut fournir de nouvelles idées à l'utilisateur et améliorer les retours obtenus lors de l'évaluation d'une IHM de référence.

En outre, deux participants ont demandé à être plus intégrés dans le processus de génération des alternatives et demandent plus de clarification des transformations appliquées :

- *"Fait le job, mais manque un peu de flexibilité (pas configurable) et l'utilisateur n'a pas la main là-dessus."*
- *"J'aurais aimé que l'outil m'explique pourquoi ce changement (gain derrière chaque alternative)".*

Grâce aux retours sur cette première question ouverte, nous concluons que GRA:FITI peut être utile pour :

- Aider à fournir des retours utilisateurs pertinents et rapidement sur une IHM proposée ; ici GRA:FITI peut être utilisé pour la génération des alternatives pour des évaluations comparatives.
- Aider les concepteurs à explorer différentes possibilités de structuration de l'IHM et à trouver des pistes d'amélioration de leurs conceptions ; ici GRA:FITI peut être utilisé comme moyen d'inspiration et d'aide à la conception.

4.2.4 Retours des utilisateurs par rapport aux différentes IHM

Lors de l'évaluation, les participants ont été invités à exprimer leurs opinions et avis et à réfléchir à voix haute par rapport à l'IHM de référence et aux alternatives générées. Nous avons constaté que les alternatives proposées ont permis aux participants de réfléchir et de critiquer essentiellement la structure de l'IHM de référence et la navigation entre les différentes pages du site web. Par exemple, certains participants ont indiqué qu'ils préfèrent les alternatives proposées à l'IHM de référence. En argument, ils indiquent qu'il est important d'avoir une vue d'ensemble du site web depuis la page de d'accueil, ce qui manque à l'IHM de référence.

En outre, nous avons analysé les réponses à la question : "Que pensez-vous globalement des IHM proposées par l'outil?". Nous avons trouvé que les participants ont critiqué l'aspect esthétique des alternatives mais ils indiquent également souvent que les alternatives fournies sont utiles pour réfléchir à la structure de l'IHM et aux façons d'organiser les éléments dans les différents espaces de travail. Ceci est un aspect important pour la conception d'une bonne IHM. En outre, nous avons observé que les participants ont exprimé leur rejet quand une IHM ne les satisfait pas.

Nous présentons ci-dessous quelques retours fournis par les participants :

- *Très basique (esthétique/graphique), donnent une idée sur la structure mais pas le design (esthétique) : pas d'importance donnée aux couleurs.*
- *Comme proposition d'architecture, ils sont bien, mais le côté esthétique est pauvre.*
- *Aident les utilisateurs à mieux comprendre ce qu'ils peuvent faire/voir dans les différentes pages.*
- *Aident à tester plusieurs façon de concevoir un site web.*

5 Discussion

Globalement, les participants ont trouvé que GRA:FITI est un outil intéressant mais que des améliorations peuvent encore y être apportées. Les participants confirment l'utilité de GRA:FITI et des alternatives de conception proposées dans le but d'optimiser le processus de conception d'une IHM. Les participants ont cependant critiqué la qualité des alternatives générées notamment en termes d'esthétique. En outre, les participants ont fourni différentes idées et suggestions d'amélioration de GRA:FITI. Certains d'entre eux ont proposé par exemple de présenter les alternatives en fonction de la taille de l'écran (petit pour un smartphone ou moyen pour une tablette etc.) en plus de la version originale afin d'évaluer l'ajustement de l'affichage par rapport à la taille d'écran du terminal utilisé en fonction des transformations proposées. Enfin, nous avons remarqué que les alternatives générées ont particulièrement favorisé des discussions sur la structure de l'IHM de référence.

" Assez puissant dans l'idée, peut s'améliorer avec une composante de Machine Learning qui permettra un apprentissage automatique sur le design et sur la structure du site. "(Retour d'un participant).

GRA:FITI est un démonstrateur de concept dont le potentiel est ici confirmé. Il mérite toutefois une exploration plus approfondie et les suggestions d'améliorations obtenues permettront d'améliorer cette première version.

Conclusion

Dans ce chapitre, nous avons présenté notre approche de génération des IHM alternatives à partir d'une IHM de référence. Ensuite nous avons présenté GRA:FITI, notre démonstrateur de concept confirmant, d'une part, la faisabilité technique de l'approche et, d'autre part, son potentiel.

Conclusion générale et perspectives

Cette thèse contribue au domaine de l'interaction Homme-Machine. Ce domaine est complexe par son caractère pluridisciplinaire. En pratique, la complexité est maîtrisée par des processus itératifs centrés utilisateurs. Dans ces processus, l'évaluation est une étape cruciale.

Cette thèse se focalise sur l'évaluation comparative d'IHM. L'étude de l'état de l'art a en effet montré que l'évaluation comparative est une approche bénéfique pour la qualité des IHM. Toutefois, le choix des alternatives pour optimiser l'évaluation n'est pas documenté. Nous nous sommes intéressés à la génération automatique de la meilleure IHM alternative pour améliorer la qualité de l'IHM tout en réduisant son coût de production.

1 Rappel des contributions

Notre travail s'est focalisé sur trois points principaux :

1. Définir la meilleure IHM alternative, c'est-à-dire caractériser les IHM alternatives possibles et déterminer celles permettant d'optimiser l'évaluation comparative.
2. Produire cette IHM alternative, c'est-à-dire définir une méthode de génération d'IHM.
3. Outiller la méthode, c'est-à-dire s'inscrire dans une démarche instrumentée pour réduire le coût de convergence vers l'IHM finale.

Les choix réalisés ont été dirigés par le pragmatisme au regard des pratiques et des coûts de mise en œuvre. Nous démontrons que l'ingénierie dirigée par les modèles offre un cadre adéquat pour obtenir la meilleure IHM alternative par transformation de l'IHM de référence. Nous en montrons la faisabilité technique par un démonstrateur de concept.

1.1 Définir la meilleure IHM alternative

L'étude de l'état de l'art montre que l'évaluation comparative d'IHM est une approche vertueuse. En particulier, Tohidi et al. concluent qu'elle :

- Favorise la critique négative,
- Permet l'élimination d'IHM non appréciées.

En revanche, le choix des IHM alternatives n'y est pas discuté.

Complémentairement, Kang et al. montrent qu'une catégorisation des IHM alternatives est possible (Kang et al. 2018). Ils distinguent deux types de similarité : d'une part, de contenu ; d'autre part, de visuel. En revanche, aucune recommandation n'y est formulée. Par ailleurs, seules les interfaces y sont étudiées, sans considération de l'interaction.

De cette analyse, nous formulons quatre principes directeurs :

Issus de l'état de l'art :

- #1 – L'évaluation comparative est une approche intéressante.
- #2 – Il est possible de caractériser les différences entre l'IHM de référence et l'IHM alternative.

Pragmatiques :

- #3 – L'IHM alternative doit être de meilleure qualité que l'IHM de référence.
- #4 – L'IHM alternative doit être obtenue à faible coût.

1.2 Production de l'IHM alternative

L'état de l'art est florissant dans la diversité d'approches étudiées. Nous les étudions selon le continuum de l'automatisation. Les processus manuels, ancrés dans les techniques de créativité consistant à combiner des phases de divergence et de convergence, présentent l'avantage de travailler sur des croquis concrets, plaçant les concepteurs à un niveau d'abstraction qui leur est familier. En revanche, la maîtrise du temps y est une difficulté majeure, la convergence pouvant être longue.

Les conceptions assistées, sur le principe de galeries d'exemples, se démontrent bénéfiques comme sources d'inspiration. Toutefois, aucune aide n'y est fournie pour le choix des IHM à retenir et l'espace de conception exploré y reste limité.

Les approches automatiques, telles que pratiquées par exemple dans les algorithmes génétiques interactifs, sont intéressantes par la vaste exploration rapide de l'espace de conception qu'elles permettent. Cependant aucune caractérisation des mutations n'existe pour orienter

l'exploration vers telle ou telle mutation à privilégier. Deux caractérisations de variations sont proposées par O'Donovan et al. (O'Donovan et al. 2015) : ils distinguent, d'une part, les améliorations de mise en page ; d'autre part, les réorganisations de pages exploitant ainsi le caractère majeur versus mineur des variations. Toutefois, le travail se limite à des interfaces, sans considération de l'interaction. Enfin, l'ingénierie dirigée par les modèles, telle qu'explorée dans le processus CAMELEON, offre un cadre de génération automatique des IHM, assurant par construction une caractérisation des variations selon le niveau d'abstraction auxquelles elles s'opèrent.

De cette analyse, nous formulons deux principes directeurs :

- #1 – Utiliser le cadre de référence CAMELEON.
- #2 – Ancrer la transformation de l'IHM de référence au niveau d'abstraction optimisant l'évaluation comparative.

1.3 L'obtenir par transformation de l'IHM de référence

Nous relatons trois études expérimentales complémentaires.

La première étude vise à répondre à la question : peut-on éviter de transformer l'IHM ? La voie explorée consiste à n'utiliser que l'IHM de référence, mais en la présentant de deux manières différentes : d'une part, en version papier ; d'autre part, en version vidéo. L'étude montre que le support impacte les résultats et que les supports sont complémentaires : l'utilisation combinée d'un prototype papier et d'un prototype vidéo permet de recueillir un plus grand nombre de retours utilisateur sur l'utilité, l'utilisabilité et l'esthétique : multiplier les supports est donc intéressant. Toutefois, ça reste insuffisant, les retours étant limités à l'IHM de référence.

La deuxième étude vise à répondre à la question : à quel niveau d'abstraction transformer alors l'IHM ? Nous montrons que le niveau d'abstraction impacte la quantité et la qualité des retours utilisateurs et que l'IHM abstraite est le niveau le plus productif. L'écart est notamment important par rapport à l'IHM concrète.

La troisième étude vise alors à répondre à la question : alors quelle transformation de l'IHM abstraite ? Nous montrons que le type de transformation impacte le nombre de retours utilisateur et que la transformation de la navigation est la plus productive.

De ces trois études, nous formulons deux principes directeurs :

- #1 – Transformer la navigation de l'IHM de référence.
- #2 – Produire une vidéo.

1.4 Démontrer la faisabilité technique

A la lumière des enseignements précédents, les principes directeurs ayant présidé à la conception de GRA:FITI sont les suivants :

- #1 – Laisser les praticiens travailler au niveau de l’IHM concrète.
- #2 – Permettre aux utilisateurs d’interagir avec le prototype.
- #3 – Utiliser l’architecture logicielle comme pivot pour la transformation de l’IHM de référence.
- #4 – Réduire le coût articulatoire.
- #5 – Permettre à l’utilisateur d’interagir avec l’IHM de référence et générer automatiquement les vidéos IHM de référence et alternatives.

L’évaluation de GRA:FITI est positive, avec un score SUS moyen de 81,25. Les retours sont encourageants :

- **Sur l’outil** : en points positifs, il apporte un support efficace pour l’évaluation comparative et est source d’inspiration pour la structuration de l’IHM ; en point de progrès, son manque d’interactivité est regretté ;
- **Sur les IHM générées** : l’incitation faite au concepteur de considérer l’interaction est bien appréciée ; en revanche, l’absence de considération de l’esthétique est regrettée.

1.5 Synthèse

Le tableau 4.2 apporte une vue intégrée des principes directeurs formulés.

Les contributions sont in fine de trois ordres :

- **Contributions à l’évaluation comparative**
 - #1 Caractérisation de l’IHM alternative optimale : variation de l’IHM abstraite dans sa navigation – utilisation du cadre de référence CAMELEON en **plasticité**.
 - #2 Définition d’une méthode de production de cette IHM : utilisation de l’architecture logicielle PAC.
 - #3 Outillage : GRA:FITI.
- **Contribution à la conception d’IHM**
 - #1 Génération d’IHM alternatives de meilleure qualité : source d’inspiration en **créativité**.
- **Contribution à l’évaluation d’IHM**
 - #1 Papier et vidéo pour couvrir utilité et utilisabilité.
 - #2 Génération automatique de vidéo.

La figure 4.21 situe ces contributions sur les processus de développement d’IHM.

En limites au travail, nous pointons les trois faiblesses suivantes :

- Faible diversité d’IHM de référence et donc de qualité d’IHM de référence.

Démarche	Principes directeurs
Définir la meilleure IHM alternative	#1 - L'évaluation comparative est une approche intéressante #2 - Il est possible de caractériser les différences entre l'IHM de référence et l'IHM alternative #3 - L'IHM alternative doit être de meilleure qualité que l'IHM de référence #4 - L'IHM alternative doit être obtenue à faible coût
Produire cette IHM alternative	#1 - Utiliser le cadre de référence CAMELEON. #2 - Ancrer la transformation de l'IHM de référence au niveau d'abstraction optimisant l'évaluation comparative
L'obtenir par transformation de l'IHM de référence	#1 - Transformer la navigation de l'IHM de référence. #2 - Produire une vidéo.
Démonstrateur de concept	#1 - Laisser les praticiens travailler au niveau de l'IHM concrète. #2 - Permettre aux utilisateurs d'interagir avec le prototype. #3 - Utiliser l'architecture logicielle comme pivot pour la transformation de l'IHM de référence. #4 - Réduire le coût articuloire. #5 - Permettre à l'utilisateur d'interagir avec l'IHM de référence et générer automatiquement les vidéos IHM de référence et alternatives.

TABLE 4.2 – Tableau comparatif.

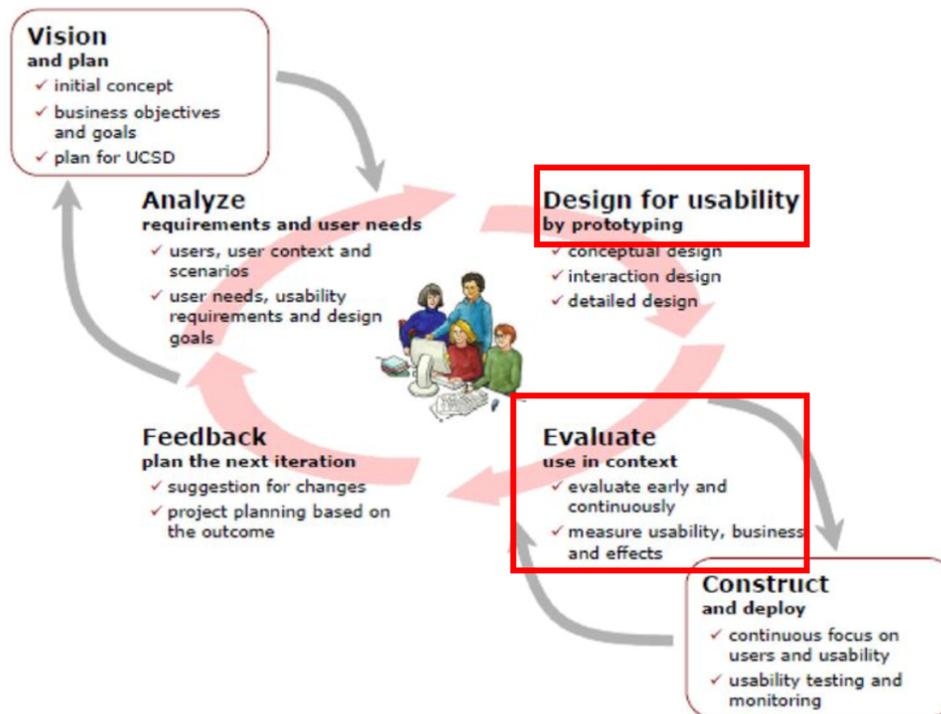


FIGURE 4.21 – Processus de développement d’IHM (Gulliksen et al. 2003).

- IHM graphiques seulement.
- Pas de qualification de l’importance des retours utilisateur.

Publications

Les contributions de cette thèse ont fait l’objet des publications suivantes :

- Hammami, H., Calvary, G., Riahi, M., Moussa, F. et Bouzit, S. (2017). “Comparative Evaluation? Yes, But With Which Alternative UI?” In proceedings of British HCI 2017 (p. 1-7).
- Hammami, H., Camara, F., Calvary, G., Riahi, M. et Moussa, F. (2020). “The Benefits of Combining Paper-and Video-Based Prototypes for User Interface Evaluation”. In : The Thirteen International Conference on Advances in Computer-Human Interactions.

2 Perspectives

En perspectives, nous souhaitons à court terme étudier en profondeur les bénéfices de l’utilisation de GRA:FITI dans le processus de conception des IHM. Deux indicateurs seront

à étudier : d'une part, le temps du processus complet (évaluation quantitative) et, d'autre part la qualité de l'IHM finale (évaluation qualitative). Cette évaluation sera faite avec et sans GRA:FITI, en faisant varier le moment auquel GRA:FITI est introduit (ou non), le degré d'expertise du concepteur et la complexité de l'application.

A moyen terme, des extensions sont envisagées pour améliorer la qualité des IHM générées :

- Impliquer le concepteur dans le processus de génération des alternatives : par exemple, lui donner la possibilité d'éditer l'IHM de référence en temps réel tout en empruntant des éléments des différentes alternatives proposées,
- Améliorer l'esthétique, par des méthodes d'apprentissage automatique pour améliorer le style de présentation des éléments transformés.
- Améliorer la qualité des IHM produites en travaillant leur esthétique et leur plasticité.
- Offrir un support à l'évaluation par la reproduction automatique des scénarios d'interaction sur les IHM alternatives.

Enfin, à long terme, une piste prometteuse est une plateforme d'évaluation comparative. L'objectif ici est de passer à l'échelle, de s'ouvrir au *crowdsourcing* pour permettre aux concepteurs de collecter rapidement et à faible coût des retours utilisateur nombreux, structurés et exploitables.

Bibliographie

- AB Tasty (28 oct. 2020). AB Tasty Plateform. URL : <https://www.abtasty.com/fr/>.
- Albert, William et Thomas Tullis (2013). *Measuring the user experience : collecting, analyzing, and presenting usability metrics*. Second Edition. Interactive Technologies. Boston : Morgan Kaufmann, p. 1-14.
- Ammar, Lassaad Ben (2019). "Towards a Uniform Model Transformation Process for Abstract User Interfaces Generation." In : *In Proceedings of the 14th International Conference on Evaluation of Novel Approaches to Software Engineering (ENASE 2019)*, p. 533-538.
- Assila, Ahlem (2016). "Une approche et un outil pour l'évaluation subjective et objective de l'utilisabilité des systèmes interactifs : application à un système d'aide à l'information voyageurs". Thèse de doctorat en informatique. Université de Valenciennes et du Hainaut-Cambrésis et Université de Sfax.
- Balbo, Sandrine (1994). "Evaluation ergonomique des interfaces utilisateur : un pas vers l'automatisation". Thèse de doctorat en informatique. Université, Joseph-Fourier-Grenoble I.
- Balsamiq Wireframes, *Quick and Easy Wireframing Tool* (Retrieved : December, 2019). URL : <https://balsamiq.com/>.
- Bastien, J. et Dominique Scapin (avr. 1992). "Ergonomic Criteria for the Evaluation of Human-Computer Interfaces • Critères Ergonomiques pour l'Évaluation d'Interfaces Utilisateurs". In : *Int. J. Hum. Comput. Interaction* 4, p. 183-196.
- Bonnardel, Nathalie (1999). "Creativity in design activities : The role of analogies in a constrained cognitive environment". In : *Proceedings of the 3rd conference on Creativity & cognition*, p. 158-165.
- Brooke, John (2013). "SUS : a retrospective". In : *Journal of usability studies* 8.2, p. 29-40.
- Buxton, Bill (2010). *Sketching user experiences : getting the design right and the right design*. Morgan kaufmann.

- Calvary, Gaëlle, Joëlle Coutaz, Olfa Dâassi, Lionel Balme et Alexandre Demeure (2004). “Towards a new generation of widgets for supporting software plasticity : the comet”. In : *IFIP International Conference on Engineering for Human-Computer Interaction*. Springer, p. 306-324.
- Calvary, Gaëlle, Joëlle Coutaz et Laurence Nigay (1997). “From Single-User Architectural Design to PAC* : A Generic Software Architecture Model for CSCW”. In : *Proceedings of the ACM SIGCHI Conference on Human Factors in Computing Systems*. CHI '97. Atlanta, Georgia, USA : Association for Computing Machinery, 242–249. DOI : [10.1145/258549.258717](https://doi.org/10.1145/258549.258717). URL : <https://doi.org/10.1145/258549.258717>.
- Calvary, Gaëlle, Joëlle Coutaz, David Thevenin, Q. Limbourg, N. Souchon, L. Bouillon et Jean Vanderdonckt (2002). “Plasticity of User Interfaces : A revised reference framework”. In : *First International Workshop on Task Models and Diagrams for User Interface Design TAMODIA '2002, Bucarest*, p. 127-134.
- Calvary, Gaëlle, Joëlle Coutaz, David Thevenin, Quentin Limbourg, Laurent Bouillon et Jean Vanderdonckt (2003). “A unifying reference framework for multi-target user interfaces”. In : *Interacting with computers* 15.3, p. 289-308.
- Coutaz, Joëlle (1987). “PAC, an Object Oriented Model for Dialog Design”. In : *Human-Computer Interaction-INTERACT*. Amsterdam : North-Holland, p. 431-436.
- Coutaz, Joëlle, Gaëlle Calvary, Alexandre Demeure et Lionel Balme (juil. 2012). *Systèmes interactifs et adaptation centrée utilisateur : la plasticité des Interfaces Homme-Machine*.
- Coyette, Adrien, Suzanne Kieffer et Jean Vanderdonckt (2007). “Multi-fidelity prototyping of user interfaces”. In : *IFIP Conference on Human-Computer Interaction*. Springer, p. 150-164.
- Daassi, Olfa (2007). “Les comets : une nouvelle génération d’Interacteurs pour la Plasticité des Interfaces Homme-Machine”. Thèse de doctorat en informatique. Université Joseph-Fourier Grenoble I.
- Degrigny, Frantz et Thierry Duval (jan. 2004). “Utilisation du modèle PAC-Amodeus pour une réutilisation optimale de code dans le développement de plusieurs versions d’un logiciel commercial”. In : p. 149-156. DOI : [10.1145/1148613.1148634](https://doi.org/10.1145/1148613.1148634).
- Dhillon, Beant, Peter Banach, Rafal Kocielnik, Jorge Peregrín Empananza, Ioannis Politis, A Rczewska et Panos Markopoulos (2011). “Visual fidelity of video prototypes and user feedback : a case study”. In : *Proceedings of HCI 2011 The 25th BCS Conference on Human Computer Interaction*, p. 139-144.
- Dicks, R Stanley (2002). “Mis-usability : on the uses and misuses of usability testing”. In : *Proceedings of the 20th annual international conference on Computer documentation*, p. 26-30.
- Dow, Steven, Julie Fortuna, Dan Schwartz, Beth Altringer, Daniel Schwartz et Scott Klemmer (2011). “Prototyping dynamics : sharing multiple designs improves exploration, group

- rapport, and results”. In : *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, p. 2807-2816.
- Dow, Steven P, Alana Glassco, Jonathan Kass, Melissa Schwarz, Daniel L Schwartz et Scott R Klemmer (2010). “Parallel prototyping leads to better design results, more divergence, and increased self-efficacy”. In : *ACM Transactions on Computer-Human Interaction (TOCHI)* 17.4, p. 1-24.
- Gardey, Juan Cruz et Alejandra Garrido (2020). “User Experience Evaluation through Automatic A/B Testing”. In : *Proceedings of the 25th International Conference on Intelligent User Interfaces Companion*, p. 25-26.
- Google Analytics* (28 oct. 2020). Google Analytics. URL : <https://marketingplatform.google.com/about/analytics/>.
- Guide complet de l’A/B testing* (28 mar. 2019). URL : <https://www.abtasty.com/fr/ab-testing/>.
- Gulliksen, Jan, Bengt Göransson, Inger Boivie, Stefan Blomkvist, Jenny Persson et Åsa Cajander (2003). “Key principles for user-centred systems design”. In : *Behaviour and Information Technology* 22.6, p. 397-409.
- Hammami, Hayet, Gaëlle Calvary, Meriem Riahi, Faouiz Moussa et Sara Bouzit (2017). “Comparative Evaluation ? Yes, But With Which Alternative UI ?” In : *In proceedings of British HCI 2017*, p. 1-7.
- Hammami, Hayet, Fatoumata Camara, Gaëlle Calvary, Meriem Riahi et Faouzi Moussa (2020). “The Benefits of Combining Paper-and Video-Based Prototypes for User Interface Evaluation”. In : *The Thirteen International Conference on Advances in Computer-Human Interactions*.
- Hartson, H Rex et Deborah Hix (1989). “Human-computer interface development : concepts and systems for its management”. In : *ACM Computing Surveys (CSUR)* 21.1, p. 5-92.
- Hartson, H.Rex, Terence S Andre et Robert C Williges (2001). “Criteria for evaluating usability evaluation methods”. In : *International journal of human-computer interaction* 13.4, p. 373-410.
- Herring, Scarlett R., Chia-Chen Chang, Jesse Krantzler et Brian P. Bailey (2009). “Getting Inspired! Understanding How and Why Examples Are Used in Creative Design Practice”. In : *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems Design*. CHI ’09. Boston, MA, USA : Association for Computing Machinery, 87–96. DOI : [10.1145/1518701.1518717](https://doi.org/10.1145/1518701.1518717). URL : <https://doi.org/10.1145/1518701.1518717>.
- Heslin, Peter A (2009). “Better than brainstorming? Potential contextual boundary conditions to brainwriting for idea generation in organizations”. In : *Journal of Occupational and Organizational Psychology* 82.1, p. 129-145.
- Houde, Stephanie et Charles Hill (1997). “What do prototypes prototype?” In : *Handbook of human-computer interaction*. Elsevier, p. 367-381.

- Ivory, Melody Y et Marti A Hearst (2001). “The state of the art in automating usability evaluation of user interfaces”. In : *ACM Computing Surveys (CSUR)*, p. 470-516.
- Johnson, Ralph E. (oct. 1997). “Frameworks = (Components + Patterns)”. In : *Commun. ACM* 40.10, 39–42. DOI : [10.1145/262793.262799](https://doi.org/10.1145/262793.262799). URL : <https://doi.org/10.1145/262793.262799>.
- Kang, Hyeonsu B, Gabriel Amoako, Neil Sengupta et Steven P. Dow (2018). “Paragon : An Online Gallery for Enhancing Design Feedback with Visual Examples”. In : *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, p. 1-13.
- Kieras, David E. et Anthony J. Hornof (2014). “Towards Accurate and Practical Predictive Models of Active-Vision-Based Visual Search”. In : *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI '14. Toronto, Ontario, Canada : Association for Computing Machinery, 3875–3884. DOI : [10.1145/2556288.2557324](https://doi.org/10.1145/2556288.2557324). URL : <https://doi.org/10.1145/2556288.2557324>.
- Kim, Huhn et Wan Chul Yoon (2005). “Supporting the cognitive process of user interface design with reusable design cases”. In : *International journal of human-computer studies* 62.4, p. 457-486.
- Kohavi, Ron et Roger Longbotham (2017). “Online Controlled Experiments and A/B Testing.” In : *Encyclopedia of machine learning and data mining* 7.8, p. 922-929.
- Kolski, Christophe et Martial GRISLIN (1995). “Méthodes d'évaluation des interfaces homme-machine, utilisation a priori et/ou a posteriori”. In : *Journée Supervision et Coopération Homme-Machine*.
- Kurosu, Masaaki (2015). “Human-Computer Interaction : Design and Evaluation”. In : *Proceedings of the 17th International Conference HCI International 2015 Los Angeles USA August 2-7 Part I*.
- Lee, Brian, Savil Srivastava, Ranjitha Kumar, Ronen Brafman et Scott R. Klemmer (2010). “Designing with Interactive Example Galleries”. In : *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* 1. CHI '10. Atlanta, Georgia, USA : Association for Computing Machinery, 2257–2266. DOI : [10.1145/1753326.1753667](https://doi.org/10.1145/1753326.1753667). URL : <https://doi.org/10.1145/1753326.1753667>.
- Leiva, Germán et Michel Beaudouin-Lafon (2018). “Montage : A video prototyping system to reduce re-shooting and increase re-usability”. In : *Proceedings of the 31st Annual ACM Symposium on User Interface Software and Technology*, p. 675-682.
- Limbourg, Quentin, Jean Vanderdonckt, Benjamin Michotte, Laurent Bouillon et Víctor López-Jaquero (2004). “USIXML : a language supporting multi-path development of user interfaces”. In : *IFIP International Conference on Engineering for Human-Computer Interaction*. Springer, p. 200-220.
- Linsey, Julie S et Blake Becker (2011). “Effectiveness of brainwriting techniques : comparing nominal groups to real teams”. In : *Design creativity 2010*. Springer, p. 165-171.

- Luther, Kurt, Jari-Lee Tolentino, Wei Wu, Amy Pavel, Brian P Bailey, Maneesh Agrawala, Björn Hartmann et Steven P Dow (2015). “Structuring, aggregating, and evaluating crowdsourced design critique”. In : *Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing*, p. 473-485.
- MacDonald, Craig M et Michael E Atwood (2014). “What does it mean for a system to be useful? An exploratory study of usefulness”. In : *Proceedings of the 2014 conference on Designing interactive systems*, p. 885-894.
- Mackay, Wendy E (1988). “Video Prototyping : a technique for developing hypermedia systems”. In : *CHI’88 Conference Companion Human Factors in Computing Systems*. T. 5. Citeseer, p. 1-3.
- Mackay, Wendy E, Anne V Ratzler et Paul Janecek (2000). “Video artifacts for design : Bridging the gap between abstraction and detail”. In : *Proceedings of the 3rd conference on Designing interactive systems : processes, practices, methods, and techniques*, p. 72-82.
- Mackay, Wendy.E (2002). “Using video to support interaction design”. In : *DVD Tutorial, The ACM CHI Conference*.
- Marsh, Richard L, Joshua D Landau et Jason L Hicks (1996). “How examples may (and may not) constrain creativity”. In : *Memory & cognition* 24.5, p. 669-680.
- Masson, Dimitri (2014). “Inspirez ! Explorez ! Soutien à la créativité en conception d’interfaces homme-machine”. Thèse de doctorat en informatique.
- Masson, Dimitri, Alexandre Demeure et Gaëlle Calvary (2010). “Magellan, an Evolutionary System to Foster User Interface Design Creativity”. In : EICS ’10. Berlin, Germany : Association for Computing Machinery, 87–92. DOI : [10.1145/1822018.1822032](https://doi.org/10.1145/1822018.1822032). URL : <https://doi.org/10.1145/1822018.1822032>.
- Masson, Dimitri, Alexandre Demeure et Gaëlle Calvary (2011). “Inspirer, explorer : Magellan, un environnement interactif evolutionnaire pour la génération d’exemples”. In : *Proceedings of the 23rd Conference on l’Interaction Homme-Machine*, p. 1-7.
- Millot, Patrick (1987). “Coopération homme-machine dans les tâches de supervision des procédés automatisés”. Thèse de doctorat en informatique. Université de Valenciennes.
- Mora, S. L (2015). *Human-Computer Interaction .Creating and comparing alternatives*. URL : <http://desarrolloweb.dlsi.ua.es/cursos/2015/hci/creating-comparing-alternatives>.
- Nielsen, Jakob (1994). “Usability inspection methods”. In : *Conference companion on Human factors in computing systems*, p. 413-414.
- Nielsen, Jakob et Jan Maurits Faber (1996). “Improving system usability through parallel design”. In : *Computer*, p. 29-35. URL : <https://www.computer.org/csdl/magazine/co/1996/02/r2029/13rRUxD9h0x>.
- Nielsen, Jakob et Thomas K. Landauer (1993). “A Mathematical Model of the Finding of Usability Problems”. In : *Proceedings of the INTERACT ’93 and CHI ’93 Conference on Human Factors in Computing Systems*. CHI ’93. Amsterdam, The Netherlands : As-

- sociation for Computing Machinery, 206–213. DOI : [10.1145/169059.169166](https://doi.org/10.1145/169059.169166). URL : <https://doi.org/10.1145/169059.169166>.
- Nigay, Laurence (1994). “Conception et modélisation logicielles des systèmes interactifs : application aux interfaces multimodales”. Thèse de doctorat en informatique. Université Joseph-Fourier-Grenoble I.
- Nishiuchi, Nobuyuki et Yutaka Takahashi (2015). “Objective evaluation method of usability using parameters of user’s fingertip movement”. In : *Transactions on Computational Science XXV*. Springer, p. 77-89.
- O’Donovan, Peter, Aseem Agarwala et Aaron Hertzmann (2015). “Designscape : Design with interactive layout suggestions”. In : *Proceedings of the 33rd annual ACM conference on human factors in computing systems*, p. 1221-1224.
- Paternò, Fabio (2004). “ConcurTaskTrees : an engineered notation for task models”. In : *The handbook of task analysis for human-computer interaction*, p. 483-503.
- Qu’est-ce que l’A/B testing, Guide complet pour tout savoir sur l’A/B testing*. (28 mar. 2019). URL : <https://www.kameleoon.com/fr/ab-testing>.
- Ritchie, Daniel, Ankita Arvind Kejriwal et Scott R. Klemmer (2011). “D.Tour : Style-Based Exploration of Design Example Galleries”. In : *Proceedings of the 24th Annual ACM Symposium on User Interface Software and Technology*. UIST ’11. Santa Barbara, California, USA : Association for Computing Machinery, 165–174. DOI : [10.1145/2047196.2047216](https://doi.org/10.1145/2047196.2047216). URL : <https://doi.org/10.1145/2047196.2047216>.
- Robb, David A, Stefano Padilla, Britta Kalkreuter et Mike J Chantler (2015). “Crowdsourced feedback with imagery rather than text : Would designers use it?” In : *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, p. 1355-1364.
- Robb, David A, Stefano Padilla, Thomas S Methven, Britta Kalkreuter et Mike J Chantler (2017). “Image-based Emotion Feedback : How Does the Crowd Feel? And Why?” In : *Proceedings of the 2017 Conference on Designing Interactive Systems*, p. 451-463.
- Rutter, Richard, Patrick H Lauke, Cynthia Waddell, Jim Thatcher, Shawn Lawton Henry, Bruce Lawson, Andrew Kirkpatrick, Christian Heilmann, Michael R Burks, Bob Regan et al. (2007). *Web accessibility : Web standards and regulatory compliance*. Apress.
- Sadler, D Royce (1989). “Formative assessment and the design of instructional systems”. In : *Instructional science* 18.2, p. 119-144.
- Schütze, Martina, Pierre Sachse et Anne Römer (2003). “Support value of sketching in the design process”. In : *Research in Engineering Design* 14.2, p. 89-97.
- Shneiderman, Ben, Catherine Plaisant, Maxine Cohen, Steven Jacobs, Niklas Elmqvist et Nicholas Diakopoulos (2016). *Designing the user interface : strategies for effective human-computer interaction*. 3rd. USA : Addison-Wesley Longman Publishing Co., Inc.
- Smith, Steven M, Thomas B Ward et Jay S Schumacher (1993). “Constraining effects of examples in a creative generation task”. In : *Memory & cognition* 21.6, p. 837-845.

- Sottet, Jean-Sébastien, Gaëlle Calvary et Jean Marie Favre (2005). “Ingénierie de l’Interaction Homme-Machine Dirigée par les modèles”. In : *IDM’05 Premières Journées sur l’Ingénierie Dirigée par les Modèles*.
- Speicher, Maximilian, Andreas Both et Martin Gaedke (2014). “Ensuring web interface quality through usability-based split testing”. In : *International conference on web engineering*. Springer, p. 93-110.
- Super-easy tool to brainstorm UI ideas* (Retrieved : December, 2019). URL : <https://www.mockflow.com/>.
- Taylor, Donald W, Paul C Berry et Clifford H Block (1958). “Does group participation when using brainstorming facilitate or inhibit creative thinking?” In : *Administrative Science Quarterly*, p. 23-47.
- Thevenin, David, Gaëlle Calvary et Joëlle Coutaz (1999). “La plasticité en interaction homme-machine”. In : *Actes IHM99, Actes de 11ème Conférence Francophone Interaction Homme-Machine (Montpellier, 22-26 novembre)*, p. 152-155.
- Todi, Kashyap, Daryl Weir et Antti Oulasvirta (2016). “Sketchplore : Sketch and Explore with a Layout Optimiser”. In : *Proceedings of the 2016 ACM Conference on Designing Interactive Systems*. DIS ’16. Brisbane, QLD, Australia : Association for Computing Machinery, 543–555. DOI : [10.1145/2901790.2901817](https://doi.org/10.1145/2901790.2901817). URL : <https://doi.org/10.1145/2901790.2901817>.
- Tohidi, Maryam, William Buxton, Ronald Baecker et A Sellen (2006a). “User sketches : a quick, inexpensive, and effective way to elicit more reflective user feedback”. In : *Proceedings of the 4th Nordic conference on Human-computer interaction : changing roles*, p. 105-114.
- Tohidi, Maryam, William Buxton, Ronald Baecker et Abigail Sellen (2006b). “Getting the right design and the design right”. In : *Proceedings of the SIGCHI conference on Human Factors in computing systems*, p. 1243-1252.
- Tran, Vi, Jean Vanderdonckt, Ricardo Tesoriero et François Beuvens (2012). “Systematic generation of abstract user interfaces”. In : *Proceedings of the 4th ACM SIGCHI symposium on Engineering interactive computing systems*, p. 101-110.
- Une plateforme d’A/B testing et de personnalisation, web et full stack* (28 oct. 2020). Kameleoon. URL : <https://www.kameleoon.com/en>.
- Vanderdonckt, Jean (1994). *Guide ergonomique des interfaces homme-machine : guide pratique pour la conception ergonomique des interfaces homme-machine pour des applications interactives*. Presses Universitaires de Namur.
- Vanderdonckt, Jean, Mathieu Zen et Radu-Daniel Vatavu (2019). “AB4Web : An on-line A/B tester for comparing user interface design alternatives”. In : *Proceedings of the ACM on Human-Computer Interaction* EICS, p. 1-28.
- Walker, Miriam, Leila Takayama et James A Landay (2002). “High-fidelity or low-fidelity, paper or computer? Choosing attributes when testing web prototypes”. In : *Proceedings*

- of the human factors and ergonomics society annual meeting*. T. 46. 5. SAGE Publications Sage CA : Los Angeles, CA, p. 661-665.
- Wiklund, Michael E, Christopher Thurrott et Joseph S Dumas (2016). “Does the fidelity of software prototypes affect the perception of usability?” In : *Proceedings of the Human Factors Society Annual Meeting*. SAGE PublicationsSage CA : Los Angeles, CA.
- Zhao, Yuxiang et Qinghua Zhu (2014). “Evaluation on crowdsourcing research : Current status and future direction”. In : *Information Systems Frontiers* 16.3, p. 417-434.
- Zouhairi, Abdallah (2010). “L’analogie, moteur de la pensee”. In : *Lectures*. URL : <https://doi.org/10.4000/lectures.1022>.
- Zwinderman, Matthijs, Rinze Leenheer, Azadeh Shirzad, Nikolay Chupriyanov, Glenn Veugen, Biyong Zhang et Panos Markopoulos (2013). “Using video prototypes for evaluating design concepts with users : a comparison to usability testing”. In : *IFIP Conference on Human-Computer Interaction*. Springer, p. 774-781.



Annexe A - Questionnaire utilisé dans l'étude expérimentale #2

Questionnaire

Le questionnaire comporte deux parties, dans la première partie les utilisateurs répondent à onze questions sur une échelle Likert de 5 points allant de "pas du tout d'accord" à "tout à fait d'accord".

Partie I

1. Pensez-vous qu'il est facile de trouver les informations pour effectuer les différentes actions ?
2. Pensez-vous que les informations dans l'IHM sont bien organisées ?
3. Pensez-vous que la navigation dans l'application se fait facilement (parcours des pages, retour menu principale, etc.) ?
4. Pensez-vous que le design de l'IHM est esthétiquement attirant ?
5. Est-ce que vous arrivez à retrouver ce que vous cherchez rapidement dans l'IHM ?
6. Pensez-vous que L'IHM est-elle claire et lisible ?
7. Est-ce que vous arrivez à vous situer facilement dans l'application ?
8. Est-ce que les informations sont bien disposées pour vous sur les écrans ?
9. Est-ce que vous vous retrouver facilement dans l'interface pour effectuer votre action

Partie II

Dans la deuxième partie du questionnaire, les utilisateurs ont été demandé à répondre aux questions suivantes :

- Quels sont les problèmes que vous voyez dans l'IHM. Pourquoi pensez-vous que c'est un problème ?
- Qu'est-ce qui vous a plu le plus dans l'IHM ?
- Si vous pouviez changer quelque chose dans l'IHM, laquelle ?
- Est-ce que vous pensez que l'IHM doit être améliorée ? Si oui comment ?



Annexe B - Questionnaire SUS

Questionnaire SUS

Le questionnaire SUS comporte les dix questions suivantes :

1. Je pense que vais utiliser le service fréquemment.
2. Je pense que le service est inutilement complexe.
3. Je pense que le service est facile d'utilisation.
4. Je pense que je vais devoir faire appel au support technique pour pouvoir utiliser ce service.
5. Je trouve que les fonctionnalités du service sont bien intégrées.
6. Je trouve qu'il y a beaucoup trop d'incohérences dans ce service.
7. Je pense que la plupart des gens apprennent très rapidement à utiliser le service
8. Je trouve le service vraiment très lourd à utiliser.
9. Je me suis senti très confiant en utilisant ce service.
10. J'ai dû apprendre beaucoup de choses avant de pouvoir utiliser ce service.

Calcul du score SUS

Les questions sont notés sur 5 points allant de "Pas du tout d'accords" à "Tout a fait d'accord". Le résultat de ce questionnaire consiste en un score SUS et est compris entre 0 et 100. Ce dernier est calculé en suivant la méthode standardisée suivante :

Pour les questions impaires, soustraire 1 au résultat donné par les utilisateurs. Si l'utilisateur a donné une note de 5, le score correspondant est 4 (5-1).

Pour les questions paires, calculer 5 moins le score donné par l'utilisateur. Si l'utilisateur a donné une note de 4, le score correspondant est 1 (5-4).

Une fois le total calculé, il doit être multiplié par 2,5 afin d'obtenir le score SUS compris entre 0 et 100.



Annexe C - Concepts fondamentaux en architecture

PAC

Le modèle Présentation-Abstraction-Contrôle (PAC) est un modèle multi-agent qui a comme principes directeurs le concept d'agent à facettes et une décomposition récursive (Coutaz 1987). Ce modèle multi-agent structure un système interactif sous la forme d'un ensemble d'unités de calcul spécialisées appelées agents. Un agent est défini comme *un système de traitement de l'information : il se distingue par un jeu d'opérations, des mécanismes d'entrée/sortie et une capacité à représenter un état* (Nigay 1994) et est capable d'initier et de réagir à des événements.

PAC est un modèle d'architecture applicable au cas des systèmes interactifs. Il structure récursivement un système interactif sous la forme d'une hiérarchie d'agents. Chaque agent est responsable d'un aspect spécifique des fonctionnalités de l'application et se compose de trois facettes : présentation (P), abstraction (A) et contrôle (C).

- **La facette présentation** : cette facette définit le comportement perceptible de l'agent pour un agent humain. Elle se charge d'exposer les modalités d'actions accessibles à l'utilisateur (le comportement d'entrée et de sortie).
- **La facette abstraction** : cette facette représente le noyau fonctionnel d'un agent, elle définit la compétence de l'agent indépendamment des considérations de présentation.
- **La facette contrôle** : cette facette est responsable de la communication entre les deux facettes présentation et abstraction et gère les relations avec les autres agents PAC (parent et fils).

Tout échange d'informations entre l'abstraction et la présentation s'effectue via la facette contrôle. C'est aussi par le biais de la facette contrôle que communiquent les différents agents PAC. En effet, quand un événement utilisateur est détecté, la communication entre les différents agents se fait comme suit : la présentation d'un agent PAC reçoit l'évènement utilisateur, elle transmet l'information à la facette contrôle, qui à son tour, prévient l'abstrac-

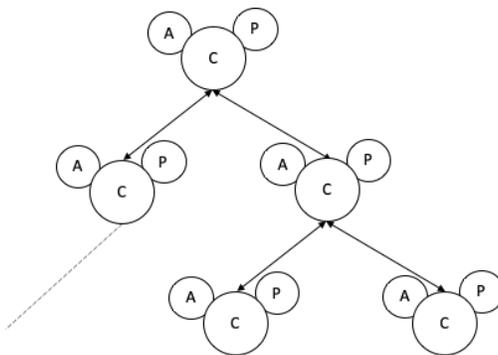


FIGURE C.1 – Hiérarchie d'agents PAC (Coutaz 1987)

tion, ensuite prévient la facette contrôle de son agent parents et ses agents fils, qui mettent à jour leurs facettes P et/ou A.

Un système interactif peut être modélisé par une architecture PAC comme l'illustre la figure C.1. PAC encourage la répartition des traitements sémantiques et syntaxiques (Nigay 1994). Il met l'accent sur une organisation modulaire hautement parallèle et répartit l'état de l'interaction entre un ensemble d'agents coopérants.