



HAL
open science

L'évaluation ergonomique de dispositifs connectés en santé pour le domicile patient dans le cadre de la chirurgie ambulatoire : vers un modèle de l'utilisabilité adapté à la e-santé

Noémie Chaniaud

► To cite this version:

Noémie Chaniaud. L'évaluation ergonomique de dispositifs connectés en santé pour le domicile patient dans le cadre de la chirurgie ambulatoire : vers un modèle de l'utilisabilité adapté à la e-santé. Psychologie. Université de Picardie Jules Verne, 2020. Français. NNT : 2020AMIE0015 . tel-03626023

HAL Id: tel-03626023

<https://theses.hal.science/tel-03626023v1>

Submitted on 31 Mar 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 de Doctorat

Spécialité Psychologie (Psychologie cognitive et ergonomie)

présentée à l'Ecole Doctorale en Sciences Humaines et Sociales (ED 586)

de l'Université de Picardie Jules Verne

par

Noémie CHANIAUD

pour obtenir le grade de Docteur de l'Université de Picardie Jules Verne

***L'évaluation ergonomique de dispositifs connectés en santé
pour le domicile patient dans le cadre de la chirurgie
ambulatoire : Vers un modèle de l'utilisabilité
adapté à la e-santé***

Soutenue le 30 novembre 2020, après avis des rapporteurs, devant le jury d'examen :

Julien NELSON, Rapporteur

Maître de Conférences HDR, Université de Paris

Sylvia PELAYO, Rapporteur

Maître de Conférences HDR, Université de Lille

Simon RICHIR, Examineur

Professeur, ENSAM Laval

Gérard VALLERY, Examineur

Professeur, Université Picardie Jules Verne

Emilie LOUP-ESCANDE, Directrice de thèse

Maître de Conférences HDR, Université Picardie Jules Verne

Olga MEGALAKAKI, Directrice de thèse

Professeure, Université Picardie Jules Verne

Thèse co-financée par le Conseil Régional des Hauts-de-France et par le FEDER (Fonds Européen de Développement Régional) dans le cadre du projet SMART ANGEL.



Remerciements

Mes premières pensées et remerciements vont aux personnes sans qui cette thèse n'aurait même pas existé. Je pense évidemment à mes directrices de thèse qui ont pris le soin de me sélectionner pour ce contrat et de m'encadrer avec bienveillance tout au long de mon travail de recherche. Merci Olga pour tous ces échanges très inspirants, pour m'avoir laissé me développer et m'épanouir. Merci Emilie pour ton investissement, pour m'avoir poussé, rassuré, m'avoir laissé cette immense liberté d'oser, proposer, me tromper (parfois...) et te laisser convaincre. Merci également de m'avoir donné l'opportunité de participer à bon nombre de conférences nationales, internationales et à des concours de vulgarisation toujours et encore. Clairement, on ne peut pas rêver mieux pour une thèse ! Et merci à vous deux de m'avoir fait confiance du début à la fin de cette thèse, mais également pour votre réactivité légendaire !! Je recevais systématiquement des retours constructifs nourris de commentaires et de (très) nombreuses corrections à chaque production. Rares sont les doctorants à avoir de tels retours aussi rapides et complets que ceux que j'ai pu avoir. Vous avez été attentives, disponibles, compréhensibles et sans vous ce travail n'aurait certainement pas abouti.

Je remercie les relecteurs officiels de ma thèse, Sylvia Pelayo et Julien Nelson, pour vos rapports bienveillants et constructifs. Je souhaiterais également remercier les autres membres du jury de ce manuscrit, Simon Richir ainsi que Gérard Vallery. Merci à vous tous pour l'intérêt que vous avez porté à mon travail. Je suis particulièrement honorée que vous ayez tous accepté de participer à ce jury de thèse de doctorat. Merci Gérard, pour m'avoir suivi tout au long de mon doctorat, en tant que membre de mon comité de suivi de thèse, mais également pour l'organisation de tous les séminaires d'ergonomie qui ont alimenté les réflexions de ma thèse.

J'aimerais remercier les membres du Projet *Smart Angel* que j'ai pu côtoyer pendant ces trois années. Pour commencer, les salariés d'Evolucare, Vincent, Alice et Pierre qui ont toujours été très intéressés par mon travail, mes retours et qui ont participé activement (financièrement et humainement) à la dernière étude de ma thèse. Merci pour votre enthousiasme, votre curiosité, votre gentillesse, vos accueils systématiquement joviaux rendant nos échanges toujours agréables et sympathiques. Je remercie également, le CHU de Nîmes, Professeur Cuvillon et Professeur Boisson, partenaire privilégié de ce projet, pour m'avoir ouvert les portes du service de chirurgie ambulatoire avec mes collègues d'Evolucare. Mes pensées vont aussi à tous ceux qui ont participé à ce projet : Annette et Cyril. Il y a aussi celles et ceux que j'ai emmenés dans les tâches particulièrement fastidieuses du double codage vidéo : Natacha, Sophie, Goran, Alexandre et Camille. Merci pour votre temps et votre motivation. Cela a été

un plaisir de travailler avec vous et j'espère que nous réitérerons ces occasions de travailler ensemble. Enfin, cette thèse ne pourrait pas aboutir sans les 235 participants mobilisés dans mes études. C'est également grâce à vous que ce travail a pu s'achever.

Mes pensées vont aussi à l'ensemble des membres de mon laboratoire, Vanessa, Mathieu, Yannick, Véronique, Béatrice, Frédérica, Harmonie, Nicolas, Élodie, Joffrey - on se boit un verre dans un verre à pied ;) - Dimitri, Perrine, Léo, Pauline, Julie, Hélios... avec une pensée particulière pour Nabila qui, il faut le rappeler, m'a sauvé la mise plus d'une fois et m'a aidé dans moult démarches. Merci infiniment ! Je voudrais également remercier les doctorants de nos séminaires très privilégiés de Gérard Vallery : Nina, Camille, Sami, Sophie. Cela a vraiment été un plaisir de passer ces trois années avec vous. Je me souviendrai éternellement de tout ce bon temps passé ensemble à Paris, Lyon, Florence... Vous êtes tous tellement compétents et pertinents dans vos remarques, mais également d'une incroyable gentillesse.

Je souhaiterais également remercier mes ami(e)s du RJCE où j'y ai trouvé un soutien sans faille, avec qui j'ai partagé mes galères communes des thèses en ergo : Artémis, Audrey, Élodie, Anaïs, Anthony, Benoît avec un petit clin d'œil pour ma copine de couture et voisine du pays de Gex, Elena. J'espère que nous pourrons reprendre rapidement nos afterworks. Je pense également à Nina et Camille qui nous ont rejoints et qui ont permis de faire d'Amiens, l'épicentre du RJCE. Et je tiens à vous dire que la relève arrive !

Mes remerciements vont également à celle et celui qui ont relu officieusement ce manuscrit avec la plus grande attention, ma mère, comme elle dirait « l'auteure de l'auteure », ainsi que par un grand romancier, j'ai nommé Claude, l'œil de lynx. Il y a également mes relecteurs anglophones préférés, Catherine et Patrick, merci encore pour vos retours, j'espère vous voir bientôt.

Une thèse ne peut être réalisée sans un soutien extérieur au milieu professionnel, j'en viens à remercier mes amis d'enfance de Franconville - Delphine, Lolo, Yoyo, Jeff, J-B, Jo - de Paris - Greg, Max, Laurel, Camille - et de Gex, Quentin & Quentin, Morgane, Antoine, Éric, Camille, Mat(t)hieu, Steven, Apolline, Elena... Toujours présents pour oublier les tracas de la thèse. Merci aussi à mon Kiné ;) Menclé pour avoir soigné (et re soigné) les conséquences de mes mauvaises habitudes de travail.

Et puis, il y a celles et ceux qu'on ne choisit pas, la famille - à mes parents, mes frères, ma sœur - à ces racines, baignant dans un humour sarcastique aux répliques percutantes et à l'animation débordante. Je vous aime tous même quand je repense à tous ces repas de famille toujours très

(trop) éveillés. Même si je ne le fais pas paraître, vous êtes pour moi un soutien d'une importance capitale. Vous savez prendre les choses avec légèreté et me rassurer dans les moments d'angoisse. Évidemment, je n'oublie pas non plus toute cette nouvelle génération de Chaniaud qui nous procure tant de joie et d'enchantement. Je ne peux pas non plus m'empêcher d'avoir une pensée pour ma belle-famille chez qui les repas sont nettement moins bruyants, mais tout aussi constructifs. J'y ai trouvé un soutien incroyable, voire quelques supporters.

Enfin, après trois années de thèse, c'est le moment de se remémorer des événements uniques qui ont marqué ce doctorat. Cette thèse aura traversé de bons moments historiques comme la coupe du monde de football en juillet 2018, où la joie, l'euphorie, la symbiose nous avait transportés. Je pense également à tous ces voyages entrepris ces trois dernières années, la Lettonie, l'Inde, L'Illinois, le Japon qui m'auront permis de penser un instant à autre chose qu'à ma thèse. Il y a également de nettement moins bonnes périodes comme la crise du coronavirus. Le premier confinement était pourtant une manière forcée d'écrire ce manuscrit, et la sensation d'un soutien national dans cette dernière ligne droite. Le deuxième, c'est aussi celui qui aura gâché la fête d'une soutenance à l'ambiance familiale inoubliable en région parisienne. La dernière page de cette histoire de thèse s'écrira donc dans la douceur et le calme des montagnes jurassiennes franco-suissees avec vue sur le plus beau panorama du Mont-Blanc où j'y ai d'ailleurs trouvé une certaine sérénité, et pour cela, toujours accompagnée d'une seule et unique personne. Je pense à celui qui a participé et a vécu tous ces événements précités avec moi, celui qui m'a porté, motivé, poussé, consolé, aidé, supporté, protégé, rassuré... aimé depuis presque une décennie, celui qui est et restera mon petit amour. Merci mon Martin.

Liste des abréviations/acronymes

ANAP : *Agence Nationale d'Appui à la Performance*
ANSM : *Agence Nationale de Sécurité du Médicament et des produits de santé*
AP-HP : *Assistance Publique - Hopitaux de Paris*
ASA : *American Society of Anesthesiologists*
ATIH : *Agence Technique de l'Information sur l'Hospitalisation*
AVAI : *Années de vie ajustées sur l'incapacité*
BPI : *Banque Publique d'Investissement*
CE : *Conformité Européenne*
CER : *Comités Ethiques et de Recherche*
CHU : *Centre Hospitalier Universitaire*
CNIL : *Commission Nationale de l'Informatique et des Libertés*
CPP : *Comités de Protection des Personnes*
CRP-CPO : *Centre de Recherche en Psychologie – Cognition, Psychisme et Organisations*
DGCCRF : *Direction Générale de la Concurrence, de la Consommation, et de la Répression des Fraudes*
DMC : *Dispositif Médical Connecté*
DMCD : *Dispositif Médical connecté à Domicile*
DMCP : *Dispositif Médical Connecté Piloté*
DMD : *Dispositifs Médicaux à Domicile*
DREES : *Direction de la Recherche, des Etudes, de l'Evaluation et des Statistiques*
HANDLED : *Home medicAl coNnected Devices usabiLity Evaluation moDel*
HAS : *Haute Autorité de la Santé*
health-ITUEM : *health IT Usability Evaluation Model*
HIS : *Health Information Systems*
HIT : *Health Informatics Technology*
HMD : *Home Medical Device*
IEA : *International Ergonomics Association*
IHM : *Interaction Humain-Machine*
INPES : *Institut National de Prévention et d'Éducation à la Santé*
INSEE : *Institut National de la Statistique et des Etudes Economiques*
INVS : *INstitut de Veille Sanitaire*
JMIR : *Journal of Medical Informatics Research*
meCUE : *Modular Evaluation of the User Experience Components*
MIS : *Modélisation, Information & Systèmes*
QEM : *Web Quality Evaluation Model*
SAM : *Self-Assessment Manikin*
SIH : *Systèmes d'Information Hospitaliers*
SIS : *Systèmes d'Information de Santé*
SUMI : *Software Usability Measurement Inventory*
SUS : *System Usability Scale, System Usability Scale*
TAM : *Technology Acceptance Model*
TI : *Technologie de l'Information*
TIC : *Technologies de l'Information et de la Communication*
UMM : *Usability Measurement Model*
UPJV : *Université de Picardie Jules Verne*
UX : *eXpérience Utilisateur*

Résumé

Le système de santé français est en crise et cela s'accroîtra probablement dans les années à venir. Certaines solutions comme la chirurgie ambulatoire sont alors investies massivement pour faire face aux besoins de santé croissants. Pourtant, cette chirurgie sans nuitée n'a pas prouvé sa rentabilité en raison d'un suivi postopératoire lacunaire. Afin d'améliorer ce suivi et optimiser la chirurgie ambulatoire, le projet pluripartenarial *Smart Angel* – ayant pour objectif de co-concevoir un dispositif médical connecté permettant aux personnels hospitaliers de suivre les patients tout au long de leur convalescence – a vu le jour. Par l'intermédiaire d'une application sur tablette associée à des objets connectés, le patient récemment opéré réalise, à son domicile, son propre suivi de santé. Ainsi, des dispositifs médicaux habituellement utilisés par du personnel qualifié et formé à ces objets, se retrouvent entre les mains de patients novices aux profils divers. Ce changement d'utilisateurs peut entraîner des risques majeurs pour la sécurité et la santé du patient si le dispositif n'est pas adapté à cette variété de profils.

En collaboration avec des partenaires industriels et hospitaliers, l'objectif de ce travail de thèse est d'assister la conception du dispositif *Smart Angel* par l'évaluation de l'utilisabilité. Afin de répondre à cet objectif, nous avons proposé un modèle d'évaluation de l'utilisabilité, le modèle HANDLED (*Home medicAl coNnected Devices usabiLity Evaluation moDel*) sur la base des métriques de l'utilisabilité (*efficacité, efficacité et satisfaction*) et d'un contexte d'usage intégré dans une démarche d'évaluation par cycles itératifs et cumulatifs. Ce modèle contient 5 niveaux associés à chaque élément du contexte d'usage (*tâches, utilisateurs, ressources, système et environnements*) que nous avons éprouvés à travers 5 études. Dans chacune des études, afin d'évaluer l'utilisabilité, nous avons mobilisé des outils et méthodes permettant une collecte de données mixtes (qualitatives, quantitatives, subjectives et objectives) dans un milieu expérimental se rapprochant progressivement d'un environnement *in situ*.

La première étude (niveau 0 du modèle HANDLED – *tâches*) est une étude exploratoire visant à analyser le besoin et à comprendre le terrain (celui du parcours patient en chirurgie ambulatoire) afin de construire les premières maquettes du dispositif *Smart Angel*. L'étude 1a (niveau 1 : *tâches – utilisateurs*) évalue l'utilisabilité des objets connectés (tensiomètre et oxymètre de pouls connectés) de *Smart Angel* en prenant en compte les caractéristiques utilisateurs. L'étude 1b (niveau 2 : *tâche – utilisateurs – ressources*) évalue l'utilisabilité du tensiomètre en prenant en compte les ressources au travers des documents procéduraux. L'étude 2 (niveau 3 : *tâches – utilisateurs – ressources – système*) évalue l'utilisabilité du prototype 1 comprenant le *système* complet de *Smart angel* : une application sur tablette associée deux objets connectés précités. Enfin, l'étude 3 évalue le prototype 2 et intègre l'ensemble des éléments de l'étude 2 en ajoutant le contexte ambulatoire et le retour au domicile (niveau 4 : *tâches – utilisateurs – ressources – système – environnements*).

Pour conclure, ces études ont permis d'implémenter le modèle HANDLED adapté aux dispositifs connectés en santé pour le domicile patient. En dernière partie de cette thèse, nous discutons des contributions théoriques au travers du modèle HANDLED et des contributions méthodologiques associées aux choix des méthodes permettant une collecte de données mixtes. Enfin, nous précisons les contraintes auxquelles ce travail a été confronté. Pour ce faire, nous présentons d'abord notre retour d'expérience sur le projet pluripartenarial *Smart Angel* et ses limites. Puis, nous présentons nos perspectives de recherche afin d'enrichir ce travail de thèse, et les études déjà mises en place qui apporteront des données concrètes dans les années à venir.

Mot-clés : Dispositifs médicaux connectés ; Dispositifs connectés en santé ; Utilisabilité ; Conception Centrée Utilisateur ; Evaluation ergonomique, E-santé.

Abstract

The French healthcare system is in crisis and this will likely worsen in the future. Some solutions such as ambulatory surgery are being invested heavily to meet the growing health needs. However, this type of surgery has not proven to be cost effective due to postoperative follow-up deficiencies. These patients may return to emergency departments or perform out-of-hospital consultations. In order to improve this postoperative follow-up and optimize ambulatory surgery, the multi-stakeholder *Smart Angel* project - whose objective is to co-design a home connected medical device allowing hospital caregivers to monitor patients throughout their convalescence - was born. Using a tablet application associated with connected objects, the recently operated patient performs his own health monitoring at home. Thus, medical devices usually used by qualified health personnel trained in these objects end up in the hands of novice patients with different profiles. This switch of users can lead to major risks for the safety and health of the patient if the device is not adapted to this variety of users' profiles.

In collaboration with industrial and hospital stakeholders, the objective of this PhD research is to assist the design of the *Smart Angel* device by evaluating its usability. To meet this objective, we have proposed a model of usability evaluation, the HANDLED model (*Home medicAl coNnected Devices usabiLity Evaluation moDel*) based on usability metrics (*effectiveness, efficiency and satisfaction*) and a context of use integrated in an evaluation approach by iterative and cumulative cycles. This model contains 5 levels associated with each element of the context of use (*tasks, users, resources, system and environments*) that we have experienced through 5 studies. In order to evaluate usability, we mobilized tools and methods allowing a mixed data collection (qualitative, quantitative, subjective and objective) in an experimental environment progressively closer to an *in situ* environment.

The first study (level 0 of the HANDLED model - *tasks*) is an exploratory study aimed at analyzing the need and understanding the field of the patient pathway in ambulatory surgery in order to build the first mock-ups of the *Smart Angel* device. Study 1a (level 1: *tasks - users*) evaluates the usability of the connected objects (connected blood pressure and pulse oximeter) of *Smart Angel* device taking into account user characteristics. Study 1b (level 2: *task - users - resources*) evaluates the usability of the blood pressure monitor by taking into account the resources through the procedural documents. Study 2 (level 3: *tasks - users - resources - system*) evaluates the usability of prototype 1 including the full *Smart Angel* system: a tablet application associated with two of the above-mentioned connected objects. Finally, study 3 evaluates prototype 2 and integrates all the elements of study 2 by adding the ambulatory context and the return home (level 4: *tasks - users - resources - system - environments*).

To conclude, these studies have allowed the implementation of the HANDLED model adapted to connected devices in health care for the patient's home. In the last part of this PhD work, we discuss the theoretical contributions through the HANDLED model and the methodological contributions associated with the choice of methods for mixed data collection. Finally, we will specify the constraints that this work has been confronted with. To do so, we first present our feedback on the multi-stakeholder *Smart Angel* project and its constraints. Then, we present our research perspectives in order to improve this PhD thesis, and the studies already implemented that will provide concrete data in the coming years.

Key-words: home medical devices; Healthy Connected Devices; Usability, User-centered design; Ergonomic evaluation; E-health.

Sommaire

INTRODUCTION GÉNÉRALE	1
1. <i>Le contexte sociétal : la naissance du projet Smart Angel</i>	1
2. <i>Le contexte pluripartenarial : le projet Smart Angel</i>	3
2.1. Le projet <i>Smart Angel</i> : un double objectif	3
2.2. Un projet de co-développement : les partenaires industriels, académiques et hospitaliers	4
2.3. Un projet de co-conception : les partenaires utilisateurs intégrés dans une démarche centrée utilisateur	6
3. <i>L'organisation du manuscrit de thèse</i>	7
PARTIE I : PARTIE THÉORIQUE	9
CHAPITRE 1 : POSITIONNEMENTS DISCIPLINAIRES, APPROCHES THÉORIQUE ET MÉTHODOLOGIQUE DE L'UTILISABILITÉ DES TECHNOLOGIES	9
1. <i>Le positionnement disciplinaire basé sur une conception centrée utilisateur : une démarche pluridisciplinaire au sein d'une activité collaborative</i>	9
1.1. La démarche de conception centrée utilisateur	9
1.2. Un positionnement pluridisciplinaire : de l'ergonomie à la psychologie cognitive en passant par les sciences de conception	10
1.2.1. L'ergonomie	10
1.2.2. La psychologie cognitive	11
1.2.3. Les sciences de la conception	12
1.3. Une thèse pluridisciplinaire dans une modalité de travail collectif	13
2. <i>Les approches théoriques et opérationnelles de l'utilisabilité : une notion centrale dans la conception centrée utilisateur</i>	13
2.1. Les concepts connexes de l'utilisabilité	13
2.1.1. L'acceptabilité pratique, sociale et située	13
2.1.2. L'utilité-valeur et l'utilité-destination	14
2.2. L'utilisabilité : quelques définitions	15
2.3. L'utilisabilité : les trois principaux modèles théoriques	15
2.3.1. Le modèle de Shackel (1991)	15
2.3.2. Le modèle de Nielsen (1993)	17
2.3.3. Le modèle proposé par la norme ISO 9142-11 (2018)	19
2.4. Les critiques des modèles et des métriques de l'utilisabilité	21
2.4.1. Le positionnement de l'utilisabilité en regard de l'acceptabilité, et l'utilité dans l'évaluation ergonomique	21
2.4.2. L'inventaire des métriques utilisées sur un recensement de modèle entre 1971 à 2018	23
2.4.3. Une utilisabilité spécifique à chaque domaine d'application	23
2.5. L'utilisabilité, des divergences disciplinaires, critérielles et applicatives	24
2.6. Synthèse des modélisations : vers une définition de l'utilisabilité et proposition d'un modèle opérationnel	25
3. <i>Les approches conceptuelles de l'utilisabilité : les mesures des métriques de l'utilisabilité</i>	26
3.1. Les mesures des trois principales métriques de l'utilisabilité	26
3.1.1. L'efficacité : définitions et mesures	26
3.1.2. L'efficacité : définitions et mesures	28
3.1.3. La satisfaction : définitions et mesures	30
3.1.4. Les liens entre l'expérience Utilisateur et l'utilisabilité	32
3.1.5. Les relations entre efficacité, efficacité et satisfaction	33
3.2. Les métriques connexes de l'utilisabilité	35
3.2.1. L'apprenabilité (ou facilité d'apprentissage) : définitions et mesures	35
3.2.2. La mémorisation : définitions et mesures	36
3.3. Synthèse des métriques de l'utilisabilité	36
3.4. Les composants du contexte d'utilisation	37
3.4.1. Les ressources	39
3.4.2. L'environnement	39
3.4.3. Caractéristiques utilisateurs	40

4.	<i>Les approches méthodologiques de l'utilisabilité : leur complémentarité dans le cycle de conception, les types, les méthodes et les outils d'évaluation</i>	40
4.1.	L'intégration des objectifs d'évaluation de l'utilisabilité dans le cycle de conception d'un produit.....	40
4.1.1.	Les objectifs du « quoi » dans l'évaluation de l'utilisabilité : une utilisabilité située	40
4.1.2.	Les objectifs du « quand » dans l'évaluation de l'utilisabilité : une utilisabilité évolutive	41
4.1.3.	Les objectifs du « quand évaluer quoi ? » : proposition d'un modèle à deux dimensions.....	41
4.2.	Trois lieux de recueil de données déterminés par trois approches épistémologiques : laboratoire, terrain, <i>in situ</i>	43
4.3.	Les types, les méthodes et les outils de l'évaluation de l'utilisabilité	44
4.3.1.	Trois types d'évaluation de l'utilisabilité	44
4.3.1.	Trois types de méthodes d'évaluation de l'utilisabilité : inspection, test utilisateur, enquête	46
4.3.2.	Le choix des bonnes méthodes et des bons outils	48
5.	<i>Synthèse du chapitre 1</i>	53

CHAPITRE 2 : L'ÉVALUATION DE L'UTILISABILITÉ DE DISPOSITIFS CONNECTÉS EN SANTÉ POUR LE DOMICILE PATIENT.....

1.	<i>Les dispositifs médicaux et le marquage de conformité européen</i>	55
1.1.	La définition des dispositifs médicaux.....	55
1.2.	La classification et la conception des dispositifs médicaux.....	55
2.	<i>La « e-santé » ou « santé connectée » : un domaine émergent non sans écueils</i>	56
2.1.	La définition de la e-santé.....	56
2.2.	La e-santé, une technologie émergente ?.....	58
2.3.	Les systèmes d'information de santé et la télésanté	59
2.3.1.	Le Système d'Information de Santé (ou Hospitaliers).....	59
2.3.2.	Les Technologies de l'Information et de la Communication ou TIC	60
2.4.	Les avantages et les freins de la e-santé	61
2.5.	Les principales difficultés de l'évaluation de l'utilisabilité de dispositifs connectés en santé.....	62
2.5.1.	L'origine des défauts d'utilisabilité des dispositifs connectés en santé utilisables par le grand public.....	62
2.5.2.	Les trois principaux défis des évaluations de l'utilisabilité des technologies en santé	63
2.5.3.	Les cadres et modèles d'évaluation de l'utilisabilité	65
3.	<i>État de l'art des cadres méthodologiques de l'évaluation de l'utilisabilité de dispositif en santé pour le domicile patient</i>	67
3.1.	La classification des études portant sur l'utilisabilité des dispositifs en santé pour le domicile patient	68
3.2.	Les résultats des études empiriques sur l'utilisabilité de dispositifs connectés en santé pour le domicile patient	71
3.2.1.	Les objets de l'évaluation	71
3.2.2.	Les revues.....	72
3.3.	Les résultats des méthodes utilisées dans les évaluations de l'utilisabilité	72
3.3.1.	Les environnements d'évaluation	72
3.3.2.	Les méthodes	74
3.3.3.	Les outils.....	75
3.3.4.	Les modèles et les normes	77
3.3.5.	Les métriques.....	78
3.3.6.	Les caractéristiques des participants	79
3.4.	Conclusion et mise en relation de l'analyse systématique avec d'autres études relatives aux évaluations des dispositifs en santé pour le domicile patient.....	80
3.4.1.	Une évolution à travers le temps	80
3.4.2.	Conclusion : Quelles évolutions pouvons-nous constater aujourd'hui ? Synthèse sur la prise en compte des recommandations des évaluations de l'utilisabilité	82
4.	<i>État de l'art sur la prise en compte du contexte d'usage dans l'évaluation de l'utilisabilité de dispositifs de m-santé</i>	84
4.1.	Les caractéristiques utilisateurs.....	85
4.1.1.	Les facteurs sociaux	87
4.1.2.	Les facteurs cognitifs	87
4.1.3.	Les facteurs émotionnels.....	90
4.2.	Les ressources.....	91
4.2.1.	Les ressources réutilisables : les documents procéduraux	91
4.2.2.	Les ressources consommables : Les formations et les programmes d'éducation thérapeutique	94
5.	<i>Synthèse du chapitre 2</i>	97

CHAPITRE 3 : PROBLÉMATIQUE ET DÉMARCHÉ DE RECHERCHE	99
1. <i>Problématique et objectif</i>	99
2. <i>Hypothèses et démarche de recherche</i>	102
2.1. Le processus de conception du dispositif <i>Smart Angel</i> en ambulatoire	102
2.2. L'intégration de notre démarche de recherche au sein du processus de conception de <i>Smart Angel</i>	104
2.3. Les choix méthodologiques.....	108
3. <i>Synthèse du chapitre 3</i>	111

PARTIE II : PARTIE EMPIRIQUE..... 112

CHAPITRE 1 : ÉTUDE EXPLORATOIRE. ANALYSE DE BESOINS ET COMPRÉHENSION DU TERRAIN..... 112

1. <i>Préambule</i>	112
1.1. Contexte et objectif industriels	112
1.2. Contexte et objectif théoriques	112
2. <i>Introduction</i>	113
2.1. La chirurgie ambulatoire en France et leurs patients	113
2.1.1. Parcours patient en ambulatoire	114
2.1.2. Les critères d'éligibilité à la chirurgie ambulatoire	114
2.1.3. Les actes chirurgicaux les plus pratiqués en ambulatoire.....	116
2.1.4. Le profil des patients en chirurgie ambulatoire	117
2.2. La conception d'un template pour personas patient adapté à la chirurgie ambulatoire	118
3. <i>Objectifs de l'étude</i>	121
4. <i>Méthode</i>	122
4.1. Recueil de données sur le terrain	122
4.1.1. Observations ouvertes et entretiens informels.....	122
4.1.2. Caractéristiques du personnel hospitalier et des patients interrogés	122
4.2. Atelier : conception de personas	123
5. <i>Résultats</i>	124
5.1. L'organisation du suivi postopératoire en constante évolution	124
5.2. Événements pour l'intégration du dispositif <i>Smart Angel</i> au domicile patient.....	126
5.3. Conception de personas et scénarii	126
6. <i>Discussion</i>	127
7. <i>Conclusion et perspectives de recherche</i>	129
8. <i>Synthèse du Chapitre 1 - Étude exploratoire</i>	131

CHAPITRE 2 : ÉTUDE 1A. EFFETS DES CONNAISSANCES EN SANTÉ D'UNE POPULATION NOVICE SUR L'UTILISABILITÉ DES DEUX DISPOSITIFS MÉDICAUX CONNECTÉS..... 132

1. <i>Préambule</i>	132
1.1. Contexte et objectif industriels	132
1.2. Contexte et objectif théoriques	133
2. <i>Introduction</i>	135
2.1. Contexte de l'étude.....	136
2.2. Évaluation de l'utilisabilité	137
2.3. L'utilisabilité et les caractéristiques utilisateurs	138
2.4. Utilisabilité, littératie en santé et connaissances antérieures en santé	139
3. <i>Objectifs et hypothèses</i>	141
4. <i>Méthodes</i>	141
4.1. Participants	141
4.2. Matériel et mesures	142
4.2.1. Dispositifs médicaux et tâches	142
4.2.2. Questionnaires	143
4.3. Procédure	144
4.4. Cotation et analyse statistique	144
4.4.1. Mesures objectives de l'utilisabilité	145
4.4.2. Mesures subjectives de l'utilisabilité	145
4.4.3. Mesure des connaissances antérieures en santé	145
4.4.4. Lectures des données de santé	145
4.4.5. Analyse de données	146

5.	<i>Résultats</i>	146
5.1.	Fiabilité Inter-juges.....	146
5.2.	Consistance interne du questionnaire de connaissances antérieures en santé.....	146
5.3.	Statistiques sur les utilisateurs : caractéristiques sociodémographiques et technophilie	147
5.4.	Résultats de l'évaluation de l'utilisabilité.....	148
5.4.1.	Tensiomètre	148
5.4.2.	Oxymètre de pouls.....	151
5.5.	Analyse par groupes/cluster (k-means).....	152
6.	<i>Discussion</i>	154
6.1.	Principales contributions	154
7.	<i>Conclusion et Perspectives de recherche</i>	157
7.1.	Perspectives de recherche	157
7.2.	Recommandations pour la conception et l'intégration des dispositifs médicaux au domicile.	157
7.3.	Conclusion	158
8.	<i>Synthèse du Chapitre – Étude 1a</i>	160

CHAPITRE 3 : ÉTUDE 1b. EFFETS DES DOCUMENTS PROCÉDURAUX D'UNE POPULATION NOVICE SUR L'UTILISABILITÉ DU TENSIOMÈTRE

	161	
1.	<i>Préambule</i>	161
1.1.	Contexte et objectif industriels	161
1.2.	Contexte et objectif théoriques.....	162
2.	<i>Introduction</i>	164
3.	<i>Objectifs et hypothèses</i>	168
4.	<i>Méthode</i>	170
4.1.	Participants.....	170
4.2.	Matériel et mesures	171
4.2.1.	Dispositif Médical: le tensiomètre	171
4.2.2.	Conditions expérimentales.....	171
4.2.3.	Questionnaires.....	172
4.3.	Procédure	172
4.4.	Cotation et analyse de données	175
5.	<i>Résultats</i>	176
5.1.	Expérience antérieure dans la manipulation du tensiomètre.....	176
5.2.	Efficacité	177
5.3.	Efficience.....	177
5.4.	Mémorisation	178
5.5.	Résultats généraux : comparaison entre les deux groupes multimodaux.....	179
6.	<i>Discussion</i>	182
6.1.	Principales observations	182
6.2.	Vers le choix du format audio.....	184
7.	<i>Conclusions et perspectives de recherche</i>	186
8.	<i>Synthèse du Chapitre 3 – Étude 1b</i>	188

CHAPITRE 4 : ÉTUDE 2. EFFETS DES CARACTÉRISTIQUES SOCIO-DEMOGRAPHIQUES ET COGNITIVES SUR L'UTILISABILITÉ DU DISPOSITIF SMART ANGEL.....

	189	
1.	<i>Préambule</i>	189
1.1.	Contexte et objectif industriels	189
1.2.	Contexte et objectif théoriques.....	190
2.	<i>Introduction</i>	193
2.1.	Utilisabilité.....	193
2.2.	Caractéristiques utilisateurs : âge, niveau d'éducation, technophilie, littératie en santé.....	194
2.2.1.	L'âge.....	194
2.2.2.	Le niveau d'éducation.....	195
2.2.3.	Technophilie (expérience TI/dispositifs médicaux)	195
2.2.4.	Littératie en santé	196
3.	<i>Objectifs et hypothèses</i>	198
4.	<i>Méthode</i>	200
4.1.	Participants.....	200

4.2.	Matériel et mesures	200
4.2.1.	Dispositif <i>Smart Angel</i>	201
4.2.2.	Les personas et scénarios	203
4.2.3.	Questionnaires	203
4.2.4.	Entretiens semi-directifs	204
4.3.	Traitement des données quantitatives	205
4.3.1.	Mesure de l'efficacité	205
4.3.2.	Mesure de l'efficience	205
4.3.3.	Mesure de la satisfaction	205
4.4.	Procédure	206
4.5.	Analyse de données	207
4.6.	Fidélité inter-juges : mesures objectives d'efficacité et d'efficience	207
5.	Résultats	208
5.1.	Analyse quantitative	208
5.1.1.	Effets des caractéristiques utilisateurs sur l'utilisabilité	208
5.2.	Analyse qualitative : Satisfaction (qualitative) des utilisateurs du dispositif <i>Smart Angel</i>	214
5.2.1.	Facilité d'utilisation perçue	214
5.2.2.	Utilité perçue	215
6.	Discussion	217
6.1.	Discussion de l'analyse quantitative	217
6.2.	Discussion de l'analyse qualitative	220
7.	Conclusions et perspectives de recherche	222
8.	Synthèse Chapitre 4 – Étude 2	224

CHAPITRE 5 : ÉTUDE 3. EFFETS DES CARACTÉRISTIQUES UTILISATEURS ET DES PÉDAGOGIES HOSPITALIÈRES EN CHIRURGIE AMBULATOIRE SUR L'UTILISABILITÉ DU DISPOSITIF *SMART ANGEL* EN SIMULATION 226

1.	Préambule	226
1.1.	Contexte et objectif industriels	226
1.2.	Contexte et objectif théoriques	227
2.	Introduction	230
2.1.	La simulation	230
2.2.	L'évaluation de l'utilisabilité	230
2.2.1.	Le choix des métriques	230
2.2.2.	Les composants du contexte d'usage	231
3.	Objectifs et hypothèses	232
4.	Méthode	234
4.1.	Participants	234
4.2.	Matériel et mesures	234
4.2.1.	Le dispositif <i>Smart Angel</i> et son implantation	234
4.2.2.	La mise en situation : le simUsanté	235
4.2.3.	Questionnaires	239
4.2.4.	Entretiens semi-directifs et focus groups	240
4.3.	Traitement des données	240
4.3.1.	Mesure objective de l'utilisabilité	240
4.3.2.	Mesure subjective de l'utilisabilité : mesure de la satisfaction	241
4.4.	Procédure	242
4.5.	Analyse de données	243
4.6.	Fiabilité inter-juge : mesures objectives d'efficacité et d'efficience	244
5.	Résultats	244
5.1.	Analyse quantitative	244
5.1.1.	Effet d'apprentissage sur l'efficacité et l'efficience	245
5.1.2.	Effet des programmes d'éducation « actif/passif »	245
5.1.3.	Effet des caractéristiques utilisateurs	249
5.2.	Analyse qualitative	252
5.2.1.	Mesure de la satisfaction : analyse thématique	252
6.	Discussion	258
6.1.	Discussion de l'analyse quantitative	258
6.2.	Discussion de l'analyse qualitative	261
6.3.	Limites	263
7.	Conclusions et perspectives de recherche	264

8. Synthèse du chapitre 5 – Étude 3	266
---	-----

DISCUSSION GÉNÉRALE 269

1. Introduction.....	269
2. Les contributions théoriques et méthodologiques de la thèse : proposition et implémentation d'un modèle de l'évaluation de l'utilisabilité des dispositifs connectés en santé pour le domicile patient (HANDLED)	272
2.1. Contributions théoriques	272
2.1.1. Les métriques objectives du modèle HANDLED	272
2.1.2. La métrique subjective du modèle HANDLED	275
2.1.3. Les composants du contexte d'usage du modèle HANDLED	277
2.1.4. L'apport de la progression des études dans le modèle HANDLED	281
2.2. Contributions méthodologiques.....	282
2.2.1. Les avantages et les inconvénients des principaux outils déployés	282
2.2.2. L'importance de la mise en situation : les personas et la simulation.....	287
3. Contraintes de la thèse : Retour d'expérience sur un projet de conception pluripartenarial.....	288
3.1. Les recommandations ergonomiques confrontées aux contraintes industrielles : un projet à deux vitesses	288
3.2. Le retard du projet <i>Smart Angel</i>	289
3.1. La hiérarchisation du partenariat de co-développement.....	291
3.2. Les évaluations ergonomiques pendant les études cliniques	291
4. Perspectives de recherche.....	292
4.1. Perspectives issues du projet <i>Smart Angel</i>	292
4.1.1. Les études ergonomiques <i>in situ</i> intégrées à l'étude clinique en cours (court terme)	292
4.1.2. Le niveau 5 du modèle HANDLED : Les changements organisationnels et relationnels dus à l'intégration du dispositif <i>Smart Angel</i> (long terme)	294
4.2. Transfert du modèle HANDLED vers d'autres conceptions de dispositifs connectés en santé pour le domicile patient : Étude sur l'utilisation de dispositifs médicaux connectés dans le cas de mesure ambulatoire de la pression artérielle (court terme)	295
5.. Synthèse de la Discussion générale	297

BIBLIOGRAPHIE 299

ANNEXES 338

<i>Annexe 1 : Revue des modèles d'utilisabilité de 1991 à 2018.....</i>	338
<i>Annexe 2 : Critères d'inclusion et d'exclusion de la population pouvant utiliser le dispositif Smart Angel</i>	341
Description de la population à étudier	341
Critères d'inclusion pour le choix des sujets.....	341
Critères de non inclusion pour le choix des sujets	342
Critères d'exclusion	342
<i>Annexe 3 : Mapping des produits de la m-santé et patterns des usages.....</i>	343
<i>Annexe 4 : Tableau d'analyse systématique des méthodes et outils utilisés dans l'évaluation de l'utilisabilité de dispositif en santé</i>	344
<i>Annexe 5 : Fiche produit iHealth Sense Tensiomètre poignet connecté BP7.....</i>	368
<i>Annexe 6 : Fiche produit Oxymètre de pouls connecté iHealth Air P03M</i>	369
<i>Annexe 7 : Prospection sur le nombre d'interventions et l'âge des patients en chirurgie ambulatoire</i>	370
<i>Annexe 8 : Présentation des personas et leurs scénarios</i>	371
Damien	371
Marianna	376
Pierre	380
Fatima.....	384
Patrick.....	388
<i>Annexe 9 : Questionnaire sur les connaissances antérieures en santé en français (40 items)</i>	392
<i>Annexe 10 : Questionnaire sur les connaissances antérieures en santé en anglais (40 items)</i>	394

<i>Annexe 11 : Détails des caractéristiques de l'utilisateur sur la base des mesures d'utilisabilité du tensiomètre et de l'oxymètre de pouls</i>	<i>396</i>
<i>Annexe 12 : Questionnaire de littératie en santé – le Newest Vital Sign</i>	<i>399</i>
<i>Annexe 13 : Questionnaire de littératie en santé – Le HLS-EU-Q16</i>	<i>400</i>
<i>Annexe 14 : Questionnaire caractéristiques sociodémographiques études 2 et 3</i>	<i>401</i>
14.1. Caractéristiques individuelles	401
14.2. Utilisation des outils numériques	402
<i>Annexe 15 : Guide d'entretien pour l'étude 2</i>	<i>404</i>
<i>Annexe 16 : Scénario Anesthésiste et rappel du déroulement d'une consultation pré-opératoire</i>	<i>405</i>
<i>Annexe 17 : Scénario Infirmier formateur du dispositif Smart Angel</i>	<i>407</i>
<i>Annexe 18 : Scénario Patient/Participant</i>	<i>408</i>
<i>Annexe 19 : Guide d'entretien et focus group pour l'étude 3</i>	<i>409</i>
<i>Annexe 20 : Grille de notation des erreurs de manipulation du dispositif</i>	<i>410</i>
<i>Annexe 21 : Nombre d'erreurs de manipulation des patients/participants lors de l'étude 3 en fonction de la grille de codage des erreurs (cf. annexe 20) et des trois essais demandés</i>	<i>413</i>
<i>Annexe 22 : Restitution des recommandations avec l'entreprise Evolucare</i>	<i>414</i>
<i>Annexe A : Liste des publications</i>	<i>417</i>

INTRODUCTION GÉNÉRALE

1. Le contexte sociétal : la naissance du projet *Smart Angel*

Le secteur hospitalier est en crise depuis le début des années 2000. D'après le rapport de 2018¹ de l'Institut National de la Statistique et des Etudes Economiques ou INSEE, la consommation des soins et de biens médicaux représente 8,9% du produit intérieur brut français en 2016 soit 198,5 milliards d'euros et ce chiffre ne fera qu'augmenter dans les années à venir. L'intensité de la demande ne déclinant pas, les hôpitaux doivent rentabiliser au maximum leurs actions ce qui n'est pas sans conséquences pour la santé des patients et du personnel hospitalier. Les études de santé en France de l'Institut National de Prévention et d'Éducation à la Santé ou INPES (Leon & Beck, 2014) et l'INstitut de Veille Sanitaire ou INVS (Eilstein, 2014), complétées par d'autres études prospectives comme les rapports de l'INSEE de 2010 alertent les politiques de santé publique sur les futurs problèmes sanitaires. La France doit se préparer à faire face à des problématiques de santé telles que le vieillissement de la population et la diminution du nombre de médecins.

La première tendance indéniable est le vieillissement de la population. Il y a une augmentation de l'âge moyen dans toutes les régions selon les scénarios centraux. D'après l'INSEE (2010), la proportion de personnes âgées de 60 ans doublera entre 2005 et 2050 : « *Cela correspond à 69 habitants âgés de 60 ans ou plus pour 100 habitants de 20 à 59 ans* ». Nous vivons de plus en plus longtemps, mais de moins en moins dans de bonnes conditions. Le premier constat alarmant de l'INVS (Eilstein, 2014) est l'augmentation du nombre d'Années de Vie Ajustées sur l'Incapacité ou AVAI² entre +8% et + 25% à l'horizon 2025 en France suivant les scénarios. Le vieillissement de la population s'accompagne d'une augmentation des maladies chroniques telles que le diabète ou les maladies cardiovasculaires, mais aussi une augmentation des pathologies neurodégénératives. Les besoins en matière de santé exploseront dans les prochaines décennies.

Dans ce même scénario, vient s'ajouter la diminution du nombre de médecins généralistes. Selon le scénario tendanciel de référence de la Direction de la Recherche, des

¹ <https://www.insee.fr/fr/statistiques>

² Indicateur créé par l'OMS en 1990 dans le cadre du Global Burden of Disease. L'acronyme AVAI est également connu sous le nom de Disability adjusted life years ou DALY.

Etudes, de l'Evaluation et des Statistiques ou DREES, à comportements constants, le nombre de médecins actifs se réduira jusqu'en 2020 (Attal-Toubert et Vanderschelden 2009). S'ajoute à cela la dégradation des relations patients-soignants. En 2025, les associations de patients seront systématiquement associées aux processus de recherche et d'innovation thérapeutique (le cas échéant de manière réglementaire). Cet horizon futur fera naître de nouveaux comportements vis-à-vis de la santé dus aux rejets du système de santé et à l'utilisation d'outils annexes, notamment associés à la e-santé pour se soigner. L'automédication devient une solution privilégiée pour éviter de passer par les longues files d'attente de l'hôpital. Les économistes attendent une explosion de l'utilisation d'outils de santé connectée (e-santé) d'ici 2025 : « leur diffusion aura un impact similaire à ce que nous avons connu avec le développement d'internet » (Allaert et al., 2016, p.1) .

Pour endiguer ces futures crises sanitaires, les politiques publiques exercent une pression sur les hôpitaux pour limiter les flux de patients dans les centres hospitaliers en augmentant drastiquement **les actes de chirurgie en ambulatoire**.

La chirurgie ambulatoire se distingue de la chirurgie conventionnelle par son absence d'hébergement de nuit (Bert et al., 2014). Les pouvoirs publics souhaitent promouvoir la chirurgie ambulatoire. Les attentes sont doubles : économiques et médicales. La chirurgie sans nuitée permettrait une amélioration de la sécurité et de la qualité des soins en limitant les risques d'exposition aux infections nosocomiales. L'activité optimisée aurait une retombée sur l'organisation hospitalière et donc sur la maîtrise des risques. Cela favorise aussi bien la qualité des prises en charge que leur prix. En raison de sa durée, la chirurgie conventionnelle est plus coûteuse que la chirurgie ambulatoire. L'économie potentielle se reflète sur les Tarifs Issus des Coûts. Ainsi, l'activité de la chirurgie ambulatoire est en nette augmentation depuis le début des années 2000. Le rapport de la DREES de 2013³ enregistre une augmentation de 40% entre 2007 et 2013 entraînant une suppression de 10% des lits de chirurgie conventionnelle. Cependant, le principal bénéficiaire de cette économie est uniquement l'assurance maladie. En effet, il serait très difficile de prouver les apports économiques que procure la chirurgie ambulatoire sur les établissements hospitaliers. Beaucoup de ces effets dépendraient de l'organisation de l'établissement. Les rapports ministériels (Bert et al., 2014) avaient donc pour

³ Source Statistique annuelle des établissements (DREES) - la distinction entre anesthésie et chirurgie ambulatoires est vue dans l'enquête SAE 2013.

objectif d'analyser en détail les conditions de croissance de la pratique ambulatoire et les freins à éviter.

Gentili et Delaunay (2014) ont pointé les échecs de la chirurgie ambulatoire. Ces deux anesthésistes ont identifié les causes de la réhospitalisation liées à la chirurgie. L'incapacité à gérer la douleur à domicile (Vial et al., 2000), ainsi que les difficultés sociales dues à l'absence d'entourage sont les principaux motifs de réadmission ou de consultation. Certaines chirurgies, par exemple des amygdales ou de la main, occasionneraient des consultations extrahospitalières dont la fréquence peut aller jusqu'à 38% des chirurgies (Ghosh & Sallam, 1994). Tous ces éléments ne font qu'alourdir la note de l'hôpital. Ainsi, la principale cause de **la non-rentabilité de la chirurgie ambulatoire semblerait être le mauvais suivi postopératoire**. C'est de ce besoin que le projet *Smart Angel* est né.

2. Le contexte pluripartenarial : le projet *Smart Angel*

2.1. Le projet *Smart Angel* : un double objectif

Le projet *Smart Angel*, débuté en avril 2017, est un Projet d'Investissement d'Avenir financé par la Banque Publique d'Investissement ou BPI. Il a la particularité d'être divisé en deux sous-projets (intra-hospitalier « retour en chambre » et ambulatoire « retour au domicile ») ayant des démarches de conception et d'évaluation différentes et ne faisant pas forcément intervenir les mêmes partenaires hospitaliers ou industriels :

- Le premier sous-projet en intra-hospitalier se focalise sur le retour en chambre post-bloc opératoire et vise une surveillance du patient alors qu'il se trouve dans sa chambre d'hôpital. Ce sous-projet implique un partenariat entre Evolucare, l'Assistance Publique - Hôpitaux de Paris ou AP-HP, deux Unités de Recherche ou UR de l'Université de Picardie Jules Verne ou UPJV (le laboratoire de Modélisation, Information & Systèmes ou MIS et le Centre de Recherche en Psychologie – Cognition, Psychisme et Organisations ou CRP-CPO) et l'entreprise InES spécialisée dans les technologies médicales.
- Le deuxième sous-projet en ambulatoire s'articule autour du suivi du patient depuis le service de chirurgie ambulatoire jusqu'à son domicile pendant sa période de convalescence. Il implique un partenariat entre Evolucare, deux Unités de Recherche de l'Université de Picardie Jules Verne (le MIS et le CRP-CPO) et le Centre Hospitalier Universitaire ou CHU de Nîmes.

Il faut savoir que ces deux sous-projets sont indépendants l'un de l'autre mais restent liées pour les démarches financières. Pour ce travail de thèse, **nous nous sommes exclusivement concentrées sur la partie relative à la chirurgie ambulatoire.**

Ces deux sous-projets s'engagent avant tout dans une démarche de conception centrée utilisateur telle que définie par la norme ISO 9241-210 (2019). **Ainsi, le projet Smart Angel peut être qualifié de projet pluripartenarial de (1) co-développement impliquant différents partenaires industriels, académiques et hospitaliers et de (2) co-conception dans la mesure où il intègre les utilisateurs finaux, patients et personnels hospitaliers.** Nous précisons ces éléments dans les deux parties qui suivent.

2.2. Un projet de co-développement : les partenaires industriels, académiques et hospitaliers

Le projet *Smart Angel* pour la partie ambulatoire a pour but le développement d'un **Dispositif Médical Connecté Piloté ou DMCP** (composé de divers dispositifs médicaux connectés et d'une application sur tablette) permettant de créer un lien virtuel entre le patient à son domicile sortant de chirurgie ambulatoire et l'hôpital. Ce dispositif permettrait à l'avenir de gérer à distance toutes complications postopératoires tout en limitant les coûts annexes de la chirurgie ambulatoire. Chacun des partenaires du projet s'engage à réaliser les lots de travaux et à répondre aux verrous scientifiques et technologiques qui leur sont adressés afin de permettre de développer puis de commercialiser le dispositif *Smart Angel* d'ici octobre 2020. Quatre parties prenantes ont été identifiées pour le projet *Smart Angel* ambulatoire :

- Une entreprise :
 - Evolucare, coordinateur du projet est une entreprise éditrice de logiciels pour la santé, à qui il incombe de développer le dispositif et de s'assurer du respect des normes pour l'accès à la conformité européenne du matériel ;
- Un hôpital :
 - Le CHU de Nîmes fournit le contenu de l'application et valide cliniquement le dispositif ;
- Des laboratoires :
 - L'UR CRP-CPO de l'UPJV d'Amiens répond aux problématiques ergonomiques et organisationnelles ;

- L'UR MIS de l'UPJV d'Amiens aborde la conception des algorithmes et de l'intelligence artificielle du dispositif.

D'après la définition des normes AFNOR (1986, p. 13), « *Le partenariat est un état d'esprit rendant possible la création entre partenaires de relations privilégiées fondées sur une recherche en commun d'objectifs à moyen ou long terme, menées dans des conditions permettant d'assurer la réciprocité des avantages* ». Ce projet se caractérise par la diversité de ses partenaires aux compétences complémentaires et aux langages variés permettant un échange de savoirs et de compétences vers un but commun d'innovation. Compte-tenu de cette analyse, nous pouvons classer le projet *Smart Angel* comme un projet **d'Open Innovation** (Chesbrough, 2003).

Différentes stratégies d'open innovation existent (sous-traitance, co-développement et partenariats d'exploration) en fonction du niveau de précision du cahier des charges et du degré de l'innovation souhaité. Pour le projet *Smart Angel*, les partenaires doivent répondre à un cahier des charges fonctionnel (Gillier, 2010) ce qui classifie ce partenariat au rang de **co-développement**. La particularité du co-développement porte sur la nature des échanges entre les partenaires, car le client est aussi fournisseur et l'objet de la collaboration est partiellement connu (Midler, Maniak, & Beaume, 2007). Les projets d'open innovation sont soumis à de nombreux freins. Aubert et Patry (2004) ont mis en évidence certaines de ces complications comme les difficultés de langage entre les partenaires ou encore les incompatibilités organisationnelles surtout dans les partenariats publics-privés.

Le projet *Smart Angel* fait intervenir des médecins, infirmiers, ingénieurs, graphistes, ergonomes, psychologues, informaticiens, tous liés au sein du même consortium. Afin de représenter cette diversité de compétences agencées vers un objectif commun (Figure 1, ci-dessous), nous avons détourné l'utilisation de la méthode du « *Matching-Building* » (Kazakçı, Gillier, & Piat, 2008, p. 28), inspirée de la théorie C-K de Hatchuel et Weil (2008), représentant les compétences et les activités de chaque partie prenante, ajoutant à son tour sa brique à la conception du produit final. Si un élément manque à la chaîne, ce sont les éléments suivants qui restent en attente. Par exemple, si le contenu de l'application n'est pas fourni par le CHU de Nîmes, le développement de l'application ne peut pas intégrer ces éléments, ce qui ne permettra pas au CRP-CPO de proposer des recommandations ergonomiques et de réaliser des améliorations sur l'application.

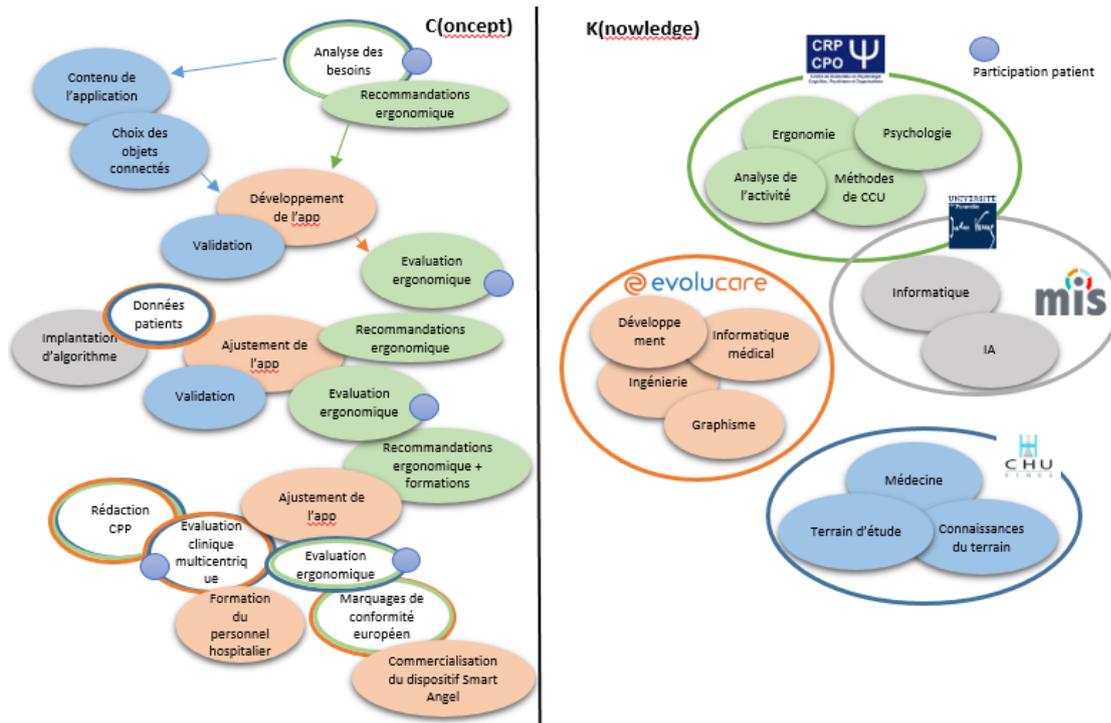


Figure 1. Représentations de la dispersion des activités des différentes parties prenantes du projet *Smart Angel* en ambulatoire à l'aide de la méthode du Matching-Building (Kazakçı, Gillier, & Piat, 2008).

2.3. Un projet de co-conception : les partenaires utilisateurs intégrés dans une démarche centrée utilisateur

Le dernier partenaire important que nous n'avons pas encore abordé est l'utilisateur final. Le projet *Smart Angel* s'engage dans une **démarche centrée utilisateur (ISO 9241-210, 2019)**. Il existe actuellement un consensus sur l'idée que la technologie doit répondre aux besoins de l'utilisateur (Piau et al., 2014) et de nombreux exemples d'approches se sont développés dans la littérature scientifique (e.g., Matthew-Maich et al., 2016 ; LeRouge & Wickramasinghe, 2013 ; Schnall, Rojas, et al., 2016a). L'objectif est d'identifier les contraintes, les exigences, les besoins, les préférences et les particularités cognitives ou sociales des futurs utilisateurs. Ces études basées sur des méthodes de conception centrées sur l'utilisateur ont montré que les dispositifs créés avaient une plus grande facilité d'utilisation parce qu'ils prenaient en compte des éléments précités de l'utilisateur final. Cela permet également de proposer des recommandations universelles pour ces dispositifs (Matthew-Maich et al., 2016 ; Zhang & Adipat, 2005). Les méthodes centrées utilisateurs prennent la forme de processus itératifs où à la fin de chaque itération, une évaluation est mise en place. Les patients, les soignants, les professionnels de la santé et d'autres personnes participent à ces évaluations, ce qui ajoute une spécificité supplémentaire à l'étude future des modèles d'évaluation de la santé.

Ces études doivent donc être collaboratives entre les patients et les professionnels de santé.

En effet, le personnel de santé est souvent absent dans les évaluations et dans l'éducation du patient (e.g., Andersen et al., 2017 ; Ten Haken et al., 2018). Ces études montrent que la priorité est principalement donnée au patient en termes d'autogestion ou d'« autosurveillance » (ou monitoring). Cependant, Andersen et al. (2017) démontrent, au moyen des données recueillies entre les thérapeutes et les patients, l'importance de la collaboration et de la relation patient-soignant sur la motivation des patients. Huvila (2016) déclare qu'on ne devrait pas parler d'autosurveillance, mais de « *co-surveillance* » (ou « *co-monitoring* »). Derrière chaque dispositif médical, il y a des soignants qui s'occupent de la santé du patient. **Cette relation doit faire partie du processus de conception et implique une approche collaborative et évaluative.** La grande complexité de ces objets nécessite un suivi rapproché d'un professionnel et une validation ergonomique à chaque étape de sa conception. Nous reviendrons sur la conception centrée utilisateur dans la partie théorique.

3. L'organisation du manuscrit de thèse

Le dispositif *Smart Angel* appartient au domaine de la santé mobile (ou m-santé) et promet de nombreux avantages pour le système de santé. Pourtant, ces technologies se heurtent à de nombreux freins : la sécurité des données, l'atteinte à la vie privée, les bouleversements organisationnels et la complexité d'utilisation. Cette dernière est particulièrement pointée du doigt par la communauté scientifique comme étant le challenge majeur de ce domaine émergent (e.g., Carayon & Hoonakker, 2019; Cifter, 2017; Jaspers, 2009; Kushniruk, 2002; Peute et al., 2008; Yen, 2010; Zhang & Adipat, 2005). A l'origine, ce matériel était prévu pour le personnel soignant formé. La démocratisation de ces produits les a conduits à être utilisés par un large public novice. Cependant, la complexité de ce matériel ne permet pas toujours cette démocratisation et peut avoir des conséquences dramatiques pour la sécurité du patient lorsque ces produits sont laissés sans supervision d'un professionnel. Cette thèse a, par conséquent, pour objectif d'accompagner la conception du dispositif de m-santé *Smart Angel* au sein du processus de conception centrée utilisateur en se focalisant sur les problématiques d'utilisabilité du système.

Ce manuscrit de thèse s'organise en trois parties : théorique, empirique et discussion générale.

La partie théorique est articulée autour de trois chapitres. L'ambition du chapitre 1 est de positionner et d'ancrer nos recherches au sein d'un cadre théorique et méthodologique de l'utilisabilité. Le chapitre 2 met en exergue la récente littérature scientifique dans le domaine de l'utilisabilité appliquée aux dispositifs connectés en santé pour le domicile patient afin de mieux comprendre les avancées actuelles sur le sujet. Le chapitre 3 présente une synthèse des éléments théoriques permettant d'argumenter notre problématique de recherche abordant les manquements de la littérature scientifique dans le domaine de la conception et de l'évaluation de l'utilisabilité des dispositifs médicaux connectés.

L'organisation de la partie empirique, composée de cinq chapitres, suit la chronologie du développement du dispositif *Smart Angel* et l'élaboration d'un modèle de l'évaluation de l'utilisabilité nommé HANDLED⁴. Le chapitre 1 (étude exploratoire) correspond à l'analyse du besoin en phase amont de conception. Cette étude a permis de comprendre le terrain dans lequel s'intégrera le dispositif *Smart Angel* au sein du parcours de soin et de proposer les premières maquettes du produit. Les chapitres 2 (étude 1a) et 3 (étude 1b) détaillent les évaluations de l'utilisabilité des objets connectés présélectionnés par le partenaire clinique qu'est le CHU de Nîmes. Précisément, nous nous sommes intéressées à l'impact des caractéristiques utilisateurs ainsi qu'à l'impact des documents procéduraux (ressources réutilisables) sur l'utilisabilité de ces objets connectés. Le chapitre 4 (étude 2) décrit l'évaluation de l'utilisabilité du prototype 1 du dispositif *Smart Angel* composé d'une première version de l'application *Smart Angel* associée aux objets connectés pré-évalués. Le chapitre 5 (étude 3) aborde l'utilisabilité du prototype 2 de *Smart Angel* en incluant l'environnement hospitalier et les formations hospitalières délivrées avant le retour du patient à son domicile.

La dernière partie de ce travail de thèse discute d'un modèle d'évaluation de l'utilisabilité adapté à la santé au travers des contributions théoriques et méthodologiques de cette thèse. Nous présentons aussi les limites et contraintes de notre travail de thèse qui nous amènent à proposer des perspectives de recherche.

⁴ HANDLED signifie « *manier* » ou « *manipuler* » en anglais.

PARTIE I : PARTIE THÉORIQUE

CHAPITRE 1 : POSITIONNEMENTS DISCIPLINAIRES, APPROCHES THÉORIQUE ET MÉTHODOLOGIQUE DE L'UTILISABILITÉ DES TECHNOLOGIES

1. Le positionnement disciplinaire basé sur une conception centrée utilisateur : une démarche pluridisciplinaire au sein d'une activité collaborative

Ce travail de thèse s'insère dans un **projet de conception pluripartenaire s'appuyant sur une conception centrée utilisateur**. Dans cette partie, nous détaillons dans un premier temps, le cadre, c'est-à-dire la démarche de la conception centrée utilisateur, dans laquelle s'inscrit notre positionnement disciplinaire que nous détaillerons dans un deuxième temps.

1.1. La démarche de conception centrée utilisateur

La conception centrée utilisateur est cadrée par la norme ISO 9241-210 (2019) la définissant comme une démarche centrale intégrant l'étude par des experts ou/et l'activité de l'opérateur humain, tout au long du processus de conception. Ce processus comprend trois phases (ISO 9241-210, 2019) impliquant les futurs utilisateurs (Figure 2) :

1. L'analyse du contexte et des spécificités des utilisateurs.
2. La définition des activités futures et le traitement des problématiques issues du contexte identifié.
3. L'évaluation de l'acceptabilité, l'utilisabilité et l'utilité des solutions de conception.

La norme ISO 9241-210 (2019) spécifie que ces trois étapes doivent être réitérées tout au long du processus de conception du produit. Un cycle itératif en conception centrée utilisateur est généralement constitué de quatre activités (Tory & Moller, 2004) :

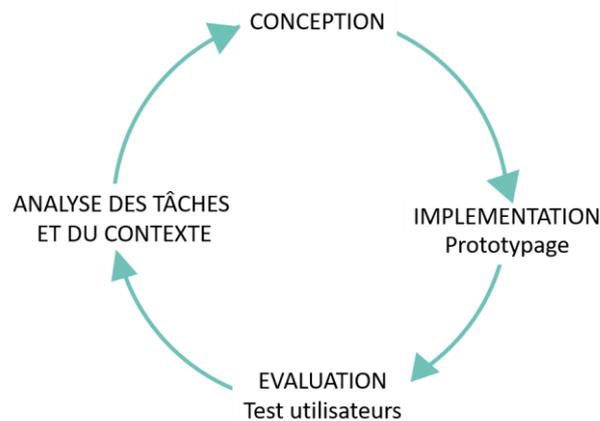


Figure 2. Le processus itératif de conception centrée utilisateur inspiré de ISO 0241-210 (2019) et Tory et Moller (2004).

1.2. Un positionnement pluridisciplinaire : de l’ergonomie à la psychologie cognitive en passant par les sciences de conception.

Afin de respecter les exigences de la conception centrée utilisateur, nous nous appuyons sur trois disciplines majeures abordant ce sujet (l’ergonomie, la psychologie cognitive et les sciences de la conception) au sein d’une activité collaborative.

1.2.1. L’ergonomie

L’ergonomie (ou *Human Factors*) est définie en 2000 par l’International Ergonomics Association ou IEA comme « *la discipline scientifique qui vise la compréhension fondamentale des interactions entre les humains et les autres composantes d’un système, et la profession qui applique principes théoriques, données et méthodes en vue d’optimiser le bien-être des personnes et la performance globale des systèmes* » (Lancry, 2009, p. 20). C’est un domaine par définition pluridisciplinaire (Sperandio, 1980). Dans son ouvrage intitulé "Ergonomie", Falzon (2004) explique qu’à l’instar des autres disciplines des sciences humaines, sociales et cognitives, **l’ergonomie se préoccupe davantage du contexte** en développant des approches holistiques impliquant à la fois les contraintes physiologiques, cognitives, environnementales et sociales. Au sein de cette discipline, plusieurs spécialisations émergent telles que l’ergonomie physique, cognitive et organisationnelle (Falzon, 2004). Cette thèse se situe inévitablement dans une approche centrée sur l’utilisateur. De ce fait, ce travail s’intègre précisément dans le champ de **l’ergonomie cognitive**. Cette science trouve ses racines dans la recherche en informatique sur l’Interaction Humain-Machine ou IHM⁵ et comme son nom

⁵ En anglais : *Human-Computer Interaction*

l'indique, se centre sur les interactions entre l'humain et son environnement de travail cognitif (Green & Hoc, 1991). L'ergonomie cognitive fait avant tout partie du domaine des sciences cognitives et trouve dans la psychologie cognitive, un appui théorique majeur. Pourtant, l'ergonomie cognitive, dont l'objectif principal est la prise en compte des facteurs humains dans le processus de conception et/ou des situations de travail complexes, se détache de ces deux précurseurs pour trouver son originalité dans ses méthodes en laboratoire et hors laboratoire. L'ergonomie cognitive fait de l'outil le prolongement de l'humain (Spérando, 2001) à travers la prise en compte des critères d'analyse ergonomique (utilité, utilisabilité et acceptabilité). Cette thèse est aussi une thèse dite « **appliquée** » **et utilise l'ergonomie cognitive** (ou psychologie ergonomique (Leplat, 1980)) de conception **à travers ses approches** (i.e., la conception centrée utilisateur) **et les critères ergonomiques** qu'elle investigate (e.g., **l'utilisabilité**). L'ergonomie a ainsi permis de nous questionner sur **le contexte d'usage du dispositif**.

1.2.2. La psychologie cognitive

Notre thèse s'appuie en grande partie sur des méthodes issues de l'ergonomie cognitive qui intègre notamment des recherches fondamentales de la psychologie cognitive. Cette dernière a pour objectif l'étude des fonctions cognitives humaines : la mémoire, le langage, le raisonnement, la perception, la résolution de problèmes, l'attention, l'intelligence et les émotions. Lemaire et Didierjean (2018, p.27) déclarent que « *l'objectif de la psychologie cognitive est de spécifier les représentations mentales et les processus opérant sur ces représentations pour accomplir les tâches cognitives auxquelles sont soumis les sujets* ». Pour étudier le fonctionnement de ces grandes fonctions, les chercheurs du domaine se servent de modèles mentaux (e.g., Baddeley, 1992; Dixon, 1987; Mayer, 2001) que l'on retrouve ensuite en ergonomie cognitive. Les méthodes utilisées en psychologie cognitive se regroupent en trois grands types d'observation du comportement (Lemaire & Didierjean, 2018) : l'observation naturelle, l'observation corrélacionnelle et l'observation expérimentale. Cette dernière couplée à la modélisation est particulièrement utilisée en psychologie cognitive **Notre travail de thèse mobilise cette approche expérimentale afin de prendre en compte les éléments cognitifs et contextuels impactant la conception (plus exactement l'utilisabilité) des dispositifs médicaux.**

Enfin, ce travail trouve également certaines inspirations dans la psychologie de la santé. Ce domaine plutôt récent de la psychologie vise à l'étude « *des différences interindividuelles en matière de santé et de maladie et leurs déterminants* » (Bruchon-Schweitzer & Siksou, 2008,

p. 28). Cette psychologie se base sur des modèles intégratifs et multifactoriels afin d'analyser l'état de santé de la population. **Nous nous sommes intéressées aux facteurs environnementaux, sociaux et cognitifs afin de comprendre les comportements en santé des patients.**

Ce positionnement est d'ailleurs visible au travers des différentes publications et communications de la thèse (cf. Annexe A : Liste des publications, p. 417) dans des colloques nationaux et internationaux aux thématiques très variées (en ergonomie : EPIQUE et IEA, en psychologie : AFPSA et SFP).

1.2.3. Les sciences de la conception

Cette thèse s'inscrit dans un contexte de conception de produit renvoyant aux sciences de la conception. Hubka et al. (1987) définissent les sciences de la conception (en anglais, *design science*) par un ensemble de connaissances et de compétences (telles que des concepts d'informations techniques et méthodologiques) appliquées au domaine de la conception. Aoussat, Christofol, et Coq (2000) représentent l'activité de conception par une approche transversale entre tous les acteurs-métiers. Ce travail de thèse est élaboré dans un contexte de travail collectif en conception. Ces approches collaboratives sont de plus en plus représentées dans les processus de conception de produit (Darses et al., 2001). Cette forme d'innovation permet d'assembler les compétences et les connaissances de nombreuses parties prenantes. Chacun possède un rôle bien délimité, néanmoins les tâches et les buts sont mis en commun. C'est ce que définit Boutigny (2004) par le terme **collaboration**. Cependant, l'utilisation combinée de différentes compétences n'est pas sans écueil (Simon, 2017). En outre, le langage, la non-proximité géographique, les centres d'intérêt sont autant de raisons de création de frustrations et de quiproquos dans cette importante équipe où le choix du partenaire est un élément majeur de la réussite d'un projet d'innovation (e.g., Chaniaud & Jean, 2017).

Ainsi, en plus de s'inscrire dans le champ de l'ergonomie et de la psychologie, ce travail de thèse prend aussi appui dans le champ disciplinaire de la conception de produit renvoyant aux disciplines des sciences de l'ingénieur, sciences de gestion et plus précisément dans les sciences de la conception (Papalambros, 2015) en s'inspirant du management de l'innovation (De Woot, 1970). **Les sciences de la conception permettent de cadrer les modalités de notre travail au sein même d'un processus de conception à travers un ensemble d'acteurs-métiers et ainsi de mieux comprendre les enjeux et prévenir les complications entre chaque partie prenante de ce processus.** C'est pourquoi, en complément d'un état de l'art

théorique, méthodologique et conception, nous avons défini un état des lieux du contexte collaboratif (cf. Introduction Générale, 2.2. Un projet de co-développement : les partenaires industriels, académiques et hospitaliers, p. 4) afin de prendre en compte l'environnement collaboratif et de rester cohérents dans nos pratiques par rapport à ce contexte partenarial.

1.3. Une thèse pluridisciplinaire dans une modalité de travail collectif

Ce travail de thèse s'intègre dans un processus de **conception centrée utilisateur (patients et soignants)** et dans un partenariat complexe **aux modalités de travail collectif** composé de divers acteurs-métiers. De plus, cette thèse se positionne aux croisements d'un ensemble de disciplines (ergonomie, psychologie et sciences de la conception), ce qui la caractérise de thèse **pluridisciplinaire**. La pluridisciplinarité est définie par Hamel (1997), en contraste à la multidisciplinarité, par l'établissement de rapport entre les disciplines s'utilisant de manière combinée sans les modifier.

Ainsi, notre travail pourrait s'appréhender de manière **multidimensionnelle** ; compte-tenu de son **positionnement pluridisciplinaire**, mais également étant donné ses modalités de **travail pluripartenarial** entre les diverses **parties prenantes (acteurs-métiers)** et les **utilisateurs finaux**. C'est sur ce positionnement disciplinaire que nous abordons la notion de l'utilisabilité et les concepts connexes, centraux dans la conception centrée utilisateur.

2. Les approches théoriques et opérationnelles de l'utilisabilité : une notion centrale dans la conception centrée utilisateur

2.1. Les concepts connexes de l'utilisabilité

2.1.1. L'acceptabilité pratique, sociale et située

L'acceptabilité est définie comme « *le degré d'intégration et d'appropriation d'un objet dans un contexte d'usage* » (Barcenilla & Bastien, 2009, p.311). L'acceptation est alors vue comme un processus d'intégration de la technologie s'insérant au sein des activités des utilisateurs et contribue à les transformer. L'acceptabilité se situe en amont de son implantation (e.g., Bobillier-Chaumon & Dubois, 2009) et est ensuite remplacée par la notion d'acceptation une fois l'implantation réussie. L'acceptabilité peut être décomposée en trois orientations théoriques différentes (Bobillier-Chaumon, 2016 ; Nielsen, 1993) :

- L'acceptabilité pratique (ou opératoire) vise à favoriser l'acceptabilité de la technologie en s'assurant de la qualité des IHM. L'acceptabilité pratique vise à assurer la compatibilité entre l'utilisateur, ses tâches et la technologie (Brangier et al., 2010) afin de favoriser l'adoption de la technologie (Bobillier-Chaumon, 2016).
- L'acceptabilité sociale a pour objectif de déterminer l'intention d'usage des utilisateurs. Sagnier et al. (2019) définissent deux orientations de modèles : ceux centrés sur la satisfaction des utilisateurs, et ceux se basant sur la perception de la technologie des utilisateurs (e.g., Technology Acceptance Model ou TAM de Davis (1989) basé sur l'utilité perçue et la facilité perçue). L'utilité perçue renvoie à « *l'intensité avec laquelle un individu croit que l'utilisation d'un système améliorera sa performance au travail* » (Bobillier-Chaumon et al., 2006, p. 249). La facilité d'utilisation perçue fait référence à « *l'intensité avec laquelle un individu croit que l'utilisation du système se fera sans difficulté ni effort supplémentaires* » (Bobillier-Chaumon et al., 2006, p. 249).
- L'acceptabilité/acceptation située (Bobillier-Chaumon, 2016) est centrée sur l'activité, le vécu et l'expérience réelle des utilisateurs avec la technologie. Cette analyse cherche à identifier les apports et les limites d'une technologie dans son contexte d'usage. Dans cette perspective, on s'intéresse concrètement à ce que « *la technologie "permet/autorise de faire" ou "oblige à faire", mais aussi ce qu'elle "empêche de faire" ou "plus comme avant" et ce, sur différentes dimensions de l'activité* » (Bobillier Chaumon, 2016, p. 12).

2.1.2. L'utilité-valeur et l'utilité-destination

L'utilité (*usefulness*) est définie par Loup-Escande et al. (2013, p.5) comme une propriété construite et évolutive « *se situant à l'interface entre le système, le sujet et la situation, correspondant à une réponse à un besoin ou à une attente de l'utilisateur* ». L'utilité peut être appréhendée sur la base de deux concepts interdépendants (Loup-Escande et al., 2013) : ce à quoi l'objet est nécessaire au sens de satisfaire un besoin créant ainsi un bénéfice ou une plus-value pour celui qui l'utilise (la « valeur »), et l'utilité associée à ce pourquoi l'objet a été conçu pour, ce à quoi l'objet est destiné (la « destination »). Ce dernier renvoie davantage à la description du système et une utilisation précise ce que Nielsen (1993) qualifie d'utilité pratique et théorique. La différence dans ces deux concepts est liée au réel. L'utilité théorique renvoie à l'objectif théorique que doit atteindre l'objet, ce pour quoi il a été conçu et l'utilité pratique renvoie à l'objectif réel que peut atteindre l'utilisateur dans un contexte particulier.

2.2. L'utilisabilité : quelques définitions

Le terme d'utilisabilité est apparu dans les années 1980 (Shackel, 1981; 1986) grâce au courant de recherche portant sur les IHM et l'engouement pour le numérique dans la société. Il désigne dans son sens le plus large la « *facilité d'utilisation/apprentissage* » voire la « *qualité d'utilisation* » (Bevan, 1995b) d'un produit ou d'un service. Norman (1987) parle « *d'ingénierie cognitive* » (« *cognitive engineering* ») pour désigner l'implication de connaissances et compétences dans l'utilisation des technologies modernes. L'utilisabilité est alors en plein essor et va progressivement constituer un vaste champ de recherche à la jonction entre plusieurs disciplines : la psychologie, l'ergonomie, l'informatique et les sciences de la conception. **En 40 ans, les définitions et les modèles d'utilisabilité vont défiler, mais aucun d'entre eux ne fera réellement consensus.** De nombreuses définitions se succéderont et se regrouperont pour former les définitions ayant servi à l'élaboration de différents standards (e.g., IEEE std. 610.12, 1990 ; ISO/IEC 9126-1, 2001 ; ISO 9241-11, 2018). Brangier et Barcenilla (2003, p. 41) déclarent que « *le contenu de la notion d'utilisabilité dépend de ce qui est mesuré et de la manière dont cette mesure est réalisée. La définition de l'utilisabilité dépend donc de la façon de l'opérationnalité* ». Chaque définition donne lieu à des métriques⁶ différentes ancrées dans des modèles théoriques. Dans toutes les définitions, l'utilisabilité est définie par sa multidimensionnalité et est composée d'une architecture sur plusieurs strates. On y trouve ses composants principaux (métriques), regroupant des sous-dimensions, ces dernières rassemblant des caractéristiques. Nous avons choisi dans un premier temps de présenter les trois principaux modèles de l'utilisabilité (Shackel, 1991 ; Nielsen, 1993 ; ISO 9241-11) puis nous avons finalisé notre revue de littérature avec l'utilisabilité opérationnelle au moyen d'un recensement des métriques constituant les modèles proposés sur les quarante dernières années (cf. Annexe 1 : Revue des modèles d'utilisabilité de 1991 à 2018, p. 338).

2.3. L'utilisabilité : les trois principaux modèles théoriques

2.3.1. Le modèle de Shackel (1991)

a. Présentation du modèle

Brian Shackel (1991, p. 24) est un précurseur de l'utilisabilité qu'il définit comme « *la capacité, en termes fonctionnels humains, à permettre une utilisation facile et effective par une*

⁶ Nous considérons le terme « métrique » comme un élément composant l'utilisabilité pouvant être mesurable (e.g., l'efficacité, l'efficience, etc.).

catégorie donnée d'utilisateurs, avec une formation et un support adapté, pour accomplir une catégorie donnée de tâches, à l'intérieur d'une catégorie spécifique de contextes ». Cette définition, comme celle de Shneiderman (1992), se centre sur la nécessité d'évaluer la vitesse d'accomplissement de la tâche et le taux d'erreurs. L'auteur considère l'utilisabilité comme un concept multidimensionnel influencé par le contexte d'utilisation incluant l'utilisateur final.

Le modèle de Shackel (Figure 3) considère l'utilisabilité sur la base de quatre métriques :

- **L'efficacité** se caractérise par la vitesse de réalisation de la tâche et d'erreurs produites. Il s'agit de la capacité à atteindre l'objectif utile du produit ;
- La **flexibilité** est définie comme l'adaptation de l'utilisateur en fonction des différents contextes d'utilisation ;
- **L'apprenabilité** (ou facilité d'apprentissage) est déterminée par le délai temporel entre la formation des utilisateurs et la mise en service du produit en fonction du nombre d'essais et de l'assistance apportée aux utilisateurs ;
- **L'attitude**, que d'autres auteurs qualifient de satisfaction, est définie par l'approbation en termes de fatigue, d'inconfort, de frustration et d'efforts.

Shackel (1991) considère ces métriques comme des propriétés intrinsèques d'un produit représenté par des aspects objectifs et subjectifs.

En résumé, l'utilisabilité, d'après Shackel (1991), est une caractéristique de conception d'un produit dépendante du contexte d'utilisation dans lequel l'interaction entre l'utilisateur et ce produit aura lieu. Ainsi, le contexte d'utilisation étant dépendant de l'utilisateur, de l'environnement, de la tâche et de l'équipement, l'utilisabilité obtenue ne sera pas transposable dans toutes les situations. Ce modèle a inspiré les modèles suivants.

b. Avantages et inconvénients

La définition de Shackel est relativement complète considérant l'utilisabilité comme objective et subjective. Cependant il manque, selon nous, certains éléments comme des précisions sur la prise en compte d'éléments du contexte (e.g., le profil des utilisateurs). De plus, Shackel (1991) met davantage l'accent sur les caractéristiques instrumentales de l'utilisabilité et délaisse toutefois les caractéristiques non instrumentales comme la satisfaction.

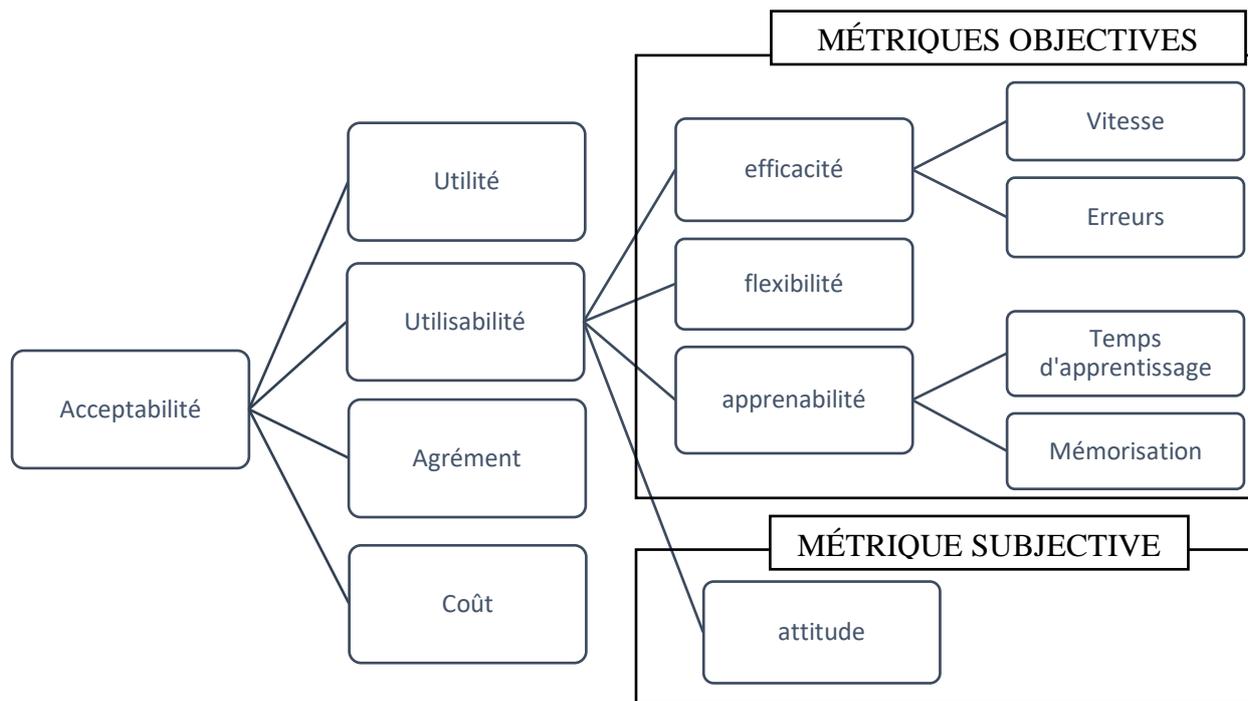


Figure 3. Modèle de l'acceptabilité des systèmes selon Shackel (1991).

2.3.2. Le modèle de Nielsen (1993)

a. Présentation du modèle

Jakob Nielsen est une figure incontournable de la conception centrée utilisateur. L'auteur considère l'utilisabilité comme la façon dont les utilisateurs vont pouvoir utiliser les fonctionnalités utiles d'un produit. Il positionne l'utilisabilité au sein d'un modèle plus large de l'acceptabilité. Tout comme Shackel (1991), Nielsen (1993) considère l'utilisabilité comme multidimensionnelle et c'est pour cette raison que l'auteur ne définit pas précisément cette notion au moyen d'une définition, mais propose un modèle opérationnel où chacune des métriques est définie.

Le modèle de Nielsen (1993) détermine l'utilisabilité sur la base de cinq métriques relativement proches de celles de Shackel (1991) qu'il faut prioriser en fonction de ses attentes (Figure 4) :

- **La facilité d'apprentissage** est définie par la vitesse d'apprentissage afin que l'utilisateur puisse accomplir les tâches du système rapidement ;
- **L'efficience** se caractérise par la vitesse de productivité (une fois le processus de manipulation du produit appris) ;

- **L'apprenabilité (ou mémorabilité)** est déterminée par la facilité à retenir les informations nécessaires à l'utilisation du produit afin que l'utilisateur puisse réutiliser le produit après un certain temps sans difficulté ;
- **La prévention des erreurs** est définie par le nombre d'erreurs et la correction des erreurs s'il y en a, lors de l'utilisation du système ;
- **La satisfaction** est définie par le caractère plaisant lors de l'utilisation afin de satisfaire l'utilisateur d'un point de vue subjectif.

Tricot et al. (2003) précisent que ce modèle doit être lu de droite à gauche, car l'utilisabilité doit être prise en compte à tous les niveaux de l'interaction. Tout comme le modèle de Shackel (1991), l'utilisabilité n'est plus définie comme résultant d'une simple interaction permettant d'accomplir la tâche principale, mais devient davantage holistique et opérationnelle.

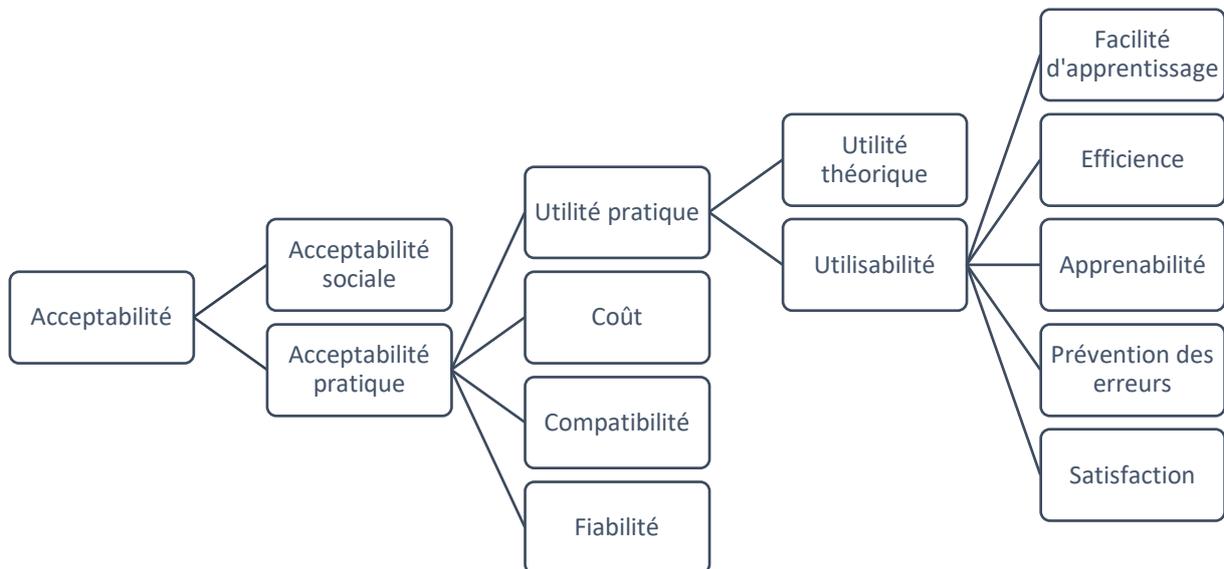


Figure 4. Modèle de l'acceptabilité des systèmes selon Nielsen (1993).

b. Avantages et inconvénients

Le modèle de Nielsen est l'un des modèles les plus cités dans la littérature. Il situe l'utilisabilité au sein d'un modèle plus complexe que nous discutons dans la partie suivante. Néanmoins, ce modèle possède quelques lacunes. Les métriques d'utilisabilité de Nielsen pourraient être enrichies en fonction des besoins d'évaluation d'un système. Par exemple, ce modèle pourrait intégrer l'efficacité (dans une métrique à part entière) comme le préconisent Shackel (1991) et la norme ISO 9241-11 (2018). De plus, le modèle de Nielsen n'intègre pas la notion située de l'utilisabilité en y ajoutant des composants contextuels comme dans la norme

ISO 9241-11 (2018) ou une métrique de flexibilité comme dans le modèle de Shackel (1991). Ce contexte est d'autant plus important, qu'il joue un rôle de modérateur sur le résultat des mesures d'utilisabilité. Par exemple, Dillon et Watson (1996) critiquent les tendances de certaines études à la non prise en compte des différences inter et intra individuelles (e.g., l'expérience, le niveau de compétences, les connaissances du domaine, et l'âge). Marcus (2002) définit cette prise en compte du contexte par la notion d'accessibilité l'associant au design universel.

2.3.3. Le modèle proposé par la norme ISO 9142-11 (2018)

a. Présentation du modèle

La norme ISO (ISO 9241-11, 1998), la plus connue et reconnue en matière d'utilisabilité s'appuie en partie sur la définition de Shackel (1991), Nielsen, (1993), Preece et al. (1994) et Constantine et Lockwood (1999). La mise à jour de cette norme (ISO 9241-11, 2018) s'appuie sur d'autres normes : ISO 13407 (1999) sur *le processus de conception centrée sur l'opérateur pour les systèmes interactifs* (révisée ISO 9241-210, 2019) et ISO/TR 18529 (2000) sur *les descriptions des processus cycle de vie centrés sur l'opérateur humain*. Ces deux standards composent aujourd'hui la norme ISO 9241-210 (2019) que nous avons présentée précédemment (cf. Partie I, Chapitre 1, 1.1. La démarche de conception centrée utilisateur, p. 9). Cependant, la définition de l'utilisabilité proposée par la norme ISO 9241-11 n'a pas changé entre 1998 et 2018. La norme ISO 9241-11 (2018, p. 2) définit l'utilisabilité comme « *Le degré selon lequel un produit peut être utilisé, par des utilisateurs identifiés, pour atteindre des buts définis avec efficacité, efficience et satisfaction, dans un contexte d'utilisation spécifié* ».

Le modèle de la norme ISO 9241-11 (2018) (Figure 5) définit l'utilisabilité sur la base de trois métriques :

- **L'efficacité** est définie par « *la précision ou degré d'achèvement selon lesquels l'utilisateur atteint des objectifs spécifiés* » (ISO 9241-11, 2018, p. 3). La norme ne spécifie pas exactement des attributs à mesurer. Elle explique en revanche que les manquements d'efficacité entraînent des erreurs ou des difficultés d'utilisation et que des fonctionnalités inutiles du système interfèrent avec la tâche de l'utilisateur. Un manque d'efficacité d'un système peut aussi entraîner des préjudices⁷ pour l'utilisateur (personne ou institution) comme pour le concepteur. Ces conséquences peuvent être des préjudices physiques, émotionnels ou pécuniers pour l'individu. Cela peut aussi

⁷ NOTE 2 de la norme ISO 9241-11 (2018, p.10)

dégrader la sécurité, la finance ou la réputation d'une entreprise (concepteur ou utilisateur) ;

- **L'efficience** est décrite comme « *le rapport entre les ressources utilisées et les résultats obtenus* » (ISO 9241-11, 2018, p. 3). Ces ressources peuvent être des ressources humaines (un effort humain pouvant être physique ou cognitif), des ressources temporelles et des ressources financières (coût engendré). Bevan et Macleod (1994) proposent une mesure mathématique de ces trois types d'efficience (humaine, temporelle et financière). Ces trois éléments seraient la résultante de l'efficacité divisée par l'effort (humain), le temps d'accomplissement (temporel) et le coût total (financier) ;
- **La satisfaction** : La définition de la norme ISO 9241-11 (2018) se concentre sur le confort des utilisateurs et des attitudes positives envers l'utilisation du système. Il s'agit du « *degré selon lequel les réactions physiques, cognitives et émotionnelles de l'utilisateur qui résultent de l'utilisation d'un système, produit ou service répondent aux besoins et attentes de l'utilisateur* » (ISO 9241-11 :2018, p.4). La mise à jour de cette norme a permis d'inscrire l'expérience utilisateur dans la notion d'utilisabilité (Bevan et al., 2015) et de prendre en compte les réactions avant, pendant et après l'utilisation du système.

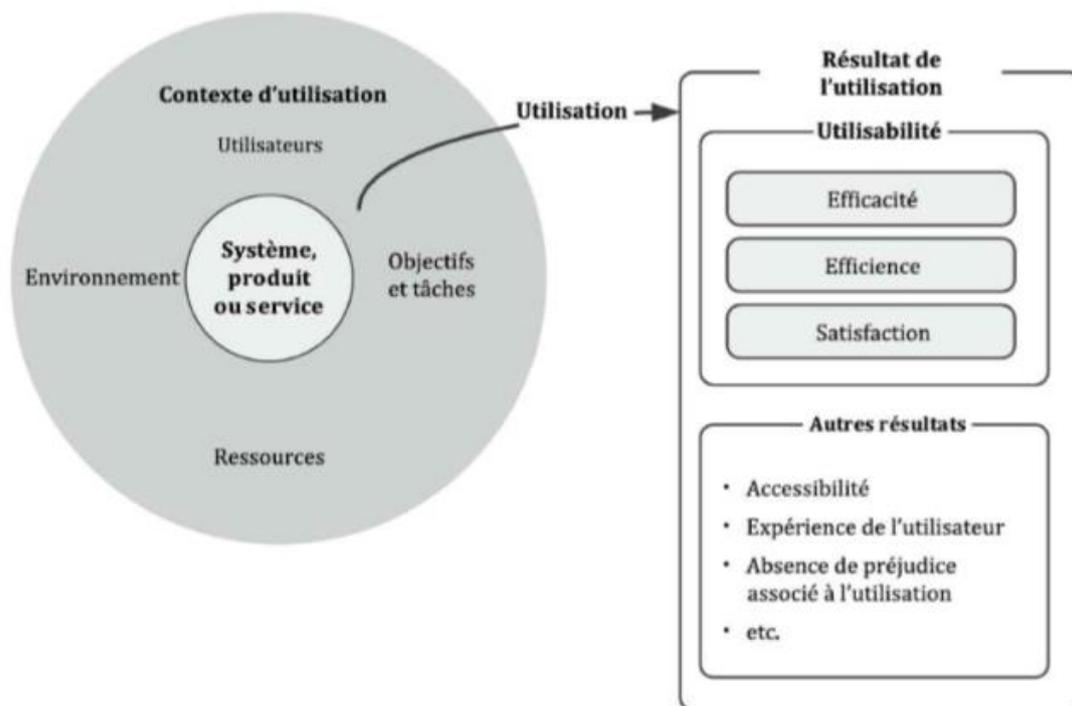


Figure 5. Représentation de l'impact des éléments du contexte d'utilisation sur l'utilisabilité d'après la norme ISO 9241-11 (2018, p. 7).

b. Avantages et inconvénients

Ce standard possède vraisemblablement le modèle le plus abouti au vu du recul et de ses mises à jour (ISO 9241-11, 1998, 2018). Sa définition n'a pas changé, mais ses concepts connexes ont évolué comme celui de l'*expérience utilisateur* relative à la composante émotionnelle prenant de plus en plus d'importance (Bevan, 2009) contrairement aux modèles de Nielsen (1993) et Shackel (1991). L'utilisabilité devient également la qualité d'une interaction plutôt que la propriété d'un produit (Bevan et al., 2015). La norme fournit aussi un détail des mesures des différentes métriques ainsi que le détail des composants du contexte d'usage peu détaillé dans les modèles précédemment décrits (le contexte intègre les tâches, l'environnement, les utilisateurs et les ressources).

Malgré ces ajouts, cette norme ne prend pas en considération les métriques d'apprenabilité (à savoir, facilité d'apprentissage et mémorisation) proposées par les modèles de Shackel (1991) et de Nielsen (1993). Enfin, cette norme qui est le standard international des exigences ergonomiques ne semble toujours pas convaincre l'ensemble de la communauté scientifique dans les domaines ayant recours à l'utilisabilité (e.g., Tractinsky, 2018; Weichbroth, 2018). Nous reviendrons sur ce débat qui anime l'utilisabilité (cf. Partie I, Chapitre 1, 2.5. L'utilisabilité, des divergences disciplinaires, critérielles et applicatives, p. 24)

2.4. Les critiques des modèles et des métriques de l'utilisabilité

2.4.1. Le positionnement de l'utilisabilité en regard de l'acceptabilité, et l'utilité dans l'évaluation ergonomique

L'utilisabilité, l'acceptabilité et l'utilité sont les trois concepts majeurs à considérer dans une évaluation ergonomique (e.g., Bastien & Tricot, 2008 ; Loup-Escande, Burkhardt, & Richir, 2013; Nielsen, 1993). Plusieurs auteurs se sont d'ailleurs intéressés aux positionnements de ces trois concepts. En fonction des auteurs, ces concepts ne sont pas hiérarchisés de la même manière.

L'utilisabilité et l'utilité sont souvent considérées comme des prérequis à l'acceptabilité (e.g., Brangier & Barcenilla, 2003 ; Shackel, 1991), à côté d'autres facteurs comme « *la capacité à être facilement pris en main* », le prix, l'esthétisme et la valeur affective et sociale. Pour ces auteurs, l'utilisabilité et l'utilité ont le même degré d'importance. Cependant Shackel (1991) s'interroge sur la place de l'utilisabilité qui tendrait à devenir une dimension centrale afin de répondre aux exigences de l'acceptabilité. Similairement à Shackel (1991), Nielsen

(1993) considère l'utilisabilité comme un facteur de l'acceptabilité d'un dispositif. Cependant, contrairement à Shackel (1991), il considère que l'utilité dite « pratique » (qualité d'utilisation) est un prérequis à l'acceptabilité pratique mais pas à l'acceptabilité sociale (ou située). Nielsen (1993) ajoute à son modèle l'idée selon laquelle l'utilisabilité et l'utilité théorique seraient des prérequis à l'utilité pratique. La norme ISO 9241-11 associe l'utilisabilité à la qualité de l'interaction produite entre l'utilisateur et le système. L'acceptabilité est alors abordée dans la notion de satisfaction. Le modèle de Ketola et Roykkee (2001) propose une évolution du modèle de Nielsen (1993). Les auteurs définissent les différents types d'acceptabilité (pratique et sociale) avant de les mettre en lien direct avec les différents types d'utilisabilité (objective et subjective). L'utilisabilité devient alors un prérequis unique à l'acceptabilité et ce n'est qu'une fois qu'un produit est acceptable et utilisable que la notion d'utilité peut être approchée. Tricot et al. (2003) critiquent ce modèle en expliquant que l'utilisabilité n'est pas la seule et unique source de problème conduisant à la non-acceptation d'un produit. Le choix et le degré d'adéquation des fonctions peuvent aussi être à l'origine des difficultés rencontrées par l'utilisateur. Dit autrement, un produit peut très bien être « facile d'utilisation », si celui-ci ne possède pas la fonctionnalité recherchée, c'est l'utilité qui est remise en cause (e.g., Blandford, Green, Furniss & Makri, 2007 ; Burkhardt & Sperandio, 2004 ; Nielsen, 1993). Cependant, une excellente utilisabilité peut également permettre aux utilisateurs de découvrir une autre utilité au produit initialement non imaginée (Tricot et al., 2003). Cela amène à une utilisation détournée d'un produit.

Pour le domaine de la e-santé, une utilisation détournée serait dangereuse pour la santé du patient. Un tensiomètre n'a pas d'autre objectif que de prendre la tension d'un patient. Il peut permettre de prévenir ou de surveiller différentes pathologies, mais l'utilité d'un tensiomètre ne pourra pas être remise en cause. Cette différence entre domaines renvoie à la notion de « flexibilité » de Shackel (1991). Par ailleurs, un patient est libre d'utiliser ou non un dispositif et si ce patient ne voit pas d'intérêt à suivre sa santé, il se pourrait que le produit ne soit pas utilisé (Barello et al., 2016; Graffigna et al., 2013). On ne remet pas en cause l'utilité du produit, on remet alors en cause l'engagement du patient dans sa propre santé, ce qui renvoie à l'acceptabilité sociale et aux modèles psychosociaux tels que le TAM. L'engagement deviendrait alors un prérequis à l'acceptation d'un dispositif au même titre que la notion d'utilisabilité. Ainsi, **les modèles présentés précédemment, bien que pertinents, sont incomplets pour s'appliquer au domaine de la e-santé.** En outre, les principales contraintes du patient peuvent intervenir par le biais de problèmes techniques ou de difficultés d'utilisation.

La sécurité passe par une utilisation fiable et simple. Cette qualité recherchée est vectrice de l'utilisabilité (Nielsen, 1993 ; Bastien et Scapin, 2001 ; Ezzedine, 2002 ; Seffah et al., 2006 ; Rubin et Chisnell, 2008).

2.4.2. L'inventaire des métriques utilisées sur un recensement de modèle entre 1971 à 2018

Shackel (1991), Nielsen (1993) et ISO 9241-11 (2018) ne sont pas les seuls modèles de l'utilisabilité. Nous proposons en annexe (cf. Annexe 1 : Revue des modèles d'utilisabilité de 1991 à 2018, p. 338) une analyse par auteurs en fonction des métriques entre 1971 et 2018. Nous avons pu recenser 27 modèles différents d'utilisabilité. Des critères communs émergent de ces modèles. Quatorze de ces auteurs incluent les métriques de « la satisfaction » ; 12 ont mentionné « l'efficacité », « l'efficience » et « la facilité d'apprentissage » dans leur modèle. Certaines métriques pourraient être regroupées : par exemple, « le taux d'erreurs » pourrait être inclus dans « l'efficacité », ou « la charge de travail » dans « l'efficience ». Shaker (1991) intègre l'efficacité dans la mesure de l'efficience et la facilité d'apprentissage peut être liée à l'apprenabilité ou à la mémorisation. Malgré l'usage des trois termes les plus cités (efficacité, efficience et satisfaction) renvoyant à la norme ISO 9241-11, les imprécisions évoquées suggèrent une difficulté des auteurs du domaine à se mettre d'accord sur un modèle commun et **des termes universels caractérisant l'utilisabilité.**

2.4.3. Une utilisabilité spécifique à chaque domaine d'application

Se rendant compte de la diversité des besoins d'utilisabilité en fonction des domaines, les chercheurs spécialisent leurs modèles en fonction des domaines d'application. Le modèle 2QCV3Q (Mich et al., 2003) évalue la qualité d'un site web sur la base du « *who-what-why-when-where-how* » et de la faisabilité. Ce modèle combiné à d'autres modèles préexistants (le modèle de la norme ISO 9126, le modèle *triangle of quality* de McCall, 1977 et le modèle Web Quality Evaluation Model ou QEM de Olsina et al. 2000) est repris par Shawgi et Noureldien (2015) dans le Usability Measurement Model ou UMM. Zhang et Adipat (2005) proposent un modèle d'utilisabilité spécialement pensé à destination des systèmes mobiles ou ubiquitaires, Rahmat, Zulzalil, Ghani, et Kamaruddin (2018) s'intéressent aux applications sur Smartphone, tandis que Koohang et Paliszkiwicz (2015) développent leurs travaux sur l'utilisabilité dans le contexte de l'éducation numérisée. Dans le domaine de la santé, Yen (2010) propose le health-ITUEM ou health IT Usability Evaluation Model basé sur les métriques de la norme ISO 9241-11 (2018) où la satisfaction est divisée en deux notions : l'utilité perçue et la facilité d'utilisation perçue. Toutes ces propositions de modèles différenciés en fonction des applications se basent

sur d'autres modèles préexistants. **Cependant, aucun de ces modèles différenciés et adaptés ne s'appuie sur une même base théorique.**

2.5. L'utilisabilité, des divergences disciplinaires, critérielles et applicatives

Cela fait de nombreuses années que les définitions de l'utilisabilité sont contestées, critiquées et qu'aucune définition ne fait l'unanimité (e.g., Abran et al., 2003; Lewis, 2014; Seffah et al., 2006). En 1990, Shackel (1990, p.31) écrivait déjà *“One of the most important issues is that there is, as yet, no generally agreed definition of usability and its measurement”*. Suivi par les propos tout en humour de Gray et Salzman (1998, p.242) : *“Attempts to derive a clear and crisp definition of usability can be aptly compared to attempts to nail a blob of Jeff-O to the wall”*. Le constat de Lewis (2014) est le même : *“That for a period of more than 30 years there is yet to be a clear and generally accepted definition”*. Jusqu'à devenir inquiétant, lorsque Tractinsky (2018, p.131) aborde la possible impasse dans lequel se trouve cette notion (*« dead end ? »*). Cette diversité de modèles n'est pas sans conséquence pour cette propriété qui *in fine* se fragilise voire se divise.

Plusieurs raisons expliqueraient cette absence de convergence :

- La diversité des termes dans des classifications (Bevan, 1995) ;
- Les différentes façons de mesurer l'utilisabilité (Folmer & Bosch, 2004) ;
- L'utilisabilité trouve son origine dans des problèmes associés à l'interaction avec l'humain, or elle est le plus souvent investiguée par des ingénieurs logiciels ayant une vision orientée sur le produit et visant à répondre à des objectifs de conception d'interfaces en termes d'efficacité mesurée par la vitesse d'exécution et de fiabilité (Nielsen, 1993) ;
- L'utilisabilité n'est pas une propriété spécifique d'une personne ou d'un produit (Dumas, 2003; Hertzum, 2010) ;
- L'utilisabilité ne peut pas être mesurée sur une simple échelle de l'utilisabilité (Dumas, 2003; Hertzum, 2010) ;
- Il semblerait pertinent de faire la différence entre l'utilisabilité « formative » et l'utilisabilité « sommative », dans la mesure où ces deux concepts n'ont pas les mêmes objectifs (cf. Partie I, Chapitre 1, 4.3.1. Trois types d'évaluation de l'utilisabilité, p. 44) (Lewis, 2014).

Les conséquences d'une absence de consensus sur la définition entraineraient alors la mort du concept d'après Tractinsky (2018). Dans un premier temps, l'auteur explique que l'utilisabilité n'est pas correctement définie par la norme ISO 9241-11. Cette incertitude planante autour de cette définition est un obstacle à l'accumulation et à l'organisation des connaissances disciplinaires. Dans un deuxième temps, il regrette le manque de standardisation des méthodes choisies ce qui aurait pour conséquences de limiter la comparaison et la reproductibilité des études d'utilisabilité. Ce manque de standardisation a des répercussions dans plusieurs domaines, dont celui de la santé. Nous développons ces conséquences dans le chapitre suivant (cf. Partie I, Chapitre 2, 2.5.2. Les trois principaux défis des évaluations de l'utilisabilité des technologies en santé, p. 63).

2.6. Synthèse des modélisations : vers une définition de l'utilisabilité et proposition d'un modèle opérationnel

L'ensemble des auteurs précités est d'accord sur un point : l'utilisabilité est multidimensionnelle. En ce sens, il ne s'agit pas d'une propriété unique et elle ne peut pas être évaluée sur une seule de ces dimensions mais sur un ensemble de critères (Nielsen, 1993 ; Abran et al., 2003; Dumas, 2003; Hertzum, 2010). L'utilisabilité se mesure par l'intermédiaire de ces métriques de manière indépendante (Nielsen, 1993). Rubin (1994, p. 18) met l'accent sur la base commune de tous les modèles précités composés pour la plupart des aspects tels que l'utilité, l'efficacité, l'apprentissage et l'attitude. Barcenilla et Bastien (2009) présentent également trois aspects universels de l'utilisabilité : (1) l'utilisateur est placé au centre du processus de conception ; (2) ce processus inclut les besoins spécifiques des populations ciblées ; (3) et l'utilisabilité doit prendre en compte les progrès de la recherche par le biais des processus de standardisation et de normalisation même si ce dernier point est particulièrement discuté dans la littérature scientifique (e.g., Peute et al., 2013; Tractinsky, 2018).

Même si la définition de l'utilisabilité reste un débat important, **nous avons fait le choix de nous inspirer des trois principaux modèles présentés (Shackel, 1991 ; Nielsen, 1993 et ISO 9241-11, 2018) pour la suite de ce travail. Ainsi, nous définissons l'utilisabilité comme la propriété interactive multidimensionnelle instrumentale et non instrumentale d'un système avec un utilisateur donné dans un contexte d'usage spécifique.**

Cette définition sera opérationnalisée par les **trois métriques principales de l'utilisabilité (*efficacité, efficience et satisfaction*)** auxquelles nous ajoutons **deux métriques**

connexes (l'apprenabilité et la mémorisation), tout en prenant en compte **le contexte d'usage** caractérisé par quatre composants (*les utilisateurs, les tâches, les ressources et les environnements*), que nous détaillons dans la partie suivante. Nous représentons ces éléments opérationnalisés dans la Figure 6.

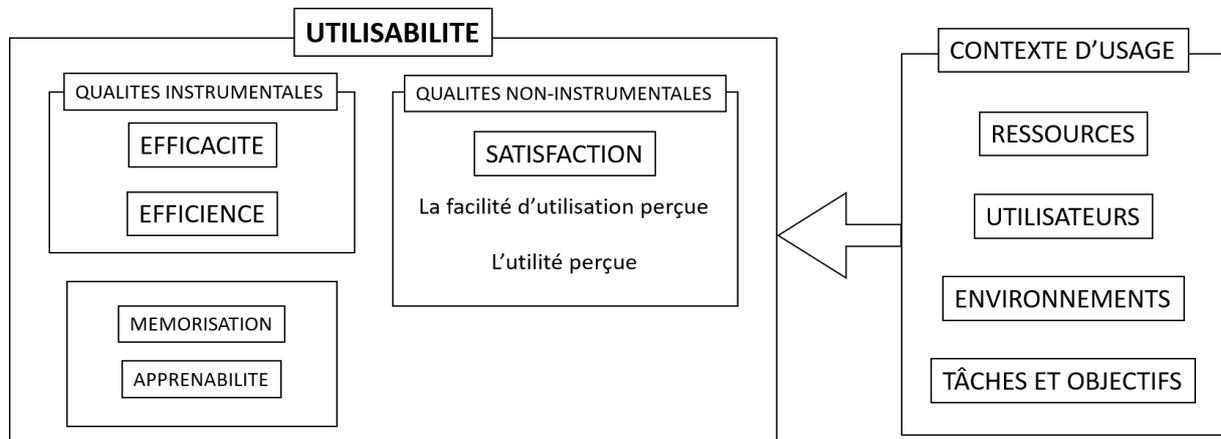


Figure 6. Représentation d'un modèle opérationnel de l'utilisabilité et de ses métriques en fonction des composants du contexte d'usage.

3. Les approches conceptuelles de l'utilisabilité : les mesures des métriques de l'utilisabilité

3.1. Les mesures des trois principales métriques de l'utilisabilité

3.1.1. L'efficacité : définitions et mesures

Seffah, Kececi, et Donyae, (2001) définissent l'efficacité comme la mesure dans laquelle le produit facilite l'utilisateur dans l'accomplissement de la tâche à laquelle il est destiné avec précision et exhaustivité tout en évitant la plupart des erreurs dans divers contextes d'utilisation. Cette notion traite de la performance, de la qualité du résultat obtenu, de l'effet attendu et non du processus qui a mené au résultat. Baber et Stanton (1996) la définissent comme la « performance acceptable » tout en mettant l'accent sur la contextualisation de cette performance. La performance acceptable peut dépendre du niveau d'expérience de l'individu et du type d'usage de l'utilisateur. Brangier et Barcenilla (2003) expliquent, dans leur ouvrage *Concevoir un produit facile à utiliser*, qu'un individu utilisant internet pour lire ses mails n'associera pas la même « acceptabilité de la performance » que celui qui utilise sa connexion pour des activités bien plus gourmandes (comme regarder Netflix si nous actualisons la citation).

D'après Dubey et Gulati (2012), l'efficacité est multidimensionnelle et composée de cinq sous-attributs :

- **La réalisation de la tâche** : ce point se rapproche le plus de la définition de l'ISO 9241-11 (2018), car il s'agit de mesurer la capacité qu'a le système à permettre aux utilisateurs d'accomplir leurs tâches et d'atteindre l'objectif ou les objectifs escompté(s) correctement. La réalisation de la tâche dépend de la quantité de tâches accomplies et de leur qualité d'accomplissement ;
- **L'opérabilité** : Ce sous-attribut correspond à la mesure dans laquelle le système fournit aux utilisateurs les fonctionnalités nécessaires dans un but d'accomplir la tâche correctement (Alonso-Ríos et al., 2009). Il comprend, d'après Dubey et Gulati (2012), la précision et la complétude ;
- **L'universalité** : Cela correspond au fait que le système peut être utilisé par tous les utilisateurs indépendamment de leurs caractéristiques physiques, mentales ou culturelles (Alonso-Ríos et al., 2009; Seffah et al., 2006). Autrement dit, l'universalité repose sur deux éléments que sont l'accessibilité et l'universalité culturelle (Dubey & Gulati, 2012) ;
- **La flexibilité** : Il s'agit de la mesure avec laquelle le système peut s'adapter à l'évolution des besoins des utilisateurs ainsi qu'à leurs préférences (Brian Shackel, 2009). Cela renvoie à l'adaptabilité et à la contrôlabilité du système (Dubey & Gulati, 2012).
- **Les erreurs** : Ce sous-attribut fait référence au nombre d'erreurs de manipulation avec le système par rapport à la tâche réalisée (Nielsen, 1994a).

Avant de vouloir évaluer l'efficacité d'un système, **il est nécessaire d'avoir fixé des objectifs afin de vérifier la bonne cohérence entre les résultats attendus et la mesure obtenue**. Par exemple, dans le cas de la manipulation d'un défibrillateur, les objectifs d'efficacité seront une utilisation sans erreur de manipulation du matériel avec comme critère, la pose des électrodes aux bons endroits afin de garantir la survie du patient. Nous présentons dans le Tableau 1 ci-dessous différentes mesures de l'efficacité.

Tableau 1. Mesures de l'efficacité.

	(Nielsen, 1994a)	(Dix et al., 2003)	(Abran et al., 2003)	(Dubey et al., 2012)	(Brangier & Barcenilla, 2003)
Mesures	Pourcentage des tâches accomplies	Pourcentage de tâches accomplies	Pourcentage des tâches accomplies	Taux d'achèvement	Réussite de la tâche
		Pourcentage des erreurs de manipulation	Pourcentage des erreurs de manipulation	Pourcentage d'utilisateurs qui accomplissent les tâches avec succès	Qualité de la performance
		Pourcentage de tâches accomplies par unité de temps	Pourcentage de tâches accomplies par unité de temps	Nombre de tâches accomplies avec succès	
				Nombre d'erreurs	
				Pourcentage des fonctions pertinentes utilisées	
				Pourcentage de la tâche achevée	
				Pourcentage de nombre d'erreurs	

3.1.2. L'efficacité : définitions et mesures

Alonso-Ríos et al. (2009, p.8) proposent une définition similaire à la norme ISO 9241-11 (2018) en définissant l'efficacité comme *“the capacity of the system to produce appropriate results in return for the resources that are invested”*. Si nous vulgarisons cette définition, cela revient à dire que plus l'effort est faible, meilleure est l'efficacité d'un système. Cependant, d'après Nielsen (1994a) et Preece et al. (1994), ces définitions ne suffisent pas, car elles ne mettent pas en exergue l'expérience des utilisateurs. Nielsen (1994a) décrit l'efficacité comme le niveau de performance atteint par les utilisateurs experts (utilisateurs qui ont appris à utiliser le système) et mesurée *“ le temps qu'il faut aux utilisateurs pour effectuer certaines tâches de tests typiques ”*. Preece et al. (1994), quant à eux, prennent en compte les utilisateurs expérimentés tout comme le débit de l'attribut, les tâches accomplies et les erreurs commises. Alonso-Ríos et al. (2009) regrettent le fait que les études d'utilisabilité ne prennent pas en compte les compétences et les connaissances des utilisateurs cibles, et suggèrent une définition de l'efficacité complète. En effet, leur définition prend en compte le temps d'exécution des tâches, l'effort mental et physique, mais aussi les caractéristiques des utilisateurs, leurs environnements et leurs ressources. Ces trois dernières notions sont présentes dans la partie

« contexte d'utilisation », soit en dehors des métriques de l'utilisabilité. Nous y reviendrons dans ce chapitre (cf. Partie I, Chapitre 1, 3.4. Les composants du contexte d'utilisation, p. 37).

Tout comme l'efficacité, il est nécessaire avant de mesurer l'efficacité d'avoir des attentes et de fixer des objectifs prédéterminés. Nous proposons dans le Tableau 2, les mesures les plus utilisées afin d'obtenir des données sur l'efficacité d'un système. Par exemple, l'efficacité d'un jeu vidéo (c'est-à-dire, le temps ou l'effort humain consacré à la résolution du jeu) est souvent liée à un certain plaisir. Il est donc nécessaire de rechercher un effort de la part de l'utilisateur. En reprenant notre exemple du défibrillateur, les objectifs d'efficacité seront une utilisation rapide *via* un effort cognitif moindre afin de ne pas perdre de temps pour augmenter les chances de survie du patient. La norme ISO 9241-11 (2018) identifie trois types d'efficacité repris par Alonso-Ríos et al., (2009) qui y ajoutent un élément supplémentaire et par Dubey et Gulati, (2012) :

- **L'efficacité dans l'effort humain ou effort de l'utilisateur** : Il s'agit de la capacité du système à produire des résultats appropriés en fonction de l'effort physique ou mental investi par l'utilisateur.
- **L'efficacité du temps d'exécution des tâches** (Alonso-Ríos et al., 2009) ou **la performance**, c'est-à-dire le temps investi par l'utilisateur dans l'exécution des actions et le temps pris par le système pour y répondre. Bass et John (2003) abordent la notion de mémoire en plus du temps investi. Dubey et Gulati (2012) intègrent trois sous-composants à cette notion : le temps d'exécution, la charge de mémoire et la complexité de la décision.
- **L'efficacité dans l'utilisation** des ressources matérielles et humaines, c'est-à-dire, la mesure des ressources utilisées pour mener à bien la tâche. Cela prend en compte le « débit » soit le retour sur investissement et « l'utilisation des commandes », soit le nombre de commandes rarement utilisées (Preece et al., 1994).
- **L'efficacité des coûts économiques ou finances** fait référence à différents types de dépenses comme le coût du système lui-même, le coût des ressources humaines, le coût de l'équipement nécessaire pour travailler avec le système et le coût des produits consommables.

Tableau 2. Mesures de l'efficacité.

	(Nielsen, 1994a)	(Dix et al., 2003)	(Abran et al., 2003)	(Chang & Dillon, 2006)	(Brangier & Barcenilla, 2003)
	Temps d'activation d'une tâche	Nombre de bonnes et de mauvaises caractéristiques rappelées par les utilisateurs	Nombre de répétitions de commandes échouées	Nombre d'objectifs/de tâches non atteintes	Taux et la nature des erreurs d'utilisateur
	Temps consacré aux erreurs	Temps d'activation d'une tâche	Fréquence d'utilisation de la documentation ou de l'aide	Période improductive	Temps pour exécuter une tâche donnée
	Pourcentage d'erreur	Temps consacré aux erreurs	Pourcentage d'erreurs	Temps nécessaire à l'exécution de la tâche	Nombre d'opérations requises pour exécuter la tâche principale et les déviations par rapport à la procédure optimale
Mesures	Fréquence d'utilisation de la documentation ou de l'aide	Pourcentage d'erreurs	Temps passé sur l'erreur	Pourcentage de la tâche non accomplie	Charge de travail
	Nombre de répétitions de commandes échouées	Fréquence d'utilisation de la documentation ou de l'aide	Temps d'accomplissement d'une tâche		
		Nombre de répétitions de commandes échouées	Nombre de bonnes et mauvaises caractéristiques rappelées par les utilisateurs		
		Nombre de commandes disponibles non sollicitées	Nombre de commandes non disponibles		

3.1.3. La satisfaction : définitions et mesures

La satisfaction est sûrement l'une des dimensions autour de laquelle se situe la plupart des désaccords. Nielsen (2014a) et Seffah et al., (2006) décrivent la satisfaction en utilisant des termes subjectifs comme l'absence d'inconfort de l'utilisation du système et la capacité du système à atteindre l'objectif de l'utilisateur. Yen (2010) proposent un modèle validé (Health – ITUEM) dans lequel la satisfaction est décomposée en deux éléments : *la facilité d'utilisation perçue* et *l'utilité perçue*. Dubey et Gulati (2012, p.333) définissent la satisfaction comme le

« degré auquel le système est agréable, confortable, attrayant et digne de confiance pour les utilisateurs »⁸. Les auteurs définissent quatre sous-attributs de la satisfaction :

- **La sympathie** (« *Likeability* ») : cela renvoie au degré selon lequel le système est apprécié par l'utilisateur. Cela implique la perception et l'opinion de l'utilisateur sur le produit (Rubin, 1994) ;
- **La confiance** : ce sous-attribut fait référence à la fidélité que le système offre à l'utilisateur (Seffah et al., 2006) et comprend la stabilité, la réputation et l'intention ;
- **Le confort** : il s'agit de la mesure selon laquelle le système crée une sensation pour l'utilisateur. Le confort comprend l'utilisation du produit (soit la réponse de l'utilisateur après l'avoir utilisé) et sa conception (soit son degré de motivation) (Dubey & Gulati, 2012) ;
- **L'attractivité** correspond à l'esthétique du système.

Brangier et Barcenilla (2003) insistent sur la difficulté à mesurer la satisfaction. Les auteurs évoquent l'utilisation d'échelles d'évaluation « subjectives » afin de mesurer cette dimension insaisissable (par exemple, le Software Usability Measurement Inventory ou SUMI). Nombreux sont les questionnaires de satisfaction qui existent. Pourtant, nombreuses sont les critiques qui leur sont associées, puisque ces questionnaires interrogent les relations complexes de l'efficacité, l'efficience et sont biaisés par la mémoire de l'utilisateur, de ses besoins, de ses caractéristiques sociales et économiques, de sa perception de l'utilité et de ses motivations. Nous présentons dans le Tableau 3 ci-dessous les mesures de la satisfaction recensées par certains auteurs.

⁸ traduit de l'anglais: "as the degree to which the software is likeable, comfortable, attractive and trustworthy for the users."

Tableau 3. Mesures de la satisfaction.

	(Nielsen, 1994a)	(Dix et al., 2003)	(Abran et al., 2003)	(Kirakowski et Corbett, 1993)	Yen (2010)
Mesures	Nombre de fois où l'utilisateur exprime sa frustration	Pourcentage de commentaires favorables et défavorables des utilisateurs	Pourcentage de commentaires favorables et défavorables des utilisateurs	Inconfort pratiqué	Utilité perçue
	Échelle d'évaluation de la satisfaction des utilisateurs à l'égard des fonctions et des caractéristiques	Nombre de fois où l'utilisateur exprime sa frustration	Nombre de fois où l'utilisateur exprime sa frustration	Affection pour le produit	Facilité d'utilisation perçue
		Échelle d'évaluation de la satisfaction des utilisateurs à l'égard des fonctions et des caractéristiques	Échelle d'évaluation de la satisfaction des utilisateurs à l'égard des fonctions et des caractéristiques	Plaisir dans l'utilisation du produit	Acceptabilité de la charge de travail lors de l'exécution de diverses tâches Mesure dans laquelle les objectifs particuliers en matière d'utilisabilité ont été atteints

3.1.4. Les liens entre l'eXpérience Utilisateur et l'utilisabilité

L'eXpérience Utilisateur ou UX⁹ est un concept relativement récent qui dépasserait l'utilisabilité traditionnelle (Petrie & Bevan, 2009), car il viendrait élargir les concepts d'utilisabilité en incluant des aspects émotionnels et temporels. Lallemand, Gronier, et Koenig (2015) mettent en opposition l'utilisabilité associée à une approche objective de l'interaction à l'expérience utilisateur liée à une approche subjective. Aujourd'hui, les définitions de l'expérience utilisateur se confrontent aux mêmes problématiques que les définitions de l'utilisabilité : il n'y a pas encore de consensus sur une définition précise (Lallemand et al., 2015). Vingt-sept définitions différentes sont référencées sur le site de « *all about UX*¹⁰ ». Nous retiendrons la définition de la norme ISO 9241-210 (2019, p.3) comme « *les perceptions et*

⁹ en anglais *User eXperience*

¹⁰ <http://www.allaboutux.org/ux-definitions>.

réponses de la personne à l'utilisation et/ou à l'utilisation anticipée d'un produit »¹¹. Il existe cependant un consensus. **L'expérience utilisateur est composée d'une dynamique temporelle et serait le résultat de l'interaction entre l'utilisateur final, le système et son contexte d'usage ce qui n'est pas très éloigné des définitions de l'utilisabilité.** Lindgaard et Dudek, (2003) ainsi que McLellan, Muddimer, et Peres (2012) montrent par ailleurs que la satisfaction des utilisateurs est liée à l'expérience de l'utilisateur avec le produit, ce qui est repris et validé par Sauro (2011). Ces auteurs mesurent la satisfaction avec le System Usability Scale ou SUS des utilisateurs expérimentés et non expérimentés avec un produit. De la même manière que la satisfaction, l'expérience utilisateur est extrêmement complexe à mesurer, car il s'agit de capter toute la complexité de l'expérience humaine. Plusieurs échelles de mesure de l'expérience utilisateur existent (e.g. AttrakDiff ; Modular Evaluation of the User Experience Components ou meCUE ; Self-Assessment Manikin ou SAM) ; elles sont basées sur différents modèles d'expérience utilisateur (e.g., Hassenzahl, 2005; Karapanos, Zimmerman, Forlizzi, & Martens, 2009; Mahlke, 2008). Tous ces auteurs incluent l'utilisabilité ou une dimension de l'utilisabilité dans leur modèle.

Finalement, si pour certains auteurs, l'expérience utilisateur inclut l'utilisabilité, pour d'autres auteurs, l'utilisabilité englobe l'expérience utilisateur au travers de la satisfaction. **Pour notre part, nous considérons que l'expérience utilisateur fait partie de l'utilisabilité.**

3.1.5. Les relations entre efficacité, efficience et satisfaction

Hornbæk et Law (2007) ont réalisé une méta-analyse mettant en commun les données de 54 études traitant des relations entre l'efficacité et l'efficience. Ils ont montré une corrélation de 0,247 (IC 95 %, +/- 0,059). Cela suggère qu'une performance efficace est également efficiente. Autrement dit, si un utilisateur exécute une tâche rapidement, ses résultats seront plus efficaces. Cela semble logique, surtout dans le cas où des erreurs de manipulation peuvent freiner la résolution d'une tâche. Les auteurs mettent en avant deux types d'erreurs : l'erreur d'achèvement de la tâche (ce que nous appelons l'erreur « **non-bloquante** ») et l'erreur en cours de route (ce que nous qualifions d'erreur « **bloquante** »). Les erreurs en cours de route (erreurs bloquantes) ont une corrélation moyenne à l'efficacité de 0,441 +/- 0,125 (14 études) ; les erreurs d'achèvement de la tâche (erreurs non-bloquantes) ont une corrélation moyenne de 0,155 +/- 0,08 (23 études). Hornbæk et Law (2007) expliquent que 87% des études démontrent une corrélation positive entre efficacité et efficience sachant que la complexité de la tâche ne

¹¹ traduit de l'anglais : "person's perceptions and responses resulting from the use and/or anticipated use of a product" (ISO, 2010, p. 3).

semble pas affecter la relation entre efficacité et efficacité. Or cet effet reste faible ($< 0,4$). **Il serait judicieux de conclure que ces deux dimensions pourraient être liées, en particulier dans le cas d'erreurs bloquantes.** Enfin, la norme ISO 9241-11 (2018) signale qu'un système peut être efficace (en consommant très peu de ressources) et en même temps être faiblement efficace en ce qui concerne l'obtention des résultats escomptés. L'inverse est aussi possible. Cependant, la norme ne présente aucune explication à ces constats.

Pour les liens entre l'efficacité et l'efficacité avec la satisfaction, Hornbæk et Law (2007) démontrent que 86 % des études (sur 39 études) ont une corrélation positive entre l'efficacité et la satisfaction, mais cette corrélation reste très faible ($0,164 \pm 0,062$) et que 81 % des études (sur 45 études) ont une corrélation positive mais faible ($0,196 \pm 0,064$) entre l'efficacité et la satisfaction. Ces mesures de corrélation sont mises en évidence dans la Figure 7 ci-dessous. Elles ont été prises en fonction du niveau de complexité du système mesuré grâce aux critères de Rasmussen (1983). Les travaux de Hornbæk et Law (2007) nous montrent que les relations entre les dimensions de l'utilisabilité sont certes corrélées positivement, mais très faiblement.

Enfin, Frøkjær Hertzum et Hornbæk (2000) avaient conclu, dans leurs travaux, à une indépendance des trois dimensions de l'utilisabilité. **Les auteurs soulignent par ailleurs, l'importance d'évaluer de manière indépendante l'efficacité, l'efficacité et la satisfaction.**

La somme des mesures de l'efficacité, l'efficacité et la satisfaction donne le profil d'utilisabilité du produit. On parle alors de **résultat de l'utilisation** (ISO 9241-11, 2018). La relation entre *l'efficacité*, *l'efficacité* et la *satisfaction* peut varier non seulement en fonction du niveau de complexité de la tâche, mais aussi relativement au contexte d'usage intégrant une multitude de facteurs.

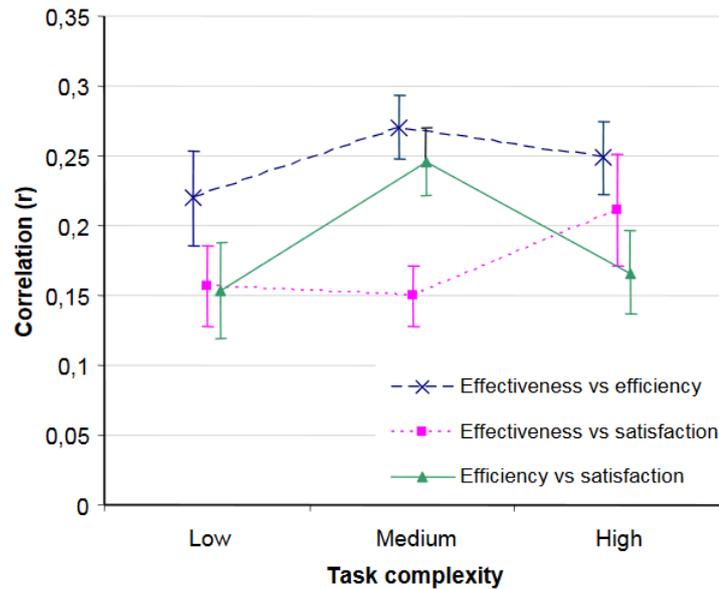


Figure 7. Représentation graphique représentant les corrélations entre les mesures d'efficacité, d'efficacité et de satisfaction en fonction de la complexité de la tâche d'après Hornbæk et Law (2007). L'évaluation de la complexité de la tâche est basée sur les travaux de Rasmussen (1983).

3.2. Les métriques connexes de l'utilisabilité

3.2.1. L'apprenabilité (ou facilité d'apprentissage) : définitions et mesures

Si l'on se base sur les définitions de Shackel (1991) et de Nielsen (1993), l'apprenabilité (en anglais *learnability*) est définie par la facilité, la rapidité et la performance avec lesquelles un novice peut atteindre un niveau de compétence suffisant pour utiliser la technologie. La facilité et la durée d'utilisation sont centrales dans l'apprenabilité. D'après Nielsen (1993), l'apprenabilité se mesure par le nombre de fonctions retenues. Ainsi, afin de mesurer l'apprenabilité d'un produit, il est nécessaire d'évaluer les performances d'un utilisateur sur plusieurs essais.

Cette notion se rapproche de l'apprentissage procédural dans lequel l'objectif est l'acquisition progressive de nouvelles compétences et connaissances grâce à une activité répétée (e.g., Anderson, 1982). Shackel (1991) et Nielsen (1993) positionnent l'apprenabilité au même niveau que l'efficacité et l'efficacité dans l'utilisabilité. Pourtant, dans certains domaines, comme celui de la santé, Larsen (2017) explique que l'utilisation de dispositifs médicaux doit être optimisée dès la première utilisation. C'est pourquoi nous positionnons cette notion comme connexe par rapport à l'efficacité, l'efficacité et la satisfaction.

3.2.2. La mémorisation : définitions et mesures

La mémorisation (en anglais *memorability*) est définie par les compétences acquises permettant une utilisation de la technologie même après une certaine période d'inutilisation. Nielsen (1993) parle de trace mnésique laissée chez l'utilisateur. La mémorisation se mesure par rappel avec une période de non-utilisation ou à une fréquence donnée.

Tout comme l'apprenabilité, nous considérons la mémorisation comme une métrique connexe de l'utilisabilité.

3.3. Synthèse des métriques de l'utilisabilité

Compte tenu des éléments précédents, nous considérons l'utilisabilité sur la base de cinq métriques (Figure 8) : deux métriques principales renvoyant aux propriétés instrumentales (l'efficacité et l'efficience), deux métriques connexes (la mémorisation et l'apprenabilité) et une métrique subjective renvoyant aux qualités non instrumentales du produit (la satisfaction) pouvant être subdivisée en deux sous-catégories : la facilité d'utilisation perçue et l'utilité perçue.

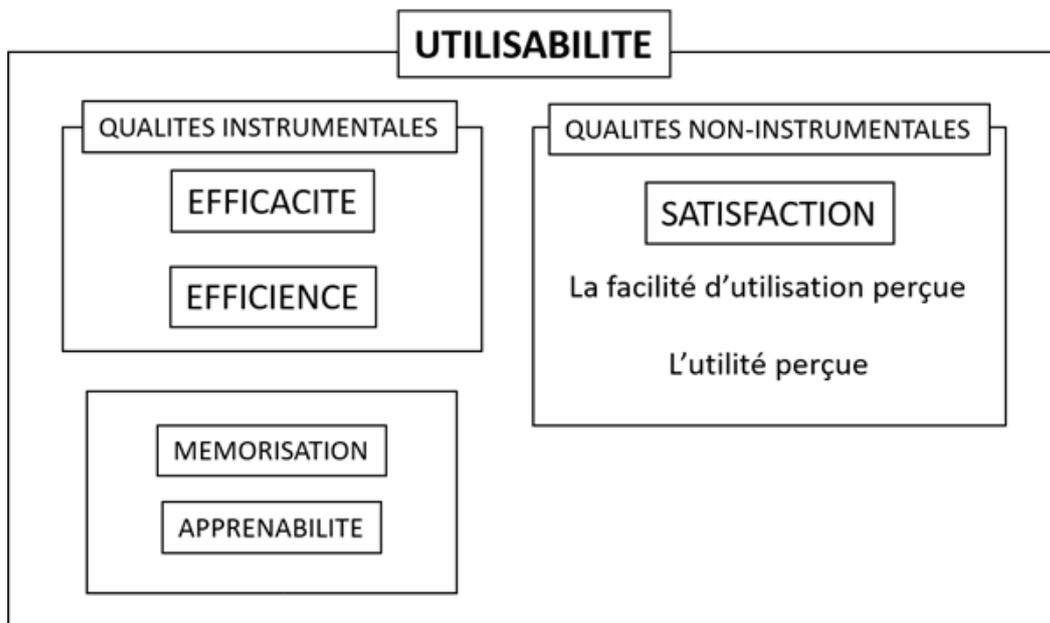


Figure 8. Synthèse des métriques d'utilisabilité mobilisées dans la thèse.

3.4. Les composants du contexte d'utilisation

L'utilisabilité n'est pas uniquement déterminée par les qualités du produit (ou système, ou service), mais aussi et surtout par le contexte dans lequel il est utilisé. Shackel et Richardson (1991) considèrent que l'utilisabilité est évaluée par l'interaction de l'utilisateur, du système et de la tâche dans un cadre spécifique. Bevan (1995b) et Bevan et Macleod (1994) mettent en évidence trois composants du contexte pouvant interférer avec le résultat de l'utilisation : les utilisateurs, les tâches et les environnements particuliers. Ainsi, pour les auteurs, il n'existe pas de produit « utilisable » ou « inutilisable ». **Tout cela dépend du contexte, puisque un produit inutilisable par des utilisateurs inexpérimentés peut tout à fait être utilisable pour des utilisateurs avertis.** Yen et Bakken (2012) complètent cette analyse en considérant cette notion sur la base de quatre composants : l'utilisateur, la tâche, le système et l'environnement. Les auteurs considèrent que la complexité des tâches, les caractéristiques utilisateurs et le soutien organisationnel sont des facteurs qui influencent l'utilisation d'un système, sans être des problèmes intrinsèques au système. De ce fait, ils doivent être différenciés des problèmes liés au système. Enfin, d'après la norme ISO 9241-11 (2018, p.14), le contexte d'utilisation est composé « *d'une combinaison d'utilisateur, d'objectifs, de tâches, de ressources et des environnements technique, physique, social, culturel et organisationnel dans lesquels est utilisé un système, produit ou service* ». La norme incorpore la notion de ressources à ces composants précités. **Ainsi, nous nous sommes servis des composants de Bevan (1995b), Yen et Bakken (2012) et ISO 9241-11 (2018) pour proposer notre propre cadre de contexte d'usage composé de l'utilisateur, de la tâche, des ressources et de l'environnement.**

Changer tout aspect pertinent du contexte d'utilisation peut modifier l'utilisabilité. Ces composants influencent l'utilisabilité (*efficience, efficacité, satisfaction, apprentissage* et mémorisation). Il est donc nécessaire d'investiguer ce contexte afin de prendre en compte l'ensemble des éléments qui orbitent autour de l'utilisabilité. Il est alors possible d'améliorer l'utilisabilité en modifiant certains aspects du contexte, comme par exemple la formation des utilisateurs, l'amélioration de l'environnement de travail en ajustant la luminosité et le bruit (Bevan & Macleod, 1994). Tous ces éléments peuvent affecter l'utilisabilité.

Il est donc nécessaire de prendre en compte chaque spécificité du contexte d'utilisation prévu. Pour cela, Bevan et Macleod (1994) expliquent qu'il faut considérer « l'adéquation de la tâche », car ces questions sont étroitement liées aux caractéristiques de la tâche, mais également « l'aptitude à l'apprentissage », « l'aptitude à l'individualisation » et la « conformité

Tableau 4. Détails des éléments du contexte à prendre en compte.

Utilisateurs	Tâche	Équipement	Environnement
Données personnelles	Répartition des tâches	Description de base	Organisationnel
Types d'utilisateurs	Nom de la tâche	Identification du produit	Environnement
Audience et utilisateurs secondaires	Objectif de la tâche	Description du produit	
	Fréquence des tâches	Principaux domaines d'application	Structure
	Durée de la tâche	Principales fonctions	Heures de travail
Compétences & connaissances	Fréquence des événements		Travail en groupe
Expérience produite	Flexibilité des tâches	Spécifications	Fonction d'emploi
Connaissance du système	Exigences physiques et mentales	Matériel informatique	Pratiques de travail
Expérience organisationnelle	Dépendances des tâches	Logiciels	Assistance
Entraînement	Sortie de tâche	Matériaux	Interruptions
Clavier et compétences de saisie	Risque résultant d'une erreur	Autres éléments	Structure de gestion
Qualifications requises			Structure de communication
Compétences linguistiques			Rémunération
Connaissances générales			
Caractéristiques personnelles			Attitudes et culture
Âge			Politique sur l'utilisation des ordinateurs
Sexe			Objectifs organisationnels
Capacités physiques			Relations industrielles
Limites physiques et incapacités			
Capacités intellectuelles			Conception du travail
Attitude			Souplesse d'emploi
Motivation			Suivi des performances
			Rétroaction sur le rendement
			Stimulation
			Autonomie
			Pouvoir discrétionnaire
			Environnement technique
			Configuration
			Matériel informatique
			Logiciels
			Documents de référence
			Environnement physique
			Conditions de travail
			Conditions atmosphériques
			Environnement auditif
			Environnement thermique
			Environnement visuel
			Instabilité environnementale
			Aménagement du lieu de travail
			Espace et mobilier
			Posture de l'utilisateur
			Lieu
			Sécurité sur le lieu de travail
			Dangers pour la santé
			Vêtements et équipements de protection

aux attentes des utilisateurs » portant sur les questions de conception étroitement liées aux caractéristiques de l'utilisateur. Les auteurs signalent qu'une attention particulière doit être accordée aux besoins des différents types d'utilisateurs visés dans l'exécution des tâches prévues dans des situations particulières. Bevan et Macleod (1994) proposent un Tableau 4 récapitulatif des différents éléments du contexte à prendre en compte.

3.4.1. Les ressources

Les ressources peuvent être divisées en deux groupes : les ressources *réutilisables* et les ressources *consommables* (ou *non renouvelables*). Les ressources *réutilisables* comprennent l'équipement ou l'information d'assistance comme les notices. Ces dernières sont facilement ajustables à travers les manuels d'utilisation composés d'instructions. Les ressources *consommables* comprennent le temps disponible, les ressources financières et les ressources humaines, intégrant notamment les formations dispensées par le personnel infirmier aux utilisateurs patients. Ces ressources sont nécessaires afin de former les utilisateurs et de dissiper le moindre doute sur les instructions à suivre pour une utilisation correcte. La norme ISO/IEC TR 9126-2 (2003) inclut la métrique de la compréhensibilité dans son modèle d'utilisabilité. Pour mesurer cette métrique, il est nécessaire d'évaluer la fréquence d'utilisation ou le temps passé à utiliser les manuels d'utilisation. Cependant, cette norme n'aborde pas la manière d'évaluer les ressources consommables. Dans le cas du projet *Smart Angel*, les patients recevront une formation hospitalière cumulée à des instructions sur tablette. Nous reviendrons sur ces notions dans les sections du chapitre suivant (cf. Partie I, Chapitre 2, 4.2.1. Les ressources réutilisables : les documents procéduraux, p. 91 et 4.2.2. Les ressources consommables : Les formations et les programmes d'éducation thérapeutique, p. 94).

3.4.2. L'environnement

L'environnement est à la fois technique, physique, social, culturel et organisationnel. L'environnement technique permet d'utiliser ou de limiter les ressources (e.g., la consommation en électricité, la batterie des produits, la connexion 4G pour permettre l'envoi des constantes du patient). L'environnement physique, quant à lui, compte l'environnement bâti (le domicile du patient ou l'hôpital) ou encore les conditions météorologiques. Enfin, les environnements social, culturel et organisationnel englobent les individus, leurs rôles, leurs relations et leurs interactions avec d'autres individus (comme les professionnels de santé ou les aidants familiaux). Ce type d'environnement prend aussi en compte le cadre législatif et la structure organisationnelle (de l'hôpital, par exemple).

3.4.3. Caractéristiques utilisateurs

Le contexte d'utilisation prend en compte les utilisateurs et leurs caractéristiques tels que leur âge, le sexe, les capacités physiques et cognitives, l'attitude et la motivation. Le projet *Smart Angel* possède déjà ses propres critères de sélection des patients/participants (cf Annexe 2 : Critères d'inclusion et d'exclusion de la population pouvant utiliser le dispositif *Smart Angel*, p. 341). Nous proposons toutefois une partie consacrée à l'influence des caractéristiques utilisateurs sur l'utilisabilité dans le chapitre suivant (cf. Partie I, Chapitre 2, 4.1. Les caractéristiques utilisateurs, p. 85).

4. Les approches méthodologiques de l'utilisabilité : leur complémentarité dans le cycle de conception, les types, les méthodes et les outils d'évaluation

4.1. L'intégration des objectifs d'évaluation de l'utilisabilité dans le cycle de conception d'un produit

La conception de produits notamment logiciels comporte trois phases pouvant être réitérées (Scholtz, 2004) : l'analyse des besoins, la conception/les tests/le développement, et l'implantation. L'utilisabilité doit être visée tout au long du cycle de conception du système et peut être évaluée à différents stades de son développement (Dumas & Redish, 1999). En effet, **l'évaluation fait partie intégrante de la conception**. Deux dimensions sont à prendre en compte lors de l'évaluation d'un système : « quand » et « quoi » évaluer. C'est ce que Yen et Bakken (2012, p. 414) appellent « *when to evaluate what* ».

4.1.1. Les objectifs du « quoi » dans l'évaluation de l'utilisabilité : une utilisabilité située

Les objectifs d'évaluation de l'utilisabilité reposent sur l'analyse des besoins dans un contexte précis (*utilisateur, tâche, ressources, environnement*) où les métriques sont hiérarchisées (*efficacité, efficacité, satisfaction, apprenabilité, mémorisation*). C'est ce que nous appelons le « quoi ». Chaque contexte et métrique pré-définis constituent des objectifs d'utilisabilité différents. Par exemple, si nous voulons concevoir un défibrillateur, nous désirons que le produit soit rapide à utiliser et manipulable sans erreur dès la première utilisation pour maximiser les chances de survie du patient en arrêt cardiaque. En revanche, si je conçois une application de fitness, l'accent sera davantage mis sur le plaisir et le confort afin de garantir l'engagement de l'utilisateur. Cela implique que nous n'aurons pas les mêmes objectifs

d'utilisabilité en fonction du produit utilisé. Certaines métriques devront être priorisées. Dans le cas du défibrillateur, nous prioriserons l'efficacité et l'efficience. Dans le cas, de l'application de fitness, nous prioriserons la satisfaction et l'expérience utilisateur. Cependant, nous avons vu qu'un dispositif est systématiquement impacté par le contexte d'usage dans lequel il est évalué. **C'est ce qui confère l'approche située à l'utilisabilité.**

4.1.2. Les objectifs du « quand » dans l'évaluation de l'utilisabilité : une utilisabilité évolutive

Les objectifs d'évaluation de l'utilisabilité dépendent aussi de l'avancement de la conception du produit (le « *quand* ») et les choix de prise en compte du contexte sont différents en fonction de cet avancement. L'évaluation de l'utilisabilité est alors composée d'un ensemble d'évaluations itératives à différents stades de développement. Ces itérations successives rendraient le produit plus adapté aux besoins utilisateurs (Friedman & Wyatt, 2005; Stead et al., 1994). De plus, au fur et à mesure des essais, les compétences et les connaissances des utilisateurs peuvent évoluer, modifiant à leur tour la perception de l'utilisabilité. Par exemple, des schèmes d'utilisation difficiles à appréhender au début peuvent devenir aisés après des utilisations successives. **C'est ce qui confère l'approche évolutive à l'utilisabilité.**

4.1.3. Les objectifs du « quand évaluer quoi ? » : proposition d'un modèle à deux dimensions

En nous inspirant du modèle de Yen et Bakken (2012) et des éléments du contexte à prendre en compte dans une évaluation, nous avons pu identifier 5 niveaux de stratification pour l'évaluation de l'utilisabilité d'un système dans sa globalité (Figure 9). Chaque niveau comprend un objectif précis :

- Le niveau 0 a pour objectif la spécification de cette tâche et l'analyse des besoins et pour niveau d'analyse, la tâche.

- Le niveau 1 a pour objectif la validation de la structure de la maquette et s'intéresse à l'interaction entre l'utilisateur et les tâches.

- Le niveau 2, que nous ajoutons au modèle de Yen et Bakken (2012), consiste à vérifier l'utilisation des ressources environnantes par l'utilisateur et intègre les ressources du système dans l'interaction humain-machine.

- Le niveau 3 a pour objectif la validation de la structure du prototype dans son système global et examine la performance des tâches pour évaluer la validation du système et l'interaction humain-machine.

- Le niveau 4 vise l'intégration du système dans les environnements naturels en intégrant les facteurs environnementaux afin d'identifier les processus de travail et les effets du produit dans des contextes réels.

- Enfin, Yen et Bakken (2012) montrent un cinquième et dernier niveau faisant référence à une utilisation routinière une fois le produit commercialisé.

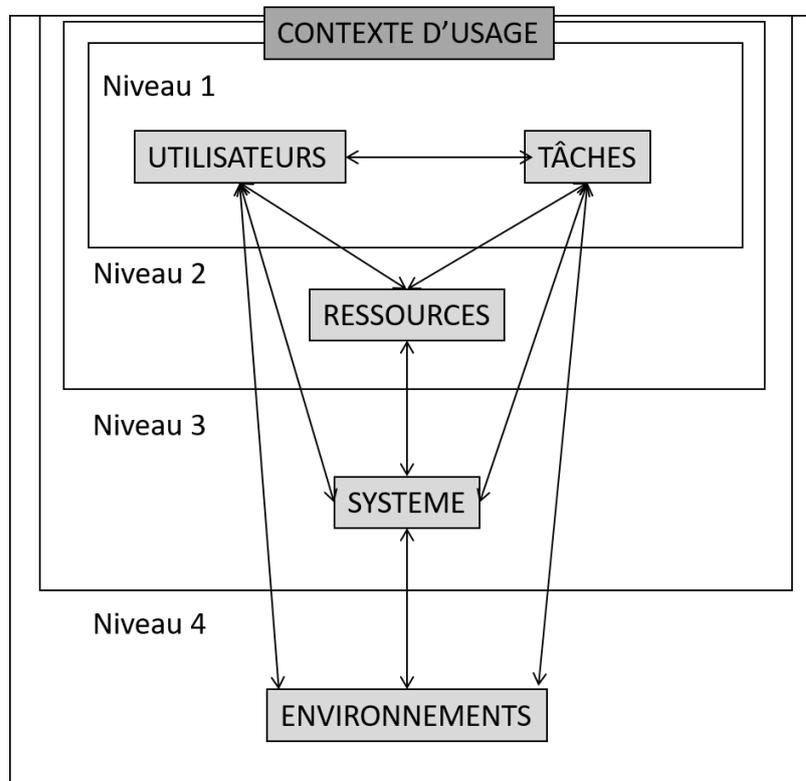


Figure 9. Représentation stratifiée de l'évaluation de l'utilisabilité inspirée de Yen et Bakken (2012).

Chaque niveau comprend un cycle itératif de conception que nous représentons dans la Figure 10. Ces niveaux sont en lien avec une approche épistémologique dans laquelle ils sont investigués (laboratoire, terrain, *in situ*). Nous définissons ces trois approches épistémologiques dans la partie suivante.

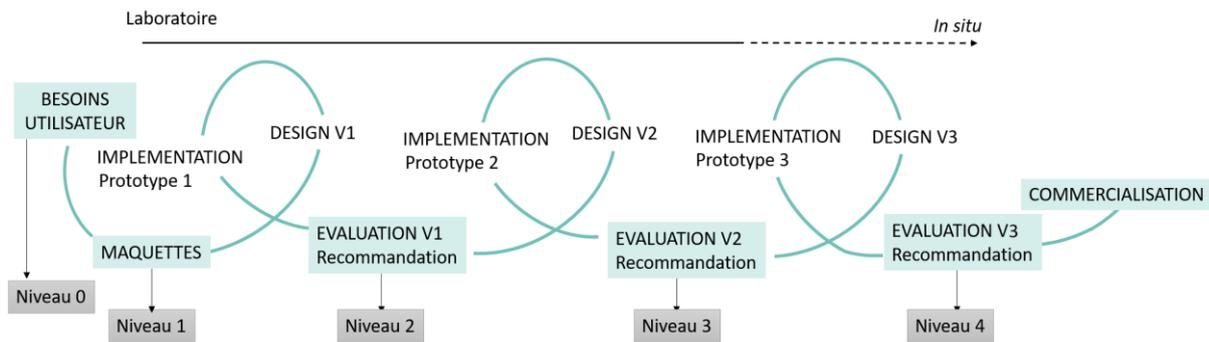


Figure 10. Représentation stratifiée de l'évaluation de l'utilisabilité en fonction du cycle de conception d'un produit.

4.2. Trois lieux de recueil de données déterminés par trois approches épistémologiques : laboratoire, terrain, *in situ*

La plupart des évaluations de l'utilisabilité se pratiquent en laboratoire (e.g., Yen & Bakken, 2012). Dans le modèle de Yen et Bakken (2012), les auteurs ne sont pas particulièrement précis sur l'utilisation du laboratoire, du terrain ou de la situation *in situ*. D'après les auteurs, l'ensemble de ces niveaux peuvent être évalués en laboratoire ou sur le terrain. Cela est probablement dû au fait que les frontières entre ces trois approches épistémologiques ne sont pas toujours bien définies. Ainsi, nous proposons de définir ce qui relève du laboratoire, du terrain et de la situation dite *in situ* (Jambon et al., 2010 ; Perrot et al., 2014). Les auteurs définissent ces trois éléments sur la base d'un continuum où le laboratoire se situerait à l'extrémité de celui-ci.

Le laboratoire se caractériserait par le contrôle de l'environnement et la simulation de l'ensemble des éléments du contexte en y apportant plus ou moins de réalisme comme ce qui est pratiqué dans les laboratoires d'utilisabilité (e.g., LOUSTIC à Rennes, LUTIN à Paris, CIC-IT à Lille, etc). La **simulation prendrait place dans l'environnement de laboratoire** dans lequel le réalisme est particulièrement travaillé. Kushniruk et al. (2013) positionnent sur le continuum la « simulation clinique » (une simulation mettant davantage l'accent sur la dimension du contexte clinique) à l'extrémité des tests en laboratoire, proche des situations de terrain. Les auteurs expliquent que cette simulation implique de tester des systèmes avec des utilisateurs représentatifs effectuant des tâches représentatives, dans des environnements représentatifs. Pour cela, deux éléments peuvent apporter des avantages considérables aux évaluations de l'utilisabilité en simulation (Kushniruk et al., 2013) : (1) la qualité du scénario d'essai écologiquement valable et (2) les simulations dans des environnements *in situ*.

À l'opposé de ce continuum, on y retrouve la **situation réelle d'usage (*in situ*) caractérisée par l'absence quasi totale de contraintes liées à l'expérimentation**. Lorsque l'utilisateur est placé dans son monde réel et dès lors que l'ensemble des éléments est naturellement présent, on considère ces pratiques comme des évaluations de « terrain » ou de situation « écologique ». Cependant, il y a une différence importante à prendre en compte entre ces deux derniers environnements.

Perrot et al. (2014) **positionnent les études de terrain à l'intersection des études en laboratoire et des études en situation réelle, les qualifiant d'extensions du laboratoire**. Il existe pourtant un consensus pour considérer que les expériences sur le terrain placent l'utilisateur dans un environnement "naturel" d'utilisation (e.g., Kjeldskov et al., 2004). Cependant, Perrot et al. (2014) soutiennent que pour considérer une étude de terrain comme une étude en situation réelle (*in situ*), l'étude ne doit pas contenir de personnels, d'équipements et de tâches assignés. En effet, Kjeldskov et al. (2005) ont par exemple démontré que la présence d'une caméra modifiait l'activité de l'utilisateur produisant une sorte de « boule de protection ». Cela a même donné lieu à l'identification de l'effet Hawthorne où le participant, se sachant observé, modifie ses pratiques et ses comportements (Macefield, 2007).

Ces trois approches ont néanmoins des intérêts complémentaires (Jake-Schoffman et al., 2017; Jambon et al., 2010). Les études en laboratoire permettent d'observer avec précision une situation en particulier, mais sont porteuses de nombreux biais. À l'inverse, si la situation réelle permet des observations limitées, elle peut révéler des comportements inattendus et évite de nombreux biais (par exemple, lié à la présence des caméras ou de l'observateur). Elle intègre également le contexte d'utilisation qui a un impact sur l'utilisabilité (ISO 9241-11 :2018).

4.3. Les types, les méthodes et les outils de l'évaluation de l'utilisabilité

4.3.1. Trois types d'évaluation de l'utilisabilité

Les évaluations de l'utilisabilité ont pour objectif de collecter des données sur l'utilisabilité d'une interface et/ou d'un produit et d'obtenir le jugement des utilisateurs permettant d'avoir un aperçu des améliorations à prendre en compte. Nous venons de voir que ces évaluations peuvent être réalisées avant, pendant et/ou après le développement de l'outil. En fonction du développement du produit, certains types d'évaluations pourraient être privilégiés. Six types d'évaluation de l'utilisabilité existent, nous les regroupons ici :

- *L'évaluation sommative vs l'évaluation formative*

Comme son nom l'indique, l'évaluation *formative* a pour but d'améliorer le produit en réalisant un diagnostic. L'évaluation réalisée permet alors de détecter les problèmes lors de la phase de conception (Hartson, 2003). L'évaluation *sommative* a pour objectifs finaux la délivrance d'une certification et la vérification de manière formelle des caractéristiques d'utilisabilité d'un produit (Nishiuchi & Takahashi, 2015). Le niveau d'utilisabilité souhaité devrait être défini lors de l'analyse des besoins et les résultats de l'évaluation sommative doivent ensuite être comparés aux objectifs initiaux. Si le niveau d'utilisabilité n'est pas atteint, il est donc nécessaire de procéder à la reconception du produit et une prochaine réévaluation de l'utilisabilité.

- *L'évaluation subjective vs l'évaluation objective*

L'objectif d'une évaluation subjective est de détecter les problèmes d'utilisabilité d'un point de vue subjectif (Nishiuchi & Takahashi, 2015). Cela est mis en forme par les opinions, les perceptions des utilisateurs ou des experts sur l'utilisabilité du produit. Un exemple de ses méthodes est l'interview, le questionnaire et/ou le focus group. Une vingtaine d'échelles promettant la mesure de l'utilisabilité existe (e.g., SUS ; UMUX ; UMUX-LITE ; NASA&TLX ; SMEQ). L'évaluation objective permet d'obtenir des résultats indépendamment des opinions personnelles de l'utilisateur (Berntsen & Dybkjær, 2010). Les observations et la collecte de données via une grille d'observation ou les méthodes d'oculométrie peuvent être utilisées. Lindgaard et Dudek, (2003), McLellan, Muddimer, et Peres (2012) conseillent d'inclure des mesures objectives et subjectives répétées dans les protocoles d'évaluation de l'utilisabilité, plus exactement de la satisfaction.

- *L'évaluation par inspection vs l'évaluation mobilisant un utilisateur réel*

Les évaluations par inspection sont réalisées la plupart du temps par des experts. Les méthodes de ces inspections comprennent l'évaluation d'heuristique, l'examen des lignes directrices, les vérifications pluralistes, les inspections de cohérence, les inspections de normes, *cognitive walkthrough*, les inspections formelles et les inspections des caractéristiques (Nielsen, 1994b). Les inspections sont généralement utilisées au début du processus de conception en évaluant des prototypes où le système ne peut pas encore être présenté à un utilisateur. Les évaluations par le biais d'un utilisateur final arrivent généralement après les phases d'inspection où un prototype est présentable et compréhensible par l'utilisateur.

Les évaluations par des utilisateurs réels sont généralement plus compliquées à mettre en place. Elles nécessitent d'avoir un utilisateur cible – un potentiel futur utilisateur (*end user*) du produit – reproduisant une série de tâches représentatives du système dont les méthodes varient (cf. Partie I, Chapitre 1, 4.3.3. Le choix des bonnes méthodes et des bons outils, p. 48). L'intérêt premier des évaluations par des utilisateurs finaux est le fait de devoir interroger une population représentative. L'inconvénient, en revanche, est le coût financier et temporel de ces évaluations. Le réalisme de l'évaluation soulève aussi des interrogations.

Nous proposons une synthèse des méthodes issues des évaluations par inspection et avec des utilisateurs finaux dans le Tableau 5.

4.3.2. Trois types de méthodes d'évaluation de l'utilisabilité : inspection, test utilisateur, enquête

Trois grands types de méthodes d'évaluation de l'utilisabilité existent : l'inspection, le test utilisateur et l'enquête. Nous avons positionné ces méthodes sur un mapping en fonction de deux axes (l'axe experts/utilisateurs finaux et l'axe *in situ*/en laboratoire correspondant au continuum du contrôle de l'environnement dans lequel est réalisée l'étude) dans lequel nous regroupons les trois grandes catégories (Figure 11).

Nous détaillons ces trois méthodes :

- **Les méthodes par inspection** requièrent les compétences d'un expert (en utilisabilité) et des outils d'évaluation. L'ensemble des méthodes par inspection (heuristiques, cognitive walkthrough, critères ergonomiques) ont pour objectif commun de détecter des difficultés d'utilisation ou des ralentissements/alourdissements du travail de l'utilisateur dans sa manipulation du produit. Cet/ces expert(s) passe(nt) en revue le système en se conformant à des critères, des règles ou des heuristiques (Nielsen, 1994b). Cette évaluation repose sur les compétences et sur le jugement de l'expert ;
- **Les tests utilisateurs** demandent la participation active d'utilisateurs cibles et se déroulent généralement dans un environnement contrôlé comme le laboratoire. Dans la plupart des cas, il est demandé aux participants d'exécuter des tâches représentatives pour lequel le produit a été conçu. L'objectif des tests utilisateurs est alors de détecter par le biais d'outils d'analyse tels que des verbalisations ou des indicateurs de performance (le nombre d'erreurs commises et/ou le temps d'exécution de la tâche) ou encore les difficultés rencontrées par le participant (Kjeldskov & Stage, 2004).

Tableau 5. Synthèse des méthodes d'évaluation ergonomique en fonction de leur type.

Types	Méthodes	Outils
Evaluation par inspection (Nielsen, 1994b)	Évaluation heuristique	Les heuristiques de Nielsen et Molich (1990) composés de 10 critères.
	Cognitive walkthrough (ou inspection cognitive)	CW Version 1 (Lewis et al., 1990) CW Version 3 : "a practitioner's guide"; (Wharton et al., 1994)
	Analyse d'action	
	Guides et critères ergonomiques	(Vanderdonckt, 1994 ; Farenc & Palanque, 1999)
Evaluation avec utilisateurs finaux	Tests d'utilisabilité	Think-aloud ; question-asking ; méthode de shadowing ; mesure de performances ; observation ; magicien d'Oz ; prototype
	Enquêtes	Traceur d'activité (cheval de Troie) ; Journal de bord ; Observations ; entretien, questionnaire, le focus group
	Modélisation analytique	
	Questionnaires	SUS ; NASA
	Monitoring de l'utilisateur	Oculométrie, Eye tracking, pupillométrie, sudation...

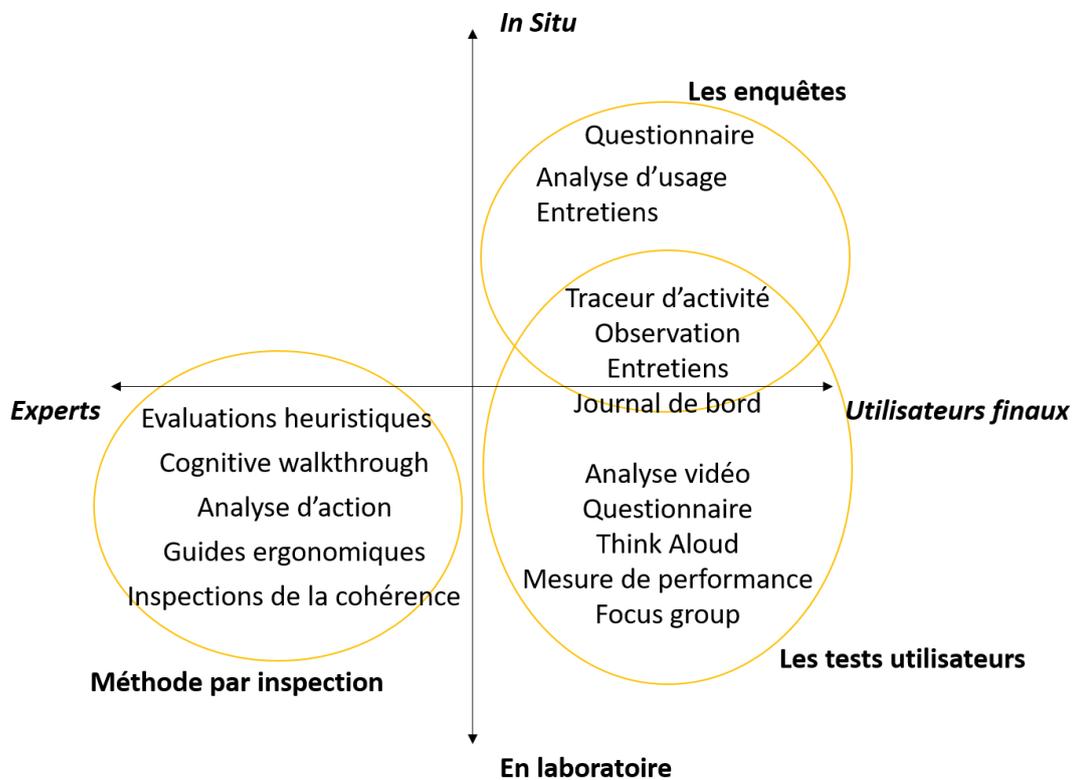


Figure 11. Représentation schématisée des trois grandes catégories d'évaluation de l'utilisabilité en fonction du lieu du déroulement de l'évaluation et du type d'évaluateurs (experts vs utilisateurs finaux).

Une fois les points bloquants identifiés, des recommandations de correction sont proposées. La méthode des tests utilisateurs nécessite de passer par un certain nombre d'étapes plus fastidieuses rendant cette méthode plus coûteuse que celle des inspections. Dumas et Redish (1999) proposent un guide pratique de mise en place de tests utilisateurs. Ces étapes sont représentées dans le schéma (Figure 12) ci-dessous. Les tests utilisateurs doivent être réalisés de manière très rigoureuse avec un échantillon suffisamment important ;

- **Les enquêtes** nécessitent également la participation des utilisateurs. En revanche, les utilisateurs sont laissés dans leur environnement naturel et aucune tâche spécifique ne leur est attribuée. L'objectif de cette méthode est de collecter les opinions et les perceptions des utilisateurs sur leur usage (Holtzblatt & Jones, 2017).

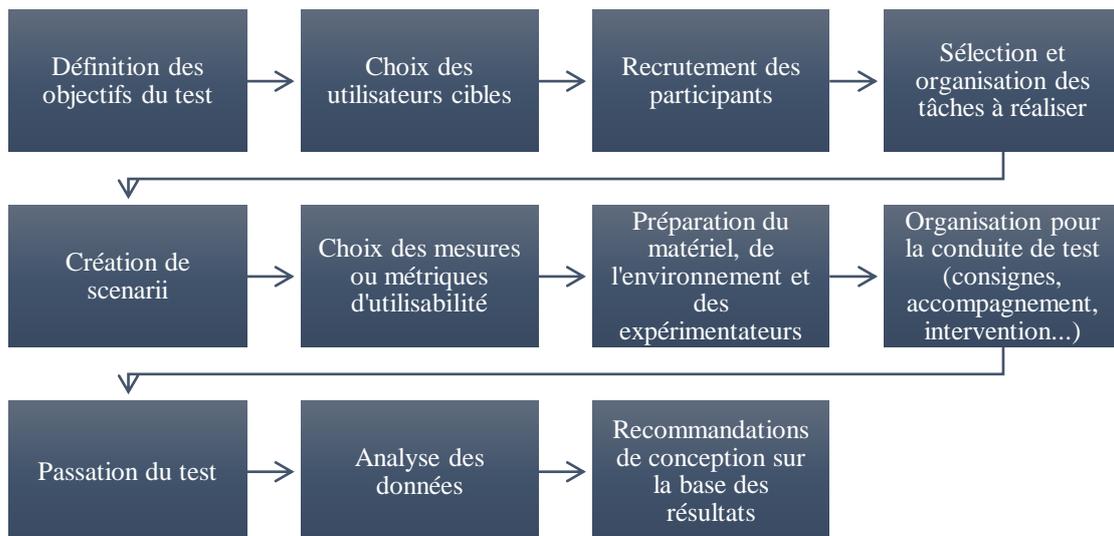


Figure 12. Déroulement d'un test utilisateur d'après Dumas et Redish (1999).

4.3.3. Le choix des bonnes méthodes et des bons outils

Il existe dans la littérature un large choix de méthodes afin d'évaluer l'utilisabilité d'un produit. Les méthodes choisies pour une évaluation dépendent majoritairement du contexte et des moyens pour réaliser l'étude. **Il est important de préciser que pour choisir les bons outils et méthodes, Ammenwerth et al. (2003), Georgsson et Staggers (2016) ainsi que Jaspers (2009) recommandent de combiner des méthodes complémentaires afin d'obtenir une vision holistique de l'utilisabilité.**

Chaque méthode et outil doit être choisi en fonction du contexte de l'évaluation, et notamment l'avancement du projet, la présence d'experts ou d'une population cible suffisamment accessible ou encore les moyens financiers.

Nous proposons ci-dessous (Tableau 6) de détailler les méthodes et outils d'évaluation de l'utilisabilité en fonction des critères suivants :

- **Pourquoi – Objectif de l'outil ;**
- **Comment – Description de l'utilisation** de l'outil et de son utilisation ;
- **Où – Lieu de l'utilisation ;**
- **Avec – Matériel (ou les matériaux)** nécessaire(s) au déploiement de cet outil ;
- **Quand – Phase d'application** de l'outil en fonction du cycle de conception du produit par rapport à notre modèle HANDLED ;
- **Critiques – Principaux avantages et limites** de l'outil ;
- **Quoi – Métriques analysées.**

Tableau 6. Synthèse des méthodes et outils d'évaluation de l'utilisabilité les plus connus.

Outils d'évaluation	Pourquoi - Objectif de l'outil	Comment - Description de l'utilisation	Où - lieu de l'utilisation	Avec Matériel(aux) nécessaire(s) et ressources	Quand - Phases d'application et niveau	Critiques - Avantages et limites	Quoi Métriques analysées
Évaluation d'experts : heuristique de ou critères ergonomiques (Bastien & Scapin, 1993; Nielsen & Molich, 1990)	Évaluer l'utilisabilité d'un système sur la base de critères/heuristiques analysés par des experts en utilisabilité.	Chaque expert est invité à analyser la qualité du système sur la base de guidelines, de critères ou d'heuristiques préalablement choisis. Puis, les experts mettent en commun leur avis et recommandation avec l'ensemble du groupe. Plusieurs types d'évaluer par des experts existent et regroupent généralement plusieurs critères identiques comme la visibilité, la gestion des erreurs, la charge mentale, le feedback...	Réalisé sur le lieu de conception du système.	Nécessite la présence d'un groupe d'expert, des critères utilisés et d'au moins une maquette du système.	Niveau 0 et 1 Dès la conception des premières maquettes.	<u>Avantages</u> : peu coûteux et simple à mettre en place, car ne mobilise pas de participant. Permet un feedback rapide des maquettes proposant des recommandations pour l'amélioration du système. <u>Limites</u> : dépend du niveau de compétences et d'expertise de l'évaluateur. Ne peuvent pas se soustraire à une évaluation plus approfondie comme les tests utilisateurs. Méthode non centrée sur l'utilisateur.	Grille d'évaluation : Compatibilité, adaptabilité, guidage, cohérence homogénéité, charge de travail, signification des codes, contrôle explicite, gestion des erreurs...
Cognitive walkthrough (inspection cognitive) (Lewis et al., 1990)	Évaluer l'utilisabilité d'un système sur la base de l'expertise de spécialiste en utilisabilité.	L'expert se met à la place de l'utilisateur pour identifier les problèmes d'utilisabilité que celui-ci pourrait rencontrer. Pour cela, il explore le système en relevant les problèmes en fonction de chaque tâche.	Réalisé sur le lieu de conception du système.	Nécessite la présence d'un groupe d'expert ainsi que le produit finalisé.	Niveau 0 et 1 Phase avancée de la conception où le produit est finalisé.	<u>Avantages</u> : peu coûteux à mettre en place, car ne mobilise pas de participant. La méthode aide à se mettre à la place de l'utilisateur et se questionner sur leurs attentes <u>Limites</u> : cette méthode s'applique très mal à des systèmes complexes. Ne peut pas se soustraire à une évaluation plus approfondie comme les tests utilisateurs. Méthode non centrée sur l'utilisateur.	Grille d'évaluation : Facilité d'utilisation, apprentissage.
Think-aloud	Comprendre le cheminement cognitif et la perception en temps réel de l'utilisateur lors de son interaction avec le système.	La méthode du think-aloud ou penser à voix haute demande de recueillir les verbalisations de l'utilisateur pendant sa manipulation. Les utilisateurs sont invités à décrire chacune de leurs pensées.	Réalisé généralement en laboratoire.	Nécessite des utilisateurs cibles, un prototype fonctionnel ou des maquettes ainsi qu'une caméra ou microphone.	Niveau 1, 2 et 3 Tout au long du processus de conception, permet l'évaluation itérative des maquettes et prototypes fonctionnels.	<u>Avantages</u> : Rapide, peu coûteux, se base sur l'observation de l'utilisateur. <u>Limites</u> : la verbalisation peut biaiser la mesure des performances, car le participant peut prendre plus son temps afin d'expliquer sa pensée (ce qui peut aussi le perturber). Les données	Efficience, efficacité.

						ont difficiles à traiter et à interpréter.	
Mesure de performance / Observation instrumentée par vidéo	Évaluer l'utilisabilité d'un système en observant les performances de l'utilisateur entrain de manipuler.	L'utilisateur cible est invité à réaliser une ou plusieurs tâche(s) précise(s) avec le système. Les évaluateurs ont préalablement préparé une grille d'analyse à prendre en compte afin de vérifier les objectifs d'utilisabilité (Le temps de manipulation, le nombre d'erreurs...).	En laboratoire/ milieu contrôlé ou en situation naturelle impliquant des consignes de tâches à l'utilisateur cible.	Caméra(s), Logiciel de codage vidéo (The Observateur, BORIS...), Un prototype ou une maquette, des utilisateurs cibles.	Niveau 1, 2, 3, 4 Tout au long du processus de conception, permet l'évaluation itérative des maquettes et prototypes fonctionnels.	<u>Avantages</u> : Permet de collecter des données objectives et fines comme le nombre de clics, ou le temps de regard vers la notice. <u>Limites</u> : méthode coûteuse et longue (temps de préparation, recrutement, passation, analyses, codage vidéo...) / Les tâches demandées éloignent l'utilisateur de son utilisation réel avec le produit.	Efficience, efficacité.
Traceur d'activité mouchard électronique	Évaluer l'utilisabilité du système à distance par le biais d'un logiciel espion implanté dans l'outil invisible à l'utilisateur cible.	Le mouchard transmet et enregistre toutes les interactions que l'utilisateur a avec le système comme par exemple le temps de réponse, le nombre de clics. Les données récoltées sont ensuite traduites et analysées.	En milieu naturel (l'utilisateur a néanmoins connaissance de ce mouchard).	Logiciel espion ou conception de lignes de code afin d'enregistrer les éléments voulus. Un prototype fonctionnel.	Niveau 4 Phase avancée de la conception où le produit est finalisé ou avec des produits déjà présents sur le marché.	<u>Avantages</u> : Permet de collecter des données objectives à distance et de manière longitudinale. <u>Limites</u> : difficile à mettre en place, car il faut le bon logiciel espion et l'autorisation de collecter les données du participant. Il est aussi difficile de déterminer les variables adéquates pour évaluer correctement la performance, surtout si le système est très complexe et à plusieurs types d'interaction. Ne permet pas d'avoir une vision du comportement de l'utilisateur.	Efficience, efficacité, apprenabilité.
Questionnaire ou échelles d'utilisabilité (Root & Draper, 1983)	Permet de collecter des données quantitatives sur la satisfaction des utilisateurs grâce à des items standardisés dans un but d'obtenir des informations sur l'utilisabilité.	Les utilisateurs répondent à l'ensemble des items les interrogeant sur la qualité perçue en se positionnant sur une échelle de Likert.	Se réalise généralement en laboratoire (sur papier), mais il est possible de le réaliser en ligne ou de le glisser dans un journal de bord.	Un questionnaire préalablement standardisé.	Niveau 1, 2, 3, 4 Peut être proposé aux utilisateurs lors de nombreux cycles itératifs de conception. Les questionnaires sont souvent associés à des tests utilisateurs.	<u>Avantages</u> : collecte de données quantitatives comparables d'un système à l'autre. La passation est simple, rapide et peut se faire à distance. L'analyse est aussi facilitée. <u>Limites</u> : peut comporter des biais de passation et ne permet pas une collecte fine des résultats ni les arguments des utilisateurs sur leurs jugements.	Satisfaction.

Outils d'évaluation	Pourquoi - Objectif de l'outil	Comment - Description de l'utilisation	Où – lieu de l'utilisation	Avec - Matériel(aux) nécessaire(s) et ressources	Quand - Phases d'application et niveau	Critiques – Avantages et limites	Quoi – Métriques analysées
Observation (Blomberg et al., 2003)	Permet d'accéder aux comportements non-verbaux dont les utilisateurs n'ont pas toujours conscience. (Blomberg et al., 2003). Il permet aussi d'analyser l'environnement naturel, physique, culturel, organisationnel et social d'un utilisateur manipulant un produit. (Isomursu et al., 2007)/.	Observer un utilisateur manipulant le produit en direct. Plusieurs techniques d'observation existent et les données collectées peuvent être différentes. Certaines nécessitent la présence d'un observateur passif (observation systématique, « fly on the wall », shadowing) ; d'autres d'un actif (enquête contextuelle et observation participante) ; ou bien sans observateur (systèmes de tracking) ; ou encore l'on va se mettre dans la peau d'un utilisateur (agent sous couverture).	Généralement utilisé en situation naturelle ou sur le terrain.	Conception au préalable d'une grille d'observation. Peut impliquer une caméra.	Niveau 1, 2, 3 L'observation peut être utilisée à tout moment dans la phase de conception d'un produit (recueil de besoins et évaluation).	<u>Avantages</u> : Permet de collecter des données en situation réelle et permet d'observer des comportements inattendus. <u>Limites</u> : difficile de répliquer les résultats ou d'observer un échantillon représentatif. La simple implantation d'un observateur est un biais.	Efficacité, efficacité.
Entretien (Flanagan, 1954; Rogers, 1942; Vermersch, 1991)	Permet de récolter des données qualitatives c'est-à-dire la conversation d'un utilisateur cible sur sa perception du produit.	Il est nécessaire de concevoir une grille d'entretien comprenant les questions prédéfinies à poser à l'utilisateur cible. Un expérimentateur interviewe séparément les utilisateurs cibles recrutés. Plusieurs types d'entretiens existent : les incidents critiques, l'entretien d'explicitation, l'auto-confrontation, l'instruction au sosie...	L'entretien peut se dérouler dans l'environnement naturel de l'utilisateur ou dans un milieu contrôlé (le laboratoire).	Conception au préalable d'une grille d'entretien et d'un microphone ou caméra.	Niveau 0, 1, 2, 3, 4 L'entretien peut être utilisé à tout moment dans la phase de conception d'un produit. Il peut s'associer avec le test utilisateur et le focus group.	<u>Avantages</u> : Le matériel de l'entretien n'est pas obligatoire et permet de collecter des données précises en fonction du contexte d'usage. <u>Limites</u> : ne permet pas d'avoir un large échantillon.	Satisfaction.
Focus group (Krueger & Casey, 2009)	Permet de récolter des données qualitatives c'est-à-dire les échanges dans un groupe.	Il faut réunir plusieurs utilisateurs cibles autour pour échanger sur leurs idées, leurs opinions, leurs perceptions du produit à évaluer. Un expérimentateur anime le groupe et analysera le contenu des échanges entre les participants.	Le focus group se réalise généralement dans un endroit pouvant regrouper ces utilisateurs dans un lieu calme.	Préparation d'un guide d'animation Microphone ou caméra.	Niveau 0, 1, 2, 3, 4 Comme l'entretien, le focus group peut être utilisé à tout moment dans la phase de conception.	<u>Avantages</u> : moins coûteux en temps que la réalisation d'entretien en face à face. Les explications de certains permettront aux autres de construire leur propre explication et pousser l'entretien plus loin qu'un face à face. <u>Limites</u> : données pouvant être imprécises ou inexacts. Ne permet pas d'aborder des questions intimes. Certains peuvent être influencés par des discours dominants.	Satisfaction.

5. Synthèse du chapitre 1

L'objectif de ce chapitre était de positionner notre recherche au sein d'une démarche de conception centrée utilisateur mobilisant une **activité collaborative et pluripartenariale**. Pour cela, nous nous appuyons sur une **approche pluridisciplinaire** (ergonomie cognitive, psychologie cognitive et sciences de la conception) en focalisant nos recherches sur **l'utilisabilité**, notions centrales de l'évaluation ergonomique, mais subissant plusieurs critiques liées à ces lacunes.

Après analyse critique des définitions et des modèles de l'utilisabilité, nous avons défini cette notion comme **la propriété interactive multidimensionnelle instrumentale et non instrumentale d'un système avec un utilisateur donné dans un contexte d'usage spécifique**. Ainsi, après étude des modèles de Shackel (1991), Nielsen (1994) et ISO 9241-11 (2018), nous avons retenu **trois métriques majeures de l'utilisabilité** (*efficacité, efficience et satisfaction*) ainsi que **deux métriques connexes** (*l'apprenabilité et la mémorisation*).

Nous avons également défini l'utilisabilité comme **située**, car dépendante de son contexte d'usage composé de quatre composants (**l'utilisateur, la tâche, les ressources et l'environnement**). L'utilisabilité est aussi **évolutive**, dans la mesure où cette propriété s'intègre dans le processus de conception d'un produit à travers des évaluations itératives. Ces évaluations peuvent être appréhendées à travers quatre niveaux d'expertises suivant **la matrice utilisateur-tâche-système-ressources-environnement**. En fonction de l'avancement du produit dans sa conception, ces évaluations deviendront de plus en plus complexes jusqu'à prendre en compte l'ensemble des éléments du contexte.

Enfin, nous avons abordé, le choix des méthodes et des outils à utiliser en fonction de l'évaluation choisie. Nous avons vu que l'ensemble des métriques de l'utilisabilité devait être évalué de manière indépendante (Frøkjær et al., 2000) en combinant une collecte de données subjectives, objectives, qualitatives et quantitatives (Jaspers, 2009).

Nous proposons dans la Figure 13, une représentation de l'utilisabilité prenant en compte l'ensemble des éléments abordés dans ce chapitre.

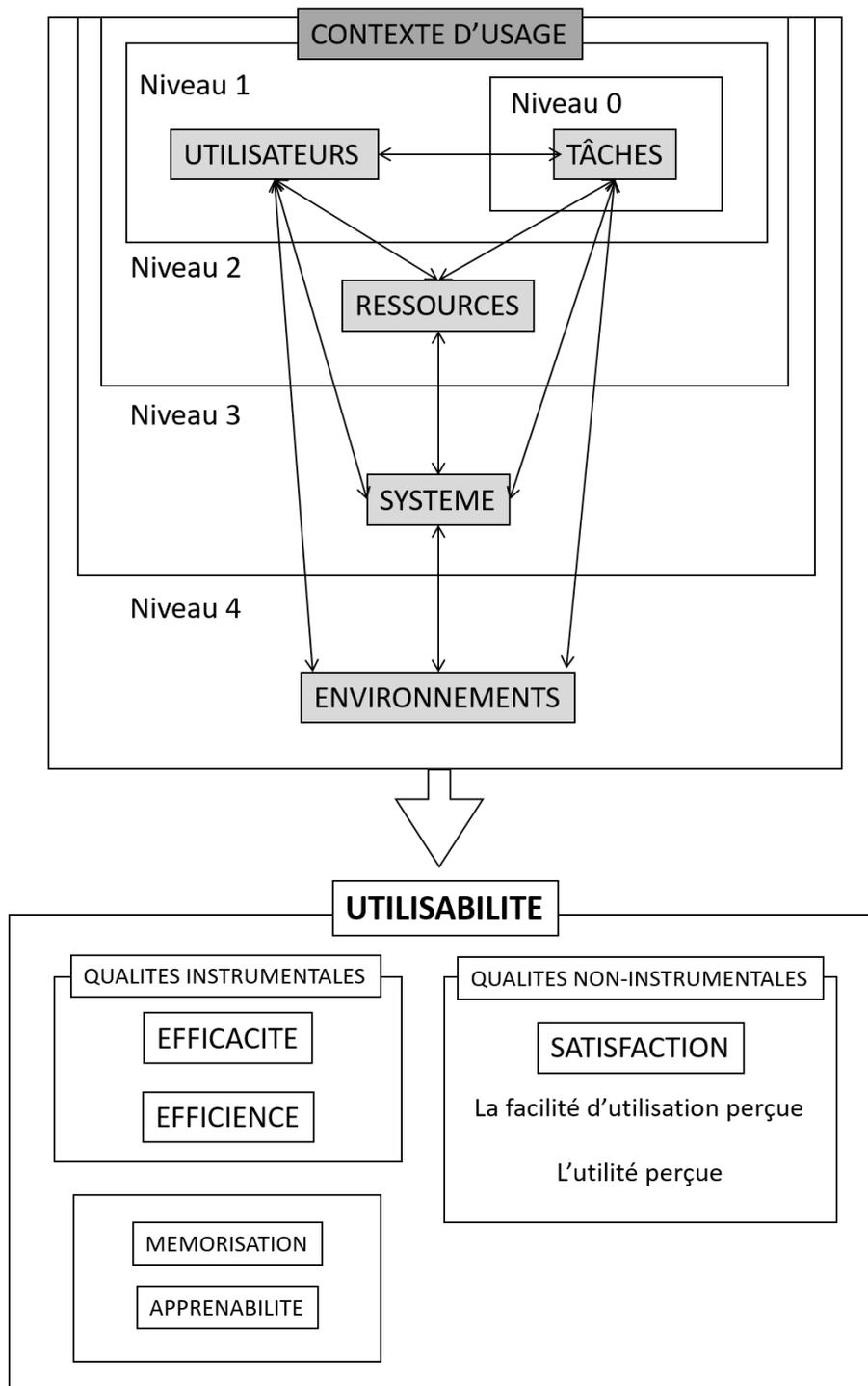


Figure 13. Proposition d'un modèle d'évaluation de l'utilisabilité incluant les métriques de l'utilisabilité et le contexte d'usage sur 4 niveaux d'analyse.

CHAPITRE 2 : L'ÉVALUATION DE L'UTILISABILITÉ DE DISPOSITIFS CONNECTÉS EN SANTÉ POUR LE DOMICILE PATIENT

1. Les dispositifs médicaux et le marquage de conformité européen

1.1. La définition des dispositifs médicaux

Le dispositif médical est une catégorie d'instrument de santé imposant une réglementation pouvant varier suivant les pays. L'article L.5211-1 issu du code de la santé publique définit, en France, le dispositif médical comme :

« Tout instrument, appareil, équipement, matière, produit, à l'exception des produits d'origine humaine, ou autre article utilisé seul ou en association, y compris les accessoires et logiciels nécessaires au bon fonctionnement de celui-ci, destiné par le fabricant à être utilisé chez l'homme à des fins médicales et dont l'action principale voulue n'est pas obtenue par des moyens pharmacologiques ou immunologiques ni par métabolisme, mais dont la fonction peut être assistée par de tels moyens. Constitue également un dispositif médical le logiciel destiné par le fabricant à être utilisé spécifiquement à des fins diagnostiques ou thérapeutiques ».

Il est destiné uniquement à des fins médicales, c'est d'ailleurs ce qui le distingue des produits issus de la e-santé pouvant s'apparenter à du « bien-être ».

1.2. La classification et la conception des dispositifs médicaux

Les dispositifs médicaux sont classés en quatre catégories en fonction des risques potentiels pour la santé. Chaque catégorie est associée à une évaluation particulière et à des contrôles spécifiques.

- **Classe I (classe de risque le plus faible)** : elle comprend par exemple, les lunettes correctrices, les véhicules pour personnes handicapées, les béquilles, etc. ;
- **Classe IIa (risque potentiel modéré/mesuré)** : il s'agit ici, entre autres, des lentilles de contact, des appareils d'échographie, des couronnes dentaires, des **tensiomètres connectés ou oxymètres connectés** ;
- **Classe IIb (risque potentiel élevé/important)** : elle comprend notamment les préservatifs, les produits de désinfection des lentilles, **le dispositif Smart Angel** ;
- **Classe III (classe de risque la plus élevée)** : elle inclut à titre d'exemples, les implants mammaires, les stents, les prothèses de hanche, etc.

La catégorisation du dispositif médical est sous la responsabilité du fabricant. En revanche, toute commercialisation dans l'Union Européenne d'un dispositif médical est soumise à des normes de sécurité et de santé. Le produit sera assigné d'un marquage Conformité Européenne ou CE si celui-ci respecte correctement l'ensemble des exigences essentielles normatives, dont l'exigence essentielle liée à l'ergonomie du dispositif médical. La norme IEC 62366-1:2015/Amd1:2020 définit le « *processus d'ingénierie à l'aptitude à l'utilisation* » (en anglais : *usability engineering process*) donne présomption de conformité à cette exigence. La norme ISO 13485 (2016) énonce « *les exigences relatives au système de management de la qualité* ». Cette norme propose une aide à la conception, au développement voire à la destruction d'un dispositif médical. La norme ISO 14971 (2019) relative à la « *gestion des risques* » aide le fabricant à identifier les dangers associés au dispositif médical. L'aptitude à l'utilisation y est définie comme une « *caractéristique de l'interface utilisateur qui établit l'efficacité, le rendement, la simplicité d'apprentissage et la satisfaction de l'utilisateur* » (IEC 62366-1:2015, p. 7). **L'utilisabilité est considérée comme une propriété majeure pour le développement des dispositifs médicaux.** Une fois sur le marché, le produit est surveillé par l'Agence Nationale de Sécurité du Médicament et des produits de santé ou ANSM ainsi que par la Direction Générale de la Concurrence, de la Consommation, et de la Répression des Fraudes ou DGCCRF.

Il est important de préciser que dans nos travaux de recherche, **nous utilisons majoritairement le terme « *dispositif connecté en santé* » et non « *dispositif médical* » étant donné que la plupart des dispositifs en cours de conception n'ont pas encore le marquage CE leur permettant d'obtenir le statut de dispositif médical.** C'est pourquoi le dispositif *Smart Angel* est actuellement un dispositif connecté en santé ambitionnant de devenir un dispositif médical.

2. La « e-santé » ou « santé connectée » : un domaine émergent non sans écueils

2.1. La définition de la e-santé

La e-santé¹² (ou santé connectée) apparue à la fin des années 1990, est définie par Eysenbach (2001, p.1) comme « *un domaine émergent à l'intersection de l'informatique*

¹² En anglais : *e-health*

médicale, de la santé publique et des affaires, qui fait référence aux services de santé et à l'information fournie ou améliorée par Internet et les technologies connexes. Dans un sens plus large, le terme caractérise non seulement un développement technique, mais aussi un état d'esprit, une façon de penser, une attitude et un engagement en faveur d'une pensée globale et en réseau pour améliorer les soins de santé aux niveaux local, régional et mondial en utilisant les technologies de l'information et des communications »¹³. Eysenbach (2001) explique que ce terme est apparu de l'industrie et du marketing afin de conserver les mêmes codes (et la même ferveur) que ceux du e-commerce. Internet fut précurseur de nouveaux défis pour l'industrie des technologies médicales. Eysenbach (2001) en recense trois : (1) la capacité des consommateurs à interagir avec leurs systèmes en ligne ; (2) l'amélioration des possibilités de transmission de données entre institutions ; (3) les nouvelles possibilités de communication entre pairs.

La e-santé se caractérise par son concept dynamique, en perpétuelle évolution technologique surfant sur les nouvelles tendances comportementales. Ce domaine englobe deux principaux concepts (Figure 14): les Systèmes d'Information (SI) de santé d'un côté et la télésanté classifiée par le Livre Blanc du Conseil national de l'Ordre des médecins (2015).

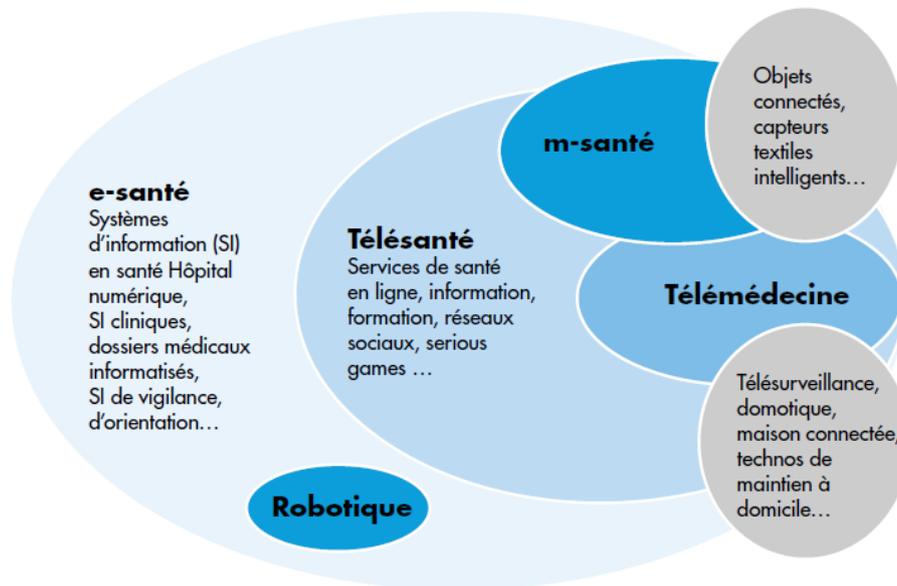


Figure 14. Définition de la e-santé dans le livre Blanc de l'ordre des médecins (2015, p. 9).

¹³ Traduit de l'anglais : "E-health is an emerging field in the intersection of medical informatics, public health and business, referring to health services and information delivered or enhanced through the Internet and related technologies. In a broader sense, the term characterizes not only a technical development, but also a state-of-mind, a way of thinking, an attitude, and a commitment for networked, global thinking, to improve health care locally, regionally, and worldwide by using information and communication technology." (Eysenbach, 2001, p.1)

2.2. La e-santé, une technologie émergente ?

En quoi la e-santé est-elle émergente si l'on reprend la définition d'Eysenbach (2001) ? Pour répondre à cette question, nous nous sommes basées sur les caractéristiques des technologies émergentes définies par Kjeldskov (2003) :

- La technologie se doit **d'être novatrice où l'avancée technologique est partiellement réalisée voire en devenir**. La e-santé est un concept récent né grâce aux développements d'internet permettant de renouveler la perception des systèmes de santé. Ces technologies ne sont pas encore pleinement exploitées, mais elles sont particulièrement prometteuses pour faire face aux difficultés sanitaires actuelles ;
- Les **usages de la technologie sont peu précis**. De nouvelles applications à la e-santé sont découvertes tous les jours. Elles facilitent l'organisation de l'hôpital, la prise de décision des médecins, le suivi des patients, l'entraide entre patients ou entre professionnels, etc. Il s'agit là d'autant de secteurs dans lesquels les usages peuvent être envisagés. Nous présentons les principaux usages de ce domaine de la santé dans la section suivante (cf. Partie I, Chapitre 2, 2.3.1. Le Système d'Information de Santé (ou Hospitaliers), p. 59 et 2.3.2. Les Technologies de l'Information et de la Communication ou TIC, p. 60) ;
- Les technologies émergentes font face à **plusieurs limites ralentissant leur application massive**. La e-santé appartient au domaine de la santé. La santé est un domaine à lui seul très complexe à modifier. En ajoutant du numérique à la santé, la sécurité des données est soumise à rude épreuve (e.g., Simon, 2017). Les études portant sur la e-santé sont aussi confrontées aux problématiques organisationnelles des hôpitaux ainsi qu'aux difficultés administratives et législatives. Nous présentons un panel de freins au développement de la e-santé dans la section suivante (cf. Partie I, Chapitre 2, 2.4. Les avantages et les freins de la e-santé, p. 61). Par ailleurs, les difficultés rencontrées dans le projet *Smart Angel* mettent en évidence ces limites ralentissant la diffusion massive du dispositif médical (cf. Partie III, Chapitre 2, 3. Contraintes de la thèse : Retour d'expérience sur un projet de conception pluripartena, p. 288) ;
- Les technologies émergentes promettent de **transformer le contexte économique et social** de leurs domaines. La e-santé est promue pour pallier les difficultés économiques

dans le domaine de la santé. Le projet *Smart Angel* répond aux difficultés de suivi à distance des patients. Son retour sur investissement devrait désengorger les hôpitaux en augmentant le nombre d'actes de chirurgie ambulatoire jugée plus économique, mais aussi en améliorant la relation patient-soignant.

En conclusion, la e-santé possède bien les quatre caractéristiques définies par Kjeldskov (2003). Cependant, nous considérerons que **la e-santé n'est pas une technologie spécifique, mais représente un domaine dans lequel sont regroupées plusieurs technologies émergentes.**

À noter qu'une technologie émergente peut être considérée comme un dispositif médical et obtenir un marquage de conformité européenne, on parle alors de **Dispositif Médical Connecté ou DMC.**

2.3. Les systèmes d'information de santé et la télésanté

La e-santé peut être décomposée en deux sous-domaines : les systèmes d'information de santé (ou hospitaliers) et les technologies de l'information et de la communication en santé (Simon et al., 2011; 2015).

2.3.1. Le Système d'Information de Santé (ou Hospitaliers)

Généralement méconnue du grand public, la e-santé comprend les Systèmes d'Information de Santé ou SIS¹⁴ (ou Systèmes d'Information Hospitaliers ou SIH) que l'on retrouve très largement dans la littérature scientifique. Le terme « système » comprend un ensemble de composants interagissant ensemble vers un but commun de sorte que cet ensemble forme un tout (e.g., la télémédecine est composée d'interfaces patient mais également médecin, elle comprend également la participation de médecins, d'administrateurs, d'informaticiens...). Pour le cas des SIS, l'objectif est alors d'améliorer la gestion des services de santé par un soutien optimal de l'information (Lippeveld et al., 2000). L'information est alors un ensemble de faits et/ou de données. De nombreux projets ont vu le jour dans un but d'échanger électroniquement des données de santé entre soignants comme le Dossier Médical Informatisé publié afin de faciliter le suivi des dossiers patients. Ces dossiers ont pu être exploités en 2011 et sont officiellement disponibles pour les patients seulement depuis novembre 2018. Un autre

¹⁴ en anglais : *Health Information Systems* ou HIS

exemple qui peine à se lancer est la prescription électronique. Son lancement a commencé début 2017 et facilite l'échange d'information entre un prescripteur et un pharmacien.

2.3.2. Les Technologies de l'Information et de la Communication ou TIC

La e-santé comprend aussi la télésanté ou les TIC appliquées à la santé. Ces dispositifs sont définis par « *l'ensemble des technologies, des réseaux et des services de soins basé sur la télécommunication et comprenant les programmes d'éducation, de recherche collaborative, de consultation ainsi que d'autres services offerts dans le but d'améliorer la santé du patient* » (Moore, 1999, p.246). Ces dispositifs sont composés de la télémédecine et de la m-santé que nous présentons ci-dessous.

a. La télémédecine

Le terme « *télé* » signifie en grec ancien « *loin* ». De ce fait, la télémédecine est définie comme un moyen de télécommunication numérique afin de créer un lien à distance entre un professionnel de santé et un patient. La télémédecine regroupe un certain nombre de concepts (Simon, 2015) :

- La téléconsultation rend possible la réalisation de consultations médicales à distance ;
- La télé-expertise aide les professionnels à obtenir l'expertise ou l'avis de confrères ;
- La télésurveillance ou télé-monitoring permet aux professionnels de santé de tracer des données physiologiques ou comportementales de manière automatisée d'un patient dans le but de les interpréter. Le patient ou les membres de la famille sont seuls responsables de la saisie et de la transmission de leurs données sans l'aide d'un fournisseur de soins de santé, comme un personnel médical ou paramédical (Paré et al., 2007) ;
- La téléassistance permet d'assister à distance des professionnels de santé entre eux dans la réalisation d'un acte médical.

b. La m-santé ou santé mobile

La m-santé (ou la santé mobile) regroupe l'ensemble des services de santé disponibles par le biais d'une application sur tablette ou smartphone. Elle peut être définie comme « *la pratique médicale et de santé publique appuyée par des appareils mobiles (e.g., téléphones mobiles, appareils de surveillance des patients et appareils sans fil)* »¹⁵ (Schnall et al., 2016, p.244). Il est aussi possible de parler de u-santé ou santé ubiquitaire (e.g., Kim et al., 2011). Ce

¹⁵ Traduit de l'anglais de : "medical and public health practice supported by mobile devices (e.g. mobile phones, patient monitoring device and wireless devices.)"

concept a vu le jour en 2011 avec l'apparition des smartphones et s'est alors fortement démocratisé. En 2012, 97 000 applications de santé ont été créées par des entreprises, des éditeurs d'information médicale ou des laboratoires pharmaceutiques (Mennecier, 2014 cité dans Allaert et al., 2016). À ces applications s'ajoutent les objets connectés de plus en plus présents sur le marché comme les montres, les balances, les fourchettes ou encore les verres connectés. Il est facile de se perdre dans ce marché particulièrement juteux et prometteur. Nous proposons en Annexe (cf. Annexe 3 : Mapping des produits de la m-santé et patterns des usages, p. 343), un benchmark présenté sous forme de mapping comprenant deux axes (curatif-préventif *versus* applications-objets connectés) recensant les applications de m-santé regroupées en 5 patterns (le bien-être/beauté, l'autonomie, le sport, les pathologies et les urgences).

2.4. Les avantages et les freins de la e-santé

Eysenbach (2001) met en avant dix avantages de la e-santé à travers ce qu'il appelle « *the 10^e's in -e-health* ». Ainsi, la e-santé révolutionnerait les comportements en santé, car cela permettrait d'éduquer, de faire circuler des informations médicales de manière plus efficiente, d'être plus éthique et d'être plus équitable dans la répartition des soins de santé sur le territoire. Cela améliorerait même la relation entre le patient et les professionnels du secteur. Schnall et al. (2016) mettent en avant la capacité de ces objets à impliquer le patient dans son traitement. La e-santé permettrait alors de rendre un patient plus autonome. Grâce à ce gain d'autonomie, le soin deviendrait alors plus simple et surtout plus efficace permettant au système de santé de se recentrer sur le patient, ce qui développerait sa réactivité et réduirait ses coûts (Laubie, 2011). Laubie (2011) met aussi en avant l'intérêt certain pour les proches de pouvoir être informés de l'état de santé de chacun. Enfin, Simon (2017) met surtout en avant l'intérêt de la e-santé pour les pays en développement.

Pourtant, les promesses de la e-santé ne semblent pas encore tenues. De nombreux freins à l'utilisation de ces outils sont aussi recensés : financiers, structurels, culturels, organisationnels, techniques ou professionnels. Simon (2017) réalise un état de l'art sur la fiabilité et la sécurité des objets connectés en santé. L'auteur conclut que la majorité des études sur les dispositifs médicaux connectés et la télémédecine ne rend actuellement pas service aux patients. De plus, ces dispositifs n'ont pas prouvé leur rentabilité économique et clinique principalement à cause des difficultés d'organisation professionnelle et du manque de coopération entre les différents acteurs-métiers (industries, hôpitaux, particuliers). Ces

manquements altèrent la sécurité et la fiabilité des patients utilisateurs. Du point de vue des particuliers, la sécurité des données ainsi que l'atteinte à la vie privée est un frein majeur à son utilisation (e.g. Allaert, 2016; Béranger, 2014; Laubie, 2011). La crainte se trouve dans leur diffusion publique ou leurs ventes vers des assureurs qui pourraient alors refuser leurs services à des patients souffrant de maladie chronique (Hall, 2014). L'accès à ces technologies est également un frein important. En effet, Greysen et al. (2014) mettent en avant l'incapacité de certains utilisateurs, surtout chez les personnes âgées, à se munir d'internet et d'un ordinateur. Les auteurs mettent également en avant le manque de compétences des utilisateurs. Les personnes âgées sont souvent au cœur de nombreuses études sur le sujet et montrent avant tout des freins psychologiques à l'utilisation de ce matériel comme le sentiment d'être trop vieux pour « s'y mettre » (Chen & Chan, 2011). Enfin, la mauvaise conception de ces matériaux est particulièrement pointée du doigt. **Ils sont jugés trop complexes à utiliser renvoyant directement à un manque d'utilisabilité** (e.g., Carayon & Hoonakker, 2019 ; Cifter, 2017; Greysen, Garcia, Sudore, Cenzer, & Covinsky, 2014 ; Huvila et al., 2016 ; Kortum & Peres, 2015). Ceci a un impact direct sur la sécurité du patient (Nielsen, 2005).

Gerdes et al. (2014, p. 1) déclarent que « *the most important requirement on such a collaborative eHealth system should be the usability of the system for all involved user groups* ». De plus, l'acceptation ou le refus définitif des technologies en santé dépendent largement du degré d'utilisabilité du système (Kushniruk, 2002). Une meilleure utilisabilité de ces dispositifs permettrait une productivité accrue, une réduction des erreurs, des formations et une assistance aux utilisateurs allégées et une meilleure acceptation générale pour les utilisateurs (Jaspers, 2009). Nous essayons de comprendre d'où provient cette complexité d'utilisation dans la section suivante.

2.5. Les principales difficultés de l'évaluation de l'utilisabilité de dispositifs connectés en santé

2.5.1. L'origine des défauts d'utilisabilité des dispositifs connectés en santé utilisables par le grand public

Après la révolution d'internet, les sociologues s'accordent à dire que la prochaine révolution viendra des objets connectés (Allaert et al., 2016). Leur nombre explose de manière exponentielle surtout dans le domaine de la santé. On peut compter aujourd'hui environ 500 000 types différents de dispositifs médicaux disponibles pour le grand public (WHO, 2016). La

plupart de ces objets connectés certifiés se retrouvent au domicile du patient sous la gestion du patient lui-même. En effet, nous passons d'un système où les frontières de l'hôpital s'étendent finalement jusqu'au domicile du patient. Ten Haken et al. (2018, p.1) décrivent ce passage de « *l'inpatient to outpatient care* ». Historiquement, ces technologies ont d'abord été conçues pour être utilisées à l'hôpital par le personnel hospitalier formé (Kaufman-Rivi et al., 2010; Weick-Brady & Lazerow, 2006). Ceci implique inévitablement des risques lorsqu'ils passent entre les mains d'utilisateurs partiellement formés. C'est alors aux concepteurs de prendre en compte ce manque de formation et d'adapter (et donc de concevoir et d'évaluer) les dispositifs à cette population.

Dans cette perspective, la norme (ISO/IEC 62366-1:2015) impose de respecter des normes d'utilisabilité et l'intégration des facteurs humains dans les processus de conception. Pourtant, les évaluations de l'utilisabilité - essentielles dans la commercialisation d'un dispositif en santé - continuent de subir certaines lacunes. En effet, depuis les vingt dernières années, nombreux sont les auteurs qui dénoncent le manque probant d'utilisabilité en ce qui concerne les dispositifs médicaux connectés (e.g., Carayon & Hoonakker, 2019 ; Cifter, 2017 ; Greysen et al., 2014 ; Horsky et al., 2005 ; Huvila et al., 2016 ; Jaspers, 2009 ; Kim & Xie, 2017 ; Pelayo et al., 2012 ; Peute et al., 2013 ; Peute et al., 2008).

Il existe actuellement deux arguments justifiant le fait que certains appareils jugés *a priori* utilisables et sûrs entraînent des erreurs d'utilisation inattendues une fois implantés dans le terrain : (1) **cela est probablement dû à l'absence d'un cadre et d'une méthode normalisée dans les études d'utilisabilité** (e.g., Brown et al., 2013 ; Georgsson & Staggers, 2016 ; Jaspers, 2009 ; Nelson et al., 2017 ; Peute et al., 2013 ; Peute et al., 2008 ; Yen, 2010 ; Yen & Bakken, 2012) et/ou (2) **à l'absence de connaissances des effets du contexte d'usage** sur l'utilisabilité (e.g., Borsci et al., 2018 ; Georgsson & Staggers, 2016 ; Grebin et al., 2018 ; Kaufman-Rivi et al., 2010). Nous développons ces deux arguments dans les sections suivantes.

2.5.2. Les trois principaux défis des évaluations de l'utilisabilité des technologies en santé

De nombreux défis afin d'améliorer l'utilisabilité de ces systèmes attendent toujours d'être résolus. Ammenwerth et al. (2003) synthétisent l'ensemble de ces problèmes en trois défis majeurs pour lesquels ils présentent les conséquences et les solutions pour chacun d'eux : (1) « la complexité de l'objet de l'évaluation » ; (2) « la complexité d'un projet d'évaluation » et (3) « la motivation de l'évaluation ». Nous détaillons ces trois défis.

Premièrement, si « l'objet » (dispositif) fait partie d'une organisation, cela signifie qu'il s'agit davantage d'une situation ou d'un processus qui doit être évalué que d'un produit unique. Pelayo et al. (2012) expliquent que l'utilisabilité d'un système dépend en effet du système lui-même, de son utilisation et de son intégration dans le contexte d'usage. Ainsi « *l'évaluation requiert non seulement une compréhension de la technologie, mais aussi des processus sociaux et comportementaux qui affectent et sont affectés par la technologie*¹⁶ » (Ammenwerth et al., 2003, p. 127).

Deuxièmement, les évaluations de l'utilisabilité de technologies en santé intègrent de nombreuses parties prenantes (e.g., médecins, personnel infirmier, patients, administration, chercheurs, industriels, organismes de financement...) et sont contraintes par des influences externes (e.g., politique de l'hôpital, législation, économie, patientèle...). Comme nous avons pu le voir précédemment (cf. Introduction générale, 2.2. Un projet de co-développement : les partenaires industriels, académiques et hospitaliers, p. 4), les projets de co-conception et multipartenariaux peuvent être source de conflits liés aux points de vue divergents sur une technologie dite « réussie » (Heathfield et al., 1999).

Troisièmement, les évaluations ergonomiques sont contraintes, entre autres, par des moyens économiques et humains (e.g., le nombre de patients recrutés). Ces deux facteurs dépendent de la motivation des parties prenantes à proposer des ressources suffisantes pour mettre en place l'évaluation. Ammenwerth et al. (2002) ont montré que la direction d'un hôpital peut parfois être réfractaire à ces évaluations par crainte de révéler des résultats négatifs. Wyatt et Spiegelhalter (1991) mettent en avant la difficulté de trouver des participants motivés (surtout lorsqu'il s'agit de personnel hospitalier). Ces participants ne sont pas rémunérés (contrairement aux essais cliniques) pour leurs efforts pouvant engendrer un temps de travail plus important ou des tâches supplémentaires comme la complétion d'un questionnaire (Ammenwerth et al., 2003).

Afin de faire face à ces trois défis, il semblerait qu'un cadre méthodologique soit nécessaire afin de standardiser ces évaluations. Rahimi et Vimarlund (2007) ont recensé les méthodes générales utilisées pour évaluer les technologies en santé ; Kushniruk et Patel (2004) présentent quant à eux une analyse méthodologique utilisée dans divers contextes (e.g., laboratoire, *in situ*). Malgré ces premiers essais de cadres et d'analyses systématiques,

¹⁶ Traduit de l'anglais : “*evaluation requires not only an understanding of computer technology, but also of the social and behavioral processes that affect and are affected by the technology*” (Ammenwerth et al., 2003, p. 127)

Kjeldskov et al. (2004) considèrent qu'il n'existe pas encore de méthodes d'évaluation de l'utilisabilité ni de techniques de recueil de données appropriées dans le domaine des IHM mobiles. De nombreux chercheurs ont alors investi ce champ pour proposer des solutions pour remédier à ces trois challenges.

2.5.3. Les cadres et modèles d'évaluation de l'utilisabilité

Depuis ces premières propositions, de nouveaux cadres d'évaluation de l'utilisabilité ont émergé. Deux laboratoires de recherche (EVALAB à Lille et HITLAB à Amsterdam) proposent un cadre d'évaluation des Technologies Informatiques en Santé ou TIS¹⁷ (Peute et al., 2013) basé sur les standards de conception centrée utilisateur, d'utilisabilité, de qualité, et d'évaluations de technologie de l'information (ISO 9241-11, 9241-210, 13407, 14155, ISO/IEC 14598-5, 9126-2,60601-1-6 et les normes issues de l'ANSI/HFES). Dans ce cadre, Peute et al. (2013) recensent les éléments nécessaires à prendre en compte lors d'une évaluation et à mentionner pour la rédaction d'un article scientifique (e.g., le type de système et ses fonctionnalités, les caractéristiques utilisateurs, les métriques utilisées...). Dans la même dynamique, Eysenbach et Group (2011) proposent le CONSORT-HEALTH. Ce cadre prend la forme d'une liste de points à contrôler (« *checklist* »). Lorsqu'un chercheur réalise une évaluation de l'utilisabilité d'une technologie en santé, il doit impérativement remplir l'ensemble des critères listés s'il veut pouvoir publier son étude dans certaines revues à comité de lecture comme le Journal of Medical Informatics Research ou JMIR. Pour cela, Eysenbach et Group (2011) ont mis en place un questionnaire en ligne¹⁸ (« google form ») retraçant l'ensemble des éléments de cette liste de contrôle. Le chercheur doit compléter la liste en remplissant le questionnaire en ligne. Nous avons réalisé cette opération lors de notre publication dans JMIR (Chaniaud et al., 2020) pour l'étude 1 (cf. Partie II, Chapitre 2). L'ensemble des éléments de la liste doit être référencé dans le questionnaire en ligne. Le recours à ce système permet de créer une base de données permettant d'énumérer les éléments composant les articles scientifiques sur l'évaluation de l'utilisabilité ou de faisabilité. Zapata et al. (2015) proposent par ailleurs, un catalogue (USB-CAT) d'applications de m-santé respectant

¹⁷ En anglais : Health Informatics Technology ou HIT

¹⁸ lien vers le questionnaire en ligne :

http://docs.google.com/forms/d/e/1FAIpQLSfZBSUp1bwOc_OimqcS64RdfIAFvvrTSkZQL2-3O8O9hrL5Sw/viewform?hl=en_US&formkey=dG1Kd2Z2Q11NSGQ0THI1azM5MS1aWWc6MA&rm=full#gid=0

la norme ISO /IEC/IEEE 29148 :2011. Parmi 133 applications auditées, 97 (77%) applications ont été jugées comme utilisables (Zapata et al., 2018).

Puis, de nombreuses autres propositions de cadres évaluatifs de l'utilisabilité dans la littérature scientifique ont émergé en se spécialisant dans des domaines précis. Par exemple, Zapata et al. (2015) proposent un cadre dans le domaine de la m-santé. Cette spécialisation par domaine est importante. En effet, les problématiques de standardisation des évaluations de l'utilisabilité pour les dispositifs en m-santé sont relativement similaires aux problématiques des systèmes ubiquitaires (Hammontree et al., 1994) et questionnent les chercheurs et concepteurs. Kaufman et al. (2003) déclaraient que peu de recherches référencées dans la littérature se sont intéressées à l'utilisabilité des technologies de soins de santé prodigués au domicile du patient, impliquant des évaluations à distance. Les auteurs élaborent alors un cadre évaluatif pour le domicile patient. Ce cadre est centré sur les qualités de l'interface et sur les aptitudes et les compétences des patients. Pour cela, les auteurs se servent de la méthode *walkthrough* permettant de collecter des données audio et vidéo (cf. Partie I, Chapitre 1, 4.3.3. Le choix des bonnes méthodes et des bons outils, p. 48). Ce cadre a été testé sur un système de télémédecine à domicile pour la population diabétique.

S'en suivent de nombreux autres cadres d'évaluation de l'utilisabilité de plus en plus spécialisés et précis (e.g., Aljaber & Gordon, 2017 ; Beatty et al., 2013 ; Reolon et al., 2016 ; Sadegh et al., 2018 ; Schmidt & De Marchi, 2017 ; Schnall et al., 2016 ; Wildenbos et al., 2015, 2018). Par exemple, Schnall, Rojas, et al. (2016a) proposent un modèle spécialement pour le recours à des tests utilisateurs dans le cadre d'une conception d'applications de m-santé. Ce modèle est basé sur une conception en trois cycles : *pertinence*, *rigueur* et *conception*. Le cycle de la *pertinence* pourrait s'apparenter à une analyse de besoins. Le cycle de *rigueur* se rapproche d'un benchmark. Enfin, le cycle de *conception* présente différentes itérations d'évaluations de l'utilisabilité de dispositif de m-santé. Dans la même lignée, Reolon et al. (2016) ont conçu un cadre pour l'évaluation de l'utilisabilité par méthodes heuristiques toujours afin d'assister la conception d'applications de m-santé. Pour cela, les auteurs ont réalisé une analyse systématique des heuristiques spécialement adaptées aux technologies en santé. Les auteurs n'ont référencé que deux articles sur le sujet. Le domaine des heuristiques des dispositifs connectés en santé reste sous-développé, pourtant, cela pourrait contribuer à l'adoption de techniques d'évaluations dites « *légères* » (Reolon et al., 2016). Wildenbos et al. (2015) ont conçu un cadre afin de soutenir l'évaluation de l'utilisabilité des dispositifs pour les personnes âgées. Puis, les mêmes auteurs (Wildenbos et al., 2018), améliorent ce cadre

(MOLD-US) en référençant dans la littérature les principaux freins par rapport à l'utilisabilité de ces outils de m-santé par les personnes âgées. Quatre catégories clés d'obstacles, liées au vieillissement, impactant l'utilisation de la m-santé, ont été identifiées : la cognition, la motivation, la capacité physique et la perception. Ce cadre a la spécificité de permettre une classification et une interprétation des données observées lors d'évaluation en fonction des obstacles intrinsèques aux problèmes d'utilisabilité rencontrés par les personnes âgées. Sadegh et al. (2018) proposent une revue de la littérature des différents cadres proposés en m-santé en tenant compte des avantages évoqués par les différentes parties prenantes. Certains cadres et modèles dans des domaines de plus en plus précis voient le jour (e.g., Aljaber & Gordon, 2017; Beatty et al., 2013; Schmidt & De Marchi, 2017). C'est le cas du modèle de Beatty et al. (2013), permettant de soutenir les conceptions et l'évaluation d'applications mobiles spécialisées dans la réadaptation cardiaque et Schmidt et De Marchi (2017) dans les « serious games ».

Au vu de la multitude de cadres proposés, il devient difficile et peu clair de se positionner et d'en privilégier un en particulier. De plus, aucun modèle ne propose de cibler les éléments du contexte à prendre en compte (e.g., Wildenbos et al., 2018) tout en assistant la conception par des évaluations itératives (e.g., Schnall, Rojas, et al., 2016a).

3. État de l'art des cadres méthodologiques de l'évaluation de l'utilisabilité de dispositif en santé pour le domicile patient

Il est nécessaire d'actualiser et de développer des connaissances fournies par ces études antérieures en prenant en compte plus explicitement les éléments qui orbitent autour des problématiques de l'utilisabilité pour les appliquer aux dispositifs de santé ubiquitaires. C'est pourquoi dans un premier temps, en nous basant sur le cadre de Peute et al. (2013) et sur les évaluations des systèmes ubiquitaires (Jambon et al., 2010), nous nous sommes demandées (1) dans quels **environnements** sont évalués ces dispositifs ; (2) quel(s) **méthode(s) et outil(s)** sont utilisés, pourquoi, quels dispositif(s) et pour quelles **tâches** ; (3) quelles **métriques** de l'utilisabilité sont prises en compte et enfin (4) quelles **populations** sont mobilisées dans les études actuelles ?

Dans un deuxième temps, nous avons voulu vérifier si les recommandations portant sur l'évaluation de l'utilisabilité de dispositif en santé étaient toujours suivies (e.g., le recours à une

méthodologie mixte associant un recueil de données objectifs *versus* subjectifs et qualitatifs *versus* quantitatifs).

Après analyse des études portant sur le sujet des dispositifs en santé, nous proposons une discussion confrontant nos résultats à d'autres analyses systématiques plus anciennes.

3.1. La classification des études portant sur l'utilisabilité des dispositifs en santé pour le domicile patient

Dans cette section, nous présentons une classification des études **portant exclusivement sur l'évaluation de l'utilisabilité de dispositifs connectés en santé pour le domicile patient**. Pour cela, nous avons réalisé une recherche bibliographique afin de référencer ce type d'études, publiées entre 2010 à aujourd'hui (2020), en fonction des objectifs de recherche et des méthodes de recherche appliquées. Pour ce faire, nous avons procédé à une lecture approfondie des résumés et des aspects méthodologiques. Nous avons sélectionné ces articles en fonction des critères d'inclusion et d'exclusion suivants :

- Les études incluses devaient proposer une évaluation de l'utilisabilité d'une application ou d'un outil médical (connecté ou non connecté) ;
- Le dispositif évalué devait avoir pour finalité d'être utilisé au domicile du patient/utilisateur ;
- Le dispositif devait être *a minima* sous une forme de prototype (i.e., à un niveau d'avancement suffisant dans le cycle de conception du produit) ou déjà commercialisé ;
- Les articles scientifiques sans comité de lecture par les pairs et les actes de conférences ont été rejetés. Nous avons aussi exclu tous les articles ayant recours à des produits/dispositifs n'ayant pas d'objectif médical/paramédical et/ou conçus pour le personnel hospitalier ;
- Nous avons rejeté les articles dans lesquels les auteurs n'abordaient pas la notion d'utilisabilité.

Afin de rechercher ces articles, nous avons utilisé les moteurs de recherche PubMed et Google Scholar avec les mots clés suivants : *HMDs, usability test, usability assess, usability evaluation, usability study, health informatics technology, patient's home, mhealth, uhealth*. Nous avons pu identifier 125 articles scientifiques, sur la base de leur titre, portant sur l'utilisabilité d'une application ou outil médical pour le domicile patient. Parmi ces 125 articles, nous en avons exclu 42, après lecture des résumés car ils ne correspondaient pas aux critères

précités. Puis, une étude a été écartée du fait de l'incompatibilité entre le résumé et le contenu de l'article et 3 autres de ces articles étaient des doublons. Plus précisément, leurs auteurs utilisaient la même population avec le même matériel, mais présentaient des éléments différents de leur évaluation de l'utilisabilité. Finalement, nous avons retenu 79 articles scientifiques classifiés et détaillés en Annexe (cf. Annexe 4 : Tableau d'analyse systématique des méthodes et outils utilisés dans l'évaluation de l'utilisabilité de dispositif en santé, p. 344). La Figure 15 ci-dessous retrace cette phase de sélection.

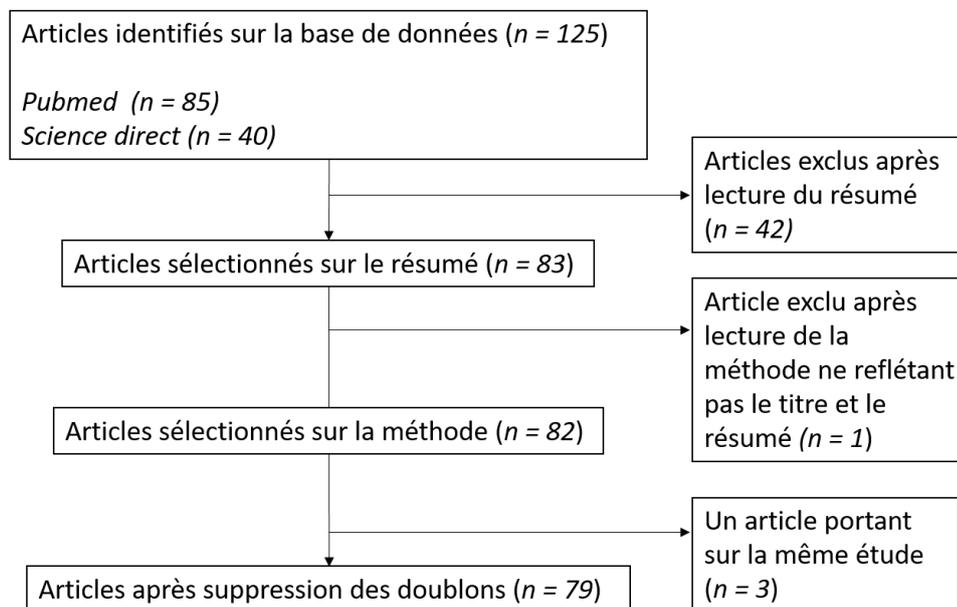


Figure 15. Méthode de sélection des articles de revue portant sur des évaluations de l'utilisabilité de dispositif en santé.

Les articles ont ensuite été classés en fonction du matériel, des méthodes, de l'environnement, des métriques, et des outils utilisés. Le Tableau 7 présente les catégories et critères de données extraites des articles. Ces critères sont basés sur des éléments déjà référencés dans notre partie théorique (cf. Partie I : Partie théorique). C'est pourquoi dans la dernière colonne du Tableau 7, nous proposons des renvois vers des emplacements spécifiques où nous avons défini ces critères.

Tableau 7. Catégorisation des études en fonction des critères présentés (détails en Annexe 4, p. 344)

Catégories	Critères	(cf. Partie, Chapitre, section, page)
Auteurs	(Nom de l'ensemble des auteurs de la publication)	-
Année	(Année de publication)	-
Revue	(Nom de la revue)	-
Objet de l'évaluation	<ol style="list-style-type: none"> 1. Nom de de/des Applications mobiles 2. Nom de la plate-forme 3. Nom de l'objet connecté 	(cf. Partie I, Chapitre 2, 2.3. Les systèmes d'information de santé et la télésanté, p. 59)
Méthodes	Test utilisateurs, heuristiques, enquêtes...	(cf. Partie I, Chapitre 1, 4.3.2. Trois types de méthodes d'évaluation de l'utilisabilité : inspection, test utilisateur, enquête, p. 46)
Environnement	<ol style="list-style-type: none"> 1. Laboratoire Expérimental (sans mise en situation) Quasi-expérimental (avec mise en situation, ex. simulation) 2. Terrain 3. <i>In situ</i> 4. <i>In situ rétrospectif</i> 	(cf. Partie I, Chapitre 1, 4.2. Trois lieux de recueil de données déterminés par trois approches épistémologiques : laboratoire, terrain, <i>in situ</i> , p. 43)
Métriques	Efficiences, efficacité, satisfaction...	(cf. Partie I, Chapitre 1, 3. Les approches conceptuelles de l'utilisabilité : les mesures des métriques de l'utilisabilité, p. 26)
Quantitative/Qualitative	Quantitative (ex. questionnaire, log file analysis) Qualitative (ex. entretien, focus group)	(cf. Partie I, Chapitre 1, 4.3.3. Le choix des bonnes méthodes et des bons outils, p. 48)
Subjectif/Objectif	Objectif (ex. observation, codage vidéo) Subjectif (ex. entretien, focus group, questionnaire)	
Outils utilisés	(ex. entretien, focus group, observation, questionnaire...)	
Population	Taille de la population et caractéristique (ex. jeune vs âgé, maladie chronique...)	(cf. Partie I, Chapitre 1, 3.4.3. Caractéristiques utilisateurs, p. 40)
Analyse de données	Description/ corrélation / Analyse de Variance / Régression	-
Résultats	Principaux résultats	-

Dans le cadre spécifique de l'évaluation de l'utilisabilité des dispositifs en santé à destination du domicile patient, nous avons classé les environnements dans lesquels ces évaluations ont été conduites (cf. Partie I, Chapitre 1, 4.2. Trois lieux de recueil de données déterminés par trois approches épistémologiques : laboratoire, terrain, *in situ*, p. 43).

- Le **laboratoire** fait référence à un contexte simulé ou non écologique, fait intervenir des évaluateurs et utilise du matériel d'observation auprès de l'utilisateur (e.g., l'hôpital, l'université...);

- Le **terrain** renvoie au contexte naturel de l'utilisateur (e.g., son domicile) au sein d'un scénario naturel et non provoqué par l'expérimentateur, mais où les méthodes utilisées rappellent constamment à l'utilisateur qu'une évaluation est en cours (e.g., l'évaluateur est présent, les consignes sont imposées, le participant est filmé, et/ou doit verbaliser ses actions...);

- La situation *in situ* correspond au contexte naturel de l'utilisateur impliquant un scénario naturel et non provoqué par l'expérimentateur. Cette situation est considérée comme « *in situ* » si les évaluateurs utilisent des méthodes n'interférant que très peu avec l'utilisateur (p. ex. l'expérimentateur n'est pas présent, l'utilisateur est libre de manipuler le matériel quand il le désire, il n'est pas filmé, n'est pas interrogé pendant l'utilisation de la technologie ...);

- Par ailleurs, il existe la situation *in situ* **rétrospective** lorsque le participant utilise le matériel librement dans un environnement naturel et est évalué rétrospectivement à cette utilisation. Cet ajustement rétrospectif ne permet pas à l'expérimentateur d'obtenir des informations directes, ce qui crée un biais rétrospectif sur les données aussi appelé le *syndrome de la sagesse après-coup* (Reason, 2013). Ce biais entraîne chez l'utilisateur une erreur de jugement liée à une rationalisation de l'évènement. Ce dernier peut atténuer les risques liés à la manipulation du matériel. Ce biais est particulièrement recensé lors d'entretiens où les participants ont tendance à considérer le matériel facile d'utilisation après-coup alors que de nombreuses erreurs de manipulation ont été détectées. Nous avons pu observer ce biais dans nos propres études lors des analyses des entretiens (cf. Partie II, Chapitre 4, 5.1. Analyse quantitative, p. 208 et Chapitre 5, 5.2. Analyse qualitative, p. 252).

3.2. Les résultats des études empiriques sur l'utilisabilité de dispositifs connectés en santé pour le domicile patient

3.2.1. Les objets de l'évaluation

Chacune des études avait pour objectif l'évaluation de l'utilisabilité d'un matériel médical que nous avons classifié en trois catégories : (1) les logiciels comprenant les applications (comme l'application *Smart Angel*) sur smartphone ou tablette fonctionnant grâce à des systèmes d'exploitation (IOS/Android/Windows) et des plateformes sur internet telles que les systèmes de télémédecine ou les dossiers patients en ligne ; (2) des objets connectés permettant une automesure *via* des capteurs se rattachant à une application telle que des montres, bracelets, balances, ou tous autres appareils; (3) les objets non connectés pouvant être

utilisés par les patients dans un but médical permettant une automesure sans avoir de lien avec une connexion internet (tels que des tensiomètres ou glucomètres non connectés).

Nous avons pu référencer 62 (78,48%) études portant sur l'utilisabilité de logiciels, 10 (12,66%) traitant des objets connectés et 7 (8,86%) portant sur des objets non connectés.

3.2.2. Les revues

Nous avons pu observer une importante diversité des revues scientifiques dans lesquelles étaient publiées les 79 articles de notre analyse. En effet, nous avons compté 36 revues différentes, 12 étant des revues de médecine clinique, 11 relevant de l'informatique médicale, 5 étant des revues d'ergonomie, 4 référencées comme des revues spécialisées dans les technologies (comprenant celles spécialisées dans la télémédecine comme « Telemedicine and e-health »), et enfin 4 revues issues d'autres domaines (e.g., éducation du patient).

Cette analyse suggère que la plupart des auteurs publient soit dans une revue spécifique à l'application du dispositif testé (i.e., l'évaluation d'un dispositif cardiaque sera publié dans « International Journal for Hemoglobin Research » ou encore une application pour gérer la prise de poids sera publiée dans « Nutrients »), soit dans une revue d'informatique médicale telle que JMIR (Journal of Medical Informatics Research) référençant à elle seule 30 articles (36% des études référencées) dans 8 sections différentes. Les revues en ergonomie (« International Journal of Mobile Human computer Interaction »; « Journal of Design Research »; « Ergonomics »; « Applied Ergonomics » et « Avances en Interaccion Humano-computadore ») ne sont visiblement pas le premier choix des chercheurs dans le domaine.

3.3. Les résultats des méthodes utilisées dans les évaluations de l'utilisabilité

Nous avons classifié les méthodes des évaluations de l'utilisabilité selon quatre variables : (1) l'environnement (e.g., Laboratoire, terrain, *in situ*...), (2) les méthodes (p. ex. tests utilisateurs, évaluations heuristiques...), (3) les métriques (e.g., efficacité, efficience, satisfaction...) et (4) les outils méthodologiques (e.g., questionnaire, focus group...). En fonction des outils et des résultats présentés, nous avons également défini s'il s'agissait d'études qualitatives et/ou quantitatives ainsi que subjectives et /ou objectives.

3.3.1. Les environnements d'évaluation

Nos résultats ont montré que 66 (69,47%) études étaient réalisées en laboratoire, 7 (7,37%) sur le terrain, 22 (23,17%) *in situ* dont 9 que nous considérons comme *in*

situ rétrospectif. Certains articles pouvaient aborder plusieurs études et comptabiliser plusieurs environnements d'évaluations. Nous avons recensé 9 articles proposant une double étude, c'est à dire une étude en laboratoire (100% des cas) et une autre *in situ* (7 cas) et/ou de terrain (2 cas).

Le premier constat en vue de ce résultat montre qu'environ les $\frac{3}{4}$ des études d'utilisabilité sont réalisées en laboratoire. L'environnement naturel prend la deuxième place. Enfin, les méthodes dites de terrain semblent sous-exploitées.

a. Les freins aux études de terrain et *in situ*

Il est évident que les complexités méthodologiques et surtout législatives des études de terrain ou *in situ* freinent leur élaboration. Les Comités de Protection des Personnes ou CPP (e.g., Peute et al., 2020), les Comités Ethiques et de Recherche ou CER, la Commission Nationale de l'Informatique et des Libertés ou CNIL sont autant de barrières connues pour la réalisation des études *in situ*, en particulier en France. S'ajoutent à ces difficultés administratives, les difficultés méthodologiques. En effet, comment faire en sorte que les patients ne se sentent pas observés et utilisent le matériel sans la moindre contrainte tout en renseignant les chercheurs ? Nous proposons une liste des outils méthodologiques utilisés dans les différents environnements cités dans la section ci-dessous (cf. Partie I, Chapitre 2, 3.3.3. Les outils, p. 75).

b. Le laboratoire est-il vraiment une bonne alternative aux environnements de terrain et *in situ* ?

Le terrain a-t-il autant de valeur que l'on voudrait lui attribuer ? Notre analyse précédente ainsi que les revues de la littérature sur les systèmes mobiles montrent que leurs évaluations se font majoritairement en laboratoire (Kjeldskov & Graham, 2003) avec 66,47% des études en laboratoire contre 23,17% *in situ*. Pourtant, nombreux sont les auteurs à plaider pour une recherche de terrain (e.g., Abowd & Mynatt, 2000; Brewster, 2002), puisque ces systèmes sont dépendants du contexte d'usage (e.g., ISO 9241-11 : 2018). Alors pourquoi il y a-t-il encore autant d'évaluations de l'utilisabilité en laboratoire ?

Premièrement, les études d'utilisabilité sur le terrain sont très coûteuses et extrêmement difficiles à réaliser. Cela relève du « *défi méthodologique* ¹⁹ » (Kjeldskov et al., 2004, p. 61). Il

¹⁹ Traduit de l'anglais : « *methodological challenge* »

est compliqué de capter l'usage sans pouvoir créer un biais par la présence de l'expérimentateur ou d'une caméra.

Deuxièmement, la valeur ajoutée du terrain reste méconnue dans le contexte médical (e.g., Kjeldskov & Stage, 2004; Nielsen, 1998; Rubin, 1994). Kjeldskov et Stage (2004) ont souhaité mesurer l'apport du terrain pour identifier et traiter les problèmes d'utilisabilité. Précisément, ils ont évalué l'utilisabilité d'un système mobile (le dossier patient électronique) en laboratoire avec 6 sujets, et sur le terrain avec 6 autres sujets. L'étude en laboratoire leur a permis d'identifier la même liste de problèmes d'utilisabilité comparativement au terrain. Les auteurs ont réitéré leur expérience avec un guide mobile pour un réseau de transport et arrivent aux mêmes constats (Kjeldskov et al., 2005).

Pour ces raisons, les auteurs ont conclu que le terrain apportait une très faible valeur ajoutée des évaluations de l'utilisabilité. Ils suggèrent de remplacer les études de terrain par des études en laboratoire utilisant des caractéristiques contextuelles telles que des scénarios et des simulations de contexte et ainsi conserver les avantages d'un environnement contrôlé tout en reproduisant de façon partielle des caractéristiques du contexte réel d'utilisation (Kjeldskov & Skov, 2003; Kjeldskov & Stage, 2004).

Cependant, cette absence de plus-value ne fait pas consensus. Cela provient des difficultés de reproductibilité de ces études (Jambon et al. 2010). De plus, ces études réalisées uniquement en laboratoire sont pertinentes si elles acquièrent une « validité écologique » permettant de définir qu'un comportement observé en laboratoire est généralisable dans un environnement naturel (Rochat et al., 2017; Schmuckler, 2001).

De notre point de vue, **ces deux approches sont complémentaires**. En effet, le laboratoire permet d'obtenir des informations très fines, mais moins fiables du fait notamment des nombreux biais générés. La situation « *in situ* » permet d'obtenir d'autres informations comme les comportements inattendus et inobservables en laboratoire (Jambon et al. , 2010).

3.3.2. Les méthodes

Pour rappel, toutes les études recensées impliquaient la participation d'utilisateurs finaux. Nous avons classifié ces études en fonction de trois types de méthodes (cf. Partie I, Chapitre 1, 4.3.2. Trois types de méthodes d'évaluation de l'utilisabilité : inspection, test utilisateur, enquête, p. 46) dans le Tableau 8 : une utilisation avec contrainte (i.e., tests utilisateurs), une utilisation libre sans consigne, sans contrainte (i.e., enquêtes) et une utilisation par inspection (i.e., évaluations heuristiques).

Le recours aux méthodes d'utilisation avec contrainte (avec consignes imposées par l'expérimentateur – test utilisateur) est majoritairement utilisé dans le cas d'un environnement contrôlé comme le laboratoire (89,29%). Quelques exceptions existent. à titre d'exemple, nous avons recensé une étude (e.g., Jeon & Park, 2018) utilisant des tests utilisateurs en milieu naturel. Si cela semble paradoxal, Jeon et Park (2018) ont trouvé une stratégie afin d'orienter le patient dans son utilisation par le biais de notifications sans interférer dans l'utilisation du dispositif.

L'utilisation libre (sans consignes imposées par l'expérimentateur – enquête) est prioritairement mise en oeuvre dans le cas du milieu *in situ* (66,67%).

L'évaluation par inspection est généralement utilisée en laboratoire (80%). Lorsque celle-ci est mise en oeuvre en milieu naturel ou sur le terrain, cela signifie qu'une première série de tests utilisateurs a été réalisée préalablement avant d'être mise en milieu naturel.

Tableau 8. Nombre d'études en fonction des méthodes utilisées et de l'environnement.

	Laboratoire	Terrain	<i>In situ</i>	Total
Utilisation avec contrainte (avec consignes) – tests utilisateurs	50	5	1	56
Utilisation libre (sans consignes) – enquêtes	8	2	20	30
Utilisation par inspection – Évaluations heuristiques	8	1	1	10
Total	66	7	22	96

3.3.3. Les outils

Chacune des sections méthodologiques des études recensées a été minutieusement inspectée afin d'identifier les outils utilisés. Ces outils ont été classifiés en deux groupes dans le Tableau 9 : les outils permettant la collecte de données subjectives ou de données objectives. Au total, 209 outils ont été utilisés sur l'analyse de 95 études répertoriées dans les 79 articles. Dit autrement, 2,2 outils en moyenne par étude ont été relevés suggérant une certaine richesse méthodologique.

a. Le choix d'outils de collecte de données subjectives *versus* objectives

Nous pouvons observer que le recours à des outils apportant des données subjectives (146) est deux fois plus utilisé que ceux apportant des données objectives (63). Les trois outils les plus utilisés permettant d'obtenir des données subjectives sont les questionnaires²⁰, les

²⁰ Le questionnaire est l'outil utilisé de manière massive dans les études d'utilisabilité. Plus d'une étude sur deux en a recours. Le questionnaire le plus répandu est le System Usability Scale ou SUS de Brooke (1995) utilisé dans 20 études. S'en suit un défilé d'acronymes : PUEU (1), QoE (1), ASQ (2), Health-ITUES (1), UEQ (1), RTLX (1), NASA (2), Acceptability e-Scale (1), PSHUQ (1), CSUQ (1), QUIS (1), Attrackdiff (2), PSSUQ (4), TAM (3), MARS (1) (et uMARS(1)), USE (2), RIP (2).

entretiens (face-à-face) et la méthode *think aloud*. Dans une moindre mesure, l'observation instrumentée par vidéo et l'observation directe restent également fortement représentées. Constantinescu et al. (2019) expliquent que parmi toutes les méthodes utilisées (observation, enregistrement vidéo, *think aloud*, entretien, questionnaire), l'observation vidéo (captures d'écran synchronisées avec des gestes vidéo) est la méthode qui a permis de recueillir le plus grand nombre d'informations.

Tableau 9. Nombre d'études en fonction des outils utilisés.

	Outils	Laboratoire	Terrain	In situ	Total
Subjectif	Entretiens focus group	5	1	3	9
	Entretien/ <i>think aloud</i> /visionnage rétrospectif	5			5
	Entretiens face à face	25	4	8	37
	Entretien téléphonique	1	1	2	4
	Tri de carte	1			1
	Cognitive walkthrough	3			3
	<i>Think aloud</i>	29	3		32
	Journal de bord	1		1	2
	Questionnaire	42	2	12	56
	Total	109	11	26	146
Objectif	Observation instrumentée par vidéo	24	3		27
	Observation instrumentée par photo	2			2
	Observation directe	20	4	1	25
	Time motion studies (TMS)			1	1
	Eye tracking	1			1
	Hierarchisation de la tâche	1			1
	Report d'erreur	1		1	2
	Traceur d'activité			3	3
	Test compréhension	1			1
	Chronomètre	1			1
	Total	51	7	5	64

Pour le choix d'outils de collecte de données subjectives, nous n'observons pas de différence entre le laboratoire, le terrain et la situation *in situ*. Le questionnaire et l'entretien sont privilégiés pour ces trois environnements.

En ce qui concerne les outils permettant la collecte de données objectives en milieu *in situ*, les observations directes et instrumentées par vidéo sont abandonnées au détriment du traceur d'activité qui représente le seul moyen d'obtenir des informations objectives sur les manipulations des utilisateurs à distance. Certains auteurs regorgent d'imagination et de stratagèmes pour obtenir des observations directes en milieu *in situ*. C'est le cas de Rantanen

et al. (2017) qui se servent de la venue des infirmières à domicile réalisant leurs visites quotidiennes, afin de vérifier la bonne utilisation d'un dispositif permettant l'observance du traitement médicamenteux chez les patients âgés, le tout combiné à un traceur d'activité.

Au total, nous avons comptabilisé 35 articles utilisant uniquement des outils permettant de collecter des données subjectives contre 4 articles utilisant uniquement des outils permettant de collecter des données objectives. Enfin, 40 articles ont utilisé les deux types d'outils (ceux permettant de recueillir des données subjectives et ceux permettant de collecter des données objectives).

b. Le choix d'outils de collecte de données qualitatives *versus* quantitatives

Intéressons-nous à présent aux outils permettant de collecter des données qualitatives *versus* quantitatives. Si l'on se réfère au Tableau 9, le questionnaire est le seul outil subjectif permettant de collecter des données quantitatives. Plus exactement, 90 outils sont associés à des données qualitatives contre 119 associés à des données quantitatives. Au total, 51 articles utilisent des outils permettant d'obtenir à la fois des données qualitatives et quantitatives comme recommandés notamment par Jaspers (2009) et Ammenwerth et al. (2003). Seuls, 17 articles n'utilisent que des outils de type qualitatif et 11 uniquement des outils de type quantitatif.

Si les études suggèrent une richesse méthodologique à travers la complémentarité des outils, nous constatons également une non standardisation des études d'utilisabilité dans le domaine de la santé. Qu'en est-il des normes et des modèles utilisés pour alimenter les raisonnements des chercheurs du domaine ?

3.3.4. Les modèles et les normes

Seuls 22 auteurs ont précisé les modèles qu'ils avaient utilisés dans les études que nous avons référencées (Tableau 10). Nous avons recensé dix modèles différents dont trois davantage cités que les autres : le modèle de Nielsen (Ayubi et al., 2014; Garcia, 2019; Krieke et al., 2012; Sparkes et al., 2012), le Health-ITUEM (Beauchemin et al., 2019; Brown et al., 2013; Cho et al., 2014; Househ et al., 2015) et la norme d'utilisabilité ISO 9241-11 (Constantinescu et al., 2019; Georgsson & Staggers, 2016b; Gunter et al., 2016; Hussain et al., 2015). Pour rappel, le Health-ITUEM (Yen, 2010) se base sur cette dernière norme ISO ainsi que sur le TAM. De plus, lorsque l'on y regarde de plus près, trois des articles proviennent de deux entités de l'Université de Columbia à New York : le département d'informatique biomédicale (Brown et

al., 2013; Cho et al., 2018) et l'école de soins infirmiers (Beauchemin et al., 2019; Cho et al., 2018). Par ailleurs, 57 auteurs n'ont pas spécifié le modèle sur lequel ils se sont basés.

Tableau 10. Nombre de modèles d'utilisabilité référencés.

Modèles	Nombre d'articles concernés	Fréquence (% sur 79 articles)
ISO/IEC 25010.31	1 (article n°3)	1,27%
Nielsen	4 (articles n°4 ;8 ;21 ;51)	5,06%
Health-ITUEM	4 (articles n°1 ; 11 ;40 ;52)	5,06%
ISO 9126	1 (article n°25)	1,27%
TAM	1 (article:n°36)	1,27%
QoE	1 (article n°37)	1,27%
STROBE guideline	1 (article n°50)	1,27%
TURF	1 (article n°63)	1,27%
NASA	1 (article n°78)	1,27%
ISO 9241-11	4 (articles n°6 ;30 ;33 ;38)	5,06%

3.3.5. Les métriques

Nous avons ensuite réalisé une analyse des métriques utilisées dans les différentes études. Si certains auteurs précisent explicitement ces métriques, nous avons référencé 26 articles dans lesquels aucune métrique n'est mentionnée. Toutefois, sur la base de la méthode et des résultats présentés dans ces articles, nous avons tenté d'inférer les métriques à partir des définitions présentées précédemment (Partie I, Chapitre 1, 3. Les approches conceptuelles de l'utilisabilité : les mesures des métriques de l'utilisabilité, p. 26). Par exemple, dans l'étude de Scott et al. (2017) (article n°74) portant sur l'évaluation d'un dispositif en chirurgie ambulatoire en milieu *in situ*, les auteurs ne précisent pas les métriques. Ils déclarent toutefois mesurer le nombre de connexions des utilisateurs sur leur application et le temps passé lors de la manipulation via un traceur d'activité et utilisent le questionnaire SUS. Nous avons donc considéré que les auteurs mesuraient ici l'efficacité ainsi que la satisfaction. D'autre part, le problème inverse a aussi pu être diagnostiqué, lorsque les auteurs citent des métriques, mais que les résultats ne traduisent pas l'utilisation de certaines métriques. Nous avons aussi pris la décision de fusionner certaines métriques mentionnées dans un vocabulaire différent, mais mesurant des variables similaires telles que l'« apprenabilité », l'« apprentissage » et la « facilité d'apprentissage ».

Ces métriques citées et inférées sont visibles sur la Figure 16. Nous avons référencé 39 métriques différentes. La métrique la plus utilisée est « la satisfaction » représentant 72,15 % des articles. En deuxième position se place l'efficacité (35,44 %) suivie de l'efficience (34,18 %). Cette importante disparité traduit également une absence de standardisation associée aux études d'utilisabilité.

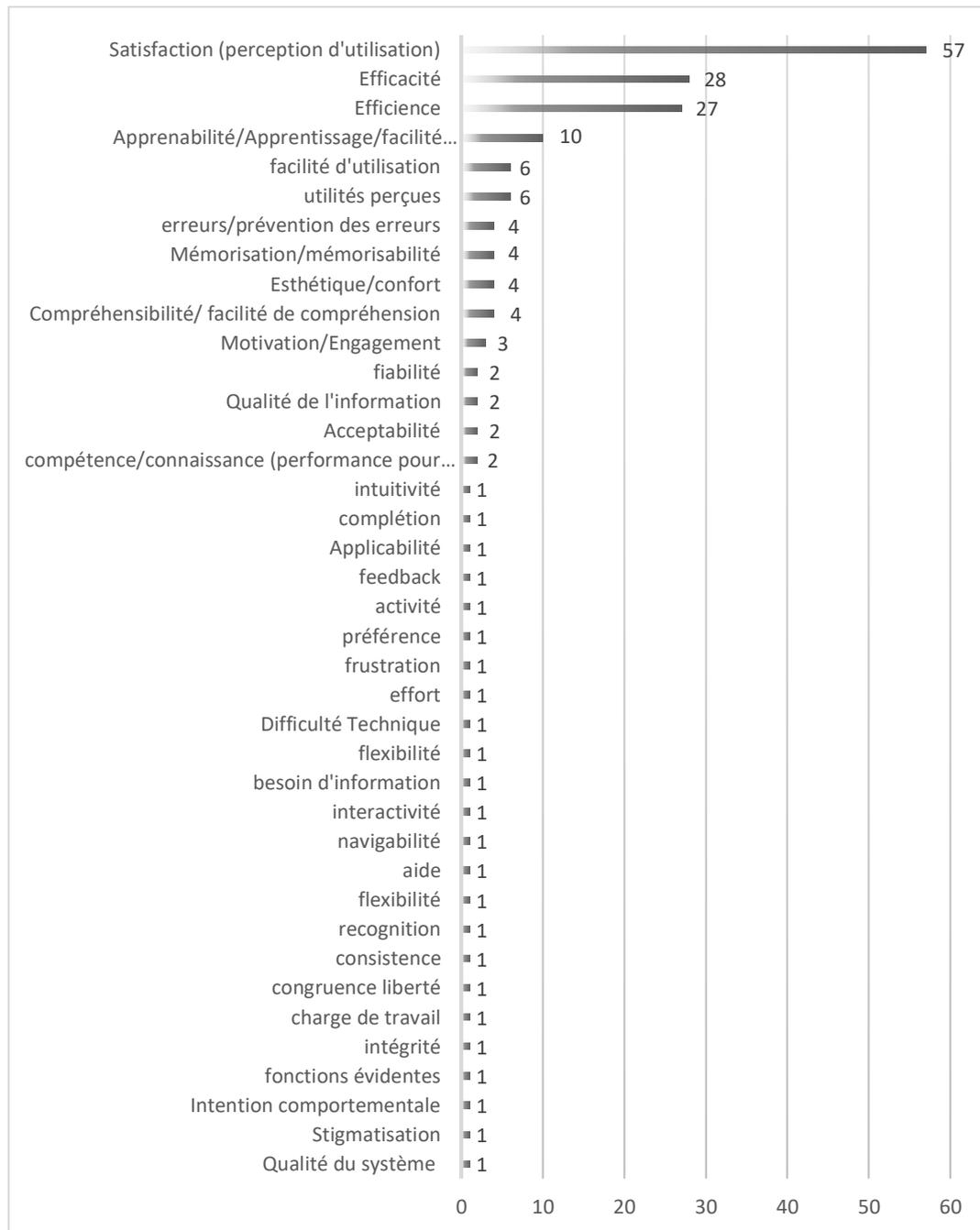


Figure 16. Fréquence des métriques référencées dans les articles sélectionnés.

3.3.6. Les caractéristiques des participants

Nous avons répertorié dans le Tableau 11, les caractéristiques utilisateurs prises en compte par les auteurs dans les 79 articles. Seize articles ont été classés comme « non spécifiés », car aucune information relative aux caractéristiques des utilisateurs n'a pu être observée. Le sexe et l'âge sont deux caractéristiques systématiquement présentes dans la définition de la population. Le niveau d'éducation est aussi une variable très souvent abordée dans la littérature

(20,25 % des études). En ce qui concerne les compétences des utilisateurs, l'expérience des Technologies de l'Information ou TI semble être la compétence la plus mesurée suivie de l'expérience relative au matériel. La littératie en santé (ou en e-santé) est sous représentée dans ces études. C'est pourtant une compétence ayant une très forte influence sur le résultat de l'utilisation (Van der Vaart et al., 2019). L'état de santé est également rarement pris en compte. Seule la pathologie diagnostiquée se détache légèrement des autres caractéristiques.

Nous avons vu que de nombreuses caractéristiques des utilisateurs sont à considérer dans l'évaluation de l'utilisabilité. Nous allons nous demander quels sont réellement les liens qu'elles entretiennent avec l'utilisabilité. Nous proposons une analyse de ces liens (cf. Partie I, Chapitre 2, 4. État de l'art sur la prise en compte du contexte d'usage dans l'évaluation de l'utilisabilité de dispositifs de m-santé, p. 84).

3.4. Conclusion et mise en relation de l'analyse systématique avec d'autres études relatives aux évaluations des dispositifs en santé pour le domicile patient

3.4.1. Une évolution à travers le temps

Ten Haken et al. (2018) et Zapata et al. (2015) ont réalisé des analyses systématiques portant sur l'utilisabilité de dispositifs en santé. Même si les critères de sélection des articles varient, nous pouvons néanmoins observer une évolution croissante depuis 2012 dans ces analyses systématiques du nombre de publications au cours de ces vingt dernières années dans le domaine de l'évaluation de l'utilisabilité des dispositifs en santé (Figure 17).

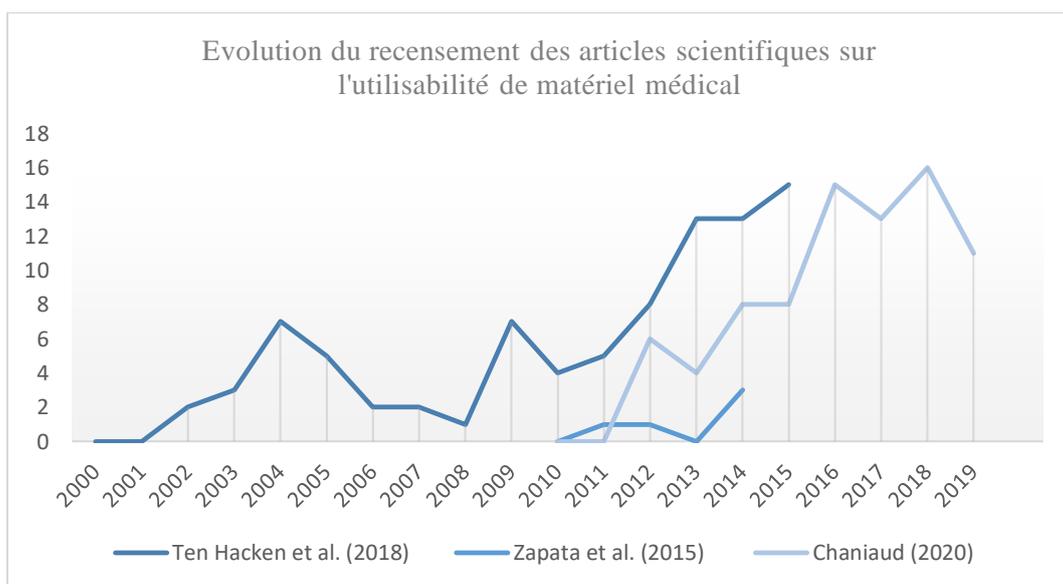


Figure 17. Évolution au cours du temps du nombre de publications portant sur l'utilisabilité de matériel médical.

Tableau 11. Nombre de variables issues des caractéristiques utilisateurs recensées.

Caractéristiques utilisateurs	Nombre d'articles concernés	Fréquence (% sur 79 articles)
Critères sociodémographiques		
Sexe	58 (articles n°1 ; 3 ; 4 ; 5 ; 6 ; 7 ; 8 ; 9 ; 10 ; 12 ; 13 ; 14 ; 16 ; 17 ; 18 ; 23 ; 24 ; 25 ; 26 ; 29 ; 30 ; 31 ; 32 ; 33 ; 34 ; 35 ; 36 ; 37 ; 38 ; 39 ; 40 ; 41 ; 42 ; 43 ; 45 ; 46 ; 47 ; 48 ; 50 ; 51 ; 52 ; 54 ; 55 ; 56 ; 60 ; 61 ; 62 ; 63 ; 65 ; 67 ; 68 ; 69 ; 70 ; 71 ; 73 ; 76 ; 78 ; 79)	73,42%
Age	58 (articles n°1 ; 3 ; 4 ; 5 ; 6 ; 7 ; 8 ; 9 ; 10 ; 12 ; 13 ; 14 ; 15 ; 16 ; 17 ; 18 ; 22 ; 23 ; 24 ; 25 ; 26 ; 29 ; 30 ; 31 ; 32 ; 33 ; 35 ; 36 ; 37 ; 38 ; 39 ; 40 ; 41 ; 42 ; 43 ; 45 ; 46 ; 47 ; 48 ; 50 ; 51 ; 52 ; 54 ; 55 ; 60 ; 61 ; 62 ; 63 ; 65 ; 67 ; 68 ; 69 ; 70 ; 71 ; 73 ; 76 ; 78 ; 79)	73,42%
Nombre de personnes vivant avec	1 (article n°76)	1,27%
Nombre d'enfants	1 (article n°56)	1,27%
Type d'assurance maladie	1 (article n°30)	1,27%
Statut socioprofessionnel	2 (articles n° 47 ; 79)	2,53%
Statut matrimonial	3 (articles n°33 ; 35 ; 73)	3,8%
Zone d'habitation	3 (articles n°3 ; 50 ; 68)	3,8%
Niveau d'éducation	16 (articles n°7 ; 17 ; 29 ; 31 ; 35 ; 39 ; 42 ; 43 ; 45 ; 47 ; 48 ; 51 ; 56 ; 60 ; 78 ; 79)	20,25%
Revenu	3 (articles n°42 ; 56 ; 79)	3,8%
Compétences		
Expérience TI	21 (articles n°3 ; 4 ; 6 ; 7 ; 8 ; 10 ; 16 ; 17 ; 21 ; 22 ; 26 ; 33 ; 36 ; 37 ; 39 ; 42 ; 43 ; 48 ; 61 ; 65 ; 79)	26,58%
Expérience dans le matériel	4 (articles n°65 ; 67 ; 70 ; 79)	5,06%
Habiletés cognitives	1 (article n°68)	1,27%
Motivation	1 (article n°37)	1,27%
Littératie en santé	2 (articles n°16 ; 38)	2,53%
Littératie en esanté	1 (article n°45)	1,27%
Littératie	1 (article n°17)	1,27%
Niveau de dextérité	1 (article n°27)	1,27%
État de santé		
Pathologies diagnostiquées	7 (articles n°6 ; 12 ; 14 ; 18 ; 35 ; 42 ; 53)	8,86%
Modalité de traitement / observance médicamenteuse	3 (articles n°12 ; 22 ; 70)	3,8%
Autogestion	1 (article n°16)	1,27%
Habitude comportementale	3 (articles n°8 ; 15 ; 22)	3,8%
Statut dépistage	1 (article n°47)	1,27%
Participation des proches	2 (articles n°16 ; 50)	2,53%
Niveau de douleur	2 (articles n°12 ; 31)	2,53%
Alimentation	1 (article n°70)	1,27%
Sommeil	1 (article n°28)	1,27%
IMC	2 (articles n°8 ; 31)	2,53%
Autre		
Non spécifié	16 (articles n°2 ; 19 ; 20 ; 30 ; 34 ; 44 ; 49 ; 57 ; 58 ; 59 ; 64 ; 66 ; 72 ; 74 ; 75 ; 77)	20,25%

Zapata et al. (2015) ont analysé 22 études portant sur l'utilisabilité de dispositifs en santé ubiquitaire (4 articles et 18 actes de conférences) entre 2010 et 2014. Dans notre propre analyse systématique, nous avons pu mettre en avant des résultats similaires à ceux de Zapata et al. (2015), comme la publication de 100% des études dans des revues d'informatiques médicales ou seulement médicales. Les auteurs ont également mis en avant la très forte proportion des études réalisées en laboratoire ayant recours en majorité à des tests utilisateurs avec une utilisation massive des questionnaires (13/22) et des entretiens (7/22) ainsi que la présence d'études ne précisant pas leurs outils de mesure (4/22). Enfin, seulement 5 études (sur 22) pourraient être considérées comme « de terrain » ou « *in situ* » selon notre vocabulaire. La seule différence est l'utilisation des métriques d'opérabilité comme première métrique recensée par Zapata et al. (2015). La satisfaction n'arrive qu'en troisième position derrière l'efficacité.

Cette différence pourrait s'expliquer par les critères d'analyses des métriques. En effet, dans notre analyse, nous avons pu observer que de nombreux articles précisaient mesurer l'efficacité. Pourtant, à la lecture des résultats, les auteurs n'ont pas présenté les données de cette métrique. Dans ces conditions, nous l'avons retiré de la liste des métriques utilisées par les auteurs. Pour le cas de l'analyse de Zapata et al. (2015), les auteurs ont probablement pris en compte l'ensemble des métriques cités par les auteurs des articles analysés. Cette information n'est pas précisée dans leur analyse.

3.4.2. Conclusion : Quelles évolutions pouvons-nous constater aujourd'hui ? Synthèse sur la prise en compte des recommandations des évaluations de l'utilisabilité

Les constats sur le manque de standardisation partagés par de nombreux auteurs (e.g., Brown et al., 2013 ; Georgsson & Staggers, 2016 ; Jaspers, 2009 ; Nelson et al., 2017 ; Peute et al., 2013 ; Peute et al., 2008 ; Yen, 2010 ; Yen & Bakken, 2012) ont-ils permis de faire évoluer les études d'utilisabilité ? Au vu de notre analyse, nous avons également pu observer cette absence de standardisation entre les études. Au total, nous comptabilisons 34 études (43,64%) ayant suivi les recommandations pour une bonne reproductibilité et pour une mesure de l'utilisabilité efficace. Cela implique qu'il y a encore trop d'études qui ne précisent pas leurs métriques, se basent sur des normes peu utilisées, ne diversifient pas leurs méthodes et ne précisent pas leurs échantillons (ou ne ciblent pas le bon).

Premièrement, nous avons vu qu'il était nécessaire de privilégier des études mixtes (qualitatives *versus* quantitatives et subjectives *versus* objectives). En effet, l'utilisation de

méthodes subjectives non couplées à des méthodes objectives comporte des biais. Powell et al. (2016) montrent, dans leur étude, d'importantes disparités entre les résultats des évaluateurs sur le recueil associé à une méthodologie subjective et à une méthodologie objective. **Le recours à des méthodes mixtes qualitatives versus quantitatives ainsi que subjectives vs objectives est fortement encouragé.**

Deuxièmement, il est nécessaire de **hiérarchiser les métriques de l'utilisabilité** en fonction du dispositif médical. A titre d'exemple, l'évaluation d'une application de gestion du diabète ne comprend pas les mêmes enjeux que la conception d'un défibrillateur. L'application relative à la gestion du diabète devra davantage mettre l'accent sur la satisfaction et l'expérience utilisateur alors que le défibrillateur aura pour principales métriques l'efficience et l'efficacité pour garantir la sécurité du patient.

Troisièmement, dans le cas de notre analyse systématique, nous ne remettons pas en question l'environnement, car certaines études en laboratoire seront poursuivies sur le terrain. Cependant, il est nécessaire de rappeler que les études en laboratoire couplées à des études *in situ* sont complémentaires. Les études en laboratoire ne représentant pas nécessairement l'utilisation naturelle des outils de santé à distance (Martínez-Pérez et al., 2015; Zhang & Adipat, 2005) ; par conséquent, des tests plus approfondis *in situ* peuvent être nécessaires. Malgré la richesse des données que peuvent fournir les tests sur le terrain, la capture de l'utilisation des applications dans un environnement dynamique rend l'observation directe difficile (Schoffman et al., 2013) et les échantillons d'utilisateurs sont souvent plus petits (Sarkar et al., 2016). Il est donc nécessaire de diversifier les environnements des tests d'utilisabilité. **Il est également possible de faire varier les environnements en laboratoire afin de remettre l'utilisateur en contexte en ayant recours à des scénarii ou des simulations** (Kjeldskov & Skov, 2003 ; Kjeldskov & Stage, 2004).

Quatrièmement, nous avons pu constater que certaines études n'ont pas interrogé de patients, mais uniquement des soignants (e.g., Brown-Johnson et al., 2015) alors que l'outil conçu (un jeu en réalité virtuelle pour promouvoir l'autogestion du cancer de la langue) était à destination des patients. A l'inverse, Ten Haken et al. (2018) regrettent dans leur analyse le peu d'études s'intéressant aux points de vue des soignants. Il est nécessaire de **cibler l'échantillon (futur utilisateur), d'avoir suffisamment de connaissances sur ce dernier et de ne pas omettre d'interroger les utilisateurs annexes comme le personnel soignant ou les aidants** lors de la conception d'un dispositif en santé.

Pour conclure, **les recherches portant sur la e-santé peuvent prendre plusieurs formes selon la question de recherche, le temps et les ressources des évaluateurs.** Il n'existe pas de méthodologie unique, car chacune d'entre elles fournit un type de données différent et présente un ensemble unique d'avantages et de limites. Pour cela, **nous encourageons les études itératives couplées à une méthodologie mixte.**

4. État de l'art sur la prise en compte du contexte d'usage dans l'évaluation de l'utilisabilité de dispositifs de m-santé

Nous avons vu que l'utilisabilité était déterminée non seulement par ses interactions entre l'utilisateur et le dispositif, mais aussi par le degré d'intégration de cette technologie pour l'exécution de tâches dans un environnement donné. En accord avec la norme ISO 9241-41, **nous considérons le contexte d'usage au sens d'utilisateur, de tâche, de ressources et d'environnement.**

Contrairement au milieu hospitalier, le domicile patient est un milieu non contrôlé ce qui implique des risques supplémentaires, puisque l'objet doit être compatible avec les connaissances, les capacités motrices, le mode de vie et l'environnement du patient utilisateur (Kaufman-Rivi et al., 2010). Cela signifie que des facteurs spécifiques concernant la mise en œuvre et l'utilisation à domicile doivent maintenant être pris en compte. Le facteur de risque le plus important se trouve dans la diversité des utilisateurs ainsi que dans le transfert dans la chaîne de soins (Ten Haken et al., 2018; Weick-Brady & Lazerow, 2006). Ten Haken et al. (2018, p. 2) soulignent : « *Comprendre l'expérience de l'utilisateur et l'information sur les événements indésirables et les incidents évités de justesse sont des aspects importants pour acquérir des connaissances sur la mise en œuvre et l'utilisation dans le contexte des soins à domicile* ²¹ ». D'après les auteurs, plusieurs éléments sont à prendre en compte dans les environnements domestiques : la diversification des utilisateurs, les niveaux de complexité des technologies médicales, les niveaux de formation, d'instruction ou d'éducation. Chacun de ces éléments peut être mis en relation avec les composants du contexte d'usage précités.

²¹ Traduit de l'anglais: "Understanding user experiences and information about adverse events and near incidents are important aspects for developing knowledge regarding implementation and use in home care setting." (Ten Haken et al., 2018, p. 2)

Une fois un dispositif conçu et utilisable, il se peut qu'une fois sur le terrain le degré d'utilisabilité soit moindre. **Cela peut être dû à la non-prise en compte d'un des quatre éléments du contexte d'usage** (e.g., Borsci et al., 2018 ; Georgsson & Staggers, 2016 ; Grebin et al., 2018 ; Kaufman-Rivi et al., 2010). Nous avons cherché à comprendre comment et pourquoi le contexte d'usage pouvait avoir un impact sur l'utilisabilité en se servant de la littérature scientifique afin de caractériser au mieux ce contexte dans le cadre du projet *Smart Angel*. Celui-ci, en effet, possède un contexte bien particulier, dans le sens où le processus d'intégration du dispositif auprès des patients est préparé à l'hôpital pour un suivi à domicile (cf. Introduction Générale, 2.1. Le projet *Smart Angel* : un double objectif, p. 3). Nous avons vu précédemment (cf. Partie I, Chapitre 1, 3.4. Les composants du contexte d'utilisation, p. 37) que ce contexte pouvait influencer l'utilisabilité. Nous avons donc sélectionné des éléments du contexte d'usage adaptés au projet *Smart Angel* que nous avons relevés dans l'analyse systématique :

- Les **caractéristiques utilisateurs** font partie des éléments du contexte d'usage les plus cités dans la littérature comme ayant un impact sur l'utilisabilité (cf. Partie I, Chapitre 2, 3.3.6. Les caractéristiques des participants, p. 79) ;
- Les **documents procéduraux** issus des manuels d'utilisation renvoyant à un apprentissage procédural considéré comme ressources réutilisables d'après la norme ISO 9241-11. (cf. Partie I, Chapitre 1, 3.4.1. Les ressources, p. 39) ;
- La **formation du patient par le personnel infirmier** peut être considérée à l'interface entre les ressources non renouvelables (ISO 9241-11) et l'environnement puisqu'appartenant à l'environnement hospitalier (cf. Partie I, Chapitre 1, 3.4.1. Les ressources, p. 39).

4.1. Les caractéristiques utilisateurs

Plusieurs chercheurs ont récemment investigué les liens entre les caractéristiques des utilisateurs et l'utilisabilité de dispositifs connectés dans le domaine de la santé (e.g. Georgsson & Staggers, 2016; Jensen et al., 2010; Liang et al., 2018; Monkman & Kushniruk, 2013; Sonderegger et al., 2016; Van der Vaart et al., 2019). Nous avons pu observer grâce à notre analyse systématique (cf. Partie I, Chapitre 2, 3.3.6. Les caractéristiques des participants, p. 79) que quatre caractéristiques utilisateurs étaient particulièrement étudiées dans la littérature scientifique :

- L'âge (e.g., Georgsson & Staggers, 2016 ; Jones & Caird, 2017 ; Kaufman et al., 2003 ; Mykityshyn et al., 2002 ; Sonderegger et al., 2016 ; Sparkes et al., 2012 ; Van der Vaart et al., 2019) ;
- La situation professionnelle et le niveau d'étude (e.g., Georgsson & Staggers, 2016 ; Liang et al., 2018 ; Van der Vaart et al., 2019) ;
- La technophilie, c'est-à-dire l'expérience antérieure dans la technologie de l'information (TI) et dans les dispositifs médicaux (e.g., Georgsson & Staggers, 2016 ; Harte et al., 2018 ; Sparkes et al., 2012) ;
- La littératie en santé (e.g., Czaja et al., 2015 ; Kaufman et al., 2003 ; Kim & Xie, 2017 ; Van der Vaart et al., 2019) incluant les connaissances en santé (e.g., Williams et al., 1998) ;
- La motivation et l'engagement des patients (e.g., O'Brien & Toms, 2008 ; Rico-Olarte et al., 2018 ; Tissier & Gronier, 2014).

Nous avons classifié ces caractéristiques utilisateurs en trois catégories de facteurs (Figure 18) : **les facteurs sociaux, les facteurs cognitifs et les facteurs émotionnels**. Dans la majorité des études, les auteurs s'accordent plus ou moins sur ces liens et cela sur des dispositifs différents. Nous détaillerons ces études dans les paragraphes suivants.

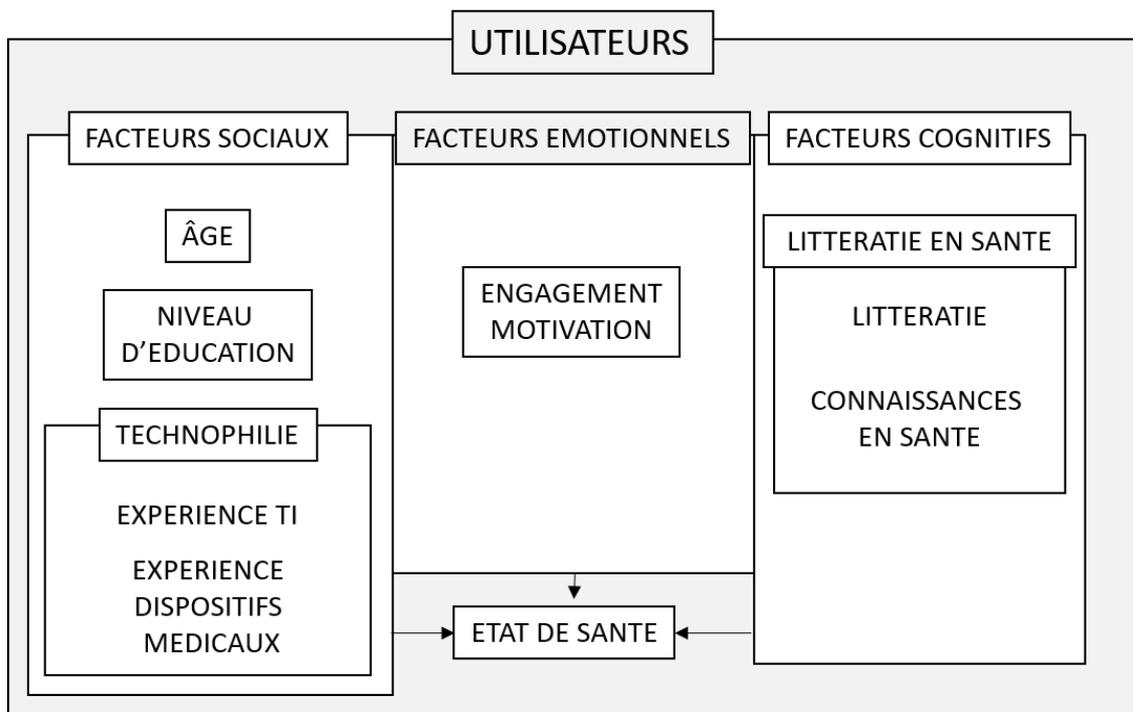


Figure 18. Représentation des caractéristiques utilisateurs en fonction des facteurs sociaux, émotionnels et cognitifs.

4.1.1. Les facteurs sociaux

a. L'âge

L'effet de l'âge sur l'utilisabilité semble faire consensus dans la littérature. Les utilisateurs plus âgés auraient plus de difficulté à utiliser des technologies en santé que les utilisateurs plus jeunes (e.g., Georgsson & Staggers, 2016 ; Jones & Caird, 2017 ; Kaufman et al., 2003 ; Mykityshyn et al., 2002 ; Sonderegger et al., 2016 ; Sparkes et al., 2012 ; Van der Vaart et al., 2019). À l'exception de la satisfaction qui ne semble pas être impactée par l'âge des utilisateurs (e.g., Liang et al., 2018). Nous détaillons l'ensemble de ces études dans la partie suivante (cf. Partie II, Chapitre 4, 2.2.1. L'âge, p. 194)

b. Le niveau d'éducation

Le niveau d'études est également une variable que l'on retrouve dans de nombreuses études traitant de l'utilisabilité. Pourtant, à notre connaissance, aucune recherche n'a prouvé ce lien (e.g., Georgsson et Staggers, 2016 ; Liang et al., 2018 et Van der Vaart et al., 2019). Nous détaillons ces études dans la partie suivantes (cf. Partie II, Chapitre 4, 2.2.2. Le niveau d'éducation, p. 195).

c. La technophilie (expérience TI/dispositifs en santé)

L'influence de l'expérience dans les technologies ainsi que l'expérience antérieure dans les dispositifs en santé sur l'utilisabilité reste plus mitigée que l'impact de l'âge (e.g., Georgsson & Staggers, 2016 ; Harte et al., 2018 ; Sparkes et al., 2012) . Nous détaillons ces études dans la partie suivante (cf. Partie II, Chapitre 4, 2.2.3. Technophilie (expérience TI/dispositifs médicaux), p. 195).

4.1.2. Les facteurs cognitifs

a. La littératie en santé

La littératie en santé est une caractéristique des utilisateurs et plus précisément une « compétence » qui semble être un élément essentiel à la bonne compréhension et manipulation d'un dispositif médical (e.g. Czaja et al., 2015 ; Holt et al., 2019 ; Jensen et al., 2010 ; Monkman et al., 2015 ; Monkman & Kushniruk, 2013 ; Van der Vaart et al., 2019). Pourtant cette caractéristique se heurte à différentes problématiques, sa définition et son évaluation, justifiant que nous y revenions ci-après.

b. La Littératie générale *versus* la littératie en santé

La littératie « générale » s'exprime différemment selon le contexte et intervient à plusieurs niveaux et dans des domaines extrêmement variés : sa santé, la gestion de son

portefeuille, sa participation politique ou au travail (Masny & Dufresne, 2007). Cette notion n'est pas acquise à l'école, mais se « *développe tout au long de sa vie dans divers contextes, ainsi que dans ses relations avec ses pairs et sa collectivité* » (Murray et al., 2005, p.95). De plus, il ne suffit pas d'avoir un bon niveau de littératie générale pour avoir un bon niveau de littératie en santé. Pour utiliser l'information en santé, il est nécessaire de connaître le fonctionnement du système de santé autant que des éléments du corps et de l'alimentation (Kickbusch, 2004). La littératie générale nécessite des compétences de base alors que la littératie en santé nécessite des compétences contextualisées en santé. Bien que ces deux littératies mobilisent des compétences différentes, il est probable que la littératie générale soit un prérequis à la littératie en santé (Rootman & Ronson, 2005). Par exemple, lorsque les chercheurs tentent d'analyser des composants de la littératie générale et de la littératie en santé, la numératie est présente dans les deux compétences (Berkman et al., 2011).

c. La/Les définition(s) de la littératie en santé

Aujourd'hui, il est possible de détailler la littératie dans différents domaines : informatique, médiatique, économique, environnementale, en santé voire en e-santé (CCA, 2007). De nombreuses définitions de la littératie en santé existent et aucune ne semble faire consensus (Huhta et al., 2018). En revanche, les chercheurs du domaine s'accordent sur le caractère multidimensionnel de cette compétence ou de ce groupement de compétences. Sørensen, Van den Broucke, Fullam, Doyle, Pelikan, Slonska, & Brand (2012, p. 3) proposent un modèle (Figure 19) et une définition consensuelle de la littératie en santé comme :

« liée à la littératie et implique des connaissances, les compétences, la motivation et la capacité d'un individu à repérer, comprendre, évaluer et utiliser des informations sur la santé lors de la prise de décisions dans les contextes des soins de santé, de la prévention des maladies et de la promotion de la santé pour maintenir ou améliorer la qualité de la vie au cours de la vie »²².

d. L'évaluation de la littératie en santé

Cette capacité (que ce soit littératie ou numératie) est particulièrement complexe à évaluer liée à sa multidimensionnalité (Sørensen et al., 2012), à l'absence de corrélation avec d'autres facteurs sociaux comme des critères socio-économiques (Zenklusen et al., 2012) et à l'absence de méthodes de mesures objectives. Nous détaillons ces éléments dans la partie suivante (cf. Partie II, Chapitre 4, 2.2.4. Littératie en santé, p. 196).

²² Traduit de l'anglais : "Health literacy is linked to literacy and entails people's knowledge, motivation and competences to access, understand, appraise, and apply health information in order to make judgments and take decisions in everyday life concerning healthcare, disease prevention and health promotion to maintain or improve quality of life during the life course" (Sørensen et al., 2012, p. 3).

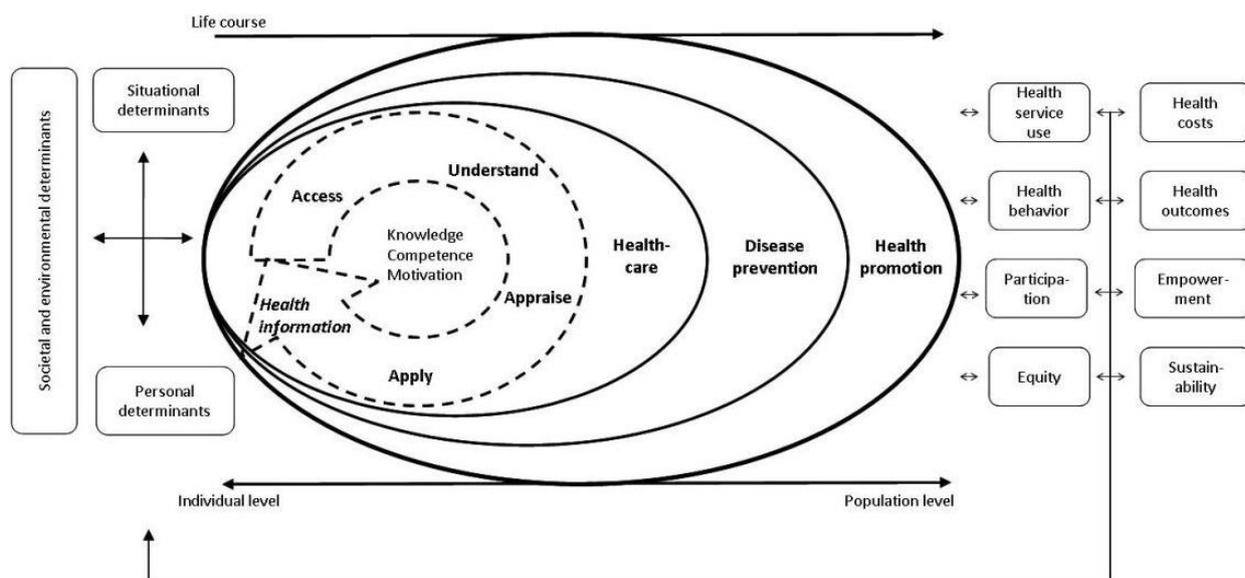


Figure 19. Modèle intégratif de la littératie en santé (Sørensen et al., 2012, p. 9).

e. Les apports de la littératie en santé

L'un des avantages (ou désavantage si cette notion n'est pas acquise) de la e-santé est justement de permettre aux patients de devenir autonomes et donc actifs dans leur propre santé (Eysenbach, 2001). Cependant, l'un des principaux précurseurs de l'autonomie dépend aussi de l'éducation thérapeutique du patient. Or, la littératie en santé est souvent mentionnée comme un déterminant à considérer dans l'éducation thérapeutique (e.g., Margat, Gagnayre, Lombrail, Andrade, & Azogui-Levy, 2017 ; Williams, Baker, Parker, & Nurss, 1998 ; Williams et al., 1998), mais aussi de bien d'autres éléments telle que la prévention (e.g., Möttus et al., 2014 ; Scott, 2002), les préoccupations envers sa santé (e.g., Rudd et al., 2007), la durée d'hospitalisation (Baker, 2002), l'observance thérapeutique (e.g., Kalichman et al., 1999), l'accès à l'information en santé (e.g., Jensen et al., 2010), l'utilisation du système de santé (e.g., Davis, 1991), l'engagement dans une démarche d'autosoins (e.g., Parker, 2000) voire l'état de santé et la guérison (e.g., Berkman et al., 2011 ; Williams et al., 1998). Les principales retombées économiques du faible niveau de littératie en santé des citoyens se reflètent sur les statistiques des hôpitaux. Les individus aux faibles niveaux de littératie en santé ont tendance à consommer davantage de soins que ceux ayant un haut niveau de littératie en santé. Il semble donc nécessaire de former les patients. Pourtant, cette solution est plus complexe qu'elle ne le laisse penser. Rares sont les patients ayant reçu une éducation qui la comprennent vraiment et savent correctement prendre des décisions vis-à-vis de leur santé. Malgré le vaste matériel

d'éducation sanitaire et les dispositifs médicaux mis à disposition, beaucoup sont rédigés pour des individus ayant un haut niveau de littératie en santé (Leichter et al., 1981).

f. La Littératie en santé et utilisabilité

Les liens entre la littératie en santé et l'utilisabilité de dispositifs connectés en santé sont, à notre connaissance, faiblement connus (e.g., Czaja et al., 2015 ; Kim & Xie, 2017, Jensen et al., 2010 ; Monkman & Kushniruk, 2013, 2015). Nous détaillons ces études dans la partie suivante (cf. Partie II, Chapitre 4, 2.2.4. Littératie en santé, p. 196).

4.1.3. Les facteurs émotionnels

a. La motivation et l'engagement

L'engagement du patient est défini par Carman et al. (2013, p. 224) comme :

« un ensemble de comportements par les patients, les membres de la famille et les professionnels de la santé et un ensemble de politiques et de procédures organisationnelles favorisant l'inclusion des patients et des membres de la famille en tant que membres actifs de l'équipe soignante et les organisations de fournisseurs avec les objectifs souhaités de l'engagement des patients et de la famille comprennent l'amélioration de la qualité et la sécurité des soins de santé »²³.

Graffigna et al. (2017) décrivent ce phénomène comme un « *processus psychosocial multidimensionnel résultant de la mise en acte cognitive, émotionnelle et comportementale* ».

En effet, ces auteurs caractérisent l'engagement à travers trois dimensions (Graffigna et al., 2013) :

- Une dimension comportementale : ce que le patient fait ;
- Une dimension cognitive : ce que le patient sait et pense, ce principe étant connecté à ce que le patient sait, comprend et comment il met du sens sur sa maladie, son traitement, ses possibles évolutions, son monitoring... ;
- Une dimension émotionnelle : ce que le patient ressent.

Les dimensions cognitives et comportementales sont en interrelation. En effet, l'utilisation d'un dispositif par un patient nécessite des capacités cognitives pour déclencher des comportements positifs envers la technologie. L'état de la maladie peut comporter des contraintes physiologiques, voire cognitives, qui engendrent des difficultés dans l'utilisation de

²³ Traduit de l'anglais : "Patient and family engagement as patients, families, their representatives, and health professionals working in active partnership at various levels across the health care system—direct care, organizational design and governance, and policy making—to improve health and health care" Carman et al. (2013, p. 224)

dispositifs médicaux limitant fortement l'engagement du patient. Il est probable que ces dimensions interdépendantes soient agencées de manière pyramidale entre elles.

b. L'engagement de l'utilisabilité

De nombreux articles témoignent de l'impact de l'engagement sur la guérison par rapport à des patients considérés comme non engagés et inactifs dans leur santé (e.g., Barelló et al., 2016 ; Greene & Hibbard, 2012 ; Hibbard et al., 2004). Pour O'Brien et Toms (2008), l'utilisabilité est liée à l'expérience d'engagement. Un manque de motivation des participants pourrait par conséquent donner des mesures plus faibles d'utilisabilité.

c. Les mesures de l'engagement

Pour mesurer l'engagement du patient, il existe des questionnaires, dont le PAM (Patient Activation Measure) réalisé par Hibbard et al. (2004). Ces questionnaires souffrent de nombreuses lacunes. Par exemple, Graffina et al. (2015) déclarent qu'un questionnaire ne permet pas de mesurer efficacement la dimension émotionnelle. De plus, actuellement, le PAM est le seul questionnaire traduit mais il est partiellement validé par Hudon et al. (2012).

4.2. Les ressources

4.2.1. Les ressources réutilisables : les documents procéduraux

a. Une définition et un modèle

Les documents procéduraux constituent l'ensemble des documents tels que des guides, notices, manuels explicatifs, recettes de cuisine, tutoriels nécessitant de réaliser une action. Heurley (1997, p. 127) propose une définition de ce type si particulier de document : « *un texte procédural est un texte dont la fonction principale est de communiquer une procédure, c'est-à-dire un ensemble d'opérations et/ou d'actions à exécuter dans le but d'atteindre un but donné* ». Apprendre une procédure à partir d'un texte est une situation d'apprentissage complexe. Les documents procéduraux sont différents des documents déclaratifs, car ils impliquent la réalisation d'une action (Heurley, 1994). Kieras et Bovair (1986) définissent ce processus comme une transformation de la connaissance déclarative en connaissance procédurale. Le type et l'organisation de l'information permettent à l'utilisateur de développer un modèle mental de la tâche à accomplir (Guthrie, Bennett & Weber, 1991). Lorsqu'un individu effectue des actions par le biais d'instructions, il a réussi à transformer une présentation linguistique et/ou graphique de l'information en une action. Cette transformation implique plusieurs processus cognitifs et se compose de 5 étapes selon le modèle de traitement procédural de l'information de Ganier,

Gombert et Fayol (2000): (1) la formation et le maintien de l'objet en mémoire de travail ; (2) l'encodage des instructions ; (3) l'encodage des caractéristiques du dispositif - selon le modèle d'Anderson (1983), cet encodage est lent et contrôlé, car il implique la mémoire de travail ; (4) l'élaboration d'une représentation intégrée de l'ensemble et le développement d'un plan d'action ; (5) enfin, la mise en œuvre de l'action. De plus, l'individu doit être capable de comparer et de régler l'état de l'appareil à chaque étape afin de réaliser la tâche (Figure 20).



Figure 20. Modèle de Ganier, Gomber et Fayol (2000) sur le traitement des documents procéduraux.

b. Les documents procéduraux et l'utilisabilité

La transformation des informations déclaratives en informations procédurales dépend des instructions contenues dans les documents, du contexte d'utilisation, de la complexité du dispositif et des connaissances et traitements effectués par l'utilisateur (Guthrie et al., 1991 ; Heurley, 1994 ; ISO 9241-11, 2018, Fayol, 2002). Cela dépend également du format de présentation des instructions (Ganier et al., 2000) ainsi que des stratégies de traitement de l'information liées à la capacité limitée de la mémoire de travail (Baddeley, 1986).

Différents formats de présentation (texte, image, audio, dynamique...) et différentes stratégies (fragmenté, retardé, atomisé...) existent. Plusieurs études ont examiné les avantages et les inconvénients de ces différents formats. Vermersch (1985) s'est intéressé aux formats de présentation des recettes de cuisine, Blaiwes (1974) a analysé la tâche d'un contrôleur de communication fictif, Ganier (2002) s'est intéressé à la cocotte-minute. Cependant, il y a un manque de recherche se concentrant sur les formats de présentation des dispositifs médicaux à domicile (Kools, van de Wiel, Ruiters, & Kok, 2006 ; Mykityshyn, Fisk, & Rogers, 2002 ; Nguyen et al., 2017). Pourtant, des instructions mal conçues peuvent induire des erreurs et/ou des défaillances dans l'exécution de l'action, voire l'abandon et le rejet de l'objet (Patel, Eisemon, & Arocha, 1990).

Des recommandations de conception des documents procéduraux existent déjà (e.g. Zhang & Adipat, 2005). Par exemple, il est nécessaire de préciser l'objectif de la lecture fixée

initialement, car le traitement d'une lecture « lire pour apprendre » et « lire pour faire » est différent (Stricht, 1985). Hayes et Simon (1974) observent une stratégie par « caractère cyclique du processus d'utilisation ». Les utilisateurs effectuent une première lecture du texte, puis réalisent ensuite de brèves séquences de lecture. Ce cycle est particulièrement visible lorsqu'un lecteur se trompe et se met à chercher la solution. Hayes et Simon (1974) déclarent que certaines consignes peuvent être lues jusqu'à onze fois durant sa résolution. Les utilisateurs ont aussi recours à ce que Vermersch (1985) appelle « *l'atomisation de l'action* », correspondant à une alternance entre l'activité de prise d'information et l'exécution de la tâche. La prise d'information, limitée par le traitement et le stockage en mémoire de travail (Baddeley, 1986) est alors fragmentée en des sessions très courtes de lecture. Jannin et al. (2019) recommandent cette stratégie dans le cadre de l'apprentissage des points de suture dans la formation des professionnels de santé.

Il est important de vérifier, par le biais d'évaluations cognitives systématiques, la bonne compréhension de ces documents (Ganier & Querrec, 2012). Plusieurs études ont porté sur l'évaluation du format de présentation des instructions (e.g., Bol, Weert, Haes, Loos, & Smets, 2015; Lemarié, Castillan, & Eyrolle, 2017; Mykityshyn, Fisk, & Rogers, 2002; Nguyen et al., 2017), mais aucune ne fait consensus sur le format le plus approprié dans le domaine de la santé.

c. Les limites des documents procéduraux

Les documents procéduraux peuvent être une aide précieuse s'ils sont correctement conçus. Le problème n'est généralement pas la conception, mais l'environnement dans lequel les utilisateurs manipulent ce type d'information procédurale. En effet, nombreux sont les auteurs à mettre en garde sur la non-lecture de ces textes (e.g., Heurley & Ganier, 2006 ; Silver & Wogalter, 1991 ; Wright, Creighton & Threlfall, 1982). Les utilisateurs liraient en dernier recours les documents après avoir tenté sans notice de faire fonctionner leur appareil, ce qui est particulièrement dangereux dans le cas des dispositifs connectés en santé. Si les utilisateurs ne lisent pas les documents, ils pourront malheureusement manipuler de façon incorrecte le dispositif et obtenir *in fine* des données erronées. La lecture des documents ainsi que la prise en compte d'autres stratégies d'apprentissage comme la formation en amont sont en ce sens capitales.

4.2.2. Les ressources consommables : Les formations et les programmes d'éducation thérapeutique

a. Les formations en milieu médical

La prise en compte des facteurs humains prend une place importante dans la sécurité des patients à l'hôpital comme à leur domicile (Pelayo et al., 2012; Weick-Brady & Lazerow, 2006). Il est donc important de considérer le contexte hospitalier incluant la formation des patients qui se retrouveront ensuite seuls chez eux. Ten Haken, Ben Allouch, et Van Harten (2018) montrent que la diversité des utilisateurs et les transferts d'informations dans la chaîne de soins sont un exemple de ces facteurs de risque à prendre en compte pour parfaire l'utilisabilité de technologie de santé. Parmi les utilisateurs de ces dispositifs, les soignants comme le personnel infirmier sont un groupe particulièrement important à former à ces dispositifs. Ten Haken et al. (2018) montrent que l'utilisation réussie des technologies médicales avancées à domicile nécessite des programmes adéquats d'éducation et de formation de la part du personnel. En effet, c'est par le personnel infirmier que se fait le transfert de l'information vers le patient dans la grande majorité des cas. Le partage de ces connaissances entre les soignants et les patients est un élément clé pour contribuer à la sécurité du patient et à la qualité des soins. Les études de Dos Santos-Fontes, Ferreiro de Andrade, Sterr et Conforto (2013) et Tearl, Cox et Hertzog (2006) ont révélé qu'une formation spécifique ou le soutien « d'un coordinateur de décharge dédié » présentent plusieurs avantages : réduire la durée du séjour à l'hôpital, augmenter les performances d'utilisation des équipements ou entraîner moins de demandes d'aide de la part des patients et/ou des familles. Il est donc nécessaire d'analyser ces transferts de connaissances pouvant impacter l'utilisabilité de dispositifs en santé.

b. Les types de formations

L'utilisation réussie des technologies médicales avancées à domicile nécessite des programmes adéquats d'éducation et de formation de la part du personnel (Ten Haken et al., 2018). Malgré ce constat, à notre connaissance, il n'existe que très peu d'études s'étant intéressées aux types de formations à l'utilisation de technologie médicale et à l'efficacité de ces programmes d'éducation pour le patient. Larsen (2017, 2018) s'est intéressé aux programmes d'éducation des patients dans le cadre de l'utilisation d'INR-monitor (International Normalized Ratio) et de l'utilisation d'une dialyse péritonéale. L'auteur explique qu'une stratégie de « *surveillance pédagogique fondée sur les erreurs*²⁴ » est majoritairement utilisée. Au lieu d'attendre que des erreurs aléatoires se produisent, le personnel infirmier oriente parfois

²⁴ error-based monitoring strategy

les patients vers des erreurs précises afin de créer des occasions d'enseignement particulières. Il s'agirait d'une pratique traditionnelle effectuée par les professionnels formés dans les hôpitaux. D'après Larsen (2017, p. 11) : « *L'erreur du patient est l'occasion pour le personnel infirmier de corriger et d'instruire les patients sur l'utilisation du moniteur INR et la vérifier le bon déroulement de la procédure* »²⁵. Cet apprentissage repose sur deux stratégies observables d'après Larsen (2017) :

- Une stratégie « active » favorisant le développement d'erreurs en incitant les patients à adopter une ligne de conduite erronée. En effet, Larsen (2017) montre qu'il arrive souvent que des erreurs soient induites par le personnel infirmier. Plutôt que d'attendre que des erreurs diverses surviennent au hasard, les infirmiers et infirmières travaillent à « *l'occasion pour s'assurer que des erreurs particulières se produisent afin d'offrir des possibilités d'enseignements précises* » (Larsen, 2017, p. 11) ;
- Une stratégie « passive », également appelée « *la surveillance pédagogique* » (Larsen, 2017, 2018), selon laquelle le personnel infirmier laisse l'erreur se développer. Cette dernière forme a l'avantage de permettre un accès immédiat à la compréhension (Lindwall & Ekström, 2012). En effet, le personnel infirmier a connaissance des compétences du patient nécessaire pour utiliser le matériel. Ils savent que ce dernier n'a pas forcément compris l'étape sur laquelle il a réalisé une erreur contrairement à la stratégie « active » où les erreurs sont induites par le personnel infirmier. Larsen (2017) a d'ailleurs montré que les patients ayant réalisé des erreurs de manipulation relevées par le personnel infirmier avaient plus de chance de comprendre les instructions que ceux pour lesquels le personnel infirmier n'avait pas relevé d'erreurs. Ce type de formation est particulièrement présente dans le monde de l'éducation. Les enseignants observent les élèves qui tentent de faire la tâche et, au besoin, corrigent leurs gestes (e.g., Hindmarsh et al., 2011; Prentice, 2007; Svensson et al., 2009).

Malgré la présence du personnel infirmier lors de la manipulation d'un dispositif par un patient novice, Larsen (2017) a également montré que les formations proposées par le personnel médical ne permettaient pas d'éradiquer toutes les erreurs de manipulation des utilisateurs patients. En effet, certaines erreurs pouvaient ne pas être vues par le personnel en charge de la formation et persister ainsi dans le temps, c'est-à-dire au domicile du patient dans le cadre de

²⁵ Traduit de l'anglais : “*Patient error provides the most frequent type of opportunity for nurses to correct and instruct the patients in the use of the INR-monitor and performance of the measuring procedure.*” (Larsen, 2017, p. 11)

la chirurgie ambulatoire. Si un patient se trompe dans la manipulation de son matériel, celui-ci a de grandes chances de les refaire tout le long de sa convalescence.

5. Synthèse du chapitre 2

Pour considérer un produit comme un **dispositif médical** (tel que le dispositif *Smart Angel*), celui-ci doit suivre une réglementation particulière en vue d'obtenir un marquage de conformité européenne afin d'obtenir le titre de dispositif médical. Parmi ces réglementations, la conception centrée utilisateur (ISO 9241-210 :2019) est une référence impliquant des évaluations de l'utilisabilité (IEC 62366-1 :2015/AMD1 :2020). Un dispositif de e-santé, s'il ne qualifie pas comme dispositif médical, n'aura pas d'obligation à se conformer à des normes spécifiques.

Pourtant, ce domaine de la e-santé peine à se développer correctement. Ceci est dû à plusieurs lacunes. L'une de ces dernières concerne **la complexité d'utilisation** de ces produits. Au vu de la littérature du domaine, deux arguments ont été énoncés pour expliquer ce problème majeur persistant :

- **L'absence d'un cadre et d'une méthode normalisée dans les études d'utilisabilité** (e.g., Brown et al., 2013 ; Georgsson & Staggers, 2016; Jaspers, 2009; Nelson et al., 2017 ; Peute et al., 2013 ; Peute et al., 2008; Yen, 2010; Yen & Bakken, 2012) ;
- **L'absence de connaissances de l'impact du contexte d'usage** sur l'utilisabilité (e.g., Borsci et al., 2018; Georgsson & Staggers, 2016; Grebin et al., 2018; Kaufman-Rivi et al., 2010).

Pour étayer le premier argument, nous avons réalisé une analyse systématique recensant 79 études portant sur les évaluations de l'utilisabilité de dispositifs connectés en santé destinés au domicile patient. Nous avons pu conclure que la plupart **des études ne se basaient pas sur des approches mixtes** associant des méthodes qualitatives *versus* quantitatives et subjectives *versus* objectives, ne s'intéressaient pas au contexte d'usage et ne précisaient pas les métriques d'utilisabilité utilisées.

Afin de confirmer les constats sur l'absence de prise en compte du contexte d'usage systématique, nous avons réalisé un état de l'art portant sur l'impact du contexte d'usage sur l'utilisabilité de dispositifs connectés en santé. Nous avons ciblé nos recherches en cohérence avec le contexte du projet *Smart Angel* et relevé trois catégories intrinsèques à ce contexte : les caractéristiques utilisateurs des patients en chirurgie ambulatoire, les ressources (réutilisable ou consommable) et l'environnement (hospitalier et domicile). Cet état de la littérature nous a permis de constater qu'il n'existait pas de consensus relatif à l'impact de ces éléments sur l'utilisabilité. Ces éléments sont présentés dans la Figure 21.

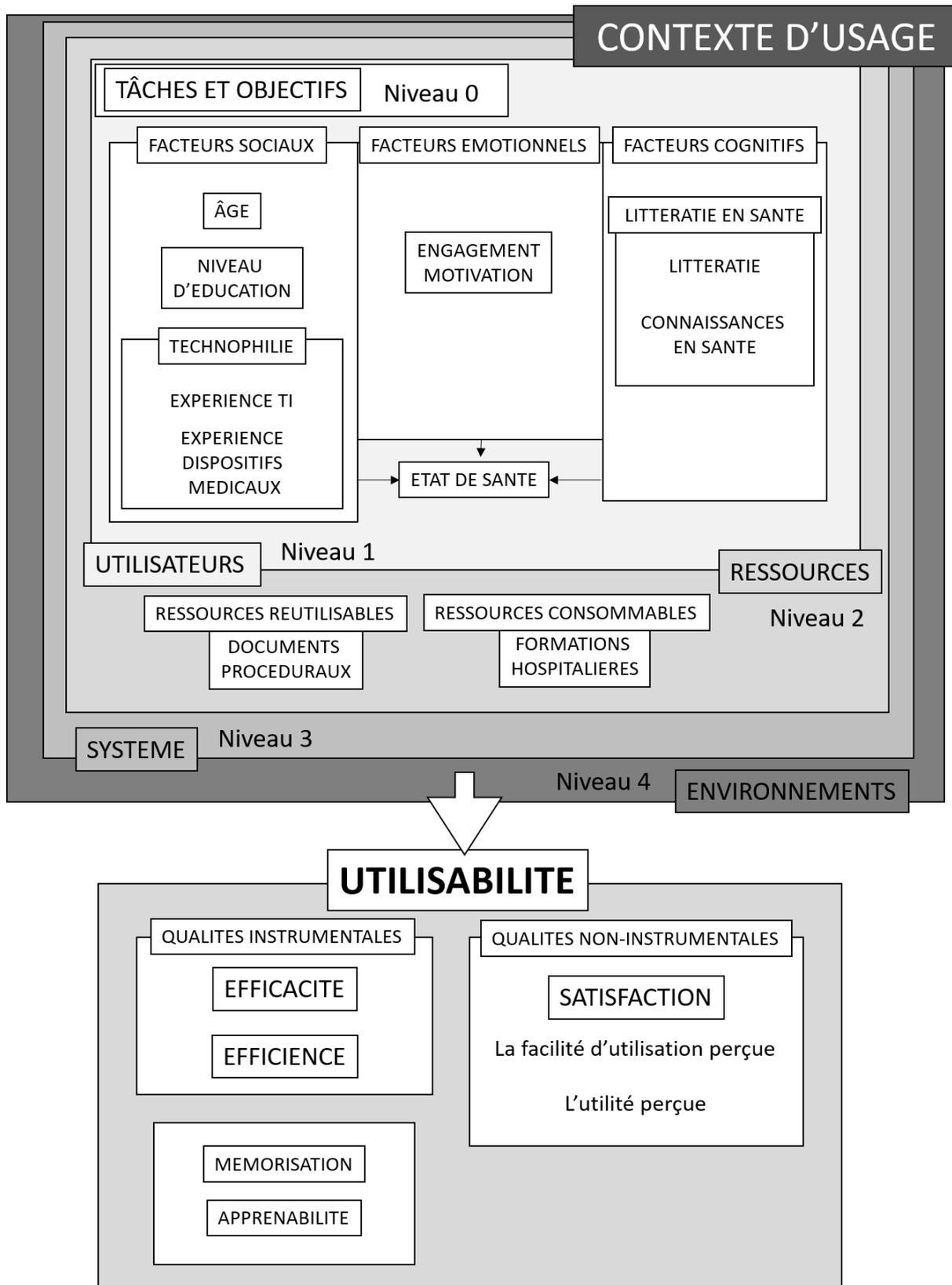


Figure 21. Synthèse des éléments du contexte d'usage (tâche, utilisateur, ressources et environnements) du modèle de l'évaluation de l'utilisabilité établi sur la base de notre revue de littérature.

CHAPITRE 3 : PROBLÉMATIQUE ET DÉMARCHE DE RECHERCHE

1. Problématique et objectif

Après une synthèse des éléments théoriques et méthodologiques permettant de rappeler les manquements de la littérature scientifique dans le domaine de la conception et de l'évaluation de l'utilisabilité des dispositifs connectés en santé, nous poserons l'objectif et la problématique de cette thèse.

Nous avons vu que l'utilisabilité était une notion centrale dans la démarche de conception centrée utilisateur (ISO 9241-210, 2019) au même titre que l'utilité et l'acceptabilité. Les chercheurs s'accordent sur le fait que l'utilisabilité est multidimensionnelle et que chacune de ses métriques doit être mesurée de manière indépendante. Cette propriété a été de nombreuses fois définie et/ou modélisée (e.g., Shackel, 1991 ; Nielsen, 1993 ; ISO 9241-11, 2018). Ces modèles ont ensuite été adaptés à des domaines d'applications variés (e.g., Yen, 2010). Les métriques récurrentes constituant ces modèles sont principalement composées de *l'efficacité*, de *l'efficience* et de la *satisfaction* même si les notions de *facilité d'apprentissage* (*apprenabilité/apprentissage*) et de *mémorisation* sont également citées. Pourtant, la notion d'utilisabilité souffre depuis sa création de l'absence de consensus sur sa définition, ce qui nuit à l'accumulation et à l'organisation des connaissances disciplinaires. Nous retrouvons ce débat au sein même des définitions des métriques pour lesquelles les auteurs ne s'accordent pas toujours. L'*efficacité* peut aussi être définie par *le taux d'erreurs* et *l'efficience* par la *charge de travail*. Enfin, la *satisfaction*, représentant la métrique la plus répandue dans les modèles d'utilisabilité, est probablement la plus complexe à définir du fait que certains auteurs n'intègrent que l'esthétique et le confort sans aborder les éléments émotionnels issus des travaux de l'expérience utilisateur. L'utilisabilité est également définie par son contexte d'usage (*la tâche, l'utilisateur, les ressources et l'environnement*) d'après la norme ISO 9241-11 (2018). Nous avons vu que chacun de ces éléments du contexte d'usage pouvait avoir un impact sur l'utilisabilité.

D'un point de vue méthodologique, les évaluations de l'utilisabilité sont aussi définies par son intégration dans le cycle de conception d'un produit. Différents types d'évaluations de l'utilisabilité peuvent être mobilisés en fonction de l'avancement du processus de conception

et il convient de définir les modalités précises de ces évaluations. La littérature suggère le recours à des méthodes mixtes associant des outils permettant un recueil de données objectives, subjectives, qualitatives et quantitatives (e.g., Jaspers et al. 2009). Malgré ces recommandations, les études d'utilisabilité restent discutables dans le domaine de la e-santé.

Nous avons vu que ce domaine émergent était en pleine croissance. Pourtant, les défauts d'utilisabilité pouvant impacter la sécurité du patient sont sans cesse mentionnés par la communauté scientifique. Afin de remédier à ces défauts, il est nécessaire d'améliorer l'utilisabilité des dispositifs connectés en santé, ce qui nécessite d'avoir un cadre cohérent et d'améliorer nos connaissances sur ce sujet. À travers une revue de la littérature, que nous avons menée, recensant 79 études sur des évaluations de l'utilisabilité de système en santé pour le domicile patient, nous avons mis en évidence que l'ensemble des connaissances créées sur le sujet sont éparpillées. En effet, ces travaux ne reposent pas systématiquement sur les mêmes définitions opérationnelles de l'utilisabilité, ce qui déstabilise et fracture davantage ce domaine déjà fragilisé. Cette carence structurelle révèle un constat clair et consensuel : le manque de connaissances structurées dans le domaine des technologies de la santé nuit *in fine* à l'acceptabilité et à l'acceptation de ces objets au domicile des patients. Ces manquements renvoient directement à un défaut de standardisation des évaluations et à des connaissances insuffisantes sur l'impact du contexte d'usage sur l'utilisabilité. Pour cela, deux arguments principaux permettant d'améliorer l'utilisabilité semblent émerger : (1) l'utilisation d'un cadre conceptuel de l'utilisabilité appliqué au domaine de la e-santé (*quoi évaluer ?*) et (2) le recours à une démarche d'évaluation standardisée intégrée au processus de conception (*quand évaluer ?*). Il est donc nécessaire d'apporter des éclaircissements sur le choix des métriques, des méthodes, des outils, du contexte d'usage sur l'utilisabilité (*comment évaluer ?*) en fonction du cycle de conception d'un dispositif connecté en santé afin de mieux cibler les principales actions à optimiser pour rendre utilisables ces dispositifs pour chacun.

L'objectif de cette thèse est, par conséquent, de proposer un cadre théorique et méthodologique traitant explicitement de l'évaluation de l'utilisabilité des dispositifs connectés en santé destinés au domicile du patient.

Pour ce faire, nous chercherons à répondre à la question de recherche suivante :

Comment assister et améliorer la conception par l'**évaluation** de l'**utilisabilité** de **dispositifs connectés en santé** pour le **patient** à son **domicile** dans un **contexte d'usage** ?

Dans cette thèse, l'**utilisabilité** est définie sur la base de **trois métriques principales indépendantes** (*l'efficacité, l'efficience et la satisfaction*) et de **deux métriques connexes indépendantes** (*l'apprenabilité et la mémorisation*) inspirées des définitions opérationnelles de la norme ISO 9241-11 (2018) et des modèles de Shackel (1991) et Nielsen (1993). Dit autrement, nous estimons que l'utilisabilité est une propriété caractérisée par des qualités instrumentales et non instrumentales.

Nous considérons qu'une **évaluation** de l'utilisabilité est systématiquement associée à des évaluations suivant une structure itérative (boucles) implantée dans le processus de conception d'un produit. Au fur et à mesure de ces itérations, l'utilisabilité du dispositif peut évoluer du fait de l'état du dispositif mais également des compétences des utilisateurs qui vont progressivement également évoluer. Cette propriété a, en ce sens, un caractère **évolutif**.

Nous définissons les **dispositifs connectés en santé** comme un instrument de santé ne comprenant pas (encore) de marquage de conformité européenne. Le dispositif étant connecté, nous considérons qu'il appartient au domaine émergent de la e-santé. Plus précisément, il s'inscrit dans le domaine de la m-santé (m pour mobile) ou de la u-santé (u pour ubiquitaire) car ces dispositifs ont pour objectif d'être utilisés au domicile du patient.

Nous définissons ce **contexte d'usage par quatre composants** : *les tâches, les utilisateurs, les ressources, et l'environnement*. Nous avons détaillé ce contexte d'usage appliqué à la santé en considérant les caractéristiques utilisateurs suivant trois facteurs (social, émotionnel et cognitif) et les ressources suivant deux facteurs (réutilisable : les documents procéduraux, et consommable : les formations hospitalières). L'utilisateur final étant le patient sortant de chirurgie ambulatoire, cela implique une grande diversité des profils de ces

utilisateurs. Nous considérons l'utilisabilité comme **située**, car le contexte d'usage impacte l'ensemble des métriques de l'utilisabilité. Le contexte d'usage *tâche-utilisateurs-système-ressources-environnement* doit être investigué à chaque étape du cycle itératif. Dans le cas de cette thèse, l'environnement se trouve être le contexte de la chirurgie ambulatoire, c'est-à-dire, le milieu hospitalier (service de chirurgie ambulatoire) suivi du domicile du patient. Cela implique que pour utiliser le dispositif connecté en santé, ce dernier peut être seul ou accompagné d'un aidant à son domicile.

Pour avancer dans cette problématique, nous avons déployé la démarche de recherche décrite ci-après.

2. Hypothèses et démarche de recherche

Une des particularités de cette thèse est qu'elle s'inscrit dans le cadre d'un projet réel de conception : le projet *Smart Angel*. Dans cette partie, nous rappelons d'abord le processus de conception du dispositif *Smart Angel* en ambulatoire en détaillant le dispositif et son contexte législatif avant d'aborder comment nous avons implanté notre modèle, que **nous avons nommé Home medicAl coNnected Devices usabiLity Evaluation moDel ou HANDLED**, dans ce processus.

2.1. Le processus de conception du dispositif *Smart Angel* en ambulatoire

Pour rappel, le projet *Smart Angel* en ambulatoire a pour objectif de développer un système de surveillance de type dispositif médical connecté (visant un dépistage précoce des complications postopératoires grâce à un kit de suivi à distance). Cette surveillance permet au patient de maintenir une connectivité avec l'hôpital depuis son domicile. En effet, les patients nouvellement opérés devront utiliser (seuls ou accompagnés) trois fois par jour pendant cinq jours le dispositif *Smart Angel*. L'hôpital suivra à distance leur état de santé *via* différentes données médicales au moyen d'un terminal infirmier. Cela permettra à l'équipe hospitalière de vérifier à distance ces données envoyées par les patients et de prendre les décisions en fonction de la criticité de celles-ci.

Le patient se verra remettre un kit composé d'appareils de surveillance (tensiomètre et oxymètre de pouls) avec application d'affichage numérique sur écran tactile. Les deux objets

connectés (tensiomètre et oxymètre de pouls) de Conformité Européenne (marquage CE) sont considérés comme des dispositifs médicaux, d’après la norme ISO 13485 (2016) et sont déjà commercialisés pour le grand public. Les fiches produits de ces deux dispositifs sont disponibles en annexe (cf. Annexe 5 : Fiche produit iHealth Sense Tensiomètre poignet connecté BP7, p. 368 et Annexe 6 : Fiche produit Oxymètre de pouls connecté iHealth Air P03M, p. 369). Les deux objets connectés sont non invasifs et appartiennent à la classe IIa (degré moyen de risque). Cependant, **l’ensemble du système *Smart Angel* sera classé IIb** (potentiel élevé de risque) après sa commercialisation. Le dispositif *Smart Angel* est avant tout un outil mettant en relation les soignants et les patients intégrant le monitoring du patient grâce aux objets connectés (permettant la collecte de données physiologiques) et une application (« *My Angel* ») développée sur tablette tactile (permettant de visualiser ces données et de collecter de données subjectives via un questionnaire en ligne). Le tout est connecté à un terminal infirmier, sur lequel le personnel hospitalier traite le signal. Cet ensemble connecté compose le système *Smart Angel* (Figure 22).

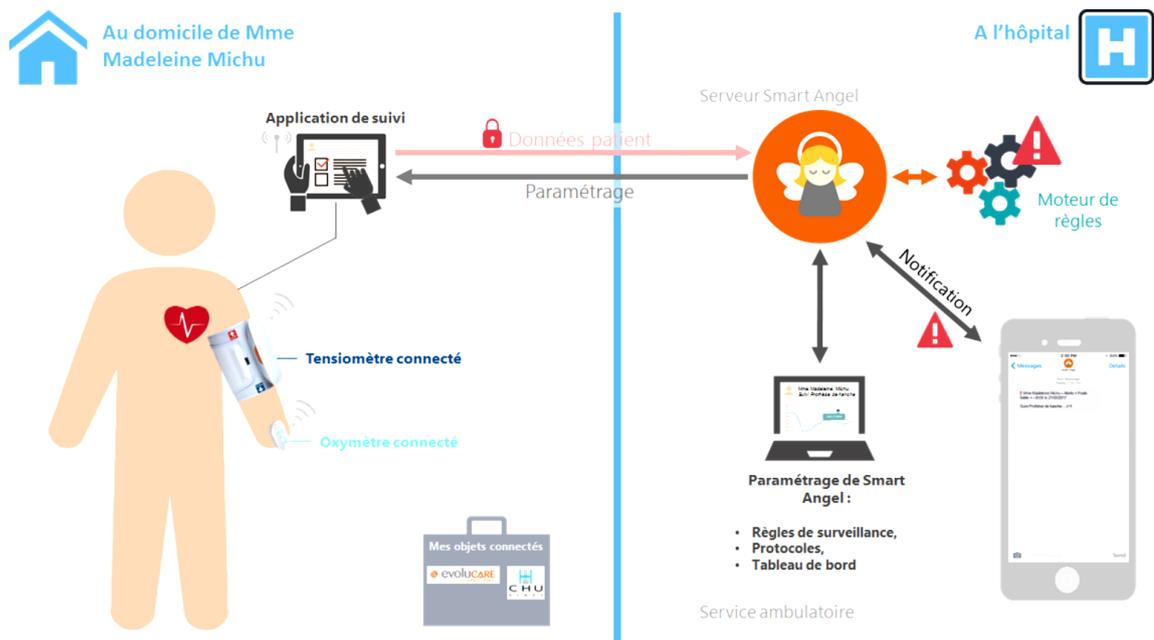


Figure 22. Présentation du système *Smart Angel* et de chacun de ses éléments : applications de suivi (*My Angel*), Tensiomètre et oxymètre de pouls connectés pour le patient à son domicile et terminal infirmier pour le personnel hospitalier à l’hôpital.

Ce dispositif connecté doit subir des évaluations ergonomiques et cliniques obligatoires imposées par la Haute Autorité de la Santé ou HAS nécessitant ainsi le recours à des Comités de Protection des Personnes ou CPP avant de pouvoir être considéré comme un dispositif médical connecté (cf. Partie I, Chapitre 2, 1.2. La classification et la conception des dispositifs médicaux, p. 55). Avant cela, des maquettes de l'application « *My Angel* » et du terminal infirmier, issues d'une analyse de besoins ont été complétées d'objets connectés sélectionnés par les partenaires du CHU de Nîmes. L'ensemble de ce dispositif constitue le prototype V1 pouvant faire l'objet d'une évaluation clinique après dépôt et validation d'un CPP. Cette étude clinique est nécessaire afin de valider cliniquement ce prototype avant sa commercialisation. Dès que le dispositif sera commercialisé, l'entreprise Evolucare envisage de continuer d'implanter des améliorations par le biais de nouvelles fonctionnalités de manière itérative avec l'objectif de rendre personnalisable le dispositif au patient en fonction de sa pathologie. La Figure 23 ci-dessous montre le déroulement du projet de conception peu de temps avant le début de la thèse jusqu'à la fin de cette dernière.

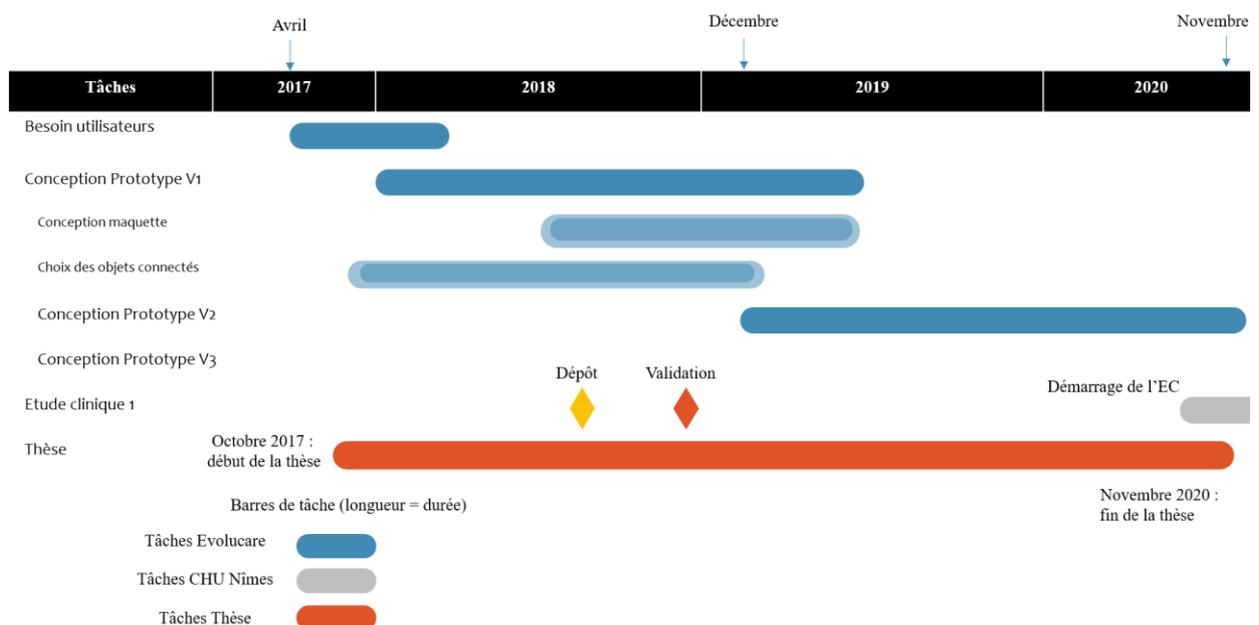


Figure 23. Processus de conception du dispositif *Smart Angel* ambulatoire en fonction du timing de la thèse.

2.2. L'intégration de notre démarche de recherche au sein du processus de conception de *Smart Angel*

Cette thèse s'intègre dans un contexte de conception de produit où les parties prenantes issues des disciplines diverses et variées se côtoient et collaborent dans un but commun, celui de proposer un produit innovant répondant aux besoins des futurs utilisateurs (Sagot et al.,

2003). Il est d'ailleurs important de préciser **que ce travail de thèse se positionne en phase amont de ce processus de conception**. À ce stade, les chargés de projet représentés par l'entreprise Evolucare possèdent un cahier des charges précis de conception co-rédigé avec les acteurs du CHU de Nîmes et du CRP-CPO et réalisent les premières maquettes et prototypes du dispositif. Ainsi, **notre travail de thèse s'intègre entre les phases de l'analyse de besoin et de l'étude clinique** (Figure 24).

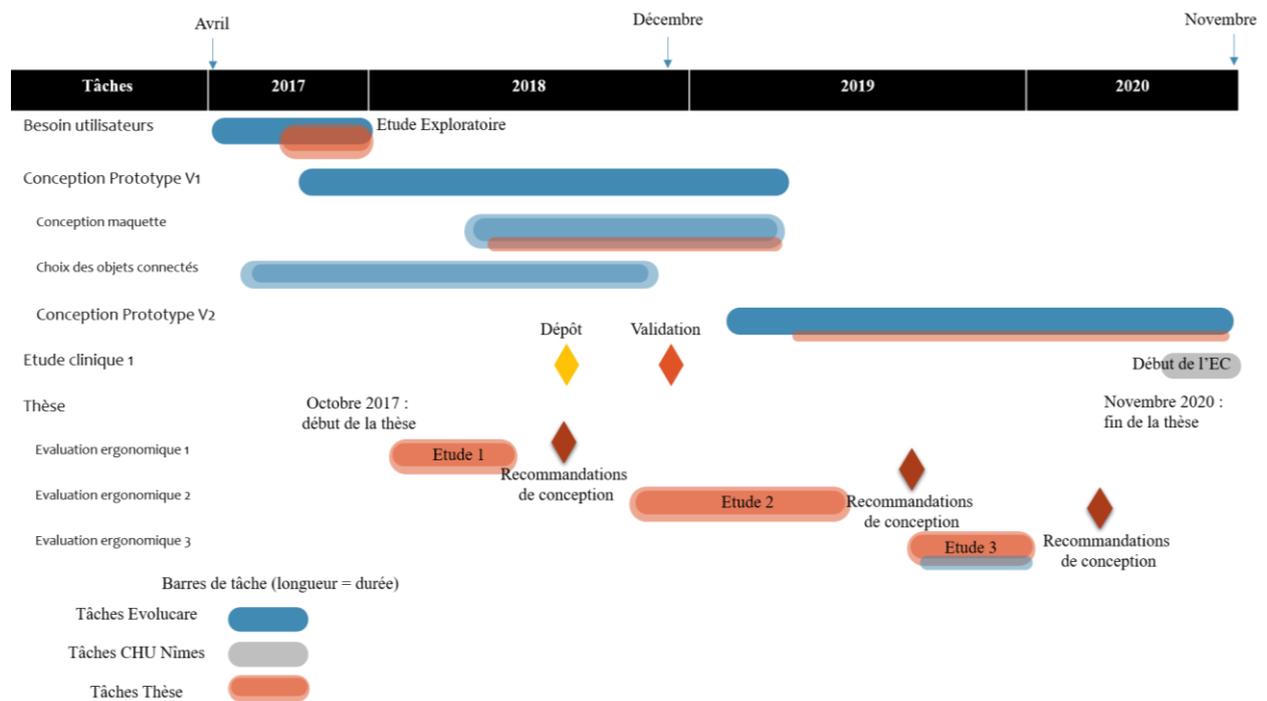


Figure 24. Processus de conception du dispositif *Smart Angel* ambulatoire en fonction des tâches et études réalisées dans la thèse.

Nous avons donc **proposé trois itérations afin d'assister l'activité de conception du dispositif *Smart Angel***. Chaque itération correspond à une étude. Chaque **étude** est associée à un niveau défini dans le modèle HANDLED que nous proposons, inspiré de celui de Yen et Bakken (2012) et de la norme ISO 9241-11 (2018). Ces niveaux sont cumulatifs. Chaque nouvel élément intégré dans une étude améliore le précédent jusqu'à inclure l'ensemble de ces éléments dans un modèle plus complet : le niveau 0 est assimilé à la **tâche** et compose l'analyse de besoin, le niveau 1 intègre les **utilisateurs** finaux dans la tâche, le niveau 2 ajoute au niveau 1 les **ressources** disponibles, le niveau 3 ajoute au niveau 2 le **système** (composé d'objet connecté, d'une tablette reliée à un terminal piloté) et le niveau 4 ajoute au niveau 3 certaines caractéristiques de l'**environnement** hospitalier. Ces études sont soutenues par une hypothèse générale intrinsèque à l'étude construite sur la base de **questions opérationnelles** de recherche

relatives aux cinq niveaux. La Figure 25 schématise ces quatre niveaux, les questions opérationnelles et les hypothèses générales associées.

Nous avons mené 5 études (dont une étude exploratoire et une étude décomposée en deux volets) en fonction de l'avancement du projet *Smart Angel* :

- **Étude exploratoire** (cf. Partie II, Chapitre 1, p. 112) : L'objectif de cette étude était d'analyser les besoins des futurs utilisateurs et d'appréhender le terrain (service de chirurgie ambulatoire) avant intégration du dispositif *Smart Angel*. Nous avons fait l'hypothèse que la tâche envisagée avec le dispositif *Smart Angel* devait être investiguée pour mieux assister l'équipe de conception. Cette étude correspond au **niveau 0** de notre modèle et n'intègre aucun des éléments du dispositif *Smart Angel* ;
- **Étude 1** : Cette étude est divisée en deux volets intégrant uniquement l'évaluation des objets connectés (dispositifs médicaux) du dispositif *Smart Angel* présélectionnés par le CHU de Nîmes.
 - o Ce premier volet, l'**étude 1a** (cf. Partie II, Chapitre 2, p. 132) correspond à l'interaction entre l'utilisateur et la tâche (**niveau 1**). Dans cette étude, nous nous sommes intéressées à l'impact des caractéristiques utilisateurs (en nous centrant sur le niveau de connaissances antérieures en santé) sur l'utilisabilité ;
 - o Ce deuxième volet, l'**étude 1b** (cf. Partie II, Chapitre 3, p. 160) correspond à l'interaction entre l'utilisateur, la tâche et les ressources. Pour cela, nous avons investigué l'impact des ressources réutilisables (les documents procéduraux pouvant se soustraire au contexte hospitalier) sur l'utilisabilité (**niveau 2**) ;
- **Étude 2** (cf. Partie II, Chapitre 4, p. 189) : Cette étude vise l'évaluation du prototype V1 de *Smart Angel* et correspond à l'interaction entre l'utilisateur, la tâche, les ressources et le système (**niveau 3**). Nous avons fait l'hypothèse que la complexité intrinsèque de l'ensemble du système *Smart Angel* impacte l'utilisabilité ;
- **Étude 3** (cf. Partie II, Chapitre 5, p. 226) : Cette dernière étude vise à l'évaluation du prototype V2 de *Smart Angel* et correspond à la dernière strate de notre modèle soit l'interaction entre l'utilisateur, la tâche, les ressources, le système et l'environnement (**niveau 4**). Nous avons investigué l'impact du contexte hospitalier sur l'utilisabilité.

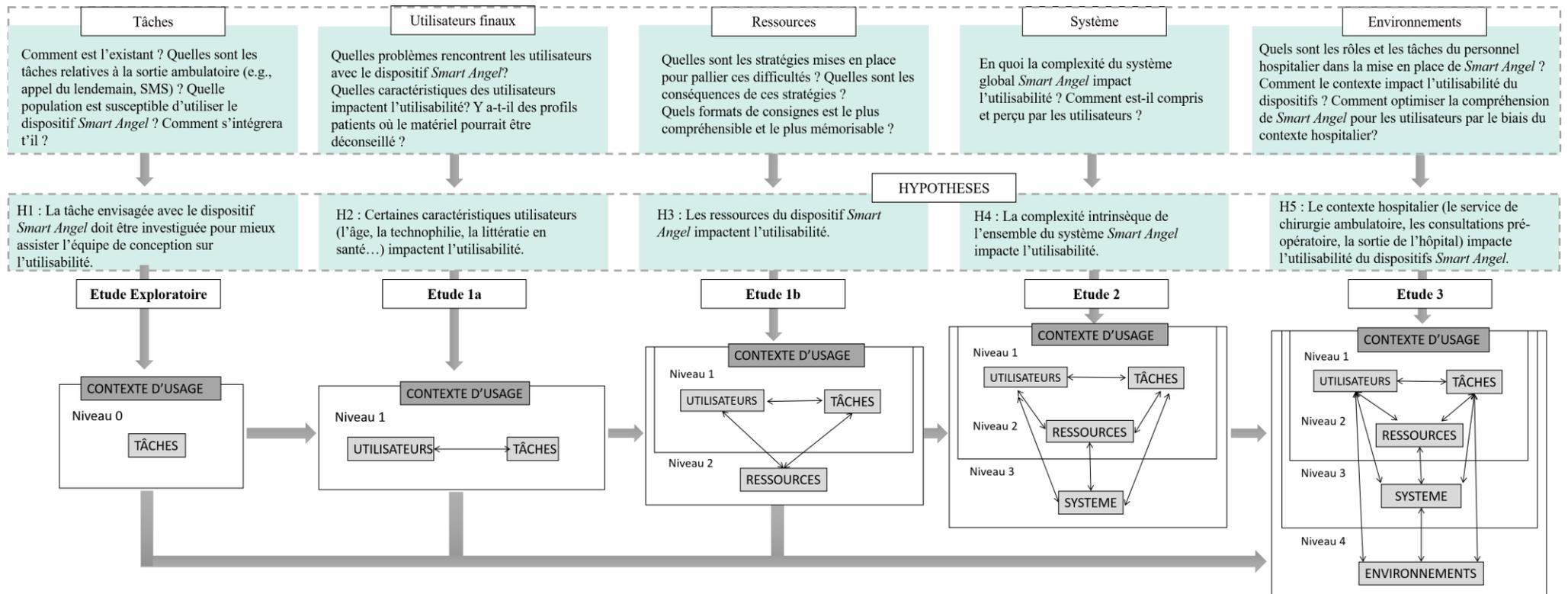


Figure 25. Représentation schématique des cinq hypothèses générales, intrinsèques à chaque étude de cette thèse, construites sur la base de questions opérationnelles de recherche relatives aux cinq niveaux (tâches, utilisateurs, ressources, système, environnements).

2.3. Les choix méthodologiques

Compte tenu des constats méthodologiques développés (cf. Partie I, Chapitre 1, 4.2. Trois lieux de recueil de données déterminés par trois approches épistémologiques : laboratoire, terrain, *in situ*, p. 43 et 4.3. Les types, les méthodes et les outils de l'évaluation de l'utilisabilité, p. 44), nous avons réalisé des choix méthodologiques pour chaque étude en précisant à la fois (1) **une mise en situation**, c'est-à-dire le contexte environnemental dans lequel se déroule l'étude (e.g., laboratoire) et (2) **des outils et méthodes permettant un recueil de données mixtes pour étudier l'utilisabilité**. Ces éléments sont détaillés dans la Figure 26.

Les études 1a et 1b se déroulent en laboratoire dans un environnement particulièrement contrôlé. Afin d'évaluer de façon indépendante les métriques de l'utilisabilité (*efficacité*, *efficience* et *satisfaction*) et de privilégier des études mixtes, les études 1a et 1b comprennent *a minima* : des observations instrumentées par vidéo visant à collecter des données quantitatives et objectives, un questionnaire (SUS ou System Usability Survey) permettant un recueil quantitatif de mesures subjectives.

L'étude 2 se déroule dans un contexte naturel pour les utilisateurs, car il s'agit de leur domicile ou de leur lieu de travail, mais ceux-là n'ont pas transité par le service de chirurgie ambulatoire. Cet environnement contrôlé est aussi considéré comme ce qui relève du « laboratoire » d'après Perrot et al. (2014). En effet, nous imposons notre présence, la présence d'une caméra et des consignes précises à l'utilisateur. Nous avons également ajouté une mise en situation par le biais de mise en scène des personas via des scénarios utilisateurs sur un format audiovisuel afin de projeter l'utilisateur dans une « *utilisation imaginaire* ». Comme pour les études 1a et 1b, l'étude 2 repose sur des observations instrumentées par vidéo visant à collecter des données quantitatives et objectives, un questionnaire (SUS) permettant un recueil quantitatif de mesures subjectives, des entretiens permettant un recueil de données qualitatives et subjectives. Nous avons ajouté à cette méthodologie, la méthode *think aloud* favorisant la collecte de données qualitatives et objectives.

Enfin, l'étude 3 se rapproche davantage d'une situation réelle dans la mesure où il s'agit d'une étude faisant appel à une « *utilisation simulée* » menée au SimUsanté d'Amiens dans lequel nous avons reconstitué, en partie, un service de chirurgie ambulatoire et son personnel. Les participants ont ensuite manipulé le dispositif dans une maison reconstituée (SimULogis). La méthodologie choisie pour cette étude s'articule autour d'observations instrumentées par vidéo visant à collecter des données quantitatives et objectives, du questionnaire SUS offrant

un recueil quantitatif de mesures subjectives, des entretiens et focus group permettant un recueil de données qualitatives et subjectives.

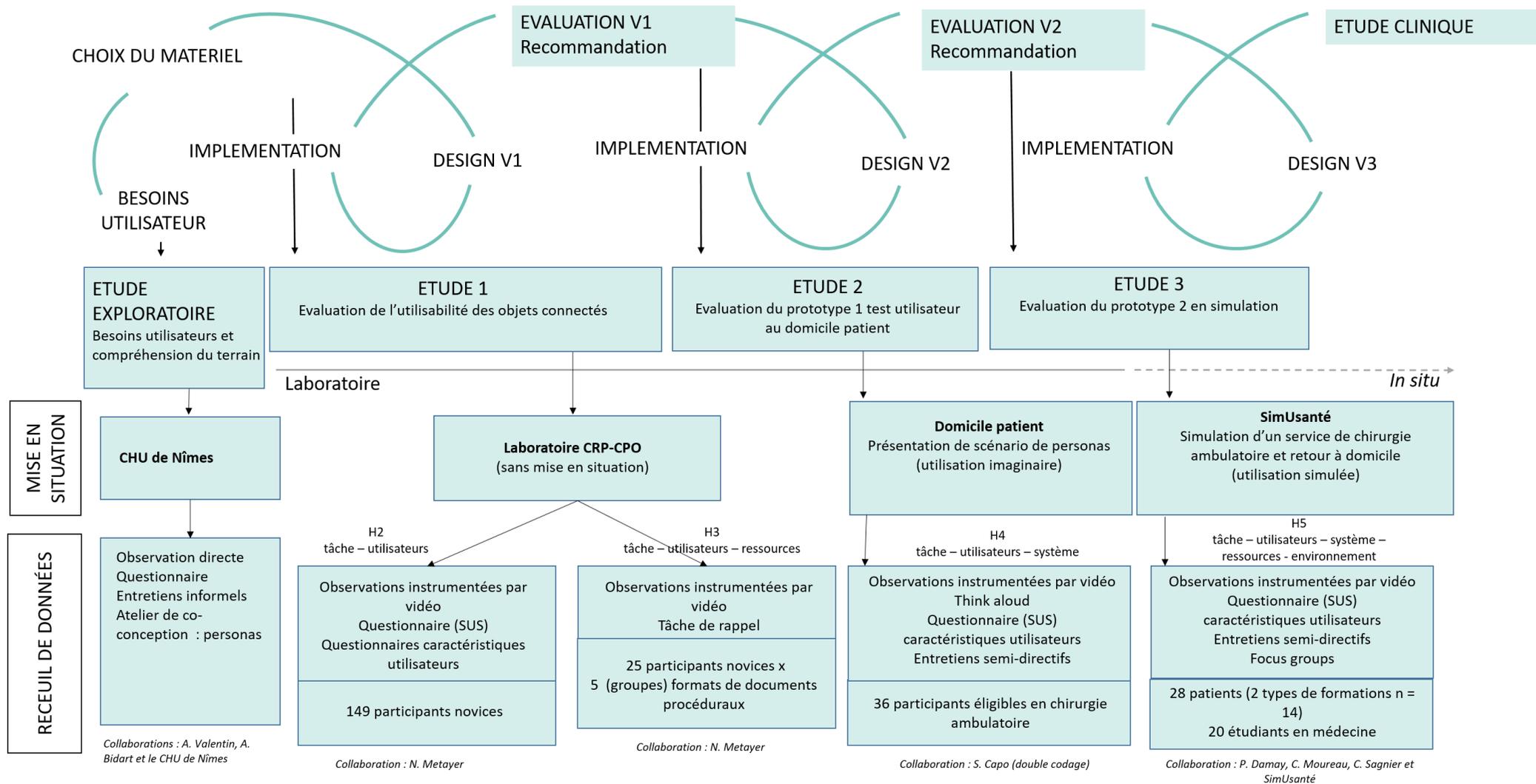


Figure 26. Démarche méthodologique de recherche intégrant les études de la thèse en fonction de la mise en situation et du recueil de données utilisés. Ces études évoluent vers un environnement de plus en plus proche d'un contexte réel dit « in situ ».

3. Synthèse du chapitre 3

Après avoir investigué les manquements de la littérature dans le domaine, nous avons pu formuler notre problématique de recherche : **Comment assister et améliorer la conception par l'évaluation de l'utilisabilité de dispositifs connectés en santé pour le patient à son domicile dans un contexte d'usage ?**

Cette problématique a été appréhendée selon une démarche de recherche basée sur le principe de « *quand évaluer quoi ?* ». Cette dernière s'articule autour de **cinq hypothèses générales de recherche**. Chacune d'entre elles oriente une de nos **cinq études** (étude exploratoire, étude 1a, étude 1b, étude 2 et étude 3).

Chaque étude intègre au fur et à mesure l'ensemble des composants du contexte d'usage (tâches, utilisateurs, ressources, système, environnements) dans des cycles itératifs d'évaluation de l'utilisabilité.

Dans chacune des études, l'utilisabilité est mesurée *a minima* sur la base des **trois métriques principales** définies dans la norme (*efficacité, efficacité et satisfaction*) et mobilise des outils et méthodes permettant **une collecte de données mixtes** dans un milieu se rapprochant progressivement d'un environnement *in situ*.

Cette stratification du contexte d'usage associée aux métriques compose le modèle d'évaluation de l'utilisabilité que nous avons nommé **HANDLED (Home medicAl coNnected Devices usabiLity Evaluation moDel)**.

PARTIE II : PARTIE EMPIRIQUE

CHAPITRE 1 : ÉTUDE EXPLORATOIRE. ANALYSE DE BESOINS ET COMPRÉHENSION DU TERRAIN

1. Préambule

1.1. Contexte et objectif industriels

L'objectif industriel de cette étude était d'apporter des connaissances du terrain afin de concevoir les premières maquettes du dispositif *Smart Angel* et d'aider à la sélection des futurs objets connectés. Pour cela, nous avons participé à l'analyse de besoins déjà initiée par la collaboration avec une consultante en ergonomie, Annette Valentin, avant le commencement de cette thèse qui compose la première étape dans la conception du dispositif (Figure 27). Nous avons réalisé des observations dans le service de chirurgie ambulatoire du CHU de Nîmes pendant 9 jours (cumulés) et des entretiens informels avec le personnel hospitalier (anesthésistes, chirurgiens, personnel infirmier, aides-soignants) et les patients. Sur la base des données collectées, nous avons pu concevoir cinq personas lors d'un atelier (Loup-Escande et al., 2019). Ces personas ont été utiles pour la conception des maquettes, mais aussi en phase d'évaluation (cf. Partie II, Chapitre 4, p. 189). Enfin, les données du terrain ont aussi permis de mieux comprendre la future intégration du dispositif dans son environnement naturel et de sélectionner des éléments clés du parcours de soin pouvant jouer un rôle crucial dans cette intégration (e.g., la consultation anesthésique). Cela a permis de recréer l'environnement en simulation (cf. Partie II, Chapitre 5, p. 226). Cette étude exploratoire nous sert ainsi de socle contextuel pour la suite des études présentées dans le cadre empirique.

1.2. Contexte et objectif théoriques

Cette étude réalisée correspond à la première étape du cycle de conception du dispositif *Smart Angel* soit le niveau 0 du processus de conception du dispositif *Smart Angel* (Figure 27). L'objectif théorique de cette étude était de réaliser des analyses prospectives sur les profils des patients en chirurgie ambulatoire afin de prendre connaissance des futurs utilisateurs du dispositif *Smart Angel*. Pour cela, nous avons réalisé une analyse démographique rétrospective couplée à une revue de la littérature sur la conception de personas dans le milieu de la santé et nous avons pu synthétiser ces données pour concevoir un template de personas patient.

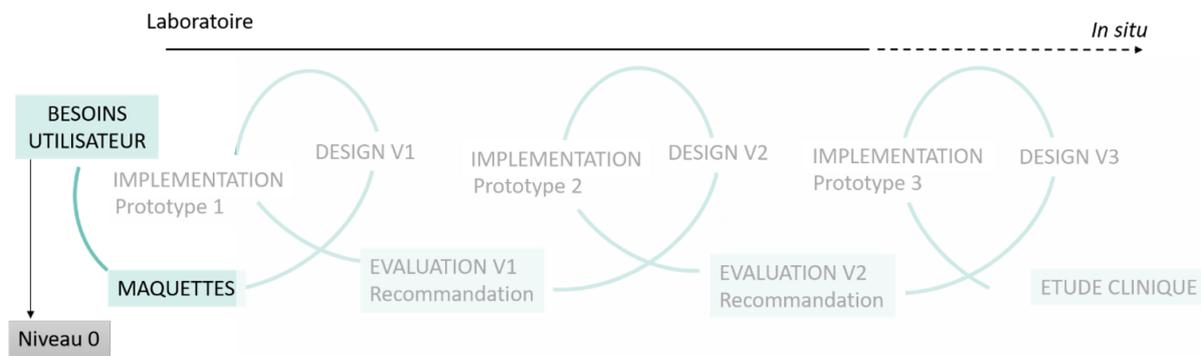


Figure 27. Étude exploratoire correspondant au niveau 0 du processus de conception du dispositif *Smart Angel*.

2. Introduction

L'une des phases amont de conception d'un dispositif technologique innovant correspond à l'analyse de besoins. L'objectif de ce type d'analyse est de proposer une innovation en adéquation avec les attentes des utilisateurs (Griffin, 2004) et de l'organisation dans laquelle cette dernière s'implantera. L'intégration de l'utilisateur dans le processus de conception d'une innovation est une des clés du succès de ces produits. Cependant, cette intégration en phase amont de conception reste controversée dans la mesure où les patients ne sont pas toujours capables de se projeter sur un matériel inexistant et peuvent parfois induire les concepteurs en erreur (Alam, 2007). C'est pourquoi nous avons dans un premier temps rassemblé des informations sur le contexte de la chirurgie ambulatoire en nous attardant sur les profils des patients dans ces services. Dans un deuxième temps, nous avons conçu un template de personas représentatif de ces patients précédemment examinés afin d'aider l'entreprise dans la conception des premières maquettes *Smart Angel*.

2.1. La chirurgie ambulatoire en France et leurs patients

La chirurgie ambulatoire est en pleine expansion d'après le rapport de la DREES (Lefebvre-Hoang & Yilmaz, 2019). Pourtant cette pratique, jugée plus rentable, souffrirait d'une mauvaise gestion postopératoire. Pour comprendre ce phénomène, nous nous sommes d'abord interrogées sur le déroulement d'un service de chirurgie ambulatoire, les critères

d'inclusion, les actes réalisés et les profils des patients que nous présentons dans les sections suivantes issues de la littérature sur le sujet.

2.1.1. Parcours patient en ambulatoire

La chirurgie ambulatoire se décompose en trois phases (Bert et al., 2014; Mahieu & Raffy-Pihan, 1997, Figure 28) : une phase avant opération pour la préparation à l'intervention, une phase le jour de l'intervention et une phase après opération pour le suivi postopératoire. Ce parcours peut avoir quelques faibles variantes en fonction de l'organisation de chaque centre hospitalier, mais se trouve être majoritairement dans cette configuration du fait de la législation en vigueur. Par exemple, le jour de l'intervention, dit « J-0 », le patient se fait opérer et ne reste pas plus de 12h à l'hôpital d'après l'article D. 6124-301-1 du code de la santé publique²⁶. Le patient repart chez lui le soir même avec des consignes postopératoires classiques pouvant varier en fonction de l'opération subie. Par exemple, un suivi par des infirmiers libéraux peut être mis en place (Bert et al., 2014 ; Mahieu & Raffy-Pihan, 1997). Les jours suivants, le personnel infirmier interroge le patient (par téléphone, ou par SMS en fonction de l'organisation hospitalière) afin de détecter toute(s) anomalie(s) ou tout/s événement(s) indésirable(s).

2.1.2. Les critères d'éligibilité à la chirurgie ambulatoire

Pour pouvoir avoir recours à la chirurgie ambulatoire, il faut d'abord être éligible en fonction des critères imposés par la Haute Autorité de la Santé ou HAS. Ceux-ci reposent sur l'analyse bénéfique/risque pour le patient, le suivi des soins, la complexité de l'intervention et de l'anesthésie. Guidat (2010) énumère, dans les annales françaises d'anesthésie et de réanimation, 71 critères d'éligibilité pour l'acceptation du patient en chirurgie ambulatoire. Dans le rapport d'évaluation technologique, l'HAS/ANAP (2014) les décrit comme des critères médicaux et psycho-socio-environnementaux.

²⁶ L'article D. 6124-301-1 du code de la santé publique dispose que les prises en charge dispensées dans les structures de chirurgie ambulatoire sont « d'une durée inférieure ou égale à douze heures » et que les « prestations délivrées équivalent par leur nature, leur complexité et la surveillance médicale qu'elles requièrent à des prestations habituellement effectuées dans le cadre d'une hospitalisation à temps complet ».

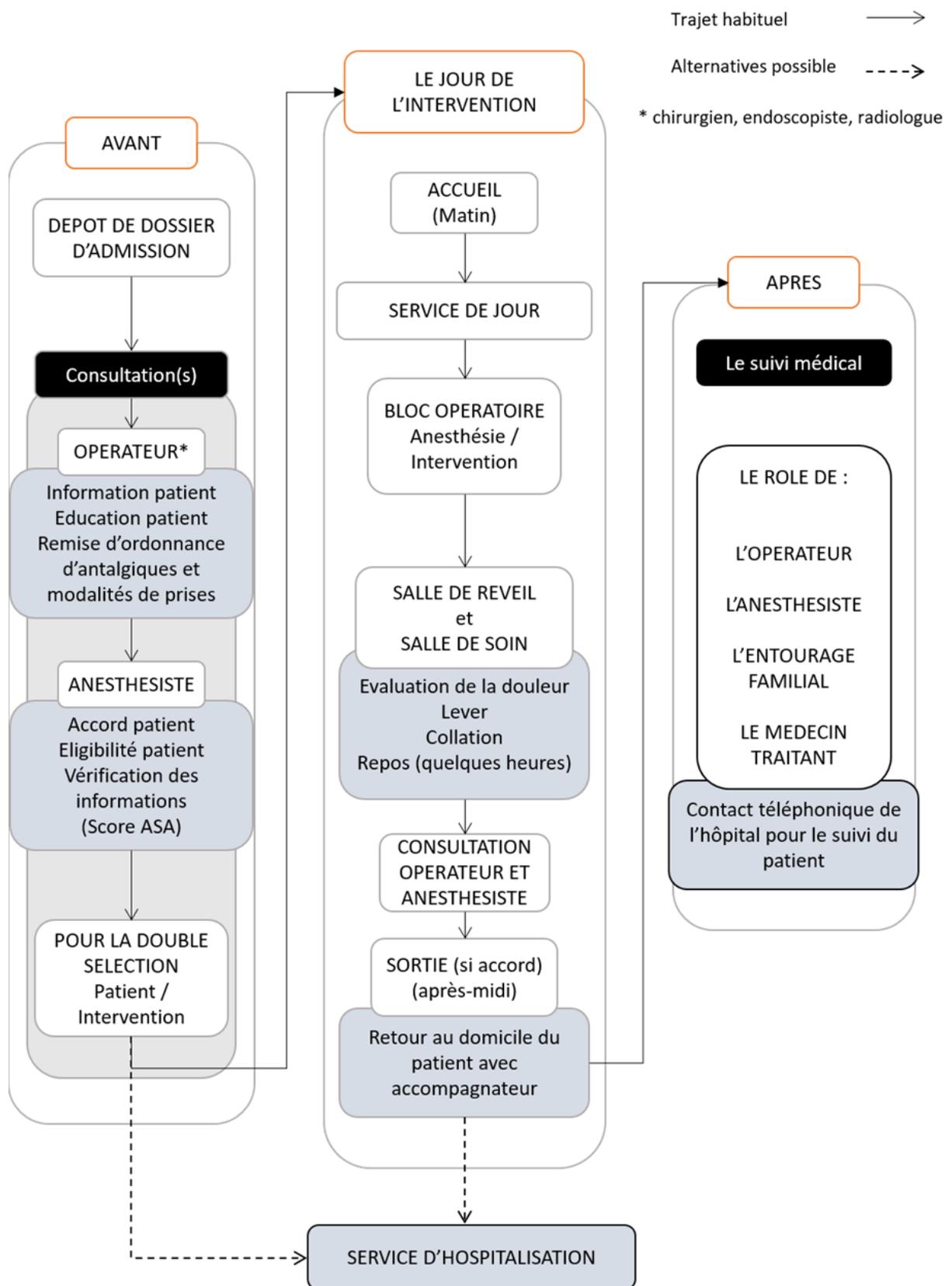


Figure 28. Parcours patient en unité de chirurgie ambulatoire (Schéma inspiré de Mahieu & Raffy-Pihan, 1997 et des données de Bert et al., 2014).

Une sélection se pratique alors sur le patient en fonction des principaux critères suivants :

- Actes réglés et programmables d'un temps maximum de 1 h 30,
- Sans complication prévisible, notamment sans risque hémorragique majeur ou cardio-respiratoire pour le patient,
- Entraînant une douleur postopératoire pouvant être contrôlée par des procédures analgésiques administrables en toute sécurité à domicile.

Le patient doit lui-même avoir des capacités minimales :

- Préservant une autonomie acceptable à domicile (déambulation minimale, alimentation),
- Compréhension suffisante de ce qui est proposé et aptitude à observer les prescriptions médicales,
- Disponibilité d'une personne accompagnante, responsable et valide,
- Éloignement géographique inférieur à 1h d'une structure de soins adaptée,
- Accès rapide à un téléphone,
- Acceptation du patient.

L'isolement social est un critère rédhibitoire à la chirurgie ambulatoire. On proposera une chirurgie conventionnelle à une personne âgée vivant seule chez elle. D'ailleurs, la chirurgie ambulatoire est largement répandue chez les sujets jeunes (75% des moins de 5 ans) d'après le rapport de l'HAS/ANAP (2014). Nous précisons ces éléments dans la section suivante.

2.1.3. Les actes chirurgicaux les plus pratiqués en ambulatoire

L'Association Française de Chirurgie Ambulatoire (AFCA, 2015) recense une liste comprenant 55 actes différents réalisables en chirurgie ambulatoire. L'Agence Régionale de la Santé (ARS, 2014) a analysé la proportion de ces actes les plus effectués en ambulatoire. On trouve (1) le canal carpien, (2) la cataracte et (3) la chirurgie du ménisque comme étant les actes les plus réalisés en ambulatoire (Figure 29). Cependant, ces résultats ne donnent pas d'information sur l'âge, ni même sur le nombre d'actes réalisés. Ils donnent la proportion de ces actes réalisés en ambulatoire. Il est possible que le nombre de chirurgies de la hernie de l'aine soit plus important que le nombre de ligatures des trompes. Il est donc nécessaire de creuser ces données statistiques pour mieux comprendre les populations présentes en chirurgie ambulatoire.

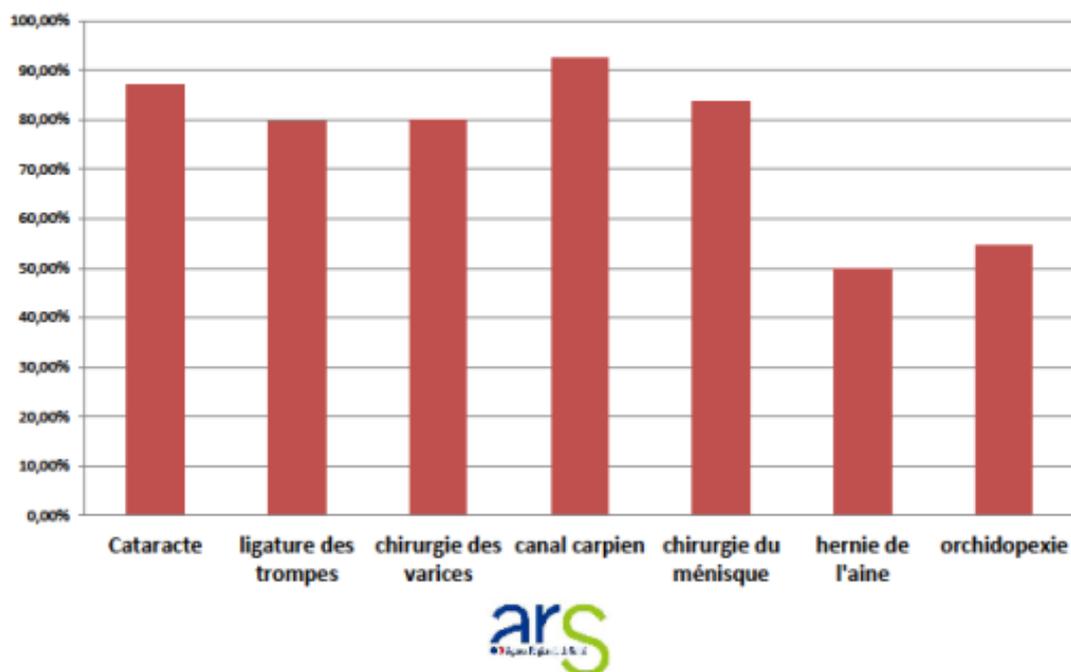


Figure 29. Principaux actes chirurgicaux en ambulatoire (ARS, 2014).

2.1.4. Le profil des patients en chirurgie ambulatoire

D'après l'Agence Technique de l'Information sur l'Hospitalisation ou ATIH, 3 740 056 actes de chirurgie ambulatoire auraient été réalisés en France dans la seule année 2018. Cependant, aucun rapport statistique n'existe à l'heure actuelle sur les profils des patients ayant majoritairement recours à ce type de chirurgie. Nous avons donc prospecté grâce à des déductions réalisées sur les rapports de santé en France (ATIH, 2016b) ainsi que sur la base de données ScanSanté²⁷ regroupant l'ensemble des interventions en chirurgie depuis 2007 jusqu'à 2018 par région de France (Tableau 12).

Notre analyse a révélé que le taux de chirurgie ambulatoire se pratique majoritairement sur des sujets jeunes (moins de 20 ans) en comparaison à la chirurgie traditionnelle. C'est-à-dire que les actes pratiqués sur les jeunes sont plus facilement éligibles à la chirurgie ambulatoire que pour les autres tranches d'âge. Cependant, ce sont les patients de 75 ans et plus qui sont proportionnellement les plus présents dans ce service (44 %) suivi par les patients de moins de 20 ans (31 %) et enfin les patients ayant entre 20 et 74 ans (24,7 %). En revanche, les actes réalisés sur les personnes âgées ne sont pas les mêmes que ceux réalisés sur les autres tranches d'âge de la population.

²⁷ <https://www.scansante.fr/applications/action-gdr-chirurgie-ambulatoire>

Tableau 12. Nombre d'actes en chirurgie ambulatoire en fonction de l'âge (source : Scansanté) et en fonction de l'éligibilité au dispositif *Smart Angel* (cf. Annexe 2 : Critères d'inclusion et d'exclusion de la population pouvant utiliser le dispositif *Smart Angel*, p. 341).

Âge	Nombre d'actes en chirurgie	%	Nombre d'interventions cataracte (approximatif)	% de la répartition des actes hors cataracte (approximatif)	Éligibilité <i>Smart Angel</i>
Moins de 20 ans	1 159 417	31,1 %	~1 159 417	~40,1 %	Non admis (> 18ans)
20 - 74 ans	923 794	24,8 %	~504 908	~17,46 %	admis
75 ans et plus	1 645 624	44 %	~1 227 240	~42,44 %	Non admis (< 80 ans)

Il est important de s'attarder sur les actes réalisés sur ces personnes âgées. En effet, certains de ces actes comme l'intervention de la cataracte, imposant l'interdiction de regarder des écrans après cette opération, sont proscrits dans le projet *Smart Angel*. Ceci pourrait modifier la population cible du futur dispositif. Les critères d'éligibilité à *Smart Angel* sont présentés en annexe (cf. Annexe 2 : Critères d'inclusion et d'exclusion de la population pouvant utiliser le dispositif *Smart Angel*, p. 341). En annexe (cf. Annexe 7 : Prospection sur le nombre d'interventions et l'âge des patients en chirurgie ambulatoire, p. 370), nous présentons le nombre d'interventions chirurgicales et la proportion pratiquée en chirurgie ambulatoire en fonction des principaux actes réalisés. La cataracte (02C05) est, par exemple, l'intervention la plus réalisée en chirurgie ambulatoire. Elle représente à elle seule 22,40 % de ces actes. De plus, nous savons qu'environ 50 % des opérations de la cataracte sont réalisées sur des patients ayant plus de 75 ans (soit 418 384 actes en chirurgie ambulatoire). L'intervention de la cataracte représente ainsi 25,42 % de l'ensemble des actes pratiqués sur les personnes âgées.

D'après le Tableau 12 et un calcul rapide, le nombre très approximatif de personnes pouvant être éligibles au dispositif *Smart Angel* ne représente que 17,46 % des patients admis en ambulatoire (taux minorant car nos calculs ne prennent pas en compte les personnes entre 18 et 20 ans et 75 et 80 ans). Sur la base de cette population, nous avons voulu prospecter sur leur profil en réalisant un template de personas que nous présentons dans la section suivante.

2.2. La conception d'un template pour personas patient adapté à la chirurgie ambulatoire

Les personas sont des archétypes des futurs utilisateurs du produit en cours de conception. Ils permettent aux concepteurs de se projeter dans les besoins et les attentes des utilisateurs

grâce à des mécanismes psychologiques (Bornet & Brangier, 2013). Ils sollicitent, par exemple, la théorie de l'esprit, la théorie des stéréotypes et peuvent engendrer certains états émotionnels (Lallemand & Gronier, 2018). Les personas permettent aussi d'augmenter l'engagement des concepteurs et de favoriser une synergie positive dans l'équipe de conception (Pruitt & Grudin, 2003). Bornet et Brangier (2013) ont recensé l'ensemble des catégories nécessaires à la conception d'un persona. Cependant, ces catégories ne sont pas adaptées au milieu de la santé et à son contexte très spécifique. En effet, les auteurs ne précisent pas les relations entre le médical, ses technologies et le persona patient. C'est pourquoi, plusieurs auteurs ont tenté de mieux comprendre les éléments à prendre en compte dans la conception de personas dans le milieu de la santé et les contextes d'utilisation de technologies en santé (e.g., Kayser et al., 2015; Andre Kushniruk & Turner, 2012; LeRouge et al., 2013). Kushniruk et Turner (2012) ont mis au point une matrice tridimensionnelle appelée "*user-task-system matrix*". Cette matrice, comme son nom l'indique, prend en compte l'utilisateur, le contexte, la/les tâche(s) pour examiner qui sont les utilisateurs des applications de soins de santé, leurs besoins et leurs exigences dans différents contextes d'utilisation. Cette matrice est reprise peu de temps après par Kayser et al. (2015) qui identifient sept domaines (trois domaines dans les compétences, deux domaines dans l'accès aux technologies et deux domaines dans l'expérience dans l'utilisation des technologies). Enfin, LeRouge, Ma, Sneha, et Tolle (2013, p.251) ont présenté dans leur étude, des éléments complétant les profils des personas traditionnels « *intégrant les structures cognitives et les comportements actuels qui orientent la pensée, le comportement futur et les demandes dans le domaine de la santé*²⁸ » appliqués à une population vieillissante. Nous avons fusionné les éléments présentés dans la Figure 30 de ces recherches afin de proposer un template de personas le plus adapté possible dans la Figure 31.

²⁸ Traduit de l'anglais : "by integrating cognitive structures and present behavior that drive health care thinking, future behavior and demand"(LeRouge et al., 2013, p. 251).

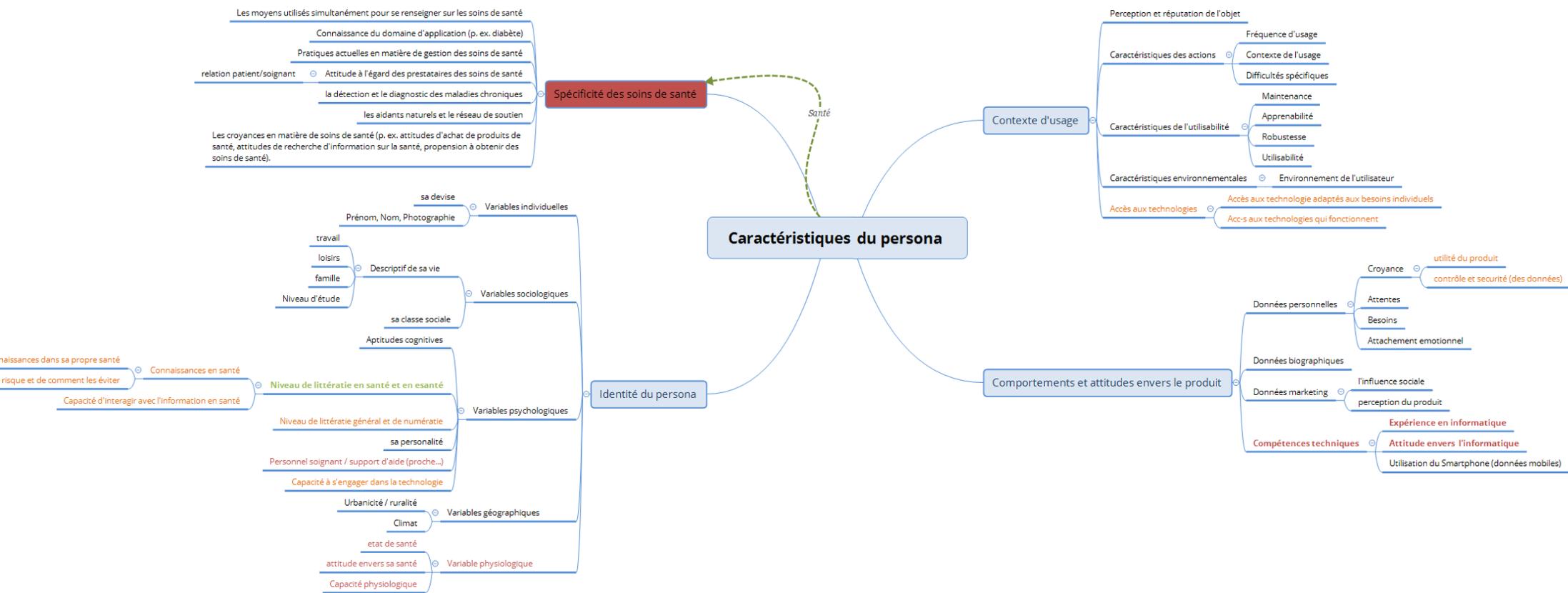


Figure 30. Caractéristiques du persona adaptées à la e-santé.

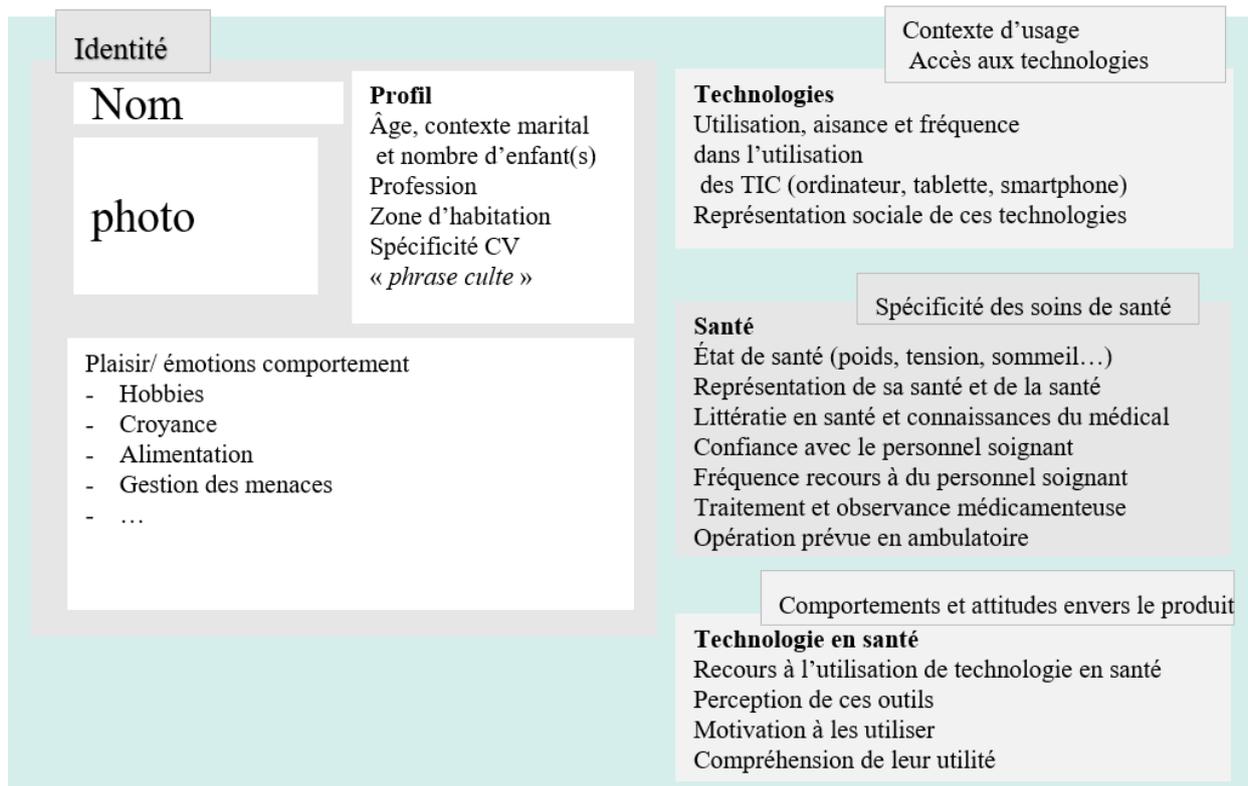


Figure 31. Template des personas patients adapté à la chirurgie ambulatoire.

3. Objectifs de l'étude

Nous avons pu prendre connaissance du parcours patient, des critères d'éligibilité pour une chirurgie ambulatoire et le pourcentage d'individus potentiellement éligibles au futur dispositif *Smart Angel*. Ces données doivent impérativement être combinées à des données de terrain. L'objectif de cette étude exploratoire **était de comprendre le terrain via une analyse de besoin et d'appréhender la population en chirurgie ambulatoire afin d'aider à la conception des premières maquettes du dispositif *Smart Angel* et de transmettre à Evolucare des données sur le fonctionnement réel du parcours ambulatoire dans lequel le futur dispositif *Smart Angel* devra s'intégrer.**

4. Méthode

La méthodologie de cette étude se divise en deux parties : le recueil de données de terrain et l'atelier permettant la conception des personas.

4.1. Recueil de données sur le terrain

Nous avons réalisé des observations (technique du *shadowing*) dans le service de chirurgie ambulatoire ainsi qu'en consultation anesthésique du CHU de Nîmes pendant 9 jours entre octobre 2017 et décembre 2018. Ce travail de terrain a été mené en collaboration avec Annette Valentin, consultante en ergonomie.

4.1.1. Observations ouvertes et entretiens informels

Les observations étaient ouvertes. Pour réaliser ces observations nous avons eu recours à la technique du *Shadowing*. Cette technique consiste à suivre les individus dans leur milieu naturel sans interférer avec leur tâche, comme son nom l'indique en étant « l'ombre » d'elles-mêmes (Czarniawska-Joerges, 2007). Nous avons consigné des verbalisations *provoquées interruptives* (Theureau & Jeffroy, 1994) et des entretiens informels dans un carnet ethnographique. Nous nous sommes intéressées à l'ensemble des tâches réalisées (visite complète du service) en focalisant plus particulièrement notre attention sur la préparation du retour au domicile du patient (appel de la veille, consigne de retour à domicile). Nous avons également assisté à trois réunions de pilotage regroupant l'ensemble des membres du projet *Smart Angel*. Nous ne disposons pas de grille d'entretien ni de grille d'observation.

4.1.2. Caractéristiques du personnel hospitalier et des patients interrogés

Nous avons également pu assister à cinq consultations anesthésiques, et pu interroger le personnel hospitalier (une infirmière chargée de téléphoner aux patients, une infirmière pédagogique expliquant les consignes avant et après l'opération, un aide-soignant, et deux anesthésistes) et des patients acceptant de nous recevoir en chambre ou à la sortie de leur consultation anesthésique. Nous présentons dans le Tableau 13 les caractéristiques des personnes interrogées.

Tableau 13. Caractéristiques du personnel hospitalier et des patients interrogés.

Personnes interrogées	Poste occupé / lieu de recueil
Anesthésiste A	Consultation (3)
Anesthésiste B	Consultation (2)
Infirmière A	Appel de la veille
Infirmière B	Éducation thérapeutique
Infirmière C	Recherche (projet <i>Smart Angel</i>)
Infirmière D	Infirmière d'état (infirmière cheffe)
Aide-soignant A	Service de chirurgie
Patient A – 20 ans	Consultation de suivi postopératoire
Patiente B – 22 ans	Consultation anesthésique
Patient C – 45 ans (environ)	Consultation anesthésique
Patiente D – 35 ans (environ)	En chambre (postopératoire)
Patient E – 45 ans (environ)	En chambre (postopératoire)
Patient F – 60 ans (environ)	En chambre (postopératoire)

4.2. Atelier : conception de personas

Nous avons réalisé un atelier de travail regroupant les membres d'Evolutcare, une ergonome ainsi que nous-mêmes afin de concevoir des personas sur la base des critères présentés en introduction de ce chapitre. Les membres de l'atelier ont complété, à l'aide des données recueillies en observation de terrain et en entretien, le template préalablement conçu (Tableau 14) sur la base de la littérature du domaine.

Tableau 14. Caractéristiques des participants au workshop pour la conception de personas.

Participants au workshop	Corps de métier
Participant 1	Consultant en ergonomie
Participant 2	Informaticien – graphiste
Participant 3	Ingénieur
Participant 4	Manager - Ingénieur

5. Résultats

Cette étude a permis dans un premier temps de collecter des données de terrain dans le service de chirurgie ambulatoire du CHU de Nîmes et dans un deuxième temps de concevoir cinq personas. Nos observations et entretiens nous ont permis de faire deux constats importants : (1) l'organisation des services de chirurgie ambulatoire est en constante évolution tout comme les moyens mis en place pour le suivi du patient post-ambulatoire, (2) certains événements du parcours patient sont à privilégier dans l'intégration du dispositif *Smart Angel*. Enfin, l'atelier de travail a permis la conception de cinq personas animés par la création de cinq scénarii mettant en action des patients archétypes dans le service de chirurgie ambulatoire (cf. Annexe 8 : Présentation des personas et leurs scénarios, p. 371). L'ensemble de ces éléments a permis à l'équipe Evolucare de proposer des premières maquettes du dispositif *Smart Angel*.

5.1. L'organisation du suivi postopératoire en constante évolution

Nous avons pu constater, entre notre première venue en octobre 2017 et notre dernière observation en décembre 2018, des modifications organisationnelles dans le suivi-post opératoire des patients sortant de chirurgie ambulatoire. Ce service du CHU de Nîmes est relativement important dans le sens où il réalise entre 8 000 et 10 000 opérations par an en ambulatoire dans des domaines très variés de la chirurgie. L'objectif de l'année 2017 du CHU était de propulser le taux d'admission en ambulatoire à 45 % sur l'ensemble des chirurgies (soit une augmentation de 5% par rapport à l'année précédente). Pour l'organisation des salles, aucun bloc opératoire n'était consacré exclusivement à la chirurgie ambulatoire. Les actes étaient répartis sur les blocs de chirurgie dans les services concernés. Puis, le service s'est peu à peu organisé afin d'avoir ses propres locaux centralisés. Enfin, depuis 2019, le CHU a investi dans un nouveau centre dédié à la chirurgie ambulatoire composé de 4 salles opératoires permettant de réaliser 20 % des actes en ambulatoire.

L'ambulatoire prend de plus en plus de place dans le CHU de Nîmes tout comme l'investissement dans le suivi du patient. Le CHU a mis en place un parcours « *patient debout* », appelé parcours patient 3D (patient Digne, Debout, Détendu). Le service s'engage dans la prise en charge des opérations de plus en plus risquées. Une toute première opération de pose de prothèse de hanche venait d'être initiée dans ce service en décembre 2017. Les demandes en termes de suivi post-ambulatoire ont alors été modifiées. Nous avons pu constater trois

modifications dans la gestion des patients en un an : (1) le recours à un personnel infirmier pour appeler le patient, (2) l'utilisation de SMS pour demander si un appel est nécessaire, (3) l'intégration d'un personnel infirmier pédagogique après la consultation anesthésique.

(1) Lors de notre première visite, un appel systématique était réalisé par un/e infirmier/ère tournant à ce poste chaque jour. Pour cela, le personnel possédait une liste de patients à appeler et pouvait passer entre 30 secondes et 10 minutes au téléphone par patient. Le personnel soignant à ce poste utilisait un protocole précis rédigé sur une fiche possédant diverses questions pré-écrites. Cependant, le personnel ne suivait pas forcément le questionnaire, mais abordait les éléments de manière pragmatique. Si le patient semblait en forme et ne nécessitait pas d'aide, l'infirmier/ière raccrochait très rapidement. En revanche, si le patient semblait en souffrance, l'interrogatoire de l'infirmier/ière devenait davantage précis jusqu'à dépasser les questions du protocole préétabli. L'infirmière à ce poste le jour de notre visite nous a déclaré : « *Quand les gens ont un problème de douleur, si elle ne passe pas c'est qu'il y a un souci, il faut contextualiser le problème et savoir pourquoi ils ont mal* ». Elle nous a ensuite expliqué que la plupart du temps, lorsque le patient est souffrant, c'est qu'il n'a tout simplement pas pris ses antidouleurs. Aucun élément statistique ne vient étayer cette affirmation. Ce personnel avait en charge d'évaluer l'état de santé du patient puis de décider de sa réhospitalisation le cas échéant.

(2) Lors de notre deuxième visite, un système automatique d'envoi de SMS avait été mis en place pour décharger le personnel infirmier des appels vers les patients. Le SMS était composé d'une simple phrase : « *Voulez-vous qu'un professionnel de santé vous appelle ?* ». Le patient devait alors répondre par SMS pour valider ou infirmer la proposition. Le point (1) se mettait en place si le patient désirait être contacté ou ne répondait pas à ce SMS.

(3) Lors de notre troisième visite, un nouveau système post consultation anesthésique avait été mis en place pour décharger les anesthésistes dans les explications aux patients. Un personnel infirmier d'éducation thérapeutique devait, après la consultation, rappeler les règles élémentaires avant (e.g., ne pas fumer, être à jeun, se raser au niveau de la zone opératoire, enlever son vernis à ongles...) et après l'opération (e.g., ne pas faire de sport après l'anesthésie, ne pas consommer d'alcool...). Ces informations étaient détaillées dans un fascicule remis au patient. Le personnel infirmier était également présent pour répondre aux questions et donner les prescriptions médicamenteuses aux patients. Le temps moyen passé avec un patient était d'environ 5 à 10 min d'après le personnel interrogé à ce poste. Lorsqu'un patient ne réalise pas convenablement ces actes décrits, celui-ci est soit renvoyé chez lui et ne peut pas être opéré,

soit le personnel hospitalier se charge de le préparer imposant ainsi une charge de travail supplémentaire. Dans les deux cas, l'hôpital est perdant si le patient ne comprend pas les principales règles avant et après une opération imposant une importante responsabilité au personnel infirmier en termes d'éducation thérapeutique.

Ainsi, ces trois constats nous montrent non seulement que le circuit de la chirurgie ambulatoire du CHU de Nîmes concorde avec nos données théoriques, mais possède également ses propres subtilités dans son organisation et que celles-ci sont en constante évolution.

5.2. Événements pour l'intégration du dispositif *Smart Angel* au domicile patient

Lors des réunions de pilotage du projet, les partenaires industriels et hospitaliers ont établi un protocole pour la future intégration du dispositif *Smart Angel*. Ils considèrent que l'anesthésiste devra choisir quel patient pourra recevoir le dispositif, puis après opération, les infirmiers devront former le patient en chambre. Nous avons donc observé des consultations anesthésiques afin d'appréhender la demande. L'anesthésiste possède un rôle central dans la chirurgie ambulatoire. Il prend la décision de proposer une chirurgie en ambulatoire ou pas. Lorsqu'il juge lui-même que le patient n'est pas apte, il ne propose pas d'ambulatoire au patient. Il juge les risques sur sa propre perception du patient, sur son âge, son état de santé (score de l'American Society of Anesthesiologists ou ASA) et ses réponses aux questions. L'acceptation du dispositif *Smart Angel* par le patient dépendra probablement de la pédagogie de l'anesthésiste à lui expliquer l'utilité de l'appareil. De plus, le choix du patient en amont est probablement essentiel pour limiter les risques d'erreurs d'utilisabilité.

Dans un deuxième temps, le personnel infirmier devra former le patient en chambre. Pour la préparation à l'étude clinique prévue dans le cadre du projet *Smart Angel*, une infirmière de recherche se chargera de former le patient. Elle nous a expliqué qu'elle comptait former le patient en 20 min et sur le principe suivant « *le patient fait et l'infirmière guide* », ce qui pourrait s'apparenter à une formation de « suivi » (Larsen, 2017). Lors de nos visites en chambre, l'état de vigilance et d'attention dépend d'un patient à l'autre.

5.3. Conception de personas et scénarii

Une fois l'ensemble des éléments du terrain pris en compte, nous avons organisé un workshop. Nous avons demandé aux participants du workshop de compléter les templates de personas présentés en amont. Six personas ont pu être créés. Un des personas pouvant être éligible en chirurgie ambulatoire ne l'était pas pour le suivi post-ambulatoire avec *Smart Angel*.

Ce persona a donc été exclu pour les scénarii mettant en scène le dispositif *Smart Angel*. Nous proposons un exemple des personas conçus dans la Figure 32. Les autres personas sont présentés en Annexe (cf. Annexe 8 : Présentation des personas et leurs scénario, p. 371).

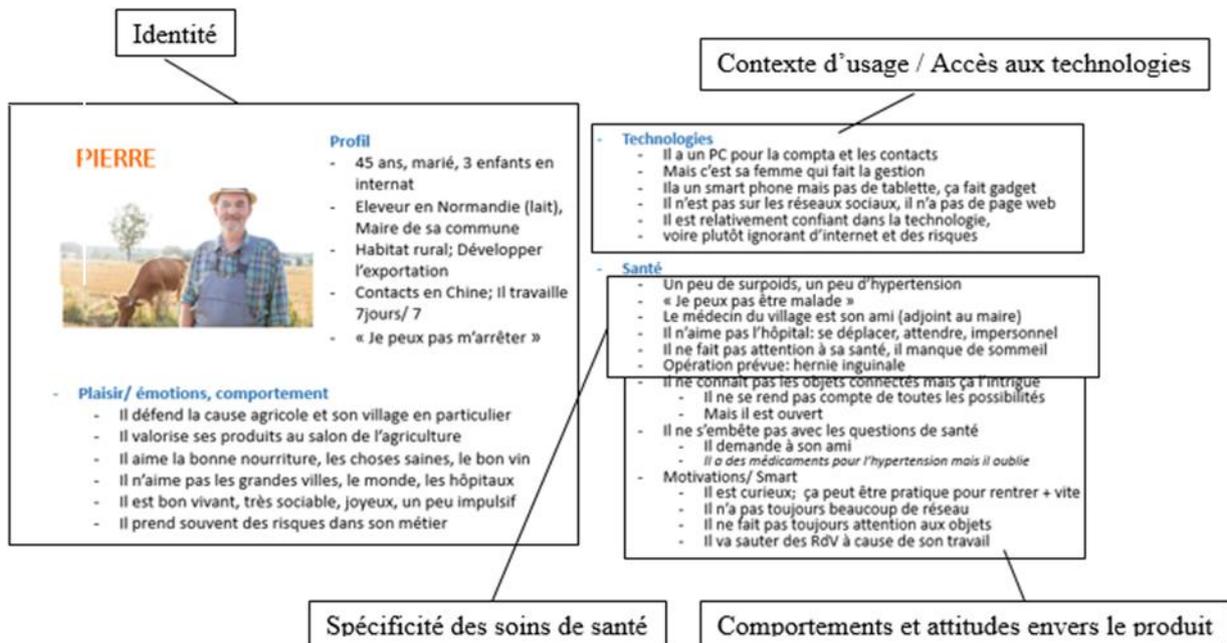


Figure 32. Exemple de persona (le cas de Pierre), construit lors de l'atelier, adapté à la e-santé.

6. Discussion

Cette étude exploratoire avait pour objectif d'analyser le besoin et de comprendre le terrain de la chirurgie ambulatoire afin d'aider à la conception des premières maquettes du dispositif *Smart Angel*. Pour cela, nous nous sommes rendues au CHU de Nîmes. Nos observations et prises de notes dans un journal ethnographique ont permis de constater les éléments suivants : l'évolution de l'organisation du service de chirurgie ambulatoire suit une progression constante ; ces évolutions sont corrélées avec des modifications dans la prise en charge du patient ; enfin, nous avons pu observer un fort investissement dans l'éducation thérapeutique mise en place par le personnel pour le bon déroulement et l'optimisation des soins.

Le terrain a également été l'occasion de comprendre quelles pourraient être les prochaines étapes à prendre en compte dans l'intégration du dispositif *Smart Angel* et appréhender son

environnement dans le contexte d'usage. Nous avons pu constater que l'anesthésiste détient probablement un rôle central dans l'acceptation du dispositif *Smart Angel* et que le personnel hospitalier sera responsable de la future formation du dispositif. En effet, cette étude a permis de révéler que le patient et ses caractéristiques ne sont pas les seuls éléments à prendre en compte dans une relation patient-soignant, le personnel hospitalier possède lui aussi des contraintes qui lui sont propres en fonction de l'organisation et de la législation dans lequel il s'intègre. Ten Hacken et al. (2018) recommandent d'interroger le personnel soignant lors de l'implantation d'une technologie de santé même si celle-ci est destinée uniquement aux patients.

Certaines limites sont à prendre en compte dans notre intervention. Nos déplacements dans le service étaient contrôlés. Un parcours approximatif était pré organisé, en fonction de nos demandes, par le personnel hospitalier, tout comme les personnes que nous avons pu interroger. Les patients étaient préalablement choisis par le personnel hospitalier. Nos visites étaient également partielles et fragmentées. En effet, dans un but d'optimiser le temps sur le terrain, nous avons dû avec les membres du projet, nous disperser dans l'hôpital afin d'observer un maximum de situations. Les informations collectées ont ultérieurement été rassemblées et centralisées entre les acteurs du terrain. Les profils présentés dans la partie ne représentent que les personnes que nous avons pu interroger. De plus, du fait de la temporalité limitée, nous n'avons pas pu prendre en compte l'extrême diversité des profils présents en chirurgie ambulatoire (Loup-Escande et al., 2019).

Par la suite, nous avons réalisé un atelier de travail avec l'équipe projet qui a abouti à la conception de cinq personas grâce aux données du terrain. Ces personas ont eu trois apports. Premièrement, ils ont favorisé une vision partagée des caractéristiques et des besoins des utilisateurs comme outil pour le développement des maquettes. Deuxièmement, ils ont également permis d'appréhender des pistes d'amélioration de l'utilisabilité jusqu'alors jamais explorées expérimentalement comme la prise en compte du niveau de littératie en santé du patient, son engagement envers les technologies, ou encore sa relation avec le personnel hospitalier. Troisièmement, les personas mis en scène à travers des scénarii ont été d'un intérêt certain pour l'évaluation du système *Smart Angel* (cf. Partie II, Chapitre 4, p. 189).

Cependant, les personas proposés contiennent certaines limites. En effet, nous manquons de connaissances sur l'impact du contexte d'usage sur l'utilisation d'outils en santé. Nous n'avons pas, à notre connaissance, de preuve empirique concernant certaines caractéristiques (e.g., littératie en santé, relation patient/soignant) intégrées dans le profil du personas, pouvant

impacter l'utilisation d'un produit connecté en santé. De plus, nous avons basé nos analyses sur un échantillon limité d'individus pouvant biaiser la construction de nos personas. Pour contrebalancer cette limite, les analyses démographiques ont permis de lisser davantage ces profils.

7. Conclusion et perspectives de recherche

Pour conclure, cette étude exploratoire, correspondant au niveau 0 du modèle HANDLED, basée sur la collecte de données démographiques prospectées ainsi que sur les observations de terrain réalisées au CHU de Nîmes (partenaire du projet), nous ont permis de mieux comprendre les besoins et les attentes suscités dans le projet *Smart Angel* et d'en connaître davantage sur la population ciblée. Ces analyses et observations nous ont également aidées à la conception de cinq personas lors d'un atelier de travail et d'une équipe pluridisciplinaire. L'ensemble de ces éléments ont permis la création des premières maquettes du dispositif *Smart Angel*.

Cette étude a également été bénéfique pour nos réflexions sur les pistes d'amélioration de l'utilisabilité du dispositif *Smart Angel*. Il semble important de prendre en compte les compétences des patients (cf. Chapitre 2, p. 132), leur état postopératoire afin de « choisir » le patient en amont et de limiter les erreurs d'utilisabilité pouvant impacter la sécurité du patient. Il est également essentiel de vérifier l'intérêt des formations délivrées par le personnel infirmier (cf. Partie II, Chapitre 5, p. 226).

Suite à cette étude, une perspective de recherche peut être suggérée. Il aurait été judicieux d'être davantage inclus dans le service et de nouer des relations avec le personnel hospitalier. Le peu de temps sur place n'a pas permis de créer une confiance suffisante afin de s'assurer de ne pas biaiser les activités du personnel par notre simple présence. Ce temps supplémentaire aurait également permis d'accumuler davantage de connaissances sur le personnel hospitalier ainsi que sur les profils des patients. En effet, nous avons accédé, de manière aléatoire, aux chambres hospitalières et aux patients qui s'y trouvaient. Un plus large échantillon investigué aurait potentiellement permis d'ajuster les personas construits. Enfin, par notre simple présence dans le CHU, nous aurions pu resserrer les liens plus étroits avec nos partenaires et entretenir la collaboration. Il est possible que la localisation géographique et la hiérarchisation entre les partenaires hospitalier (Nîmes) et industriels (Amiens) aient été un frein à cette proximité sociale pourtant nécessaire dans un projet collaboratif (e.g., Du Chatenier et al., 2010). Nous

reviendrons plus en détails sur ces éléments dans la discussion générale de la thèse (cf. Discussion Générale, 3.1. La hiérarchisation du partenariat , p. 291).

8. Synthèse du Chapitre 1 - Étude exploratoire

L'objectif de cette étude exploratoire, correspondant au niveau 0 (tâche) de notre modèle HANDLED, **était de comprendre le terrain via une analyse de besoin et d'appréhender la population en chirurgie ambulatoire afin d'apporter des données réelles à Evolucare avant l'implantation du dispositif *Smart Angel*. Ces données permettront *in fine* de concevoir les premières maquettes du dispositif.**

Pour cela, nous avons utilisé une double méthodologie : **une analyse du terrain et un atelier de conception de personas.**

Premièrement, nous avons réalisé des **observations directes et ouvertes** ainsi que des **entretiens informels** avec le personnel hospitalier dans le service de chirurgie ambulatoire ainsi qu'en consultation anesthésique au CHU de Nîmes pendant 9 jours cumulés.

En combinant une analyse théorique du parcours patient en chirurgie ambulatoire et notre analyse de terrain, **nous avons pu assimiler le fonctionnement et l'organisation du suivi postopératoire et appréhender la future intégration du dispositif *Smart Angel* au sein de ce service.** À la suite de cette investigation, deux facteurs semblent particulièrement importants à retenir comme perspectives de recherche pour la suite de notre travail de thèse afin d'assurer la sécurité du patient lors de l'utilisation du dispositif *Smart Angel* : **l'identification du patient apte à l'utilisation du dispositif** et la **pédagogie permettant d'expliquer l'utilité de l'appareil par l'anesthésiste.** Ces éléments issus du contexte environnemental du dispositif *Smart Angel* ont servi de base pour la reconstitution de scénarios en simulation lors de l'étude 3 (cf. Partie II, Chapitre 5, p. 226).

Deuxièmement, nous nous sommes centrées sur **le profil du futur utilisateur** du dispositif *Smart Angel*. Pour cela, nous avons mis en place un **atelier de conception de personas** avec **l'équipe pluridisciplinaire** du projet *Smart Angel*. En combinant une **analyse démographique et prospective** issue de différents rapports gouvernementaux, une **analyse théorique des personas patient** et notre **analyse de terrain**, nous avons pu créer **un template de personas patient adapté au futur dispositif *Smart Angel*.** À la suite de l'atelier de conception, cinq personas ont pu être créés. Ces personas ont été utilisés lors de l'évaluation du prototype 1 de *Smart Angel* lors de l'étude 2 (cf. Partie II, Chapitre 4, p. 189).

CHAPITRE 2 : ÉTUDE 1a. EFFETS DES CONNAISSANCES EN SANTÉ D'UNE POPULATION NOVICE SUR L'UTILISABILITÉ DES DEUX DISPOSITIFS MÉDICAUX CONNECTÉS

1. Préambule

Ce présent chapitre est tiré d'un article scientifique (Chaniaud et al., 2020)²⁹ accepté dans *Journal of Medical Informatics and Research* dans la section *mhealth & uhealth*. Ce travail a été réalisé en collaboration avec Natacha Métayer, ingénieure de recherche sur le Projet *Smart Angel*.

1.1. Contexte et objectif industriels

Lors de la conception de cette étude, aucun prototype du dispositif *Smart Angel* n'était opérationnel (Figure 33). Le CHU de Nîmes propose cependant deux objets connectés à implanter dans le dispositif pour la prise des constantes quotidiennes : un tensiomètre de poignet et un oxymètre de pouls. Ces deux objets sont déjà commercialisés, disponibles pour le grand public et possèdent un marquage de Conformité Européenne (appartenant à la classe IIa – risque potentiel modéré). L'objectif industriel de cette étude est d'évaluer l'utilisabilité afin de juger de l'*efficacité*, l'*efficience* et de la *satisfaction* (ISO 9241-11 : 2018) de ces deux produits.

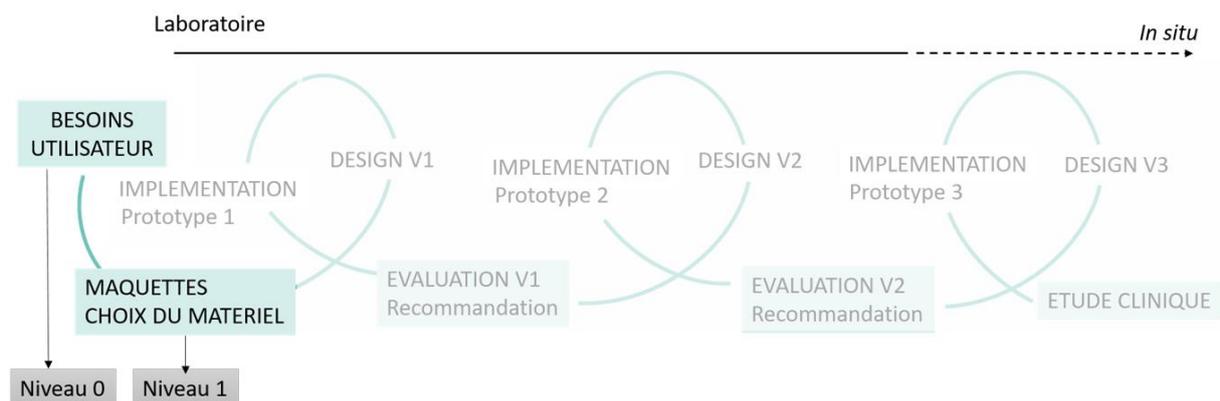


Figure 33. Étude 1a correspondant au niveau 1 du processus de conception du dispositif *Smart Angel*.

²⁹ Chaniaud, N., Métayer, N., Megalakaki, O., & Loup-Escande, E. (2020). Effect of prior health knowledge of novice users on Usability Results: Usability Study. *JMIR mhealth & uhealth*, 8(9). 10.2196/17983. <https://mhealth.jmir.org/2020/9/e17983/pdf>

1.2. Contexte et objectif théoriques

Cette étude correspond à la première itération dans le cycle de conception du dispositif *Smart Angel* soit le niveau 1 du processus de conception du dispositif *Smart Angel* (Figure 33). C'est pourquoi nous étudions uniquement l'impact de certaines caractéristiques utilisateurs (technophilie et connaissances en santé) d'une population homogène (jeune et novice) sur l'utilisabilité dans un milieu contrôlé. La littérature scientifique met en avant la littératie en santé comme potentielle piste privilégiée impactant l'utilisabilité, même si aucune étude n'a alors prouvé son lien empirique (Kim & Xie, 2017). Or, lors de la construction de cette étude, aucune traduction des questionnaires de littératie en santé reconnus et adaptés n'était validée en français (e.g., Mansfield et al., 2018; Rouquette et al., 2018). Nous avons fait le choix de nous intéresser aux connaissances antérieures en santé, car ce facteur est intégré dans la définition de la littératie en santé (e.g., Sørensen, Brouck et al., 2012). Cette relation a d'ailleurs été prouvée empiriquement (Williams et al., 1998). De ce fait, nous nous attendions à ce que les connaissances en santé aient un impact sur l'utilisabilité d'un dispositif médical. La Figure 34 modélise les différents éléments évalués dans cette étude basée sur la matrice : *tâches – utilisateurs*.

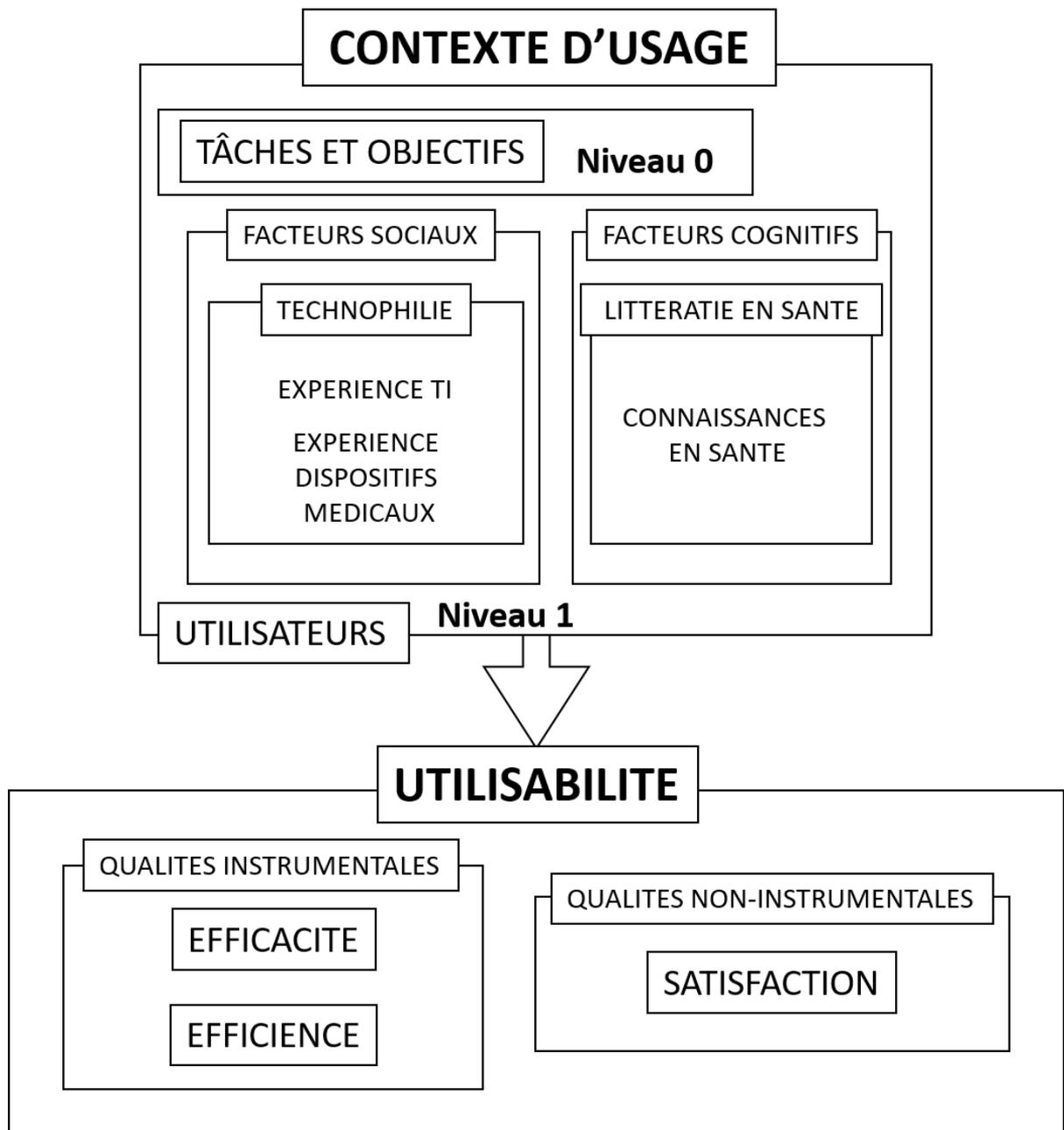


Figure 34. Attributs éprouvés lors de l'étude 1a du contexte d'usage sur l'utilisabilité (*efficacité, efficacité et satisfaction*).

2. Introduction

Les Dispositifs Médicaux à Domicile³⁰ ou DMD sont de plus en plus souvent prescrits par les professionnels de santé afin de décongestionner les hôpitaux et constituent potentiellement un moyen rentable de répondre aux besoins croissants en matière de soins de santé (Bitterman, 2011 ; Conseil national de la recherche, 1996). Les DMD destinés au grand public doivent être adaptés à tous les types de population, quel que soit l'environnement dans lequel ils sont utilisés (Huvila et al., 2016 ; Yu et al., 2017). Certains d'entre eux peuvent être très complexes à utiliser (e.g., Mykityshyn et al., 2002). Cependant, si un DMD est mal conçu, les patients peuvent commettre des erreurs d'utilisation qui pourraient sérieusement affecter leur sécurité (e.g., Carayon & Hoonakker, 2019 ; Cifter, 2017 ; Greysen et al., 2014 ; Huvila et al., 2016 ; Kortum & Peres, 2015 ; Tarricone et al., 2017).

C'est pourquoi un certain nombre de projets en cours portent sur l'évaluation et l'amélioration de la conception de ces dispositifs médicaux (e.g., Tarricone et al., 2017). Les méthodologies centrées sur l'utilisateur (e.g., LeRouge & Wickramasinghe, 2013 ; Matthew-Maich et al., 2016 ; Schnall et al., 2016) sont de plus en plus utilisées à mesure que de nouvelles normes de qualité et de sécurité émergent (ISO/IEC 62366-1:2015), afin d'éviter les erreurs de conception. Le marquage "Conformité Européenne" (CE) est utilisé pour prouver la sécurité et l'utilisabilité (ISO 9241-11:2018) par trois métriques l'*efficacité*, l'*efficience* et la *satisfaction* dans le contexte spécifique de l'utilisation de l'appareil. Malgré toutes ces normes, des problèmes d'utilisabilité persistent. Les analyses systématiques des dix dernières années ont constamment tiré la sonnette d'alarme (Jaspers, 2009 ; Lyles et al., 2014 ; Peute et al., 2013 ; Peute et al., 2008 ; Tractinsky, 2018 ; Yen & Bakken, 2012) en signalant l'absence de cadre et de méthode normalisée dans les études d'utilisabilité. Même si la définition de l'utilisabilité reste toujours un débat important (e.g., Borsci et al., 2018), sa norme a été mise à jour (ISO 9241-11:2018). Bevan et al., (2015) mettent en avant l'importance de qualifier l'utilisabilité comme le résultat d'une interaction plutôt que la propriété d'un produit. Cette nouvelle norme (ISO 9241-11:2018) définit l'utilisabilité par son contexte d'usage et comprend 4 composants : *objectifs et tâches, ressources, environnement, utilisateurs*. Ces composantes influencent l'utilisabilité (composée par l'*efficacité*, l'*efficience* et la *satisfaction*) et il est donc nécessaire de savoir comment. Dans cette étude, nous nous concentrons sur les caractéristiques des

³⁰ Traduit de l'anglais : Home Medical Device ou HMD

utilisateurs (en contrôlant les trois autres composants) en nous basant sur la métrique de l'utilisabilité de la norme ISO 9241-11 (2018).

Plusieurs chercheurs ont récemment étudié les liens entre les caractéristiques des utilisateurs et l'utilisabilité des dispositifs connectés en santé (e.g., Georgsson & Staggers, 2016 ; Jensen et al., 2010 ; Liang et al., 2018 ; Monkman & Kushniruk, 2013 ; Sonderegger et al., 2016 ; Van der Vaart et al., 2019). Quatre caractéristiques des utilisateurs semblent être particulièrement étudiées dans la littérature scientifique :

- L'âge : les jeunes utilisateurs obtiennent de meilleurs résultats que les utilisateurs plus âgés (e.g., Georgsson & Staggers, 2016 ; Jones & Caird, 2017 ; Kaufman et al., 2003 ; Mykityshyn et al., 2002 ; Sonderegger et al., 2016 ; Sparkes et al., 2012 ; Van der Vaart et al., 2019 ; Yu et al., 2017) ;
- L'expérience dans le domaine des technologies de l'information (TI) (ou technophilie), c'est-à-dire une expérience préalable dans le domaine des technologies et des dispositifs médicaux : les experts en technologie obtiennent de meilleurs résultats que les novices (e.g., Georgsson & Staggers, 2016 ; Harte et al., 2018 ; Sparkes et al., 2012 ; Yu et al., 2017) ;
- La motivation : les utilisateurs plus motivés surpassent les utilisateurs moins motivés (e.g., O'Brien & Toms, 2008 ; Rico-Olarte et al., 2018 ; Tissier & Gronier, 2014) ;
- Les connaissances antérieures en santé : les utilisateurs ayant un niveau élevé de connaissances antérieures en santé obtiennent de meilleurs résultats que les utilisateurs ayant un faible niveau de connaissances antérieures en santé (e.g., Czaja et al., 2015 ; Kaufman et al., 2003 ; Van der Vaart et al., 2019).

Si certaines études établissent les relations entre ces caractéristiques et l'utilisabilité, il n'existe que peu d'études examinant le lien entre la littératie en santé (e.g., Czaja et al., 2015 ; Kaufman et al., 2003 ; Kim & Xie, 2017 ; Van der Vaart et al., 2019) ou la littératie en e-santé (e.g., Vaart et al., 2019) - en particulier les connaissances antérieures en santé - et l'utilisabilité. L'objectif de la présente étude est donc d'examiner les effets des connaissances antérieures en santé sur l'utilisabilité des DMD.

2.1. Contexte de l'étude

L'étude décrite dans ce document fait partie du projet *Smart Angel*, qui vise à fournir un suivi individuel à domicile pour les patients ayant subi une chirurgie ambulatoire. Cette

surveillance est effectuée pendant une semaine, ce qui permet au patient de rester en contact avec l'hôpital depuis son domicile. Ce dispositif est considéré comme une technologie de e-santé. L'e-santé (ou santé connectée) est définie par Eysenbach (Eysenbach, 2001) comme un domaine émergent réunissant différentes disciplines telles que l'informatique médicale, la santé publique et les entreprises. La e-santé offre une occasion majeure de remédier aux lacunes des systèmes de santé actuels et de soutenir les professionnels de la santé et les patients en les rendant acteurs de leur propre santé (Eysenbach, 2001 ; Hamine et al., 2015).

Le kit *Smart Angel* est composé de dispositifs de surveillance (tensiomètre et oxymètre de pouls) connectés à une application numérique utilisable sur tablette tactile. Ces deux dispositifs ont été choisis par les partenaires hospitaliers du projet *Smart Angel* sur la base d'un benchmark. Les critères de sélection étaient la présence d'une certification européenne (CE) et une utilisation jugée aisée par les professionnels médicaux. Il n'y a aucun lien entre notre laboratoire et la société qui fabrique ces dispositifs. Ainsi, les futurs patients se retrouveront dans un contexte postopératoire, c'est-à-dire potentiellement avec des douleurs, des nausées, à domicile avec leur dispositif et devront l'utiliser (seul ou accompagné) trois fois par jour pour fournir des relevés de santé actualisés à l'hôpital. Les besoins recueillis sur le terrain en chirurgie ambulatoire ont permis de sélectionner deux dispositifs connectés qui assureront la surveillance du patient : le tensiomètre et l'oxymètre de pouls, disponibles sur le marché public et portant le marquage CE (Conformité Européenne).

2.2. Évaluation de l'utilisabilité

Selon la norme ISO 9241-11 (2018), l'utilisabilité est définie comme "*le degré selon lequel un produit peut être utilisé par des utilisateurs spécifiés pour atteindre des objectifs spécifiés avec efficacité, efficience et satisfaction dans un contexte d'utilisation spécifié*". Ce cadre permet de stabiliser l'utilisabilité autour de trois dimensions principales (l'efficacité, l'efficience et la satisfaction) qui sont largement utilisées dans le domaine de la e-santé (e.g., Berkman et al., 2011 ; Georgsson & Staggers, 2016 ; Lyles et al., 2014 ; Triantafyllidis et al., 2015 ; Yen & Bakken, 2012). Certains auteurs soutiennent ces normes (Bevan, 2009) et l'importance d'évaluer les mesures d'utilisabilité par ces trois composantes de manière indépendante et de recueillir des données à la fois subjectives et objectives (Frøkjær et al., 2000 ; Georgsson & Staggers, 2016 ; Lyles et al., 2014).

Selon la norme ISO 9241-11 (2018), l'efficacité est définie comme la précision et l'exhaustivité avec laquelle les utilisateurs atteignent les objectifs spécifiés. Elle est généralement mesurée sur trois points : (1) les erreurs ou les difficultés d'utilisation ; (2) les éléments de sortie inutiles qui interfèrent avec la tâche de l'utilisateur ; (3) les décisions inappropriées prises sur la base de données de sortie inexactes ou incomplètes. L'efficience est définie comme les ressources utilisées par rapport aux résultats obtenus (les ressources typiques comprennent le temps, l'effort humain, les coûts et les matériaux). L'efficience comprend "le temps utilisé", c'est-à-dire le temps passé à essayer d'atteindre un objectif. La satisfaction est définie comme la mesure dans laquelle les réponses physiques, cognitives et émotionnelles de l'utilisateur qui résultent de l'utilisation d'un système, d'un produit ou d'un service répondent aux besoins et aux attentes de l'utilisateur. La satisfaction est évaluée par des réactions physiques (sentiments de confort ou d'inconfort) et cognitives (attitudes, préférences et perceptions) d'après la norme ISO (9241-11:2018).

2.3. L'utilisabilité et les caractéristiques utilisateurs

Une meilleure compréhension de ce qui influence l'utilisabilité est d'une importance cruciale pour améliorer la conception des dispositifs médicaux. Un nombre croissant de chercheurs prennent en compte les caractéristiques démographiques des utilisateurs qu'ils interrogent dans leurs études d'utilisabilité, telles que l'âge, le sexe, le niveau d'éducation, l'expérience informatique, le type de maladie (e.g., Ayubi et al., 2014 ; Mirkovic et al., 2014 ; Sheehan & Lucero, 2015 ; Sparkes et al., 2012). Cependant, très peu d'entre eux recueillent les compétences des utilisateurs telles que la connaissance antérieure en santé ou dans ces dispositifs (e.g., Constantinescu et al., 2019 ; Van der Vaart et al., 2019). Selon Borsci, Uchegbu et al. (2018) et Grebin et al. (2018), le manque d'attention portée aux facteurs humains est l'une des raisons de la lenteur de l'adoption des innovations médicales. Ces auteurs ont proposé de mieux comprendre les facteurs qui influencent ces processus de décision afin de mieux comprendre les capacités de résilience des individus.

Nous avons pu identifier, à partir de la littérature scientifique, quatre variables principales des caractéristiques des utilisateurs qui influencent directement l'utilisabilité sur une population "en bonne santé". Nous proposons ci-dessous une analyse non exhaustive axée sur les caractéristiques des utilisateurs qui influencent l'utilisabilité (Figure 35). Ces quatre variables principales sont : l'âge (e.g., Georgsson & Staggers, 2016 ; Jones & Caird, 2017 ;

Kaufman et al., 2003 ; Mykityshyn et al., 2002 ; Sonderegger et al., 2016 ; Sparkes et al., 2012 ; Van der Vaart et al., 2019 ; Yu et al., 2017), la technophilie (e.g., Georgsson & Staggers, 2016 ; Harte et al., 2018 ; Sparkes et al., 2012 ; Yu et al., 2017), la littératie en santé (e.g., Czaja et al., 2015 ; Kaufman et al., 2003 ; Kim & Xie, 2017 ; Vaart et al., 2019 ; Van der Vaart et al., 2019) et la motivation (e.g., O'Brien & Toms, 2008 ; Rico-Olarte et al., 2018 ; Tissier & Gronier, 2014). Par exemple, une étude de Loorbach, Karreman et Steehouder (2007) a montré que des instructions manuelles plus motivantes amélioreraient l'efficacité et l'efficience (mais pas la satisfaction) d'un téléphone portable chez une population âgée (60-70 ans). La motivation est directement liée à l'engagement (activation) qui est également directement lié à la santé du patient (Greene & Hibbard, 2012). En outre, l'âge (e.g., Sonderegger et al., 2016 ; Sparkes et al., 2012 ; Van der Vaart et al., 2019 ; Yu et al., 2017) et la technophilie (e.g., Georgsson & Staggers, 2016 ; Harte et al., 2018) ont un impact sur l'utilisabilité. En revanche, l'influence du niveau d'éducation et des situations professionnelles reste ambiguë (e.g., Georgsson & Staggers, 2016 ; Liang et al., 2018 ; Van der Vaart et al., 2019).

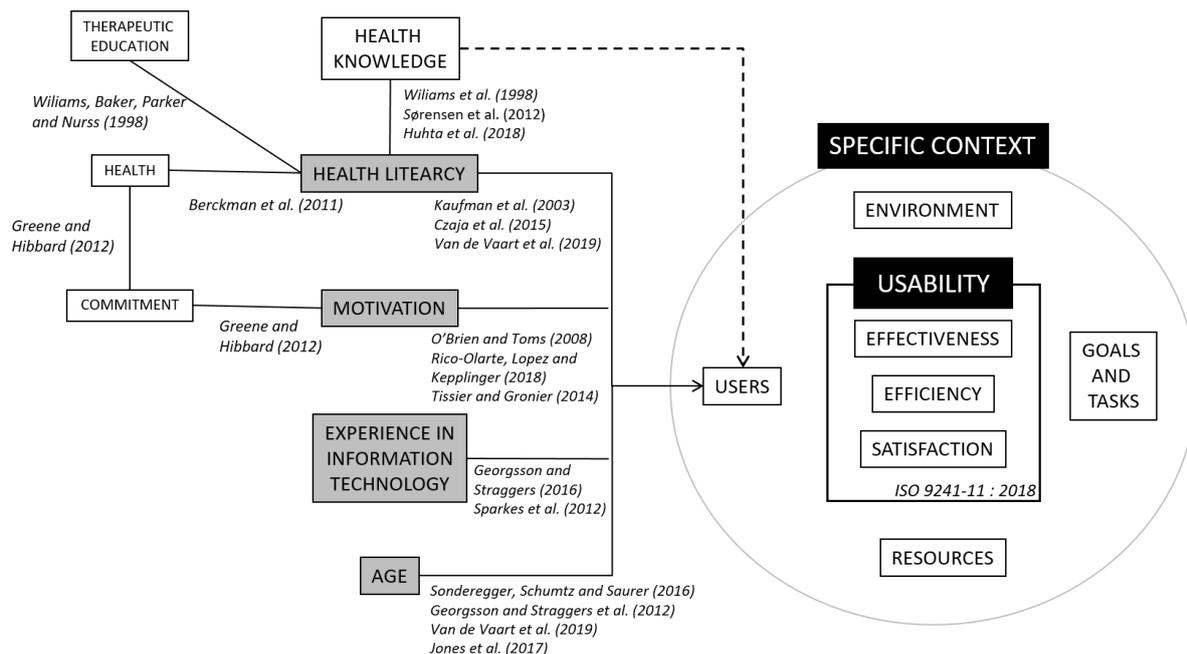


Figure 35. L'influence des quatre principales caractéristiques de l'utilisateur en gris (âge, technophilie, motivation et littératie en santé) tirées de la littérature scientifique sur l'utilisabilité (efficacité, efficience et satisfaction) et qui ont un impact sur la santé des patients. Les traits pleins sont issus de la littérature publiée tandis que les traits pointillés sont hypothétiques.

2.4. Utilisabilité, littératie en santé et connaissances antérieures en santé

La littératie en santé semble être un domaine sur lequel de nombreux chercheurs se concentrent afin d'améliorer la conception des dispositifs médicaux (Czaja et al., 2015 ; Kim

& Xie, 2017 ; Monkman et al., 2015 ; Monkman & Kushniruk, 2013) et ainsi améliorer la santé des patients (Berkman et al., 2011). Czaja et al. (2015) ont montré que les personnes ayant un faible niveau de littératie en santé éprouvaient des difficultés à remplir les dossiers de santé électroniques personnels. Mackert, Mabry-Flynn, Champlin, Donovan et Pounders (2016) ont constaté que les patients ayant un faible niveau de littératie en santé utilisaient moins les technologies de l'information sur la santé. Kim et Xie (2017) ont réalisé une étude systématique de l'impact de la littératie en santé sur les technologies de la santé. Après une analyse de 74 études, les auteurs concluent que le principal obstacle à l'accès et à l'utilisation des informations de santé en ligne pour les personnes ayant un faible niveau de littératie en santé est fortement lié à l'utilisabilité. Un faible niveau de littératie en santé peut limiter l'accès des patients à l'éducation thérapeutique (e.g., Margat et al., 2014 ; Williams et al., 1998) et par conséquent diminuer leur engagement (e.g., Sarkar et al., 2006), pouvant même être visible sur leur état de santé (e.g., Paasche-Orlow et Wolf, 2007).

Même avec un niveau élevé de littératie "générale", une personne peut avoir des difficultés à obtenir, comprendre et utiliser des informations sur la santé (Kickbusch, 2001). La littératie en santé est particulièrement complexe à évaluer, car des études ont montré qu'elle n'est pas liée à l'âge, au revenu ou à l'éducation des patients (Van der Vaart et al., 2011) et il n'y a toujours pas de consensus sur sa définition (Huhta et al., 2018). Toutefois, la plupart des définitions incluent les connaissances antérieures en santé dans le concept de la littératie en santé (Huhta et al., 2018 ; Sørensen, Van den Broucke, et al., 2012). Williams et al. (1998) ont démontré le lien entre les connaissances et la littératie en santé par le biais d'expérimentations avec des patients souffrant de maladies chroniques. Il semblerait qu'il soit nécessaire d'utiliser les connaissances sur le fonctionnement du corps et des traitements pour utiliser les informations sur la santé. Ainsi, il est possible de suggérer que de meilleures connaissances en santé pourraient conduire à une meilleure utilisabilité des équipements médicaux.

Les connaissances sont souvent définies comme une croyance vraie et justifiée. Une réponse correcte ou incorrecte est interprétée comme signifiant simplement qu'une personne sait ou ne sait pas quelque chose. Il est nécessaire d'intégrer "the test-taker's certainty"³¹ (Hunt, 2003) dans le test afin de prendre en compte toutes les dimensions de son savoir. Nous cherchons à mesurer spécifiquement les connaissances sur le corps humain en relation avec les dispositifs médicaux choisis, c'est-à-dire à capter la compréhension du fonctionnement de la pression

³¹ Traduction en français : « *la certitude du participant* »

artérielle en relation avec l'utilisation d'un tensiomètre et du fonctionnement de l'oxygénation du sang en relation avec l'utilisation de l'oxymètre de pouls.

3. Objectifs et hypothèses

Comme nous l'avons mentionné précédemment, des études théoriques sont nécessaires pour mieux comprendre comment les caractéristiques des utilisateurs influencent l'utilisabilité afin de concevoir des dispositifs plus sûrs pour les patients. L'âge, la technophilie, le niveau de littératie en santé, la motivation ont prouvé leur influence sur l'utilisabilité (Figure 35). À notre connaissance, aucune étude n'a encore mis en évidence le lien empirique entre les connaissances en santé et l'utilisabilité. Nous proposons donc d'analyser cette relation en contrôlant les trois autres variables (i.e., en ayant recours à un environnement contrôlé et une population jeune, en bonne santé et familière de la technologie), afin de limiter l'impact sur l'utilisabilité.

L'objectif de notre étude expérimentale est d'examiner la relation entre les connaissances antérieures en santé et l'utilisabilité des DMD. Nous supposons que les participants ayant de bonnes connaissances antérieures en santé utiliseront ces dispositifs avec une meilleure efficacité (H1), efficacité (H2) et satisfaction (H3) dans leur utilisation.

4. Méthodes

4.1. Participants

Cent cinquante-trois étudiants de premier cycle en psychologie (âge moyen = 20,72 ans ; ET = 1,65 ; rang : 18 - 32 ans ; 41 hommes ; 112 femmes) de l'Université de Picardie Jules Verne à Amiens (France) ont participé à l'expérience. Ils ont été recrutés à l'université entre mars et avril 2018 et ont été informés qu'ils allaient devoir tester des technologies de la santé. Quatre participants ont dû être retirés de l'analyse en raison de problèmes techniques. Tous les participants étaient de langue maternelle française et ont signé un formulaire de consentement éclairé. Les données recueillies sur les participants étaient anonymes. Cette recherche était conforme au code d'éthique de l'American Psychological Association. Un examen et une approbation complets n'étaient pas nécessaires selon les directives de nos institutions et les réglementations nationales. Les participants n'ont reçu aucune compensation financière pour leur participation. Nous avons choisi cette population pour éviter les biais liés à l'âge

(Sonderegger et al., 2016), et pour avoir des capacités homogènes en technophilie (Georgsson & Staggers, 2016). En outre, ce type de population a très peu d'expérience dans l'utilisation de dispositifs médicaux, ce qui permet de stabiliser nos résultats et d'éviter les biais liés à la population.

4.2. Matériel et mesures

4.2.1. Dispositifs médicaux et tâches

Les participants ont utilisé deux appareils médicaux destinés au grand public (Figure 36) : un oxymètre de pouls sans fil (iHealth Oximeter PO3) et un tensiomètre de poignet sans fil mesurant la pression artérielle (iHealth BP7). Le tensiomètre, une fois la mesure terminée, affiche la pression systolique (SYS) et diastolique (DIA) en mmHg, ainsi que le pouls à côté d'un pictogramme en forme de cœur et de l'icône PUL. L'oxymètre de pouls indique le taux d'oxygène dans le sang en % SpO2, ainsi que la fréquence du pouls en PR bpm (battement par minute). Les mesures et les indicateurs de chaque appareil restent affichés pendant quelques secondes.



Figure 36. À gauche (a) le tensiomètre de poignet sans fil (iHealth BP7) et, à droite (b) l'oxymètre de pouls (iHealth Oximeter PO3).

Les participants devaient accomplir les 10 tâches du tensiomètre et les 9 tâches de l'oxymètre de pouls présentées ci-dessous (Tableau 15).

Tableau 15. Tâches de l'utilisateur avec un tensiomètre et un oxymètre de pouls.

	Tensiomètre	Oxymètre de pouls
Tâche 1	Allumer l'appareil	
Tâche 2a	Positionner correctement l'appareil	
Tâche 2b	Incliner l'avant-bras en suivant les flèches lumineuses pour aider à trouver le bon angle	
Tâche 3	Commencer la prise de mesure	
Tâche 4	Rester immobile pendant la prise de mesure	
Tâche 5	Inscrire la mesure sur une feuille	
Tâche 6	Interpréter les symboles des appareils (par exemple, sys pour systolique, dya pour diastolique et bpm pour battement par minute)	Interpréter les symboles sur les appareils (par exemple SPO2 pour le niveau d'oxygène et bpm pour le battement par minute)
Tâche 7	Interpréter les résultats	
Tâche 8	Enlever le dispositif	
Tâche 9	Éteindre le dispositif	

4.2.2. Questionnaires

Les participants ont été invités à répondre à des questionnaires avant et après avoir accompli la tâche, de la manière suivante :

- *Questionnaire sur les connaissances antérieures en santé (avant la tâche)* : Pour évaluer les connaissances antérieures, nous avons utilisé un questionnaire conçu en collaboration avec des professionnels de la médecine et de l'éducation (cf. Annexe 9 : Questionnaire sur les connaissances antérieures en santé en français (40 items), p. 392). Les participants devaient indiquer un certain nombre d'affirmations comme étant vraies, fausses ou je ne sais pas (e.g., le cœur agit comme une pompe). Nous proposons une version traduite (cf. Annexe 10 : Questionnaire sur les connaissances antérieures en santé en anglais (40 items), p. 394) du questionnaire en anglais ;

- *Informations personnelles / Questionnaire sur l'utilisation de la technologie (avant la tâche)* : Ce questionnaire comportait trois informations personnelles (âge, sexe, niveau d'éducation) et deux éléments, adaptés d'Agarwal et Prasad (Agarwal & Prasad, 1998) relatifs à l'utilisation et à la volonté d'explorer les innovations en matière de technologies de l'information par le

participant (e.g., « *laquelle de ces technologies utilisez-vous et à quelle fréquence ?* ») Sur une échelle de Likert à 5 points, les réponses possibles allaient de « *Jamais* » à « *Très souvent* » ;

- *System Usability Survey (SUS) (après la tâche)* : Cet instrument en 10 points, visant à enregistrer des évaluations subjectives de l'utilisabilité (Brooke, 1996 ; Lewis & Sauro, 2009), est un outil "*quick and dirty*" avec cinq options de réponse allant de « *Tout à fait d'accord* » à « *Tout à fait en désaccord* ». Nous avons utilisé la version modifiée décrite dans (Bangor et al., 2008). Cette version a encore été modifiée en changeant le mot "système" par "dispositif médical". Ce type de changement n'a pas d'impact sur la validité ou la fiabilité de l'instrument d'enquête (Sauro, 2011).

4.3. Procédure

Pour les participants, la session a été divisée en deux parties. Chaque partie a duré environ 15 minutes. Dans la première partie, les participants ont rempli le questionnaire Informations personnelles / Utilisation de la technologie, suivi du questionnaire sur les connaissances antérieures en santé (cf. Annexe 9 : Questionnaire sur les connaissances antérieures en santé en français (40 items), p. 392). Ce temps a également permis aux participants de se reposer pour la mesure de la pression artérielle qui a suivi. Dans la deuxième partie, ils ont été invités par les expérimentateurs à manipuler dans un ordre aléatoire deux DMD (l'oxymètre de pouls et le tensiomètre). La seule instruction donnée aux participants était d'effectuer une mesure sur eux-mêmes et d'enregistrer les résultats sur leur feuille de remise. Les participants ont été filmés pendant la session. L'expérience a été menée avec un participant par session dans une pièce calme avec un minimum de distraction. Ils n'ont reçu aucune formation ou démonstration préalable de la part de l'expérimentateur. L'expérimentateur n'est intervenu qu'en cas de problèmes techniques (e.g., un problème de batterie). À la fin de chaque manipulation, les participants ont répondu à des questions sur leur mesure, y compris des questions sur la façon d'interpréter les données, et ils ont rempli le questionnaire SUS (voir ci-dessus).

4.4. Cotation et analyse statistique

Les vidéos ont été analysées par deux experts évaluateurs à l'aide du logiciel BORIS (Behavioral Observation Research Interactive Software). *L'efficacité*, *l'efficience* et la *satisfaction* ont été mesurées comme suit selon la norme ISO 9241-11: 2018.

4.4.1. Mesures objectives de l'utilisabilité

Afin de mesurer l'*efficacité* et l'*efficience*, nous avons utilisé les mêmes paramètres que les études de Georgsson et Staggars (2016) et de Sheehan et Lucero (2015).

La mesure de l'efficacité a été analysée en évaluant le nombre d'erreurs de manipulation (e.g., le participant ne met pas le brassard de surveillance dans la bonne position). Quatre erreurs principales ont été détectées pour le tensiomètre et trois pour l'oxymètre (e.g., le participant n'accroche pas le dispositif dans la bonne position). Une grille de notation a été utilisée pour identifier les erreurs de manipulation. Les participants répétaient parfois la même erreur plusieurs fois, et nous avons enregistré le nombre cumulatif d'erreurs. Le nombre d'erreurs est compris entre 0 et 4, sauf lorsque les participants ont répété successivement la même erreur. La moyenne de toutes les erreurs a ensuite été calculée.

La mesure de l'efficacité a été analysée par le temps de traitement (en secondes). Le temps de manipulation a été mesuré à partir du moment où les participants touchent le dispositif pour la première fois jusqu'au moment où ils l'éteignent à nouveau après avoir pris la mesure.

4.4.2. Mesures subjectives de l'utilisabilité

Afin de mesurer la satisfaction, nous avons utilisé l'échelle d'utilisation du système (SUS) comme d'autres auteurs (e.g., Georgsson & Staggars, 2016 ; Liang et al., 2018 ; Scott et al., 2017). Les scores ont été calculés selon les directives de Brooke (1996). Le score SUS s'étend de 0 à 100, les scores les plus bas indiquant une plus faible utilisabilité.

4.4.3. Mesure des connaissances antérieures en santé

Ce questionnaire (cf. Annexe 9 : Questionnaire sur les connaissances antérieures en santé en français (40 items), p. 392) comprend 40 items et est divisé en deux parties, les 20 premiers items concernent la pression artérielle et les 20 autres items concernent l'oxygénation du sang. Chaque réponse/item correcte du questionnaire sur les connaissances antérieures en santé permettait de comptabiliser 1 point, et chaque réponse incorrecte n'apportait pas de point, ce qui donne un score total compris entre 0 et 40.

4.4.4. Lectures des données de santé

Nous avons examiné si les participants étaient capables de lire correctement leur pression artérielle et s'ils étaient capables d'annoter ces deux premières mesures. Les données de santé enregistrées par les DMD ne sont pas explicitement indiquées. Par exemple, les participants

doivent lire trois mesures sur le tensiomètre (les deux premières sont les pressions systolique et diastolique et la troisième indique le pouls).

4.4.5. Analyse de données

Les résultats ont été calculés à l'aide de SPSS® version 22 (IBM Corporation, 2013). Les résultats descriptifs ont été comparés aux connaissances antérieures en santé des participants. Afin d'établir des comparaisons entre les caractéristiques des utilisateurs, leurs performances et leur satisfaction, nous avons comparé le genre, l'âge, les connaissances antérieures en santé et la technophilie (expérience dans la technologie et dans les dispositifs médicaux) par rapport à l'*efficacité*, l'*efficience* et le score moyen au SUS des deux dispositifs (tensiomètre et oxymètre de pouls). L'échantillon ne suit pas une distribution normale ; c'est pourquoi le coefficient de corrélation sur les rangs (Spearman's Roh) a été utilisé entre les métriques de la norme ISO et les connaissances antérieures en santé. Nous avons ensuite établi trois groupes (clusters) de connaissances (niveaux faible, moyen et élevé) en utilisant les k-means. Le test non paramétrique de Kruskal-Wallis a été utilisé pour analyser ces groupes.

5. Résultats

5.1. Fiabilité Inter-juges

Nous avons utilisé le Coefficient de Corrélation Intra-classe (CCI) pour vérifier la fiabilité des données quantitatives inter-juges (Elie & Colombet, 2011 ; Shrout & Fleiss, 1979). Un double codage a été effectué sur 33 % des données vidéo. La mesure CCI pour les erreurs des tensiomètres était de 0,962 avec un intervalle de confiance à 95 % de 0,933 à 0,979 ($F(47,47) = 26,461, p < ,001$). La mesure CCI pour le temps de manipulation (pour le moniteur de pression sanguine) était de 0,995 avec un intervalle de confiance à 95 % de 0,992 à 0,997 ($F(47,47) = 261,275, p < ,001$). La mesure de CCI pour les erreurs de l'oxymètre de pouls était de 0,936 avec un intervalle de confiance à 95 % de 0,887 à 0,964 ($F(47,47) = 15,732, p < ,001$). La mesure CCI du temps de manipulation (pour l'oxymètre de pouls) était de 0,995 avec un intervalle de confiance à 95 % de 0,992 à 0,997 ($F(47,47) = 261,275, p < ,001$).

5.2. Consistance interne du questionnaire de connaissances antérieures en santé

La méthode d'évaluation communément utilisée pour évaluer la validité d'un test est la méthode test-retest. Cette méthode consiste à administrer deux fois le même test aux mêmes

personnes, mais à des temps différents. Si le test est fiable, la mesure de la corrélation entre le test et le test-retest sera élevée. Cependant, cette méthode favorisant la stabilité de l'évaluation semble inappropriée dans le cas de notre questionnaire de connaissances. En effet, les connaissances peuvent fluctuer dans le temps. Par conséquent, une méthode d'évaluation de la fiabilité qui n'implique pas un test-retest est nécessaire dans le cas de notre étude. La cohérence interne est notamment estimée par le coefficient alpha de Cronbach. Elle est considérée comme bonne lorsque la valeur alpha de Cronbach est supérieure à 0,70 (Nunnally, 1978). Le questionnaire sur les connaissances antérieures en santé (cf. Annexe 9 : Questionnaire sur les connaissances antérieures en santé en français (40 items), p. 392) a été pré-testé sur une population de 68 étudiants de premier cycle en psychologie, ce qui nous a permis de confirmer une distribution standardisée des résultats (moyenne = 20, médiane = 20 (50 % de réponses correctes) et ET = 5,3). L'alpha de Cronbach est de 0,72 pour l'échelle de connaissance de la pression artérielle et de 0,76 pour l'échelle de connaissance de l'oxygénation du sang, ce qui est correct pour les deux dimensions du questionnaire.

5.3. Statistiques sur les utilisateurs : caractéristiques sociodémographiques et technophilie

Parmi les 149 participants (Tableau 16), nous avons décidé de ne retenir que ceux qui ont utilisé avec succès les dispositifs médicaux afin de ne pas biaiser les résultats d'efficacité (i.e., 137 participants pour le tensiomètre et 147 participants pour l'oxymètre de pouls). En effet, si un utilisateur manipule l'appareil pendant très longtemps avant d'abandonner définitivement, il n'est pas possible de prendre en compte ces mesures de temps de manipulation (*efficacité*), car cela ne correspondra pas aux critères de mesure de cette variable.

Nous avons commencé par analyser l'impact des caractéristiques de l'utilisateur (e.g., les critères sociodémographiques, l'expérience dans la technologie, l'expérience dans les dispositifs médicaux) en fonction de l'utilisabilité. Nous avons ensuite analysé la corrélation entre les connaissances antérieures en santé du participant et l'utilisabilité pour tester nos hypothèses (efficacité H1, efficacité H2, satisfaction H3) : (1) pour le tensiomètre, (2) pour l'oxymètre de pouls, puis (3) les analyses par groupes de niveau.

L'âge et le niveau d'éducation n'ont pas eu d'impact sur les résultats de l'utilisation en raison des profils homogènes des étudiants à l'exception du genre pour l'efficacité ($U = 1686,5$, $p = ,02$) et la satisfaction ($U = 2756$, $p = ,01$) pour le cas de l'oxymètre de pouls. Le test de

Kruskal-Wallis n'a révélé aucune différence significative entre la technophilie et l'utilisabilité (*efficacité* : $\chi^2 = 0,51, p = ,77$; *efficience* : $\chi^2 = 0,64, p = ,73$ et *satisfaction* : $\chi^2 = 0,24, p = ,88$) pour le tensiomètre. Pour l'oxymètre de pouls, aucune différence significative n'a été observée entre l'expérience informatique et le résultat de l'utilisabilité (*efficacité* : $\chi^2 = 1,83, p = ,4$; *efficience* : $\chi^2 = 1,68, p = ,43$). Cependant, les participants qui ont déjà utilisé un oxymètre de pouls sont significativement plus satisfaits, car ils ont un meilleur score SUS (*satisfaction* : $\chi^2 = 6,036 ; p = ,049$). Les détails de ces mesures sont disponibles en Annexe (cf. Annexe 11 : Détails des caractéristiques de l'utilisateur sur la base des mesures d'utilisabilité du tensiomètre et de l'oxymètre de pouls, p. 396).

5.4. Résultats de l'évaluation de l'utilisabilité

5.4.1. Tensiomètre

a. Mesure de l'utilisabilité

Cent trente-sept utilisateurs novices ont pu utiliser le dispositif avec succès, c'est-à-dire qu'ils ont obtenu une mesure de leur pression artérielle (Tableau 17). Douze participants ont abandonné, face à la complexité de l'utilisation de l'appareil. Ce résultat indique un manque d'efficacité. Pour le reste de l'échantillon, les utilisateurs ont commis environ 0,77 erreur (ET = 1,49). Les utilisateurs ont manipulé le tensiomètre pendant 260,91 secondes (ET = 107,12). Le score SUS moyen était de 72,4 (ET = 21,07), ce qui indique un score "satisfaisant" (Bangor et al., 2008). En ce qui concerne les résultats descriptifs, ce dispositif semble avoir un manque probant d'utilisabilité pour une population n'ayant aucune formation préalable. De plus, parmi les participants ayant réussi à utiliser correctement le tensiomètre, seulement 12 participants (9 %) ont pu lire correctement leurs mesures de tension artérielle.

Tableau 16. Les caractéristiques sociodémographiques et la technophilie (expérience dans la technologie de l'information et dans les dispositifs médicaux) sur un échantillon de 149 participants.

Variable (N=149)		n(%)
Caractéristiques sociodémographiques		
Âge, moyenne (écart-type)		20,72 (1,65)
Genre	Homme	37 (24,83)
	Femme	112 (75,17)
Niveau d'éducation	1 ^{ère} année	2 (1,34)
	2 ^{ème} année	88 (59,06)
	3 ^{ème} année	59 (39,59)
Expérience dans les technologies de l'information		
Cluster de l'expérience dans les technologies de l'information	Élevé	59 (43,1)
	Modéré	47 (34,3)
	Faible	31 (22,6)
À l'aise avec l'ordinateur	Élevé	88 (59,06)
	Modéré	60 (40,27)
	Faible	1 (0,67)
Fréquence d'utilisation de l'ordinateur	Tous les jours (très souvent)	103 (69,13)
	Plusieurs fois par semaine (souvent)	39 (26,17)
	Quelques fois (rarement)	6 (4,03)
	Jamais	0 (0)
À l'aise avec un téléphone portable/smartphone	Élevé	120 (80,54)
	Modéré	29 (19,46)
	Faible	0 (0)
Fréquence dans l'utilisation d'un téléphone/smartphone	Tous les jours (très souvent)	125 (83,89)
	Plusieurs fois par semaine (souvent)	19 (12,75)
	Quelques fois (rarement)	4 (2,68)
	Jamais	0 (0)
À l'aise avec les objets connectés	Élevé	10 (6,71)
	Modéré	31 (20,8)
	Faible	108 (72,48)

Fréquence dans l'utilisation des objets connectés	Tous les jours (très souvent)	3 (2,01)
	Plusieurs fois par semaine (souvent)	4 (2,68)
	Quelques fois (rarement)	46 (30,87)
	Jamais	95 (63,76)
Expérience dans l'utilisation des dispositifs médicaux		
Tensiomètre	Oui	86 (57,72)
	Non	63 (42,28)
Oxymètre de pouls	Oui	20 (13,42)
	Non	129 (86,58)

Tableau 17. Résultats des mesures d'utilisabilité du tensiomètre et corrélation de Spearman entre l'utilisabilité et les connaissances antérieures en santé.

Tensiomètre (N=137)	Efficacité (erreurs)	Efficience (secondes)	Satisfaction (SUS)	Lecture de la mesure
Moyenne	0,77	260,91	72,4	9%
ET	1,49	107,12	21,07	0,28
Rang (min-max)	0 - 8	90,5 - 681	27,5 - 100	0 - 1
Corrélation de Spearman	$r = -0,191$	$r = -0,104$	$r = 0,146$	$r = 0,302$
Sig. (unilateral)	$p = ,026^*$	$p = ,23$	$p = ,088$	$p < ,001^{***}$

*. Les corrélations sont significatives à ,05.

***. Les corrélations sont significatives à ,001.

Tableau 18. Résultats des mesures d'utilisabilité de l'oxymètre de pouls et corrélation de Spearman entre l'utilisabilité et les connaissances antérieures en santé.

Oxymètre de pouls (N=147)	Efficacité (erreurs)	Efficience (secondes)	Satisfaction (SUS)	Lecture de la mesure
Moyenne	0,99	158,42	71.52	65 %
ET	0,92	75,75	17.29	0,48
Rang (min-max)	0 - 6	24,8 – 458,8	15-100	0 - 1
Correlation de Spearman	$r = -0,263$	$r = -0,2062$	$r = 0,195$	$r = 0,018$
Sig. (unilateral)	$p = ,001^{**}$	$p = ,45$	$p = ,017^*$	$p = ,82$

*. Les corrélations sont significatives à ,05.

** . Les corrélations sont significatives à ,01.

b. Corrélation entre l'utilisabilité et les connaissances antérieures en santé

En moyenne, les participants ont obtenu un score de 10,16/20 (ET = 2,95, rang = 1 - 16), soit 50,8 % de réponses correctes au questionnaire sur les connaissances antérieures en santé. La corrélation de Spearman entre les connaissances sur la pression artérielle et les mesures d'utilisabilité montre des corrélations faibles et négatives, mais significatives entre le nombre d'erreurs et les connaissances des participants ($r = -0,191, p = ,026$). Les participants ayant un niveau élevé de connaissances en santé ont commis moins d'erreurs. Ces résultats ne sont toutefois pas significatifs pour l'efficacité ($r = -0,104, p = ,22$) et la satisfaction ($r = 0,146, p = ,08$). La majorité des participants n'ont pas su enregistrer leurs données. Nous avons observé que seuls les participants ayant un niveau élevé de connaissances en santé étaient capables d'enregistrer correctement leur pression artérielle ($r = 0,302, p < ,001$). Il s'agit d'une corrélation moyenne, positive et significative qui illustre les limites de ce dispositif médical. Bien que la plupart des participants aient su manipuler le tensiomètre, ils ont eu du mal à lire les résultats.

5.4.2. Oxymètre de pouls

a. Mesure de l'utilisabilité

Cent quarante-sept participants ont pu utiliser l'oxymètre de pouls avec succès (Tableau 18). Seuls deux participants ont abandonné parce qu'ils n'ont pas utilisé l'appareil correctement, et leurs résultats ont été exclus de l'analyse. L'oxymètre semble donc plus facile à utiliser que le tensiomètre. L'oxymètre de pouls a été abandonné par seulement 2 participants, en revanche, les participants ont commis plus d'erreurs ($M = 0,99, ET = 0,92$). Le score SUS moyen était de 71,52 (ET = 17,29), ce qui signifie un score "satisfaisant" (Bangor et al., 2008). En outre, les relevés de données étaient assez bons. En moyenne, 65 % des participants ont pu lire correctement leur taux d'oxygène.

b. Corrélation entre l'utilisabilité et les connaissances antérieures en santé

Pour les connaissances antérieures en santé sur l'oxygénation du sang, les participants ont obtenu en moyenne un score de 5,28/20 (ET = 2,77, rang = 0 - 15), ce qui correspond à 26,3 % de bonnes réponses. La corrélation de Spearman entre les connaissances de la pression sanguine et les métriques de l'utilisabilité norme ISO (ISO 9241-11, 2018) montre des résultats similaires pour l'oxymètre de pouls par rapport au tensiomètre. Il existe une corrélation faible, négative et significative entre le nombre d'erreurs et les connaissances des participants ($r = -0,263, p = ,001$). On peut en déduire que les participants ayant de meilleures connaissances en santé font moins d'erreurs et sont donc plus efficaces dans la manipulation du dispositif. On observe également un effet significatif entre la satisfaction et les connaissances antérieures en

santé et des participants ($r = 0,195, p = ,017$). Cependant, nous n'observons pas de corrélation sur la lecture des données de santé et les connaissances antérieures en santé ($r = 0,018, p = ,82$).

5.5. Analyse par groupes/cluster (k-means)

Compte tenu des faibles corrélations observées, nous avons envisagé la possibilité d'un effet de seuil. À cette fin, nous avons créé des groupes (k-means), en séparant les participants en trois groupes d'affectation selon leur niveau de connaissances antérieures en santé. Nous avons ensuite examiné les trois groupes de connaissances (faible, moyen, élevé) en fonction de leurs résultats en termes *d'efficacité*, *d'efficience* et de *satisfaction* pour le tensiomètre (Tableau 19) et pour l'oxymètre de pouls (Tableau 20).

En ce qui concerne l'efficacité, le test de Kruskal-Wallis a révélé une différence significative entre le nombre d'erreurs de manipulation de l'oxymètre de pouls (efficacité) et les différents groupes de connaissances ($\chi^2 = 10,89, p = ,004$). Le groupe Faible ($M = 1,3, ET = 0,88$) fait significativement plus d'erreurs que les deux autres groupes. Nous avons ensuite effectué la même analyse avec l'efficience. L'efficience ne révèle aucune différence significative ($\chi^2 = 1,97, p = ,37$). Toutefois, cet effet de seuil se retrouve aussi dans la satisfaction basée sur le score SUS ($\chi^2 = 7,452, p = ,024$). Ces résultats ne sont pas transférables au tensiomètre.

Tableau 19. Mesures des connaissances antérieures en santé par groupes en fonction des métriques de l'utilisabilité pour le tensiomètre (test de Kruskal-Wallis).

Tensiomètre (N=137)	Efficacité (erreurs)	Efficienc e (secondes)	Satisfaction (SUS)	Lecture de la mesure
Faible (N = 26 ; moyenne = 8,73/40 ; ET = 2,51)	M = ,92	M = 286,42	M = 70,38	M = 0
	ET = 1,44	ET = 110,16	ET = 16,46	ET = 0
	Rang = 0 - 7	Rang = 112,1 – 534,3	Rang = 27,5 – 100	Rang = 0 - 0
	M = 0,86	M = 257,53	M = 72,39	M = 0,05
Élevé (N = 45 ; moyenne = 20,93/40 ; ET = 2,9)	ET = 1,62	ET = 119,41	ET = 16,46	ET = 0,21
	Rang = 0 - 8	Rang = 90,5 – 681,04	Rang = 27,5 – 97,5	Rang = 0 - 1
	M = ,56	M = 251,14	M = 73,82	M = ,2
	ET = 1,01	ET = 83,72	ET = 16,44	ET = ,41
Kruskal-Wallis	Rang = 0 - 4	Rang = 127,5 – 494,2	Rang = 40 - 100	Rang = 0 - 1
	$\chi^2 = 2,50$ $p = ,29$	$\chi^2 = ,92$ $p = ,23$	$\chi^2 = 1,01$ $p = ,6$	$\chi^2 = 3,47$ $p = ,18$

*. Les corrélations sont significatives à ,05.

**. Les corrélations sont significatives à ,01.

Tableau 20. Mesures des connaissances antérieures en santé par groupes en fonction des métriques de l'utilisabilité pour l'oxymètre (test de Kruskal-Wallis).

Oxymètre de pouls (N=147)	Efficacité (erreurs)	Efficienc e (secondes)	Satisfaction (SUS)	Lecture de la mesure
Faible (N= 44 ; Moyenne = 9,77/40 ; ET = 2,76)	M = 1,34	M = 168,38	M = 69,86	M = 0,89
	ET = 0,88	ET = 76,98	ET = 16,38	ET = 0,32
	Range = 0 - 4	Range = 48,2 – 385,4	Range = 32 - 95	Range = 0 - 1
Moyen (N= 29 ; Moyenne = 14,1/40 ; ET = ,77)	M = 0,69	M = 163,68	M = 71,45	M = 0,86
	ET = 0,6	ET = 77,47	ET = 23,23	ET = 0,35
	Range = 0 - 7	Range = 71,4 – 452,5	Range = 15 - 98	Range = 0 - 1
Élevé (N= 74 ; moyenne = 19,31/40 ; ET = 3,28)	M = 0,92	M = 150,44	M = 72,73	M = 0,91
	ET = 1,04	ET = 74,49	ET = 15,16	ET = 0,29
	Rang = 0 - 6	Rang = 78 - 458,8	Rang = 30 - 100	Rang = 0 - 1
Kruskal-Wallis	$\chi^2 = 10,89$ $p = ,004^{**}$	$\chi^2 = 1,97$ $p = ,37$	$\chi^2 = 7,452$ $p = ,024^*$	$\chi^2 = 0,031$ $p = ,98$

*. Les corrélations sont significatives à ,05.

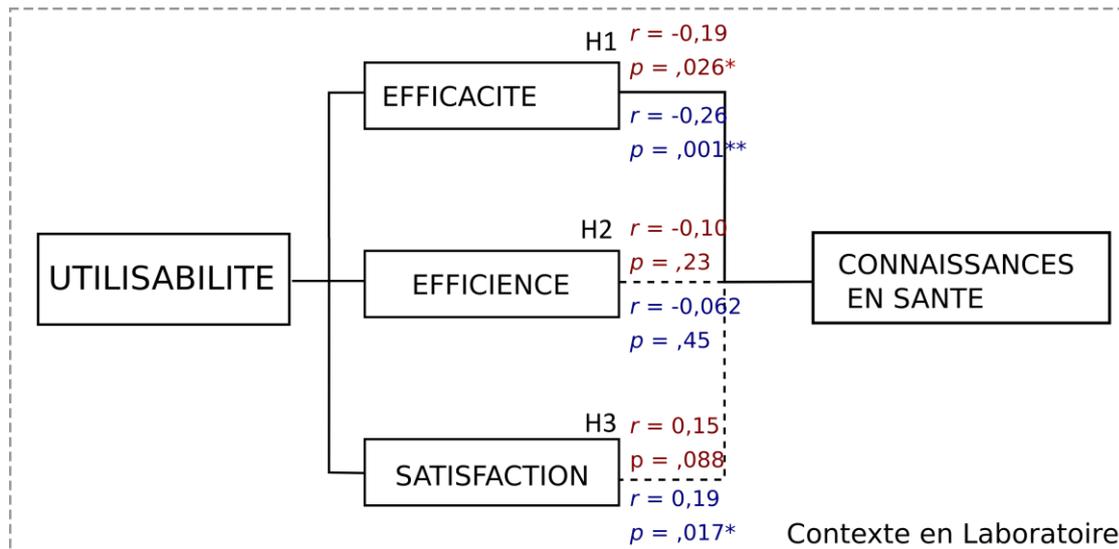
**. Les corrélations sont significatives à ,01.

6. Discussion

6.1.Principales contributions

L'objectif de cette étude était d'explorer comment les connaissances antérieures en santé, considérées comme faisant partie de la littératie en santé (Huhta et al., 2018 ; Sørensen, Broucke, et al., 2012 ; Williams et al., 1998), pouvaient avoir un impact sur le résultat de l'utilisabilité (*efficacité, efficience et satisfaction*) des DMD. Les résultats soutiennent l'hypothèse centrale de la présente étude, à savoir que de meilleures connaissances antérieures en santé entraînent une meilleure utilisabilité. Les participants ayant un niveau élevé de connaissances antérieures en santé sont plus efficaces que ceux ayant un faible niveau. Plus précisément, les participants ayant des connaissances sur le fonctionnement de la pression artérielle dans le corps humain font significativement moins d'erreurs de manipulation lorsqu'ils utilisent le tensiomètre. Les connaissances sur le fonctionnement du corps, comme la pression artérielle, aideraient à mieux comprendre le fonctionnement du tensiomètre et à prévenir les erreurs de posture. C'est le cas pour les deux dispositifs testés, cependant seul un effet de seuil est visible dans le cas de l'oxymètre. Les tests effectués par les utilisateurs indiquent également que le tensiomètre est plus difficile à utiliser que l'oxymètre. La lecture du résultat d'une mesure semble être intuitive dans le cas de l'oxymètre, mais pas dans le cas du tensiomètre. La majorité des participants n'ont pas su lire leur tension ce qui indique qu'ils n'ont pas pu l'interpréter. Ainsi, les participants peuvent se servir du matériel correctement, mais ils ont besoin d'aide pour comprendre et interpréter leurs données physiologiques.

À la lumière des observations ci-dessus, nos résultats semblent valider notre hypothèse (Figure 37). Les connaissances antérieures en santé ont un impact sur l'efficacité (H1) et partiellement sur la satisfaction (H3). Cependant, la seconde hypothèse (H2) concernant le lien entre les connaissances antérieures en santé et le temps de manipulation (efficience) de ces deux dispositifs ne peut être validée, car aucun lien significatif n'a pu être observé.



*. Les corrélations sont significatives à ,05.

** . Les corrélations sont significatives à ,01.

Figure 37. Synthèse des résultats de l'influence des connaissances antérieures en santé sur l'utilisabilité (ISO 9241-11:2018). Les résultats indiqués en rouge correspondent au tensiomètre, et ceux en bleu à l'oxymètre de pouls.

Ces résultats peuvent s'expliquer par l'adaptation du modèle de Monkman et Kushniruk (Monkman et al., 2015) en remplaçant le niveau littératie en santé par le niveau de connaissances en santé. L'oxymètre requiert un niveau de connaissances inférieur de la part du sujet. Quant au tensiomètre, cet appareil se porte au poignet, ce qui contredit une croyance répandue selon laquelle la pression artérielle ne peut être surveillée que par le bras. Les connaissances liées à cet appareil exigent donc un plus grand effort de compréhension du fonctionnement d'un tensiomètre, et impliquent une connaissance plus approfondie. Selon le modèle de Monkman et Kushniruk, si les « exigences en matière de connaissances en e-santé » d'un appareil dépassent « le niveau de connaissances en e-santé des consommateurs », l'adoption de l'appareil est alors compromise. Dans ce cadre, la limite de compréhension de l'oxymètre de pouls est susceptible de se situer entre le groupe Faible et le groupe Moyen, ce qui explique l'effet de seuil (Figure 38). Cette interprétation est également soutenue par Paasche-Orlow et Wolf (2007), qui ont précédemment observé un effet de seuil entre la connaissance de la santé et les résultats pour la santé. Cependant, le tensiomètre, contrairement à l'oxymètre de pouls, nécessite un niveau élevé de connaissances dans son utilisation et dans la lecture de ses résultats. Selon le modèle de Monkman et Kushniruk (2015), l'adoption, la valeur et l'utilisation réussie sont liées à la fois au domaine de l'utilisabilité et à celui de l'utilité.

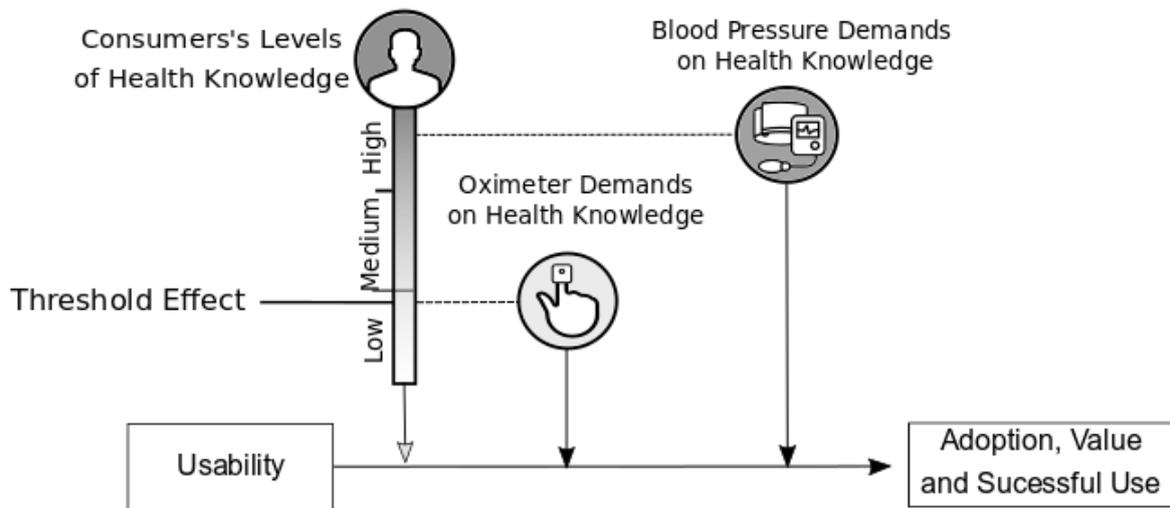


Figure 38. Adaptation du modèle de Monkman et Kushniruk (Monkman et al., 2015) appliqué à l'utilisation du tensiomètre et de l'oxymètre de pouls.

Toutefois, cette étude comprend certaines limites. Premièrement, nous avons réalisé cette expérience sur une population jeune et en bonne santé probablement pas ciblée par l'utilisation de dispositifs médicaux à domicile. Or, la motivation peut affecter nos résultats. Pour O'Brien et Toms (2008), l'utilisabilité est liée à l'expérience d'engagement. Un manque de motivation des participants pourrait donner des mesures plus faibles d'utilisabilité. Deuxièmement, seules les performances des participants qui ont réussi à collecter leurs données de santé avec les appareils ont été analysées. Si ces participants ayant échoué avaient été inclus, les mesures du nombre d'erreurs auraient certainement augmenté et le score SUS aurait diminué. Troisièmement, le questionnaire sur les connaissances antérieures en santé a été conçu sur la base d'un niveau universitaire afin de mieux discriminer les connaissances des étudiants. Cependant, dans la population nationale, la distribution est probablement beaucoup plus large que dans une population universitaire, car cet échantillon n'est pas représentatif de la population nationale. Par exemple, cet échantillon n'inclut pas les personnes ayant des difficultés de langage, de lecture et d'écriture. Ce lien mal compris entre les connaissances antérieures en santé et l'utilisabilité pourrait être approfondi à l'avenir, notamment sur des populations plus représentatives des utilisateurs finaux (telles que les populations plus âgées, sujettes à des pathologies, etc.). Quatrièmement, le choix des équipements a été fait par des professionnels de santé habitués à manipuler des équipements médicaux. Ils ont jugé par eux-mêmes de la simplicité du matériel, même s'il était apparemment complexe à utiliser. Ce constat souligne donc l'importance de tester ce type d'équipement sur des utilisateurs novices, de promouvoir la conception universelle (e.g., Gibbons et al., 2014) et soulève des questions sur la législation

relative à la conformité européenne. Il est possible que certains équipements soient plus simples à utiliser que ceux choisis ici. Ainsi, nous encourageons les études sur les dispositifs médicaux qui sont déjà sur le marché afin d'étudier les éventuelles difficultés d'utilisation par le grand public (Zapata et al., 2015).

7. Conclusion et Perspectives de recherche

7.1. Perspectives de recherche

Il s'agit de la première étude/itération visant à détecter les premiers problèmes d'utilisabilité de deux dispositifs médicaux n'incluant qu'une partie du contexte d'usage. Dans ce contexte d'usage complet, le patient se trouvera dans un contexte postopératoire (avec des douleurs, des nausées, du stress) et aura souvent besoin de l'aide des membres de sa famille pour utiliser l'appareil à domicile. Cette première étude du projet *Smart Angel* sert de base pour comparer les mesures d'utilisabilité sur d'autres contextes spécifiques manipulant les quatre composants de l'utilisabilité (*utilisateurs, tâches, environnement, ressources*). Une première perspective de recherche est utilisée pour définir comment le contexte de la chirurgie ambulatoire influe sur l'utilisabilité du système. Si un patient ayant un niveau élevé de connaissances antérieures en santé commet des erreurs d'utilisation avec un appareil, le contexte d'apprentissage, y compris la douleur et/ou le stress, peut être la cause de cette erreur et donc d'un mauvais résultat d'utilisabilité.

Une deuxième perspective de recherche se concentre sur l'acceptabilité du dispositif, pour les patients, mais aussi pour le personnel médical. Au cours de cette période, les chercheurs se concentreront sur l'impact de la mise en œuvre du système sur le parcours du patient et sur l'organisation de l'hôpital. À cette fin, une étude hospitalière est menée sur la population des patients en chirurgie ambulatoire. Il est nécessaire d'approfondir nos connaissances en matière d'ergonomie sur les facteurs qui influencent l'acceptabilité afin d'améliorer la sécurité des patients.

7.2. Recommandations pour la conception et l'intégration des dispositifs médicaux au domicile.

Trois types de recommandations basées sur les résultats obtenus peuvent être proposées. Premièrement, les résultats de l'étude actuelle pourraient être utilisés à l'hôpital pour améliorer le suivi des patients impliquant des dispositifs médicaux et pour éviter les erreurs d'utilisation. Ils pourraient déterminer quels dispositifs seraient les plus adaptés au profil de l'individu en

termes de compréhension de la santé (Chaniaud et al., 2018). Ces informations, sur les connaissances antérieures en santé du patient, aideraient les médecins à prendre la décision de recommander un dispositif particulier à ce patient. Toutefois, cela nécessiterait des études supplémentaires sur une population plus représentative de la population nationale.

Deuxièmement, nous avons pu constater, à partir de nos résultats, que les connaissances spécifiques au corps humain (e.g., la fonction cardiaque, la pression sanguine et l'oxygénation du sang) permettaient de réduire le nombre d'erreurs d'utilisation, même si l'individu n'a aucune expérience des dispositifs médicaux. Par conséquent, une éducation thérapeutique basée sur le fonctionnement du corps humain, la maladie, la manière de soigner et le fonctionnement des dispositifs pour répondre aux besoins du patient, serait la clé pour rendre les DMD plus utilisables (Williams et al., 1998). Les professionnels de la santé pourraient fournir aux patients des explications anatomiques et physiologiques adaptées aux DMD choisis sur le fonctionnement du corps afin de prendre ces mesures correctement et de préserver la sécurité des patients.

Troisièmement, nous recommandons aux concepteurs de faire attention aux termes choisis. Il est important de disposer d'une méthodologie rigoureuse lors d'évaluation de l'utilisabilité et de suivre les lignes directrices relatives à l'utilisation d'un langage clair et simple dans les communications sur les soins de santé (e.g., Ruel et al., 2018). Tout comme les professionnels de santé, les concepteurs peuvent également ajouter aux instructions des informations ludiques sur le fonctionnement du corps humain.

7.3. Conclusion

En conclusion, notre étude comporte deux contributions principales. Premièrement, un intérêt scientifique à fournir des connaissances théoriques sur les facteurs impactant l'utilisabilité. En effet, nos résultats indiquent que les connaissances antérieures en santé impactent l'efficacité des DMDs.

Deuxièmement, notre étude a un intérêt appliqué pour aider les concepteurs et le personnel médical à cibler l'importance de fournir des connaissances spécifiques sur le sujet afin de les aider à comprendre le fonctionnement du dispositif. Les résultats montrent qu'il est possible de bien adapter les DMDs à un faible niveau de littératie en santé des patients. C'est le cas, par exemple, de l'oxymètre utilisé dans cette étude : les participants ne connaissaient pas l'appareil, et pourtant ils étaient capables de l'utiliser et de lire leurs résultats de santé.

Il convient de noter que l'étude présentait certaines limites. Notre échantillon était constitué de jeunes adultes ayant un niveau d'éducation élevé et des connaissances en santé probablement supérieures à la moyenne nationale. Il semble évident que l'utilisabilité doit être évaluée sur des groupes d'utilisateurs plus importants et plus diversifiés.

8. Synthèse du Chapitre – Étude 1a

L'étude 1a correspond au **niveau 1** (tâches – utilisateurs) de notre modèle HANDLED, représentant la première itération du cycle de conception du dispositif *Smart Angel*. Cette étude a été conçue afin **d'analyser la relation entre l'utilisabilité de deux dispositifs médicaux à domicile (tensiomètre et oxymètre de pouls)** sélectionnés par le CHU de Nîmes pour intégrer le futur dispositif *Smart Angel* et **les connaissances antérieures en santé par une approche expérimentale**.

Pour cela, nous avons réalisé un test utilisateur auprès de 149 étudiants. Une **approche mixte** (subjective et objective), utilisant divers outils, a été utilisée (observation directe et instrumentée par vidéo, questionnaires). Tout d'abord, les participants ont rempli un questionnaire afin **d'évaluer leurs connaissances antérieures en santé et ont ensuite utilisé les deux dispositifs médicaux de manière randomisée**. Les mesures d'efficience et d'efficacité ont été recueillies par observation instrumentée par vidéo. Les mesures de satisfaction ont été recueillies par questionnaire SUS. Des analyses corrélationnelles et des clusters ont été effectués pour tester la relation entre les caractéristiques sociodémographiques, les connaissances antérieures en santé et l'utilisabilité.

Les résultats indiquent un manque d'utilisabilité pour les deux dispositifs. Pour le tensiomètre (N = 137 participants), les utilisateurs ont commis environ 0,77 erreur (ET = 1,49) et la moyenne du score au SUS était de 72,4 (ET = 21,07), ce qui est considéré comme "satisfaisant". **Douze participants ont abandonné l'utilisation du tensiomètre**. L'oxymètre de pouls (N = 147) **semble plus facile à utiliser**, car il ne compte que **2 abandons**, même si les participants ont commis plus d'erreurs (M = 0,99, ET = 0,92). Le score SUS moyen était de 71,52 (ET = 17,29), ce qui indique un score "satisfaisant". Les résultats ont également montré **une faible corrélation négative et significative uniquement entre l'efficacité des deux dispositifs** (tensiomètre : $r = -0,191$, $p = ,026$; et oxymètre de pouls : $r = -0,263$, $p = ,001$) et les connaissances antérieures en santé. Plus subtilement, nous avons identifié expérimentalement **l'existence d'un niveau seuil** ($\chi^2 = 10,89$, $p = ,004$) sur les connaissances antérieures en santé pour utiliser correctement l'oxymètre de pouls, mais qui fait défaut pour le tensiomètre.

Ces résultats révèlent l'importance de prendre en compte les niveaux de compétences des utilisateurs afin d'adapter le dispositif à ce niveau.

CHAPITRE 3 : ÉTUDE 1b. EFFETS DES DOCUMENTS PROCÉDURAUX D'UNE POPULATION NOVICE SUR L'UTILISABILITÉ DU TENSIONNOMÈTRE

1. Préambule

Ce présent chapitre est tiré d'un article scientifique³² actuellement soumis dans *Cognitive Processing*. Ce travail a été réalisé en collaboration avec Natacha Métayer, ingénieure de recherche sur le Projet *Smart Angel*.

1.1. Contexte et objectif industriels

Le contexte industriel reste le même que dans le chapitre précédent dans la mesure où les objets connectés ont été sélectionnés par le CHU de Nîmes et des premières maquettes introduisant les instructions pour manipuler le tensiomètre et l'oxymètre de pouls sont en cours de conception (Figure 39). Dans un premier temps, nous avons assisté nos partenaires industriels dans la rédaction des instructions puis nous avons voulu les évaluer : premièrement sur la base d'heuristiques (Ganier, 2002) et deuxièmement lors de tests utilisateurs, afin de proposer le/les format(s) le/les plus adapté(s) et d'améliorer l'utilisabilité du dispositif.

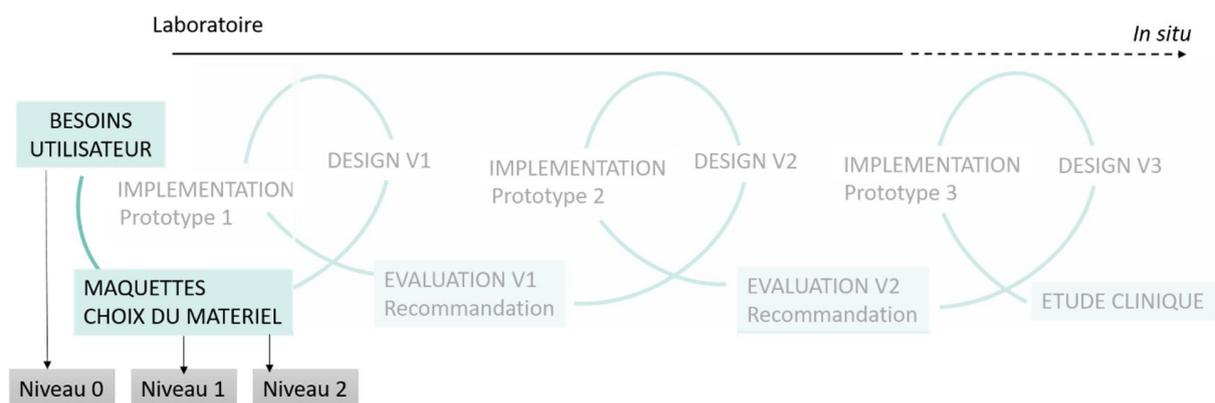


Figure 39. Étude 1b correspondant au niveau 2 du processus de conception du dispositif *Smart Angel*.

³² Chaniaud, N., Métayer, N., Loup-Escande, E., & Megalakaki, O. (Accepted). Effects of mode presentation on the execution of Home Medical Device instructions: the case of a blood pressure monitor. (Revue *Cognitive Processing*)

1.2. Contexte et objectif théoriques

Cette étude réalisée correspond au niveau 2 du processus de conception du dispositif *Smart Angel* (Figure 39). Nous avons fait le choix de nous concentrer sur les documents procéduraux, ressources réutilisables (ISO 9241-11), car la réglementation du dispositif médical (classé IIa) impose la présence d'une notice d'utilisation (§5.7.3. et 5.8. IEC 62366-1). Cette notice fait partie du contexte d'usage et plus exactement des ressources réutilisables et impacte l'utilisabilité. Cette notice est composée d'instructions (aussi appelés documents procéduraux) sollicitant un apprentissage procédural. L'apprenant utilise alors les instructions présentées sous des formats variés (image, texte, audio) dans le but de réaliser une action motrice (Anderson, 1982) afin de faire fonctionner le tensiomètre. Dans cette présente étude, nous avons fait varier les formats de présentation des documents procéduraux (unimodal : texte, image, audio ; multimodal : texte / audio avec image) afin de déterminer si l'un d'eux pouvait améliorer l'utilisabilité du tensiomètre.

Cette étude a porté sur les deux objets connectés, mais les résultats de l'oxymètre de pouls n'ont révélé aucun effet significatif sur l'utilisabilité (voir Figure 40 tirée de Chaniaud et al., 2019). Cela implique que l'oxymètre de pouls peut être utilisé sans notice d'utilisation. C'est pourquoi nous avons fait le choix de présenter dans cette étude uniquement les résultats du tensiomètre.

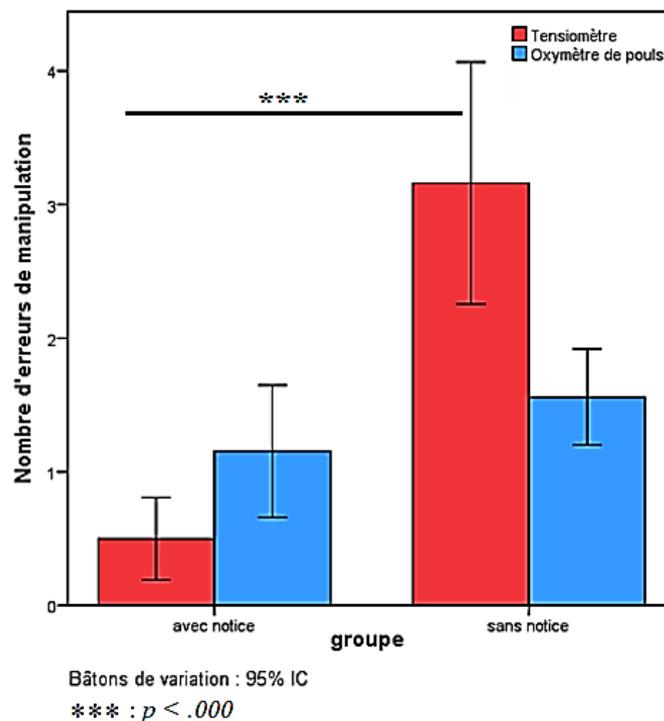


Figure 40. Graphique en barres avec intervalle de confiance représentant le nombre d'erreurs de manipulation du tensiomètre et de l'oxymètre de pouls en fonction de la présence d'une notice – tiré de Chaniaud et al., 2019).

Dans l'étude précédente (cf. Partie II, Chapitre 2, p. 132), nous avons déjà évalué la satisfaction de ces objets. C'est pourquoi nous n'avons pas réévalué cette métrique. Cependant, nous avons souhaité détailler davantage les mesures de l'efficacité en nous intéressant au temps de lecture des instructions. Nous avons alors décomposé cette métrique en trois mesures : temps de lecture des instructions, temps d'exécution et temps total.

La Figure 41 modélise les différents éléments évalués dans cette étude basée sur la matrice : *tâches – utilisateurs – ressources* .

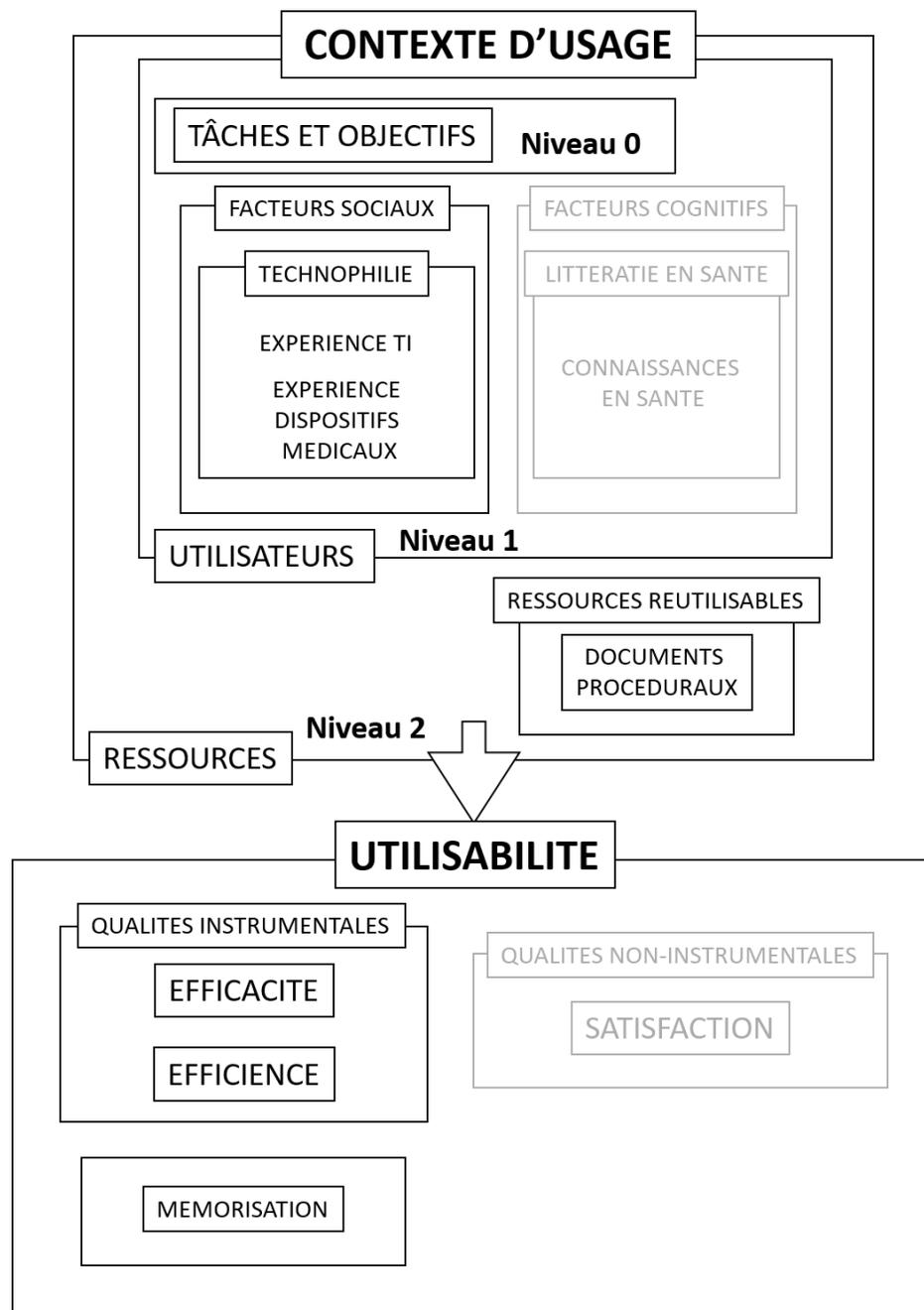


Figure 41. Attributs éprouvés lors de l'étude 1b du contexte d'usage sur l'utilisabilité (efficacité, efficience et mémorisation).

2. Introduction

Les documents multimédias sont actuellement utilisés dans de nombreux domaines à des fins diverses (enseignement, procédures, etc.). Ils peuvent être composés d'images, de textes, d'audio, de vidéo ou de combinaisons de ceux-ci présentés sous différents formats (fragmenté, différé, etc.). Alors que la technologie numérique envahissait notre environnement, les chercheurs se sont concentrés sur le traitement des documents multimédias (e.g., Jamet, 2012). Le modèle de Mayer (2001) appelé CTML (Cognitive theory of multimedia learning) est probablement le plus utilisé dans ce domaine. Ce modèle définit la manière dont les individus apprennent à partir de mots et/ou d'images. Mayer (2001) émet l'hypothèse que l'utilisation "d'un double canal" permet d'augmenter les performances. L'auteur suppose que les individus disposent de canaux séparés pour le traitement des informations verbales et visuelles, en fonction des capacités limitées de la mémoire de travail (Baddeley, 1992). L'information sélectionnée et traitée par ces processus de double codage est combinée avec les connaissances préalables déjà disponibles dans la mémoire à long terme, ce qui permet la construction du modèle mental (verbal et pictural). Ce processus d'intégration de nouvelles connaissances dans les connaissances antérieures n'est pas séquentiel, mais il fonctionne de manière itérative avant d'être intégré dans la mémoire à long terme (Mayer, 2003). Selon ce modèle, lorsque des informations multimédias et/ou multimodales sont présentées, les éléments contenus dans ces différentes sources d'information seront traités par deux canaux différents. Chacun de ces canaux possède des ressources cognitives distinctes, la répartition de la charge de traitement entre ces deux canaux est possible, ce qui évite la surcharge cognitive. L'utilisation de deux canaux de traitement soulage la charge cognitive, car il y a alors une répartition de la charge de traitement entre les deux canaux. C'est pourquoi la présentation d'un texte ou d'une image uniquement sera moins intégrée dans la mémoire qu'un texte combiné à une image, car les textes avec images nécessitent deux canaux de traitement différents. C'est le principe du multimédia. Cependant, la présentation d'un texte écrit avec une image (présentation exclusivement visuelle) sera moins intégrée dans la mémoire qu'un texte oral avec une image. En effet, les mots écrits sont traités à la fois par le canal visuel/image et ensuite par le canal auditif/verbal. En revanche, un texte oral est directement traité par le canal auditif, de sorte que dès que l'information pertinente est sélectionnée, la charge de traitement est répartie entre les deux canaux, ce qui facilite le traitement des informations orales combinées à des informations picturales. C'est le *principe de modalité*.

Le modèle de Mayer pourrait être validé dans le cas de l'apprentissage déclaratif (Butcher, 2006), mais n'est pas nécessairement applicable à l'apprentissage procédural (e.g., Lemarié, Castillan, & Eyrolle, 2017) qui dépend fortement du contexte d'utilisation, y compris des caractéristiques de l'utilisateur (e.g., Ganier, Gombert, & Fayol, 2000). Dans le cas de l'apprentissage procédural, le type et l'organisation des informations permettent à l'utilisateur de développer un modèle mental de la tâche à accomplir (Guthrie, Bennett, & Weber, 1991). L'apprenant utilise des instructions écrites et/ou picturales transformées en actions motrices pour faire fonctionner l'équipement, ce que Anderson (1982) appelle l'apprentissage procédural. La spécificité de l'apprentissage procédural par rapport à l'apprentissage déclaratif est que les documents procéduraux doivent être compris, interprétés et mémorisés jusqu'à ce qu'une action soit effectuée (Heurley, 1994) alors que l'apprentissage déclaratif ne correspond qu'aux connaissances de l'apprenant. Il existe différentes stratégies de présentation de l'information pour améliorer les performances des apprenants dans le cas de l'apprentissage procédural comme l'atomisation de l'action (e.g., Jannin, Ganier, & De Vries, 2019). Cette fragmentation de l'information est décrite par Vermersch (1985) comme une alternance entre l'activité de collecte d'informations et l'exécution de la tâche. Cette stratégie consiste à réduire le nombre d'éléments que l'apprenant doit traiter simultanément en segmentant l'information, réduisant ainsi les tâches à des segments gérables (Fraser, Ayres, & Sweller, 2015).

De nombreuses études ont examiné les avantages de différents formats pour différents types d'objets. Par exemple, Vermersch (1985) s'est intéressé aux formats de présentation des recettes de cuisine, Ganier (2002) à celui de la cocotte-minute, Mykityshyn, Fisk, et Rogers (2002) à la manipulation d'un glucomètre ou plus récemment à l'apprentissage dans l'élaboration d'une coiffure (Lemarié et al., 2017). Tous ces auteurs ont pu souligner l'importance du contexte d'utilisation (Ganier et al., 2000), y compris par exemple les caractéristiques de l'utilisateur (connaissances préalables) et pas seulement le format de présentation. Cependant, ces études antérieures montrent des résultats contradictoires selon les formats de présentation des instructions. Certaines études ont montré que la combinaison du texte et de l'image permettait une action plus rapide (i.e., plus efficace, Ganier et al., 2000) et un meilleur guidage dans l'exécution des instructions, en particulier dans le cas des utilisateurs âgés (Paas, Van Gerven, & Tabbers, 2005). Mykityshyn et al. (2002), qui ont travaillé sur la manipulation d'un dispositif médical (glucomètre) sur une population jeune ($M_{\text{age}} = 19,6$, $ET = 1,9$) et âgée ($M_{\text{age}} = 68,5$, $ET = 2,8$), ont également montré que le format multimodal est meilleur (i.e., plus efficace) que le format unimodal pour les deux groupes d'âge sans formation préalable.

D'autres études ont montré que les instructions multimédias peuvent être plus efficaces lorsque les informations verbales sont présentées de manière auditive plutôt que visuelle (e.g., Morrell et al., 1990 ; Savage et al., 1991 ; Tabbers et al., 2004 ; Wright et al., 2008). Morrell, Park et Poon (1990) ont montré que les personnes âgées ($M_{\text{age}} = 71,38$, $ET = 6,19$) font plus d'erreurs (i.e., moins efficaces), mais sont plus rapides (i.e., plus efficaces) dans un format multimodal (images et texte) que dans une description unimodale qui n'inclut que l'audio dans le contexte de la surveillance des médicaments. Pour les jeunes adultes ($M_{\text{age}} = 19,16$, $SD = 1,05$), le contraire du groupe des personnes âgées est observé. Ils font moins d'erreurs (i.e., plus efficaces), mais sont plus lents (i.e., moins efficaces) dans un format multimodal (image et texte) que dans une description unimodale (audio uniquement). Savage, Kemp, et Payne (1991) ont mené une étude mesurant l'efficacité, l'efficacité et la satisfaction des "messages d'aide" (instructions de procédure) d'un poste téléphonique présentés sous forme auditive ou visuelle sur une interface. Les auteurs ont montré que les participants ($N=36$ employés, âge non spécifié) présentaient nettement moins d'erreurs lorsque les appareils possédaient des instructions au format auditif. Toutefois, aucune différence significative n'a été constatée entre les formats pour le temps d'exécution des tâches. Les auteurs expliquent leurs résultats en suggérant que les informations présentées dans un format auditif persisteraient plus longtemps dans la mémoire à court terme et seraient moins vulnérables aux interférences que les informations présentées visuellement. Ce résultat pourrait également s'expliquer par le principe de modalité présenté ci-dessus (Mayer, 2001). Lorsque les deux canaux sensoriels (auditif et visuel) sont utilisés, il y a un meilleur partage de la charge cognitive que lorsqu'un seul des canaux est utilisé. L'utilisation de procédures auditives permet alors d'éviter la saturation des canaux visuels déjà investis par la tâche à accomplir dans le cas d'une manipulation. Ainsi, les individus pourraient mieux reconstruire les relations spatiales lors de l'écoute des phrases. Cela permet d'améliorer le processus d'apprentissage de différentes manières, car, premièrement, cela augmente l'efficacité du produit en réduisant l'effort mental et le temps de résolution des problèmes (Jeung, Chandler, & Sweller, 1997 ; Mousavi, Low, & Sweller, 1995) ; deuxièmement, cela améliore l'efficacité avec de meilleurs résultats aux tests (Kalyuga, Sweller, & Chandler, 2000) et troisièmement, cela permet une meilleure mémorisation et un meilleur transfert des informations (Mayer & Moreno, 1998). Au contraire, Wright et al. (2008) et Savage et al. (1991) affirment que l'audio réduit les erreurs, mais n'a pas d'impact sur la vitesse de manipulation.

Toutefois, les performances de l'utilisateur en matière d'apprentissage procédural dépendent de plusieurs paramètres tels que le niveau de difficulté de la tâche ou l'expérience de l'utilisateur. Par exemple, Lemarié, Castillan et Eyrolle (2017) ont mis en place une expérience dans laquelle deux groupes de niveau (novices et experts en coiffure) devaient reproduire une coupe de cheveux spécifique après avoir été exposés à des instructions écrites ou à des images, ou aux deux. Ils ont constaté que le groupe d'experts (N = 13 femmes, $M_{age} = 21,87$) exposé à des instructions écrites et à des images obtenait de meilleurs résultats. En revanche, le groupe des novices (N = 13 femmes, $M_{age} = 19,07$) a obtenu de meilleurs résultats lorsqu'il a été exposé uniquement à des instructions écrites. La combinaison de textes et d'images peut être restrictive pour l'apprentissage en raison de leur séparation physique sur le document (Jamet & Arguet, 2008). Pour le lecteur, cela nécessite de faire des allers-retours entre les deux sources d'information, ce qui entraîne un phénomène de partage attentionnel coûteux, voire inutile (Chandler & Sweller, 1996 ; Sweller & Chandler, 1994). C'est ce que Mayer (2001) appelle la *contiguïté spatiale* (voir Ginns, 2006, pour une revue).

Dans l'étude menée par Van Hooijdonk et Krahmer (2008), les auteurs ont montré, sur un exercice d'apprentissage pour la prévention des microtraumatismes répétés (RSI) auprès d'un échantillon de 30 jeunes adultes, que le seul format texte était plus bénéfique que le seul format image pour la pratique des exercices en phase d'apprentissage ainsi que pour la performance (quel que soit le niveau de l'exercice effectué). En revanche, les auteurs ont montré que les images permettent un temps d'apprentissage plus court, et un temps d'exécution plus rapide, mais que la performance diminue si l'exercice est difficile à réaliser. Van Hooijdonk et Krahmer (2008) ont expliqué leurs résultats en disant que la lecture d'un texte exige que les individus transforment l'information verbale en information visuo-spatiale, alors que l'image est déjà sous une forme visuo-spatiale et nécessite donc moins de transformation. L'encodage à partir de photographies ou d'images permettrait d'avoir des représentations préalablement établies. Cette représentation permettrait d'effectuer une tâche moins exigeante et moins coûteuse que l'élaboration de représentations à partir du seul texte (Kools, van de Wiel, Ruiter, & Kok, 2006 ; Stone & Glock, 1981).

La divergence des résultats observés dans les études ci-dessus pourrait s'expliquer par diverses limitations méthodologiques. Par exemple, Lemarié et al. (2017) limitent le temps d'exécution à dix minutes, ce qui ne permet pas d'obtenir des informations sur l'efficacité. Jamet et Arguel (2008), Mykityshyn et al. (2002) et Morrell, Park et Poon (1990) ont utilisé un paradigme d'exécution retardée où les utilisateurs doivent d'abord consulter l'information puis

la restituer (ce qui ne mesure que le rappel et la compréhension). Ces paradigmes d'exécution retardée et de limitation du temps d'exécution ne sont pas applicables dans une situation *in situ* (Vermersch, 1985 ; Jannin et al. 2019). En outre, cela ne fournit pas d'informations sur l'efficacité des documents procéduraux. Enfin, il est donc conseillé d'avoir une vue d'ensemble (efficacité, efficience et mémorisation) afin de juger de la qualité des documents procéduraux. Compte tenu de ces limites et de ces résultats contradictoires, on peut donc se demander si ces résultats sont généralisables à d'autres domaines d'application tels que les dispositifs médicaux connectés pour le domicile, qui fait l'objet de la présente étude. Il est important de préciser qu'à notre connaissance, il existe peu d'études sur le sujet. Du moins, les chercheurs s'intéressent de plus en plus à l'apprentissage procédural dans le but de concevoir des environnements d'apprentissage efficaces, mais la plupart de ces recherches sont limitées au domaine de l'éducation (Jannin, Ganier, & De Vries, 2019 ; Van Hooijdonk & Krahmer, 2008 ; Fraser, Ayses, & Sweller, 2015 ; Paas & Sweller, 2012). Peu d'études se concentrent sur les dispositifs à haut risque, comme dans le domaine des soins de santé, pour lesquels l'évaluation des instructions fait souvent défaut. Une mauvaise conception des instructions peut avoir un impact négatif sur la réussite de l'utilisation des dispositifs médicaux, entraînant des erreurs de manipulation ce qui est un risque pour la sécurité des patients (Cifter, 2017 ; Kushniruk, 2002). Depuis l'expansion de la e-santé (ou *santé connectée*), la complexité de l'utilisation de ces produits devient préoccupante au vu des conséquences catastrophiques (e.g., Cifter, 2017 ; Huvila et al., 2016 ; Kortum & Peres, 2015). Ces mauvaises utilisations en font un sujet majeur auquel il faut trouver des solutions.

3. Objectifs et hypothèses

Nous avons vu dans la revue de la littérature, l'importance d'étudier les documents procéduraux, ainsi que le fait que peu d'études ont investi le cas des dispositifs médicaux à domicile. Pourtant, ces études sont essentielles en raison des conséquences désastreuses pour la santé du patient (e.g., Kortum & Peres, 2015). Nous avons également vu que la réussite de l'utilisation du dispositif dépend des documents procéduraux et de leurs formats, mais aussi des caractéristiques des dispositifs et des utilisateurs (e.g., l'expérience ou l'âge). En outre, les résultats de la littérature semblent contradictoires et les modèles théoriques existants ne semblent pas s'adapter aux documents procéduraux dans le domaine de la santé. À notre connaissance, aucun travail n'est vraiment précis dans le cas d'informations présentées sous un

format unimodal ou multimodal, surtout si les utilisateurs sont novices, ce qui est souvent le cas dans le domaine des dispositifs médicaux. De plus, les études mentionnées ci-dessus ne prennent pas en compte les mêmes métriques ou ne prennent en compte qu'une seule métrique (apprentissage, efficacité, efficience, préférences, mémorisation...), ce qui empêche la standardisation de ces études. Il est en effet nécessaire d'évaluer les documents procéduraux sur plusieurs métriques complémentaires (Ganier, 2002 ; Jaspers, 2009) afin d'avoir une approche holistique (e.g., l'efficacité, l'efficience et la mémorisation). Nous avons choisi de nous concentrer sur trois mesures : l'efficacité pour mesurer le risque d'erreur de manipulation et assurer la sécurité des patients, l'efficience pour mesurer les ressources nécessaires à la manipulation du matériel, et la mémorisation pour vérifier si les procédures peuvent être facilement mémorisées en fonction des différents formats présentés.

Ainsi, on peut se demander quel pourrait être le format le plus efficace, le plus efficient et le mieux mémorisé pour un dispositif médical complexe à utiliser ? La présente étude vise à mieux comprendre l'influence du format de présentation des documents procéduraux dans la manipulation d'un tensiomètre pour les utilisateurs novices grâce à des mesures holistiques précitées. Plus précisément, nous examinons quel format de présentation parmi les formats texte seul, image seule, audio seul, texte avec images et audio avec images est le plus efficace, le plus efficient et le mieux mémorisé lorsque les novices doivent suivre des instructions.

Dans la présente étude, nous avons choisi de présenter un tensiomètre, car il est de plus en plus démocratisé auprès d'un large public (Cifter, 2017). Cet appareil semble toujours problématique à utiliser. Sur la base de nos instructions adaptés du manuel d'utilisation original du constructeur, l'utilisateur doit suivre 12 étapes (Figure 42) avant de pouvoir prendre sa tension artérielle. De plus, les utilisateurs ne reçoivent pas de *feedback* afin d'être informés du positionnement correct ou incorrect de l'appareil. Nous savons que ce dispositif ne peut pas se substituer à la présence d'un manuel d'utilisation (Chaniaud, Loup-Escande, Métayer, & Megalakaki, 2019). Pour cette raison, nous nous concentrerons sur la première utilisation de l'appareil, qui est un test crucial pour la suite de son utilisation. Jannin, Ganier, et De Vries (2019) recommandent aux participants d'essayer plusieurs fois avant d'acquérir une nouvelle compétence (cas de l'apprentissage procédural pour réaliser des points de suture). Nous considérons que dans le cas de la manipulation de dispositifs médicaux destinés au domicile du patient, il est fondamental que l'utilisateur ne commette aucune erreur de manipulation dès la première utilisation (Larsen, 2017). Ainsi afin d'éviter ces erreurs dès la première utilisation, il est important de disposer d'instructions solides, fiables et correctement présentées. C'est

pourquoi nous ne mesurerons pas l'apprentissage, car nous nous sommes concentrées sur la première utilisation. Enfin, nous avons choisi une population jeune et donc novice pour nous assurer que les performances ne sont pas influencées par l'âge (Boucheix, Lowe, & Bugajska, 2015 ; Sonderegger et al., 2016) ou par la technophilie (Georgsson & Staggers, 2016) même si cette population en bonne santé peut être désintéressée par les dispositifs médicaux, car la motivation est une variable qui peut influencer les performances (Rico-Olarte, López, & Kepplinger, 2018 ; Tissier & Gronier, 2014).

Nous allons donc expérimenter chacun des formats ci-dessus pour tester nos hypothèses :

Dans la première hypothèse, nous suggérons que l'utilisation d'informations présentées en format unimodal (groupe texte seulement ; groupe image seulement ; groupe audio seulement) favorise un meilleur résultat d'efficacité, d'efficience et de mémorisation que les informations présentées en format multimodal (audio avec image ; texte avec image) sur un jeune utilisateur novice (hypothèse 1 ; e.g., Jamet & Arguel, 2008 ; Lemarié et al., 2017).

Pour la deuxième hypothèse, nous nous concentrons sur des conditions unimodales. Nous supposons que les participants seront plus efficaces dans leurs manipulations et auront une meilleure mémorisation si la procédure est présentée sous forme audio plutôt que sous forme de texte et d'images (hypothèse 2 ; e.g., Lemarié et al., 2017 ; Wright et al., 2008 ; Savage et al., 1991).

Les deux hypothèses précédentes portent sur les conditions multimédias ou unimodales séparément. Qu'en est-il des formats avec des entrées sensorielles différentes (auditives et visuelles) ? Selon le principe de modalité de Mayer (2003), dans la troisième hypothèse, nous supposons que l'efficacité, l'efficience et la mémorisation du dispositif seront meilleures pour les groupes ayant un format uniquement audio que pour les groupes ayant des formats visuels (comprenant du texte uniquement et des images uniquement et du texte avec des images) (hypothèse 3 ; e.g., Morrell et al., 1990 ; Savage et al., 1991 ; Wright et al., 2008).

4. Méthode

4.1. Participants

Cent vingt-quatre étudiants de premier cycle en psychologie (âge moyen = 20,70 ans, ET = 1,725 ; tranche d'âge = 18 - 32 ans ; 32 hommes ; 92 (74,2%) femmes) de l'Université de Picardie Jules Verne à Amiens (France) ont participé à l'expérience. Ils ont été répartis au hasard

en cinq groupes selon le format de présentation du manuel d'utilisation fourni : texte seulement (N = 26), image seulement (N = 24), audio seulement (N = 24), texte avec images (N = 25), audio avec images (N = 25). Deux participants du groupe "images seulement" et un participant du groupe "audio seulement" ont été exclus des analyses parce qu'ils n'ont pas accompli la tâche demandée.

Tous les participants étaient de langue maternelle française et ont signé un formulaire de consentement éclairé. Cette recherche était conforme au code d'éthique de l'American Psychological Association. Un examen et une approbation complets n'étaient pas nécessaires selon les directives de notre institution et les réglementations nationales. Les participants n'ont reçu aucune compensation financière pour leur participation. De plus, ce type de population a très peu d'expérience dans l'utilisation de dispositifs médicaux, ce qui permet de stabiliser nos résultats et d'éviter les biais liés à la population présentée ci-dessus.

4.2. Matériel et mesures

4.2.1. Dispositif Médical: le tensiomètre

Les participants ont utilisé un tensiomètre de poignet sans fil mesurant la pression artérielle (iHealth BP7) destiné à être utilisé par le grand public et labellisé de conformité européenne (CE). Cet appareil, une fois la mesure terminée, affiche la pression systolique (SYS) et diastolique (DIA) en mmHg, ainsi que le pouls à côté d'un pictogramme en forme de cœur et de l'étiquette PUL. Les mesures effectuées par chaque appareil restent affichées pendant quelques secondes. Le tensiomètre peut être utilisé avec une application tactile. Un manuel d'utilisation comportant douze étapes et deux illustrations est initialement fourni avec l'appareil.

4.2.2. Conditions expérimentales

Nous avons mis en place trois conditions unimodales : 1) texte seul (contenant le même texte initial que celui du fabricant et découpé en 12 étapes) ; 2) image seule (12 illustrations ont été dessinées par un graphiste professionnel pour illustrer les 12 étapes décrites du fabricant) ; 3) audio seule, une voix féminine sans intonation a été enregistrée lisant exactement le même texte que pour la condition texte seul ; et deux conditions multimodales : 4) texte avec images (Figure 42) et 5) texte auditif avec images. Dans tous les cas, le texte était le même, contenant 12 phrases, chacune correspondant à une action. Pour les deux conditions multimodales, les illustrations étaient les mêmes que pour la condition unimodale avec images.

4.2.3. Questionnaires

Avant le début de la tâche, les participants devaient remplir un questionnaire comprenant des *informations personnelles* et des *informations sur l'utilisation de la technologie*. Ce questionnaire comprenait : (1) quatre informations personnelles (âge, genre, niveau d'éducation, expérience avec un tensiomètre) et (2) deux éléments, adaptés d'Agarwal et Prasad (1998), concernant l'utilisation et la volonté du participant d'explorer les innovations des technologies de l'information. Par exemple, cela comprenait des questions telles que "*Parmi ces technologies, laquelle utilisez-vous et à quelle fréquence*" sur une échelle de Likert en 5 points, allant de *Jamais* à *Très souvent*.

4.3. Procédure

L'expérience a été menée dans une salle calme avec un minimum de distraction, comptabilisant un participant par session. Les participants n'ont reçu aucune formation ou démonstration préalable de la part de l'expérimentateur. L'expérimentateur n'est intervenu qu'en cas de problèmes techniques (e.g., un problème de batterie). Les participants ont été filmés pendant la session. Pour les participants, la session a été divisée en trois parties, chacune d'elles prenant environ 10 minutes :

Dans la première partie, les participants ont rempli le questionnaire *Informations personnelles / Utilisation de la technologie*.

Dans la deuxième partie, ils ont été invités à manipuler le tensiomètre. Les participants reçoivent des instructions en fonction de leur groupe (image, texte, etc.) et ils peuvent consulter les instructions autant de fois qu'ils le souhaitent. En d'autres termes, les participants peuvent manipuler librement le tensiomètre pendant la consultation des instructions. Il n'y avait pas de limite de temps. Ils pouvaient utiliser le matériel comme en condition *in situ*. Les instructions étaient présentées sur un clavier tactile. Chaque action (composée d'une phrase ou/et d'une image) était présentée sur une seule page. Le participant devait faire glisser son doigt sur la page sur le clavier tactile pour accéder à l'étape suivante. Pour le format audio, un pictogramme "audio" est affiché à l'écran. Le participant doit appuyer sur le pictogramme pour écouter l'action à effectuer. C'est le même processus pour la condition audio avec l'image. Le participant peut revenir sur l'enregistrement, le réécouter ou l'arrêter à tout moment. Douze étapes ont été demandées pour exécuter les instructions. Une douzième étape (Figure 43), consistant à écrire sur une feuille de papier les résultats affichés sur le tensiomètre, a été ajoutée à la notice initiale.

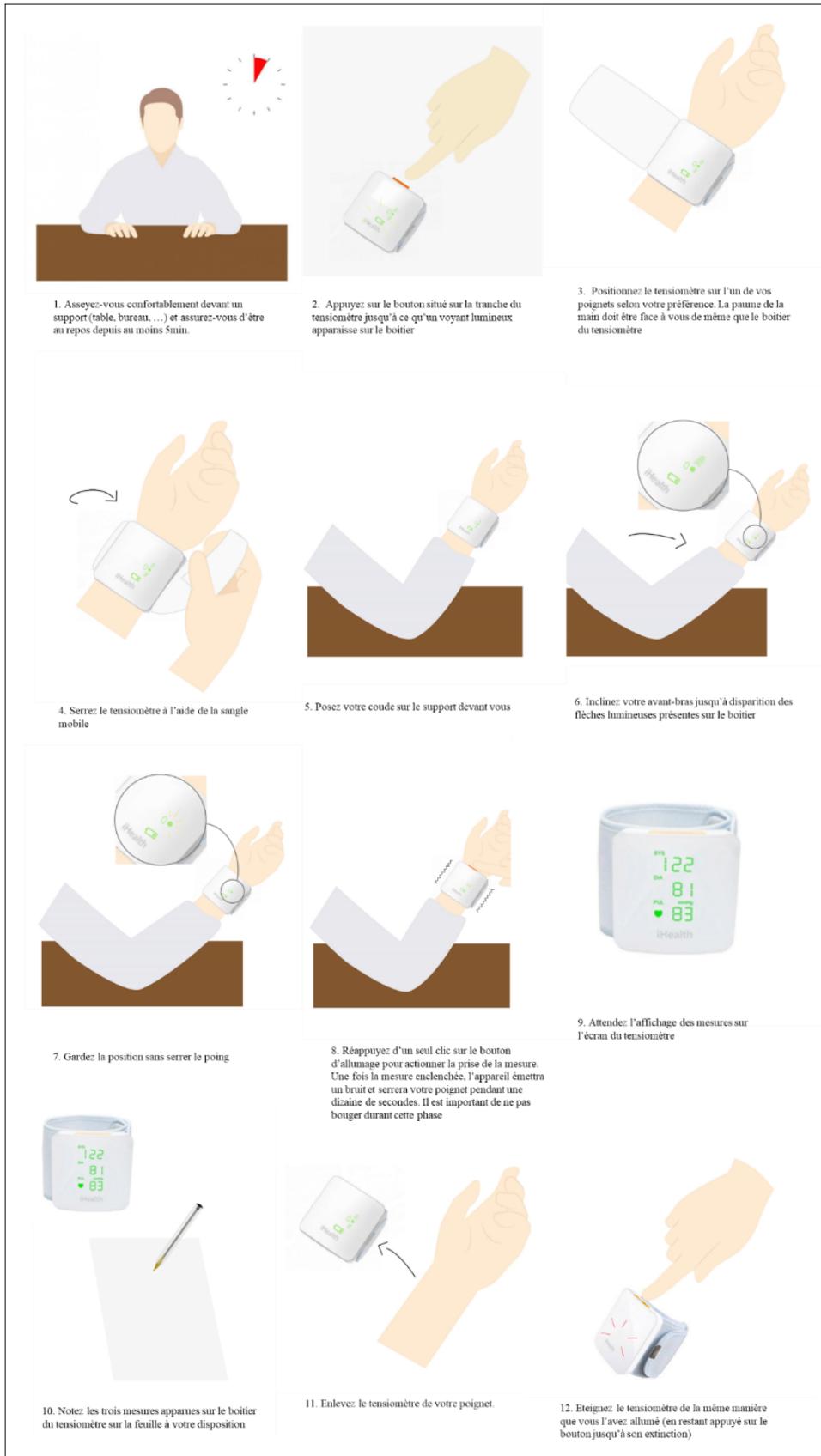


Figure 42. Exemple d'instructions de la condition texte avec images.

Les tâches de l'utilisateur consistaient à (1) se détendre pendant au moins cinq minutes, (2) allumer l'appareil, (3) positionner correctement l'appareil (le boîtier du tensiomètre étant tourné vers lui), (4) accrocher correctement l'appareil (bien serré), (5) positionner son corps par rapport au tensiomètre, (6) incliner l'avant-bras à l'aide d'un système de flèches lumineuses pour aider à trouver le bon angle du coude pour commencer la mesure, (7) commencer la mesure, (8) ne pas bouger pendant la mesure, (9) attendre que les mesures s'affichent, (10) retirer l'appareil, (11) éteindre l'appareil, (12) écrire le résultat affiché sur le tensiomètre. La Figure 43 montre une catégorisation de ces tâches selon les étapes de la procédure en repérant les erreurs de manipulation.

Lors de la troisième et dernière partie, les participants doivent rappeler les instructions données.

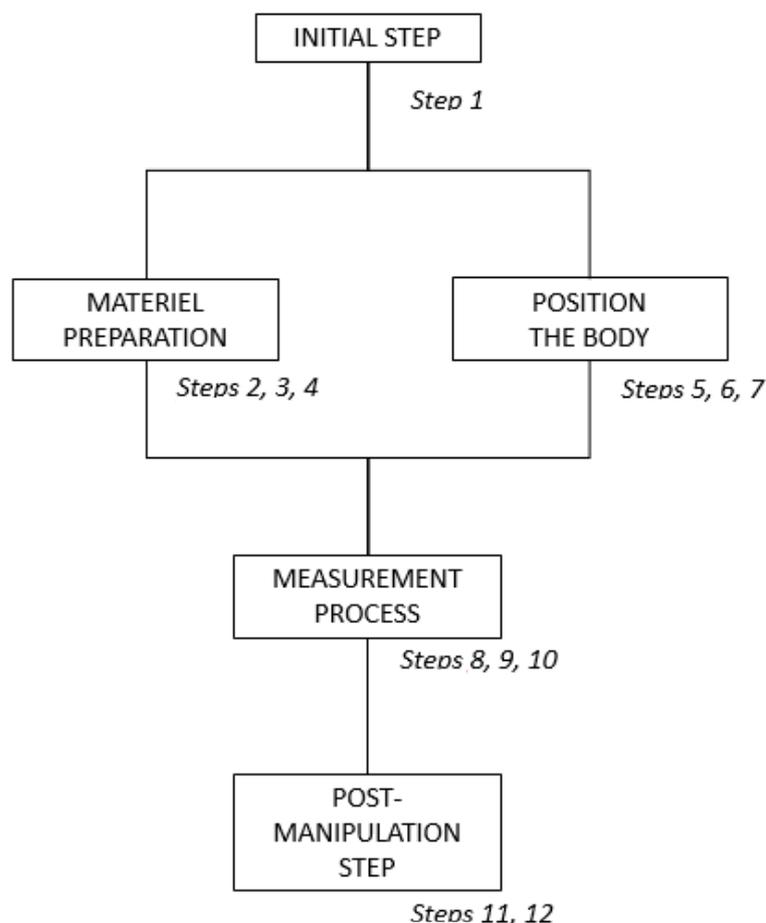


Figure 43. Représentation schématique de la procédure lors de l'utilisation du tensiomètre.

4.4. Cotation et analyse de données

Pour analyser les données, nous avons utilisé les mesures de l'efficacité, de l'efficience et de la mémorisation.

L'efficacité a été analysée en évaluant le nombre d'erreurs de fonctionnement. Quatre erreurs principales ont été détectées pour le tensiomètre : une mauvaise position sur le poignet (le participant pose le moniteur sur le bras), une mauvaise direction (le participant fait une erreur dans la direction), un mauvais serrage (le moniteur n'est pas assez serré sur le poignet), une mauvaise inclinaison (le participant n'incline pas l'avant-bras au niveau du cœur). Une grille de codage a été utilisée pour identifier les erreurs de manipulation. Les participants répétaient parfois la même erreur plusieurs fois, dans ces conditions nous avons enregistré le nombre cumulé d'erreurs.

Pour évaluer *l'efficience*, trois mesures ont été effectuées. *Le temps total de manipulation* mesuré à partir du moment où les participants ont touché l'appareil pour la première fois jusqu'au moment où ils l'ont éteint à nouveau après avoir pris la mesure. *Le temps de consultation* est mesuré à partir du moment où le participant regarde les instructions présentées sur la tablette tactile jusqu'au moment où il détourne le regard. *Le temps d'exécution* est mesuré à partir du moment où le participant touche le tensiomètre jusqu'au moment où il cesse de regarder ou de manipuler l'appareil.

Pour la *mémorisation*, il a été demandé aux participants de verbaliser en détail comment ils ont utilisé le tensiomètre. Les douze étapes du manuel étaient attendues (de "*être au repos pendant au moins 5 min*" jusqu'à "*l'appareil est éteint*"). Chaque étape rappelée par les participants a été évaluée à l'aide de cinq sous-mesures (objectifs, actions, lieux, caractérisation de l'action, conditions de réussite) tirées des transcriptions dactylographiées de leur rappel verbal. Par exemple, pour la première étape, si un participant expliquait "*J'ai d'abord allumé le tensiomètre*", les expérimentateurs lui donnaient un point pour l'action "*J'ai allumé*", et la caractérisation de l'action "*allumé*". Cependant, si le participant expliquait : "*J'ai allumé le tensiomètre en appuyant sur le bouton de ce côté et un voyant s'est allumé*", les expérimentateurs lui ont donné quatre points : un point pour l'action "*J'ai allumé*" et pour la caractérisation de l'action "*allumé*", un point pour la localisation "*sur le bouton*", un point pour la condition de succès "*un voyant s'est allumé*".

Les vidéos ont été analysées par deux experts évaluateurs à l'aide de BORIS (Behavioral Observation Research Interactive Software). Nous avons utilisé le Coefficient Corrélation Intra-

classe (CCI) pour vérifier la fiabilité inter-juges des données quantitatives (Elie & Colombet, 2011 ; Shrout & Fleiss, 1979). Un double codage a été effectué sur 33 % de la collecte de données vidéo. La mesure CCI moyenne pour les erreurs des tensiomètres était de 0,962 avec un intervalle de confiance à 95 % de 0,933 à 0,979 ($F(47,47) = 26,461, p < .001$). La mesure CCI moyenne pour le temps de manipulation (pour le moniteur de pression sanguine) était de 0,995 avec un intervalle de confiance à 95% de 0,992 à 0,997 ($F(47,47) = 261,275, p < .001$). Cela confirme la forte concordance entre les juges.

Les résultats ont été calculés à l'aide de SPSS® version 22 (IBM Corporation, 2013). Afin d'établir des comparaisons entre les caractéristiques et les performances des utilisateurs, nous avons comparé le genre, l'âge, et la technophilie (expérience dans les technologies de l'information et expérience dans le tensiomètre) par rapport à *l'efficacité*, *l'efficience* et la *mémorisation*. L'échantillon ne suit pas la distribution normale ; c'est pourquoi des tests non paramétriques ont été utilisés (test de Mann Whitney et Kruskal-Wallis). Lorsque notre échantillon suivait une distribution normale, des ANOVA ont été utilisées.

5. Résultats

Toutes les données sur *l'efficacité*, *l'efficience* et la *mémorisation* sont présentées dans le Tableau 21 ci-dessous.

5.1. Expérience antérieure dans la manipulation du tensiomètre

Aucun des participants n'a déclaré avoir déjà utilisé un tensiomètre de poignet ou même observé quelqu'un prendre sa mesure de pression artérielle au poignet. Cependant, 54,8 % des participants ont déclaré avoir déjà utilisé ou observé la manipulation d'un tensiomètre classique (au bras). En outre, pour les 68 participants ayant une expérience antérieure dans l'utilisation d'un tensiomètre, cela n'a eu aucune influence sur les résultats. Les mesures ANOVA ne révèlent aucun effet sur le temps total de manipulation : $F(1,122) = 1,94, p = ,166$; temps d'exécution : $F(1,122) = 1,944, p = ,166$; temps de consultation : $F(1,122) = 1,058, p = ,306$; nombre d'erreurs : $F(1,122) = 0,146, p = ,703$, ou tâche de rappel : $F(1,122) = 0,01, p = ,918$ entre les participants ayant une expérience antérieure et les participants sans expérience.

5.2. Efficacité

Le test de Mann-Whitney ne révèle aucune différence significative sur le nombre d'erreurs de manipulation (Mann-Whitney $U = 1667$, $p = ,64$) entre les groupes unimodaux (composés d'un groupe "texte seul", d'un groupe "image seule" et d'un groupe "audio seul") et les groupes multimodaux (composés d'un "texte avec image" et d'un "audio avec image"). Toutefois, le groupe "audio seul" présente un nombre d'erreurs nettement inférieur à celui des deux groupes multimodaux combinés (Mann-Whitney $U = 836$, $p = ,002$). Ce n'est pas le cas pour le groupe "texte seul" par rapport aux deux groupes multimodaux combinés (Mann-Whitney $U = 616,5$, $p = ,66$). C'est l'inverse pour le groupe "image seule". Les deux groupes multimodaux combinés présentent un nombre d'erreurs significativement inférieur à celui du groupe "image seule" (Mann-Whitney $U = 426,5$, $p = ,022$).

Le test de Kruskal-Wallis effectué sur tous les groupes unimodaux révèle une différence significative (nombre d'erreurs de manipulation : $\chi^2(2) = 15,34$, $p < ,001$). Le groupe "audio seul" est significativement plus efficace que le groupe "image seule" (Mann-Whitney $U = 416$, $p = ,002$) et le groupe "texte seul" (Mann-Whitney $U = 248$, $p = ,105$), mais il n'y a pas de différence significative entre le groupe "image seule" et le groupe "texte seul" (Mann-Whitney $U = 233$, $p = ,22$).

Le groupe "audio seul" fait beaucoup moins d'erreurs de manipulation que le groupe "visuel unimodal" (groupe "image seule" + groupe "texte seul") (Mann-Whitney $U = 370,50$, $p = ,008$) et le groupe "visuel multimodal" (texte avec image) (Mann-Whitney $U = 207$, $p = ,05$). Il n'y a pas de différence significative entre le groupe "audio seul" et le groupe "audio multimodal" (audio avec image) (Mann-Whitney $U = 334$, $p = ,347$).

5.3. Efficience

Le test de Mann-Whitney ne révèle aucune différence significative dans le résultat de l'efficience (temps de consultation : Mann-Whitney $U = 1482$, $p = ,169$; temps d'exécution : Mann-Whitney $U = 1445$, $p = ,116$; temps total : Mann-Whitney $U = 1659$, $p = ,667$) entre les groupes unimodaux (composés par le groupe "texte seul", le groupe "image seule" et le groupe "audio seul") et les groupes multimodaux (composés par "texte avec image" et "audio avec image"). En revanche, le groupe "audio seul" est significativement moins efficace que les deux groupes multimodaux combinés pour le temps d'exécution (Mann-Whitney $U = 894$, $p = ,021$), mais il n'y a pas de différence significative pour le temps de consultation (Mann-Whitney $U =$

578, $p = ,799$). Le groupe "image seule" passe significativement moins de temps sur les instructions (Mann-Whitney $U = 818$, $p = ,012$) que les deux groupes multimodaux combinés, ce qui n'est pas le cas pour le temps d'exécution (Mann-Whitney $U = 628$, $p = ,746$) et le temps total (Mann-Whitney $U = 728$, $p = ,139$). Le groupe "texte seul" n'est pas significativement plus efficient que les deux groupes multimodaux (temps de consultation : Mann-Whitney $U = 699$, $p = ,592$; temps d'exécution : Mann-Whitney $U = 659$, $p = ,922$; temps total : Mann-Whitney $U = 735$, $p = ,352$).

Le test de Kruskal-Wallis effectué sur tous les groupes unimodaux révèle une différence significative (temps de consultation : $\chi^2(2) = 6,217$, $p = ,045$, temps d'exécution : $\chi^2(2) = 12,116$, $p = ,002$, temps total : $\chi^2(2) = 15,340$, $p < ,001$). Le groupe "image seule" est significativement plus efficient que le groupe "audio seul" (temps de consultation : Mann-Whitney $U = 174$, $p = ,019$; temps d'exécution : Mann-Whitney $U = 150$, $p = ,004$; temps total : Mann-Whitney $U = 172$, $p = ,017$), mais il ne diffère pas du groupe "texte seul" (temps de consultation : Mann-Whitney $U = 223$, $p = ,084$; temps d'exécution : Mann-Whitney $U = 248$, $p = ,214$; temps total : Mann-Whitney $U = 298$, $p = ,786$). Le groupe "audio seul" est le groupe le moins efficient par rapport au groupe "texte seul" (temps d'exécution : Mann-Whitney $U = 510$, $p < ,001$; temps total : Mann-Whitney $U = 497$, $p < ,001$) sauf pour le temps de consultation (Mann-Whitney $U = 360$, $p = ,351$).

Le groupe "audio seul" est nettement moins efficient que le groupe visuel unimodal (groupe "image seule" et groupe "texte seul") (temps de consultation : Mann-Whitney $U = 373$, $p = ,028$; temps d'exécution : Mann-Whitney $U = 242$, $p < ,001$; temps total : Mann-Whitney $U = 205$, $p < ,001$) et le groupe visuel multimodal ("texte avec image" (temps d'exécution : Mann-Whitney $U = 520$, $p < ,001$; temps total : Mann-Whitney $U = 502$, $p < ,001$ sauf pour le temps de consultation : Mann-Whitney $U = 359$, $p = ,077$).

5.4. Mémorisation

Les groupes multimodaux mémorisent les instructions nettement mieux que les groupes unimodaux (tâche de rappel : Mann-Whitney $U = 1242$, $p = ,008$), voir les résultats généraux des comparaisons 2 par 2.

Le test de Kruskal-Wallis effectué sur tous les groupes unimodaux révèle une différence significative (tâche de rappel $\chi^2(2) = 16,676$, $p < ,001$). Le groupe "texte seul" mémorise mieux (tâche de rappel : Mann-Whitney $U = 110$, $p < ,001$) la procédure d'utilisation du tensiomètre

que le groupe "image seule". Le groupe "audio seul" mémorise mieux (tâche de rappel : Mann-Whitney $U = 121,5$, $p = ,001$) que le groupe "image seule", mais a des performances similaires à celles du groupe "texte seul" (Mann-Whitney $U = 323$, $p = ,823$). Les deux groupes multimodaux combinés mémorisent les instructions significativement mieux que le groupe "image seule" (Mann-Whitney $U = 998$, $p < ,001$) et, mais il n'y a pas de différence avec le groupe "texte seul" (Mann-Whitney $U = 716$, $p = ,469$).

Il n'y a pas de différence significative dans la mémorisation des instructions entre le groupe "audio seul" et le groupe "visuel unimodal" (groupe "image seule" + groupe "texte seul") (tâche de rappel : Mann-Whitney $U = 436,5$, $p = ,154$) et le groupe "visuel multimodal" ("texte avec image") (tâche de rappel : Mann-Whitney $U = 279,5$, $p = ,94$).

5.5. Résultats généraux : comparaison entre les deux groupes multimodaux

Le test de Mann-Whitney ne révèle aucune différence significative dans le résultat du nombre d'erreurs de manipulation (Mann-Whitney $U = 271,5$, $p = ,335$) et dans le résultat de la mémorisation (Mann-Whitney $U = 405,5$, $p = ,071$) entre les deux groupes multimodaux : groupe "audio avec image" et groupe "texte avec image". Le groupe "texte avec image" est significativement plus efficace que le groupe "audio avec image" (temps de consultation : Mann-Whitney $U = 439$, $p = ,014$; temps d'exécution : Mann-Whitney $U = 512$, $p < ,001$; temps total : Mann-Whitney $U = 530$, $p < ,001$).

Tableau 21. Moyens, écarts types et rang en fonction de l'efficacité, l'efficience et la mémorisation de la tâche pour tous les groupes.

	Efficacité	Efficience		Mémorisation		
	Erreurs de manipulation	Temps de consultation (secondes)	Temps d'exécution (secondes)	Temps total (secondes)	Tâche de rappel [sur 47]	
Unimodal	Texte seul (N=26)	<i>M</i> = 0,50	<i>M</i> = 112,93	<i>M</i> = 112,21	<i>M</i> = 248,19	<i>M</i> = 18,08
		<i>ET</i> = 0,762	<i>ET</i> = 41,97	<i>ET</i> = 64,57	<i>ET</i> = 98,65	<i>ET</i> = 4,45
		Rang = 0 – 2	Rang = 31,03 – 201,4	Rang = 33,46 – 333,9	Rang = 152,2 – 572,9	Rang = 7 – 25
	Image seule (N= 22)	<i>M</i> = 0,77	<i>M</i> = 90,55	<i>M</i> = 105,24	<i>M</i> = 213,37	<i>M</i> = 13,36
		<i>ET</i> = 0,869	<i>ET</i> = 48,25	<i>ET</i> = 66,49	<i>ET</i> = 108,11	<i>ET</i> = 3,27
		Rang = 0 – 3	Rang = 28,57 – 217,4	Rang = 28,02 – 258,2	Rang = 90,5 – 520,1	Rang = 8 – 21
	Audio seul (N = 23)	<i>M</i> = 0,13	<i>M</i> = 127,56	<i>M</i> = 161,27	<i>M</i> = 325,53	<i>M</i> = 17,83
		<i>ET</i> = 0,344	<i>ET</i> = 47,83	<i>ET</i> = 52,08	<i>ET</i> = 92,47	<i>ET</i> = 6
		Rang = 0 – 1	Rang = 64,6 – 224,9	Rang = 92,68 – 292,1	Rang = 237,47 – 550,8	Range = 7 – 32

Total des groupes unimodaux (texte, image, audio)		<i>M</i> = 0,46	<i>M</i> = 110,74	<i>M</i> = 125,94	<i>M</i> = 262,45	<i>M</i> = 16,54
(N = 71)		<i>ET</i> = 0,734	<i>ET</i> = 47,64	<i>ET</i> = 65,44	<i>ET</i> = 108,67	<i>ET</i> = 5,12
		Rang = 0 – 3	Rang = 28,5 – 224,9	Rang = 28,02 – 333,9	Rang = 90,50 – 572,9	Rang = 7 – 32
Multimodal	Texte avec image (N =25)	<i>M</i> = 0,42	<i>M</i> =102,45	<i>M</i> = 83,44	<i>M</i> = 206,91	<i>M</i> = 18,08
		<i>ET</i> = 0,584	<i>ET</i> = 29,3	<i>ET</i> = 28,99	<i>ET</i> = 48,37	<i>ET</i> = 6,2
		Rang = 0 – 2	Rang = 50,5 – 152,6	Rang = 29,56 – 153,2	Rang = 127,58 – 320,5	Rang = 9 –30
	Audio avec image (N =25)	<i>M</i> = 0,28	<i>M</i> =129,67	<i>M</i> =124,64	<i>M</i> = 279,36	<i>M</i> = 20,68
	<i>ET</i> = 0,46	<i>ET</i> = 29,09	<i>ET</i> = 39,67	<i>ET</i> = 61,51	<i>ET</i> = 4,92	
	Rang = 0 – 1	Rang = 81,09 – 208,9	Rang = 75,77 – 269,4	Rang = 190,9 – 430,07	Rang = 12 – 30	
Total Multimodal (Texte + image, Audio + image)		<i>M</i> = 0,35	<i>M</i> = 116,34	<i>M</i> = 104,45	<i>M</i> = 243,88	<i>M</i> = 19,41
(N = 49)		<i>ET</i> = 0,522	<i>ET</i> = 31,99	<i>ET</i> = 40,28	<i>ET</i> = 66,32	<i>ET</i> = 5,68
		Rang = 0 – 2	Rang = 50,50 – 208,9	Rang = 29,56 – 269,4	Rang = 127,58 – 430,07	Rang = 9 – 30

6. Discussion

Cette étude s'est concentrée sur l'impact des différents formats de présentation des instructions d'un tensiomètre sur une situation d'apprentissage procédural avec des utilisateurs novices. L'objectif principal était de déterminer quel format serait le plus approprié pour la manipulation de ce tensiomètre afin de garantir la sécurité du patient. C'est-à-dire, lequel des cinq formats présentés est le plus efficace pour limiter les erreurs de manipulation dès la première utilisation, le plus efficient et évite la surcharge cognitive et quel format permet une meilleure mémorisation des procédures.

6.1. Principales observations

Notre première hypothèse concernait le fait que les groupes unimodaux (groupe "texte seul" ; groupe "image seule" ; groupe "audio seul") seront plus performants que les groupes multimodaux ("texte avec image" et "audio avec image") en termes d'*efficacité*, d'*efficience* et de *mémorisation*. Les résultats ne valident pas notre hypothèse, car aucun résultat significatif n'a été observé sur les mesures d'*efficacité* et d'*efficience* entre le groupe unimodal et le groupe multimodal. La seule différence significative observée montre une meilleure mémorisation pour les formats multimodaux en comparaison avec les formats unimodaux.

En ce qui concerne la mesure de mémorisation, il est important de souligner que les faibles résultats de mémorisation sont en partie dus au groupe "image seule". En effet, il semble qu'il soit plus difficile de donner des détails - comme demandé dans la tâche de rappel mesurant la mémorisation – si une seule image est présentée par étape. Ces effets observés sur le groupe "image seule" (difficulté de mémorisation et grand nombre d'erreurs) sont compatibles avec les résultats de Kalyuga et al. (2000) et Lemarié et al. (2017) obtenus sur une population inexpérimentée. Ces résultats suggèrent que le format ne semble pas avoir d'impact sur l'efficience et l'efficacité. Comme les recherches précédentes, nous avons observé que le niveau d'expertise des participants peut avoir un impact plus important (Lemarié et al., 2017 ; Kalyuga et al., 2000).

Ainsi, ces comparaisons nous amènent à penser que, compte tenu des disparités entre les groupes (unimodaux et multimodaux), il est nécessaire de se concentrer sur ces différences entre l'ensemble des groupes séparément.

La **deuxième hypothèse** est que le groupe "audio seul" sera plus performant que le groupe "texte seul" et le groupe "image seule" en termes *d'efficacité*, *d'efficience* et de *mémorisation*. Les résultats valident partiellement notre hypothèse. En effet, le groupe "audio seul" est significativement **plus efficace**, car il fait beaucoup moins d'erreurs que les deux autres groupes ("image seule" et "texte seul"). Cependant, le groupe "audio seul" est le **moins efficient** (i.e., le temps d'exécution et le temps total) par rapport au groupe "image seule" et "texte seul". En revanche, le groupe "image seule" est nettement plus efficient que le groupe "audio seul" et on peut distinguer une tendance sur le groupe "texte seul". À l'exception du groupe "image seule", il n'y a pas de différence significative entre les résultats du groupe "audio seul" et du groupe "texte seul" concernant la mémorisation. Les résultats d'efficacité et d'efficience du groupe "audio seul" sont compatibles avec les résultats de Savage et al. (1991) et Wright et al. (2008).

Une explication possible des résultats de la mémorisation est que les images ne permettent pas d'accéder à une verbalisation explicite, contrairement au groupe contenant du texte (audible ou écrit) où les participants se sont souvenus de fragments de la procédure écrite. De plus, les meilleurs résultats en matière d'efficience peuvent s'expliquer par les modèles d'apprentissage multimédia (Mayer, 2001). Nous postulons qu'une construction d'un modèle mental incluant des informations visuelles-spatiales sera élaborée en mémoire de travail sur la base des informations sélectionnées. Cependant, l'image fournit déjà des informations visuo-spatiales, ce qui réduit le temps de traitement, car ces informations sont déjà traitées par le format dans lequel elles sont présentées. Cela peut expliquer pourquoi le groupe "image seule" est le plus efficient. Toutefois, cela n'explique pas pourquoi le groupe "audio seul" est le plus efficace. C'est pourquoi il est nécessaire d'étudier les différences entre le groupe audio (ou auditif) et les groupes visuels.

La **troisième hypothèse** stipulait que les participants du groupe "audio seul" auraient de meilleurs résultats en termes *d'efficacité*, *d'efficience* et de *mémorisation* que les groupes visuels unimodaux (groupe "texte seul" + groupe "image seule") et le groupe visuel multimodal ("texte avec image"). Notre hypothèse est partiellement validée à nouveau, car le groupe "audio seul" est plus efficace (i.e., réalise moins d'erreurs de manipulation), mais significativement moins efficient (i.e., est plus lent, plus spécifiquement pour la lecture des instructions et l'exécution des tâches) que les groupes visuels unimodaux et le groupe visuel multimodal (sauf pour le temps de consultation). La raison pour laquelle le groupe "audio seul" est le moins efficient pourrait s'expliquer par le fait que les groupes avec texte (audible ou écrit) nécessitent

la transformation d'informations linguistiques en informations visuelles-spatiales, ce qui demande plus de temps pour la construction du modèle mental. Par conséquent, le groupe texte seul devrait suivre les mêmes résultats, mais ce groupe n'est pas plus efficace, ni plus efficient que le groupe image seul ou les groupes multimodaux.

Nos résultats sont donc cohérents avec ceux de Savage et al. (1991) et Wright et al. (2008) ainsi qu'avec ceux de Morrell et al. (1990), mais uniquement pour le groupe des personnes âgées et non pour le groupe des jeunes.

Les résultats obtenus sur les trois hypothèses montrent un effet particulier sur le groupe "audio seul". En effet, le groupe "audio seul" est probablement le plus efficace par rapport aux autres groupes, même s'il présente des inconvénients. Nous proposons d'examiner ce format plus en détail.

6.2. Vers le choix du format audio

Nous observons dans les résultats que la mise en œuvre d'un format audio dans les instructions de procédure réduit le nombre d'erreurs de fonctionnement par rapport à d'autres formats de présentation d'instructions unimodales ou multimodales, ce qui est crucial dans le domaine des dispositifs médicaux. Plusieurs raisons peuvent expliquer ces résultats.

Tout d'abord, **le format audio a obligé les participants à écouter et à concentrer leur attention sur les instructions**, contrairement aux autres groupes. Par rapport aux autres groupes, les instructions audio n'avaient pas d'autre indication sur l'écran de la tablette qu'un pictogramme "écoute". Les participants n'ont pas d'autres points de repère que d'écouter l'action avant de commencer à manipuler. En outre, ils sont encouragés à écouter les informations jusqu'au bout. Inversement, les groupes contenant du texte et/ou des images peuvent à tout moment arrêter la lecture et revenir en arrière si une erreur s'est produite. Des études antérieures montrent également que les documents procéduraux ne sont pas toujours lus par les participants dans la vie réelle (Silver & Wogalter, 1991 ; Wright, Creighton & Threlfall, 1982), surtout s'ils nécessitent un effort cognitif important (Fu & Gray, 2006). Ces auteurs expliquent que les utilisateurs ne lisent les documents qu'en dernier recours, après avoir essayé sans instructions. On peut penser que le même phénomène se produit avec les autres groupes unimodaux et multimodaux (image, texte et texte/audio avec images).

Une raison supplémentaire à la première, qui explique en partie le succès de ce groupe audio seul, est le fait que **ce groupe a été obligé de faire des pauses dans sa manipulation**. Lorsque les participants sont confrontés à une image et/ou un texte, ils peuvent concentrer leur attention sur un point de l'image ou une partie du texte. Dans le cas du groupe audio, le participant est obligé d'écouter tous les détails des instructions. Une fois la phrase écoutée, dans la plupart des cas, les participants ont effectué leurs actions et ont ensuite tourné la page sur la tablette pour activer l'écoute d'une autre phrase. Ce résultat corrobore les résultats de Fleury et Jamet (2015) qui ont montré que les interruptions peuvent améliorer l'attention de l'utilisateur et donc accroître l'efficacité du logiciel en réduisant la durée des tâches, notamment dans la phase de vérification post-reconnaissance. Les interruptions offrent la possibilité de corriger les erreurs pendant la reconnaissance plutôt qu'après, ce qui est plus facile, car cela nécessite moins de ressources cognitives.

Enfin, la troisième raison est que **le groupe audio uniquement a permis de couper artificiellement l'information, en plus de permettre des pauses et de forcer les participants à écouter avec attention**. Nous pouvons affirmer que le groupe "audio seul" a forcé l'utilisateur à suivre cette stratégie de traitement de l'information en atomisant l'action (Vermersch, 1985). Le format audio a l'avantage de permettre au participant de traiter les détails l'un après l'autre. Une fois que l'utilisateur a traité un élément de la procédure, il revient au manuel pour obtenir le détail suivant de la procédure qui fragmente l'information et la conduit vers une stratégie d'atomisation de l'action. Pour les autres formats visuels, les participants ont tendance à choisir une stratégie d'essai-erreur qui est cognitivement rentable, mais peu fiable en termes de taux d'erreur. Ce type de traitement par atomisation améliore les performances en termes de rappel (Blankenship & Dansereau, 2000), d'apprentissage (Jannin et al., 2019 ; Wright, Hull, & Black, 1990), de compréhension (Bétrancourt, Dillenbourg, & Montarnal, 2003) et de transfert d'informations (Jamet & Arguel, 2008).

Cependant, nos résultats montrent également que le groupe "audio seul" est moins efficace lors de la première utilisation du tensiomètre. En effet, nous avons observé que **le temps passé à écouter les instructions et à les exécuter est plus important que pour les autres groupes unimodaux**. Prendre plus de temps pour consulter/exécuter les actions est probablement une bonne pratique, car elle permet de réduire les erreurs de manipulation, ce qui suggère que les participants ont un engagement cognitif plus important (Sinatra, Heddy, & Lombardi, 2015). En effet, l'apprentissage procédural requiert un effort cognitif important, mais nécessaire, c'est pourquoi les utilisateurs ne lisent les instructions qu'en dernier recours. Il est

bien connu que les apprenants optent toujours pour le moindre effort cognitif (Eiriksdottir & Catrambon, 2011 ; Ganier, 2004 ; Szlichcinski, 1979). C'est pourquoi la stratégie « essayer-erreur » est souvent la méthode privilégiée par les utilisateurs (Rettig, 1991). Mais il n'est pas souhaitable que cette méthode soit appliquée aux dispositifs médicaux. Nous nous trouvons donc devant un dilemme : soit proposer une instruction exigeant un effort important de la part de l'individu et/ou forcer le participant à faire cet effort, mais en évitant les erreurs de manipulation, soit proposer des instructions exigeant un traitement moindre de l'information, mais augmentant le nombre d'erreurs de manipulation. La deuxième option peut conduire les utilisateurs à avoir recours à la stratégie d'essai-erreur, ce qui peut être dangereux dans le domaine des dispositifs en santé. Les utilisateurs n'ont pas nécessairement de retour d'information pour identifier leurs erreurs et risquent un apprentissage procédural incorrect et la répétition de leurs erreurs (Larsen, 2017).

Ainsi, les résultats obtenus valident le modèle de Mayer. L'utilisation combinée d'informations verbales et picturales accélère l'acquisition de connaissances. Cependant, nous ne pouvons pas remettre en cause le modèle de Mayer, nous pouvons débattre de l'intérêt d'optimiser trop rapidement l'acquisition de connaissances dans l'application de dispositifs spécifiques tels que ceux utilisés dans le domaine des soins de santé.

7. Conclusions et perspectives de recherche

Notre étude montre la complexité de l'utilisation de cet appareil et l'augmentation du nombre d'erreurs si l'utilisateur n'utilise pas l'appareil correctement (surtout s'il ne lit pas les instructions dans sa totalité). Nos résultats mettent en évidence les avantages d'une instruction audio obligeant les futurs utilisateurs à entendre les instructions en évitant la stratégie d'essais et d'erreurs, particulièrement dangereuse, et en facilitant des stratégies plus appropriées, telles que l'atomisation de l'action.

Pour conclure, le format audio semble avoir un impact intéressant sur la stratégie de traitement des instructions par rapport aux autres formats. Ce mode de traitement peut être recommandé dans la conception des procédures des dispositifs en santé afin de limiter les erreurs de manipulation et de garantir la santé et la sécurité du patient.

Cependant, afin de valider la significativité de nos résultats, cette étude devrait être répétée dans des conditions différentes. Tout d'abord, cette étude devrait évaluer les performances de l'utilisateur sur l'ensemble du dispositif (y compris Bluetooth) en tenant compte de l'*efficacité*, de l'*efficience*, de la *mémorisation* mais aussi de l'*apprentissage* en évaluant les performances longitudinalement afin de tester les résultats de l'étude de Larsen (2017) suggérant que les participants qui commettent une première erreur ont de fortes chances de répéter cette erreur dans la tâche suivante. Deuxièmement, il aurait été intéressant de proposer un format dynamique (e.g., un format audiovisuel) bien que groupe "audio avec image" soit le plus proche de ce format. Berney et Bétrancourt (2016) démontrent dans une méta-analyse l'impact positif des formats dynamiques sur l'apprentissage global par rapport aux formats statiques. Enfin, il est important, tout comme la mesure de l'utilisabilité (ISO 9241-11:2018), de prendre en compte les caractéristiques de l'utilisateur (âge, motivation, niveau de littératie en santé, expérience, etc.), les ressources, la tâche et la complexité de l'appareil à manipuler. Par exemple, cette étude doit être étendue à une plus grande partie de la population, y compris à une population d'âge et de pathologie variables. D'autres facteurs tels que l'engagement d'un individu envers la santé pourraient influencer la performance d'un individu dans l'utilisation d'un appareil (Barello et al., 2016) par rapport à l'utilisation d'un appareil domestique tel qu'une cocotte-minute (Ganier, 2002). Chacun de ses éléments influence les mesures de performance et les documents de procédure doivent être adaptés à chaque cas spécifique.

8. Synthèse du Chapitre 3 – Étude 1b

L'étude 1b constitue le deuxième volet de l'étude 1 et correspond au **niveau 2** (tâches – utilisateurs – ressources) du modèle HANDLED. Cette étude avait pour **objectif de mieux comprendre l'influence des ressources réutilisables matérialisées sous des formats de présentation de documents procéduraux** lors de la manipulation d'un tensiomètre pour les utilisateurs novices.

Pour cela, cinq formats (unimodals : texte, audio ou image et multimodal : texte/audio avec images) ont été testés sur 124 participants novices répartis de manière aléatoire dans 5 groupes. Nous avons examiné quel format de présentation était le plus efficace, le plus efficient et le mieux mémorisé lorsque les novices doivent suivre des instructions.

Nos résultats suggèrent **que le format audio est le plus efficace** par rapport aux quatre formats et pourtant, c'est également **le moins efficient des formats**. Les résultats des autres conditions n'ont montré aucun impact réel sur les performances de l'utilisateur.

Afin de privilégier la sécurité du patient, nous considérons qu'**il est préférable d'opter pour un format permettant d'optimiser l'efficacité au détriment de l'efficience dans le cas des dispositifs connectés en santé pour le domicile patient**. En effet, ces résultats ont montré que le format audio présentait plusieurs avantages permettant de forcer les participants à utiliser une stratégie adaptée à la situation, celle de l'atomisation de l'action. Cette stratégie incite l'utilisateur à fragmenter l'information et permet d'éviter la saturation de la mémoire de travail.

Cette absence de différences dans les performances est néanmoins importante et montre la nécessité d'adapter chaque instruction aux contextes d'apprentissage et d'usage, c'est-à-dire en fonction des utilisateurs, de l'environnement, des ressources et de la complexité de la tâche proposée.

CHAPITRE 4 : ÉTUDE 2. EFFETS DES CARACTÉRISTIQUES SOCIO-DEMOGRAPHIQUES ET COGNITIVES SUR L'UTILISABILITÉ DU DISPOSITIF *SMART ANGEL*

1. Préambule

Ce présent chapitre est tiré d'un article scientifique³³ actuellement soumis dans *JMIR Human factors*.

1.1. Contexte et objectif industriels

Un premier prototype de l'application *Smart Angel* connectée aux deux objets présélectionnés a vu le jour sur la base des maquettes et du choix du matériel (Figure 44). Le CHU de Nîmes a préféré proposer un deuxième tensiomètre de poignet (de la même marque), car le précédent a été jugé trop complexe à utiliser. La particularité de ce tensiomètre par rapport à l'ancien est l'absence de repère pour positionner l'avant-bras (*via* des flèches lumineuses présentes sur le boîtier) et un affichage des données qui ne se fait plus sur le boîtier, mais uniquement sur l'application. L'objectif industriel de cette étude est d'évaluer l'utilisabilité du système dans sa globalité en se basant sur les métriques principales de l'utilisabilité (l'*efficacité*, l'*efficience* et la *satisfaction*) auprès d'une population cible du dispositif. Cette étude a permis de faire remonter des recommandations de conception pour ce premier prototype telles que les difficultés de connexion avec le Bluetooth. Nous avons aussi pu observer que ce nouveau tensiomètre diminue le nombre d'abandons, mais augmente le nombre d'erreurs entraînant des mesures erronées de la tension.

³³ Chaniaud, N., Megalakaki, O., Capo, S., Loup-Escande, E. (under review). Effects of user characteristics on the usability of home connected medical device: The case of *Smart Angel*.

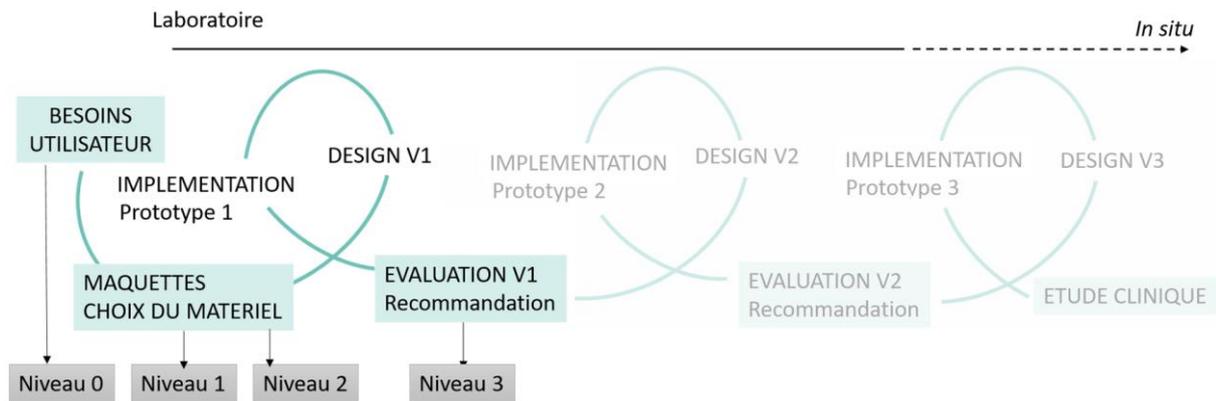


Figure 44. Étude 2 correspondant au niveau 3 du processus de conception du dispositif *Smart Angel*.

1.2. Contexte et objectif théoriques

Cette étude réalisée correspond au niveau 3 du processus de conception du dispositif *Smart Angel* (Figure 44). Ce chapitre peut être divisé en deux parties : une partie quantitative qui a fait l'objet d'un article scientifique et une partie qualitative. De ce fait, ce chapitre possède deux objectifs théoriques.

Le premier objectif suit les précédents en identifiant les éléments contextuels pouvant impacter l'utilisabilité (e.g., Borsci et al., 2018; Grebin, Echeveste, Magnago, Tanure, & Pulgati, 2018) en intégrant, pour la première fois, le dispositif *Smart Angel* dans sa globalité. Nous nous sommes donc intéressées aux caractéristiques intrinsèques du dispositif sur l'utilisabilité. Pour cela, nous avons choisi d'étudier l'impact de quatre caractéristiques utilisateurs sur l'ensemble du système *Smart Angel* : l'âge, le niveau d'éducation, la technophilie et la littératie en santé. L'âge et le niveau d'éducation ont pu être analysés grâce à la diversité de la population interrogée. La littératie en santé (Sørensen, et al., 2012) a été mesurée à l'aide de deux questionnaires standardisés (Newest Vital Sign, Weiss et al. 2005, et Health literacy Scale, Sørensen, et al. 2015) récemment validés et traduits en français (e.g., Mansfield et al., 2018; Rouquette et al., 2018). Afin de nous rapprocher davantage d'un contexte *in situ*, nous avons réalisé cette étude au domicile des participants ou sur leur lieu de travail. De plus, nous leur avons affecté un persona représenté à travers un scénario diffusé avant la manipulation que nous avons conçu lors d'un workshop grâce à l'étude exploratoire (cf, Partie II, Chapitre 1, p. 112).

Le deuxième objectif était de réaliser une analyse qualitative pour laquelle nous nous sommes intéressées à deux éléments : (1) la facilité d'utilisation perçue et (2) l'utilité perçue

(Yen, 2010). Premièrement, lors de l'étude 1b (cf. Partie I, Chapitre 3, p. 160), nous avons discuté des possibilités de survenue des erreurs de manipulation. Parmi ces arguments, nous avons émis la possibilité que les utilisateurs ne lisaient pas (ou peu) les instructions (e.g., Hoc, 1987; Silver & Wogalter, 1991; Wogalter et al., 2016; Wright et al., 1982). Afin de vérifier cette hypothèse, nous leur avons demandé de « penser à voix haute » (« *think aloud* ») pendant la manipulation, puis nous avons réalisé un entretien semi-directif. Nous présentons un détail des explications des utilisateurs à la survenue de leurs erreurs. Deuxièmement, nous nous sommes intéressées à l'utilité perçue en interrogeant les participants sur le ratio bénéfiques/risques perçue d'un tel dispositif.

La Figure 45 modélise les différents éléments évalués dans cette étude basée sur la matrice : *tâches – utilisateurs – ressources - système*.

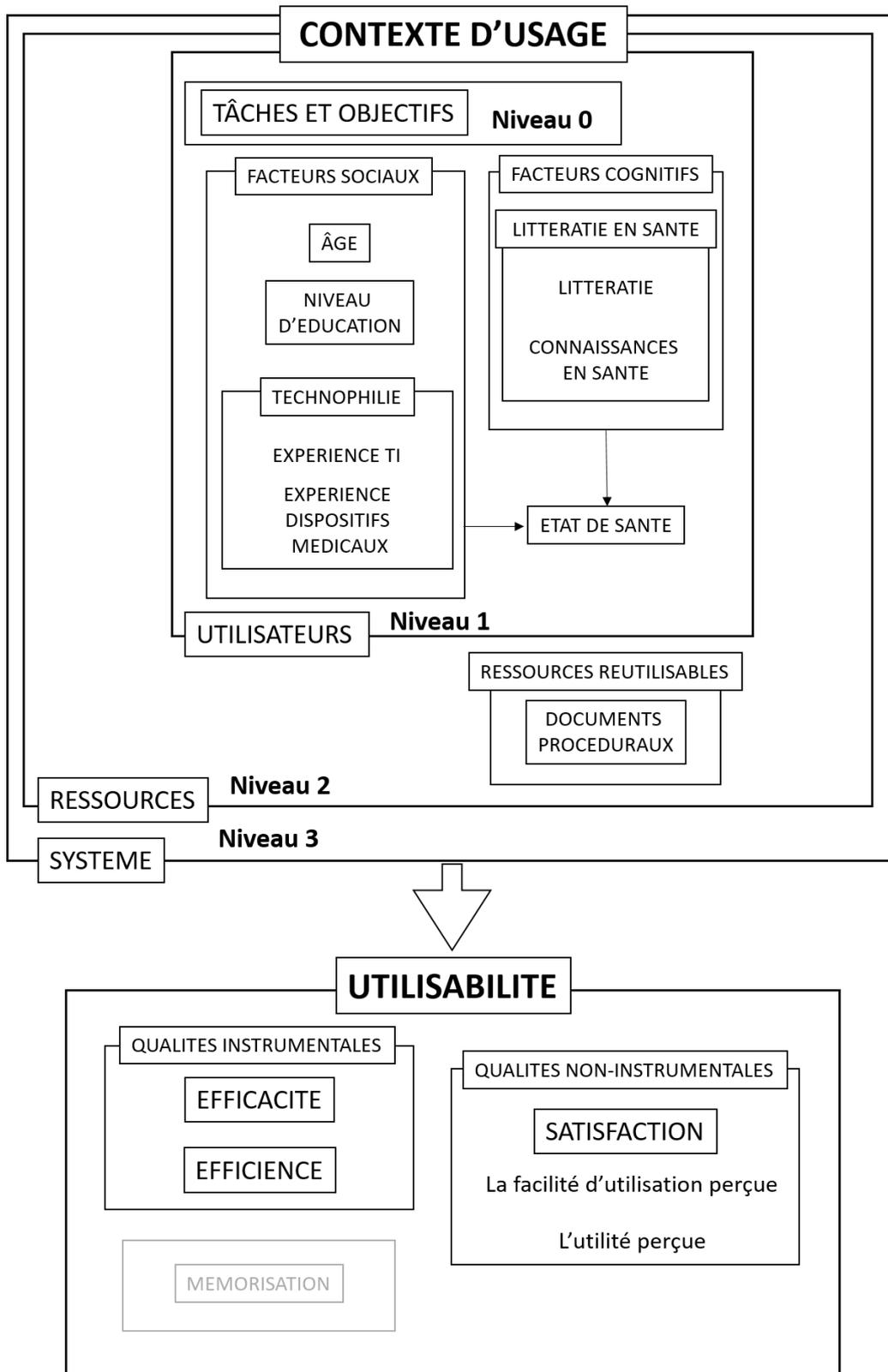


Figure 45. Attributs éprouvés lors de l'étude 2 du contexte d'usage sur l'utilisabilité (efficacité, efficacité et satisfaction).

2. Introduction

La chirurgie ambulatoire est en pleine croissance ces dernières années. Les opérations pratiquées sont de plus en plus complexes et dangereuses pour le patient devant gérer sa convalescence à son domicile. Le dispositif *Smart Angel* est un futur dispositif médical connecté à domicile, spécialement conçu pour prévenir des complications post-chirurgicales relatives à la chirurgie ambulatoire. L'objectif de ce dispositif est de faciliter le retour au domicile du patient en maintenant une liaison avec l'hôpital. Pour cela, le patient doit manipuler ce dispositif après son opération dès son retour à domicile en envoyant trois fois par jour l'ensemble de ses constantes, pendant une semaine, avant de restituer le matériel au Centre Hospitalier. Ce suivi postopératoire permettrait aussi aux patients de mieux gérer leur convalescence en évitant les retours souvent fréquents dans les services d'urgence ou en consultations extrahospitalières (Ghosh & Sallam, 1994). Actuellement, ce dispositif est à un stade précoce de conception. Comme tout dispositif médical, cet outil doit suivre des normes de sécurité et de qualité (ISO 14971), ainsi que d'utilisabilité (ISO 9241-11) afin de répondre aux exigences de la Conformité Européenne (marquage CE) en vue d'être commercialisé. Pourtant, encore aujourd'hui, le déploiement de ces dispositifs connectés en santé est freiné entre autres par leur complexité d'utilisation renvoyant directement à un manque d'utilisabilité (Carayon & Hoonakker, 2019; Chaniaud et al., 2019; Cifter, 2017; Greysen et al., 2014; Huvila et al., 2016; H. Kim & Xie, 2017), impactant ainsi la sécurité du patient.

2.1. Utilisabilité

L'utilisabilité peut être définie par la norme ISO 9241-11 (2018, p.2) comme « *Le degré selon lequel un produit peut être utilisé, par des utilisateurs identifiés, pour atteindre des buts définis avec efficacité, efficience et satisfaction, dans un contexte d'utilisation spécifié* ». Cette notion toujours discutée par la communauté scientifique comporte trois dimensions distinctes ISO 9241-11 (2018) :

- *Efficacité* : l'exactitude et l'exhaustivité avec lesquelles les utilisateurs atteignent certains objectifs ;
- *Efficience* : la relation entre l'exactitude et les ressources dépensées pour les atteindre ;
- *Satisfaction* : le confort des utilisateurs et l'évaluation positive de leur interaction.

L'utilisabilité définie par ces trois dimensions (*efficacité, efficacité* et *satisfaction*) est liée à son contexte d'usage (ISO 9241-11, 2018), lui-même étant caractérisé par quatre composants (la *tâche*, l'*environnement*, les *ressources*, les *utilisateurs*).

2.2. Caractéristiques utilisateurs : âge, niveau d'éducation, technophilie, littératie en santé

Plusieurs chercheurs ont récemment étudié les liens entre les caractéristiques des utilisateurs et l'utilisabilité de dispositifs connectés dans le domaine de la santé (e.g., Georgsson & Staggers, 2016; Jensen et al., 2010; Liang et al., 2018; Monkman & Kushniruk, 2013; Sonderegger et al., 2016; Van der Vaart et al., 2019). Quatre caractéristiques utilisateurs sont particulièrement étudiées dans la littérature scientifique : l'âge (e.g., Georgsson & Staggers, 2016; Jones & Caird, 2017; Kaufman et al., 2003; Mykityshyn et al., 2002; Sonderegger et al., 2016; Sparkes et al., 2012; Van der Vaart et al., 2019) ; le niveau d'éducation (e.g., Georgsson & Staggers, 2016; Liang et al., 2018; Van der Vaart et al., 2019) ; la technophilie, c'est-à-dire l'expérience dans la technologie de l'information (TI) et l'expérience antérieure dans les dispositifs médicaux (e.g., Georgsson & Staggers, 2016; Harte et al., 2018; Sparkes et al., 2012) ; et la littératie en santé (Czaja et al., 2015; Kaufman et al., 2003; H. Kim & Xie, 2017; Van der Vaart et al., 2019).

Dans la majorité des études, les auteurs s'accordent plus ou moins sur ces liens et cela sur des dispositifs différents. Nous détaillons ces études ci-dessous.

2.2.1. L'âge

De nombreux auteurs se sont intéressés à l'influence de l'âge sur l'utilisabilité de dispositifs connectés dans le domaine de la santé. La plupart de ces auteurs s'accordent sur l'impact de l'âge sur l'utilisabilité. Par exemple, Georgsson et Staggers (2016) investiguent l'utilisabilité d'une application de gestion du diabète fonctionnant sur smartphone en utilisant les métriques de la norme ISO 9241-11 (*efficacité, efficacité* et *satisfaction*). Les auteurs mettent en évidence que le groupe plus jeune (30-49ans) fait moins d'erreurs (i.e., plus efficace), est plus rapide (i.e., plus efficient) et est plus satisfait (score au questionnaire SUS de 88.33 contre 77.14) que le groupe plus âgé (50-69ans). Sparkes et al. (2012) s'intéressent à l'utilisabilité des tests cardiaques à distance et montrent que l'âge des participants influence leurs capacités dans l'installation du matériel. Les sujets jeunes semblent plus à l'aise que les sujets plus âgés. Jones et Caird (2017) se sont intéressés à l'utilisation d'un glucomètre. Ils ont

montré que les sujets plus jeunes avaient moins de difficultés et réalisaient moins d'erreurs (i.e., efficaces) que les sujets plus âgés. Mykityshyn et al. (2002) se sont également intéressés à l'utilisation d'un glucomètre. Ils ont constaté que les sujets jeunes étaient plus rapides (i.e., efficaces) que les sujets plus âgés quel que soit le format d'instructions fourni (écrit + dessiné VS vidéo). Van der Vaart et al. (2019) évaluent l'utilisabilité d'une application pour la surveillance des symptômes de 32 narcoleptiques. Les auteurs montrent que l'utilisabilité (mesurée en termes de quantité de tâches complétées et de problèmes rencontrés) est corrélée de manière modérée et positive avec l'âge ainsi qu'avec le niveau de littératie en e-santé.

Toutefois, Liang et al. (2018) n'ont montré aucun lien entre l'âge et la satisfaction mesurée via le score du questionnaire Système Usability Scale (SUS) dans leur étude sur l'évaluation de sept dispositifs de santé utilisés par le grand public (e.g., montres connectées) conduite auprès d'un échantillon de 388 participants. De même, Jensen et al. (2010) ne démontrent aucun lien entre l'utilisabilité et l'âge des participants en ce qui concerne l'accès et l'utilisation d'information de santé en ligne. Les auteurs expliquent que cette absence de résultat est probablement due au contraste des niveaux de littératie en santé qui auraient pris le dessus sur les autres variables.

2.2.2. Le niveau d'éducation

Le niveau d'éducation est également une variable que l'on retrouve dans de nombreuses évaluations de l'utilisabilité. Pourtant, à notre connaissance, aucune étude n'a prouvé ce lien. En effet, Georgsson et Staggers (2016), Liang et al. (2018) et Van der Vaart et al. (2019) ont tous révélé une absence de lien entre le niveau d'éducation des participants et l'utilisabilité (*efficacité, efficience et satisfaction*).

2.2.3. Technophilie (expérience TI/dispositifs médicaux)

Concernant l'influence de l'expérience dans les technologies ainsi que l'expérience antérieure dans les dispositifs médicaux sur l'utilisabilité, les auteurs ont obtenu des résultats divergents. En effet, Georgsson et Staggers (2016) montrent que les plus expérimentés dans la technologie (ce que les auteurs appellent « IT/computer experience ») réalisent moins d'erreurs (i.e., efficaces), sont plus rapides (i.e., efficaces) et sont plus satisfaits de l'application de gestion du diabète (+ 5 points pour le score du Système Usability Scale). À l'inverse, Harte et al. (2018) ont réalisé des analyses de régression entre l'expérience des technologies et le score au System Usability Scale (SUS) et n'ont trouvé aucun effet significatif lors de l'évaluation d'une application de santé sur Smartphone. Enfin, Sparkes et al. (2012) ont montré que la

familiarité avec les technologies semblait avoir une influence sur la bonne installation de leur dispositif.

2.2.4. Littératie en santé

a. Définition de la littératie en santé

La littératie en santé est une caractéristique des utilisateurs qui aurait un effet sur l'utilisabilité des dispositifs médicaux (e.g., Czaja et al., 2015; Jensen et al., 2010; Van der Vaart et al., 2019). En raison de son approche multidimensionnelle, cette caractéristique est toutefois complexe à définir et est source de difficultés dans son évaluation. Sørensen, Van den Broucke, Fullam, Doyle, Pelikan, Slonska, et Brand (2012, p.3) la définissent comme : « *Les connaissances, les compétences, la motivation et la capacité d'un individu à repérer, comprendre, évaluer et utiliser des informations sur la santé lors de la prise de décisions dans les contextes des soins de santé, de la prévention des maladies et de la promotion de la santé pour maintenir ou améliorer la qualité de la vie au cours de la vie* ». Cette notion est pourtant souvent mentionnée comme un déterminant à considérer dans l'éducation thérapeutique (Margat et al., 2017; Williams et al., 1998), la prévention (Möttus et al., 2014), l'observance thérapeutique, l'accès à l'information en santé (Jensen et al., 2010) voire la guérison (Berkman et al., 2011; Williams et al., 1998). Pourtant, aucune étude à notre connaissance n'a évalué le niveau de littératie en santé des Français, sur le plan national.

b. Evaluation de la littératie en santé

En termes d'évaluation, la littératie en santé est particulièrement difficile à mesurer pour au moins deux raisons. La première raison concerne sa spécificité multidimensionnelle (Sørensen et al., 2012). La seconde raison est le fait que la littératie en santé n'est pas liée aux critères socio-économiques comme il serait intuitif de le penser (Zenklusen et al., 2012).

La littératie en santé se mesure sur un continuum variant de très faible à très élevé d'après le Conseil Supérieur de l'Éducation du Québec (CSE, 2013). Actuellement, il existe deux grandes méthodes de mesure de la littératie en santé (Jordan et al., 2011) : 1) des méthodes par questionnaire(s) au moyen desquelles on évalue les capacités d'un individu et 2) des méthodes auto-rapportées par lesquelles on observe directement les comportements d'un individu face à un professionnel de santé. Actuellement, peu d'outils existent en français comparativement aux 51 instruments anglophones recensés par Haun, Valerio, McCormack, Sørensen et Paasche-Orlow (2014). Les instruments les plus souvent utilisés et cités font partie des méthodes par questionnaire : il s'agit du TOFHLA (Parker et al., 1995), du REALM (*The*

Rapid Estimate of Adult Literacy in Medicine de Murphy et al., 1993), du HLS-EU (*European Health Literacy Survey* de Sørensen et al., 2015) et du NVS (*Newest Vital Sign* de Weiss et al., 2005). Cependant, ces instruments comportent plusieurs limites. Parmi ces instruments, le REALM correspond davantage à un test de lecture qu'à un test de compréhension puisqu'il est demandé aux participants de lire des termes médicaux. La version courte du TOFHLA (i.e., S-TOFHLA : le Short Test of Functional Health Literacy in Adults), qui permet d'évaluer le niveau de compréhension des répondants, semble plus adaptée à la culture suisse qu'à la culture française (Connor et al., 2013). En effet, il est fait référence au système de la « *caisse maladie* » suisse et à l'envoi de certains documents ne s'appliquant pas au modèle de la sécurité sociale française. De plus, la validité du S-TOFHLA fait actuellement objet d'une certaine controverse en raison des incohérences dans l'interprétation des items qui le composent (Houston et al., 2018). Un autre instrument proposé dans la littérature, le NVS (Weiss et al., 2005) présente une forte corrélation (Cronbach $\alpha > .76$) avec la mesure du S-TOFHLA (Baker et al., 1999). Il permet également d'évaluer une partie des compétences cognitives des répondants (Lecture Écriture Compréhension Numératie). Enfin, le HLS-EU repose sur le modèle multidimensionnel de la littératie de Sørensen et al. (2012). Cet outil a mis en évidence d'importantes lacunes dans huit pays européens dans la mesure où environ 1 personne sur 2 aurait un niveau de littératie en santé limité, voire faible (Sørensen et al., 2015).

c. Utilisabilité et littératie en santé

Dans le contexte des technologies de la santé telles que les dispositifs médicaux connectés prenant de plus en plus de place dans la vie du patient, les études portant sur les liens entre la littératie en santé et l'utilisabilité restent encore rares et/ou exploratoires. Monkman et Kushniruk (2013) proposent une évaluation de l'utilisabilité en prenant en compte la littératie en santé à travers la conception et la validation de critères heuristiques. Pour cela, les auteurs ont adapté un ensemble de lignes directrices existantes pour la conception de sites Web spécifiques à la santé afin de rendre le contenu plus compréhensible par les utilisateurs ayant un faible niveau de littératie en santé. Ces mêmes auteurs proposent un modèle (Figure 46) appelé « *The consumer Health Information System Adoption Model* » (Monkman & Kushniruk, 2015, p. 28). Ils considèrent que la compréhension d'un système se traduit par deux facteurs : une combinaison entre l'utilisabilité, l'utilité et la littératie en e-santé. Finalement, à notre connaissance, aucune étude expérimentale n'a permis de caractériser empiriquement des liens entre la littératie en santé et l'utilisabilité au sens d'efficacité, efficacité et satisfaction.

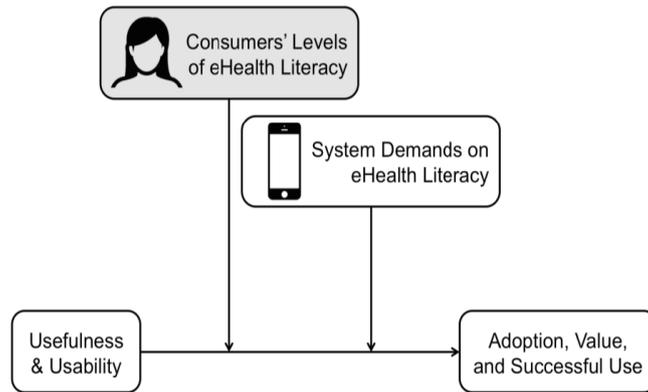


Figure 46. Modèle de littératie en e-santé en fonction de l'utilisabilité d'un système : *"The consumer Health Information System Adoption Model"* (Monkman & Kushniruk, 2015, p. 28)

Czaja et al. (2015) ont pu montrer, en utilisant un système de dossiers de santé personnels électroniques, que les populations ayant un faible niveau de littératie éprouvaient davantage de difficultés à se servir de ces outils. Kim et Xie (2017) ont réalisé une revue systématique des articles s'intéressant à l'impact d'un faible niveau en littératie en santé sur l'utilisation de dispositif de e-santé. Sur la base de 74 études, les auteurs concluent que l'obstacle majeur à l'accès et à l'utilisation de l'information en ligne sur la santé pour les individus à faible niveau de littératie est fortement lié à l'utilisabilité du site web. Jensen et al. (2010) ont observé que les participants ayant de faibles niveaux en littératie en santé (mesurées à l'aide du REALM) utilisaient moins les technologies de la santé. Celles ayant de faibles niveaux en numératie en santé (mesurées à l'aide du TOFHLA) auraient, quant à elles, un accès restreint à ces technologies. Ce dernier résultat va dans le sens de ceux de Kaufman et al. (2003) qui concluent également que le faible niveau de numératie peut être un obstacle à l'utilisation d'un système de télémédecine. Finalement, à notre connaissance, aucune étude expérimentale n'a permis de caractériser empiriquement des liens entre la littératie en santé et l'utilisabilité au sens d'efficience, efficacité et satisfaction.

3. Objectifs et hypothèses

Nous avons vu que la complexité d'utilisation des dispositifs connectés en santé résidait essentiellement dans des problèmes d'utilisabilité (e.g., Kim & Xie, 2017), et ce d'autant plus qu'ils doivent être utilisables par des patients ayant des profils diversifiés (Weick-Brady &

Lazerow, 2006 ; Ten Haken, Ben Allouch, et Van Harten, 2018). Ainsi, **l'objectif principal de cette étude est de comprendre en quoi la complexité intrinsèque de l'ensemble du système *Smart Angel* impacte l'utilisabilité.** Pour cela il est essentiel de prendre en compte les éléments qui composent ce système incluant chaque élément du système (application sur tablette et objets connectés) et des profils patients divers. En ce sens, la prise en compte des caractéristiques utilisateurs incluant l'âge, le niveau d'éducation, la technophilie (expérience TI et dans les dispositifs médicaux) et la littératie en santé sont des facteurs importants à considérer dans la conception d'un dispositif médical connecté à destination du domicile patient, tel que le dispositif *Smart Angel*. Or, à notre connaissance, aucune étude impliquant l'ensemble de ces caractéristiques précitées n'a été réalisée. De plus, les liens entre ces caractéristiques et l'utilisabilité restent encore peu exploités dans la littérature. Ainsi, **ce travail se décompose en deux sous-objectifs reliés à deux types d'analyses : quantitative et qualitative.** Le premier objectif, à travers une analyse quantitative de ce travail, est de mieux comprendre les relations qu'entretiennent les quatre caractéristiques utilisateurs (i.e., âge, niveau d'éducation, technophilie et littératie en santé) et l'utilisabilité (ISO 9241-11 : 2018 : *efficacité, efficacité et satisfaction*) d'un dispositif médical à destination du domicile patient.

Pour ce faire, nous avons formulé quatre hypothèses :

- (H1) les utilisateurs âgés seront moins efficaces, efficaces et satisfaits du dispositif médical connecté *Smart Angel* que les jeunes utilisateurs (e.g., Georgsson & Staggers, 2016 ; Jones & Caird, 2017 ; Mykityshyn et al., 2002 ; Van der Vaart et al., 2019) ;
- (H2) Les utilisateurs ayant un faible niveau de technophilie (expérience IT et dispositifs médicaux) seront moins efficaces, efficaces et satisfaits du dispositif médical connecté *Smart Angel* que ceux ayant un niveau élevé de technophilie (e.g., Georgsson & Staggers, 2016 ; Harte et al., 2018 ; Sparkes et al., 2012) ;
- (H3) Le niveau d'éducation n'aura pas d'effet sur l'efficacité, l'efficacité et la satisfaction du dispositif médical connecté *Smart Angel* (e.g., Georgsson & Staggers, 2016 ; Liang et al., 2018 ; Van der Vaart et al., 2019) ;
- (H4) Les utilisateurs ayant un niveau faible de littératie en santé (mesurée par les scores de NVS et de HLS-EU) seront moins efficaces, efficaces et satisfaits du dispositif médical connecté *Smart Angel* que ceux ayant un niveau élevé de littératie en santé (e.g., Jensen et al., 2010 ; Kaufman et al., 2003).

Le deuxième objectif, à travers une analyse qualitative, est de comprendre comment les utilisateurs perçoivent la *facilité d'utilisation* et l'*utilité perçue* du dispositif, deux caractéristiques de la satisfaction d'après Yen (2010).

- (H5) La facilité d'utilisation perçue : Les erreurs de manipulation proviennent principalement d'une mauvaise lecture des documents procéduraux (Silver & Wogalter, 1991) ;
- (H6) L'utilité perçue : La sécurité et la fiabilité des données (e.g., Allaert, 2016 ; Béranger, 2014 ; Laubie, 2011), la responsabilisation (Eysenbach, 2001), l'apport pour les zones isolées (Simon, 2017) pour les proches (Laubie, 2011) et la complexité d'utilisation (e.g., Carayon & Hoonakker, 2019 ; Cifter, 2017 ; Greysen, Garcia, Sudore, Cenzer, & Covinsky, 2014) sont les principaux freins et avantages du dispositif *Smart Angel*.

4. Méthode

4.1. Participants

Trente-six participants, soit 17 femmes et 19 hommes (respectivement, 47,22 % et 52,78 %), âgés de 20 à 64 ans (moy = 40,75 ans, ET = 14,45) ont participé à cette étude. Les critères d'inclusion de cette étude étaient les suivants : le participant devait avoir une connexion 4G à son domicile, être âgé de moins de 70 ans, être éligible à la chirurgie ambulatoire, ne pas être seul à son domicile. L'ensemble des participants était de langue maternelle française et a signé un formulaire de consentement après avoir été informé du déroulement de l'étude. L'étude est conforme aux recommandations éthiques de la déclaration d'Helsinki. Les participants ont été recrutés sur la base du volontariat et aucun dédommagement ne leur a été attribué. Les passations se sont déroulées au domicile ou sur le lieu de travail du participant.

4.2. Matériel et mesures

Le matériel de notre expérience est composé : 1) du dispositif *Smart Angel*, 2) de personas et de leurs scénarios, 3) de questionnaires (deux questionnaires évaluant le niveau de littératie en santé, à savoir, le NVS et le HLS-EU, un questionnaire relatif aux données sociodémographiques, et un questionnaire d'évaluation de la satisfaction, à savoir le SUS) et 4) d'entretiens semi-directifs.

4.2.1. Dispositif *Smart Angel*

Le dispositif *Smart Angel* (Figure 47) est composé d'une tablette Samsung 9 pouces comprenant l'application *Smart Angel* ainsi que de deux objets connectés disponibles pour le grand public ayant une certification européenne : un tensiomètre de poignet (iHealth BP7) permettant d'obtenir une mesure de tension et un oxymètre (iHealth Oximeter PO3) mesurant la saturation en oxygène et de pouls. Afin d'utiliser le dispositif *Smart Angel* conçu par Evolucare Technologies, il est nécessaire d'entrer dans l'application *Smart Angel* et de réaliser un « rendez-vous » médical numérique depuis une application sur tablette.

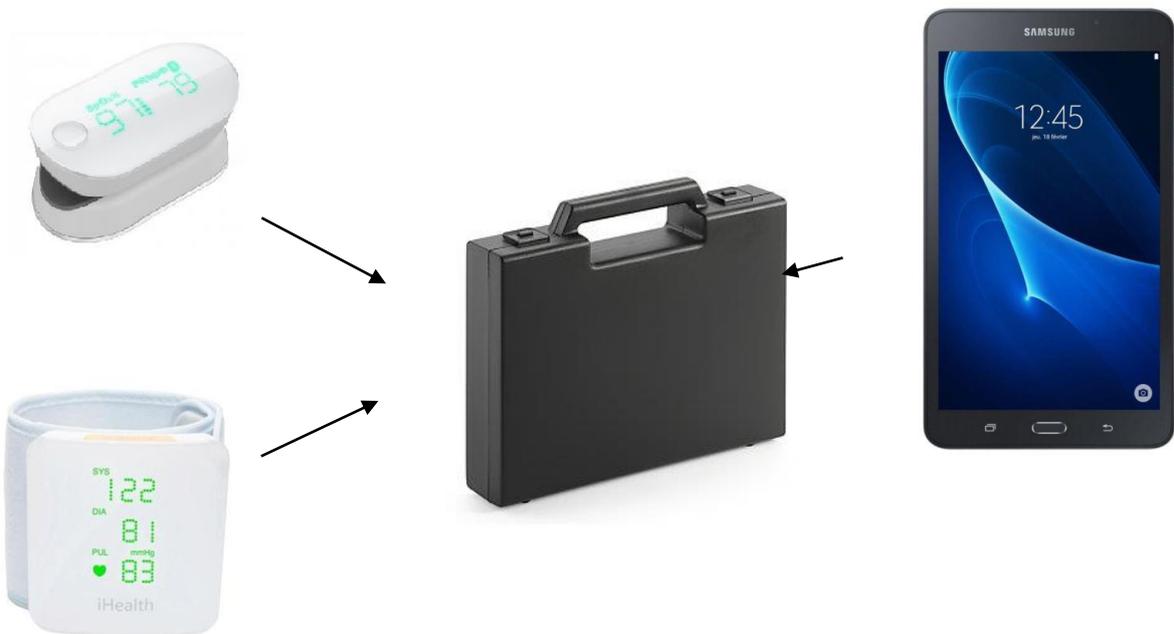


Figure 47. L'ensemble des composants du dispositif *Smart Angel* : un oxymètre de pouls (iHealth Oximeter PO3) en haut à gauche, un tensiomètre de poignet (iHealth BP7) en bas à gauche et une tablette comprenant l'application *Smart Angel* à droite.

L'application *Smart Angel* est constituée de (Figure 47) :

- Une procédure d'utilisation des objets connectés (tensiomètre et oxymètre de pouls) grâce à laquelle le patient découvre les informations les unes à la suite des autres et peut enclencher la connexion puis la prise de mesure au moyen de ces objets. Une fois la prise de tension ou la prise d'oxygénation réalisée, ses données de santé s'affichent sur une jauge colorée (de vert à orange) en fonction du niveau de gravité de la constante collectée ;

- Un formulaire dans lequel l'utilisateur répond à un questionnaire comprenant divers items relatifs à l'état de santé générale, la douleur, le sommeil et la nausée. Ces items sont présentés soit au format QCS³⁴ (e.g., « *comment vous-sentez-vous aujourd'hui ? bien, pas bien, pas bien du tout* ») soit sur une échelle de Likert (e.g., « *Évaluez votre douleur sur une échelle de 1 à 10* »).

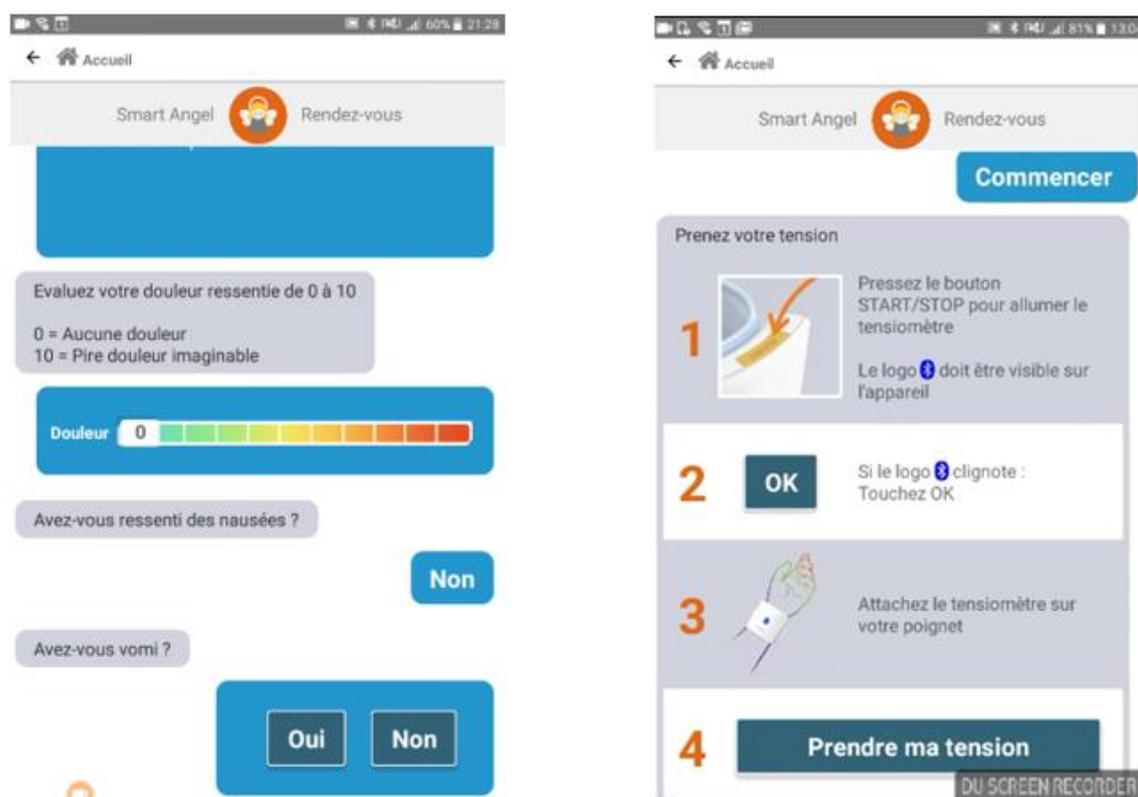


Figure 48. Capture d'écran de l'application *Smart Angel* fabriquée par Evolucare Technologies. À gauche le formulaire permettant d'avoir un aperçu de l'état de santé subjectif, et à droite la procédure d'utilisation du tensiomètre.

Les procédures d'utilisation du tensiomètre et de l'oxymètre de pouls sont intégrées à l'application (Figure 48). Elles comprennent du texte et des images pour chaque étape de la manipulation. Pour les deux objets connectés, le participant doit d'abord avoir une bonne position du corps, connecter le matériel puis l'installer correctement sur lui-même, lancer la mesure, enlever et éteindre le matériel. Une représentation schématique de la procédure d'utilisation du matériel est présentée sur la Figure 49 ci-dessous.

³⁴ QCS = questions à choix simple

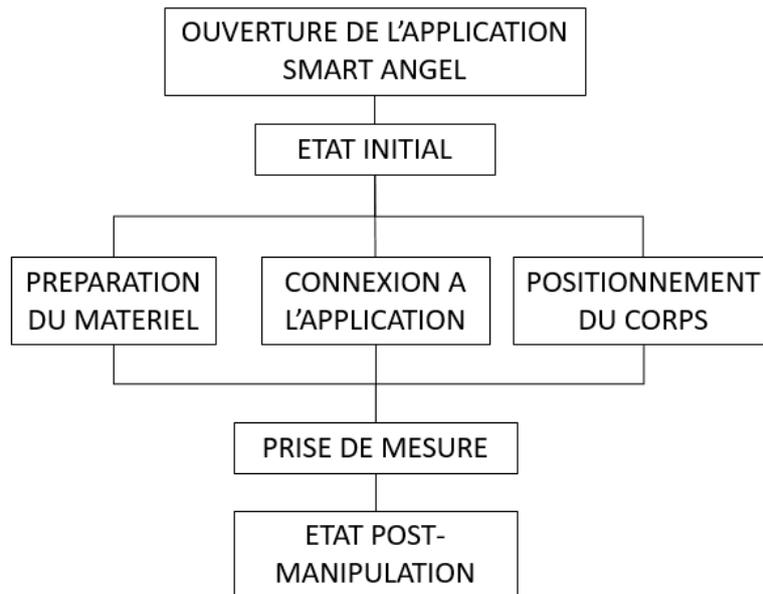


Figure 49. Représentation schématique des grandes étapes de l'utilisation du dispositif *Smart Angel*.

4.2.2. Les personas et scénarios

Cinq personas ainsi que leurs scénarios (cf. Annexe 8 : Présentation des personas et leurs scénarios, p. 371) ont été construits sur la base de relevés statistiques des types d'actes de chirurgie ambulatoire en France (ATIH, 2016a) et d'observations réalisées sur le terrain (Loup-Escande et al., 2019). Généralement utilisée en phase de conception, la méthode des personas sollicite la théorie de l'esprit, la théorie des stéréotypes et peut provoquer certains états émotionnels (Lallemand & Gronier, 2015). Les mises en situation des personas (les scénarios) étaient présentées aux participants sous forme d'un dessin animé (audiovisuel). Tous les scénarios ont été construits de la même manière. Seuls le type d'opération et la cause de l'opération changeaient en fonction du persona.

4.2.3. Questionnaires

a. Mesures du niveau de littératie en santé

Étant donné les limites relatives aux questionnaires de littératie en santé traduits et validés en français, nous avons choisi d'utiliser deux questionnaires de littératie en santé afin d'avoir une vision holistique de cette compétence multidimensionnelle : le Newest Vital Sign (NVS) et le Health Literacy Suvey (HLS-EU-Q16) :

- **Newest Vital Sign (NVS)** de Weiss et al. (2005) traduit en français par Mansfield et al. (2018) est une épreuve validée évaluant la capacité des patients à comprendre la lecture

ainsi qu'à manipuler des chiffres (numératie). Les participants doivent se servir d'une étiquette nutritionnelle de crème glacée afin de répondre à six questions (e.g., « *Si je suis allergique aux arachides, puis-je manger cette crème glacée ? réponse : non, car la glace contient des traces d'huile d'arachides* »). Le NVS (cf. Annexe 12 : Questionnaire de littératie en santé – le Newest Vital Sign, p. 399) permet d'identifier les patients à faible niveau de littératie en santé en les classant par niveau : entre 0 et 1 point, niveau « faible », entre 2 et 3 points, niveau « limité », entre 4 et 6 points, niveau « correct » de littératie en santé. La fiabilité inter-item du NVS dans cette étude est bonne (Nunnally, 1978) : le α de Cronbach est de 0,883 ;

- **Health Literacy Survey – Europe (HLS-EU-Q16)** de Sørensen et al. (2013) traduit en français par Rouquette et al. (2018), est la version courte du questionnaire Health Literacy Survey développé par Sørensen et al. (2012). Cette version est composée de 16 items dont 13 évaluent les quatre types de compétences de la littératie en santé : la capacité d'accéder, de comprendre, d'évaluer et d'appliquer l'information sur la santé. Le répondant doit évaluer sa propre capacité à accéder à l'information (e.g., « *Indiquez, sur une échelle de très facile à très difficile, dans quelle mesure il est facile pour vous de comprendre les consignes de votre médecin ou de votre pharmacien sur la manière de prendre vos médicaments ?* »). Quatre catégories de réponses sont proposées sur une échelle de Likert à quatre points allant (de « *très facile* » à « *très difficile* »). Pour calculer le score total, les réponses « *facile* » et « *très facile* » ont un point par item, les réponses « *difficile* » et « *très difficile* » ne gagnent pas de point (cf. Annexe 13 : Questionnaire de littératie en santé – Le HLS-EU-Q16, p. 400). La somme totale des items (de 0 à 16 points) permet de classer les répondants en trois catégories : (0 à 8 points) littératie en santé faible (9 à 12 points) limitée (13 à 16 points) correcte. La fiabilité inter-item du HSL-EU dans cette étude est bonne (Nunnally, 1978) : le α de Cronbach est de 0,803.

b. Mesures sociodémographiques (âge, niveau d'éducation, technophilie...)

Ce questionnaire comprenait des détails personnels : âge, sexe, niveau d'éducation, zone d'habitat, technophilie (expérience TI et dispositifs médicaux), expérience de l'hôpital (cf. Annexe 14 : Questionnaire caractéristiques sociodémographiques études 2 et 3, p. 401).

4.2.4. Entretiens semi-directifs

Les entretiens semi-directifs avaient pour objectif d'étudier la perception et les potentielles intentions d'usage des utilisateurs par rapport au dispositif dans sa globalité et surtout de comprendre d'où pouvaient provenir les erreurs de manipulation. Pour cela, lors de l'entretien

avec le participant nous suivions un guide d'entretien (cf. Annexe 15 : Guide d'entretien pour l'étude 2, p. 404) composé de quatre thématiques : la perception du dispositif ; l'acceptation du dispositif ; les erreurs ; des recommandations de conception.

4.3. Traitement des données quantitatives

4.3.1. Mesure de l'efficacité

La mesure de l'*efficacité* a été évaluée en comptabilisant le nombre d'erreurs de manipulation, comme le fait de ne pas mettre le brassard du tensiomètre dans la bonne position. Cinq catégories d'erreurs ont été identifiées en ce qui concerne l'utilisation du tensiomètre ((i) le participant ne positionne pas le tensiomètre à l'envers ; (ii) dans le sens de la paume de la main (iii) ne positionne pas l'avant-bras correctement (iv) bouge pendant la prise de mesure (v) ne connecte pas le Bluetooth du tensiomètre à la tablette) ; et quatre catégories relatives à l'usage de l'oxymètre de pouls ont été repérées ((i) le participant ne positionne pas l'oxymètre dans le bon sens (ii) n'insère pas son doigt jusqu'au capteur (iii) retire l'oxymètre trop tôt pendant la prise de mesure (iv) ne connecte pas le Bluetooth de l'oxymètre à la tablette) ; un type d'erreurs concernant la tablette a été observé (le participant ne rentre pas dans le « rendez-vous » de l'application). Une grille de notation a été utilisée pour identifier ces erreurs de manipulation. Lorsque le participant réalisait plusieurs tentatives, nous avons enregistré le nombre cumulatif d'erreurs.

4.3.2. Mesure de l'efficience

La mesure de l'*efficience* est basée sur les temps de manipulation des différents outils du dispositif concernant trois mesures : la mesure de manipulation du tensiomètre, la mesure de manipulation de l'oxymètre de pouls et la mesure de manipulation totale du dispositif incluant l'ensemble du « rendez-vous ». Ces temps ont été mesurés à partir du moment où les participants ont touché l'appareil (tensiomètre, oxymètre de pouls ou tablette) pour la première fois jusqu'au moment où ils l'ont éteint après avoir pris la mesure.

4.3.3. Mesure de la satisfaction

a. Satisfaction quantitative : System Usability Survey

La satisfaction a été mesurée à l'aide du System Usability Survey (SUS). Ce questionnaire "*quick and dirty*" (Brooke, 1996) est composé de 10 items avec cinq options de réponse sur une échelle de Likert allant de « *pas du tout d'accord* » à « *tout à fait d'accord* » et

permet l'évaluation subjective de l'utilisabilité (Lewis & Sauro, 2009). Nous avons utilisé la version réadaptée et validée décrite par Bangor et al. (2008) dans lequel nous avons remplacé le terme "système" par le terme "*dispositif médical*"³⁵. Les scores ont été calculés selon les recommandations de Brooke (1996) et varie de 0 à 100. Les scores les plus faibles indiquent une utilisabilité faible.

b. Satisfaction qualitative : Analyse par thématique du discours think aloud et des entretiens semi-directifs

Le recours à la méthode du *think aloud* et l'entretien semi-directif ont permis de mieux comprendre la satisfaction du dispositif *Smart Angel*. Pour cela, nous avons retranscrit l'intégralité des discours du think aloud et de l'entretien semi-directif. Nous avons combiné ces deux types de données qualitatives et les avons analysées par thématique (Braun & Clarke, 2006) en nous inspirant du modèle Health-ITUEM (Yen, 2010). Braun et Clarke (2006, p.6) définissent l'analyse thématique comme "*a method for identifying, analysing, and reporting patterns (themes) within data*".

4.4. Procédure

La passation de cette expérience durait en moyenne 50 minutes. Le participant était d'abord invité à choisir parmi 5 personas proposés afin de lui permettre de se projeter dans les besoins des utilisateurs futurs du dispositif *Smart Angel* (Bornet & Brangier, 2013). Le persona choisi devait être en concordance avec au moins son âge, sa profession et l'opération antérieurement subie du participant. Ensuite, l'expérimentateur réalisait une démonstration devant le participant d'environ 3min à propos de l'utilisation du dispositif *Smart Angel* en énumérant les informations relatives à la bonne manipulation (e.g., « *le tensiomètre doit toujours se trouver au niveau du cœur* ») sur lui-même. Le participant devait, par la suite, compléter trois questionnaires : le questionnaire relatif aux données sociodémographiques, le *Health Literacy Survey – Europe* (HLS-EU-Q16) et le *Newest Vital Sign* (NVS). Puis, le participant devait manipuler le dispositif *Smart Angel* en réalisant d'abord une mesure de tension, suivie d'une mesure de la saturation en oxygène et enfin en répondant au questionnaire de santé générale. Il était invité à penser à voix haute (*think aloud*) pendant toute la manipulation. Il n'avait aucune limite de temps. Il était filmé pendant la durée de la manipulation. L'expérimentateur ne pouvait intervenir qu'en cas de problème technique (e.g.,

³⁵ Ce type de modification n'a aucune incidence sur la validité ou la fiabilité de l'instrument d'enquête (Sauro, 2011).

problème de batterie). Enfin, après celle-ci, le participant devait répondre au SUS et était invité à discuter avec l'expérimentateur pour un entretien semi-directif.

4.5. Analyse de données

Les vidéos ont été analysées à l'aide de BORIS (Behavioral Observation Research Interactive Software, <http://www.boris.unito.it/>) et ont permis d'obtenir des données sur l'*efficacité* et l'*efficience*.

Les données quantitatives ont été analysées à l'aide de SPSS® version 22 (IBM Corporation, 2013). Chaque caractéristique utilisateur a été comparée systématiquement aux composants de l'utilisabilité comprenant l'*efficacité*, l'*efficience* et la *satisfaction*. Pour la mesure de la littératie en santé, nous avons d'abord analysé le résultat au HLS-EU-Q16 puis le résultat au NVS. Des corrélations bivariées, des ANOVAs et tests t de Student ont été réalisés lorsque l'échantillon répondait aux critères d'homoscédasticité, tandis que des tests non paramétriques (Kruskal-Wallis et Mann-Whitney) ont été opérés lorsque l'échantillon ne répondait pas à ces critères.

Les données qualitatives ont été analysées à l'aide de NVivo 12 (QSR International Pty Ltd, 2019) en suivant les 6 étapes de l'analyse thématique (Braun & Clarke, 2006) : familiarisation avec les données, codification, thématiques (arborescence), révision des thèmes, définition des thèmes, contextualisation avec la littérature scientifique. Pour cela, nous nous sommes inspirées du modèle ITUEM-health (Yen, 2010) positionnant la satisfaction selon deux thématiques principales : la facilité d'utilisation perçue et l'utilité perçue. C'est sur cette base que nous avons créé notre propre arborescence.

4.6. Fidélité inter-juges : mesures objectives d'efficacité et d'efficience

Nous avons utilisé l'Intra-Class Corrélation (ICC) afin de vérifier la fiabilité inter-juge pour les données quantitatives (Elie & Colombet, 2011 ; Shrout & Fleiss, 1979). Un double codage à 33 % des données vidéos collectées a été réalisé. La mesure moyenne ICC pour le temps de manipulation total (*efficience*) est de 0,978 avec 95 % d'intervalle de confiance à 0,918 à 0,994 ($F(11,11) = 45,436, p < ,001$). La mesure moyenne ICC pour le temps de manipulation (*efficience*) du tensiomètre est de 0,988 avec 95 % d'intervalle de confiance compris entre 0,954 et 0,997 ($F(11,11) = 81,635, p < ,001$). La mesure moyenne ICC pour le temps de manipulation

(*efficience*) de l'oxymètre de pouls est de 0,956 avec 95 % d'intervalle de confiance 0,838 à 0,988 ($F(11,11) = 22,955, p < ,001$). La mesure moyenne ICC pour le temps de manipulation (*efficience*) de la tablette est de 0,906 avec 95 % d'intervalle de confiance compris entre 0,652 et 0,975 ($F(11,11) = 10,688, p < ,001$). La mesure moyenne du nombre d'erreurs de manipulation (*efficacité*) est de 0,952 avec un intervalle de confiance à 95 % compris entre 0,842 et 0,985 ($F(11,11) = 20,789, p < ,001$).

5. Résultats

5.1. Analyse quantitative

5.1.1. Effets des caractéristiques utilisateurs sur l'utilisabilité

Les relations entre les caractéristiques utilisateurs (voir Tableau 22) et les composants de l'utilisabilité (i.e., le nombre d'erreurs de manipulation : *efficacité* ; le temps de manipulation : *efficience* et le score au SUS : *satisfaction*) ont été analysées de façon systématique.

a. L'âge

L'âge ($M = 40,75, ET = 14,45, \text{rang} = 20 - 64$ ans) est significativement corrélé (positivement et faiblement) au nombre d'erreurs de manipulation (*efficacité* : $r = 0,359 ; p = ,032$) et au temps de manipulation (*efficience* : $r = 0,357 ; p = ,033$). Il n'y a en revanche pas de corrélation significative entre l'âge et le score au SUS (*satisfaction* : $r = -0,138 ; p = ,424$). De plus, il est important de souligner que l'âge n'est pas corrélé avec le niveau de littératie du HLS-EU-Q16 ($r = 0,013 ; p = ,942$) ni du NVS ($r = -0,013 ; p = ,942$) (Figure 50).

b. Technophilie (Expérience TI et dispositifs médicaux)

L'expérience TI n'a montré aucun impact sur le nombre d'erreurs de manipulation (*efficacité* : $F(5,30) = 1,229 ; p = ,32$) ; le temps de manipulation (*efficience* : $F(5,30) = 1,39 ; p = ,256$). En revanche, il y a un effet significatif entre l'expérience TI et le score au SUS (*satisfaction* : $\chi(3) = 8,671 ; p = ,034$).

L'expérience antérieure relative à l'utilisation de dispositifs médicaux permettant la prise de sa propre prise de tension n'a pas eu d'influence sur le nombre d'erreurs de manipulation (*efficacité* : $t = 0,443 ; \text{ddl} = 34 ; p = ,661$) ; sur le temps de manipulation (*efficience* : $t = 0,39 ; \text{ddl} = 34 ; p = ,555$), ni sur le score au SUS (*satisfaction* : Mann-Whitney $U = 104 ; p = ,188$). L'expérience antérieure relative à l'utilisation de dispositifs médicaux

Tableau 22. Analyses descriptives des caractéristiques des utilisateurs, leurs expériences en santé, dans les dispositifs médicaux et dans la technologie.

<i>Variable (N=36)</i>	<i>n(%)</i>	<i>Efficacité Moy (ET)</i>	<i>Effizienz Moy (ET)</i>	<i>Satisfaction Moy (ET)</i>	
Caractéristiques					
<i>Âge</i>	40,75 ans (14,45)	36 (100)			
<i>Genre</i>	Homme	19 (52,78)	1,21 (1,27)	362,09 (144,16)	87,24 (11,18)
	Femme	17 (47,22)	2,06 (1,25)	373,91 (126,3)	81,03 (11,73)
<i>Niveau d'éducation</i>	Secondaire (seconde, 1 ^{re} Terminale), niveau Baccalauréat ou Brevet professionnel	5 (13,9)	2,8 (1,64)	337,96 (89,67)	77 (9,75)
	Supérieur 1er cycle (Bac +1 ou 2)	11 (30,6)	1,36 (1,1)	334,99 (106,25)	87,05 (13,82)
	Supérieur 2e cycle (Bac +3 ou 4)	11 (30,6)	1,64 (1,2)	412,95 (198,6)	82,73 (13,34)
	Supérieur 3e cycle (Bac +5 et plus)	9 (25)	1,22 (1,3)	368,77 (80,88)	86,94 (5,97)
<i>Zones</i>	Rural	6 (16,7)	1 (0,89)	362,92 (76,48)	88,75 (6,85)
	Semi-urbain	5 (13,9)	1,8 (2,05)	339,11 (72,3)	87 (11,37)
	Urbain	25 (64,9)	1,72 (1,24)	374,52 (154,78)	82,7 (12,62)
<i>Persona choisi</i>	Persona 1	8 (22,2)	Fortement corrélé à l'âge des participants		
	Persona 2	8 (22,2)			
	Persona 3	8 (22,2)			
	Persona 4	4 (11,1)			
	Persona 5	8 (22,2)			
Expérience en santé					
<i>Opération(s)</i>	Oui	32 (88,9)	1,59 (1,21)	376,18 (136,24)	85,39 (11,72)
	Non	4 (11,1)	1,75 (2,22)	299,55 (106,9)	75,62 (8)
<i>Opération(s) en ambulatoire</i>	Oui	18 (50)	1,39 (1,33)	367,15 (142,09)	86,11 (11,8)
	Non	18 (50)	1,83 (1,29)	368,19 (130)	82,5 (11,66)
<i>Souffrant d'une maladie chronique</i>	Oui	11 (30,6)	1,27 (1,35)	386,68 (184,62)	81,14 (15,26)

	Non	25 (69,4)	1,76 (1,3)	359,3 (108,77)	85,7 (9,8)
Expérience dans les dispositifs médicaux					
<i>Prise de tension</i>	Oui	24 (66,7)	1,54 (1,32)	361,43 (131,26)	63,3 (11,22)
	Non	12 (33,3)	1,75 (1,36)	380,15 (145,06)	86,25 (12,9)
<i>Prise d'oxygénation du sang</i>	Oui	5 (13,9)	0,4 (0,55)	309,16 (70,87)	89 (8,02)
	Non	31 (86,1)	1,81 (1,3)	377,11 (140,31)	83,56 (12,12)
Expérience dans les technologies de l'information (TI)					
<i>À l'aise avec la tablette/l'ordinateur/ le téléphone</i>	Très à l'aise avec la technologie	23 (63,9)	1,35 (1,23)	360,16 (120,21)	86,85 (11,24)
	Plutôt à l'aise avec la technologie	11 (30,6)	2,27 (1,42)	401,34 (166,78)	78,18 (11,78)
	Moyennement à l'aise avec la technologie	2 (5,6)	1 (0)	268,81 (33,95)	88,75 (5,3)
	Plutôt pas à l'aise avec la technologie	0 (0)	-	-	-
	Pas du tout à l'aise avec la technologie	0 (0)	-	-	-
<i>Fréquence d'utilisation des technologies</i>	Tous les jours (très souvent)	5 (13,9)	1 (1,22)	314,42 (48,65)	92,5 (6,85)
	Plusieurs fois par semaine (souvent)	12 (33,3)	1,25 (1,36)	364,46 (133,26)	90,42 (6,47)
	Une fois de temps en temps (rarement)	17 (47,2)	2 (1,27)	364,08 (99,16)	79,26 (11,38)
	Occasionnellement (très rarement)	1 (2,8)	3 (-)	856,33 (-)	55 (-)
	Jamais	1 (2,8)	1 (-)	244,8 (-)	85 (-)

permettant la prise de son taux d'oxygène met en évidence un effet significatif sur le nombre d'erreurs de manipulation (*efficacité* : $t = 2,359$; ddl = 34 ; $p = ,024$; $\eta^2 = 0,14$), mais cet effet n'est pas significatif sur le temps de manipulation (*efficacité* : $t = 1,052$; ddl = 34 ; $p = ,3$) ainsi que sur le score au SUS (*satisfaction* : $t = -0,965$; ddl = 34 ; $p = ,341$).

c. Le niveau d'éducation

Le niveau d'éducation n'a aucun impact sur l'utilisabilité en termes de nombre d'erreurs de manipulation (*efficacité* : $F(3,32) = 1,889$; $p = ,151$) ; de temps de manipulation (*efficacité*

: $F(3,32) = 0,698$; $p = ,56$; et relativement au score au SUS (*satisfaction* : $F(3,32) = 1,076$; $p = ,373$).

d. La littératie en santé

Des analyses systématiques entre le niveau de littératie (HLS-EU-Q16 et NVS) et chacun des composants de l'utilisabilité (*efficacité*, *efficience* et *satisfaction* d'après ISO 9241-11, 2018) ont été réalisées. Une représentation descriptive des résultats obtenus aux deux questionnaires de littératie en santé est présentée dans le Tableau 23 ci-dessous.

Tableau 23. Statistiques descriptives des questionnaires HLS-EU-Q16 et NVS.

Groupes	Effectifs (%)
Health literacy Survey (M = 12,97/16 ; ET = 2,952 ; rang = 5 - 16)	
Littératie en santé faible	3 (8,3)
Littératie en santé limitée	9 (25)
Littératie en santé correcte	24 (66,7)
Newest Vital Sign (M = 4,17/6 ; ET = 2,223 ; rang = 0 - 6)	
Littératie en santé faible	6 (16,7)
Littératie en santé limitée	7 (19,4)
Littératie en santé correcte	23 (63,9)

e. Résultats du questionnaire HLS-EU-Q16

Il n'existe pas de corrélation significative entre les résultats de la version européenne du questionnaire Health Literatie Survey (HLS-EU-Q16) et l'utilisabilité que ce soit en termes de nombre d'erreurs de manipulation (*efficacité* : $r = 0,34$; $p = ,844$), de temps de manipulation (*efficience* : $r = -0,40$; $p = ,816$) et relativement au score au SUS (*satisfaction* : $r = 0,144$; $p = ,402$).

Après analyse des corrélations, les participants ont été clustérisés en fonction les mesures du HLS-EU-Q16 (Tableau 24) en suivant les recommandations de Sørensen et al. (2013). Aucune différence intergroupe n'a pu être observée entre les résultats au HLS-EU-Q16 et l'utilisabilité (Tableau 24) que ce soit en termes de nombre d'erreurs de manipulation (*efficacité* : $F(2,33) = 0,277$; $p = ,76$), de temps de manipulation (*efficience* : $F(2,33) = 0,015$; $p = ,985$) et au score au SUS (*satisfaction* : $F(2,33) = 0,483$; $p = ,621$).

Tableau 24. Analyses du résultat au Health Literacy Survey-Europe-16 (HLS-EU-Q16) en fonction de l'utilisabilité (efficacité, efficacité et satisfaction).

Health Literacy Survey – Europe – 16 items N = 36	Efficacité – nombre d’erreurs Moyenne (ET)	Efficiency - temps de manipulation Moyenne (ET)	Satisfaction (score SUS)
Littératie en santé faible (N = 3)	1,67 (2,08)	373,26 (88,76)	83,33 (3,82)
Littératie en santé limitée (N = 9)	1,89 (1,27)	373,35 (98,53)	81,11 (14,53)
Littératie en santé correcte (N = 24)	1,50 (1,28)	364,84 (152,73)	85,62 (11,3)
<i>ANOVA</i>	F (2,33) = 0,277 <i>p</i> = ,76	F (2,33) = 0,015 <i>p</i> = ,985	F (2,33) = 0,483 <i>p</i> = ,621

f. Résultats du questionnaire NVS

Il existe une corrélation significative de taille moyenne entre les résultats de la version française du questionnaire Newest Vital Sign (NVS) et l'utilisabilité en termes de nombre d'erreurs de manipulation (*efficacité* : $r = -0,417$; $p = ,011$), de temps de manipulation (*efficacité* : $r = -0,38$; $p = ,022$) et le score au SUS (*satisfaction* : $r = 0,45$; $p = ,006$). Autrement dit, plus un participant à un haut niveau de littératie en santé (mesuré à l'aide du NVS) moins il réalise d'erreurs de manipulation (i.e., plus il est *efficace*), plus il sera rapide dans la manipulation (i.e., plus il sera *efficient*) et plus sa note au score au SUS sera élevée (i.e., plus il sera *satisfait*).

Après analyse des corrélations, les participants ont été clustérisés en fonction des mesures du NVS (Tableau 23) en suivant les recommandations de Weiss et al. (2005). Aucune différence intergroupe n'a pu être observée entre le niveau de littératie au NVS et l'utilisabilité (Tableau 25) à l'exception du nombre d'erreurs (*efficacité* : $\chi^2 = 6,679$; $p = ,035$).

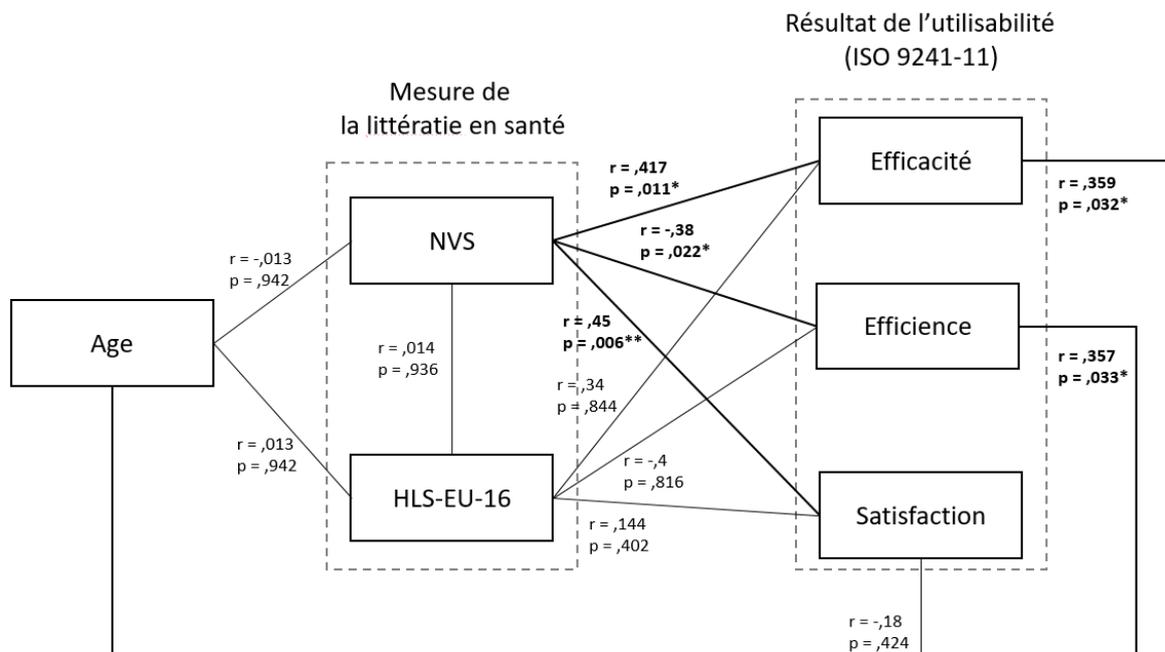
Tableau 25. Analyses du résultat au Newest Vital Sign (NVS) en fonction de l'utilisabilité (efficacité, efficience et satisfaction).

Newest Vital Sign (NVS) N = 36	Efficacité – nombre d'erreurs Moyenne (ET)	Efficience - temps de manipulation Moyenne (ET)	Satisfaction (score SUS) Moyenne (ET)
Littératie en santé faible (N = 6)	2,67 (0,816)	463 (165,18)	77,08 (14,27)
Littératie en santé limitée (N = 7)	1,71 (0,756)	387,72 (219,2)	80,71 (15,05)
Littératie en santé correcte (N = 23)	1,30 (1,43)	336,7 (75,79)	87,28(9,07)
ANOVA	Impossible	Impossible	F (2,33) = 2,392 p = ,107
Test de Kruskal-Wallis (Khi2)	$\chi^2 = 6,679$; p = ,035*	$\chi^2 = 3,07$; p = ,21	$\chi^2 = 2,618$; p = ,270

** . Les corrélations sont significatives à ,01.

* . Les corrélations sont significatives à ,05.

Une analyse intergroupe plus approfondie met en évidence (Figure 50) un effet significatif entre les groupes « littératie en santé faible » et « littératie en santé correcte » en fonction du nombre d'erreurs de manipulation (*efficacité* : Mann-Whitney U = 27 ; p = ,022).



** . Les corrélations sont significatives à ,01.

* . Les corrélations sont significatives à ,05.

Figure 50. Représentation schématique des corrélations entre l'âge, la mesure de la littératie en santé des participants et de l'utilisabilité (ISO 9241-11).

5.2. Analyse qualitative : Satisfaction (qualitative) des utilisateurs du dispositif *Smart Angel*

L'analyse thématique réalisée sur la base des retranscriptions des discours en *think aloud* et des entretiens semi-directifs ont permis de proposer une arborescence de la satisfaction des utilisateurs à l'égard du dispositif *Smart Angel* sur la base de deux thèmes principaux : Facilité d'utilisation perçue (divisée en deux sous-thèmes : Besoin d'information et Erreurs/prévention) et l'utilité perçue (divisée en deux sous-thèmes : avantages et freins). La Figure 51 représente cette arborescence. Nous présentons chacun des thèmes dans les parties suivantes.

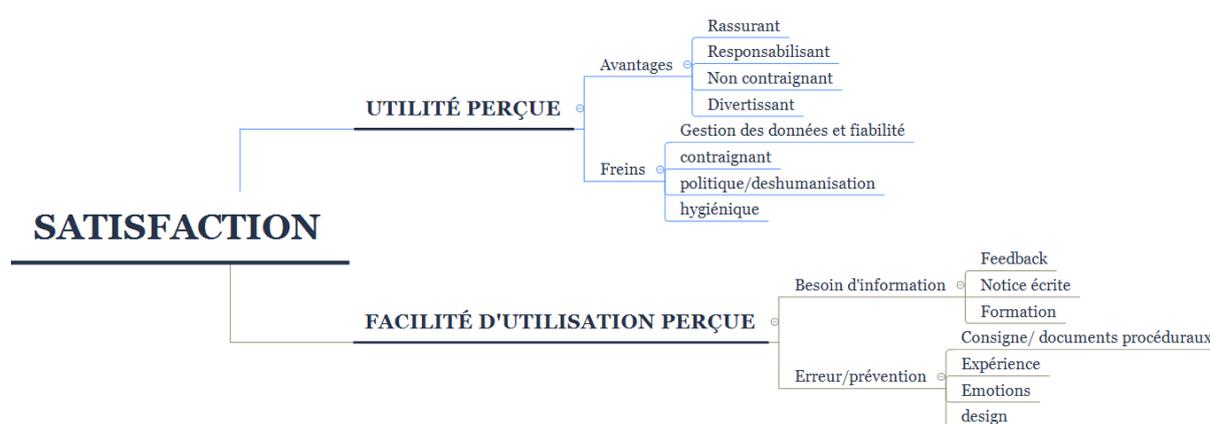


Figure 51. Synthèse de l'analyse qualitative de la satisfaction décomposée en deux sous parties : l'utilité perçue et la facilité d'utilisation perçue.

5.2.1. Facilité d'utilisation perçue

La facilité d'utilisation perçue se décompose en deux sous parties : le besoin d'information et les erreurs/préventions. Nous avons pu mettre en évidence d'où provenaient les principales erreurs d'utilisation (erreurs/prévention) du matériel et une manière d'y remédier (besoin d'information).

a. Erreurs et prévention

Le thème erreur/prévention a pu faire émerger quatre explications pouvant argumenter la survenue des erreurs d'utilisation. Nous avons demandé aux participants de nous expliquer rétrospectivement la survenue de leurs erreurs. Nous avons pu analyser que la non-lecture de la notice, l'expérience antérieure, les émotions et le mauvais design de l'appareil étaient quatre arguments principaux pour expliquer cet évènement. La première explication est la plus largement répandue (douze participants dont 14 références), il s'agit de la non-lecture des

instructions ou leur survol. Nous avons essayé de comprendre pourquoi les utilisateurs ne les lisaient pas. Cinq participants ont déclaré ne jamais lire les instructions « par principe » ou par « personnalité » (« *je ne lis pas les notices, c'est un trait de caractère (rire)* » participant 30) ou perçoivent le matériel avec une très haute facilité d'utilisation (« *C'est quand même super simple ! Donc je me suis un peu trop précipité, j'ai pas vraiment lu les consignes en mettant l'appareil* » participant 7). Par ailleurs, il est important de préciser que six participants ont jugé les instructions comme suffisantes et claires alors qu'ils avaient réalisé au moins une erreur de manipulation. Trois participants expliquent avoir été biaisés par leur propre croyance induisant des erreurs liées à leurs expériences antérieures avec leur propre matériel, leur vécu chez un praticien ou à l'hôpital. Pensant connaître le matériel, ils n'ont pas pris la peine de s'arrêter sur les instructions (« *j'ai subodoré que c'était le même que le mien c'est ça l'erreur que j'ai faite. Moi le mien effectivement tu le mets d'abord tu l'enclenches et de toute façon la lecture se fait sur le cadran, là le fait que ça soit pas sur l'appareil, mais sur la tablette donc là j'ai mal suivi les instructions donc c'est moi.* » participant 17). La peur et l'angoisse de casser le matériel ou de mal faire a parfois eu raison de sept participants (13 références) (« *je me suis précipitée sur les appareils sans m'occuper de la tablette alors que tout était bien dit dans la tablette, mais j'ai comme une espèce de panique* » participant 16). Enfin, le mauvais design est aussi une source d'erreur pour 18 participants (35 références). La taille de la police, l'imprécision ou le manque d'indications et les schémas trop petits sont reprochés afin d'expliquer certaines survenues d'erreurs.

b. Besoins d'informations

Les participants ont pu proposer des éléments d'amélioration afin d'éviter la survenue de leurs erreurs de manipulation. Outre l'agrandissement de la police et des schémas, trois participants (4 références) réclament davantage de feedback (« *non, mais j'ai fait tout à l'envers toute à l'heure parce que ce n'était pas clair.* » participant 6) dans l'application et la présence d'une notice écrite pour cinq participants (12 références). En revanche, lorsque nous leur demandions de juger la formation qu'ils avaient reçue, aucun ne l'a jugé négativement et tous l'ont considérée comme satisfaisante.

5.2.2. Utilité perçue

L'utilité perçue du matériel a été analysée à travers deux sous parties : ses avantages et ses freins. Les participants devaient décrire ce qu'ils avaient compris de l'intérêt de ce système et pourquoi ils l'utiliseraient s'ils étaient amenés à s'en servir.

a. Avantages

Plusieurs avantages ont été mis en avant. Le produit a été jugé rassurant, responsabilisant, non contraignant et divertissant. La plupart des participants ont décrit le dispositif comme rassurant pour soi, mais surtout pour ses proches. L'intérêt d'un lien permettant de se sentir entouré par du personnel hospitalier est davantage mis en avant que l'intérêt médical de pouvoir agir rapidement. Les participants ont aussi compris que le dispositif pouvait avoir un très large intérêt pour leur entourage. Ils se sont majoritairement projetés en tant qu'aidant (pour leur conjoint, parents, grands-parents) et non en tant que patient (« *Tu vois, j'aurais bien aimé que mon mari ait ça oui ! C'est pour l'entourage que je trouve que c'est bien. Parce qu'ils savent quoi faire.* » participant 4). Le dispositif permet aussi de se responsabiliser envers sa santé (« *moi je trouve que c'est même se responsabiliser par rapport à sa santé et tout* » Participant 5). Le dispositif a aussi été jugé comme non contraignant, car il est rapide à utiliser, non invasif, et permet de gagner du temps, car il évite de se déplacer à l'hôpital. Or cette réponse est fautive. Le dispositif permet un suivi postopératoire, mais ne peut pas se substituer à un retour à l'hôpital si celui-ci est nécessaire. Enfin, un élément stimulant l'utilisation du matériel et particulièrement mis en avant par les participants est sa capacité à divertir (« *Je trouve ça cool de pouvoir prendre les mesures soi-même* » Participant 13). Les termes « *cool* » et « *rigolo* » ont particulièrement été utilisés pour désigner le matériel. Cependant, ce point n'est pas toujours positif d'un point de vue médical, car les participants ne prennent pas toujours au sérieux le matériel et semblent davantage curieux d'essayer (« *C'est marrant ton truc c'est un gadget !* » Participant 6).

b. Freins

Nous avons pu identifier quatre freins principaux dont certains peuvent être rédhibitoires comme la fiabilité des données et leur gestion. Par exemple, le tensiomètre a été pointé du doigt comme produit non fiable. Cette méfiance a probablement été causée par les erreurs de manipulation ayant altéré les données. Le tensiomètre a été jugé imprécis ce qui a provoqué une angoisse chez deux participants hypertendus (« *Juste le tensiomètre si on avait du matos ultra performance Ouais à 200 % le truc est utile. Là moi je suis à deux doigts d'aller à la pharmacie pour aller me faire vérifier ma tension* » participant 17). La gestion des données de santé a aussi été un sujet sensible pour sept participants (9 références) (« *Mais maintenant, les mecs de Google ils connaissent ma santé une fois que j'ai connecté ce truc.* » Participant 1). Le dispositif a également été jugé comme contraignant pour plusieurs raisons : son coût pour la société et son caractère polluant et invasif. Douze participants nous ont demandé qui payait pour ce

service et combien ? Le produit a également été jugé polluant, car composé de batteries par un participant et enfin vingt-deux participants ont jugé le dispositif comme invasif, car pouvant être « *bruyant* » (i.e., notification), « *gênant* » (i.e., batterie,), « *intrusif* » (i.e., rigueur dans son utilisation), « *fragile* » voire « *douloureux* ». Trois participants ont mis en avant un intérêt politique, car le produit permet de rentabiliser la chirurgie ambulatoire et de la promouvoir. Les quatre patients ayant mis en avant cet argument sont majoritairement contre la chirurgie ambulatoire et préféreraient être hospitalisés en cas d'opération. Enfin, trois participants ont jugé le dispositif comme peu hygiénique (« *là tout de suite je me dis un matériel qui a déjà été utilisé j'aurais peur qu'il soit dégueulasse. Ils sont lavés désinfectés, mais là le truc d'hygiène qui passe de main en main là franchement ça me dégoûte moi* » Participant 25).

6. Discussion

6.1. Discussion de l'analyse quantitative

L'objectif de cette étude était de mieux comprendre les relations entre quatre caractéristiques utilisateurs (âge, niveau d'éducation, technophilie et littératie en santé) et l'utilisabilité (ISO 9241-11, 2018) comprenant l'*efficacité*, l'*efficience* et la *satisfaction* relatives à l'utilisation du dispositif *Smart Angel*. Pour ce faire, des données sociodémographiques ont été collectées, le niveau de littératie a été investigué au moyen du Health Literacy Survey - Europe (Sørensen, et al., 2012) et du Newest Vital Sign (Weiss et al., 2005), et des mesures d'utilisabilité ont été effectuées (erreurs et temps de manipulation, questionnaire SUS).

Nous avons formulé quatre hypothèses selon lesquelles l'âge, la technophilie, et la littératie en santé auraient un impact sur l'utilisabilité, contrairement au niveau d'éducation qui n'en aurait pas.

Notre première hypothèse stipulait que les utilisateurs plus âgés seraient moins efficaces, efficaces et satisfaits du dispositif par rapport aux jeunes utilisateurs. Nous pouvons partiellement valider cette hypothèse. En effet, les résultats montrent que plus les individus sont jeunes, moins ils réalisent d'erreurs de manipulation (i.e., ils sont plus efficaces) et manient plus rapidement (i.e., ils sont plus efficaces) le matériel. En revanche, nous n'avons pas observé de différence entre l'âge des sujets et le score au SUS (*satisfaction*). L'ensemble de ces résultats est en accord avec les précédents travaux (Jones & Caird, 2017; Liang et al. (2018), Mykityshyn et al. (2002) et Van der Vaart et al. (2019). En effet, ces travaux montrent que les utilisateurs

plus jeunes sont plus efficaces (exemple du glucomètre de Jones et Caird, 2017) et efficaces (exemple du glucomètre de Mykityshyn et al., 2002 et de l'application pour les narcoleptiques de Van der Vaart et al., 2019) par rapport aux utilisateurs plus âgés, et cela avec une corrélation positive et moyenne (Van der Vaart et al., 2019). Cependant, les utilisateurs plus jeunes sont autant satisfaits (score du System Usability Scale) du dispositif que les utilisateurs plus âgés ce qui est en concordance avec les résultats de Liang et al. (2018) et en désaccord avec ceux de Georgsson et Staggers (2016).

Notre deuxième hypothèse s'intéressait à l'absence de relation entre le niveau d'éducation et l'utilisabilité. Les résultats confirment cette hypothèse, car aucun lien significatif n'a été trouvé entre le niveau d'éducation des participants et l'utilisabilité que ce soit en termes d'*efficacité*, d'*efficience* et de *satisfaction*. Ces résultats vont également dans le sens de ceux de Georgsson et Staggers (2016), Liang et al. (2018) et Van der Vaart et al. (2019).

Notre troisième hypothèse s'intéressait à la technophilie (expérience des technologies de l'information et des dispositifs médicaux). Les résultats permettent de valider partiellement cette hypothèse, car aucun lien n'a pu être observé entre l'expérience TI et l'utilisabilité en termes d'*efficacité* et d'*efficience*. En revanche, les participants technophiles ont significativement donné un meilleur score au SUS (*satisfaction*) que les participants ayant un faible niveau de technophilie. Si ces résultats sont en accord avec ceux de Harte et al. (2018), ils sont en contradiction avec les travaux de Georgsson et Staggers, (2016) et Sparkes et al. (2012). Nous expliquons ces résultats par une représentation assez homogène de l'expérience TI en fonction de l'âge des participants dans notre échantillon. Nous pensons que ces items (Agarwal & Prasad, 1998) mettent en exergue la représentation subjective de l'utilisation de technologie (par rapport à son âge) que sa réelle performance dans l'utilisation du matériel. Il est possible que des personnes plus âgées pensent manipuler convenablement une tablette sans pour autant utiliser d'autres fonctionnalités accessibles dans l'outil. Elles se jugeraient alors tout à fait technophiles, car efficaces dans l'utilisation courante de ces technologies. Cependant, on ne connaît pas leur réelle capacité d'adaptation aux technologies. Par exemple, si une mise à jour est réalisée sur une des applications couramment utilisées, il est possible que cela déstabilise la manipulation faite par ces individus.

Ensuite, nous avons pu observer un lien entre l'expérience des dispositifs médicaux et l'utilisabilité. En revanche, l'expérience antérieure dans l'utilisation d'un tensiomètre n'a pas eu d'impact sur l'utilisabilité. Inversement, l'expérience antérieure dans l'utilisation d'un oxymètre de pouls a eu un effet significatif sur l'*efficacité*. Les participants ayant déjà manipulé

un oxymètre de pouls réalisent significativement moins d'erreurs que ceux n'en ayant jamais manipulé. En revanche, l'expérience antérieure dans l'utilisation d'un oxymètre de pouls n'a pas d'effet sur l'*efficacité* et la *satisfaction*. La totalité des sujets déclarant avoir déjà utilisé un oxymètre de pouls a aussi déjà manipulé un tensiomètre. Ce résultat laisse à penser que l'utilisation antérieure de l'oxymètre de pouls combiné à celui du tensiomètre faciliterait la manipulation du dispositif *Smart Angel* en termes d'efficacité. Nous pensons que les participants ayant l'habitude d'utiliser ce type de dispositif complexe ont tendance à se préoccuper des problématiques de santé, ce qui peut être la preuve d'une forte implication du patient pour sa propre santé (Greene & Hibbard, 2012).

Enfin, la quatrième hypothèse postulait que la littératie en santé avait un effet sur l'utilisabilité (*efficacité*, *efficacité* et *satisfaction*). Les scores du HLS-EU-Q16 n'ont montré aucun effet sur l'utilisabilité (Figure 50). À l'inverse, les scores du NVS ont mis en évidence un effet significatif sur le nombre d'erreurs de manipulation (*efficacité*), le temps de manipulation (*efficacité*) et le score au SUS (*satisfaction*). Ce qui est en accord avec les résultats des études précédentes (Czaja et al., 2015 ; Kim & Xie, 2017, Jensen et al., 2010). En effet, des corrélations significatives et de taille moyenne entre le score au NVS et chacune des dimensions de l'utilisabilité ont été observées (Figure 50). Cela suggère que plus les participants ont un haut niveau de littératie, moins ils réalisent d'erreurs de manipulation (i.e., plus ils sont efficaces) d'une part, plus ils sont rapides (i.e., plus ils sont efficaces) et plus le score au SUS sera élevé (i.e., plus ils sont satisfaits) d'autre part. En revanche, après regroupement par cluster des participants tel que recommandé par Weiss et al. (2005), on observe un effet significatif entre le niveau de littératie au NVS et le nombre d'erreurs (*efficacité*), mais aucun lien entre le temps de manipulation (*efficacité*) et le score au SUS (*satisfaction*). Les participants ayant un niveau correct de littératie ont réalisé significativement moins d'erreurs que ceux ayant un niveau faible ou limité de littératie.

Il est important de préciser que les résultats du HLS-EU et du NVS sont contradictoires entre eux et montrent bien le caractère complexe de l'évaluation de la littératie en santé. En outre, nos résultats suggèrent que le HLS-EU interrogeant le participant sur ses capacités subjectives à avoir accès à l'information en santé et à prendre des décisions, constitue un biais important dans la mesure de la littératie en santé. Certains participants peuvent affirmer n'avoir aucune difficulté à utiliser des informations de santé, mais aucune vérification sur la bonne conduite de ces informations ne le confirme. À l'inverse, l'instrument NVS semble plus adapté à la récolte d'informations sur les capacités cognitives des sujets, car il s'agit d'une épreuve

permettant de récolter des informations sur la réflexion des participants à la lecture d'une étiquette alimentaire, constituant ainsi une évaluation plus objective de la littératie en santé.

6.2. Discussion de l'analyse qualitative

Dans un deuxième temps, nous avons réalisé une analyse thématique de la satisfaction sur la base des discours « *think aloud* » et des entretiens semi-directifs réalisés avec chaque participant. Nous avons divisé la satisfaction (qualitative) en deux thématiques suivant le modèle de Yen (2010) : la facilité d'utilisation perçue et l'utilité perçue. Nous avons formulé deux hypothèses par rapport à ces deux thématiques.

Notre cinquième hypothèse s'intéressait à la facilité d'utilisation et plus exactement à la survenue des erreurs. Cette hypothèse est validée. En effet, nous avons pu observer qu'une grande majorité des participants ne lisaient pas (ou partiellement) les instructions. Cette observation rejoint les résultats des travaux précédents (e.g., Hoc, 1987; Silver & Wogalter, 1991; Wogalter et al., 2016; Wright et al., 1982) ainsi que nos propres observations (cf. Partie II, Chapitre 3, p. 160). En effet Silver et Wogalter (1991) ont montré que la majorité des utilisateurs ne lisent pas les instructions accompagnant les produits, mêmes pour ceux ayant un risque important pour l'utilisateur (e.g., pesticides). Quatre arguments pourraient expliquer cette absence de lecture. Premièrement, nous avons pu observer que le dispositif était décrit comme « *divertissant* » ce qui montre un risque perçu faible du matériel. Si le produit n'est pas perçu comme dangereux, les instructions seront moins lues (Wogalter & Barlow, 1990). Deuxièmement, les utilisateurs ont expliqué que la non-lecture des instructions faisait partie de leur personnalité. Cela peut s'expliquer par le fait que la plupart des utilisateurs préfèrent apprendre par l'action (e.g., Carroll, 1990; Hoc, 1987), ce qui est cognitivement moins coûteux (Fu & Gray, 2006). Troisièmement, les participants ont décrété que les instructions étaient mal conçues (e.g., Helyar, 1992). Quatrièmement, certains participants ont aussi expliqué avoir été induits en erreur à cause de leur expérience antérieure dans un produit similaire ou lors d'une expérience médicale. Wright et al. (1982) ont montré que le caractère familier d'un produit peut limiter la lecture des instructions. Deux autres arguments sont référencés dans la littérature, que nous n'avons pas retrouvé dans notre analyse, expliquant le désintérêt pour la lecture des instructions : la forme du produit (Silver & Wogalter, 1991) et/ou la présentation des instructions (Silver & Wogalter, 1991 ; Wogalter & Barlow, 1990). L'apparence simpliste de l'appareil pourrait être étudiée afin de vérifier un potentiel lien avec le risque perçu. Un

remodelage de la forme des objets connectés permettrait potentiellement de les rendre plus « médicaux » en ajoutant des notions de design émotionnel ou l'ingénierie Kansei (Nagamachi, 1995) dans le cycle de conception et d'évaluation du produit.

Notre sixième hypothèse s'intéressait à l'utilité perçue et plus exactement aux avantages et aux freins du dispositif. Nous avons considéré que la sécurité et la fiabilité des données (e.g., Allaert, 2016; Béranger, 2014; Laubie, 2011), la responsabilisation (Eysenbach, 2001), l'apport pour les zones isolées (Simon, 2017) pour les proches (Laubie, 2011) seraient les principaux freins et avantages du dispositif *Smart Angel*. Nous pouvons valider partiellement notre hypothèse. En effet, nous retrouvons les arguments redondants à la littérature scientifique : intérêt pour les proches (Laubie, 2011), responsabilisation des utilisateurs (Eysenbach, 2001) et crainte de la sécurité et la fiabilité des données (e.g., Allaert, 2016; Béranger, 2014; Laubie, 2011). Pourtant contrairement à la littérature (e.g., Carayon & Hoonakker, 2019; Cifter, 2017; Greysen, Garcia, Sudore, Cenzer, & Covinsky, 2014), ces dispositifs ne sont pas jugés comme complexes par les utilisateurs ce qui les rend d'autant plus dangereux. Nous avons pu mettre en évidence d'autres freins à l'utilisation du matériel comme le lien politique, la déshumanisation, l'écologie, la fragilité du système ou encore l'hygiène de ces dispositifs passant de patient en patient. Ces éléments ne sont pas, à notre connaissance, représentés dans la littérature scientifique. Pour les avantages, nous avons pu remarquer que les participants ont très rarement abordé l'intérêt médical du dispositif. Ils ont parfois déclaré que le produit était rassurant, car cela permet de conserver un lien avec l'hôpital, mais ont très peu mentionné directement l'intérêt médical. Lorsque nous leur demandions s'ils accepteraient le dispositif s'ils étaient en situation, la réponse principale était « cela dépend (du contexte, de l'opération, du coût) ». Nous pouvons alors nous demander si les participants ont réellement compris l'utilité-destination (Loup-Escande et al., 2013) du dispositif. Majoritairement, les participants ne comprenaient pas l'intérêt du dispositif si l'opération subie était peu lourde puisque les risques à domicile étaient faibles, mais ne comprenaient pas non plus l'intérêt de l'ambulatoire si l'opération subie était lourde puisqu'ils déclaraient préférer ne pas être opérés en ambulatoire. Le dispositif devient alors un substitut à l'hospitalisation lors d'opérations à risque élevé.

7. Conclusions et perspectives de recherche

En conclusion, la présente étude apporte des éléments de compréhension théoriques sur les effets des caractéristiques utilisateurs (e.g., âge, expérience, niveau d'éducation, littératie en santé) à travers l'utilisation de personas sur l'utilisabilité (*efficacité, efficacité et satisfaction* d'après la norme ISO 9241-11) dans le cas précis du dispositif médical connecté *Smart Angel*. Cette étude comporte un apport méthodologique dans la mesure où elle a permis de révéler les différences de collecte de données entre le Newest Vital Sign et le Health Literacy Survey – Europe - 16 prouvant ainsi l'importance de poursuivre les recherches dans le domaine des outils de mesure de la littératie en santé. De plus, ces résultats nous permettent de mieux comprendre l'importance de l'impact de la technophilie des personnes plus âgées ayant un niveau correct de littératie en santé sur l'utilisabilité.

À la suite de cette étude, quatre perspectives de recherche peuvent être suggérées.

Premièrement, la pertinence de la méthode des personas en phase d'évaluation d'un prototype n'a actuellement jamais été prouvée. Cette méthode est classiquement utilisée en phase de conception par des concepteurs (ergonomes, designers, ingénieurs voire des futurs utilisateurs), mais beaucoup plus rarement mobilisée dans un cadre d'évaluation. Afin de valider cette méthode dans ce nouveau contexte d'usage en phase d'évaluation, il serait nécessaire de reproduire cette étude en y ajoutant un groupe contrôle soit un groupe sans présentation des personas.

Deuxièmement, la formation réalisée par l'expérimentateur pourrait être adaptée en fonction des niveaux de littératie des participants. En effet, la principale difficulté dans l'utilisation d'un dispositif médical est la compréhension des procédures et celles-ci ne peuvent avoir lieu si la formation en amont n'est pas suffisante (Larsen, 2017). Il conviendrait sûrement d'adapter la formation en fonction des âges et des niveaux de littératie. La démonstration réalisée par l'expérimentateur peut être suffisante pour des groupes ayant un niveau adéquat de littératie en santé. À l'inverse, pour les groupes au niveau faible ou limité de littératie en santé, un enseignement plus poussé devrait être envisagé.

Troisièmement, le choix des questionnaires est une étape cruciale dans la mesure du niveau de littératie en santé. En effet, nous avons pu observer une importante disparité dans les résultats entre le HLS-EU-Q16 et le NVS. Comme expliqué en discussion, ces deux questionnaires ne semblent pas évaluer les mêmes dimensions de la littératie en santé. Des travaux visant à

comprendre ce qu'évaluent précisément les questionnaires de littératie en santé mériteraient d'être réalisés.

Enfin, au-delà de la littératie en santé, il conviendrait aujourd'hui de mesurer le niveau de littératie en *e-santé* (e.g., Van der Vaart et al., 2019). Malheureusement, aucun questionnaire valide en français n'existe sur ce sujet. Ainsi, des traductions/adaptations plus systématiques de ces outils devraient être envisagées dans des études ultérieures.

8. Synthèse Chapitre 4 – Étude 2

Cette étude 2 correspond au **niveau 3** (tâches – utilisateurs – ressources – système) de notre modèle HANDLED. **L'objectif principal de cette étude est de comprendre en quoi la complexité intrinsèque de l'ensemble du système *Smart Angel* impacte l'utilisabilité.** Afin de répondre à cet objectif nous avons divisé ce travail en deux analyses (quantitatives et qualitatives) répondant chacune à deux sous-objectifs.

L'analyse quantitative a permis d'analyser les relations qu'entretiennent les quatre caractéristiques utilisateurs (i.e., âge, niveau d'éducation, technophilie et littératie en santé) et l'utilisabilité (ISO 9241-11 : 2018 : *efficacité, efficacité et satisfaction*) du dispositif *Smart Angel*. L'analyse qualitative a permis de mieux comprendre comment les utilisateurs percevaient le dispositif *Smart Angel* au travers de deux attributs de la satisfaction : la *facilité d'utilisation* et *l'utilité perçue* issus du modèle de Yen (2010).

Pour cela, nous avons réalisé un test utilisateur auprès de 36 participants âgés de 20 à 64 ans (moy = 40,75 ans, ET = 14,45). **Une approche mixte** (subjective, objective, qualitative et quantitative) utilisant divers outils a été utilisée (observation directe et instrumentée par vidéo, *think aloud*, questionnaires et entretiens). Les participants ont, tout d'abord, visionné un scénario de personas leur ressemblant, puis ils ont rempli un questionnaire afin d'évaluer leur niveau de littératie en santé et de technophilie, enfin ils ont manipulé le prototype 1 du dispositif *Smart Angel* tout en pensant à voix haute (*think aloud*).

Les mesures d'efficacité et d'efficacité ont été recueillies par observation instrumentée par vidéo. Les mesures de satisfaction ont été recueillies par questionnaire SUS et entretiens semi-directifs. Une analyse thématique, basée sur les retranscriptions des entretiens, a été effectuée afin de proposer une arborescence de la satisfaction.

Sur l'ensemble du système *Smart Angel*, les utilisateurs ont commis environ 1,63 erreurs (ET = 1,26) et la moyenne du score au SUS était de 84,13 (ET = 11,45), ce qui est considéré comme "acceptable". **Aucun abandon n'a été signalé.** Les résultats ont montré **une corrélation modérée et significative entre l'efficacité ($r = 0,417$; $p = ,011$), l'efficacité ($r = -0,38$; $p = ,022$) et la satisfaction ($r = 0,45$; $p = ,006$)** du dispositif et la littératie en santé (NVS). Les résultats ont également révélé une corrélation modérée, positive et significative entre l'âge et l'efficacité ($r = 0,359$; $p = ,032$), et entre l'âge et l'efficacité ($r = 0,357$; $p = ,033$). Comme pour l'étude 1a, nous avons identifié l'existence d'un effet de seuil entre la

littératie en santé et l'efficacité. En effet, il y a un effet significatif entre les groupes « littératie en santé faible » et « littératie en santé correcte » en fonction du nombre d'erreurs de manipulation (efficacité : Mann-Whitney $U = 27$; $p = ,022$).

L'analyse thématique a révélé, premièrement, qu'une grande majorité des participants ne lisaient pas (ou partiellement) les instructions liées à un faible risque perçu, un coût cognitif trop élevé, une expérience antérieure sur des outils similaires, des instructions peu attractives et mal conçues. Deuxièmement, nous nous sommes intéressées à l'utilité perçue et plus exactement aux avantages et aux freins du dispositif. Nous avons pu retrouver des arguments redondants à la littérature scientifique : intérêt pour les proches, responsabilisation des utilisateurs, sécurité et fiabilité des données. Nous avons également révélé des arguments absents de la littérature scientifique : le lien politique, la déshumanisation, l'écologie, la fragilité du dispositif et l'hygiène du dispositif.

Les résultats de cette étude ont permis de proposer des recommandations pour la conception d'un deuxième prototype testé lors de l'étude 3.

CHAPITRE 5 : ÉTUDE 3. EFFETS DES CARACTÉRISTIQUES UTILISATEURS ET DES PÉDAGOGIES HOSPITALIÈRES EN CHIRURGIE AMBULATOIRE SUR L'UTILISABILITÉ DU DISPOSITIF *SMART ANGEL* EN SIMULATION

1. Préambule

Ce travail a été réalisé avec l'aide de Camille Sagnier, ingénieure d'étude ; Goran Loiseau, Alexandre Lamotte, Stagiaires de Licence 3 Psychologie Parcours Psychologie du Travail, Ergonomie, Psychologie Sociale ; Cyril Moureau, post-doctorant sur le Projet *Smart Angel* et Pierre Damay, UX designer à Evolucare.

1.1. Contexte et objectif industriels

L'étude précédente (cf. Partie II, Chapitre 4, p. 189) a permis de faire remonter un certain nombre de recommandations de conception prises en compte dans une deuxième version du prototype (Figure 52). À la suite de cette étude, les objets connectés n'ont pas changé, seuls certains formats de présentation des instructions et des schémas ont été modifiés et ajoutés. Les industriels ont rédigé le dossier permettant d'obtenir la certification européenne. Une évaluation de ce dernier prototype avant l'étude clinique viendrait nourrir ce dossier. Ainsi, l'objectif industriel de cette étude est d'évaluer l'utilisabilité du système dans sa globalité auprès d'une population cible du dispositif dans un environnement le plus représentatif possible.

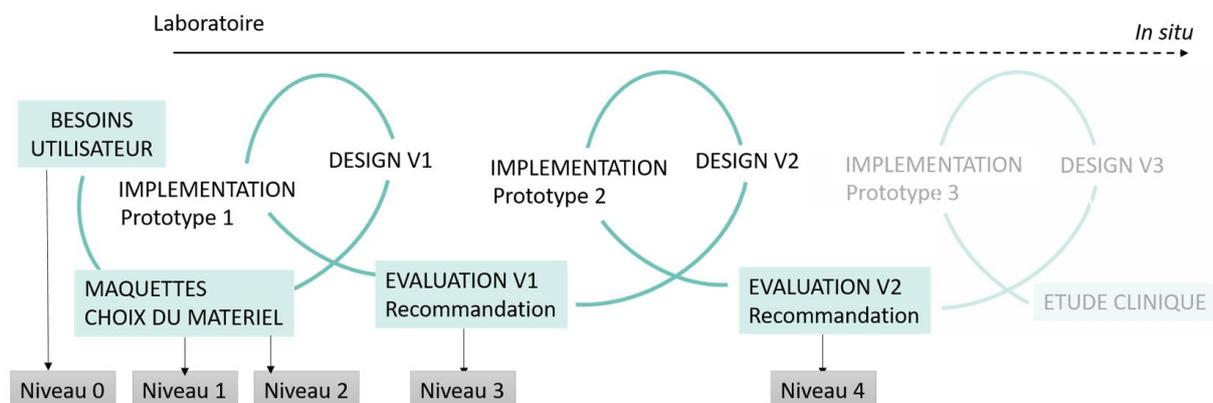


Figure 52. Étude 3 correspondant au niveau 4 du processus de conception du dispositif *Smart Angel*.

1.2. Contexte et objectif théoriques

Cette étude réalisée correspond au niveau 3 du processus de conception du dispositif *Smart Angel* (Figure 52). Comme le chapitre précédent, cette étude mobilise deux approches (quantitative et qualitative) s'appuyant sur deux objectifs théoriques. Le premier objectif cherche à identifier les éléments contextuels pouvant impacter l'utilisabilité (e.g., Borsci et al., 2018; Grebin, Echeveste, Magnago, Tanure, & Pulgati, 2018), tout en complexifiant le contexte d'usage *via* la considération de « l'environnement » jusque-là exclu des études précédentes. Pour cela, nous avons eu recours à la simulation (Kushniruk et al., 2013) en reproduisant des éléments clés du parcours patients identifiés dans l'étude exploratoire (cf. Partie II, Chapitre 1, p. 112).

Au vu des résultats des précédentes études, nous avons été amenés à réaliser quatre choix méthodologiques. Premièrement, dans l'étude 1b (cf. Partie II, Chapitre 3, p. 160), nous avons discuté de l'importance d'évaluer la première utilisation. Nous pensons que l'utilisateur peut apprendre à manipuler le matériel avec des erreurs et ainsi les reproduire lors d'autres manipulations (Larsen, 2017). Afin de vérifier cette hypothèse, nous avons mesuré l'*apprenabilité* (métrique citée dans les modèles de Nielsen, 1993 et Shackel, 1991). Deuxièmement, pour prendre conscience de l'impact du système sur l'utilisabilité, nous avons décidé d'observer indépendamment chaque élément du système (la tablette, l'oxymètre de pouls et le tensiomètre) contrairement à l'étude 2 (cf. Partie II, Chapitre 4, p. 189). Troisièmement, cet ensemble de nouvelles observations a été ajouté à celles qui nous ont semblé les plus pertinentes dans les précédentes études comme la prise en compte des facteurs cognitifs, émotionnels et sociaux, mais aussi des analyses vidéos fines mesurant le temps de lecture des instructions comme dans l'étude 1b (cf. Partie II, Chapitre 2, p. 160). Quatrièmement, nous avons également voulu inclure l'impact des formations hospitalières sur l'utilisabilité jusqu'alors seulement abordé dans notre étude exploratoire (cf. Partie II, Chapitre 1, p. 112). Nous avons vu lors de cette étude exploratoire que l'infirmière de recherche, en charge de mettre en place le protocole pour l'étude clinique, comptait utiliser une formation « passive ». Nous souhaitons donc vérifier l'impact de cette formation en randomisant nos passations avec une autre formation. Pour cela, nous avons choisi la formation « active » (Larsen, 2017). Nous détaillons ces formations en introduction de ce chapitre (cf. Partie II, Chapitre 5, 2. Introduction, p. 230).

Le deuxième objectif, issu de l'analyse qualitative, était d'étudier l'impact de l'environnement sur les décisions des participants. Pour cela, nous nous sommes intéressées de

manière exploratoire à la satisfaction dite « qualitative » en nous basant sur le modèle de Yen (2010) composé de : (1) la facilité d'utilisation perçue et (2) l'utilité perçue. Pour la facilité d'utilisation perçue, nous avons modifié deux éléments méthodologiques par rapport à l'étude précédente (cf. Partie II, Chapitre 4, p. 189) s'appuyant pourtant sur le même modèle (Yen, 2010). D'une part, nous ne pouvions pas appliquer une méthode *think aloud* (contrairement à l'étude précédente) puisque le patient/participant se trouvait seul dans la pièce. Pour pallier ces informations manquantes, nous avons pris note des erreurs commises pendant les manipulations (grâce aux retranscriptions vidéo en direct en salle de gestion des manipulations) et nous avons interrogé le patient/participant lors d'un entretien semi-directif. D'autre part, afin d'obtenir des informations supplémentaires sur le dispositif, nous avons aussi interrogé les étudiants en médecine acceptant de participer à un entretien ou focus group comme le recommandent Ten Haken et al. (2018).

La Figure 53 modélise les différents éléments évalués dans cette étude basée sur la matrice : *tâches – utilisateurs – ressources – système – environnements*.

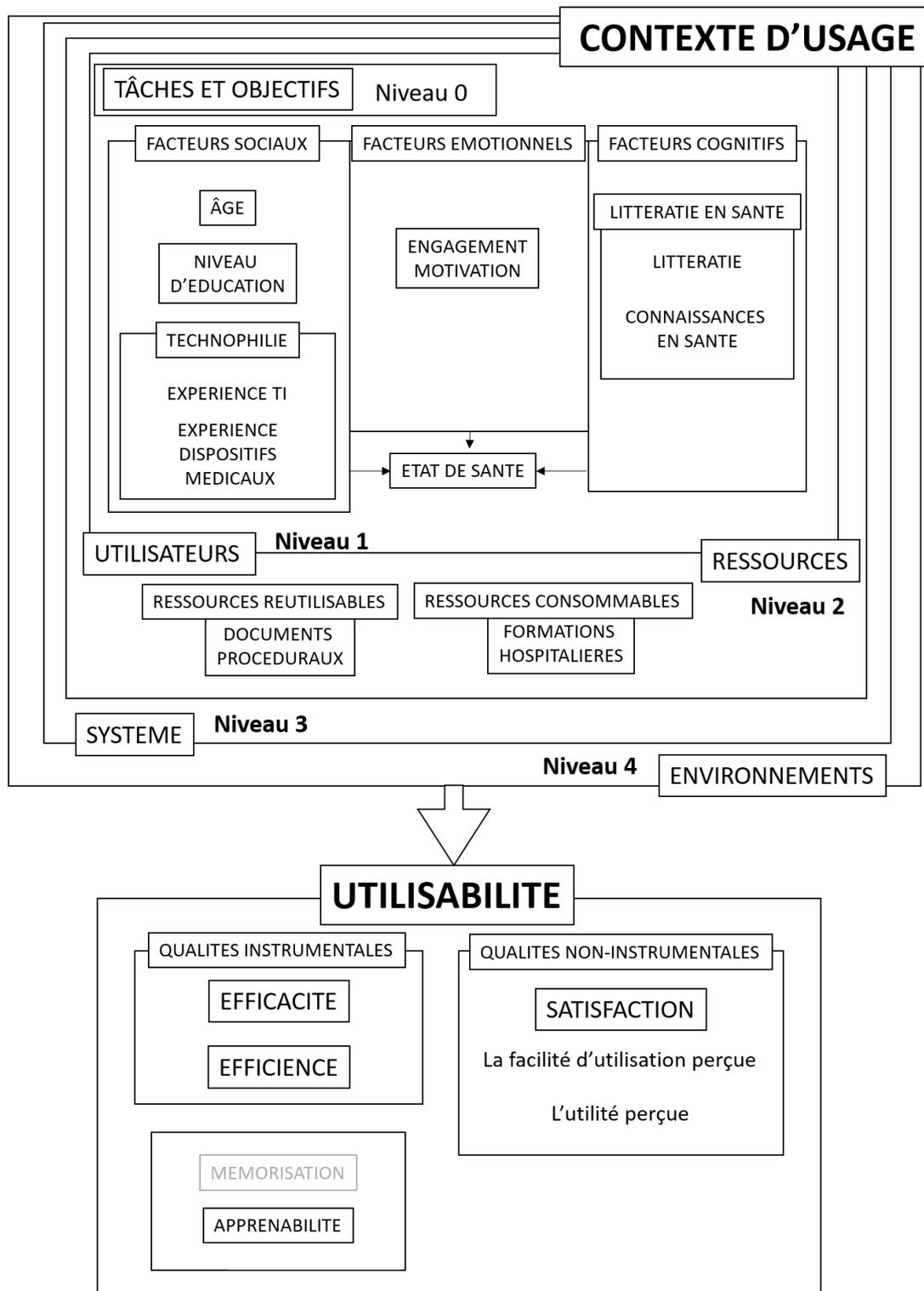


Figure 53. Attributs éprouvés lors de l'étude 3 du contexte d'usage sur l'utilisabilité (*efficacité, efficacité, satisfaction et apprenabilité*).

2. Introduction

L'utilisabilité est au centre des préoccupations pour l'acceptabilité des dispositifs connectés en santé (Kushniruk, 2002). Plusieurs chercheurs préconisent des évaluations en milieu *in situ* (e.g., Abowd & Mynatt, 2000; Brewster, 2002). Cependant, l'utilisation d'un prototype dans un environnement *in situ* peut être dangereux pour le patient puisque celui-ci peut réaliser des mesures erronées ou se blesser avec l'appareil pouvant entraîner un rejet de la technologie. De plus, les évaluations de l'utilisabilité de systèmes ubiquitaires et médicaux sont très complexes à mettre en œuvre ce qui est en partie lié aux contraintes méthodologiques (Kjeldskov et al., 2004, Jambon et al., 2010). En effet, l'observation directe de manipulation de ces dispositifs est primordiale pour limiter les erreurs de manipulation. Or, en milieu *in situ*, ces observations sont limitées.

2.1. La simulation

Afin de pallier ces difficultés, le recours à la simulation semble être une solution adaptée. Kushniruk et al. (2013) expliquent que la simulation peut apporter des avantages considérables aux évaluations de l'utilisabilité si celle-ci est correctement utilisée. Les auteurs recommandent la conception de scénario d'essai écologiquement valable, de qualité et le développement de la simulation dans des environnements *in situ*. Cet environnement virtuel permet d'appréhender l'environnement futur du dispositif.

2.2. L'évaluation de l'utilisabilité

2.2.1. Le choix des métriques

Les évaluations de l'utilisabilité des dispositifs connectés en santé doivent également être standardisées, multidimensionnelles et holistiques (e.g., Ammenwerth et al., 2003; Georgsson & Staggers, 2016; Jaspers, 2009; Nelson et al., 2017; Peute et al., 2013; Peute et al., 2008; Yen & Bakken, 2012). Pour cela, il est important de varier les outils permettant des mesures subjectives et objectives, ainsi que qualitatives et quantitatives (Ammenwerth et al., 2003; Jaspers, 2009). Ces choix permettront la mise en place d'une méthodologie solide et mixte. Il est également nécessaire de mesurer indépendamment l'ensemble de des métriques de l'utilisabilité (Nielsen, 1994b). Le choix de ces métriques est aussi largement discuté. Certaines

études d'utilisabilité dans le domaine de la santé se basent sur les métriques de la norme ISO 9241-11 comprenant l'efficacité, l'efficience et la satisfaction (e.g., Constantinescu et al., 2019; Georgsson & Staggers, 2016; Gunter et al., 2016; Hussain et al., 2015). D'autres études se basent sur les métriques de Nielsen (1993) comprenant la facilité d'apprentissage, l'efficience, l'apprenabilité, la prévention des erreurs et la satisfaction (e.g., Ayubi et al., 2014; Garcia, 2019; Krieke et al., 2012; Sparkes et al., 2012). Ces deux modèles sont similaires à l'exception de l'intégration de la métrique de *l'apprenabilité* (ou *facilité d'apprentissage*) présente dans le modèle de Nielsen (1993) et absente de la norme ISO 9241-11. Le choix des métriques de l'utilisabilité dans le domaine des dispositifs médicaux reste, à notre connaissance, très peu étudié. Par exemple, Larsen (2017) considère que la première utilisation est la plus importante car toute erreur de manipulation pourrait être reconduite lors d'autres utilisations. De ce fait, la notion *d'apprenabilité* n'apporterait pas d'éléments pertinents dans une évaluation de l'utilisabilité.

La satisfaction, métrique présente dans la plupart des modèles d'utilisabilité (cf. 2.4.2. Partie I, Chapitre 1, 2.4.2. L'inventaire des métriques utilisées sur un recensement de modèle entre 1971 à 2018, p. 23) est également une métrique sur laquelle orbitent de nombreux débats. Cette métrique est constituée, d'après le modèle mHealth IT usability de Yen (2010), de la facilité d'utilisation perçue et de l'utilité perçue. À notre connaissance, peu d'études se sont intéressées à la composition de la satisfaction dans le domaine des dispositifs connectés en santé.

2.2.2. Les composants du contexte d'usage

Il est également nécessaire, lors d'une évaluation de l'utilisabilité, d'inclure les composants du contexte d'usage (incluant les ressources, les utilisateurs et les environnements) impactant l'utilisabilité. Nous avons vu que l'environnement peut être simulé.

a. Les ressources

Les ressources peuvent être de deux types d'après la norme ISO 9241-11 (2018) : réutilisables (e.g., documents procéduraux) et consommables (e.g., programme d'éducation thérapeutique). Les ressources sont dites réutilisables lorsque l'utilisateur peut relire et revenir autant de fois qu'il le désire sur ces instructions afin d'utiliser le système (cf. Partie II, Chapitre 3, p. 160). A l'inverse, les ressources consommables ne sont présentées qu'une seule fois à l'utilisateur. Ce dernier doit alors garder en mémoire ces informations pour utiliser le dispositif. C'est le cas de la formation au dispositif *Smart Angel* qui sera dispensée aux utilisateurs lors de

leur réveil (cf. Partie II, Chapitre 1, 5.2. Événements pour l'intégration du dispositif *Smart Angel* au domicile patient, p. 126). Larsen (2017) met en évidence deux types de stratégies de formation : « active » *versus* « passive » (cf. Partie I, Chapitre 2, 4.2.2. Les ressources consommables : Les formations et les programmes d'éducation thérapeutique, p. 94). La stratégie « active » consiste, pour le personnel hospitalier en charge de la formation, à créer des erreurs de manipulation devant le patient tout en expliquant la présence de l'erreur. Le personnel hospitalier énumère alors l'ensemble des erreurs à ne pas commettre avec le matériel. A l'inverse, la stratégie « passive » consiste à laisser le patient découvrir et manipuler seul le matériel et corriger en temps réel ses propres erreurs de manipulation. À notre connaissance, aucune étude ne s'est intéressée à l'utilisation de stratégie de la pratique d'enseignement (« active/passive ») en prenant en compte l'ensemble des caractéristiques de l'utilisateur sur l'utilisabilité incluant les métriques de l'*efficacité*, l'*efficience*, la *satisfaction* et l'*apprenabilité* d'un dispositif médical pour le domicile patient.

b. Les caractéristiques utilisateurs

Les caractéristiques utilisateurs sont également essentielles à prendre en compte. Nous avons vu précédemment que l'âge, la littératie en santé et la technophilie pouvaient impacter l'utilisabilité de dispositifs connectés en santé (e.g., Chaniaud et al., 2020; Georgsson & Staggers, 2016; Harte et al., 2018; Jones & Caird, 2017; Kaufman et al., 2003; Liang et al., 2018; Mykityshyn et al., 2002; Sparkes et al., 2012; Van der Vaart et al., 2019; Wong et al., 2009). Il est donc nécessaire de vérifier l'impact de chacun de ces éléments sur l'utilisabilité du dispositif.

3. Objectifs et hypothèses

Le recours à un environnement simulé afin d'évaluer l'utilisabilité de dispositif connecté en santé semble être une solution méthodologique intéressante pour pallier les difficultés du milieu *in situ*. Cet environnement correspond au domicile du patient – comme nous l'avons pris en compte dans l'étude 2 (cf. Partie II, Chapitre 4) – mais également ce qui précède l'utilisation du dispositif, soit le parcours de soin du patient lors de son intervention en ambulatoire (cf. Partie II, Chapitre 1, 2.1.1. Parcours patient en ambulatoire, p. 114). Il est donc nécessaire de procéder à une simulation de l'ensemble du parcours patient pour parfaire cette évaluation. Pour cela, il est également nécessaire de réaliser une évaluation mixte, holistique et

standardisée de l'utilisabilité tout en prenant en compte l'ensemble des composants du contexte d'usage. Ainsi, l'objectif principal de cette étude est de **mieux comprendre les relations qu'entretiennent le contexte d'usage à travers les caractéristiques utilisateurs (littératie en santé, âge, technophilie), les ressources (le programme d'éducation « actif » et « passif »), le système et l'environnement (simulé) sur l'utilisabilité (efficacité, efficience, satisfaction et apprenabilité) du dispositif *Smart Angel* dans sa globalité.**

Pour ce faire, nous avons formulé six hypothèses :

- (H1) le nombre d'essais (*apprenabilité*) n'a pas d'impact sur l'efficacité et l'efficience du dispositif *Smart Angel* (e.g., Larsen, 2017) ;
- (H2) Les utilisateurs ayant reçu un programme d'éducation « actif » seront plus efficaces, efficaces, satisfaits du dispositif *Smart Angel* que ceux qui auront reçu un programme d'éducation « passif » (e.g., Larsen, 2017) ;
- (H3) Les utilisateurs âgés seront moins efficaces, efficaces et satisfaits le dispositif *Smart Angel* que les jeunes utilisateurs (e.g., Georgsson & Staggers, 2016; Jones & Caird, 2017; Mykityshyn et al., 2002; Van der Vaart et al., 2019) ;
- (H4) Les utilisateurs ayant un niveau faible de littératie en santé (mesurée par les scores de NVS et de HLS-EU) seront moins efficaces, efficaces et satisfaits du dispositif *Smart Angel* que ceux ayant un niveau élevé de littératie en santé (e.g., Jensen et al., 2010; Kaufman et al., 2003) ;
- (H5) Le niveau de technophilie n'aura pas d'effet sur l'efficacité, l'efficience et la satisfaction du dispositif *Smart Angel* (e.g., Georgsson & Staggers, 2016; Liang et al., 2018; Van der Vaart et al., 2019).

Dans un deuxième objectif, nous avons poussé davantage nos recherches sur la satisfaction du dispositif à travers deux notions (l'utilité perçue et la facilité d'utilisation perçue) en nous basant sur une analyse qualitative grâce à une analyse thématique du discours des utilisateurs (patients et personnels hospitaliers). Nous avons pu formuler une hypothèse :

- (H6) L'environnement (simulé) impacte le jugement (i.e., la satisfaction) des patients/participants sur le dispositif *Smart Angel*.

4. Méthode

4.1. Participants

Vingt-huit patients/participants (17 femmes et 11 hommes, Moy = 30,79 ans ; ET = 12,64 ; rang = 19 – 64 ans) et vingt étudiants en médecine ont participé à cette étude. Les critères d'inclusion des patients/participants étaient les suivants : le participant devait avoir une connexion 4G à son domicile, être âgé de moins de 75 ans, être éligible à la chirurgie ambulatoire, ne pas être seul à son domicile. Les critères d'inclusion des étudiants étaient les suivants : les étudiants devaient avoir réalisé des stages dans des services similaires (e.g., consultations généralistes) et/ou devaient être au minimum en 6^e année de médecine pour pouvoir endosser le rôle d'anesthésiste et au minimum en 3^e année pour former le patient au dispositif *Smart Angel*. L'ensemble des participants (patients et étudiants en médecine) était de langue maternelle française et a signé un formulaire de consentement après avoir été informé du déroulement de l'étude. L'étude est conforme aux recommandations éthiques de la déclaration d'Helsinki. Les patients/participants ont été recrutés à l'aide d'annonces en ligne ou dans les journaux locaux. Les étudiants en médecine ont été recrutés par la diffusion d'annonces réalisée par l'administration du CHU d'Amiens. Seuls, les patients/participants recevaient un dédommagement en chèque cadeau d'une valeur de 20€ pour leur contribution. Les passations se sont déroulées au simUsanté (centre de formation par simulation des étudiants de médecins) et simUlogis (domicile simulé dans le centre de formation) au CHU d'Amiens.

4.2. Matériel et mesures

Le matériel de notre expérience est composé : 1) du dispositif *Smart Angel*, 2) d'une mise en situation composée de trois scénarios différents (la consultation anesthésique, le programme d'éducation et le parcours patient), 3) de questionnaires (deux questionnaires évaluant le niveau de littératie en santé, à savoir, le NVS et le HLS-EU-16, un questionnaire relatif aux données sociodémographiques, et un questionnaire d'évaluation de la satisfaction, à savoir le SUS) et 4) d'entretiens semi-directifs ou de focus groups lorsque les participants le permettaient.

4.2.1. Le dispositif *Smart Angel* et son implantation

Le dispositif *Smart Angel* est constitué des mêmes éléments que lors de l'étude 2 (cf. Partie II, Chapitre 4, p. 189). Seules des modifications de formes, de couleurs, d'information sur les

procédures et schémas ont été implémentées à l'application afin d'améliorer l'utilisabilité du système.

L'intégration du dispositif *Smart Angel* comme envisagée pour l'étude clinique initialement prévue est présentée dans la Figure 54. Ce schéma suit le parcours patient en ambulatoire comme présenté dans l'étude exploratoire (cf. Partie II, Chapitre 1, 2.1. La chirurgie ambulatoire en France et leurs patients, p. 113). Nous avons ajouté les explications, la formation, l'utilisation et le retour du matériel à l'hôpital. Les encadrés bleus (e.g., *Explication du matériel Smart Angel et choix du patient*) représentent les événements d'interaction entre le patient et le matériel.

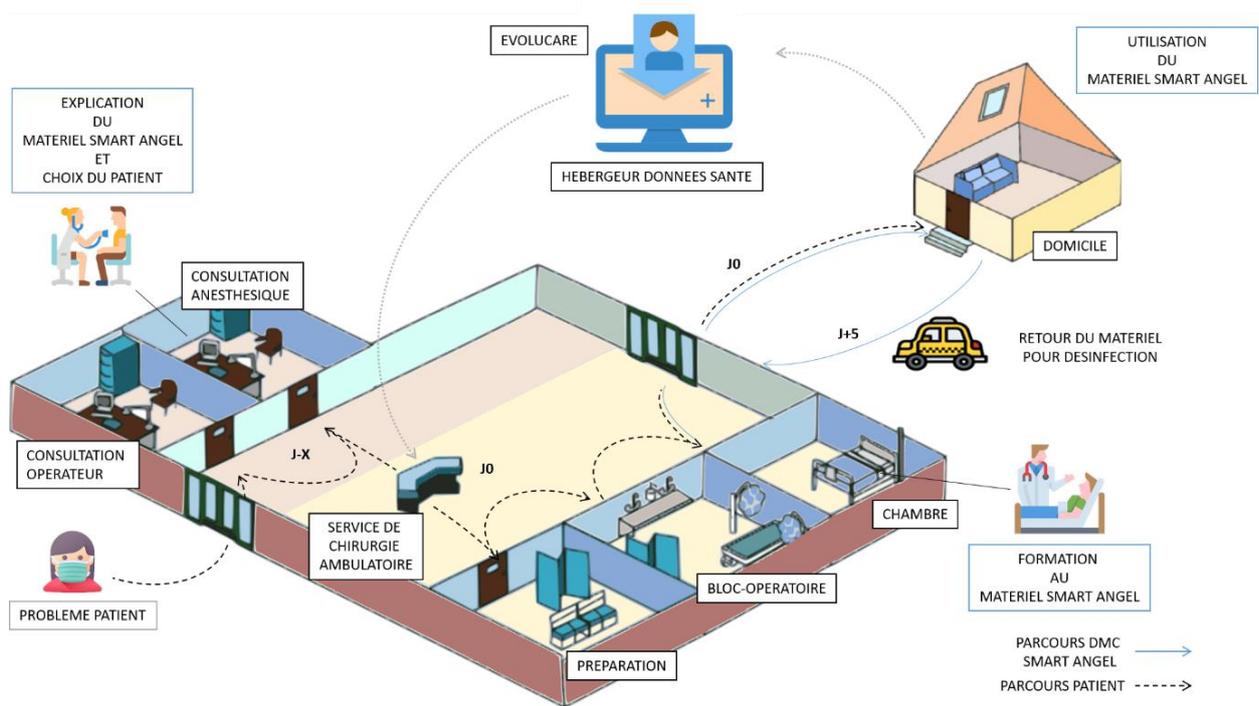


Figure 54. Représentation schématique de l'implantation du dispositif *Smart Angel* au sein de l'hôpital et du domicile patient. Le circuit en pointillés noirs correspond au parcours patient et le circuit bleu correspond au parcours du dispositif *Smart Angel*.

4.2.2. La mise en situation : le simUsanté

Le simUsanté est un hôpital simulé implanté dans le CHU d'Amiens, comprenant un ensemble de salles hautement réelles habituellement utilisées dans la formation du personnel hospitalier. Nous avons recréé une partie du parcours de chirurgie ambulatoire en nous concentrant sur des événements clés de l'implantation du dispositif *Smart Angel* à l'hôpital que nous avons pu identifier dans l'étude exploratoire ayant pour objectif l'analyse des besoins et

la compréhension du terrain (cf. Partie II, Chapitre 1, 5.2. Événements pour l'intégration du dispositif *Smart Angel* au domicile patient, p. 126). Ces événements clés sont la consultation anesthésique, la formation en chambre et le domicile du patient. Nous détaillons chacun de ces trois événements à travers la construction de scénarios dans les parties suivantes.

a. La consultation anesthésique : le scénario anesthésiste

Une fois l'acte en ambulatoire programmé et approuvé par le médecin et le patient, le matériel est présenté oralement au patient pour la première fois lors de la consultation anesthésique. C'est l'anesthésiste qui décide de le proposer au patient en fonction de critères cognitifs et d'éligibilité du dispositif (cf. Annexe 2 : Critères d'inclusion et d'exclusion de la population pouvant utiliser le dispositif *Smart Angel*, p. 341). Si l'anesthésiste propose au patient et qu'il accepte d'avoir le dispositif *Smart Angel* pendant sa convalescence, il ne pourra tester le matériel qu'après sa sortie du bloc opératoire, soit en chambre. L'anesthésiste a donc un rôle crucial pour argumenter sur l'utilité du dispositif et ses fonctionnalités auprès du patient.

L'étudiant en médecine jouant le rôle d'anesthésiste devait suivre un protocole précis. Premièrement, celui-ci était formé par le personnel Evolucare à l'utilisation du matériel *Smart Angel*. Puis, il devait réaliser une consultation anesthésique en suivant le protocole SFAR en salle de consultation (Figure 55). Cette salle était équipée d'un stéthoscope, d'une balance, de tensiomètres électriques et manuels et de deux caméras accrochées aux murs. Pour cela, il suivait un questionnaire préanesthésique³⁶ utilisé classiquement en consultation afin d'évaluer et de classifier le patient en fonction des risques anesthésiques avant de procéder à une rapide auscultation (prise de tension manuelle et poids). Le questionnaire interrogeait les patients/participants sur leurs antécédents médicaux, les traitements en cours et leurs habitudes de vie. Enfin, l'anesthésiste devait discuter avec le patient/participant du dispositif *Smart Angel* et le convaincre d'accepter le matériel. L'étudiant en médecine pouvait réaliser plusieurs consultations en fonction du temps qu'il nous allouait. Le scénario anesthésiste est disponible en annexe (cf. Annexe 16 : Scénario Anesthésiste et rappel du déroulement d'une consultation pré-opératoire, p. 405).

³⁶ Ce questionnaire proposé par ASSPRO Scientifique est disponible à l'adresse suivante :

https://www.google.com/url?sa=t&rct=j&q=&esrc=s&source=web&cd=&ved=2ahUKEwibubezzf_qAhX1BGM BHbDpBN0QFjABegQIBBAB&url=http%3A%2F%2Fwww.assproscientifique.fr%2Fmedias%2Fuploads%2Fquestionnaire%2Fquestionnaire-pre-anesthesie.docx&usg=AOvVaw0ueizSnL5RchtS9EYIv9hc

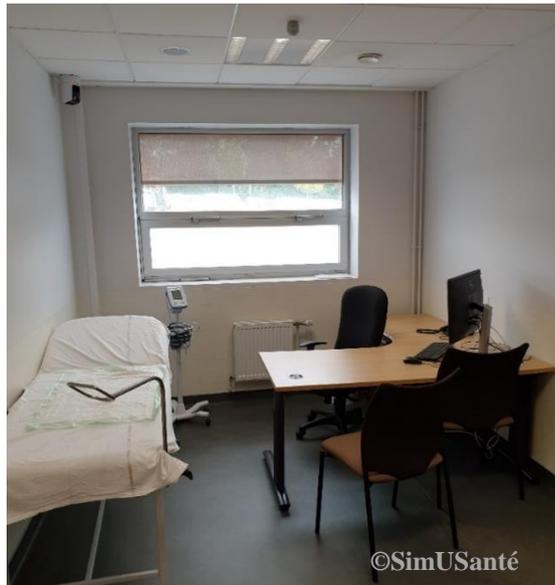


Figure 55. Photographie de la salle de consultation anesthésique © SimUSanté.

b. Les programmes d'éducation à travers des formations (« actif » et « passif »)
: le scénario infirmier

Les étudiants en médecine jouant le rôle d'infirmier forment le patient dans la chambre d'hôpital au dispositif *Smart Angel*. Avant cela, ils suivent préalablement une formation par le personnel Evolucare à l'utilisation du matériel *Smart Angel* puis à la manière dont ils doivent eux-mêmes former les patients/participants. Toutes les manipulations doivent se faire sur le patient/participant. Deux types de formations étaient proposés aux étudiants en médecine jouant le rôle des infirmiers : la formation « active » où l'infirmier informe sur toutes les erreurs à ne pas faire même si le patient/participant manipule correctement (e.g., « *tu vois là, il n'aurait surtout pas fallu que tu le mettes comme ça ou comme ça* ») ; la formation « passive » où l'infirmier visualise la manière dont le patient/participant manipule le matériel et n'intervient qu'en cas d'erreur.

Le type de formation « actif/passif » est assigné de manière randomisée. Une fois l'ensemble des informations données, l'étudiant en médecine jouant le rôle de l'infirmier se rend en chambre d'hôpital (Figure 56) pour former à son tour le patient/participant. La chambre d'hôpital est munie de deux caméras accrochées aux murs. Le scénario infirmier est disponible en annexe (cf. Annexe 17 : Scénario Infirmier formateur du dispositif *Smart Angel*, p. 407).



Figure 56. Photographie de la chambre hospitalière avec lit médicalisé © SimUSanté.

c. Le parcours patient : Scénario patient

À l'arrivée du patient/participant, un scénario lui est remis décrivant le parcours du patient (e.g., « *il y a trois semaines, vous vous plaigniez de douleur dans l'abdomen, vous vous êtes donc rendu chez votre médecin généraliste qui vous a envoyé faire des examens complémentaires...* ») et un faux problème médical lui est attribué, une hernie inguinale (un des actes les plus pratiqués en chirurgie ambulatoire ; cf. Partie II, Chapitre 1, 2.1.3. Les actes chirurgicaux les plus pratiqués en ambulatoire, p. 116). Le patient attend dans une salle d'attente l'arrivée de l'anesthésiste pour la réalisation de la consultation anesthésique, puis les expérimentateurs préparent le patient pour la chambre hospitalière. Le patient/participant est alors installé préalablement dans le lit médicalisé revêtu d'une blouse hospitalière. Une fois installé, un infirmier rentre dans la chambre pour former le patient au dispositif *Smart Angel*. Une fois la formation terminée, les examinateurs accompagnent le patient/participant au simUlogis (à l'étage supérieur) avec le matériel *Smart Angel*. Le participant peut s'installer dans la chambre ou dans le salon simulés (Figure 57) pour manipuler le dispositif. Deux caméras étaient accrochées aux murs dans chaque salle. Afin de respecter l'anonymat, les participants étaient tous renommés, Monsieur ou Madame Dupont. Le scénario patient est disponible en annexe (cf. Annexe 18 : Scénario Patient/Participant, p. 408)



Figure 57. Photographie du salon du SimUlogis © SimUSanté.

4.2.3. Questionnaires

a. Mesures du niveau de littératie

Deux questionnaires ont été utilisés pour mesurer le niveau de littératie en santé : le Newest Vital Sign (NVS) et le Health Literacy Suvey (HLS-EU-Q16) :

- **Newest Vital Sign (NVS)** (Weiss et al., 2005) traduit en français par Mansfield et al. (2018) est une épreuve validée évaluant la capacité des patients à comprendre la lecture et ainsi qu'à manipuler des chiffres (numératie). Les participants doivent se servir d'une étiquette nutritionnelle de crème glacée afin de répondre à six questions (*e.g.*, *Si je suis allergique aux arachides, puis-je manger cette crème glacée ? réponse : non, car la glace contient des traces d'huile d'arachides*). Le NVS permet d'identifier les patients à faible niveau de littératie en santé en les classant par niveau : entre 0 et 1 point, niveau « faible », entre 2 et 3 points, niveau « limité », entre 4 et 6 points, niveau « correct » de littératie en santé (cf. Annexe 12 : Questionnaire de littératie en santé – le Newest Vital Sign, p. 399). La fiabilité inter-item du NVS dans cette étude est bonne (Nunnaly, 1978): le α de Cronbach est de 0,705 ;
- **Health Literacy Survey – Europe (HLS-EU-Q16)** (Sørensen et al., 2013) traduit en français par Rouquette et al. (2018), est la version courte du questionnaire Health Literacy Survey développé par Sørensen et al. (2012). Cette version est composée de 16 items dont 13 évaluent les quatre types de compétences de la littératie en santé : la capacité d'accéder, de comprendre, d'évaluer et d'appliquer l'information sur la santé. Le répondant doit évaluer sa propre capacité à accéder à l'information (*e.g.*, « Indiquez, sur une échelle de très

facile à très difficile, dans quelle mesure il est facile pour vous de comprendre les consignes de votre médecin ou de votre pharmacien sur la manière de prendre vos médicaments ?»). Quatre catégories de réponses sont proposées sur une échelle de Likert à quatre points allant (de « *très facile* » à « *très difficile* »). Pour calculer le score total, les réponses « *facile* » et « *très facile* » ont un point par item, les réponses « *difficile* » et « *très difficile* » ne gagnent pas de point. La somme totale des items (de 0 à 16 points) permet de classer les répondants en trois catégories : (0 à 8 points) littératie en santé faible (9 à 12 points) limitée (13 à 16 points) correcte (cf. Annexe 13 : Questionnaire de littératie en santé – Le HLS-EU-Q16, p. 400). La fiabilité inter-item du HSL-EU dans cette étude est bonne (Nunnally, 1978) : le α de Cronbach est de 0,833.

b. Mesures sociodémographiques

Ce questionnaire comprenait des détails personnels : âge, sexe, niveau d'éducation, zone d'habitat, technophilie (expérience TI et dispositifs médicaux), expérience de l'hôpital (cf. Annexe 14 : Questionnaire caractéristiques sociodémographiques études 2 et 3, p. 401).

4.2.4. Entretiens semi-directifs et focus groups

Nous avons pu interroger les 28 patients/participants et 6 étudiants en médecine. Les entretiens semi-directifs et focus groups avaient pour objectif d'étudier l'utilité perçue et la facilité d'utilisation perçue du dispositif. Pour cela, lors de l'entretien avec le patient/participant et/ou l'/les étudiant(s) en médecine, nous suivions un guide d'entretien composé de cinq thématiques : la perception du dispositif ; l'acceptation du dispositif ; la formation ; la mise en situation ; les recommandations de conception. Ce guide d'entretien est disponible en annexe (cf. Annexe 19 : Guide d'entretien et focus group pour l'étude 3, p. 409).

4.3. Traitement des données

4.3.1. Mesure objective de l'utilisabilité

a. Mesure de l'efficacité

L'*efficacité* a été mesurée en comptabilisant le nombre d'erreurs de manipulation. Treize erreurs de manipulation différentes ont été identifiées en ce qui concerne l'utilisation du tensiomètre ; cinq erreurs relatives à l'usage de l'oxymètre de pouls ont été repérées : et deux types d'erreurs concernant la tablette ont été observés. Nous présentons en annexe (cf. Annexe 17 : Grille de notation des erreurs de manipulation du dispositif, p. 410) le détail de ces erreurs sur une grille de codage intégrant le type de l'erreur (erreurs bloquantes *versus* non-bloquantes)

en fonction de son impact sur la variation de la mesure (faible, modérée, importante). Lorsque le participant réalisait plusieurs tentatives, nous avons enregistré le nombre cumulatif d'erreurs.

b. Mesures de l'efficience

La mesure de l'*efficience* est basée sur les temps de manipulation des trois produits présents dans la mallette *Smart Angel* à savoir le temps de manipulation du tensiomètre, de l'oxymètre de pouls et de la tablette ainsi que le temps de manipulation total du dispositif incluant l'ensemble du « rendez-vous ». Ces temps ont été mesurés à partir du moment où les participants ont touché l'appareil (tensiomètre, oxymètre de pouls ou tablette) pour la première fois jusqu'au moment où ils l'ont éteint après avoir pris la mesure. Contrairement aux études 1b et 2 (cf. Partie II, Chapitre 3, p. 160 et Chapitre 4, p. 189), les temps de chargement du tensiomètre et de l'oxymètre lors de la prise de mesure (i.e., entre le lancement de la prise de mesure et l'affichage des données) ont été soustraits au temps d'exécution, ce qui n'est pas le cas pour la mesure du temps total.

c. Mesure de l'apprenabilité

L'*apprenabilité* a été investiguée au moyen de la comparaison des mesures d'efficacité et d'efficience entre les trois essais attendus.

4.3.2. Mesure subjective de l'utilisabilité : mesure de la satisfaction

a. Mesure quantitative : Questionnaire

La satisfaction a été quantitativement mesurée à l'aide du System Usability Survey (SUS). Ce questionnaire "*quick and dirty*" (Brooke, 1996) est composé de 10 items avec cinq options de réponse sur une échelle de Likert allant de « *pas du tout d'accord* » à « *tout à fait d'accord* » et permet l'évaluation subjective de l'utilisabilité (Lewis & Sauro, 2009). Nous avons utilisé la version réadaptée et validée décrite par Bangor et al. (2008) dans lequel nous avons remplacé le terme "*système*" par le terme "*dispositif médical*"³⁷. Les scores ont été calculés selon les recommandations de Brooke (1996) et varient de 0 à 100. Les scores les plus faibles indiquent une utilisabilité perçue faible.

b. Mesure qualitative : Entretiens semi-directifs et méthode « penser à haute voix »

Les entretiens semi-directifs ont permis de mieux cerner la dimension qualitative de la satisfaction du dispositif *Smart Angel*. Pour cela, nous avons retranscrit l'intégralité des

³⁷ Ce type de modification n'a aucune incidence sur la validité ou la fiabilité de l'instrument d'enquête (Sauro, 2011).

discours issus de la méthode « penser à voix haute » (*think aloud*) et de l'entretien semi-directif. Nous avons combiné ces deux types de données qualitatives et les avons analysés par thématique (Braun & Clarke, 2006). Braun et Clarke (2006, p.6) définissent l'analyse thématique comme “*a method for identifying, analysing, and reporting patterns (themes) within data*”. Nous avons créé notre propre arborescence en nous inspirant du modèle Health-ITUEM (Yen, 2010).

4.4. Procédure

La simulation durait en moyenne 2 heures pour les patients/participants. Elle était structurée en quatre phases (d'environ 30 min chacune ; présentées dans la Figure 58) liées à quatre lieux différents : (1) la consultation anesthésique (2) la chambre hospitalière (3) le SimUlogis (domicile simulé) et (4) l'entretien. À l'arrivée de chaque participant (patients ou étudiants en médecine), un scénario était remis et attribué en fonction de leur rôle attribué. Le personnel hospitalier se rendait d'abord en formation avec un membre d'Evolucare.

- (1) Les patients/participants étaient appelés par un anesthésiste pour leur consultation. À la fin de la consultation, l'anesthésiste présentait le dispositif *Smart Angel* sans le montrer au patient/participant ;
- (2) Les patients/participants étaient ensuite conduits en chambre hospitalière où ils étaient divisés en deux groupes randomisés : ceux recevant une formation « active » et ceux recevant une formation « passive » par un(e) infirmier/ère. Une mallette *Smart Angel* était alors remise au patient/participant jusqu'à la fin de la passation. Les étudiants en médecine (anesthésiste et infirmier) pouvaient réaliser jusqu'à quatre consultations et formations successives ;
- (3) Les patients/participants sont ensuite été conduits au SimUlogis, un domicile simulé où ils pouvaient s'installer dans un salon ou dans une chambre pour réaliser leurs manipulations au calme et seuls avec le dispositif *Smart Angel*. Les patients/participants répondaient aux questionnaires entre chaque essai. Trois essais étaient attendus : (i) après le premier essai, le participant devait répondre au questionnaire sociodémographique et au premier questionnaire de littératie en santé (*Health Literacy Survey – Europe, HLS-EU-Q16*), (ii) après le deuxième essai, le participant devait répondre au deuxième questionnaire de littératie en santé (*Newest Vital Sign, NVS*), (iii) après le dernier essai, le participant était invité à répondre au questionnaire de satisfaction (*System Usability Survey, SUS*). Lors d'un

essai de manipulation du dispositif *Smart Angel*, la procédure était toujours la même, le participant devait d'abord réaliser une mesure de tension, suivie d'une mesure de la saturation en oxygène et enfin répondre au questionnaire de santé générale dans l'application sur tablette. Il n'avait aucune limite de temps. L'expérimentateur ne pouvait intervenir qu'en cas de problème technique (e.g., problème de batterie) ;

- (4) Une fois les trois essais de manipulation réalisés, le patient/participant et les étudiants en médecine (s'ils acceptaient) ayant travaillé avec lui se rendaient avec l'examineur à un atelier de type « focus group ».

L'ensemble de ces phases a été filmé et visionné en direct sous deux angles différents par les expérimentateurs pour s'assurer du bon déroulement de chaque phase. Un minimum de trois examinateurs était nécessaire pour la réussite des passations.

4.5. Analyse de données

Les vidéos ont été analysées à l'aide de BORIS (Behavioral Observation Research Interactive Software, <http://www.boris.unito.it/>) et ont permis d'obtenir des données sur l'efficacité et l'efficacité.

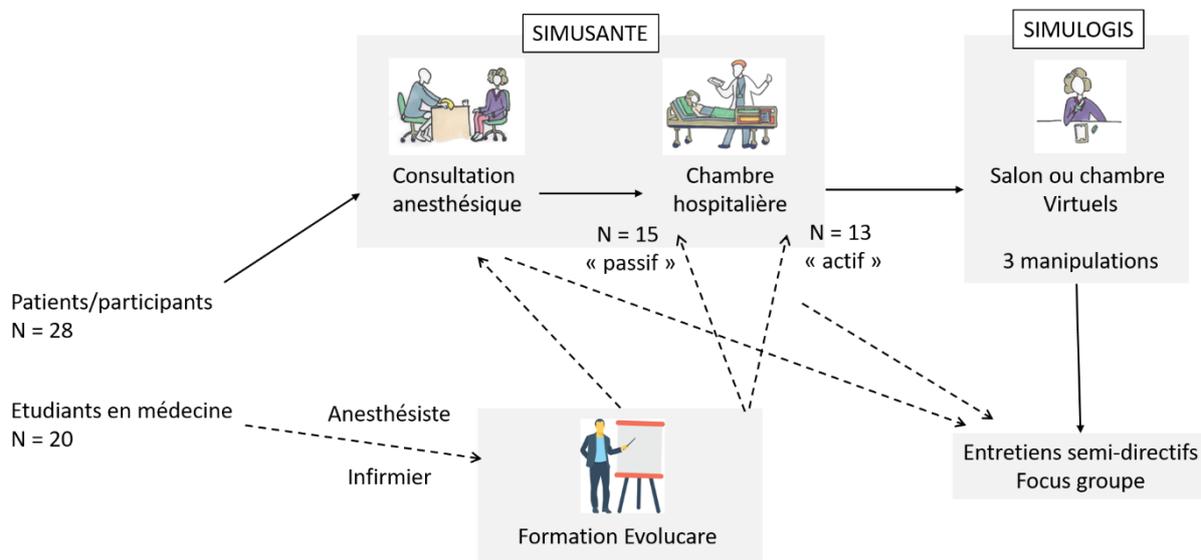


Figure 58. Représentation schématique de la procédure de l'étude au SimUsanté.

Les données quantitatives ont été analysées à l'aide de SPSS® version 22 (IBM Corporation, 2013). Chaque caractéristique utilisateur a été comparée systématiquement aux composants de l'utilisabilité comprenant l'*efficacité*, l'*efficience*, l'*apprenabilité* et la *satisfaction*. Des corrélations bivariées, des ANOVAs et tests t de Student ont été réalisés lorsque l'échantillon répondait aux critères d'homoscédasticité, tandis que des tests non paramétriques (Kruskal-Wallis et Mann-Whitney) ont été opérés lorsque l'échantillon ne répondait pas à ces critères. La taille de l'effet est mesurée selon les principes de Cohen (1988).

Les données qualitatives ont été analysées à l'aide de NVivo 12 (QSR International Pty Ltd, 2019) en suivant les étapes 6 de l'analyse thématique (Braun & Clarke, 2006) : familiarisation avec les données, codification, thématiques (arborescence), révision des thèmes, définition des thèmes, contextualisation avec la littérature scientifique.

4.6. Fiabilité inter-juge : mesures objectives d'efficacité et d'efficience

Nous avons utilisé l'Intra-Class Corrélation (ICC) afin de vérifier la fiabilité inter-juge pour les données quantitatives (Elie & Colombet, 2011; Shrout & Fleiss, 1979). Un double codage à 33 % des données vidéos collectées a été réalisé. La mesure moyenne ICC pour le temps de manipulation total (*efficience*) pour les trois manipulations est de .996 avec 95 % d'intervalle de confiance compris entre 0,991 et 0,998 ($F(29,29) = 227,312, p < ,001$). La mesure moyenne ICC pour le temps de manipulation (*efficience*) pour les trois manipulations du tensiomètre est de 0,978 avec 95 % d'intervalle de confiance compris entre 0,953 et 0,989 ($F(29,29) = 44,857, p < ,001$). La mesure moyenne ICC pour le temps de manipulation (*efficience*) pour les trois manipulations de l'oxymètre de pouls est de 0,925 avec 95 % d'intervalle de confiance compris entre 0,861 et 0,925 ($F(29,29) = 13,42, p < ,001$). La mesure moyenne ICC pour le temps de manipulation (*efficience*) pour les trois manipulations de la tablette est de 0,993 avec 95 % d'intervalle de confiance compris entre 0,985 et 0,997 ($F(29,29) = 140,573, p < ,001$).

5. Résultats

5.1. Analyse quantitative

Nous avons dû exclure deux participants lors des mesures incluant les trois essais de manipulation. En effet, un participant n'a pas réalisé le troisième essai, un autre participant a abandonné lors du troisième essai sur l'oxymètre de pouls. Ainsi afin d'éviter un biais dans les

mesures d'*efficience*, nous avons préféré retirer ce participant. Nous avons cependant inclus les deux participants pour les mesures portant sur le tensiomètre et sur les mesures ne s'intéressant qu'à la première utilisation.

5.1.1. Effet d'apprentissage sur l'efficacité et l'efficience

Le nombre d'essais (*apprenabilité*) n'a pas eu d'effet significatif sur le nombre d'erreurs (*efficacité*) total ($F(2,75) = 1,628, p = ,203$) et plus précisément sur le nombre d'erreurs réalisé lors de la manipulation du tensiomètre ($F(2,75) = 1,093, p = ,341$), de l'oxymètre de pouls ($F(2,80) = 0,347, p = ,708$) et de la tablette ($\chi^2 = 1,893, p = ,388$). La Figure 59 présente l'évolution du nombre d'erreurs en fonction des trois essais demandés. Nous proposons un détail des erreurs réalisées en annexe (cf. Annexe 21 : Nombre d'erreurs de manipulation des patients/participants lors de l'étude 3 en fonction de la grille de codage des erreurs (cf. annexe 20) et des trois essais demandés, p. 413). Dans ce détail, il est possible d'observer, de manière descriptive, que cela dépend des erreurs de manipulation. Par exemple, on observe que les patients/participants ayant positionné le brassard du tensiomètre sur le mauvais bras, réitèrent cette erreur dans les autres essais. A l'inverse, on observe que les patients/participants corrigent leur position en posant leur coude sur un support lors des autres essais.

Le nombre d'essais (*apprenabilité*) est significativement différent sur le temps de manipulation (*efficience*) total ($\chi^2(2) = 18,969, p < ,001$) comprenant le temps de manipulation du tensiomètre ($F(2,75) = 12,11, p < ,001, \eta^2 = 0,24$), de l'oxymètre de pouls ($F(2,75) = 6,862, p = ,002, \eta^2 = 0,15$) et de la tablette ($\chi^2(2) = 29,012, p < ,001$). La Figure 60 présente l'évolution du temps moyen de manipulation du tensiomètre, de l'oxymètre de pouls et de la tablette en fonction des trois essais demandés.

5.1.2. Effet des programmes d'éducation « actif/passif »

a. Validation de la randomisation

Afin de vérifier la bonne répartition issue de la randomisation des patients/participants dans les groupes, nous avons vérifié le niveau de littératie en santé des deux groupes. Nous avons pu observer que la répartition des groupes n'a pas d'effet sur le niveau de littératie en santé lié au score HLS ($F(1,26) = 0,059, p = ,811$). En revanche, la randomisation semble être très proche d'un effet sur le niveau de littératie en santé lié au score de NVS ($F(1,26) = 3,96, p = ,057$).

b. Type de formation « actif/passif »

Système : Tensiomètre, oxymètre de pouls et application Smart Angel

Le type de formation (actif/passif) n'a pas d'effet significatif lors du premier essai sur le nombre d'erreurs (*efficacité* : U de Mann-Whiney = 77, $p = ,363$), le temps de manipulation (*efficacité* : U de Mann-Whiney = 55, $p = ,052$) et sur le score SUS (*satisfaction* : $F(1,26) = 0,9$, $p = ,767$) et sur les trois essais moyennés sur le nombre d'erreurs (*efficacité* : $F(1,24) = 0,692$, $p = ,414$) et sur le temps de manipulation (*efficacité* : $F(1, 24) = 0,045$, $p = ,834$).

Tensiomètre

Le type de formation (actif/passif) n'a pas d'effet significatif lors du premier essai sur le nombre d'erreurs (*efficacité* : U de Mann-Whiney = 83, $p = ,525$) et le temps de manipulation du tensiomètre (*efficacité - temps de manipulation* : $F(1,26) = 1,603$, $p = ,217$; *temps d'exécution* : $F(1,26) = 2,694$, $p = ,113$; *temps d'instruction* : $F(1,26) = 0,081$, $p = ,778$).

Oxymètre de pouls

Le type de formation (actif/passif) n'a pas d'effet significatif lors du premier essai sur le nombre d'erreurs (*efficacité* : $F(1,26) = 0,124$, $p = ,728$) et le temps de manipulation de l'oxymètre de pouls (*efficacité - temps de manipulation* : $F(1,26) = 0,535$, $p = ,471$; *temps d'exécution* : $F(1,26) = 0,366$, $p = ,55$; *temps d'instruction* : $F(1,26) = 1,869$, $p = ,183$).

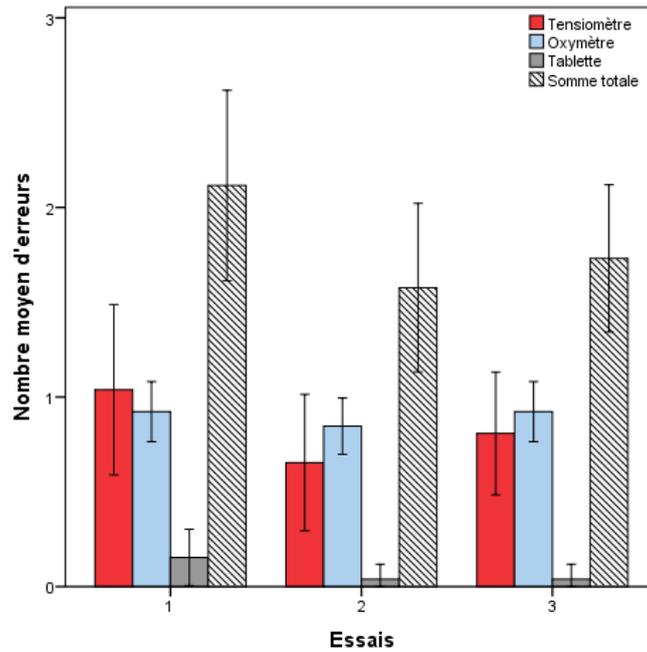
Tablette

Le type de formation (actif/passif) n'a pas d'effet significatif lors du premier essai sur le nombre d'erreurs (*efficacité* : U de Mann-Whiney = 83, $p = ,525$) et le temps de manipulation de la tablette (*efficacité* : U de Mann-Whiney = 86, $p = ,618$).

c. L'âge

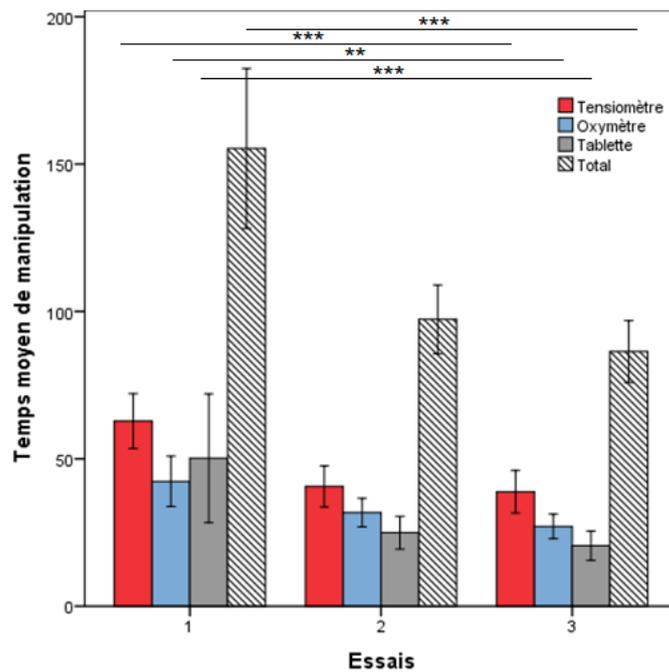
Système : Tensiomètre, oxymètre de pouls et application Smart Angel

L'âge ($M = 30,79$, $ET = 12,64$, rang = 19 – 64 ans) n'est pas corrélé avec le nombre d'erreurs de manipulation (*efficacité* : $r = 0,327$, $p = ,09$) et avec le score au SUS (*satisfaction* : $r = 0,047$, $p = ,812$), mais est significativement corrélé (positivement et fortement) avec le temps de manipulation total (*efficacité* : $r = 0,589$, $p = ,001$) ce qui signifie que les personnes jeunes manipulent plus rapidement le dispositif que les personnes plus âgées. Nous pouvons observer que l'âge est corrélé significativement et modérément avec le score au HLS-EU-16 ($r = 0,419$, $p = ,027$), mais pas avec le score de NVS ($r = -0,302$, $p = ,118$).



Bâtons de variation : 95% IC

Figure 59. Graphique en barre du nombre moyen d'erreurs réalisées sur le tensiomètre, l'oxymètre de pouls et la tablette en fonction des trois essais demandés (N = 26).



Bâtons de variation : 95% IC

** . Les corrélations sont significatives à ,01.
 ***. Les corrélations sont significatives à ,001.

Figure 60. Temps moyen de manipulation du tensiomètre, de l'oxymètre de pouls et de la tablette en fonction des trois essais demandés (N = 26).

Tensiomètre

L'âge n'est pas corrélé avec le nombre d'erreurs (*efficacité* : $r = 0,234$, $p = ,23$) et le temps de manipulation du tensiomètre (*efficacité – temps de manipulation total* : $r = 0,319$, $p = ,098$; *temps d'exécution* : $r = 0,11$, $p = ,576$), mais l'est avec le temps de lecture des instructions (*temps de consultation* : $r = 0,447$, $p = ,017$).

Oxymètre de pouls

L'âge n'est pas corrélé avec le nombre d'erreurs (*efficacité* : $r = -0,096$, $p = ,627$) et le temps de manipulation de l'oxymètre de pouls (*efficacité – temps de manipulation total* : $r = 0,338$, $p = ,078$, *temps de consultation* : $r = 0,216$, $p = ,269$ et *temps d'exécution* : $r = 0,127$, $p = ,52$).

Tablette

L'âge est significativement corrélé (positivement et moyennement) avec le nombre d'erreurs (*efficacité* : $r = 0,431$, $p = ,022$) et avec le temps de manipulation de la tablette (*efficacité* : $r = 0,472$, $p = ,011$).

d. La littératie en santé : Health literacy Scale (HLS-EU-16)

Système : Tensiomètre, oxymètre de pouls et application Smart Angel

Le niveau de littératie en santé (score au HLS-EU-16) n'est pas corrélé avec le nombre d'erreurs (*efficacité* : $r = 0,239$, $p = ,22$), le temps de manipulation total (*efficacité* : $r = 0,346$, $p = ,072$), et le score au SUS (*satisfaction* : $r = -0,033$, $p = ,866$)

Tensiomètre

Le niveau de littératie en santé (score au HLS-EU-16) n'est pas corrélé avec le nombre d'erreurs (*efficacité* : $r = 0,126$, $p = ,524$) et le temps de manipulation du tensiomètre (*efficacité – temps de manipulation total* : $r = 0,222$, $p = ,256$, *temps d'exécution* : $r = 0,184$, $p = ,347$, *temps d'instruction* : $r = 0,18$, $p = ,359$).

Oxymètre de pouls

Le niveau de littératie en santé (score au HLS-EU-16) n'est pas corrélé avec le nombre d'erreurs (*efficacité* : $r = -0,072$, $p = ,715$) et le temps de manipulation de l'oxymètre de pouls (*efficacité – temps de manipulation total* : $r = 0,156$, $p = ,429$, *temps d'exécution* : $r = 0,064$, $p = ,746$, *temps d'instruction* : $r = -0,076$, $p = ,703$).

5.1.3. Effet des caractéristiques utilisateurs

Tableau 26. Analyse descriptive des caractéristiques des utilisateurs (N = 28), leur expérience en santé, dans les dispositifs médicaux et leur niveau de technophilie et de littératie en santé (score du HLS-EU-16 et du NVS), lors du premier essai.

<i>Variable (N=28)</i>		<i>n(%)</i>	<i>Efficacité Moy (ET) (1er essai)</i>	<i>Effizienz Moy (ET) (1er essai)</i>	<i>Satisfaction Moy (ET)</i>
<i>Caractéristiques</i>					
<i>Âge</i>	30,79 ans (12,64)	28 (100)			
<i>Genre</i>	Homme	11 (39,3)	2,27 (1,35)	173,63 (95,53)	87,73 (11,77)
	Femme	17 (60,7)	2 (1,32)	159,97 (77,9)	87,53 (10,93)
<i>Niveau d'éducation</i>	Secondaire (seconde, 1re Terminale), Baccalauréat ou Brevet professionnel	3 (10,71)	3 (2,65)	190,47 (21,59)	83 (18,25)
	Supérieur 1er cycle (Bac +1 ou 2)	12 (42,86)	2,08 (1,1)	147,48 (78,49)	91,42 (7,73)
	Supérieur 2e cycle (Bac +3 ou 4)	5 (17,86)	1,6 (1,52)	172,92 (83)	87 (5,66)
	Supérieur 3e cycle (Bac +5 et plus)	8 (28,57)	2,13 (,99)	177,97 (110,88)	84 (14,66)
<i>Zones</i>	Rural	4 (14,29)	12,25 (1,26)	125,7 (27,95)	96,5 (1,73)
	Semi-urbain	4 (14,29)	3,5 (,58)	140,68 (29,98)	91 (4,24)
	Urbain	20 (71,43)	1,8 (1,28)	178,2 (95,29)	85,15 (12,032)
<i>Groupe formations</i>	Actif	13 (46,4)	2,38 (1,56)	198,83 (105,59)	86,92 (11,34)
	Passif	15 (53,6)	1,87 (1,06)	136,32 (45,21)	88,2 (11,16)
<i>Expérience en santé</i>					
<i>Opération(s)</i>	Oui	23 (82,14)	2,04 (1,36)	157,83	88,7
	Non	4 (14,29)	2,5 (1,29)	(69,98) 204,99	(9,54) 80,75
	Je ne sais pas	1	2	(156,01) 196,71	(18,94) 90
<i>Type d'hospitali- sation</i>	Avec hospitalisation	11 (39,29)	2,36 (1,5)	179,78 (81,54)	86,36 (11,9)
	Ambulatoire (sans	4 (14,29)	1,25 (0,5)	97,9	95,25

	hospitalisation)			(7,06)	(3,78)
	Les deux	7 (25)	2,29 (1,25)	147,69 (54,82)	87,71 (6,73)
	Je ne sais pas	2 (7,14)	1 (1,41)	203,31 (9,34)	92,5 (3,54)
<i>Complication chirurgicale</i>	Oui	2 (7,14)	2,12 (1,33)	113,27 (14,03)	83 (21,21)
	Non	24 (84,71)	1,5 (0,71)	159,72 (68,5)	89,75 (8,18)
<i>Souffrant d'une maladie chronique</i>	Oui	10 (35,71)	1,8 (1,32)	146,29 (51,43)	89,20 (5,41)
	Non	18 (64,29)	2,28 (1,32)	175,93 (97,06)	86,72 (13,27)
Expérience dans les dispositifs médicaux					
<i>Prise de tension</i>	Oui	23 (82,14)	2,6 (1,34)	195,99 (88,7)	89 (6,48)
	Non	5 (17,86)	2 (1,31)	158,68 (83,23)	87,3 (11,91)
<i>Prise d'oxygénation du sang</i>	Oui	1 (3,57)	4	107,44	95
	Non	27 (96,43)	2,04 (1,28)	167,49 (84,61)	87,23 (11,17)
Expérience dans les technologies de l'information (TI)					
<i>Technophilie</i>	Faible	3 (10,71)	3,33 (1,15)	308,41 (135,62)	79 (20,81)
	Medium	15 (53,6)	2 (1,46)	160,63 (71,7)	88,87 (9,58)
	Fort	10 (35,71)	1,9 (,99)	129,49 (26,36)	88,3 (9,99)
Littératie en santé – HLS-EU-16 et NVS					
<i>HLS-EU-16 (Health literacy scale)</i>	Faible	6 (21,42)	2 (1,09)	132,24 (21,9)	86,33 (7,26)
	Limité	8 (28,57)	1,38 (0,92)	140,29 (34,34,6)	88,88 (6,29)
	Correct	14 (50)	2,57 (1,45)	193,84 (109,54)	87,43 (14,51)
<i>NVS (Newest Vital Sign)</i>	Faible	4 (14,29)	4 (0,82)	219,64 (90,1)	86 (16,25)
	Limité	6 (21,42)	2,33 (1,63)	170 (132,62)	85,83 (16,36)
	Correct	18 (64,29)	1,61 (0,85)	151,72 (60,41)	88,56 (8,13)

Tablette

Le niveau de littératie en santé (score au HLS-EU-16) est corrélé significativement (positivement et moyennement) au nombre d'erreurs (*efficacité* : $r = 0,416$, $p = ,027$), mais il n'est pas corrélé avec le temps de manipulation de la tablette (*efficacité* : $r = 0,258$, $p = ,185$).

e. La littératie en santé : Newest Vital Sign (NVS)

Système : Tensiomètre, oxymètre de pouls et application Smart Angel

Le niveau de littératie (score au NVS) est significativement corrélé (négativement et fortement) avec le nombre d'erreurs (*efficacité* : $r = -0,593$, $p = ,001$) et avec le temps de manipulation total (*efficacité* : $r = -0,495$, $p = ,007$), mais il ne l'est pas avec le score au SUS (*satisfaction* : $r = 0,061$, $p = ,757$).

Tensiomètre

Le niveau de littératie (score au NVS) est significativement corrélé (négativement) avec le nombre d'erreurs (*efficacité* : $r = -0,584$, $p = ,001$) et le temps de manipulation du tensiomètre (*efficacité – temps de manipulation total* $r = -0,381$, $p = ,046$), mais il n'est pas corrélé avec le temps d'exécution et le temps d'instruction (*efficacité : temps d'exécution* : $r = -0,294$, $p = ,128$; *temps d'instruction* : $r = -0,345$, $p = ,072$).

Oxymètre de pouls

Le niveau de littératie (score au NVS) n'est pas corrélé avec le nombre d'erreurs (*efficacité* : $r = -0,197$, $p = ,315$) et le temps de manipulation de l'oxymètre de pouls (*efficacité : temps de manipulation total* : $r = -0,165$, $p = ,4$; *temps d'exécution* : $r = 0,081$, $p = ,682$, *le temps d'instruction* : $r = -0,191$, $p = ,33$).

Tablette

Le niveau de littératie en santé (score au NVS) n'est pas corrélé avec le nombre d'erreurs (*efficacité* : $r = -0,123$, $p = ,532$) et le temps de manipulation de la tablette (*efficacité* : $r = -0,25$, $p = ,185$).

f. La technophilie

Système : Tensiomètre, oxymètre de pouls et application Smart Angel

Le niveau de technophilie est significativement corrélé (négativement et modérément) au temps de manipulation du total (*efficacité* : $r = -0,427$, $p = ,024$). Le niveau de technophilie n'est pas corrélé au nombre d'erreurs (*efficacité* : $r = -0,115$, $p = ,561$) et au score au SUS (*satisfaction* : $r = 0,149$, $p = ,448$).

Tensiomètre

Le niveau de technophilie n'est pas corrélé avec le nombre d'erreurs (*efficacité* : $r = -0,112$, $p = ,572$) et le temps de manipulation du tensiomètre (*efficacité – temps de manipulation* : $r = -0,243$, $p = ,213$; *temps d'exécution* : $r = -0,121$, $p = ,539$; *temps d'instruction* : $r = -0,3$, $p = ,121$).

Oxymètre de pouls

Le niveau de technophilie est significativement corrélé (positivement et faiblement) au nombre d'erreurs (*efficacité* : $r = ,376$, $p = ,049$), mais n'est pas corrélé avec le temps de manipulation de l'oxymètre de pouls (*efficacité - temps de manipulation* : $r = -0,301$, $p = ,12$; *temps d'exécution* : $r = -0,229$, $p = ,24$; *temps d'instruction* : $r = -0,305$, $p = ,115$).

Tablette

Le niveau de technophilie est quasiment corrélé significativement avec le nombre d'erreurs (*efficacité* : $r = -0,373$, $p = ,051$) et corrélé significativement avec le temps de manipulation de la tablette (*efficacité* : $r = -0,464$, $p = ,013$).

5.2. Analyse qualitative

5.2.1. Mesure de la satisfaction : analyse thématique

L'analyse thématique réalisée sur la base des retranscriptions des entretiens semi-directifs et des focus group a permis de proposer une arborescence de la satisfaction des utilisateurs à l'égard du dispositif *Smart Angel* sur la base de deux thèmes principaux : la facilité d'utilisation perçue et l'utilité perçue. La Figure 62 représente cette arborescence. Nous présentons chacun des thèmes dans les parties suivantes.

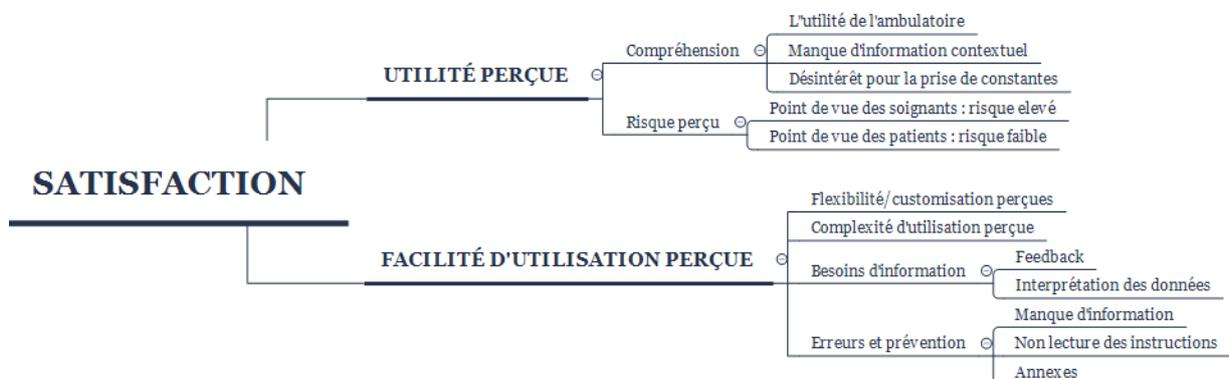


Figure 61. Synthèse des thèmes (utilité perçue et facilité d'utilisation perçue d'après le modèle de Yen, 2010) et les sous-thèmes de la satisfaction mis en évidence grâce à l'analyse qualitative.

Pour chaque thématique évoquée, nous présentons le nombre de participants ayant abordé cette thématique et/ou cette sous-thématique suivi, entre parenthèse, du nombre de fois qu'elles ont été mentionnées dans le discours (e.g., X participants ont mentionné Y fois la thématique ce qui correspond à Y références).

a. Facilité d'utilisation perçue

La facilité d'utilisation est composée de quatre sous-thématiques qui nous permettent de structurer nos résultats ci-après : la flexibilité/customisation perçue, la complexité d'utilisation perçue ; le besoin d'information ; les erreurs et la prévention.

Flexibilité/customisation perçue

Le dispositif a été jugé positivement par cinq participants (9 références) mettant en avant les avantages de flexibilité du dispositif comme la rapidité avec lequel il se manipule, sa mobilité et sa taille adaptée (« *Là c'est bien on peut le prendre avec nous c'est transportable, c'est une petite sacoche, moi je trouve que c'est bien fait* » participant 15). Deux participants (3 références) ont en revanche décrit un manque de flexibilité du matériel (« *ce qui me rendrait service, c'est une application sur mon téléphone* » participant 11).

Complexité d'utilisation perçue

Pour les 34 participants (patients et étudiants en médecine) interrogés, tous ont jugé le dispositif *Smart Angel* très facile d'utilisation même s'ils avaient réalisé une erreur pendant la passation (« *Il m'a paru très facile à utiliser, fin voilà c'est ludique c'est quand même pas très compliqué à mettre en route etc...* » participant 26 ; « *C'est clair, même sur la tablette, c'est fluide, c'est clair, ça passe pas par quatre chemins* » participant 15). Le dispositif semble, par conséquent, facile à utiliser au premier abord. Pourtant lorsque les participants étaient invités à préciser leur choix de manipulation et plus précisément d'expliquer la présence d'une erreur, une des causes majeures énoncées était l'ambiguïté des instructions ou le manque d'information. Nous présentons ces éléments dans la section « erreurs et prévention ».

Besoins d'information : feedback et interprétation des données

Deux types d'informations ont manqué d'après les participants : l'absence de feedback et d'éléments pour interprétation des données physiologiques. Deux participants (3 références) regrettent le manque de feedback du matériel. Il aurait par exemple voulu s'assurer que leurs données sont envoyées à l'hôpital (« *Tu fais ta mesure il y a rien qui dit que l'info a été envoyée à l'hôpital* » participant 27) ou avoir un moyen de s'assurer de la bonne utilisation du matériel (« *comme je le fais moi-même et qu'en plus c'est électronique je me dis je peux me tromper je*

ne suis pas médecin, si ça se trouve je le prends mal » participant 11). Sept patients/participants ont déclaré ne pas savoir interpréter les résultats obtenus (« *Ouais par contre je comprends pas ce que c'est les chiffres* » participant 12) notamment en ce qui concerne l'utilisation de l'oxymètre de pouls (« *Par contre je trouve qu'elle aurait pu m'expliquer ce que c'était l'oxygénation du sang* » participant 14).

Erreurs et prévention

Nous avons questionné les participants pour comprendre la survenue des erreurs de manipulation. Deux explications majeures (le manque d'information et la non-lecture des instructions sont mis en évidence) ainsi que quatre explications annexes ont été référencées. La Figure 62 représente en proportion les références données afin d'argumenter la survenue d'une ou plusieurs erreurs. Nous allons préciser chacun des arguments énoncés sachant que les participants pouvaient donner plusieurs explications.

L'explication la plus répertoriée était les lacunes en matière de design produit et en rapport avec les instructions données. Nous avons divisé cette explication en trois éléments : le manque d'information, le manque de feedback sur le positionnement et l'ambiguïté de certaines instructions. En effet, dix-huit participants (35 références) ont reproché au matériel d'être mal conçu. Six participants (7 références) se sont plaints de certains éléments pouvant induire en erreur comme l'imprécision des schémas (« *Après il y a un dessin de main (montre sa paume) ça pourrait être plus clair parce qu'on ne sait pas si c'est le bras droit dans le sens (montre l'autre sens que la paume de la main), Il y a des trucs comme ça plus pour le tensiomètre* » infirmier 24) et la taille de la police (« *OK ça et ça je pense qu'il faudrait le mettre en plus gros* » participant 15). Deux participants (3 références) pensent que leurs erreurs auraient pu être évitées par un meilleur feedback sur le matériel comme par exemple sur le serrage du bracelet du tensiomètre (« *Je trouve que ça c'est trop court et il n'est pas assez souple [...] Est-ce qu'il est bien serré ? Pour moi il n'a jamais été serré de la même façon. Et d'ailleurs je n'ai pas les mêmes résultats* » participant 18) et le bouton de l'oxymètre (« *oui j'ai mal appuyé dessus, mais comment savoir ?* » participant 19). Enfin, des indications auraient pu être ajoutées pour limiter les incertitudes de manipulation d'après dix participants (14 références).

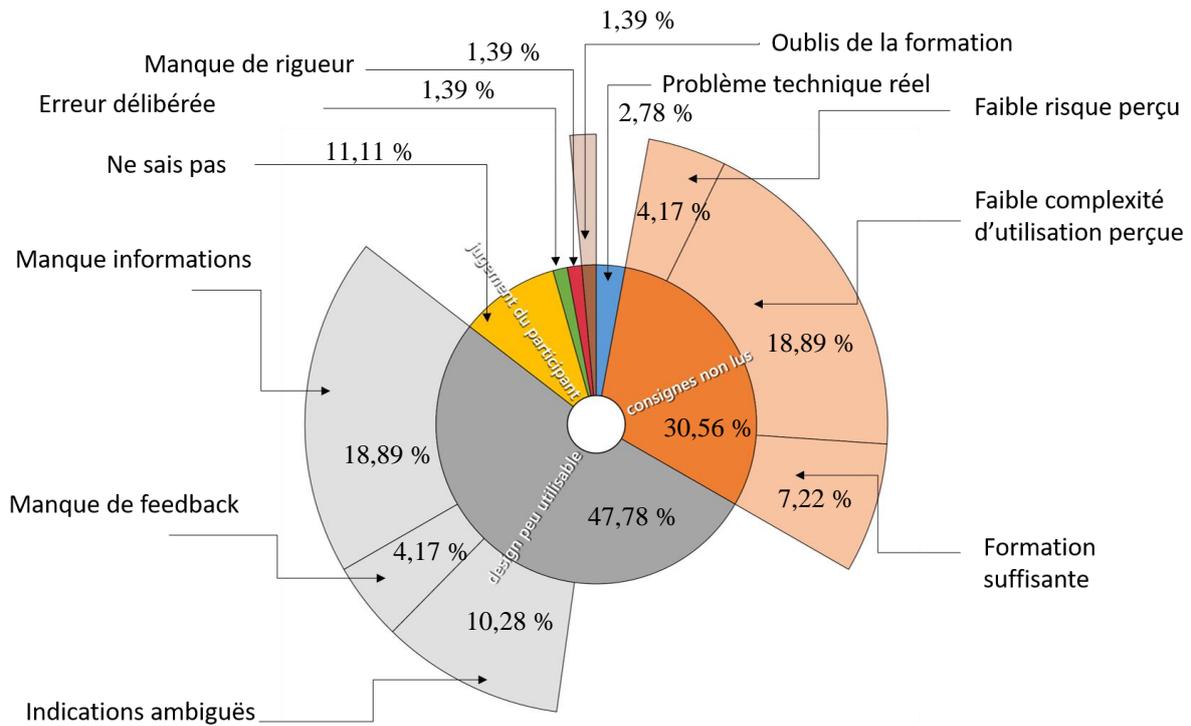


Figure 62. Représentation graphique de la proportion des références fournies par les personnes interrogées afin de justifier la survenue d'une ou plusieurs erreurs.

La deuxième explication majeure est la non-lecture des instructions. Nous avons essayé de comprendre pourquoi les participants ne lisaient pas les consignes. Nous avons pu expliquer ce constat à travers trois éléments : la faible complexité d'utilisation perçue, le faible risque perçue et le jugement de la formation. Treize participants (22 références) ont déclaré ne pas avoir lu les consignes. Onze participants (14 références) ont considéré que le dispositif était simple à utiliser et ne requéraient pas la lecture des instructions pour être utilisé. Parmi ces participants, trois ont décrit une volonté d'aller vite, un empressement dans la manipulation (« *Je me suis jeté tout de suite sur les trucs et pas sur la tablette* » participant 27). Cinq participants (5 références) ont déclaré ne pas avoir ressenti le besoin de lire les instructions, car la formation a été considérée comme suffisante pour utiliser le matériel (« *non non, j'ai pas lu, j'ai mémorisé ce que m'a dit l'infirmière, pas de soucis* » participant 14). Enfin trois participants (3 références) ont mis en évidence un faible risque perçue ce qui signifie que le dispositif ne présente pour eux aucun danger. Nous retrouvons dans cet argumentaire, un manque de sérieux dans la manipulation dévoilant le dispositif davantage comme un divertissement plutôt qu'un dispositif médical et ainsi prendre le risque de se tromper (« *je me suis dit ah j'ai envie d'essayer ce truc ça a l'air trop cool "Commencer" commencer" et hop j'ai mis le truc Mais c'est vrai que j'ai pas lu* » participant 27). Enfin trois participants ont déclaré ne pas avoir lu délibérément

les instructions (« *Donc j'ai pas vraiment lu, je savais ce qu'il fallait faire [...] au risque parfois peut être de cliquer sur le mauvais truc* » participant 5).

Enfin, plusieurs explications annexes ont pu être répertoriées. Premièrement, cinq participants (8 références) n'ont pas su détailler pourquoi ils avaient réalisé une erreur de manipulation. Dans ce cas, tous ont décrit un problème technique. Or à la lecture de la vidéo, aucun problème technique n'est survenu. Nous avons pu constater que certains participants pouvaient bouger sans s'en rendre compte pendant la manipulation ce qui pourrait expliquer leurs erreurs. Deuxièmement, les participants ont manqué de rigueur lors de la répétition des essais (« *C'est un peu, enfin l'utiliser une fois c'est sympa, on fait plus attention puis après ça devient rébarbatif, Et on fait moins attention* » participant 17). Troisièmement, un participant n'a volontairement pas respecté les instructions, ne sachant pas les conséquences de ce non-respect (« *[position du tensiomètre] au début à droite, car j'avais mon élastique ici* » participant 28). Quatrièmement, un participant a décréé ne plus se souvenir des consignes données lors de la formation et de ce fait avoir confondu certains gestes (« *C'est vrai que c'est pas compliqué, mais on peut se dire ah mince qu'est-ce qu'elle m'a dit (en parlant de l'infirmière) on se trompe un peu* » participant 1).

Enfin, pour deux participants, il y a eu un réel problème technique (e.g., l'application plante) que nous n'avons pas comptabilisé comme une erreur de manipulation.

b. Utilité perçue

L'utilité perçue a été décomposée en deux sous thématiques qui nous permettent de structurer nos résultats autour de la compréhension et du risque perçu.

Compréhension

Nous avons divisé cette section en trois éléments : l'utilité de l'ambulatoire, le manque d'information contextuelle, et le désintérêt pour la prise de constantes.

Lorsque nous avons interrogé les participants sur l'utilité du matériel et les avantages qu'ils pouvaient en tirer pour eux, nous avons pu répertorier plusieurs arguments pouvant être relativement éloignés de l'utilité réelle du matériel. En effet, trois patients/participants semblaient davantage convaincus par le principe de la chirurgie ambulatoire ou l'intérêt de limiter les déplacements que par le suivi postopératoire qu'offrait le dispositif (« *C'est l'aspect de ne pas retourner à l'hôpital qui m'intéresse* » participant 11, « *Si les résultats sont communiqués directement je pense ça évite d'avoir des allers-retours à faire pour des*

consultations » participant 3). Or, même sans dispositif, le patient n'aurait pas eu besoin de se déplacer.

Ces erreurs de jugement dans l'utilité du matériel se traduisent aussi par le manque d'information reçu par rapport aux contextes d'utilisation. Dix-sept participants (30 références) ont posé des questions supplémentaires lors de l'entretien afin de mieux comprendre le déroulé de l'utilisation du matériel. Par exemple, ils ne savaient pas toujours comment se déroulerait le rendu du dispositif, combien de « rendez-vous » les attendaient par jour et pendant combien de temps, qui appeler en cas de problème... (« *Mais la question que je ne savais pas c'était combien de temps ça durait le suivi et la question aussi c'était qu'est-ce que je fais de la tablette à la fin ?* » participant 10). Par ailleurs, six patients/participants ont posé leur question à l'infirmier et non pas à l'anesthésiste qui était pourtant là pour expliquer l'ensemble des éléments (« *J'ai pas tout à fait osé, parce que je ne savais pas encore comment ça allait concrètement, comment concrètement ça allait s'organiser [...] Mais sinon j'aurais posé beaucoup plus de questions* » participant 4). Ces participants ont par ailleurs eu des explications qu'ils n'étaient pas en mesure de comprendre (« *En fait, il [l'anesthésiste] m'a parlé de l'oxymètre, j'ai fait "oui", je savais pas ce que c'était (rires)* » participant 4). Au risque pour un participant de refuser le matériel par manque d'information contextuelle si celui-ci se retrouvait à nouveau dans cette situation (« *OK donc imaginons que demain vous vous fassiez opérer et que vous passiez devant l'anesthésiste, vous acceptez de prendre le dispositif?* » expérimentateur « *si c'est bien expliqué oui (rires)* » participant 13).

Nous avons pu constater un manque de compréhension sur l'intérêt des prises de constantes médicales. Neuf participants (12 références) nous ont rapporté ne pas savoir pourquoi il fallait prendre sa tension et son taux d'oxygène (« *Mais je pense que si je suis mal dans mon coin, prendre ma tension c'est pas la première idée qui me viendrait à l'esprit. C'est d'ailleurs ce que j'ai dit au monsieur qui est venu je lui ai dit j'ai mal (montre le ventre) pourquoi vous voulez prendre ma tension là (montre le bras). Ça n'a pas de rapport quoi* » participant 25). Lorsque nous avons précisé l'intérêt de ces prises de constantes, l'utilité du matériel s'est renforcée tout comme l'intention d'usage pour les participants (« *mais en sachant que ça surveille l'hémorragie par exemple. Je pense que je l'aurai plus utilisé* » participant 5).

Risque perçu

Nous avons observé que le risque perçu n'est pas le même pour les personnels hospitaliers que pour les patients. Les soignants ont tendance à considérer l'objet comme une annexe pour la santé et en cas de problème restent vigilants et déclarent contacter un professionnel (« *Je*

pense qu'il faut être à l'aise avec ça et bien comprendre comment ça fonctionne Et aussi à l'inverse ne pas se reposer que là-dessus Si on se sent mal il faut quand même appeler un médecin » étudiants en médecine (rôle infirmier) 24). De plus, ils semblent conscients des risques liés aux suivis postopératoires (« ça peut faire peur vis-à-vis des risques qu'il y a derrière. Mais je trouve que c'est quand même intéressant » étudiants en médecine (rôle infirmier) 4).

Le raisonnement est différent pour les patients/participants qui perçoivent un risque moindre (« *Après je ne sais pas ce qui peut se passer après une opération* » participant 26). Ce faible risque perçu est mis en évidence lorsque nous avons questionné les patients sur leur potentielle rigueur dans l'utilisation. Huit participants (9 références) ont déclaré qu'ils ne réaliseraient probablement pas tous les rendez-vous comme cela est imposé (« *Bon ben je me serais obligé puisque c'était quand même mon suivi après l'opération, Mais c'est vrai que j'aurais tendance à être négligente* » participant 1).

6. Discussion

6.1. Discussion de l'analyse quantitative

L'objectif principal de cette analyse quantitative était de mieux comprendre les relations qu'entretiennent le contexte d'usage à travers les caractéristiques utilisateurs et le programme d'éducation (« actif » et « passif ») sur l'utilisabilité (efficacité, efficience, satisfaction et apprenabilité) du dispositif *Smart Angel* évaluée par une analyse quantitative.

Pour ce faire, une simulation d'un service de chirurgie ambulatoire et du retour au domicile a été mise en place associée à la collecte de données sociodémographiques et cognitives (le niveau de littératie mesuré au moyen du Health Literacy Survey - Europe de Sørensen, et al., 2012, et du Newest Vital Sign de Weiss et al., 2005) et des mesures d'utilisabilité (erreurs et temps de manipulation, questionnaire SUS). Nous avons formulé cinq hypothèses selon lesquelles l'apprenabilité, le programme d'éducation, l'âge, la littératie en santé et la technophilie auraient un impact sur l'utilisabilité.

Notre première hypothèse stipulait que le nombre d'essais n'aurait pas d'impact sur l'efficacité et l'efficience du dispositif *Smart Angel* (e.g., Larsen, 2017). Nous pouvons partiellement valider notre hypothèse. Le nombre d'essais (*apprenabilité*) n'a pas eu d'effet significatif sur le nombre d'erreurs (*efficacité*). Le détail en annexe (cf. Annexe 21 : Nombre

d'erreurs de manipulation des patients/participants lors de l'étude 3 en fonction de la grille de codage des erreurs (cf. annexe 20) et des trois essais demandés, p. 413) montre que certaines erreurs ne sont pas corrigées systématiquement par les utilisateurs. Ces dernières se reproduisent lors d'autres manipulations impactant la sécurité du patient et la fiabilité des données envoyées. Ces résultats sont en accord avec ceux de Larsen et al. (2017). Cependant, nous avons observé que cela dépend de l'erreur. Ainsi, une évaluation plus poussée des erreurs réitérées après plusieurs essais serait nécessaire pour les détecter. En revanche, nous avons pu observer que le nombre d'essais (*apprenabilité*) améliore significativement l'efficacité du dispositif et cela pour chacun des éléments du système (tensiomètre, oxymètre de pouls, tablette). Ces résultats impliquent que les utilisateurs se familiarisent avec le matériel au fur et à mesure des essais, mais ils ne corrigent pas systématiquement leurs erreurs de manipulation, ce qui est d'autant plus dangereux pour la sécurité du patient. Ainsi, les données collectées sur l'utilisabilité lors de la première utilisation sont une source d'informations cruciale afin de prévenir les futures erreurs de manipulation.

Notre deuxième hypothèse s'intéressait au programme d'éducation au travers de formations. Nous pensions que les utilisateurs ayant reçu un programme de formation « actif » seraient plus *efficaces*, *efficaces* et *satisfaits* du dispositif *Smart Angel* que ceux ayant reçu un programme de formation « passif ». Nous ne pouvons pas valider notre hypothèse. Le type de formation n'a montré aucun effet significatif sur l'utilisabilité et sur aucun élément du système en particulier. Cela ne signifie pas pour autant que ces formations ne sont pas utiles. Au contraire, elles ont été jugées essentielles lors de l'analyse du discours. Nous pensons que deux arguments peuvent expliquer cette absence de résultat. Premièrement, l'invalidation de la randomisation montre que les groupes sont probablement inégaux sur le niveau de littératie en santé. En effet, il semblerait que le groupe passif ait un meilleur niveau (presque significatif) de littératie en santé (score au NVS) que le groupe actif. Deuxièmement, les formations ont été réalisées par des étudiants en médecine et non par un expérimentateur attiré et expert dans les passations expérimentales permettant d'apporter une rigueur dans l'application des instructions des deux formations. Cela implique que les formations reçues par les patients/participants peuvent varier au sein d'un même groupe. De plus, nous avons pu observer que la manière de présenter différait d'un étudiant en médecine à l'autre au sein du groupe « actif ». Certains poussaient davantage les explications en précisant les raisons de l'ensemble des gestes à réaliser. Par exemple, ils expliquaient où positionner le brassard du tensiomètre en fonction de

la localisation du capteur sur l'appareil qui doit se retrouver le plus proche possible des artères situées dans le poignet.

Notre troisième hypothèse stipulait que les utilisateurs âgés seraient moins efficaces, efficaces et satisfaits le dispositif *Smart Angel* que les jeunes utilisateurs (e.g., Georgsson & Staggers, 2016; Jones & Caird, 2017; Mykityshyn et al., 2002; Van der Vaart et al., 2019). Nous pouvons partiellement valider notre hypothèse. En effet, les patients/participants plus jeunes manipulent le dispositif plus rapidement (sont *plus efficaces*) que les patients/participants plus âgés sur le tensiomètre et la tablette. Précisément, nous avons pu observer que le temps de lecture des instructions du tensiomètre était plus long chez les participants âgés que chez les participants jeunes. Les patients/participants plus âgés semblent en revanche avoir davantage de difficulté (en termes d'*efficacité* et d'*efficacités*) dans la manipulation de la tablette. En effet, les patients/participants plus âgés ont significativement réalisé davantage d'erreurs et ont été significativement plus longs dans l'utilisation de cet outil. En revanche, nous n'observons pas de différences sur l'efficacité, l'efficacité et la satisfaction dans la manipulation de l'oxymètre de pouls. Ce résultat implique que le traitement de la tablette est différent des objets connectés en fonction de l'âge. Cela montre aussi l'importance d'évaluer l'utilisabilité de chaque élément qui compose un système. Les résultats relatifs aux objets connectés (tensiomètre et oxymètre de pouls) sont en désaccord avec les travaux de Jones et Caird (2017), Mykityshyn et al. (2002) et Van der Vaart et al. (2019) qui ont montré que les utilisateurs plus jeunes sont plus efficaces et efficaces. En revanche, ces résultats sont en accord avec ceux de Liang et al. (2018) car l'âge n'a pas d'impact sur la satisfaction.

Notre quatrième hypothèse s'intéressait au niveau de littératie en santé. Nous pensions que les utilisateurs ayant un faible niveau de littératie en santé seraient moins efficaces, efficaces et satisfaits du dispositif *Smart Angel* que ceux ayant un niveau élevé de littératie en santé (e.g., Jensen et al., 2010; Kaufman et al., 2003). Pour cela nous avons mesuré le niveau de littératie à l'aide de deux questionnaires : le HLS-EU-16 et le NVS. Les scores du HLS-EU-16 n'ont montré aucun effet sur l'utilisabilité à l'exception du nombre d'erreurs sur la tablette. Les participants ayant un score élevé au HLS-EU-16 ont réalisé moins d'erreurs sur la tablette que les participants ayant un faible score. À l'inverse, les scores du NVS ont mis en évidence un effet significatif (corrélation forte) sur le nombre d'erreurs de manipulation (*efficacité*), le temps de manipulation (*efficacité*) sur l'ensemble du système *Smart Angel*. Ce qui est en accord avec les résultats des études précédentes (Czaja et al., 2015 ; Kim & Xie, 2017, Jensen et al., 2010). En revanche, nous n'observons aucun effet significatif sur la satisfaction. Après analyse

des éléments séparément, nous pouvons constater que les participants ayant un score élevé au NVS réalisent significativement moins d'erreurs (i.e., sont plus efficaces) et sont plus rapides (i.e., sont plus efficaces) dans la manipulation du tensiomètre uniquement. En effet, nous n'observons aucune différence significative entre le niveau de littératie en santé et la manipulation de l'oxymètre de pouls et la tablette. Pour expliquer ces résultats, nous savons que le tensiomètre est plus complexe à manipuler que l'oxymètre de pouls. Il semblerait que la formation proposée en amont a permis d'éviter les erreurs de manipulation de l'oxymètre, mais que certaines, dues à la complexité du tensiomètre, n'ont pas pu être évitées. Les utilisateurs se sont donc basés sur leur littératie en santé afin de prendre des décisions dans la manipulation de ce produit. Enfin, cela expliquerait probablement pourquoi le niveau de littératie en santé n'a pas eu d'impact sur l'utilisation de la tablette puisqu'il ne s'agit pas d'un produit de santé.

Notre cinquième hypothèse s'intéressait à la technophilie (e.g., Georgsson & Staggers, 2016; Liang et al., 2018; Van der Vaart et al., 2019). Les résultats permettent de valider partiellement cette hypothèse. En effet, plus les patients/participants avaient un haut niveau de technophilie, plus ils manipulaient significativement rapidement (i.e., plus efficaces) le dispositif. En revanche, nous n'observons pas de différence sur l'efficacité et la satisfaction. Enfin, si nous regardons les outils de manière indépendante, nous observons que cet effet est principalement valable pour la manipulation de la tablette. En effet, les patients/participants ayant un haut niveau de technophilie ont manipulé plus rapidement la tablette que ceux avec un faible niveau de technophilie de même que pour l'efficacité même si nous n'observons qu'une tendance (non significative). En revanche, nous avons pu observer un nombre significativement plus important d'erreurs pour les patients/participants ayant un haut niveau de technophilie par rapport aux patients/participants ayant un bas niveau sur l'oxymètre de pouls. Un haut niveau de technophilie ne semble pas avoir d'impact sur l'utilisation d'objet connecté en santé, voire pourrait être un désavantage pour l'efficacité de l'oxymètre, mais un avantage pour l'utilisation d'outil informatique comme la tablette tactile.

6.2. Discussion de l'analyse qualitative

Dans un deuxième objectif, nous avons poussé davantage nos recherches sur le jugement des participants sur le dispositif à travers deux notions (l'utilité perçue et la facilité d'utilisation perçue) dans un milieu simulé. Pour cela, nous avons effectué une analyse thématique basée

sur des données qualitatives issues des discours des participants (patients et personnel hospitalier).

Même si cette analyse subjective ne repose pas sur une validation quantitative, notre sixième hypothèse semble validée. En effet, l'activité simulée a permis de mettre en exergue un certain nombre d'éléments jusqu'alors inobservables dans les études précédentes propulsant ainsi l'environnement comme prépondérant dans l'utilisabilité du dispositif *Smart Angel*. En effet, nous avons pu observer de très importantes lacunes des patients/participants à comprendre le contexte du dispositif *Smart Angel* : l'interprétation des données, l'utilisation de certains objets et l'organisation du système. Ces carences sont dues à un manque d'informations données par le personnel hospitalier renvoyant directement aux lacunes de la pédagogie hospitalière. Nous expliquons, sur la base de deux arguments, pourquoi l'environnement (simulé) a pu impacter les décisions et le jugement du patient (i.e., la satisfaction) dans l'utilisation du dispositif *Smart Angel*.

Premièrement, grâce à l'intégration de consultation anesthésique, faisant partie de l'environnement naturel du parcours patient en chirurgie ambulatoire, nous avons pu observer que l'anesthésiste n'a pas su expliquer le contexte et les enjeux du dispositif. Il semblerait que le manque d'explications de l'anesthésiste sur les enjeux du matériel impacte directement le risque perçu et l'utilité perçue du matériel. En effet, le vocabulaire utilisé par ce dernier a été jugé trop complexe par certains patients.

Deuxièmement, de la même manière que l'anesthésiste, l'infirmier/ère n'a pas su expliquer aux patients les gestes les plus importants à retenir et n'a pas apporté d'information supplémentaire pour détailler l'importance de la prise de tension et l'interprétation de ces données. D'ailleurs, les patients/participants ont davantage jugé négativement la formation « passive » liée aux manques d'interaction patient-soignant. Pourtant, cette formation sera utilisée lors de l'étude clinique devant débiter peu de temps après cette étude. Il s'agit pourtant d'une piste intéressante afin d'améliorer les relations patients-soignants. Nous avons d'ailleurs pu observer un désintérêt pour le matériel pour le patient lorsque les échanges patient/soignant étaient moindres. Il est possible que le manque d'expérience et/ou de retours d'expérience des étudiants en médecine ait abaissé le niveau des formations présentées et des explications données. Il semble pourtant important de former le personnel hospitalier à ces questions.

6.3. Limites

La pédagogie hospitalière semble avoir un réel impact sur l'utilisabilité du produit et est donc essentielle avant, pendant, et après que le patient manipule le dispositif. Cependant, trois limites peuvent être mentionnées.

La première limite réside dans le fait que nos analyses de discours se basent exclusivement sur des entretiens semi-directifs. En effet, des analyses de l'interaction entre les patients, le personnel hospitalier et les industriels pourraient être envisagées afin de mieux comprendre les échanges d'informations. Il est possible que certains éléments aient été cités par le personnel hospitalier et n'aient pas été mémorisés par les patients lors de la consultation ou de la formation. Cela pourrait remettre en cause la quantité d'informations donnée par le personnel hospitalier aux patients. Il est aussi possible que les formations de l'industriel Evolucare au personnel hospitalier aient été insuffisantes.

Les deux autres limites concernent la simulation en elle-même. Nous avons vu que ce type d'environnement possède un intérêt majeur pour proposer des évaluations de l'utilisabilité se rapprochant d'un contexte *in situ* (Kushniruk et al., 2013). Cependant, nous ne pouvons pas confirmer la « validité écologique » (e.g., Rochat et al., 2017) c'est-à-dire, le fait que les comportements observés dans un environnement simulé soient généralisables à l'environnement dans lequel un produit sera utilisé dans la "vraie vie".

De plus, outre la complexité de sa mise en place et son coût financier et humain, nous avons trouvé que les participants étaient moins critiques, créatifs et expressifs par rapport aux entretiens réalisés dans l'étude précédente (cf. Partie II, Chapitre 4, 5.2. Analyse qualitative : Satisfaction (qualitative) des utilisateurs du dispositif *Smart Angel*, p. 214). Il est possible que cette différence ne soit pas due qu'à l'amélioration du produit, mais aussi à l'environnement dans lequel se trouvent les participants. La plupart d'entre eux ont déclaré avoir été anxieux pendant la simulation liée entre autres à l'environnement hospitalier particulièrement anxiogène. Cette anxiété que l'on retrouve dans un environnement proche de l'*in situ* pourrait aussi impressionner les participants et limiter leurs projections. À l'inverse, dans l'étude précédente, les participants se trouvaient dans un environnement rassurant pour eux, puisqu'il s'agissait de leur domicile ou de leur lieu de travail. L'intrusion était moindre, ce qui a probablement mis davantage en confiance les participants, décuplé leur créativité et leurs capacités à se projeter. Cela les aurait incités à détailler chacun des éléments non appréciés dans

le dispositif. Ce phénomène ne semble pas être décrit, à notre connaissance, dans la littérature scientifique portant sur les études d'utilisabilité ayant recours à la simulation.

7. Conclusions et perspectives de recherche

Les analyses qualitatives sont une source très importante d'informations sur l'utilisabilité du matériel et nous ont permis, couplées aux données quantitatives, de proposer quatre perspectives de recherche.

Premièrement, nous avons pu constater des similitudes sur les explications données pour argumenter la survenue des erreurs. Parmi ces explications, la non-lecture des notices est de nombreuses fois abordée, mais cela reste en moins grande proportion que dans l'étude 2 (cf. Partie II, Chapitre 4, p. 189). Cette sous-représentation est probablement due à la non mise en place de la méthode de « penser à voix haut » (*think aloud*) permettant de saisir dans l'instant les pensées des utilisateurs. Cette méthode est donc un atout majeur pour analyser et comprendre la survenue des erreurs. Il est important de noter que les observations instrumentées par vidéos ont montré que la lecture des instructions se faisait de manière très rapide leur conférant en ce sens un statut de « rappel ». Les instructions servent essentiellement en cas d'oubli des procédures abordées lors de la formation. La formation est, par conséquent, d'une importance capitale.

Deuxièmement, l'analyse thématique est une activité complexe à traiter et l'arborescence proposée sur la satisfaction peut être remise en question, car certaines thématiques se chevauchent. C'est le cas de la complexité d'utilisation perçue et du risque perçu où les frontières entre ces deux propriétés semblent très perméables. Cependant, le modèle de Yen (2010) semble prometteur pour détailler les sous-métriques de la satisfaction jusqu'alors particulièrement discutée dans la littérature scientifique.

Troisièmement, il serait intéressant d'explorer certaines pistes d'amélioration de la pédagogie hospitalière. Par exemple, les explications apportées par le personnel infirmier lors de la formation du matériel au patient ne devraient pas être seulement descriptives, mais aussi explicatives afin d'argumenter pourquoi le dispositif doit être mis dans ce sens, ou sur ce bras.

Quatrièmement, nous avons pu retrouver les mêmes besoins que dans l'étude 2 (cf. Partie II, Chapitre 4, p. 189) permettant de proposer des recommandations de conception. Certaines

de ces remarques étaient parfois identiques à celles de l'étude 2 comme la taille de la police trop petite. Il serait intéressant de comprendre pourquoi certaines recommandations ne sont pas prises en compte par l'entreprise partenaire.

Pour conclure, nous sommes en accord avec Ten Hacken et al. (2018) pour décrire le rôle du personnel hospitalier comme central dans l'utilisabilité des dispositifs en santé pourtant à destination du patient. Il semble nécessaire de l'introduire au maximum dans les études d'utilisabilité – notamment à travers des simulations dans le cas où l'accès au terrain est difficile – car ce dernier fait partie de l'environnement, composant majeur du contexte d'usage.

8. Synthèse du chapitre 5 – Étude 3

Cette étude 3 correspond au **niveau 4** (tâches – utilisateurs – ressources – système – environnement) de notre modèle HANDLED. L'objectif principal de cette étude est de **mieux comprendre les relations qu'entretiennent le contexte d'usage à travers les caractéristiques utilisateurs (littératie en santé, âge, technophilie), les ressources (le programme d'éducation « actif » et « passif »), le système et l'environnement (simulé) sur l'utilisabilité (efficacité, efficience, satisfaction et apprenabilité) du dispositif Smart Angel.**

Pour cela, nous avons réalisé un test utilisateur auprès de 28 patients/participants (Moy = 30,79ans ; ET = 12,64 ; rang = 19 – 64 ans). Nous les avons introduits dans une simulation d'un service de chirurgie ambulatoire et du retour au domicile basée sur nos observations de terrain (cf. Partie II, Chapitre 1). Cette simulation a été associée à une **approche mixte de l'utilisabilité** (subjective, objective, qualitative et quantitative) *via* des observations directes, les observations instrumentées par vidéo, des questionnaires, entretiens semi-directifs et à la collecte de données sociodémographiques et cognitives (le niveau de littératie mesuré au moyen du Health Literacy Survey - Europe de Sørensen, et al., 2012, et du Newest Vital Sign de Weiss et al., 2005). Les participants ont, tout d'abord, consulté un anesthésiste, puis ont reçu en chambre hospitalière une formation à la manipulation du prototype 2 du dispositif *Smart Angel*. Enfin, ils ont manipulé le dispositif trois fois d'affilée entrecoupées de questionnaires, dans une habitation simulée. Des entretiens et focus group ont été enregistrés à la fin de la passation, systématiquement, avec le patient/participant et, quand cela était possible, avec le personnel hospitalier.

Les mesures d'efficience et d'efficacité ont été recueillies par observation instrumentée par vidéo. Les mesures de satisfaction ont été recueillies par questionnaire SUS et entretiens semi-directifs. Une analyse thématique, basée sur les retranscriptions des entretiens, a été effectuée afin de proposer une arborescence de la satisfaction.

Nous n'avons décelé aucun abandon lors de la première utilisation. Pour le tensiomètre, les utilisateurs ont commis en moyenne 1,11 erreurs (ET = 1,09). Les patients/participants ont commis en moyenne 0,93 erreur (ET = 0,42) sur l'oxymètre de pouls et 0,25 erreur (ET = 0,5) sur la tablette lors de la première utilisation. Cela implique que les utilisateurs ont réalisé en moyenne 2,13 erreurs (ET = 1,33) de manipulation et le score SUS moyen est de 87,63 (ET = 11,35) pour le système, ce qui indique un score "acceptable".

Nous avons formulé **six hypothèses selon lesquelles l'apprenabilité, le programme d'éducation, l'âge, la littératie en santé et la technophilie, l'environnement (simulé) auraient un impact sur l'utilisabilité.**

Premièrement, nos résultats suggèrent que **l'apprenabilité semble être une métrique connexe à l'utilisabilité de dispositif en santé.** En effet, les utilisateurs pouvant réitérer leurs erreurs de manipulation après plus d'essais, cela implique que la première utilisation d'un dispositif connecté en santé est primordiale.

Deuxièmement, nous n'avons pas pu nous assurer de l'impact des programmes d'éducation « actif » et « passif » du fait d'une mauvaise randomisation des groupes.

Troisièmement, nous avons pu observer que les utilisateurs plus âgés ne semblaient pas avoir plus de difficultés avec les dispositifs médicaux connectés (tensiomètre et oxymètre de pouls) par rapport aux utilisateurs plus jeunes, contrairement à ce que nous pensions. En revanche, nous observons **des difficultés sur l'utilisation de la tablette pour les utilisateurs âgés** contrairement aux utilisateurs jeunes.

Quatrièmement, nous nous étions intéressées à l'impact du niveau de littératie en santé (HSL-EU-Q16 et NVS) (cf. Partie II, Chapitre 4). Les scores du HLS-EU-Q16 n'ont montré aucun effet sur l'utilisabilité à l'exception du nombre d'erreurs sur la tablette. À l'inverse, les scores du NVS ont mis en évidence un effet significatif (corrélation forte) sur le nombre d'erreurs de manipulation (*efficacité*), le temps de manipulation (*efficacité*) sur l'ensemble du système *Smart Angel*. De plus, nous avons de nouveau identifié une corrélation particulièrement élevée sur la littératie en santé pour utiliser correctement le tensiomètre ($r = -0,584, p = ,001$) mais qui fait défaut pour l'oxymètre de pouls ($r = -0,123, p = ,532$), ce qui est l'inverse de nos observations lors de l'étude 1a (cf. Partie II, Chapitre 1a). **Comme dans l'étude précédente, nos résultats valident le lien entre la littératie en santé, la technophilie et l'âge des utilisateurs et l'utilisabilité.**

Cinquièmement, nous avons évalué l'impact de la technophilie sur l'utilisabilité du dispositif *Smart Angel*. Nous avons pu mettre en avant que les utilisateurs technophiles avaient tendance à être plus performants dans l'utilisation de la tablette par rapport aux utilisateurs technophobes. Nous avons également pu observer que les utilisateurs technophiles avaient tendance à regarder les instructions plus rapidement que les utilisateurs technophobes ce qui pourrait expliquer leur plus grande propension à réaliser des erreurs de manipulation sur l'oxymètre de pouls.

Sixièmement, nous nous sommes intéressées, *via* une analyse thématique basée sur les retranscriptions des entretiens et focus group, à l'impact de l'environnement (simulé) sur l'utilisabilité du dispositif. Nous avons pu retrouver des résultats similaires à nos analyses issues de l'étude 2 (cf. Partie II, Chapitre 4). De plus, **nous avons pu mettre en évidence que certains évènements du contexte pouvaient modifier l'utilisabilité**. C'est le cas des explications fournies par l'anesthésiste lors de la consultation pré-anesthésique afin d'aider le patient à comprendre l'utilité du matériel. De plus, les patients/participants nous ont rapporté l'importance de la qualité des explications apportées lors de la formation à la manipulation du dispositif par l'infirmier/ère. Des explications détaillant le « pourquoi » de la manipulation permettraient à l'utilisateur de retenir plus aisément les procédures.

Cette étude a permis de proposer des recommandations de conception du dispositif et de formation pour le personnel hospitalier avant la proposition d'un prototype pour l'étude clinique, dernière étude avant la commercialisation de *Smart Angel*. Cette étude a également servi de base afin de préciser les éléments à observer dans le service de chirurgie ambulatoire (e.g., la consultation anesthésique) pour l'analyse des modifications organisationnelles et relationnelles lors de l'implantation du dispositif (cf. discussion générale, 4.1.2. Le niveau 5 du modèle HANDLED : Les changements organisationnels et relationnels dus à l'intégration du dispositif *Smart Angel*, p. 294).

DISCUSSION GÉNÉRALE

1. Introduction

Ce travail de thèse avait pour objectif d'**assister la conception par l'évaluation de l'utilisabilité du dispositif connecté en santé *Smart Angel* dans un contexte post-ambulatoire**. Pour rappel, cette thèse s'inscrit dans le cadre d'un projet réel de conception : le projet *Smart Angel* de co-développement comprenant des partenaires industriels, académiques et hospitaliers. Afin de répondre à l'objectif de notre travail de recherche, nous avons conduit cinq études que nous avons intégrées au processus de conception du dispositif *Smart Angel*. Ces études ont permis d'éprouver le modèle HANDLED proposé afin d'**assister le processus de conception du dispositif *Smart Angel* en y ajoutant trois itérations supplémentaires**.

Le modèle HANDLED (Figure 63) est un modèle d'évaluation de l'utilisabilité adapté à la conception de dispositifs connectés en santé pour le domicile patient. Ce modèle est basé sur une définition opérationnelle de l'utilisabilité composée de métriques principales (*efficacité, efficience, satisfaction*) et connexes (*apprenabilité, mémorisation*). L'utilisabilité est également considérée comme une propriété située et évolutive. De ce fait, ce modèle, particulièrement inspiré de la norme ISO 9241-11 (2018) et du modèle de Yen et Bakken (2012), adapté à un contexte d'usage (*tâche, utilisateurs, ressources, système, environnements*) intégré une démarche d'évaluation en cinq niveaux itératifs est associés à cinq études de notre travail de thèse. La particularité de ce modèle réside dans ces itérations cumulatives. En effet, ce modèle fonctionne comme une pyramide inversée, dans laquelle chaque niveau viendrait s'ajouter au précédent jusqu'à proposer une évaluation complète. Nous observons donc une progression du contexte d'usage au fur et à mesure des études présentées.

La première étude, correspondant au niveau 0 (tâche) du modèle HANDLED, est une étude exploratoire visant à analyser le besoin et à comprendre le terrain, celui du parcours patient en chirurgie ambulatoire, afin de construire les premières maquettes du dispositif *Smart Angel*. Les études 1 (a et b), 2 et 3 permettent d'évaluer l'utilisabilité de manière itérative dans laquelle chaque étude représente une boucle d'itération. L'étude 1a évalue l'utilisabilité des objets connectés (tensiomètre et oxymètre de pouls connectés) de *Smart Angel* en prenant en compte les caractéristiques utilisateurs (niveau 1 : tâche – utilisateur). L'étude 1b évalue l'utilisabilité du tensiomètre en prenant en compte les ressources au travers des documents

procéduraux (niveau 2 : tâche – utilisateur – ressources). L'étude 2 évalue l'utilisabilité du prototype 1 comprenant le système complet de *Smart Angel* (niveau 3 : tâche – utilisateur – ressources – système) : une application sur tablette associée aux deux objets connectés précités. Enfin, l'étude 3 évalue le prototype 2 et intègre l'ensemble des éléments de l'étude 2 en ajoutant le contexte ambulatoire et le retour au domicile (niveau 4 : tâche – utilisateur – ressources – système – environnement).

Chacune de ces études a permis d'implémenter, par niveaux cumulatifs, le modèle HANDLED que nous proposons de discuter, dans cette partie, en mettant en perspective l'ensemble des études de ce travail de thèse. Pour cela, nous allons, dans un premier temps, discuter du contenu du modèle HANDLED à travers des métriques utilisées (*quoi évaluer ?*), puis des différentes itérations successives permettant la conception du contexte d'usage de ce modèle basé sur une stratification en cinq niveaux (*quand évaluer ?*). Dans un deuxième temps, nous aborderons les apports méthodologiques de la thèse à travers les méthodes et les outils mis en place. En effet, toutes les études mobilisaient une méthodologie mixte, associant une diversité d'outils permettant un recueil de données objectives, subjectives, qualitatives et quantitatives (*comment évaluer ?*). Enfin, nous préciserons les contraintes auxquelles ce travail a été confronté. Pour ce faire, nous présentons d'abord notre retour d'expérience sur le projet pluripartenarial *Smart Angel* et ses limites. Puis, nous présentons nos perspectives de recherche afin d'enrichir ce travail de thèse, mais également les études déjà mises en place qui apporteront des données concrètes dans les 2 à 3 années à venir.

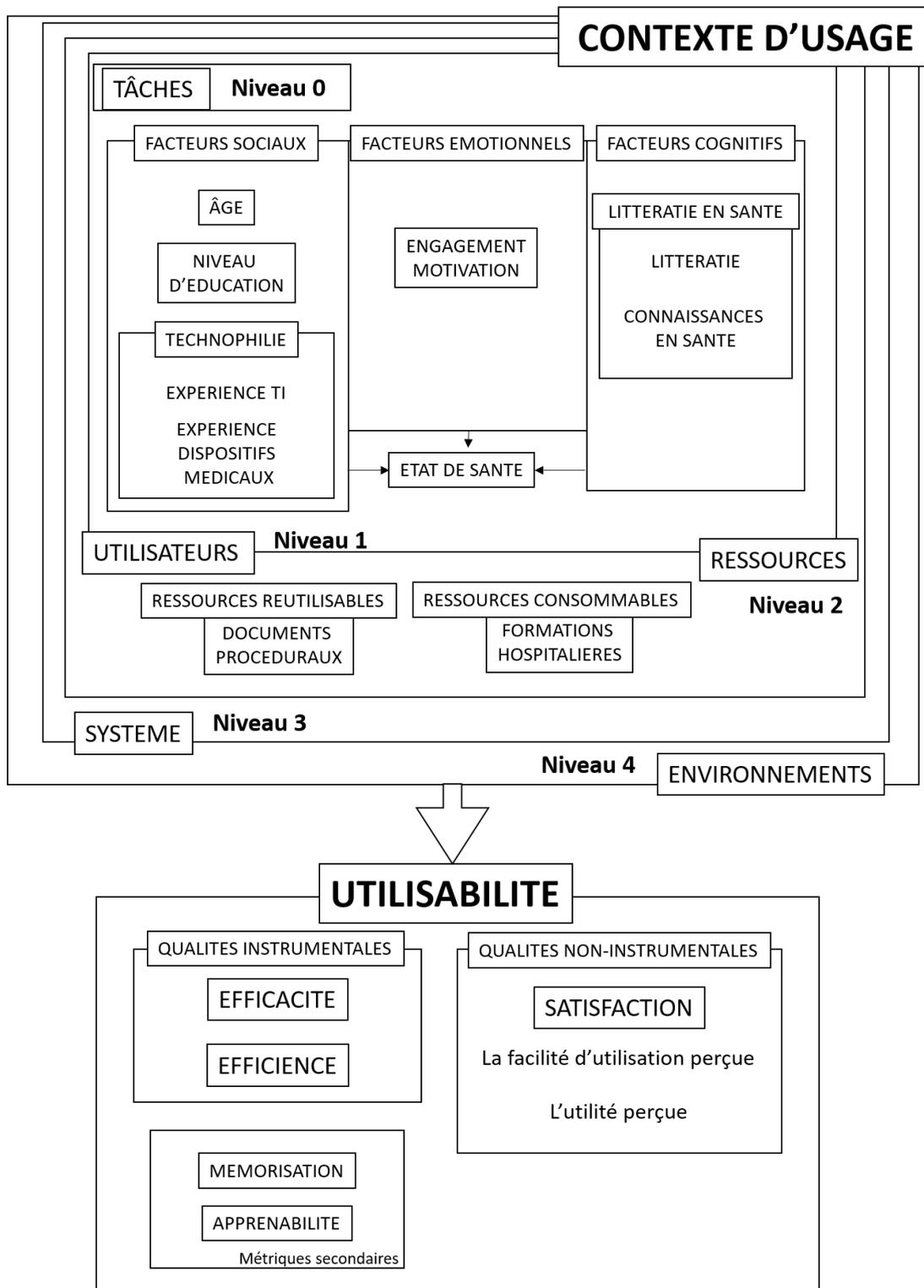


Figure 63. Home medical coNnected Devices usability Evaluation moDel (HANDLED) développé en partie théorique et éprouvé en partie empirique, composé des métriques d'utilisabilité et d'un contexte d'usage. Ce modèle est basé sur cinq niveaux cumulatifs en fonction des tâches, des utilisateurs, des ressources, du système et des environnements.

2. Les contributions théoriques et méthodologiques de la thèse : proposition et implémentation d'un modèle de l'évaluation de l'utilisabilité des dispositifs connectés en santé pour le domicile patient (HANDLED)

2.1. Contributions théoriques

2.1.1. Les métriques objectives du modèle HANDLED

- a. Les contradictions des mesures de l'efficacité au travers du nombre d'abandons et du nombre d'erreurs de manipulation : l'importance des mesures holistiques de l'efficacité

Nos résultats, issus des études 1, 2 et 3 (cf. Partie II, Chapitres 2, 3, 4 et 5), mettent en exergue l'importance de prendre en compte l'efficacité dans les évaluations de l'utilisabilité afin de garantir la sécurité du patient. Cette métrique nous renseigne plus précisément sur l'utilisation fiable et correcte du dispositif. En effet, les erreurs de manipulation – associées en partie aux mesures de l'efficacité – peuvent engendrer des données erronées, directement envoyées au centre hospitalier associé, rendant le dispositif *Smart Angel* inefficace voire douteux. Le personnel soignant, en charge de surveiller à distance la santé du patient, peut observer de fausses alertes ou, dans le pire des cas, ne pas prendre en considération de réelles problématiques médicales. De ce fait, nous avons été particulièrement attentives à cette métrique révélant les problèmes futurs de sécurité du patient.

Pourtant, nous pouvons observer, dans nos études expérimentales 1, 2 et 3 (cf. Partie II, Chapitres 2, 3, 4 et 5), **des résultats contradictoires dans les mesures de l'efficacité** du dispositif *Smart Angel*. Pour rappel, l'efficacité est définie comme la mesure selon laquelle le produit facilite l'utilisateur dans l'accomplissement de la tâche qu'il doit réaliser avec précision et exhaustivité tout en évitant la plupart des erreurs dans divers contextes d'utilisation (Seffah, Kececi, & Donyaee, 2001). Dans les études 1, 2 et 3 (cf. Partie II, Chapitres 2, 3, 4 et 5), nous avons systématiquement mesuré l'efficacité sur la base de deux indicateurs : le nombre d'abandons et le nombre d'erreurs de manipulation. **Nous avons pu observer que le nombre d'abandons n'a fait que décroître entre les études 1** (cf. Partie II, Chapitres 2 et 3), **2 et 3** (cf. Partie II, Chapitres 4 et 5). Cet indicateur nous a permis d'appréhender l'utilisation du matériel en visualisant le nombre d'utilisateurs réussissant à accomplir la tâche demandée : celle de prendre sa mesure physiologique et de répondre aux questionnaires issus de l'application *My Angel*. La totalité de l'échantillon d'utilisateurs de l'étude 3 (cf. Partie II, Chapitre 5) a pu réaliser la tâche prévue. Au vu de ce résultat, il semblerait naturel de penser

que l'efficacité du dispositif s'est améliorée en fonction des différents prototypes présentés. Pourtant, à l'inverse du nombre d'abandons, **le nombre d'erreurs de manipulation n'a fait que croître entre les études 1** (cf. Partie II, Chapitres 2 et 3), **2 et 3** (cf. Partie II, Chapitres 4 et 5). Nous pouvons nous demander pour quelles raisons nous avons observé une décroissance du nombre d'abandons alors que le nombre d'erreurs de manipulation a augmenté.

Au vu de ce constat, nos résultats semblent paradoxaux. Nous proposons **une explication à ce paradoxe par la présence d'erreurs « non-bloquantes »** détectées dans l'ensemble de nos études expérimentales (1, 2 et 3). Pour rappel, Hornbæk et Law (2007) décrivent la présence de deux types d'erreurs de manipulation : l'erreur en cours de route que nous qualifions d'erreur « bloquante » car cette dernière va impacter l'atteinte de la tâche prévue pouvant mener à un abandon ; et l'erreur d'achèvement de la tâche que nous appelons l'erreur « non-bloquante » car cette dernière ne gêne pas l'utilisateur pour atteindre la tâche prévue (cf. Partie I, Chapitre 1, 3.1.5. Les relations entre efficacité, efficacité et satisfaction, p. 33). Nous avons pu observer dans les études 1a et 1b (cf. Partie II, Chapitres 2 et 3) un nombre important d'abandons causé par la présence d'un tensiomètre créant des erreurs « bloquantes » pour l'utilisateur. Ces erreurs sont dues à un système de flèches lumineuses imposant une certaine inclinaison de l'avant-bras pour pouvoir lancer la prise de mesure. Si cette inclinaison n'est pas respectée, la mesure ne se déclenche pas, l'atteinte de la tâche ne peut pas se faire et l'utilisateur se retrouve alors bloqué et peut abandonner s'il ne trouve pas de solution. Au vu du nombre d'abandons lors de l'étude 1, les parties prenantes du projet *Smart Angel* ont pris la décision de changer le tensiomètre. Sur ce dernier modèle que nous avons utilisé pour les études 2 et 3 (cf. Partie II, Chapitres 4 et 5), ce système de flèche lumineuse est absent. Il est seulement inscrit dans la notice de positionner le tensiomètre à hauteur du cœur. Si l'utilisateur n'incline pas convenablement l'avant-bras, il peut lancer la mesure, mais celle-ci sera erronée. L'erreur devient « non-bloquante », car l'utilisateur atteint artificiellement la tâche demandée. Les erreurs « bloquantes » et « non-bloquantes » ainsi que leurs variations sur la mesure sont détaillées en annexe (cf. Annexe 20 : Grille de notation des erreurs de manipulation du dispositif, p. 410).

Ce constat met en évidence l'importance d'évaluer l'efficacité sur plusieurs mesures (i.e., le nombre d'abandons et le nombre d'erreurs de manipulation) **afin d'obtenir une vision holistique de l'efficacité** lors des évaluations de l'utilisabilité d'un dispositif connecté en santé.

b. Faut-il privilégier l'efficacité au détriment de l'efficacité ?

L'efficacité et l'efficacité ont une relation complexe et aucune de ces deux métriques ne doit être négligée. Lors de l'étude 1b (cf. Partie II, Chapitre 3), nos résultats ont révélé une

opposition entre les mesures d'efficacité et d'efficience. Pour rappel, l'efficience est définie par les ressources utilisées par rapport aux résultats obtenus d'après la norme ISO 9241-11 (cf. Partie I, Chapitre 1, 3.1.2. L'efficience : définitions et mesures, p. 28). Les résultats de l'étude 1b (cf. Partie II, Chapitre 3) ont montré que certains formats de présentation (e.g., audio) augmentaient le temps de manipulation relatif à l'efficience et diminuaient le nombre d'erreurs de manipulation et le nombre d'abandons relatifs à l'efficacité. À l'inverse, les autres formats comme ceux utilisés dans l'application *Smart Angel* composée d'images et de texte diminuaient le temps de manipulation et augmentaient le nombre d'erreurs. Afin d'améliorer l'efficacité, faudrait-il détériorer l'efficience ?

Nous avons pu constater dans l'étude 3 (Cf. Partie II, Chapitre 5) que le nombre d'erreurs de manipulation entre les différents essais ne diminue pas significativement contrairement au temps de manipulation relatif à l'efficience. Ce résultat est dû au nombre important d'erreurs « non-bloquantes » détectées, comme nous l'avons vu dans la section précédente. L'utilisateur ne corrige pas ses erreurs « non-bloquantes » de manipulation et les réitère systématiquement biaisant inconsciemment ses prises de mesure (Cifter, 2017). En revanche, au fur et à mesure des essais, l'utilisateur semble apprendre à utiliser le dispositif avec plus d'aisance et en sollicitant moins de ressources, car son temps de manipulation diminue (cf. Partie II, Chapitre 5, 5.1.1. Effet d'apprentissage sur l'efficacité et l'efficience, p. 245). Ces résultats vont dans le sens de la méta-analyse de Hornbæk et Law (2007) mettant en relation les mesures d'efficacité et d'efficience. Les auteurs montrent une très faible corrélation positive entre les mesures d'efficience et d'efficacité lors d'erreurs « non-bloquantes » et une corrélation modérée entre les mesures d'efficience et d'efficacité lors d'erreurs « bloquantes ». **Ce constat implique qu'il serait pertinent, dans un premier temps, d'optimiser l'efficacité au détriment de l'efficience qui s'améliorerait au fur et à mesure des essais.**

Cependant, **en réalité, il ne s'agirait pas de détériorer l'efficience, mais d'encourager un engagement cognitif chez l'utilisateur afin d'optimiser l'efficacité.** En effet, dans la discussion de l'étude 1b (cf. Partie II, Chapitre 3), nous expliquons nos résultats paradoxaux entre l'efficacité et l'efficience, par le fait que le format audio a permis de « forcer » l'utilisateur à s'engager cognitivement (Sinatra, Heddy, & Lombardi, 2015) en ayant recours à une stratégie d'atomisation de l'action et ainsi contraint à lire/écouter les procédures. Cela s'est alors traduit par une augmentation du temps de manipulation mais également par une diminution du nombre d'erreurs. Nous confirmons ces observations par les analyses qualitatives des études 2 et 3 (Partie II, Chapitres 4 et 5) mettant en exergue l'absence de lecture des procédures par les

utilisateurs. En effet, les utilisateurs tendent toujours à réduire leur coût cognitif lors de la manipulation d'un produit (e.g., Eiriksdottir & Catrambone, 2011; Fu & Gray, 2006; Ganier, 2004). Ainsi, un engagement cognitif conséquent dans la tâche (Sinatra, Heddy, & Lombardi, 2015) permettrait de réduire significativement le nombre d'erreurs de manipulation. Ainsi, les résultats des études 1b (cf. Partie II, Chapitre 3) et 3 (Partie II, Chapitre 5) suggèrent qu'un temps de manipulation important – traduisant une faible efficacité – ne signifie pas systématiquement qu'un produit soit inutilisable si ce dernier met en évidence peu d'erreurs de manipulation relative à une bonne efficacité. Cela témoigne, en revanche, d'un effort cognitif important, par l'utilisateur, pour limiter les erreurs.

En revanche, ces analyses témoignent de deux constats majeurs relatifs aux relations entre l'efficacité et l'efficacité : **l'importance de détecter le type d'erreurs de manipulation telles que les erreurs non-bloquantes invisibles aux mesures d'efficacité** (Hornbæk et Law, 2007) et **d'encourager l'engagement cognitif de l'utilisateur, par le biais de formats de procédures adaptés, afin de réduire significativement ces erreurs et prévenir la sécurité du patient.**

2.1.2. La métrique subjective du modèle HANDLED

a. La prise en compte du risque perçu dans la satisfaction

Les résultats des analyses qualitatives des études 2 (cf. Partie II, Chapitre 4) et 3 (cf. Partie II, Chapitre 5) suggèrent d'intégrer le risque perçu dans les attributs de la satisfaction. Pour rappel, dans nos études, nous avons tenté d'appréhender cette métrique au travers des attributs de Dubey et Gulati (2012) comprenant la sympathie, la confiance, le confort, l'attractivité, et au travers du modèle de Yen (2010) comprenant l'utilité perçue et la facilité d'utilisation perçue. Nos analyses thématiques réalisées sur des données extraites d'entretiens et de focus groups nous ont menées à intégrer « le risque perçu » dans la satisfaction. Ce nouvel élément, lié au contexte médical, pouvait modifier la perception de l'utilisateur jusqu'à altérer ses comportements. En effet, lors de l'étude 3 (cf. Partie II, Chapitre 5), certains utilisateurs ont décrit le dispositif comme « rigolo » ou « sympathique » et n'avaient pas conscience des risques postopératoires contrairement aux professionnels de santé mobilisés pour l'étude. Dans nos études 2 et 3 (cf. Partie II, Chapitres 4 et 5), les utilisateurs ont affirmé qu'ils ne seraient probablement pas rigoureux lors de l'utilisation du dispositif s'ils venaient à devoir l'utiliser pendant une semaine. Les utilisateurs ont également expliqué ne pas lire les instructions étant donné que le produit semblait facile d'utilisation et sans risque apparent pour la sécurité du patient. Ces deux constats montrent un faible risque perçu du dispositif *Smart*

Angel et suggèrent que cette métrique aurait un rôle important dans l'orientation des choix comportementaux des utilisateurs. Cottin-Joshi et Vallery (2015) ont montré, par exemple, que le personnel hospitalier adapte ses équipements de protection en fonction du risque perçu. Ces professionnels sont en revanche experts dans leur métier et dans la perception du risque ce qui n'est pas le cas de nos utilisateurs novices. **L'intégration de cet attribut à la satisfaction semble particulièrement adaptée aux évaluations d'utilisabilité des dispositifs de santé.** De plus, le risque perçu est, à notre connaissance, absent des modèles d'utilisabilité et de la littérature scientifique sur les dispositifs en santé.

b. L'impact de l'utilité perçue sur la satisfaction

Les résultats des analyses qualitatives des études 2 (cf. Partie II, Chapitre 4) et 3 (cf. Partie II, Chapitre 5) ont montré que l'utilité perçue possède un rôle majeur dans la satisfaction des dispositifs en santé. Plusieurs modèles théoriques intègrent déjà l'utilité perçue dans la satisfaction des dispositifs en santé (e.g., Yen et al., 2010). Cependant, ces modèles ne précisent pas en quoi cette utilité perçue est nécessaire afin d'améliorer la satisfaction. Dans nos études 2 et 3 (cf. Partie II, Chapitres 4 et 5), intégrant une analyse thématique des entretiens et focus group, nous avons pu observer que l'utilité-destination (Loup-Escande et al., 2013) du matériel n'était pas toujours comprise. Nous avons relevé trois arguments montrant que **ces ambiguïtés, associées à une mauvaise utilité perçue, peuvent avoir un impact sur la satisfaction du matériel pouvant engendrer jusqu'à un rejet du dispositif par l'utilisateur.** Premièrement, les analyses qualitatives des études 2 et 3 (cf. Partie II, Chapitres 4 et 5) ont montré que les utilisateurs ont majoritairement mis en avant l'apport d'un gain de temps par rapport à l'intérêt médical. Pour rappel, l'objectif principal du dispositif *Smart Angel* est de prévenir les complications post-ambulatoires. Deuxièmement, dans l'étude 3 (cf. Partie II, Chapitre 5), les utilisateurs ont été nombreux à questionner les expérimentateurs sur l'intérêt d'évaluer la tension ou l'oxygénation alors que l'opération se situait au niveau digestif. Pour rappel, les utilisateurs possédaient tous le même scénario dans lequel ils souffraient d'une hernie inguinale (cf. Annexe 15 : Scénario Patients/participants, p. 408). Plus précisément, ils ne comprenaient pas l'intérêt de prendre leur tension alors qu'ils avaient des douleurs dans le bas ventre. Troisièmement, nous avons également pu constater que des utilisateurs préféraient refuser le matériel lorsqu'ils n'en voyaient pas l'utilité. Ce n'est qu'après des explications sur les potentiels risques opératoires que les participants ont été convaincus et accepteraient volontiers de s'en resservir si le contexte ambulatoire l'imposait. **Ces trois**

arguments montrent l'importance de vérifier l'utilité perçue dans le cadre de la satisfaction du matériel.

2.1.3. Les composants du contexte d'usage du modèle HANDLED

- a. Les caractéristiques utilisateurs (niveau 1) : La littératie en santé, une compétence majeure pour améliorer l'utilisabilité des dispositifs en santé

Les résultats de nos études ont montré que les caractéristiques utilisateurs (l'âge, la technophilie, la littératie en santé) impactent l'utilisabilité du matériel. Les études 1a, 2 et 3 (cf. Partie II, Chapitres 2, 4 et 5) ont montré que la littératie en santé impactait significativement l'efficacité du matériel davantage que les autres variables (e.g., âge, technophilie). Ces résultats vont dans le sens de ceux de Van der Vaart et al. (2019) déclarant que cette compétence possède un très fort impact sur l'utilisabilité.

Plus précisément, l'étude 1a (cf. Partie II, Chapitre 2) a mis en évidence l'impact des connaissances antérieures en santé sur l'utilisabilité de l'oxymètre de pouls et **a révélé la présence d'un effet de seuil (Chaniaud et al., 2020). Cela suggère qu'il existerait un niveau minimal requis permettant d'utiliser le dispositif.** Si l'utilisateur ne possède pas ce niveau, il réaliserait significativement plus d'erreurs ou échouerait dans l'utilisation contrairement aux utilisateurs ayant le niveau requis. Ces résultats étaient visibles avec l'oxymètre de pouls. En revanche, nous avons considéré que cet effet n'était pas visible avec le tensiomètre du fait de sa difficulté d'utilisation trop importante. Aucun de nos utilisateurs ne possédait le niveau de connaissances requis pour avoir une utilisation correcte de cet objet.

Dans le même sens que l'étude 1a, (cf. Partie II, Chapitre 2), les études 2 et 3 (cf. Partie II, Chapitres 4 et 5) ont montré que la littératie en santé mesurée à l'aide de questionnaire d'évaluation (NVS) – et non auto-rapporté (HLS) – impactait l'efficacité du dispositif *Smart Angel*. Contrairement à l'étude 1a, les résultats de l'étude 3 (cf. Partie II, Chapitre 5) ont mis en évidence un effet de seuil sur l'utilisabilité du tensiomètre et non de l'oxymètre de pouls. Entre ces deux études, les modifications apportées au système, aux ressources (i.e., instructions), aux environnements (i.e., contexte hospitalier permettant de recevoir une formation spécifique par un personnel en soin infirmier) auraient permis d'améliorer l'utilisabilité du dispositif et d'abaisser le niveau requis pour utiliser ces objets connectés. En outre, les résultats de l'étude 3 mettent en exergue une importante facilité d'utilisation de l'oxymètre de pouls liée à un faible nombre d'erreurs de manipulation comparé à l'étude 1a.

Afin d'illustrer cette relation entre l'efficacité et la littératie en santé, que nous venons de détailler à partir des résultats des études 1a et 3, nous présentons un schéma qualitatif (Figure 64) de cet effet de seuil. Une importante marge d'erreur en fonction de la complexité du dispositif pourrait expliquer nos résultats contradictoires entre les études 1a et 3. Plus précisément, afin d'utiliser correctement un dispositif connecté en santé, l'utilisateur doit systématiquement avoir le niveau minimal de littératie en santé exigé en fonction de la complexité d'utilisation du dispositif. Ces résultats vont dans le sens du modèle de Monkman et al., (2015) proposant un modèle similaire sur le niveau de littératie en e-santé. Ainsi, le tensiomètre de pouls du dispositif *Smart Angel*, lié à sa complexité d'utilisation, demande un niveau de littératie en santé plus important que l'oxymètre de pouls.

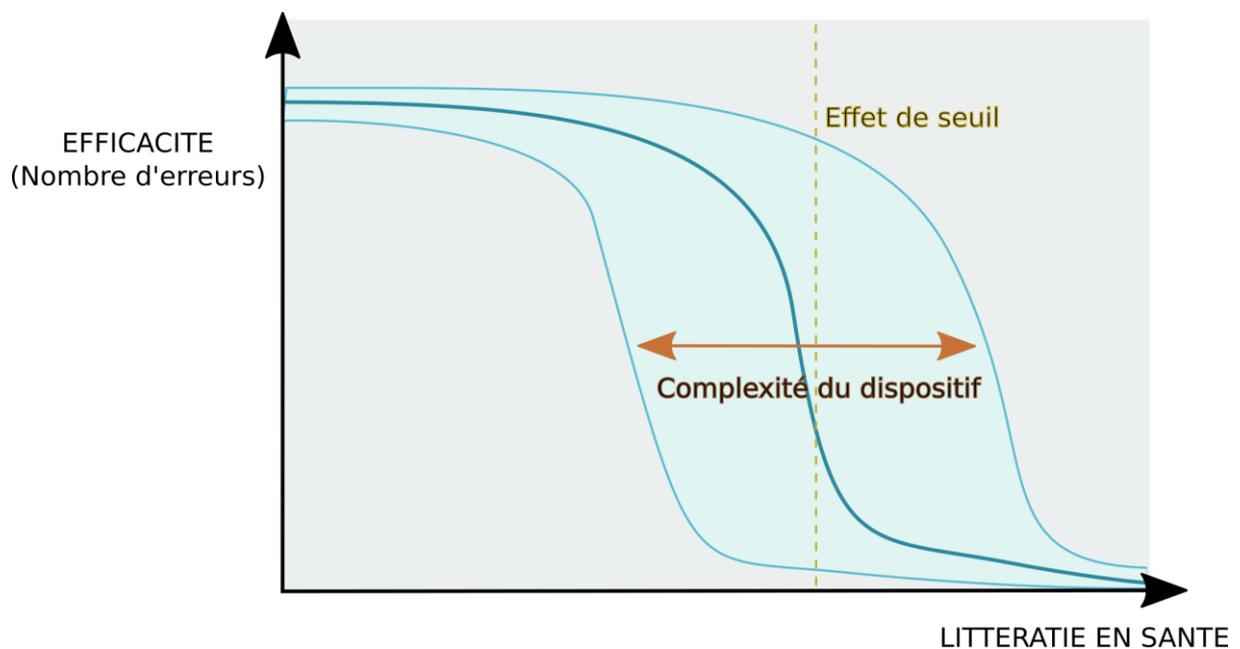


Figure 64. Représentation schématique qualitative de l'effet du niveau de la littératie en santé sur l'efficacité.

b. Les ressources (niveau 2) : l'importance de la formation hospitalière par rapport aux procédures

Nos travaux de thèse ont mobilisé deux types de ressources : réutilisables (i.e., les documents procéduraux) dans l'étude 1b (cf. Partie II, Chapitre 3) et consommables (i.e., les formations hospitalières) dans l'étude 3 (cf. Partie II, Chapitre 5). Lors de l'étude 1b (cf. Partie II, Chapitre 3), nous avons pu observer l'importance des ressources réutilisables à travers la lecture des procédures pour utiliser le dispositif correctement. En effet, nous avons pu constater que les utilisateurs lisant attentivement les procédures réalisaient moins d'erreurs de

manipulation que ceux ne les lisant pas. Les analyses qualitatives relatives à l'étude 2 (cf. Partie II, Chapitre 4) et à l'étude 3 (cf. Partie II, Chapitre 5) ont confirmé nos observations quantitatives. Les utilisateurs ont expliqué la survenue de leurs erreurs, en grande proportion, par la non-lecture des procédures. Ces analyses qualitatives ont également permis de révéler que les utilisateurs ont préféré investir les ressources réutilisables (démonstration et formation hospitalière) au détriment des ressources consommables (les procédures). Cette préférence viendrait de leur coût cognitif : les ressources consommables ont servi de rappel, alors que les ressources réutilisables ont servi de référence. Lors de l'étude 1b (cf. Partie II, Chapitre 3), il a fallu trouver des stratégies, telles que le recours à des formats audio, afin de forcer les utilisateurs à lire ces procédures. Ainsi, nos résultats suggèrent que si le concepteur n'intègre pas de format adapté à une lecture forcée des procédures, il est préférable d'investir dans les formations hospitalières (i.e., ressources consommables) plutôt que dans la conception de procédures (i.e., ressources réutilisables) puisque dans tous les cas, l'utilisateur sera tenté de ne pas lire ces dernières. De plus, si des ressources consommables (e.g., programmes d'éducation thérapeutique) sont proposées aux patients utilisateurs, alors le concepteur se doit de vérifier qu'elles apportent l'ensemble des éléments nécessaires à l'utilisateur afin de manipuler le dispositif correctement.

c. Le système (niveau 3) : l'impact du système sur les utilisateurs

Les résultats de l'étude 3 (cf. Partie II, Chapitre 5), détaillant l'efficacité et l'efficience de chaque élément du système *Smart Angel* (i.e., comprenant une application sur tablette et deux objets connectés : un tensiomètre et un oxymètre) ont montré qu'il existait des différences importantes dans la complexité intrinsèque de ces objets. Les résultats de l'étude 2 (cf. Partie II, Chapitre 4) montrent que les différents objets du dispositif n'étaient pas considérés de la même manière par les utilisateurs. En effet, dans l'étude 3 (cf. Partie II, Chapitre 5), nous avons pu observer que certains utilisateurs, en particulier les plus âgés, pouvaient être en difficulté avec la tablette et non avec les objets connectés. À l'inverse, d'autres utilisateurs, à savoir les plus technophiles, pouvaient montrer une facilité certaine dans l'utilisation de la tablette et davantage de difficultés avec les objets connectés. Ces résultats suggèrent que les objets présents dans le système peuvent être perçus et utilisés de manière inégale. Ainsi, les utilisateurs ayant des difficultés sur l'un ou l'autre des dispositifs ne peuvent pas manipuler correctement le système dans sa globalité. Afin de répondre à ces problématiques, il semble important de détailler les données pour chaque élément du système lors des évaluations de l'utilisabilité.

d. L'environnement (niveau 4) : l'importance de la pédagogie hospitalière dans un contexte post-chirurgical

Les observations et analyses prospectives de l'étude exploratoire (cf. Partie II, Chapitre 1) ont permis de comprendre le parcours patient en chirurgie ambulatoire. Nous avons également pu investiguer, au cours de cette étude, les futures modifications apportées à ce parcours pour l'implantation du dispositif *Smart Angel* au sein du service : l'anesthésiste, lors de la consultation pré-anesthésique serait le premier à aborder le dispositif avec le patient ; puis, ce dernier recevrait en chambre, après son opération, une formation spécifique à l'utilisation du matériel. Nous avons pu confirmer le fait, lors de l'étude 3 (cf. Partie II, Chapitre 5) où nous avons reconstitué et simulé un service de chirurgie ambulatoire, que cette première consultation anesthésique, issue du contexte hospitalier, était la clé pour la bonne intégration du dispositif. En effet, les analyses qualitatives de l'étude 3 ont révélé que le patient n'est pas en mesure de comprendre l'utilité-destination du matériel seul. Afin de percevoir pleinement cette utilité, il a besoin de pédagogie mise en avant par le personnel hospitalier et plus spécifiquement de la part de l'anesthésiste lors de leur première rencontre en consultation.

Nos résultats suggèrent que cette pédagogie hospitalière issues de *l'environnement*, pourtant éloignée de l'interaction entre l'utilisateur et le dispositif, impacte l'utilisabilité. En effet, pour rappel, nos résultats issus de l'étude 1b (cf. Partie II, Chapitre 3) ont montré que l'engagement cognitif accordé au dispositif jouait un rôle majeur pour motiver la lecture de la notice et améliorer l'utilisabilité du matériel (cf. Discussion générale, 1.1.1.a. Les contradictions des mesures de l'efficacité au travers du nombre d'abandons et du nombre d'erreurs de manipulation, p. 272). De la même manière, lors de l'étude 3 (cf. Partie II, Chapitre 5), nos résultats ont mis en évidence que la consultation anesthésique permettait aux patients de se créer une représentation mentale du dispositif *Smart Angel* ce qui pourrait modifier cet engagement futur envers le matériel. Si le patient n'est pas convaincu de l'utilité du matériel alors ce dernier n'investira pas cognitivement le dispositif impactant à son tour l'utilisabilité du dispositif *Smart Angel*. Ce constat met en exergue comment des éléments issus de *l'environnement*, n'étant pas forcément associés directement à l'utilisation du matériel, peuvent en impacter son utilisation.

Ainsi, afin de parfaire cette pédagogie et améliorer l'utilisabilité du dispositif, le personnel hospitalier doit s'exprimer dans un langage compréhensible pour le patient. Lors des entretiens semi-directifs menés dans le cadre de l'étude 3 (cf. Partie II, Chapitre 5), certains patients ont déclaré ne pas comprendre ce qu'expliquait l'anesthésiste. Par exemple, l'anesthésiste utilisait

le terme « d'oxymètre de pouls » afin de décrire le contenu de la mallette *Smart Angel*. Ce terme pouvait être inconnu par le patient. Ce dernier était alors incapable de se créer une représentation du matériel et d'investir l'utilité du produit. Il est alors nécessaire de vérifier que les explications de l'anesthésiste sont suffisamment accessibles (e.g., en se basant sur le niveau de littératie en santé) pour le patient.

2.1.4. L'apport de la progression des études dans le modèle HANDLED

Le modèle HANDLED (Figure 63) a été éprouvé sur la base de cinq études basées sur cinq niveaux du contexte d'usage. Chaque étude est un cumul du contexte d'usage de la précédente étude. Ce modèle est basé sur des métriques complémentaires permettant une évaluation holistique de l'utilisabilité et un contexte d'usage inspiré de la norme ISO 9241-11 et du modèle de Yen et Bakken (2012). **Le cumul du contexte d'usage, au travers de niveaux (*tâches, utilisateurs, ressources, système, environnements*) permettant une progression des évaluations itératives de l'utilisabilité, est l'atout majeur de ce modèle.** En effet, cela permet de mieux comprendre l'impact de chacun de ces éléments en contrôlant progressivement les variables implantées dans les études.

La première étude exploratoire (cf. Partie II, Chapitre 1), représentant le niveau 0 de notre modèle HANDLED, a permis de comprendre le terrain et les futures tâches imposées par le dispositif *Smart Angel*. Cette étude a servi de base pour comprendre l'environnement dans lequel sera intégré le dispositif.

L'étude 1a (cf. Partie II, Chapitre 2), représentant le niveau 1 du modèle HANDLED évalue l'interaction entre l'*utilisateur* et la *tâche* au travers de l'utilisation de certains éléments du dispositif (tensiomètre et oxymètre de pouls).

L'étude 1b (cf. Partie II, Chapitre 3), correspondant au niveau 2 du modèle, a montré l'impact des *ressources* (réutilisables) sur la *tâche* en fonction de l'*utilisateur* sur le tensiomètre du dispositif.

L'étude 2 (cf. Partie II, Chapitre 4), représentant le niveau 3 du modèle, a révélé l'impact des éléments composants *Smart Angel* sur l'utilisabilité du *système* complet en prenant en compte l'*utilisateur* et ses *ressources* réutilisables. Cette étude a permis, dans un premier temps, de valider les relations entre l'utilisabilité et certaines caractéristiques utilisateurs (littératie en santé) déjà investiguées lors de l'étude 1a. Dans un deuxième temps, cette étude a validé les

observations de l'étude 1b sur l'impact de la non-lecture des instructions (ressources réutilisables) pas les utilisateurs.

L'étude 3 (cf. Partie II, Chapitre 5), correspondant au niveau 4 du modèle, s'intéressent à l'impact de l'*environnement* hospitalier et du retour au domicile sur l'utilisabilité du *système* complet en prenant en compte les caractéristiques *utilisateurs* et les *ressources* réutilisables et consommables (issus de la création de l'environnement). Comme pour l'étude 2, cette étude met en évidence le lien incontournable entre certaines caractéristiques utilisateurs et notamment de la littératie en santé et l'utilisabilité. Deuxièmement, les observations réalisées lors de l'étude 3 ont également montré, comme les études 1b et 2, les relations entre les ressources réutilisables et l'utilisabilité au vu de la non-lecture des instructions. Ces deux éléments que nous venons de citer (*utilisateurs* et *ressources* réutilisables) issus du contexte ne semblent pas être influencés par l'*environnement*. En revanche, lors de cette dernière étude, nous avons pu observer que les utilisateurs ont davantage investi les ressources consommables (formations hospitalières) au détriment des ressources réutilisables améliorant l'utilisabilité du dispositif.

Ainsi, la démarche du modèle HANDLED permet d'observer les variations entre les composants constituant le contexte d'usage et l'utilisabilité aux différents stades de l'évaluation. En effet, l'ensemble des éléments constituant le contexte d'usage a pu être testé de manière individuelle et combinée consolidant ainsi les liens entre chacun de ces composants.

2.2. Contributions méthodologiques

2.2.1. Les avantages et les inconvénients des principaux outils déployés

Nous avons défini, dans notre partie théorique (cf. Partie I, Chapitre 1, 2.6. Synthèse des modélisations : vers une définition de l'utilisabilité et proposition d'un modèle opérationnel, p. 25), l'utilisabilité comme une dimension évaluant les propriétés instrumentales et non instrumentales d'un produit. De ce fait, il est important d'avoir recours à des études mixtes associant des recueils de données qualitatives, quantitatives, subjectives et objectives (e.g., Ammenwerth et al., 2003 ; Jaspers, 2009). Afin d'obtenir un tel recueil, il est nécessaire de mobiliser une combinaison d'outils. Actuellement, peu de preuves empiriques permettent de juger de l'intérêt des outils et des méthodes ou de leurs combinaisons pour l'évaluation de l'utilisabilité des dispositifs de santé (Marcilly et al., 2016). Cependant, dans l'ensemble de nos

études, nous avons mobilisé une importante variété d'outils et de méthodes. Nous proposons de décrire les apports et les inconvénients de chaque outil et méthode utilisés :

- **L'observation directe** est utile afin de détecter la présence d'erreurs de manipulation et de préciser le type d'erreurs « bloquantes » et « non-bloquantes » (cf. Discussion générale, a. Les contradictions des mesures de l'efficacité au travers du nombre d'abandons et du nombre d'erreurs de manipulation, p. 272). Cependant, son intérêt dépend de la présence de l'examineur. En effet, cette méthode est d'autant plus pertinente si l'examineur est absent du lieu dans lequel manipule l'utilisateur, car ce dernier se sent moins observé. Cette gêne due à la présence d'une tierce personne peut modifier le comportement de l'utilisateur (Kjeldskov et al., 2005). Dans les études 1 et 2 (cf. Partie II, Chapitres 2, 3 et 4), l'examineur était présent dans la pièce et les utilisateurs ont rapporté se sentir surveillés. En revanche, dans l'étude 3 (cf. Partie II, Chapitre 5), l'examineur observait au travers d'une vitre sans tain ou grâce à une retranscription en direct des caméras de surveillance de la pièce. Les utilisateurs nous ont révélé ne pas avoir eu la sensation d'être observés. Il serait donc intéressant **de distinguer l'observation directe visible de l'observation directe invisible** (ou discrète). De plus, nos observations issues des études 2 et 3 (cf. Partie II, Chapitres 4 et 5) ont montré que l'observation directe est pertinente si elle est combinée à des entretiens ou des focus groups ultérieurs à la manipulation, car cela permet à l'examineur de prendre note de certaines difficultés rencontrées pendant la manipulation qui pourront ensuite être discutées et approfondies grâce aux données subjectives. Ainsi, nous avons davantage compris la survenue d'erreurs de manipulation grâce à cette combinaison de méthode.
- **L'observation instrumentée par vidéo** utilisée dans l'ensemble de nos études expérimentales (études 1, 2 et 3) est une méthode efficace pour mesurer finement l'efficiency et l'efficacité. En effet, cette méthode permet d'observer le temps d'exécution et de lecture des instructions. Cela permet aussi de mieux détecter certaines erreurs de manipulation pouvant être imperceptibles lors de l'observation directe. En revanche, elle est davantage coûteuse humainement, par rapport à l'observation directe, du fait des traitements des données nécessitant une grille de notation et un double codage par au moins deux expérimentateurs. Cette méthode demande aussi du matériel (caméra fixée au mur ou sur trépied). Le positionnement de la caméra ne doit pas être négligé

pour observer l'utilisateur. Si ce dernier se détourne de la caméra, les données vidéos ne seront pas exploitables ce qui peut être fréquent lorsque l'expérimentateur est absent de la pièce dans lequel se déroule la passation. Il est parfois nécessaire d'avoir une pièce équipée de plusieurs caméras. Ainsi, comme pour l'observation directe, l'angle et l'emplacement plus ou moins discret de la caméra peuvent biaiser l'étude. En effet, lors des études 1 et 2 (cf. Partie II, Chapitres 2, 3 et 4), la caméra était posée sur un trépied proche de l'utilisateur. Dans l'étude 3 (cf. Partie II, Chapitre 5), différentes caméras étaient accrochées aux murs. La plupart des utilisateurs nous ont dit avoir oublié les caméras pendant la passation de cette dernière étude. Ainsi, tout comme l'observation directe, il serait intéressant de **différencier l'observation instrumentée par vidéo visible de l'observation instrumentée par vidéo discrète**. Cette méthode est, à notre sens, incontournable dans les évaluations de l'utilisabilité.

- **Les entretiens semi-directifs et les focus groups** nous ont permis de mesurer la satisfaction. Ces méthodes qualitatives semblent particulièrement pertinentes afin de comprendre la perception de l'utilisateur sur la facilité d'utilisation perçue et l'utilité perçue. En effet, les analyses qualitatives des études 2 et 3 (cf. Partie II, Chapitres 4 et 5) basées sur les entretiens et sur les focus groups ont permis de réaliser des arborescences de la facilité d'utilisation perçue et de l'utilité perçue. Comme attendu, les entretiens semi-directifs permettent d'avoir une vision plus large et un corpus plus riche et nuancé que le questionnaire (e.g., Combessie, 2007). Par exemple, nous avons pu mettre en exergue la thématique du risque perçu jusque-là absente de la littérature scientifique (cf. Discussion générale, a. La prise en compte du risque perçu dans la satisfaction, p. 275). Ces analyses ont également permis de mettre en évidence les incompréhensions liées à l'utilité-destination ce qui impacte l'utilisabilité du dispositif (cf. Discussion générale, b. L'impact de l'utilité perçue sur la satisfaction, p. 276). Cependant, l'analyse des entretiens et focus group est une tâche fastidieuse et complexe. La difficulté des analyses des entretiens réside dans le travail d'interprétation et de conceptualisation de ces données. Le risque est de tomber dans une analyse uniquement descriptive (Richards, 2015). Afin d'éviter cette erreur, plusieurs méthodes de traitement des entretiens existent (e.g., analyse thématique, straussienne, constructiviste... ; Jacques et al., 2014). Pour notre part, nous avons opté pour un traitement par analyse thématique de Braun et Clarke (2006) en six étapes.

- **Think aloud (direct)** permet d'avoir des informations sur les mesures d'efficacité et plus exactement de comprendre la survenue des erreurs et la manière de résoudre ces erreurs. Par exemple, lors de l'étude 2 (cf. Partie II, Chapitre 4), lorsqu'un utilisateur s'arrêtait de manipuler, nous aurions pu interpréter cela comme une pause. Pourtant, grâce à la méthode *think aloud*, l'utilisateur a expliqué qu'il était à la recherche d'un élément sur l'interface de la tablette.

Cependant, cette méthode s'avère très contraignante pour l'utilisateur. En effet, lors de l'étude 2 (cf. Partie II, Chapitre 4), nous avons dû plusieurs fois rappeler la consigne de « *penser à voix haute* » à l'utilisateur qui oubliait rapidement de préciser ce qu'il faisait. Cette méthode, bridant l'utilisateur, biaise les mesures d'efficacité, car ce dernier doit penser à détailler ses actions en plus de les réaliser correctement. C'est pourquoi le recours à cette méthode ne nous permet pas de comparer les mesures d'efficacité de l'étude 2 (cf. Partie II, Chapitre 4) avec les autres études expérimentales.

- **Le questionnaire** est la méthode la plus référencée dans notre revue systématique sur les cadres méthodologiques de l'évaluation de l'utilisabilité de dispositif connecté en santé pour le domicile patient (cf. Partie I, Chapitre 2, 3.3.3. Les outils, p. 75), plusieurs questionnaires d'utilisabilité existent (e.g., SUS ; UMUX ; UMUX-LITE ; NASA&TLX ; SMEQ). Pour notre part, nous avons privilégié le *System Usability Scale* (SUS) de Brooke (1996) pour sa rapidité d'administration (10 items). Cet outil permet d'obtenir un score « d'utilisabilité perçue » sur 100 (Lewis & Sauro, 2009). En fonction de ce score, il est possible de classer le dispositif évalué (e.g., Bangor et al., 2008). Cet outil est également intéressant afin de réaliser des comparaisons entre les évaluations de l'utilisabilité. Ainsi, cet outil a permis de révéler une progression de l'utilisabilité perçue entre les différentes études expérimentales. Lors de l'étude 1, évaluant les objets connectés du dispositif, nous observons un score moyen de 72,4 (ET = 21,07), pour l'étude 2, évaluant le prototype 1, un score moyen de 84,13 (ET = 11,45), enfin pour l'étude 3, évaluant le prototype 2, un score moyen de 87,63 (ET = 11,35). Cette progression montre les apports des différentes modifications apportées au dispositif au contexte d'usage. Cet outil est également utile afin de réaliser des corrélations ou régressions avec d'autres variables quantitatives (e.g., littératie en santé, âge...)

En revanche, ce type d'outil ne peut, en aucun cas, se soustraire à d'autres méthodes d'évaluation de l'utilisabilité. Comme nous venons de le voir, cet outil donne accès à

un score global et ne permet pas de cibler avec précision et exhaustivité les problèmes d'utilisabilité.

- **La tâche de rappel**, utilisée lors de l'étude 1b (cf. Partie II, Chapitre 3), permet de vérifier la bonne intégration des instructions ainsi que de vérifier les potentielles erreurs de manipulation pouvant être réitérées lors d'autres essais. En effet, si un utilisateur énonce mal les procédures, il est possible que celui-ci oublie ou confonde certaines instructions cruciales lors d'une prochaine utilisation. Or, l'étude 3 (cf. Partie II, Chapitre 5) a montré que les utilisateurs lisaient encore moins les instructions après la première utilisation et de nouvelles erreurs de manipulation pouvaient survenir. Ainsi, la tâche de rappel est une méthode efficace afin d'analyser les éléments non retenus par les utilisateurs.

Cependant, cette méthode semble peu fiable. En effet, lors de l'étude 1b (cf. Partie II, Chapitre 3), nous avons montré que les utilisateurs ayant les procédures sous format image retenaient significativement moins les notices. Cela était dû au fait que les utilisateurs n'avaient pas accès aux termes employés dans les instructions comme cela pouvait être le cas pour les formats écrits ou audios. De plus, la cotation permettant de mesurer la mémorisation peut également être remise en question. Nous avons utilisé une grille de cotation gratifiant davantage les utilisateurs précis dans leurs réponses. Par exemple, l'utilisateur précisant l'emplacement du bouton sur lequel il devait appuyer gagnait un point supplémentaire par rapport à celui qui ne le précisait pas. Ces limites ne permettent pas d'avoir une réelle mesure de la mémorisation.

Ainsi pour conclure, deux combinaisons de méthodes et d'outils nous semblent particulièrement pertinentes pour une évaluation de l'utilisabilité en milieu contrôlé :

1. L'observation directe discrète associée à de l'observation instrumentée par vidéo couplée à des entretiens (ou à des focus groups) dans le cas d'un prototype abouti.
2. Le *Think aloud* rétrospectif associé à de l'observation instrumentée par vidéo couplée à des entretiens dans le cas d'un prototype peu abouti.

Ces deux combinaisons permettent la collecte de données qualitatives et quantitatives ainsi que subjectives et objectives sur l'ensemble des métriques de l'utilisabilité. De plus, les observations directes permettent de nourrir les entretiens et les données vidéos et d'obtenir de solides informations sur l'efficacité et l'efficience. D'autres méthodes et outils existent dans la littérature et pourraient compléter notre retour d'expérience. Nous pensons, par exemple, au

traceur d'activité ou au journal de bord qui semblent être deux méthodes pertinentes pour évaluer l'utilisabilité à distance dans un contexte *in situ* que nous n'avons pas pu mettre en place dans cette thèse. Nous abordons ces deux méthodes dans les perspectives (cf. Discussion générale, 4.1.1. Les études ergonomiques *in situ* intégrées à l'étude clinique en cours, p. 292).

2.2.2. L'importance de la mise en situation : les personas et la simulation

a. L'intérêt des personas en phase d'évaluation

Afin d'aider les participants à se projeter dans l'utilisation du dispositif *Smart Angel* lors de l'étude 2, nous avons réutilisé les personas utilisés généralement en tant que méthode prospective (Bornet & Brangier, 2013; Lallemand & Gronier, 2015). Lors de l'étude 2 (cf. Partie II, Chapitre 4), nous avons pu observer des réactions très positives des utilisateurs se mettant à la place des personas présentés afin d'expliquer leur ressenti et leur perception du dispositif *Smart Angel*. Ainsi, le recours à des personas semble également prometteur en phase d'évaluation.

Même si les personas ont pu prouver leur efficacité (e.g., Brangier et al., 2012), il est important de garder à l'esprit la limite principale de cette méthode : sa validité. Bornet et Brangier (2013) déclarent que « *cette démarche se démarque donc des procédures classiques, plus rigoureuses et linéaires, ce qui soulève des questions au niveau de la validité des personas, de la méthodologie associée et de la manière dont les personas sont perçus par les concepteurs* » (Bornet & Brangier, 2013, p. 119). Ainsi, la validité des personas peut être remise en question du fait d'une reproductibilité complexe.

b. L'intérêt de la simulation en phase amont de conception

La simulation, intégrant certains éléments du réel, que nous avons mise en place dans l'étude 3 (cf. Partie II, Chapitre 5), bien que partielle, a permis aux patients/participants de se projeter dans l'utilisation du dispositif *Smart Angel*. En effet, le lieu réel (l'hôpital) et la présence du personnel hospitalier ont permis aux utilisateurs de se sentir immergés dans un environnement hospitalier particulièrement angoissant pour certains. Lors des entretiens, les utilisateurs ont déclaré se forcer à se rappeler qu'ils étaient dans une simulation et non dans un réel parcours chirurgical. La plupart ayant déjà vécu un séjour hospitalier ont déclaré revivre cet événement pendant la simulation. De plus cette méthode, bien qu'initiant une démarche expérimentale et contrôlée est un moyen pertinent de se rendre compte des problèmes d'utilisabilité lors de l'intégration du matériel (Kushniruk et al., 2013). Ainsi, la simulation en

milieu hospitalier semble être une alternative intéressante pour les évaluations de l'utilisabilité et également pour des études portant sur l'acceptabilité/acceptation en phase amont de conception du produit. En effet, l'étude 3 (cf. Partie II, Chapitre 5) a servi de base à la construction d'une grille d'observation pour des études de terrain, portant sur l'impact de l'intégration du dispositif sur l'organisation hospitalière, menées par Cyril Moureau, postdoctorant sur le projet *Smart Angel*.

Cependant, la simulation d'un service entier (comme celui de la chirurgie ambulatoire) reste très complexe à mettre en œuvre, tout comme la réelle prise en compte des délais entre la prise de rendez-vous et l'opération. Cela nécessite de recruter et de former le personnel hospitalier à la simulation et au dispositif *Smart Angel* mais cela demande également une importante organisation en termes de logistique afin de gérer les salles, les patients et le personnel.

3. Contraintes de la thèse : Retour d'expérience sur un projet de conception pluripartenarial

3.1. Les recommandations ergonomiques confrontées aux contraintes industrielles : un projet à deux vitesses

Nos résultats ont montré que le nombre d'erreurs de manipulation n'avait pas particulièrement baissé au fur et à mesure des études. Nous avons pu observer que ces erreurs pouvaient être les mêmes d'une étude à l'autre, ce qui indique que nous n'avons probablement pas proposé les bonnes recommandations, ou qu'elles n'ont pas été prises en compte par l'équipe de développeurs. Pourtant, après chaque étude, nous réalisons une restitution des résultats et déposons un livrable de recommandations de conception pour le prochain prototype, auprès du partenaire industriel.

Nous avons alors rédigé un rapport des différentes recommandations (cf. Annexe 22 : Restitution des recommandations avec l'entreprise Evolucare, p. 414) en fonction des études réalisées et nous avons pu observer que la plupart n'ont pas été intégrées dans la dernière version du prototype (e.g., clarté des schémas, taille de la police, choix du tensiomètre...). Nous nous sommes alors interrogées sur la bonne prise en compte de ces recommandations. Sont-elles toujours comprises par les développeurs ? Sont-elles trop complexes et/ou trop coûteuses

à mettre en place ? Sont-elles en inadéquation avec la culture de l'entreprise et/ou du cahier des charges initialement prévus par l'entreprise ?

Nous avons interrogé l'entreprise lors d'une réunion de restitution, afin de comprendre leurs attentes et leurs projets par rapport à ces recommandations proposées mais non intégrées. En Annexe (cf. Annexe 22 : Restitution des recommandations avec l'entreprise Evolucare, p. 414), nous avons listé les commentaires de nos partenaires industriels par rapport à chaque recommandation énoncée. Nous avons noté que la plupart des réponses de nos partenaires à la non-prise en compte des recommandations était « *c'est une recommandation importante, nous allons le faire* ». L'entreprise semble en ce sens convaincue de l'importance d'implanter ces modifications majeures au sein des futures mises à jour du dispositif, cependant, ces recommandations nécessitaient un coût humain qu'elle ne pouvait se permettre dans la temporalité de la thèse. Les modifications demandées étaient trop importantes par rapport au nombre d'employés disponibles sur le projet.

Le déploiement de nos études, pourtant retardé par l'élaboration des prototypes, a été plus rapide que le développement de l'application. La planification initiale du projet intégrant le calendrier d'une thèse, pourtant établie initialement avec l'accord de l'ensemble des partenaires, a été durement retardée au fur et à mesure de l'avancement du projet. Par exemple, nous n'avons pas pu intégrer les résultats d'études *in situ* dans ce manuscrit bien que nous soyons co-auteur du protocole de recherche (cf. Discussion générale, 4.1.1. Les études ergonomiques *in situ* intégrées à l'étude clinique en cours (court terme), p. 292). Nous avons par conséquent dû nous orienter vers des études expérimentales dans des milieux contrôlés (e.g., la simulation) afin de mener à bien ce travail de thèse. Les résultats de ces études ont permis de proposer des recommandations utiles à l'entreprise pour le développement des prototypes du système, la réflexion autour de l'implantation du dispositif ainsi que la mise en pratique de la formation pour le personnel hospitalier.

3.2. Le retard du projet *Smart Angel*

Le projet *Smart Angel* comptabilise aujourd'hui environ deux années et trois mois de retard sur le planning initial. Ce retard a été provoqué par quatre éléments majeurs.

Premièrement, l'application *Smart Angel* n'a pas été développée par l'entreprise dans la temporalité initialement prévue. Cette absence de développement est due à un manque de ressources humaines chez Evolucare, mais également aux manques d'informations sur le

potentiel contenu de l'application remonté par le CHU de Nîmes. Midler et al. (1997, p. 10) décrivent ce phénomène comme « *la mise au point d'un ajustement harmonieux* ». Il s'agit d'un temps d'adaptation et d'ajustement entre les différents partenaires nécessaires à chaque conception dans des projets d'open innovation.

Deuxièmement, la rédaction et la validation du Comité de Protection des Personnes ou CPP, encadré par la loi Jardé issue du code de la santé publique³⁸, ont été plus longues que prévu. Cette difficulté projectariale est relativement connue dans le système français (e.g., Peute et al., 2020). Ce CPP a pu être validé en décembre 2018 au lieu d'avril 2018.

Troisièmement, les financements prévus (en particulier pour le CHU de Nîmes) pour mener les études cliniques du projet *Smart Angel* n'ont pas été débloqués dans les temps bien que le CPP de l'étude ambulatoire avait été approuvé. Ce retard de financement est dû à des difficultés réglementaires sur le deuxième projet *Smart Angel* (intra-hospitalier). Pour rappel, *Smart Angel* est divisé en deux projets indépendants, mais co-dépendants financièrement. Afin de débloquent les fonds nécessaires aux études cliniques, les deux CPP devaient être validés. Or, actuellement, le CPP n'ayant pas été déposé pour ce deuxième projet, aucun fond n'a été débloquent pour financer l'étude clinique multicentrique finalement lancée timidement sur un seul centre hospitalier en mars 2020.

Quatrièmement, le COVID-19 a stoppé l'étude clinique qui a repris en juin 2020. L'entreprise Evolucare a adapté ses plannings pour le développement de l'application en ambulatoire retardant alors les travaux du CRP-CPO.

L'ensemble de ces retards a poussé les responsables projets à revoir les plannings. Initialement, il devait y avoir deux études cliniques en ambulatoire. Au vu des contraintes temporelles et financières, la deuxième étude a été annulée. En fonction des aléas du projet, il a fallu adapter le travail de thèse et la démarche méthodologique. En effet, nos expérimentations auraient dû être intégrées aux évaluations cliniques (cf. Discussion générale, 4.1.1. Les études ergonomiques *in situ* intégrées à l'étude clinique en cours, p. 292). Nous présentons le planning prévu pour la suite de ce travail de thèse en Figure 65.

³⁸ Code de la Santé Publique, Livre Ier, Titre II : Recherches impliquant la personne humaine. Loi Jardé n° 2012-300 du 5 mars 2012 modifiée par ordonnance n° 2016-800 du 16 juin 2016, décret d'application modificatif n° 2017-884 du 9 mai 2017 et textes subséquents incluant les arrêtés du 12 avril 2018.

3.1. La hiérarchisation du partenariat de co-développement

Nous avons vu en introduction générale de ce travail de thèse (cf. Introduction Générale, 2.2. Un projet de co-développement : les partenaires industriels, académiques et hospitaliers, p. 4) que le projet *Smart Angel* peut être qualifié de partenariat de conception ou de co-développement (Midler et al., 1997, 2007). Pourtant, les membres du CRP-CPO ont très peu été en contact direct avec les autres partenaires du projet tel que le CHU de Nîmes. Cela n'a pas aidé à créer un lien privilégié avec le CHU éloignant le terrain des études ergonomiques. Nous faisons remonter systématiquement nos résultats, nos publications ou nos besoins à l'entreprise qui se chargeait ensuite de les diffuser aux autres partenaires. Nous n'étions pas régulièrement conviées aux réunions sur le pilotage du projet entre Evolucare et le CHU de Nîmes alors que nous étions co-auteur du CPP.

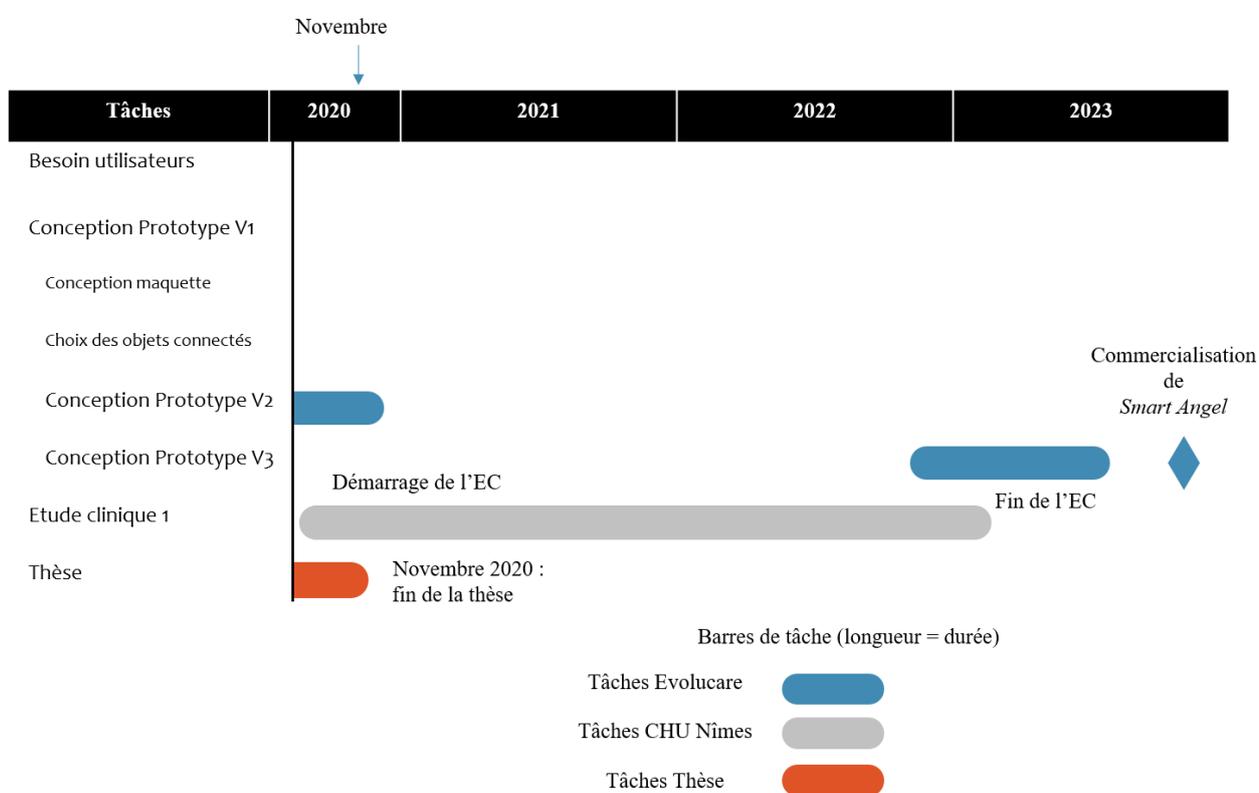


Figure 65. Suite et fin du processus de conception de *Smart Angel* après la soutenance de thèse.

3.2. Les évaluations ergonomiques pendant les études cliniques

À l'origine du projet, des évaluations ergonomiques auraient dû se dérouler pendant les études cliniques conduites au CHU de Nîmes permettant d'avoir un retour *in situ* du matériel. De plus, il faut savoir que l'étude clinique n'est pas menée par le personnel hospitalier de chirurgie ambulatoire, mais par une équipe de recherche en charge de former les patients,

distribuer, stocker et désinfecter le matériel. Le temps alloué par ce personnel de recherche pour former les patients n'est peut-être pas représentatif du temps réel dont dispose le personnel hospitalier en situation *in situ*. **De ce fait, à aucun moment, le matériel ne pourra être testé dans un environnement parfaitement *in situ* avant sa commercialisation.** Au vu du retard de la mise en place de l'étude clinique, nous n'avons pas pu l'inclure dans notre travail de thèse. Cependant, en tant que co-auteur du CPP, nous avons intégré des outils permettant l'évaluation de l'utilisabilité dans ce contexte clinique. Il existe par ailleurs une contrainte non négligeable à ce protocole : nous ne pouvons à aucun moment être en contact avec le patient. Nous avons dû envisager des solutions méthodologiques alternatives pour avoir des indications sur l'utilisabilité du dispositif *Smart Angel*. Nous présentons les outils et méthodes mises en place dans les perspectives de recherche ci-dessous.

4. Perspectives de recherche

Les perspectives proposées ont pour ambition d'améliorer le modèle HANDLED. Nous avons divisé cette section en deux parties : les perspectives faisant suite au projet *Smart Angel*, et les perspectives extérieures au projet.

4.1. Perspectives issues du projet *Smart Angel*

4.1.1. Les études ergonomiques *in situ* intégrées à l'étude clinique en cours (court terme)

L'étude clinique initiée par le CHU de Nîmes, actuellement en cours, permettra de collecter des données cliniques et ergonomiques probablement d'ici 2022. Cette étude clinique multicentrique devrait intégrer jusqu'à 1000 patients en ambulatoire. En effet, nous avons intégré au CPP, validé en décembre 2018, un protocole pour l'évaluation ergonomique du dispositif *Smart Angel* permettant la collecte de données *in situ*. Pour obtenir des données sur l'efficience et l'efficacité, nous avons déjà développé sur mesure, avec l'aide de nos partenaires industriels (Evolucare), un traceur d'activité. Les informations recueillies permettront d'avoir des indications sur le nombre de connexions, le temps passé sur chaque tâche (entre chaque étape), le nombre de clics à l'écran sans gêner l'utilisateur. En revanche, il est impossible d'avoir des informations sur les données physiologiques relevées. Afin d'obtenir des données

sur la satisfaction, nous avons conçu un journal de bord. Chaque page du journal correspond à une journée de convalescence. Le patient devra indiquer s'il a eu besoin d'aide pendant la manipulation, si quelqu'un d'autre que lui a utilisé le matériel et s'il a eu des difficultés dans la manipulation. Les données du journal et du traceur d'activité seront ensuite liées par l'identifiant patient.

Pour bâtir ce protocole, nous nous sommes appuyées sur nos données issues de notre analyse systématique portant sur les cadres méthodologiques de l'évaluation de l'utilisabilité de dispositif en santé pour le domicile patient (cf. Partie I, Chapitre 2, 3.3.3. Les outils, p. 75). Nous avons recensé que les principaux outils utilisés, pour ce type d'étude *in situ*, étaient le questionnaire, l'entretien et le traceur d'activité. Ne pouvant pas mettre au point des entretiens, nous avons opté pour le journal de bord.

Ces données permettront de garantir la validité écologique (e.g., Rochat et al., 2017) de nos études 2 et 3. En effet, nous pourrions généraliser nos résultats, en comparant nos analyses issues de ces études, à ceux de l'étude clinique, grâce aux retours d'expérience des patients inclus dans l'étude clinique.

Les évaluations à distance ont l'avantage de capter l'usage dans son environnement naturel (e.g., Jambon et al., 2010), composant le niveau 4 de notre modèle d'évaluation de l'utilisabilité. Ces évaluations peuvent intégrer les comportements inattendus comme les utilisations détournées. De plus, les jugements des participants (indicateur de la satisfaction) ne font plus partie de la sphère projective, mais sont ancrés dans le réel et le présent. Par exemple, cela permettrait de prendre en compte des éléments du contexte hospitalier que nous n'avons pas pu investiguer, pendant ce travail de thèse, comme l'impact de la douleur ou d'une gêne issues de l'opération sur l'utilisabilité.

Néanmoins, ces évaluations possèdent une limite majeure : elles ne permettent pas de mesurer l'efficacité du matériel. L'absence de mesure de cette métrique, jugée pourtant cruciale dans les évaluations de l'utilisabilité, est liée au fait qu'il est très difficile de juger de la situation s'il n'est pas possible de l'observer directement. En effet, nous ne pouvons pas mesurer le nombre d'abandons puisque plusieurs raisons peuvent expliquer qu'un patient n'utilise pas le matériel. De la même manière, cette évaluation ne permet pas d'évaluer le nombre d'erreurs de manipulation. Nous ne pouvons que prospecter sur une potentielle erreur liée à un temps de manipulation plus long entre deux étapes, mais de nouveau, plusieurs explications peuvent aussi interférer sur cet indicateur. Du fait de cette limite, nous pensons que les études *in situ* ne

peuvent pas se substituer aux études en milieu contrôlé. Il est nécessaire d’avoir une vision holistique de l’efficacité ainsi qu’une liste des erreurs de manipulation du matériel, avant de proposer une étude dans un environnement *in situ*. Cela corrobore notre opinion sur la complémentarité entre les études *in situ* et les études contrôlées en laboratoire (cf. Partie I, Chapitre I, 4.2. Trois lieux de recueil de données déterminés par trois approches épistémologiques : laboratoire, terrain, *in situ*, p. 43).

4.1.2. Le niveau 5 du modèle HANDLED : Les changements organisationnels et relationnels dus à l’intégration du dispositif *Smart Angel* (long terme)

Le modèle HANDLED pourrait contenir un niveau 5 (Figure 66), c’est-à-dire une évaluation de l’utilisabilité dans un usage routinier comme indiquée dans le modèle de Yen et Bakken (2012). Pour cela, il faudra attendre la commercialisation du dispositif *Smart Angel* afin d’observer cette routine. Ce niveau 5 permettrait également de rendre compte de l’impact des modifications relationnelles et organisationnelles au sein d’un centre hospitalier sur l’utilisabilité du dispositif *Smart Angel* lors de son intégration.

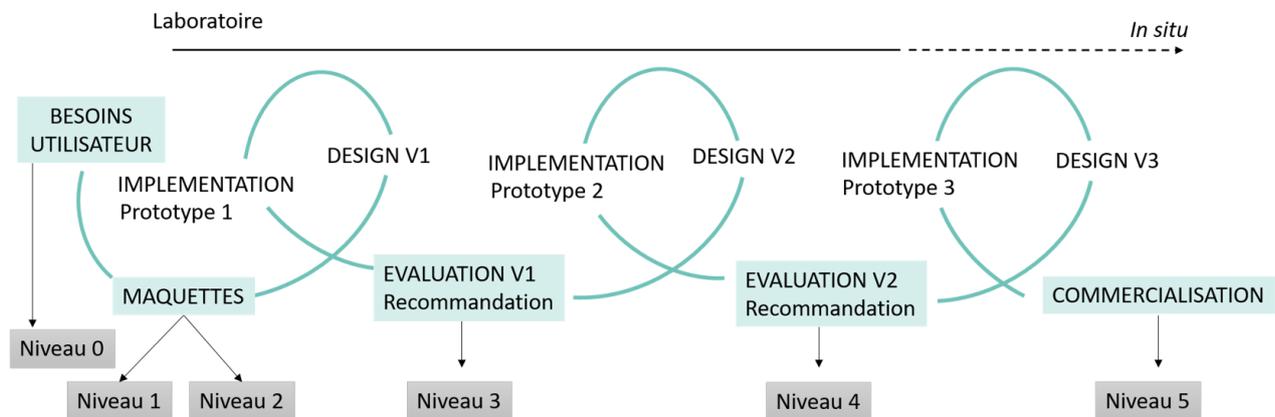


Figure 66. Niveau 5 du processus de conception du dispositif *Smart Angel*.

L’observation de ces modifications dues à l’intégration du dispositif, est particulièrement intéressante afin d’analyser, premièrement, la façon dont il modifie les relations de soin entre le patient lui-même avec sa maladie. En effet, les technologies en santé promettent de responsabiliser le patient sur sa santé (Eysenbach, 2001). Cette responsabilisation permettrait d’augmenter « l’activation » et « l’engagement » du patient dans sa santé (Greene & Hibbard, 2012). Barcello et al. (2013) ont montré que des patients engagés ont davantage de chance de guérir de leur maladie puisque ces derniers la comprennent. Ainsi, le patient réalisant des actes

médicaux pratiqués anciennement par le personnel soignant sur lui-même, le patient sera davantage investi dans sa guérison et le personnel soignant aura également plus de temps à consacrer à d'autres patients. Ce gain de temps serait alors prometteur afin d'améliorer la relation de soin patient/soignant (Eysenbach, 2001). Cet engagement n'a, actuellement, jamais été mesuré dans nos études et pourrait avoir un impact positif sur l'utilisabilité puisque le patient pourrait être davantage investi dans sa réussite (cf. Discussion générale, b. Faut-il privilégier l'efficacité au détriment de l'efficience ?, p. 273).

Dans un deuxième temps, l'intégration du dispositif *Smart Angel* modifierait également la relation patient/soignant due aux transformations des conditions de travail du personnel soignant. Comme nous avons pu l'observer dans l'étude exploratoire (cf. Partie II, Chapitre 1) le dispositif *Smart Angel* permettrait, à terme, d'alléger le poste du personnel en charge des appels dans le service de chirurgie ambulatoire et préviendrait les réhospitalisations ainsi que les consultations extrahospitalières. Il serait donc intéressant, lors de la commercialisation du dispositif d'observer le lien patient/soignant avant et après la mise en service du produit. Des données quantitatives sur l'impact de l'intégration du dispositif sur les réhospitalisations pourront être disponibles lors des retours de l'étude clinique.

4.2. Transfert du modèle HANDLED vers d'autres conceptions de dispositifs connectés en santé pour le domicile patient : Étude sur l'utilisation de dispositifs médicaux connectés dans le cas de mesure ambulatoire de la pression artérielle (court terme)

Suite à la collaboration entre Emilie Loup-Escande et Maxime Gignon, Professeur de médecine au CHU d'Amiens dans le cadre de l'encadrement de la thèse d'exercice en médecine de Assia Cherfouh, médecin généraliste, nous avons proposé une étude afin de vérifier la capacité de notre modèle HANDLED (Figure 63) à être opérationnel sur d'autres contextes que celui de la chirurgie ambulatoire. Cette thèse d'exercice a pour objectif d'évaluer l'acceptabilité sociale et située d'un tensiomètre connecté dans un contexte de Mesure Ambulatoire de la Pression Artérielle ou MAPA auprès des médecins généralistes en milieu rural. Assia Cherfouh proposerait alors les tensiomètres issus du projet *Smart Angel* et l'application Android du concepteur originel (ihealth) à des patients en attente d'un diagnostic d'hypertension. Ces futurs utilisateurs du tensiomètre connecté doivent réaliser trois mesures de tension dans une journée avant de retourner le matériel. Dans ce cadre, nous avons proposé un carnet de bord qui sera

distribué aux patients lors de la consultation avec le généraliste, afin d'appréhender l'utilisabilité du tensiomètre connecté et de son application dans le contexte MAPA. L'étude aurait dû débuter en mars 2020. En raison de la crise sanitaire, elle a été retardée.

5. Synthèse de la Discussion générale

Ce travail de thèse avait pour objectif d'**assister la conception par l'évaluation de l'utilisabilité du dispositif connecté en santé *Smart Angel* dans un contexte post-ambulatoire. Afin de répondre à notre problématique, nous avons implémenté sur 5 études, par niveaux cumulatifs, le modèle HANDLED. Cette discussion comprend trois parties : les contributions, les contraintes et les perspectives.**

Premièrement, nous avons discuté des contributions de ce modèle. Ce travail de thèse met en avant plusieurs apports théoriques :

- Évaluer chaque métrique de l'utilisabilité de manière holistique, en se basant sur plusieurs indicateurs (e.g., nombre d'abandons et nombre d'erreurs de manipulation), en particulier l'efficacité, en précisant le type d'erreurs (« bloquantes » et « non-bloquantes ») ;
- Encourager l'engagement cognitif chez l'utilisateur afin d'optimiser l'efficacité au détriment de l'efficience lors de la première utilisation ;
- Intégrer des mesures du risque perçu et de l'utilité perçue dans les mesures de satisfaction.
- Proposer des études cumulatives lors d'évaluation de l'utilisabilité, intégrant de manière systématique des éléments du contexte d'usage (*tâches, utilisateurs, ressources, système, environnements*) sur différents niveaux.

Cette thèse apporte également des contributions méthodologiques :

- Différencier les méthodes d'observation directe visible de l'observation directe invisible, de même que pour les méthodes d'observation instrumentée par vidéo visible de l'observation instrumentée par vidéo discrète ;
- Intégrer des personas en phase d'évaluation ainsi que de la simulation en phase amont de conception.

Deuxièmement, nous avons exposé les contraintes de cette thèse dans lequel nous argumentons sur le retard du projet *Smart Angel* ainsi que l'absence de contact privilégié avec nos partenaires hospitaliers, l'impossibilité de réaliser d'étude *in situ* et ses répercussions sur les choix méthodologiques de ce travail de thèse.

Troisièmement, nous présentons les protocoles de recherche de deux futures études actuellement en cours qui apporteront des données concrètes dans les 2 à 3 années à venir : une étude clinique portant sur le dispositif *Smart Angel* et une étude permettant de valider le transfert du modèle HANDLED vers un contexte différent de celui de la chirurgie ambulatoire, celui du diagnostic de l'hypertension artérielle. Afin d'améliorer ce modèle, nous proposons également un cinquième niveau correspondant à l'usage routinier du produit (après sa commercialisation) et à l'intégration de nouvelles caractéristiques dans le contexte d'usage tel que les changements organisationnels et relationnels comprenant l'impact des relations patients/soignants ainsi que l'engagement du patient dans sa santé sur l'utilisabilité du dispositif *Smart Angel*.

BIBLIOGRAPHIE

- Abowd, G., & Mynatt, E. (2000). Charting Past, Present and Future Research in Ubiquitous Computing. *ACM Transactions on Computer-Human Interaction*, 7(1), 29-58.
- Abran, A., Khelifi, A., Suryan, W., & Seffah, A. (2003). Usability Meanings and Interpretations in ISO Standards. *Software Quality Journal*, 11(4), 325-338. <https://doi.org/10.1023/A:1025869312943>
- Abts, N., Hernandez, A., Caplan, S., Hettinger, A. Z., Larsen, E., & Lewis, V. R. (2014). When human factors and design unite: using visual language and usability testing to improve instructions for a home-use medication infusion pump. *Proceedings of the International Symposium on Human Factors and Ergonomics in Health Care*, 3(1), 254-260. <https://doi.org/10.1177/2327857914031041>
- AFCA. (2015). *Chirurgie ambulatoire & PMSA* (p. 13). http://www.chirurgie-ambulatoire.org/uploads/6/4/6/4/64646507/chirurgie_ambulatoire_pmsi_oct2015.pdf
- Agarwal, R., & Prasad, J. (1998). A Conceptual and Operational Definition of Personal Innovativeness in the Domain of Information Technology. *Information Systems Research*, 9(2), 204-215. <https://doi.org/10.1287/isre.9.2.204>
- Agnisarman, S. O., Chalil Madathil, K., Smith, K., Ashok, A., Welch, B., & McElligott, J. T. (2017). Lessons learned from the usability assessment of home-based telemedicine systems. *Applied Ergonomics*, 58, 424-434. <https://doi.org/10.1016/j.apergo.2016.08.003>
- Alam, I. (2004). Interacting with Customers in the New Product Development Process. In K. B. Kahn (Éd.), *The PDMA Handbook of New Product Development* (1^{re} éd., p. 249-262). Wiley. <https://doi.org/10.1002/9780470172483.ch16>
- Alanzi, T., Istepanian, R., & Philip, N. (2016). Design and Usability Evaluation of Social Mobile Diabetes Management System in the Gulf Region. *JMIR Research Protocols*, 5(3), e93. <https://doi.org/10.2196/resprot.4348>
- Albrecht, L., Wood, P. W., Fradette, M., McAlister, F. A., Rabi, D., Boulanger, P., & Padwal, R. (2018). Usability and Acceptability of a Home Blood Pressure Telemonitoring Device Among Community-Dwelling Senior Citizens With Hypertension: Qualitative Study. *JMIR Aging*, 1(2), e10975. <https://doi.org/10.2196/10975>
- Aljaber, T., & Gordon, N. (2017). A Guidance and Evaluation Approach for mHealth Education Applications. In P. Zaphiris & A. Ioannou (Éds.), *Learning and Collaboration Technologies. Technology in Education* (p. 330-340). Springer International Publishing. https://doi.org/10.1007/978-3-319-58515-4_26
- Allaert, F. (2016). Ethical and Technological Evaluation of Security of Connected Objects and Health Applications: How to Move Towards Certification. *Value in Health*, 19(7), A507-A508. <https://doi.org/10.1016/j.jval.2016.09.935>

- Allaert, F.-A., Mazen, N.-J., Legrand, L., & Quantin, C. (2016). Les enjeux de la sécurité des objets connectés et applications de santé. *Journal de gestion et d'économie médicales*, 34(5), 311. <https://doi.org/10.3917/jgem.165.0311>
- Alnasser, A., Kyle, J., Alkhalifah, A., & Marais, D. (2018). Relationship Between Evidence Requirements, User Expectations, and Actual Experiences: Usability Evaluation of the Twazon Arabic Weight Loss App. *JMIR Human Factors*, 5(2), e16. <https://doi.org/10.2196/humanfactors.9765>
- Alnosayan, N., Chatterjee, S., Alluhaidan, A., Lee, E., & Feenstra, L. H. (2017). Design and Usability of a Heart Failure mHealth System: A Pilot Study. *JMIR Human Factors*, 4(1), e9. <https://doi.org/10.2196/humanfactors.6481>
- Alonso-Ríos, D., Vázquez-García, A., Mosqueira-Rey, E., & Moret-Bonillo, V. (2009). Usability: A Critical Analysis and a Taxonomy. *International Journal of Human-Computer Interaction*, 26(1), 53-74. <https://doi.org/10.1080/10447310903025552>
- Ammenwerth, E., Gräber, S., Herrmann, G., Bürkle, T., & König, J. (2003). Evaluation of health information systems-problems and challenges. *International Journal of Medical Informatics*, 71(2-3), 125-135. [https://doi.org/10.1016/s1386-5056\(03\)00131-x](https://doi.org/10.1016/s1386-5056(03)00131-x)
- Ammenwerth, E., Mansmann, U., Mahler, C., Kandert, M., & Eichstädter, R. (2002). Are quantitative methods sufficient to show why wards react differently to computer-based nursing documentation? *Studies in Health Technology and Informatics*, 90, 377-381.
- Anderson, J. R. (1982). Acquisition of cognitive skill. *Psychological Review*, 89(4), 369-406. <https://doi.org/10.1037/0033-295X.89.4.369>
- Anderson, J. R. (1983). A spreading activation theory of memory. *Journal of Verbal Learning and Verbal Behavior*, 22(3), 261-295. [https://doi.org/10.1016/S0022-5371\(83\)90201-3](https://doi.org/10.1016/S0022-5371(83)90201-3)
- Andersen, S. B., Rasmussen, C. K., & Frøkjær, E. (2017). Bringing Content Understanding into Usability Testing in Complex Application Domains—a Case Study in eHealth. In A. Marcus & W. Wang (Eds.), *Design, User Experience, and Usability: Theory, Methodology, and Management* (pp. 327-341). Springer International Publishing. https://doi.org/10.1007/978-3-319-58634-2_25
- Aoussat, A., Christofol, H., & Coq, M. L. (2000). The new product design - a transverse approach. *Journal of Engineering Design*, 11(4), 399-417. <https://doi.org/10.1080/09544820010000971>
- Arora, S., Moclair, B., Murphy, K., Jaffe, J. S., & Kaminetsky, J. C. (2018). Summative Usability Evaluation of the SCTE-AI Device: A Novel Prefilled Autoinjector for Subcutaneous Testosterone Administration. *The Journal of Sexual Medicine*, 15(12), 1707-1715. <https://doi.org/10.1016/j.jsxm.2018.09.017>
- ATIH. (2016a). *Etat des lieux 2015 sur l'activité de chirurgie ambulatoire : Programme national chirurgie ambulatoire*. https://www.atih.sante.fr/sites/default/files/public/content/1504/note_gdr_chirambu_2

- ATIH. (2016b). *Rapport d'Activité 2016* (p. 52). Agence Technique de l'Information sur l'Hospitalisation.
https://www.atih.sante.fr/sites/default/files/public/content/2196/rapport_dactivite_atih_2016.pdf
- Aubert, B. A., & Patry, M. (2004). *Les partenariats public-privé : une option à découvrir* (N° 2004rb-02; CIRANO Burgundy Reports). CIRANO.
<https://ideas.repec.org/p/cir/cirbur/2004rb-02.html>
- Ayubi, S. U. A., Parmanto, B., Branch, R., & Ding, D. (2014). A Persuasive and Social mHealth Application for Physical Activity: A Usability and Feasibility Study. *JMIR MHealth and UHealth*, 2(2), e25. <https://doi.org/10.2196/mhealth.2902>
- Baber, C., & Stanton, N. A. (1996). Human error identification techniques applied to public technology: predictions compared with observed use. *Applied Ergonomics*, 27(2), 119-131. [https://doi.org/10.1016/0003-6870\(95\)00067-4](https://doi.org/10.1016/0003-6870(95)00067-4)
- Baddeley, A. (1986). *Working memory*. Clarendon Press/Oxford University Press.
- Baddeley, A. (1992). Working memory. *Science*, 255(5044), 556-559.
<https://doi.org/10.1126/science.1736359>
- Baker, D. W., Gazmararian, J. A., Williams, M. V., Scott, T., Parker, R. M., Green, D., Ren, J., & Peel, J. (2002). Functional Health Literacy and the Risk of Hospital Admission Among Medicare Managed Care Enrollees. *American Journal of Public Health*, 92(8), 1278–1283. <https://doi.org/10.2105/AJPH.92.8.1278>
- Baker, D. W., Williams, M. V., Parker, R. M., Gazmararian, J. A., & Nurss, J. (1999). Development of a brief test to measure functional health literacy. *Patient Education and Counseling*, 38(1), 33-42. [https://doi.org/10.1016/s0738-3991\(98\)00116-5](https://doi.org/10.1016/s0738-3991(98)00116-5)
- Bangor, A., Kortum, P. T., & Miller, J. T. (2008). An Empirical Evaluation of the System Usability Scale. *International Journal of Human-Computer Interaction*, 24(6), 574-594. <https://doi.org/10.1080/10447310802205776>
- Barbara, A. M., Dobbins, M., Haynes, R. B., Iorio, A., Lavis, J. N., Raina, P., & Levinson, A. J. (2016). The McMaster Optimal Aging Portal: Usability Evaluation of a Unique Evidence-Based Health Information Website. *JMIR Human Factors*, 3(1), e14.
<https://doi.org/10.2196/humanfactors.4800>
- Barcenilla, J., & Bastien, J. M. C. (2009). L'acceptabilité des nouvelles technologies : quelles relations avec l'ergonomie, l'utilisabilité et l'expérience utilisateur ? *Le travail humain*, Vol. 72(4), 311-331. <https://www.cairn.info/revue-le-travail-humain-2009-4-page-311.htm>
- Barello, S., Triberti, S., Graffigna, G., Libreri, C., Serino, S., Hibbard, J., & Riva, G. (2016). eHealth for Patient Engagement: A Systematic Review. *Frontiers in Psychology*, 6.

<https://doi.org/10.3389/fpsyg.2015.02013>

- Bass, L., & John, B. E. (2003). Linking usability to software architecture patterns through general scenarios. *Journal of Systems and Software*, 66(3), 187-197.
[https://doi.org/10.1016/S0164-1212\(02\)00076-6](https://doi.org/10.1016/S0164-1212(02)00076-6)
- Bastien, J. M. Christian, & Scapin, D. L. (1993). *Ergonomic criteria for the evaluation of human-computer interfaces* (p. 79). INRIA. <https://hal.inria.fr/inria-00070012/document>
- Bastien, J. M. C. et Scapin, L. (2001). Evaluation des systèmes d'information et critères ergonomiques. In Kolski, C. (ed s.) : Environnement évolués et évaluation de l'IHM. Interaction Homme Machine pour les SI, vol(2), pages 53-80, Paris : Hermès.
- Bastien, J. M. C., & Tricot, A. (2008). *L'évaluation ergonomique des documents électroniques*. Presses Universitaires de France. <https://www.cairn.info/ergonomie-des-documents-electroniques--9782130559924-page-205.htm>
- Batsis, J. A., Zagaria, A., Kotz, D. F., Bartels, S. J., Boateng, G. G., Proctor, P. O., Halter, R. J., & Carpenter-Song, E. A. (2018). Usability evaluation for the Amulet Wearable Device in rural older adults with obesity. *Gerontechnology*, 17(3), 151-159.
<https://doi.org/10.4017/gt.2018.17.3.003.00>
- Beatty, Fukuoka Yoshimi, & Whooley Mary A. (2013). Using Mobile Technology for Cardiac Rehabilitation: A Review and Framework for Development and Evaluation. *Journal of the American Heart Association*, 2(6), e000568.
<https://doi.org/10.1161/JAHA.113.000568>
- Beauchemin, M., Gradilla, M., Baik, D., Cho, H., & Schnall, R. (2019). A Multi-step Usability Evaluation of a Self-Management App to Support Medication Adherence in Persons Living with HIV. *International Journal of Medical Informatics*, 122, 37-44.
<https://doi.org/10.1016/j.ijmedinf.2018.11.012>
- Ben-Zeev, D., Kaiser, S. M., Brenner, C. J., Begale, M., Duffecy, J., & Mohr, D. C. (2013). Development and usability testing of FOCUS: a smartphone system for self-management of schizophrenia. *Psychiatric Rehabilitation Journal*, 36(4), 289-296.
<https://doi.org/10.1037/prj0000019>
- Béranger, J. (2014). Vers une médecine connectée, mesurée et personnalisée centrée sur les données et les big data médicaux. *23èmes journées nationales Athods - Perpignan*, 4.
- Berkman, N. D., Sheridan, S. L., Donahue, K. E., Halpern, D. J., & Crotty, K. (2011). Low Health Literacy and Health Outcomes: An Updated Systematic Review. *Annals of Internal Medicine*, 155(2), 97. <https://doi.org/10.7326/0003-4819-155-2-201107190-00005>
- Berney, S., & Bétrancourt, M. (2016). Does animation enhance learning? A meta-analysis. *Computers & Education*, 101, 150-167.
<https://doi.org/10.1016/j.compedu.2016.06.005>

- Berntsen, N. O., & Dybkjær, L. (2010). *Multimodal Usability*. Springer-Verlag.
<https://www.springer.com/la/book/9781848825529>
- Berry, D. L., Halpenny, B., Bosco, J. L. F., Bruyere, J., & Sanda, M. G. (2015). Usability evaluation and adaptation of the e-health Personal Patient Profile-Prostate decision aid for Spanish-speaking Latino men. *BMC Medical Informatics and Decision Making*, *15*, 56. <https://doi.org/10.1186/s12911-015-0180-4>
- Bert, T., Hausswalt, P., Weill, M., D'Autume, C., Dupays, S., & Perlberg, J. (2014). Perspectives du développement de la chirurgie ambulatoire en France, 288.
- Bétrancourt, M., Dillenbourg, P., & Montarnal, C. (n.d.). *Computer Technologies in Powerful Learning Environments: The Case of Using Animated and Interactive Graphics for Teaching Financial Concepts*.
- Bevan, N. (1995a). Usability is Quality of Use. In Y. Anzai, K. Ogawa, & H. Mori (Éds.), *Advances in Human Factors/Ergonomics* (Vol. 20, p. 349-354). Elsevier.
[https://doi.org/10.1016/S0921-2647\(06\)80241-8](https://doi.org/10.1016/S0921-2647(06)80241-8)
- Bevan, N. (1995b). Measuring usability as quality of use. *Software Quality Journal*, *4*(2), 115-130. <https://doi.org/10.1007/BF00402715>
- Bevan, N. (2009). International Standards for Usability Should Be More Widely Used. *J. Usability Studies*, *4*(3), 106–113. <http://dl.acm.org/citation.cfm?id=2835587.2835588>
- Bevan, N., Carter, J., & Harker, S. (2015). ISO 9241-11 Revised: What Have We Learnt About Usability Since 1998? In M. Kurosu (Éd.), *Human-Computer Interaction: Design and Evaluation* (p. 143-151). Springer International Publishing.
- Bevan, N., & Macleod, M. (1994). Usability measurement in context. *Behaviour & Information Technology*, *13*(1-2), 132-145.
<https://doi.org/10.1080/01449299408914592>
- Blaiwes, A. S. (1974). Formats for presenting procedural instructions. *Journal of Applied Psychology*, *59*(6), 683–686. <https://doi.org/10.1037/h0037492>
- Blandford, A., Green, T. R. G., Furniss, D., & Makri, S. (2007). Evaluating system utility and conceptual fit using CASSM. *International Journal of Human-Computer Studies*, *66*(6), 1-26.
- Blankenship, J., & Dansereau, D. F. (2000). The Effect of Animated Node-Link Displays on Information Recall. *The Journal of Experimental Education*, *68*(4), 293–308.
<https://doi.org/10.1080/00220970009600640>
- Blomberg, J., Burrell, M., & Guest, G. (2003). *The Human-computer Interaction Handbook* (J. A. Jacko & A. Sears (éds.); p. 964–986). L. Erlbaum Associates Inc.
<http://dl.acm.org/citation.cfm?id=772072.772133>
- Bobillier-Chaumon, M.-E. (2016). L'acceptation située des technologies dans et par l'activité : premiers étayages pour une clinique de l'usage. *Psychologie du travail et*

des organisations, 22(1). <https://doi.org/10.1016/j.pto.2016.01.001>

- Bobillier-Chaumon, M.-É., & Dubois, M. (2009). L'adoption des technologies en situation professionnelle : quelles articulations possibles entre acceptabilité et acceptation ? *Le travail humain*, Vol. 72(4), 355-382. <https://www.cairn.info/revue-le-travail-humain-2009-4-page-355.htm?contenu=resume>
- Bobillier Chaumon, M.-E., Dubois, M., & Retour, D. (2006). L'acceptation des nouvelles technologies d'information: le cas des systèmes d'information en milieu bancaire. *Psychologie du Travail et des Organisations*, 12(4), 247-262. <https://doi.org/10.1016/j.pto.2006.05.002>
- Bol, N., Weert, J. C. van, Haes, H. C. de, Loos, E. F., & Smets, E. M. (2015). The Effect of Modality and Narration Style on Recall of Online Health Information: Results From a Web-Based Experiment. *Journal of Medical Internet Research*, 17(4). <https://doi.org/10.2196/jmir.4164>
- Bornet, C., & Brangier, E. (2013). La méthode des personas : principes, intérêts et limites. *Bulletin de psychologie*, Numéro 524(2), 115-134. <https://www.cairn.info/revue-bulletin-de-psychologie-2013-2-page-115.htm>
- Borsci, S., Uchegbu, I., Buckle, P., Ni, Z., Walne, S., & Hanna, G. B. (2018). Designing medical technology for resilience: integrating health economics and human factors approaches. *Expert Review of Medical Devices*, 15(1), 15-26. <https://doi.org/10.1080/17434440.2018.1418661>
- Boucheix, J.-M., Lowe, R. K., & Bugaiska, A. (2015). *Age Differences in Learning from Instructional Animations*. <https://doi.org/10.1002/acp.3131>
- Boutigny, E. (2004). Coopération dans l'entreprise et compétence collective. *Actes du 15e Congrès de l'AGRH*, 3, 1328-1342.
- Brangier, E., & Barcenilla, J. (2003). *Concevoir un produit facile à utiliser - Éric Brangier, Javier... - Éditions Eyrolles* (Editions d'Organisation). <https://www.editions-eyrolles.com/Livre/9782708129009/concevoir-un-produit-facile-a-utiliser>
- Brangier, É., Bornet, C., Bastien, J. M. C., & Vivian, R. (2012). Effets des personas et contraintes fonctionnelles sur l'idéation dans la conception d'une bibliothèque numérique. *Le travail humain*, Vol. 75(2), 121-145. <https://www.cairn.info/revue-le-travail-humain-2012-2-page-121.htm>
- Brangier, É., Hammes-Adelé, S., & Bastien, J.-M. C. (2010). Analyse critique des approches de l'acceptation des technologies : de l'utilisabilité à la symbiose humain-technologie organisation. *Revue Européenne de Psychologie Appliquée*, 60(2), 129-146.
- Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2), 77-101. <https://doi.org/10.1191/1478088706qp063oa>
- Brewster, S. (2002). Overcoming the Lack of Screen Space on Mobile Computers. *Personal and Ubiquitous Computing*, 6(3), 188-205. <https://doi.org/10.1007/s007790200019>

- Brooke, J. (1996). SUS-A quick and dirty usability scale. *Usability evaluation in industry*, 189(194), 4–7.
- Brown, W., Yen, P.-Y., Rojas, M., & Schnall, R. (2013). Assessment of the Health IT Usability Evaluation Model (Health-ITUEM) for evaluating mobile health (mHealth) technology. *Journal of Biomedical Informatics*, 46(6), 1080-1087. <https://doi.org/10.1016/j.jbi.2013.08.001>
- Brown-Johnson, C. G., Berrean, B., & Cataldo, J. K. (2015). Development and usability evaluation of the mHealth Tool for Lung Cancer (mHealth TLC): A virtual world health game for lung cancer patients. *Patient Education and Counseling*, 98(4), 506-511. <https://doi.org/10.1016/j.pec.2014.12.006>
- Bruchon-Schweitzer, M., & Siksou, M. (2008). La psychologie de la santé. *Le Journal des psychologues*, n° 260(7), 28-32. <https://www.cairn.info/revue-le-journal-des-psychologues-2008-7-page-28.htm>
- Burkhardt, J.-M., & Sperandio, J.-C. (2004). Ergonomie et conception informatique. In P. Falzon (Ed.), *Ergonomie* (pp. 437-450). Paris, France : Presses universitaires de France.
- Burnay, E., Cruz-Correia, R., Jacinto, T., Sousa, A. S., & Fonseca, J. (2012). Challenges of a Mobile Application for Asthma and Allergic Rhinitis Patient Enablement—Interface and Synchronization. *Telemedicine and e-Health*, 19(1), 13-18. <https://doi.org/10.1089/tmj.2012.0020>
- Butcher, K. R. (2006). Learning from text with diagrams: Promoting mental model development and inference generation. *Journal of Educational Psychology*, 98(1), 182–197. <https://doi.org/10.1037/0022-0663.98.1.182>
- Carayon, P., & Hoonakker, P. (2019). Human Factors and Usability for Health Information Technology: Old and New Challenges. *Yearbook of Medical Informatics*, 28(1), 71-77. <https://doi.org/10.1055/s-0039-1677907>
- Carman, K. L., Dardess, P., Maurer, M., Sofaer, S., Adams, K., Bechtel, C., & Sweeney, J. (2013). Patient and family engagement: a framework for understanding the elements and developing interventions and policies. *Health Affairs (Project Hope)*, 32(2), 223-231. <https://doi.org/10.1377/hlthaff.2012.1133>
- Carmody, J. K., Denson, L. A., & Hommel, K. A. (2019). Content and Usability Evaluation of Medication Adherence Mobile Applications for Use in Pediatrics. *Journal of Pediatric Psychology*, 44(3), 333-342. <https://doi.org/10.1093/jpepsy/jsy086>
- Carroll, J. M. (1990). *The Nurnberg funnel: designing minimalist instruction for practical computer skill*. MIT Press.
- CCA. (2007). *État de l'apprentissage au Canada : Pas le temps de s'illusionner*. Conseil canadien sur l'apprentissage. <https://www.ccl-cca.ca/early-childhood-learning-knowledge-centre-a-path-towards-improved-early-learning/>

- Chandler, P., & Sweller, J. (1996). Cognitive Load While Learning to Use a Computer Program. *Applied Cognitive Psychology*, 10(2), 151–170. [https://doi.org/10.1002/\(SICI\)1099-0720\(199604\)10:2<151::AID-ACP380>3.0.CO;2-U](https://doi.org/10.1002/(SICI)1099-0720(199604)10:2<151::AID-ACP380>3.0.CO;2-U)
- Chang, E., & Dillon, T. S. (2006). A usability-evaluation metric based on a soft-computing approach. *IEEE Transactions on Systems, Man, and Cybernetics - Part A: Systems and Humans*, 36(2), 356-372. <https://doi.org/10.1109/TSMCA.2005.851349>
- Chaniaud, N., Jean, C. (2017) Open Innovation : Une nouvelle méthode pour sélectionner de jeunes partenaires innovants. CONFER 2017, Séville, Espagne.
- Chaniaud, N., Loup-Escande, E., & Megalakaki, O. (2018). Design, Understanding and Usability Evaluation of Connected Devices in the Field of Health: Contribution of Cognitive Psychology and Ergonomics. *Proceedings of the 20th Congress of the International Ergonomics Association (IEA 2018)*, 582-591. https://doi.org/10.1007/978-3-319-96098-2_72
- Chaniaud, N., Loup-Escande, E., Métayer, N., & Megalakaki, O. (2019). Effets des manuels d'utilisation sur l'utilisabilité des dispositifs médicaux : Les cas d'un tensiomètre et d'un oxymètre de pouls. *EPIQUE 2019*, 1-8.
- Chaniaud, N., Métayer, N., Megalakaki, O., & Loup-Escande, É. (2020). Effect of prior health knowledge of novice users on Usability Results: Usability Study. *JMIR MHealth and UHealth*, 8(9). [10.2196/17983](https://doi.org/10.2196/17983) <https://mhealth.jmir.org/2020/9/e17983/pdf>
- Chaniaud, N., Métayer, N., Loup-Escande, É., & Megalakaki, O. (2021). Impact of the format of user instructions on the handling of a wrist blood pressure monitor. *Cognitive Processing*. [10.1007/s10339-020-01006-1](https://doi.org/10.1007/s10339-020-01006-1)
- Chatenier, E. du, Verstegen, J. A. A. M., Biemans, H. J. A., Mulder, M., & Omta, O. S. W. F. (2010). Identification of competencies for professionals in open innovation teams. *R&D Management*, 40(3), 271-280. <https://doi.org/10.1111/j.1467-9310.2010.00590.x>
- Chen, K., & Chan, A. H. S. (2011). A review of technology acceptance by older adults. *Gerontechnology*, 10(1), 1-12. <https://doi.org/10.4017/gt.2011.10.01.006.00>
- Chesbrough, H. W. (2003). *Open Innovation: The New Imperative for Creating and Profiting from Technology*. Harvard Business Press.
- Cho, H., Powell, D., Pichon, A., Thai, J., Bruce, J., Kuhns, L. M., Garofalo, R., & Schnall, R. (2018). A Mobile Health Intervention for HIV Prevention Among Racially and Ethnically Diverse Young Men: Usability Evaluation. *JMIR MHealth and UHealth*, 6(9), e11450. <https://doi.org/10.2196/11450>
- Cho, H., Yen, P.-Y., Dowding, D., Merrill, J. A., & Schnall, R. (2018). A multi-level usability evaluation of mobile health applications: A case study. *Journal of Biomedical Informatics*, 86, 79-89. <https://doi.org/10.1016/j.jbi.2018.08.012>
- Cho, J., Park, D., & Lee, H. E. (2014). Cognitive Factors of Using Health Apps: Systematic Analysis of Relationships Among Health Consciousness, Health Information

- Orientation, eHealth Literacy, and Health App Use Efficacy. *Journal of Medical Internet Research*, 16(5), e125. <https://doi.org/10.2196/jmir.3283>
- Chong, I., Cha, J., Peng, F., & Yu, D. (2019). Usability of a Home-Use Blood Pressure Monitor System: An Evaluation of the Device and Instructional Materials. *Proceedings of the International Symposium on Human Factors and Ergonomics in Health Care*, 8(1), 147-151. <https://doi.org/10.1177/2327857919081036>
- Cifter, A. S. (2017). Blood Pressure Monitor Usability Problems Detected Through Human Factors Evaluation. *Ergonomics in Design*, 25(3), 11-19. <https://doi.org/10.1177/1064804617691397>
- Combessie, J.-C. (2007). II. L'entretien semi-directif. *Reperes*, 5e éd., 24-32. <https://www.cairn.info/la-methode-en-sociologie--9782707152411-page-24.htm>
- Connor, M., Mantwill, S., & Schulz, P. J. (2013). Functional health literacy in Switzerland—Validation of a German, Italian, and French health literacy test. *Patient Education and Counseling*, 90(1), 12-17. <https://doi.org/10.1016/j.pec.2012.08.018>
- Constantine, L. ., & Lockwood, L. A. . (1999). *Software for Use: A Practical Guide to the Models and Methods of Usage-Centered Design*. Reading.
- Constantinescu, G., Kuffel, K., King, B., Hodgetts, W., & Rieger, J. (2019). Usability testing of an mHealth device for swallowing therapy in head and neck cancer survivors. *Health Informatics Journal*, 25(4), 1373-1382. <https://doi.org/10.1177/1460458218766574>
- Cottin-Joshi, I., & Vallery, G. (2015). *Conception et prévention des risques en milieu hospitalier : Les apports de l'ergonomie*. 10.
- Cruz Zapata, B., Fernández-Alemán, J. L., Toval, A., & Idri, A. (2018). Reusable Software Usability Specifications for mHealth Applications. *Journal of Medical Systems*, 42(3), 45. <https://doi.org/10.1007/s10916-018-0902-0>
- CSE. (2013). *Un engagement collectif pour maintenir et rehausser les compétences en littératie des adultes* (p. 193). Conseil Supérieur de l'Éducation. <http://www.specs-csn.qc.ca/site-com/publications-divers/Conseil-superieur-education/CSE-avis-litteratie-des-adultes.pdf>
- Czaja, S. J., Zarcadoolas, C., Vaughn, W. L., Lee, C. C., Rockoff, M. L., & Levy, J. (2015). The Usability of Electronic Personal Health Record Systems for an Underserved Adult Population. *Human Factors*, 57(3), 491. <https://doi.org/10.1177/0018720814549238>
- Czarniawska-Joerges, B. (2007). *Shadowing: And Other Techniques for Doing Fieldwork in Modern Societies*. Copenhagen Business School Press DK.
- Darses, F., Détienne, F., & Visser, W. (2001). Assister la conception : perspectives pour la psychologie cognitive ergonomique. *Proceedings of EPIQUE*, 1-10. <http://sfpsy.org/spe-grape/Actes-epique-2001-article-darses.pdf>

- Daugherty, B. L., Schap, T. E., Ettienne-Gittens, R., Zhu, F. M., Bosch, M., Delp, E. J., Ebert, D. S., Kerr, D. A., & Boushey, C. J. (2012). Novel Technologies for Assessing Dietary Intake: Evaluating the Usability of a Mobile Telephone Food Record Among Adults and Adolescents. *Journal of Medical Internet Research*, *14*(2), e58. <https://doi.org/10.2196/jmir.1967>
- Davidson, T. M., Yuen, E. K., Felton, J. W., McCauley, J., Gros, K. S., & Ruggiero, K. J. (2014). Feasibility assessment of a brief, web-based behavioral activation intervention for adolescents with depressed mood. *International Journal of Psychiatry in Medicine*, *48*(1), 69-82. <https://doi.org/10.2190/PM.48.1.f>
- Davis, F. D. (1989). Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology. *MIS Quarterly*, *13*(3), 319. <https://doi.org/10.2307/249008>
- Davis, T. C., Crouch, M. A., Long, S. W., Jackson, R. H., Bates, P., George, R. B., & Bairnsfather, L. E. (1991). Rapid assessment of literacy levels of adult primary care patients. *Family Medicine*, *23*(6), 433-435.
- Dillon, A., & Watson, C. (1996). User analysis in HCI—the historical lessons from individual differences research. *International Journal of Human-Computer Studies*, *45*(6), 619-637. <https://doi.org/10.1006/ijhc.1996.0071>
- Dix, A., Finlay, J., Abowd, G., & Russell, B. (2003). *Human-Computer Interaction* (Prentice Hall).
- Dixon, P. (1987). The structure of mental plans for following directions. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *13*(1), 18-26. <https://doi.org/10.1037/0278-7393.13.1.18>
- Dos Santos-Fontes, R. L., Ferreiro de Andrade, K. N., Sterr, A., & Conforto, A. B. (2013). Home-Based Nerve Stimulation to Enhance Effects of Motor Training in Patients in the Chronic Phase After Stroke: A Proof-of-Principle Study. *Neurorehabilitation and Neural Repair*, *27*(6), 483-490. <https://doi.org/10.1177/1545968313478488>
- Dubey, S. K., & Gulati, A. (2012). *Integrated Model for Software Usability*. *4*(03), 9.
- Dubey, S. K., Rana, A., & Sharma, A. (2012). *Usability Evaluation of Object Oriented Software System using Fuzzy Logic Approach*. <https://doi.org/10.5120/6208-8778>
- Dumas, J. S. (2003). *The Human-computer Interaction Handbook* (J. A. Jacko & A. Sears (éds.); p. 1093-1117). L. Erlbaum Associates Inc. <http://dl.acm.org/citation.cfm?id=772072.772141>
- Dumas, J. S., & Redish, J. C. (1999). *A Practical Guide to Usability Testing* (1st éd.). Intellect Books.
- Eiriksdottir, E., & Catrambone, R. (2011). Procedural Instructions, Principles, and Examples: How to Structure Instructions for Procedural Tasks to Enhance Performance, Learning, and Transfer. *Human Factors*, *53*(6), 749-770.

<https://doi.org/10.1177/0018720811419154>

- Elie, C., & Colombet, I. (2011). Méthodes d'estimation de la reproductibilité. *Sang Thrombose Vaisseaux*, 23(3), 138-145. <https://doi.org/10.1684/stv.2011.0592>
- Eysenbach, G. (2001). What is e-health? *Journal of Medical Internet Research*, 3(2), e20. <https://doi.org/10.2196/jmir.3.2.e20>
- Eysenbach, Gunther, & Group, C.-E. (2011). CONSORT-EHEALTH: Improving and Standardizing Evaluation Reports of Web-based and Mobile Health Interventions. *Journal of Medical Internet Research*, 13(4), e126. <https://doi.org/10.2196/jmir.1923>
- Ezzedine, H. (2002). Méthodes et Modèles de Spécification et Evaluation des Interfaces Homme-Machine dans les systèmes industriels complexes. Mémoire d'Habilitation diriger des recherches. Université de Valenciennes et du Hainaut-Cambrésis, Valenciennes, décembre.
- Falzon, P. (2004). *1. Nature, objectifs et connaissances de l'ergonomie*. Presses Universitaires de France. <https://www.cairn.info/ergonomie--9782130514046-page-15.htm>
- Flanagan, J. C. (1954). The critical incident technique. *Psychological Bulletin*, 51(4), 327-358. <https://doi.org/10.1037/h0061470>
- Fleury, S., & Jamet, E. (2015). Supervised Automatic Interpretation of Technical Documents: When Interruption is a Time Saver. *Perceptual and Motor Skills*, 120(1), 67-83. <https://doi.org/10.2466/22.25.PMS.120v11x9>
- Folmer, E., & Bosch, J. (2004). Architecting for usability: a survey. *Journal of Systems and Software*, 70(1-2), 61-78. [https://doi.org/\(...\)0164-1212\(02\)00159-0](https://doi.org/(...)0164-1212(02)00159-0)
- Fraser, K. L., Ayres, P., & Sweller, J. (2015). Cognitive Load Theory for the Design of Medical Simulations. *Simulation in Healthcare: Journal of the Society for Simulation in Healthcare*, 10(5), 295-307. <https://doi.org/10.1097/SIH.0000000000000097>
- Friedman, C. P., & Wyatt, J. (2005). *Evaluation Methods in Biomedical Informatics*. Springer Science & Business Media.
- Frøkjær, E., Hertzum, M., & Hornbæk, K. (2000). *Measuring usability: are effectiveness, efficiency, and satisfaction really correlated?* 345-352. <https://doi.org/10.1145/332040.332455>
- Fu, M. R., Axelrod, D., Guth, A. A., Rampertaap, K., El-Shammaa, N., Hiotis, K., Scagliola, J., Yu, G., & Wang, Y. (2016). mHealth self-care interventions: managing symptoms following breast cancer treatment. *mHealth*, 2. <https://doi.org/10.21037/mhealth.2016.07.03>
- Fu, W.-T., & Gray, W. D. (2006). Suboptimal tradeoffs in information seeking. *Cognitive Psychology*, 52(3), 195-242. <https://doi.org/10.1016/j.cogpsych.2005.08.002>
- Fuller-Tyszkiewicz, M., Richardson, B., Klein, B., Skouteris, H., Christensen, H., Austin, D., Castle, D., Mihalopoulos, C., O'Donnell, R., Arulkadacham, L., Shatte, A., & Ware,

- A. (2018). A Mobile App-Based Intervention for Depression: End-User and Expert Usability Testing Study. *JMIR Mental Health*, 5(3), e54. <https://doi.org/10.2196/mental.9445>
- Fung, C. H., Igodan, U., Alessi, C., Martin, J. L., Dzierzewski, J. M., Josephson, K., & Kramer, B. J. (2015). Human factors/usability barriers to home medical devices among individuals with disabling conditions: in-depth interviews with positive airway pressure device users. *Disability and Health Journal*, 8(1), 86-92. <https://doi.org/10.1016/j.dhjo.2014.06.002>
- Ganier, F. (2004). Factors affecting the processing of procedural instructions: implications for document design. *IEEE Transactions on Professional Communication*, 47(1), 15-26. <https://doi.org/10.1109/TPC.2004.824289>
- Ganier, F., & Querrec, R. (2012). TIP-EXE: A Software Tool for Studying the Use and Understanding of Procedural Documents. *IEEE Transactions on Professional Communication*, 55(2), 106-121. <https://doi.org/10.1109/TPC.2012.2194600>
- Ganier, F. (2002). Évaluer l'efficacité des documents techniques procéduraux : un panorama des méthodes. *Le travail humain*, Vol. 65(1), 1-27. <https://www.cairn.info/revue-le-travail-humain-2002-1-page-1.htm>
- Ganier, F., Gombert, J.-E., & Fayol, M. (2000). Effets du format de présentation des instructions sur l'apprentissage de procédures à l'aide de documents techniques. *Le Travail Humain*, 63(2), 121–152. JSTOR. <https://www.jstor.org/stable/40660254>
- Garcia, M. B. (2019). A Speech Therapy Game Application for Aphasia Patient Neurorehabilitation - A Pilot Study of an mHealth App. *International Journal of Simulation: Systems, Science & Technology*, 20(2), 51-59. https://www.academia.edu/40133381/A_Speech_Therapy_Game_Application_for_Aphasia_Patient_Neurorehabilitation_-_A_Pilot_Study_of_an_mHealth_App
- Georgsson, M., & Staggers, N. (2016a). Quantifying usability: an evaluation of a diabetes mHealth system on effectiveness, efficiency, and satisfaction metrics with associated user characteristics. *Journal of the American Medical Informatics Association*, 23(1), 5-11. <https://doi.org/10.1093/jamia/ocv099>
- Georgsson, M., & Staggers, N. (2016b). An evaluation of patients' experienced usability of a diabetes mHealth system using a multi-method approach. *Journal of Biomedical Informatics*, 59, 115-129. <https://doi.org/10.1016/j.jbi.2015.11.008>
- Georgsson, M., Staggers, N., Årsand, E., & Kushniruk, A. (2019). Employing a user-centered cognitive walkthrough to evaluate a mHealth diabetes self-management application: A case study and beginning method validation. *Journal of Biomedical Informatics*, 91, 103110. <https://doi.org/10.1016/j.jbi.2019.103110>
- Gerdes, M., Smaradottir, B., & Fensli, R. W. (2014). *End-to-End Infrastructure for Usability Evaluation of eHealth Applications and Services*.

- Ghosh, S., & Sallam, S. (1994). Patient satisfaction and postoperative demands on hospital and community services after day surgery. *The British Journal of Surgery*, *81*(11), 1635-1638.
- Gordon, J. S., Armin, J. S., Cunningham, J. K., Muramoto, M. L., Christiansen, S. M., & Jacobs, T. A. (2017). Lessons learned in the development and evaluation of RxCoach™, an mHealth app to increase tobacco cessation medication adherence. *Patient Education and Counseling*, *100*(4), 720-727. <https://doi.org/10.1016/j.pec.2016.11.003>
- Graffigna, G., Barello, S., & Bonanomi, A. (2017). The role of Patient Health Engagement Model (PHE-model) in affecting patient activation and medication adherence: A structural equation model. *PLoS ONE*, *12*(6). <https://doi.org/10.1371/journal.pone.0179865>
- Graffigna, G., Barello, S., & Riva, G. (2013). Technologies for patient engagement. *Health Affairs (Project Hope)*, *32*(6), 1172. <https://doi.org/10.1377/hlthaff.2013.0279>
- Grebin, S. Z., Echeveste, M. E. S., Magnago, P. F., Tanure, R. L. Z., & Pulgati, F. H. (2018). Analytical strategy for assessment of usability of medical devices from the user's perspective: a study of hemodialysis patients. *Cadernos De Saude Publica*, *34*(8), e00074417. <https://doi.org/10.1590/0102-311X00074417>
- Green, T. R. G., & Hoc, J.-M. (1991). What is cognitive ergonomics? *Le Travail Humain*, *54*(4), 291-304. JSTOR. <https://www.jstor.org/stable/40657604>
- Greene, J., & Hibbard, J. H. (2012). Why Does Patient Activation Matter? An Examination of the Relationships Between Patient Activation and Health-Related Outcomes. *Journal of General Internal Medicine*, *27*(5), 520-526. <https://doi.org/10.1007/s11606-011-1931-2>
- Greysen, S. R., Garcia, C. C., Sudore, R. L., Cenzer, I. S., & Covinsky, K. E. (2014). Functional Impairment and Internet Use among Older Adults: Implications for meaningful use of patient portals. *JAMA Internal Medicine*, *174*(7), 1188. <https://doi.org/10.1001/jamainternmed.2014.1864>
- Griffin, A. (2004). Obtaining Customer Needs for Product Development. In K. B. Kahn (Éd.), *The PDMA Handbook of New Product Development* (1^{re} éd., p. 211-227). Wiley. <https://doi.org/10.1002/9780470172483.ch14>
- Grindrod, K. A., Li, M., & Gates, A. (2014). Evaluating user perceptions of mobile medication management applications with older adults: a usability study. *JMIR mHealth and uHealth*, *2*(1), e11-e11. <https://doi.org/10.2196/mhealth.3048>
- Guidat, A. (2010). Prise en charge anesthésique des patients en hospitalisation ambulatoire. *Annales Françaises d'Anesthésie et de Réanimation*, *29*, 67-72.
- Gunter, R., Fernandes-Taylor, S., Mahnke, A., Awoyinka, L., Schroeder, C., Wiseman, J., Sullivan, S., Bennett, K., Greenberg, C., & Kent, K. C. (2016). Evaluating Patient Usability of an Image-Based Mobile Health Platform for Postoperative Wound

- Monitoring. *JMIR MHealth and UHealth*, 4(3), e113.
<https://doi.org/10.2196/mhealth.6023>
- Guthrie, J. T., Bennett, S., & Weber, S. (1991). Processing procedural documents: A cognitive model for following written directions. *Educational Psychology Review*, 3(3), 249–265.
<https://doi.org/10.1007/BF01320078>
- Hamine, S., Gerth-Guyette, E., Faulx, D., Green, B. B., & Ginsburg, A. S. (2015). Impact of mHealth chronic disease management on treatment adherence and patient outcomes: a systematic review. *Journal of Medical Internet Research*, 17(2), e52.
<https://doi.org/10.2196/jmir.3951>
- Hammontree, M., Weiler, P., & Nayak, N. (1994). Remote usability testing. *Interactions*, 1(3), 21-25. <https://doi.org/10.1145/182966.182969>
- Harrison, R., Flood, D., & Duce, D. (2013). Usability of mobile applications: literature review and rationale for a new usability model. *Journal of Interaction Science*, 1(1), 1.
<https://doi.org/10.1186/2194-0827-1-1>
- Harte, R., Glynn, L., Rodríguez-Molinero, A., Baker, P. M., Scharf, T., Quinlan, L. R., & ÓLaighin, G. (2017). A Human-Centered Design Methodology to Enhance the Usability, Human Factors, and User Experience of Connected Health Systems: A Three-Phase Methodology. *JMIR Human Factors*, 4(1), e8.
<https://doi.org/10.2196/humanfactors.5443>
- Harte, R., Hall, T., Glynn, L., Rodríguez-Molinero, A., Scharf, T., Quinlan, L. R., & ÓLaighin, G. (2018). Enhancing Home Health Mobile Phone App Usability Through General Smartphone Training: Usability and Learnability Case Study. *JMIR Human Factors*, 5(2), e18. <https://doi.org/10.2196/humanfactors.7718>
- Hartson, R. (2003). Cognitive, physical, sensory, and functional affordances in interaction design. *Behaviour & Information Technology*, 22(5), 315-338.
<https://doi.org/10.1080/01449290310001592587>
- HAS/ANAP. (2014). Rapport d'évaluation technologique : Elements d'appréciation en vue de la prise en charge d'un patient. 198. https://www.has.sante.fr/upload/docs/application/pdf/2014-07/rapport_chir_ambu_axe2_vf.pdf
- Hassenzahl, M. (2005). The Thing and I: Understanding the Relationship Between User and Product. In M. A. Blythe, K. Overbeeke, A. F. Monk, & P. C. Wright (Éds.), *Funology: From Usability to Enjoyment* (p. 31-42). Springer Netherlands.
https://doi.org/10.1007/1-4020-2967-5_4
- Hatchuel, A., & Weil, B. (2008). C-K design theory: an advanced formulation. *Research in Engineering Design*, 19(4), 181. <https://doi.org/10.1007/s00163-008-0043-4>
- Haun, J. N., Valerio, M. A., McCormack, L. A., Sørensen, K., & Paasche-Orlow, M. K. (2014). Health Literacy Measurement: An Inventory and Descriptive Summary of 51 Instruments. *Journal of Health Communication*, 19(sup2), 302-333.
<https://doi.org/10.1080/10810730.2014.936571>

- Hayes, J. R., & Simon, H. A. (1974). Understanding Written Problem Instructions. In L. W. Gregg (Ed.), *Knowledge and Cognition* (pp. 167-200). New York: LEA Press.
- Heathfield, H., Hudson, P., Kay, S., Mackay, L., Marley, T., Nicholson, L., Peel, V., Roberts, R., & Williams, J. (1999). Issues in the multi-disciplinary assessment of healthcare information systems. *Information Technology & People*, 12(3), 253-275.
<https://doi.org/10.1108/09593849910278277>
- Helyar, P. S. (1992). Products Liability: Meeting Legal Standards for Adequate Instructions. *Journal of Technical Writing and Communication*, 22(2), 125-147.
<https://doi.org/10.2190/6KCA-5D22-00Q9-A5FG>
- Hertzum, M. (2010). Images of Usability. *International Journal of Human-Computer Interaction*, 26(6), 567-600. <https://doi.org/10.1080/10447311003781300>
- Heurley, L. (1994). *Traitement de textes proceduraux : etude de psycholinguistique cognitive des processus de production et de comprehension chez des adultes non experts* [Thesis, Dijon]. <http://www.theses.fr/1994DIJOL001>
- Heurley, L. (1997). Vers une définition du concept de texte procédural : le point de vue de la psycholinguistique. *Les cahiers du Français Contemporain*, 109-133.
- Hibbard, J. H., Stockard, J., Mahoney, E. R., & Tusler, M. (2004). Development of the Patient Activation Measure (PAM): Conceptualizing and Measuring Activation in Patients and Consumers. *Health Services Research*, 39(4 Pt 1), 1005-1026.
<https://doi.org/10.1111/j.1475-6773.2004.00269.x>
- Hindmarsh, J., Reynolds, P., & Dunne, S. (2011). Exhibiting understanding: The body in apprenticeship. *Journal of Pragmatics*, 43(2), 489-503.
<https://doi.org/10.1016/j.pragma.2009.09.008>
- Hoc, J.-M. (1987). *Cognitive Psychology of Planning*.
<https://doi.org/10.1017/S0263574700006366>
- Holt, K. A., Karnoe, A., Overgaard, D., Nielsen, S. E., Kayser, L., Røder, M. E., & From, G. (2019). Differences in the Level of Electronic Health Literacy Between Users and Nonusers of Digital Health Services: An Exploratory Survey of a Group of Medical Outpatients. *Interactive Journal of Medical Research*, 8(2), e8423.
<https://doi.org/10.2196/ijmr.8423>
- Holtzblatt, K., & Jones, S. (2017). Contextual Inquiry: A Participatory Technique for System Design. In D. Schuler & A. Namioka, *Participatory Design : Principles and Practices*. Erlbaum Associates Inc. /paper/Contextual-Inquiry%3A-A-Participatory-Technique-for-Karen-Sandra/bd9f531421f8357b4f6fdadd1cd0e0c3e7c1a5b5
- Hornbæk, K., & Law, E. L.-C. (2007). Meta-analysis of Correlations Among Usability Measures. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 617-626. <https://doi.org/10.1145/1240624.1240722>

- Horsky, J., Kuperman, G. J., & Patel, V. L. (2005). Comprehensive analysis of a medication dosing error related to CPOE. *Journal of the American Medical Informatics Association: JAMIA*, 12(4), 377-382. <https://doi.org/10.1197/jamia.M1740>
- Househ, M. S., Shubair, M. M., Yunus, F., Jamal, A., & Aldossari, B. (2015). The Use of an Adapted Health IT Usability Evaluation Model (Health-ITUEM) for Evaluating Consumer Reported Ratings of Diabetes mHealth Applications: Implications for Diabetes Care and Management. *Acta Informatica Medica: AIM: Journal of the Society for Medical Informatics of Bosnia & Herzegovina: Casopis Drustva Za Medicinsku Informatiku BiH*, 23(5), 290-295. <https://doi.org/10.5455/aim.2015.23.290-295>
- Housten, A. J., Lowenstein, L. M., Hoover, D. S., Leal, V. B., Kamath, G. R., & Volk, R. J. (2018). Limitations of the S-TOFHLA in measuring poor numeracy: a cross-sectional study. *BMC Public Health*, 18. <https://doi.org/10.1186/s12889-018-5333-9>
- Hubka, V., & Ernst Eder, W. (1987). A scientific approach to engineering design. *Design Studies*, 8(3), 123–137. [https://doi.org/10.1016/0142-694X\(87\)90035-4](https://doi.org/10.1016/0142-694X(87)90035-4)
- Hudon, C., Lambert, M., Almirall, J., & Cook, M. (2012). The reliability of the French version of the Patient Activation Measure. *40th Annual meeting of North American Primary Care Research Group*.
- Huhta, A.-M., Hirvonen, N., & Huotari, M.-L. (2018). Health Literacy in Web-Based Health Information Environments: Systematic Review of Concepts, Definitions, and Operationalization for Measurement. *Journal of Medical Internet Research*, 20(12), e10273. <https://doi.org/10.2196/10273>
- Hunt, D. P. (2003). The concept of knowledge and how to measure it. *Journal of Intellectual Capital*, 4(1), 100–113. <https://doi.org/10.1108/14691930310455414>
- Hussain, A., Mkpojiogu, E. O. C., & Hussain, Z. (2015). Usability evaluation of a web-based health awareness portal on smartphone devices using ISO 9241-11 model. *Jurnal Teknologi*, 77(4). <https://doi.org/10.11113/jt.v77.6035>
- Huvila, I., Ek, S., Enwald, H., Eriksson, K., Hirvonen, N., & Känsäkoski, H. (2016). Taking Health Information Behaviour into Account in the design of e-health services. *Scientific papers*, 8(4), 153-165.
- Isaković, M., Sedlar, U., Volk, M., & Bešter, J. (2016). *Usability Pitfalls of Diabetes mHealth Apps for the Elderly* [Research article]. *Journal of Diabetes Research*. <https://doi.org/10.1155/2016/1604609>
- IEC 62366-1:2015/Amd1:2020. (2020). *Medical devices -- Part 1: Application of usability engineering to medical devices - Amendment 1*. ISO. <https://www.iso.org/fr/standard/73007.html>
- ISO 13407:1999 (1999). *Processus de conception centrée sur l'opérateur humain pour les systèmes interactifs*. ISO. <https://www.iso.org/fr/standard/21197.html>

- ISO 13485:2016 (2016). *Dispositif médicaux — Systèmes de management de la qualité — Exigences à des fins réglementaires*. ISO. <https://www.iso.org/fr/standard/59752.html>
- ISO 14155:2020 (2020). *Investigation clinique des dispositifs médicaux pour sujets humains Bonne pratique clinique*. ISO. <https://www.iso.org/fr/standard/71690.html>
- ISO 14971:2019. (2019). *Dispositifs médicaux — Application de la gestion des risques aux dispositifs médicaux*. ISO. <https://www.iso.org/fr/standard/72704.html>
- ISO 9126-1:1991 (1991). *Génie du logiciel — Qualité des produits*. ISO. <https://www.iso.org/fr/standard/16722.html>
- ISO 9241-11:2018 (2018). *Ergonomie de l'interaction homme-système — Partie 11 : Utilisabilité — Définitions et concepts*. ISO. <https://www.iso.org/fr/standard/63500.html>
- ISO 9241-210:2019 (2019). *Ergonomie de l'interaction homme-système — Partie 210 : Conception centrée sur l'opérateur humain pour les systèmes interactifs*. ISO. <https://www.iso.org/fr/standard/77520.html>
- ISO/IEC 14598-5:1998 (1998). *Technologies de l'information — Evaluation du produit logiciel — Partie 5: Procédés pour les évaluateurs*. ISO. <https://www.iso.org/fr/standard/24906.html>
- ISO/IEC 60601-1-6:2010 (2010). *Appareils électromédicaux — Partie 1-6 : exigences générales pour la sécurité de base et les performances essentielles — Norme collatérale : aptitude à l'utilisation*. <https://www.boutique.afnor.org/norme/nf-en-60601-1-6/appareils-electromedicaux-partie-1-6-exigences-generales-pour-la-securite-de-base-et-les-performances-essentielles-norme-collate/article/738110/fa165887>
- ISO/TR 18529:2000 (2000). *Ergonomie — Ergonomie de l'interaction homme/système — Descriptions des processus cycle de vie centrées sur l'opérateur humain*. <https://www.iso.org/fr/standard/33499.html>
- ISO/IEC TR 9126-2:2003 (2003). *Génie du logiciel — Qualité des produits — Partie 2 : Métrologie externe*. ISO. <https://www.iso.org/fr/standard/22750.html>
- Isomursu, M., Tähti, M., Väinämö, S., & Kuutti, K. (2007). Experimental evaluation of five methods for collecting emotions in field settings with mobile applications. *International Journal of Human-Computer Studies*, 65(4), 404-418. <https://doi.org/10.1016/j.ijhcs.2006.11.007>
- Jake-Schoffman, D. E., Silfee, V. J., Waring, M. E., Boudreaux, E. D., Sadasivam, R. S., Mullen, S. P., Carey, J. L., Hayes, R. B., Ding, E. Y., Bennett, G. G., & Pagoto, S. L. (2017). Methods for Evaluating the Content, Usability, and Efficacy of Commercial Mobile Health Apps. *JMIR MHealth and UHealth*, 5(12), e190. <https://doi.org/10.2196/mhealth.8758>

- Jambon, F., Mandran, N., Meillon, B., & Perrot, C. (2010). Evaluation des systèmes mobiles et ubiquitaires: proposition de méthodologie et retours d'expérience. *Journal d'Interaction Personne-Système*, 1(1), 1-34. <https://hal.inria.fr/hal-01058933>
- Jamet, É. (2012). *La compréhension des documents multimédias : de la cognition à la conception*. Solal.
- Jamet, É., & Arguel, A. (2008). La compréhension d'un document technique multimédia peut-elle être améliorée par une présentation séquentielle de son contenu ? *Le travail humain*, Vol. 71(3), 253–270. <https://www.cairn.info/revue-le-travail-humain-2008-3-page-253.htm>
- Jannin, L., Ganier, F., & De Vries, P. (2019). Atomized or delayed execution? An alternative paradigm for the study of procedural learning. *Journal of Educational Psychology*, No Pagination Specified-No Pagination Specified. <https://doi.org/10.1037/edu0000357>
- Jaspers, M. W. M. (2009). A comparison of usability methods for testing interactive health technologies: Methodological aspects and empirical evidence. *International Journal of Medical Informatics*, 78(5), 340-353. <https://doi.org/10.1016/j.ijmedinf.2008.10.002>
- Jensen, J. D., King, A. J., Davis, L. A., & Guntzville, L. M. (2010). Utilization of internet technology by low-income adults: the role of health literacy, health numeracy, and computer assistance. *Journal of Aging and Health*, 22(6), 804-826. <https://doi.org/10.1177/0898264310366161>
- Jeon, E., & Park, H.-A. (2018). Development of the IMB Model and an Evidence-Based Diabetes Self-management Mobile Application. *Healthcare Informatics Research*, 24(2), 125-138. <https://doi.org/10.4258/hir.2018.24.2.125>
- Jeung, H.-J., Chandler, P., & Sweller, J. (1997). The Role of Visual Indicators in Dual Sensory Mode Instruction. *Educational Psychology*, 17(3), 329–345. <https://doi.org/10.1080/0144341970170307>
- Jibb, L. A., Cafazzo, J. A., Nathan, P. C., Seto, E., Stevens, B. J., Nguyen, C., & Stinson, J. N. (2017). Development of a mHealth Real-Time Pain Self-Management App for Adolescents With Cancer: An Iterative Usability Testing Study. *Journal of Pediatric Oncology Nursing*, 34(4), 283-294. <https://doi.org/10.1177/1043454217697022>
- Jones, J., & Caird, J. K. (2017). The Usability of Blood Glucose Meters: Task Performance Differences Between Younger and Older Age Groups. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 61(1), 604-608. <https://doi.org/10.1177/1541931213601636>
- Jordan, J. E., Osborne, R. H., & Buchbinder, R. (2011). Critical appraisal of health literacy indices revealed variable underlying constructs, narrow content and psychometric weaknesses. *Journal of Clinical Epidemiology*, 64(4), 366-379. <https://doi.org/10.1016/j.jclinepi.2010.04.005>
- Kalichman, S. C., Ramachandran, B., & Catz, S. (1999). Adherence to combination antiretroviral therapies in HIV patients of low health literacy. *Journal of General*

- Internal Medicine*, 14(5), 267–273. <https://doi.org/10.1046/j.1525-1497.1999.00334.x>
- Kalyuga, S., Sweller, J., & Chandler, P. (2000). Incorporating learner experience into the design of multimedia instruction. *Faculty of Education - Papers (Archive)*, 126–136. <https://doi.org/10.1037/0022-0663.92.1.126>
- Karapanos, E., Zimmerman, J., Forlizzi, J., & Martens, J.-B. (2009). User Experience over Time: An Initial Framework. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 729–738. <https://doi.org/10.1145/1518701.1518814>
- Kaufman, D. R., Patel, V. L., Hilliman, C., Morin, P. C., Pevzner, J., Weinstock, R. S., Goland, R., Shea, S., & Starren, J. (2003). Usability in the real world: assessing medical information technologies in patients' homes. *Journal of Biomedical Informatics*, 36(1-2), 45-60.
- Kaufman-Rivi, D., Collins-Mitchell, J., & Jetley, R. (2010). Design considerations for medical devices in the home environment. *Biomedical Instrumentation & Technology, Suppl Home Healthcare*, 21-26.
- Kayser, L., Kushniruk, A., Osborne, R. H., Norgaard, O., & Turner, P. (2015). Enhancing the Effectiveness of Consumer-Focused Health Information Technology Systems Through eHealth Literacy: A Framework for Understanding Users' Needs. *JMIR Human Factors*, 2(1), e9. <https://doi.org/10.2196/humanfactors.3696>
- Kazakç1, A. O., Gillier, T., & Piat, G. (2008). Investigating co-innovation in exploratory partnerships: An analytical framework based on design theory. In E. P.-www.erima.estia.fr (Éd.), *European Research on Innovation and Management Alliance* (Vol. 2, p. 119). <https://hal.archives-ouvertes.fr/hal-00417247>
- Ketola, P., & Roykkee, M. (2001). The Three Facets of Usability In Mobile Handsets. *CHI 2001*.
- Kieras, D. E., & Bovair, S. (1986). The acquisition of procedures from text: A production-system analysis of transfer of training. *Journal of Memory and Language*, 25(5), 507–524. [https://doi.org/10.1016/0749-596X\(86\)90008-2](https://doi.org/10.1016/0749-596X(86)90008-2)
- Kickbusch, D. I. (2004). *Improving Health Literacy in the European Union: towards a Europe of informed and active health citizens*. 16.
- Kim, H., & Xie, B. (2017). Health literacy in the eHealth era: A systematic review of the literature. *Patient Education and Counseling*, 100(6), 1073-1082. <https://doi.org/10.1016/j.pec.2017.01.015>
- Kim, J., Kim, S., Kim, H., Kim, K., Yang, S., & Shin, Y. (2011). Development of Implementation Strategies for u-Health Services Based on the Healthcare Professionals' Experiences. *Telemedicine and e-Health*, 17(2), 80-87. <https://doi.org/10.1089/tmj.2010.0124>
- Kirakowski, J. and Corbett, M. (1993) *SUMI User handbook Human Factors Research Group / Course Hero*. <https://www.coursehero.com/file/p2snihq/Kirakowski-J-and-Corbett->

- Kjeldskov, J., Graham, C., Pedell, S., Vetere, F., Howard, S., Balbo, S., & Davies, J. (2005). Evaluating the usability of a mobile guide: The influence of location, participants and resources. *Behaviour & Information Technology*, 24(1), 51-65. <https://doi.org/10.1080/01449290512331319030>
- Kjeldskov, J., & Skov, M. B. (2003). Creating Realistic Laboratory Settings: Comparative Studies of Three Think-Aloud Usability Evaluations of a Mobile System. *Proceedings of the 9th IFIP TC13 International Conference on Human-Computer Interaction (Interact2003)*, 663-670. <https://vbn.aau.dk/en/publications/creating-realistic-laboratory-settings-comparative-studies-of-thr>
- Kjeldskov, J. (2003). *Human-Computer Interaction Design for Emerging Technologies: Virtual Reality, Augmented Reality and Mobile Computer Systems*. <https://www.semanticscholar.org/paper/Human-Computer-Interaction-Design-for-Emerging-and-Kjeldskov/de5650db51b237907ce46e3b30fa2a32ee606a57>
- Kjeldskov, Jesper, & Graham, C. (2003). A Review of Mobile HCI Research Methods. In L. Chittaro (Éd.), *Human-Computer Interaction with Mobile Devices and Services* (p. 317-335). Springer Berlin Heidelberg.
- Kjeldskov, J., Skov, M. B., Als, B. S., & Høegh, R. T. (2004). Is It Worth the Hassle? Exploring the Added Value of Evaluating the Usability of Context-Aware Mobile Systems in the Field. In S. Brewster & M. Dunlop (Éds.), *Mobile Human-Computer Interaction - MobileHCI 2004* (p. 61-73). Springer. https://doi.org/10.1007/978-3-540-28637-0_6
- Kjeldskov, J., & Stage, J. (2004). New techniques for usability evaluation of mobile systems. *International Journal of Human-Computer Studies*, 60(5), 599-620. <https://doi.org/10.1016/j.ijhcs.2003.11.001>
- Koohang, A., & Paliszkievicz, J. (2015). E-Learning Courseware Usability: Building a Theoretical Model. *Journal of Computer Information Systems*. <https://www.tandfonline.com/doi/abs/10.1080/08874417.2015.11645801>
- Kools, M., Van de Wiel, M. W. J., Ruiter, R. A. C., & Kok, G. (2006). Pictures and text in instructions for medical devices: Effects on recall and actual performance. *Patient Education and Counseling*, 64(1-3), 104-111. <https://doi.org/10.1016/j.pec.2005.12.003>
- Kortum, P., & Peres, S. C. (2015). Evaluation of Home Health Care Devices: Remote Usability Assessment. *JMIR Human Factors*, 2(1), e10. <https://doi.org/10.2196/humanfactors.4570>
- Krieke, L. Van der Emerencia, A. C., Aiello, M., & Sytema, S. (2012). Usability Evaluation of a Web-Based Support System for People With a Schizophrenia Diagnosis. *Journal of Medical Internet Research*, 14(1), e24. <https://doi.org/10.2196/jmir.1921>

- Krishnamurti, T., Davis, A. L., Wong-Parodi, G., Fischhoff, B., Sadovsky, Y., & Simhan, H. N. (2017). Development and Testing of the MyHealthyPregnancy App: A Behavioral Decision Research-Based Tool for Assessing and Communicating Pregnancy Risk. *JMIR MHealth and UHealth*, 5(4), e42. <https://doi.org/10.2196/mhealth.7036>*
- Krueger, R. A., & Casey, M. A. (2009). *Focus Groups: A Practical Guide for Applied Research*. SAGE.
- Kushniruk, A., Nohr, C., Jensen, S., & Borycki, E. M. (2013). From Usability Testing to Clinical Simulations: Bringing Context into the Design and Evaluation of Usable and Safe Health Information Technologies. Contribution of the IMIA Human Factors Engineering for Healthcare Informatics Working Group. *Yearbook of Medical Informatics*, 8, 78-85.
- Kushniruk, A. W., & Patel, V. L. (2004). Cognitive and usability engineering methods for the evaluation of clinical information systems. *Journal of Biomedical Informatics*, 37(1), 56-76. <https://doi.org/10.1016/j.jbi.2004.01.003>
- Kushniruk, A. (2002). Evaluation in the design of health information systems: application of approaches emerging from usability engineering. *Computers in Biology and Medicine*, 32(3), 141-149.
- Kushniruk, Andre, & Turner, P. (2012). A framework for user involvement and context in the design and development of safe e-Health systems. *Studies in Health Technology and Informatics*, 180, 353-357.
- Lallemant, C., & Gronier, G. (2015). *Methodes de design UX: 30 méthodes fondamentales pour recevoir et évaluer les systèmes interactifs*. Editions Eyrolles.
- Lallemant, C., Gronier, G., & Koenig, V. (2015). User experience: A concept without consensus? Exploring practitioners' perspectives through an international survey. *Computers in Human Behavior*, 43, 35-48. <https://doi.org/10.1016/j.chb.2014.10.048>
- Lancry, A. (2009). *L'Ergonomie* (Presses Universitaires de France).
- Larsen, T. (2017). Nurses' instruction of patients in the use of INR-monitors for self-management of cardio-vascular conditions: Missed instructional opportunities. *Patient Education and Counseling*, 100(4), 673-681. <https://doi.org/10.1016/j.pec.2016.10.001>
- Larsen, T. (2018). Nurses' elicitation of patient error as a practice in training end-stage renal patients in automated home peritoneal dialysis. *Sociology of Health & Illness*, 40(5), 807-827. <https://doi.org/10.1111/1467-9566.12721>
- Laubie, R. (2011). Le patient connecté ou les métamorphoses de la santé. *L'Expansion Management Review*, N° 143(4), 24-31. <https://www.cairn.info/revue-l-expansion-management-review-2011-4-page-24.htm>
- Le livre Blanc du Conseil national de l'Ordre des médecins. (2015). *Santé Connectée : de la e-santé à la santé connectée* (p. 36p). <https://www.conseil-national.medecin.fr/sites/default/files/external-package/edition/lu5yh9/medecins->

- Lefebvre-Hoang, I., & Yilmaz, E. (2019). *Etat des lieux des pratiques de chirurgie ambulatoire en 2016* (Les dossiers de la DREES N° 41; p. 1-30). <https://drees.solidarites-sante.gouv.fr/IMG/pdf/dd41.pdf>
- Leichter, S. B., Nieman, J. A., Moore, R. W., Collins, P., & Rhodes, A. (1981). Readability of Self-Care Instructional Pamphlets for Diabetic Patients. *Diabetes Care*, 4(6), 627-630. <https://doi.org/10.2337/diacare.4.6.627>
- Lemaire, P., & Didierjean, A. (2018). *Introduction à la psychologie cognitive*. De Boeck Supérieur.
- Lemarié, J., Castillan, L., & Eyrolle, H. (2017). Effects of expertise and multimedia presentation on the enactment and recall of procedural instructions. *Psychologie Française*, 62(4), 351-359. <https://doi.org/10.1016/j.psfr.2016.07.002>
- Leon, C., & Beck, F. (2014). *Les comportements de santé des 55-85 ans. Analyses du Baromètre santé 2010*. INPES.
- Leplat, J. (1980). *La psychologie ergonomique*. PUF.
- LeRouge, C., Ma, J., Sneha, S., & Tolle, K. (2013). User profiles and personas in the design and development of consumer health technologies. *International Journal of Medical Informatics*, 82(11), e251-e268. <https://doi.org/10.1016/j.ijmedinf.2011.03.006>
- Lewis, C., Polson, P. G., Wharton, C., & Rieman, J. (1990). Testing a Walkthrough Methodology for Theory-based Design of Walk-up-and-use Interfaces. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 235–242. <https://doi.org/10.1145/97243.97279>
- Lewis, J. R. (2014). Usability: Lessons Learned ... and Yet to Be Learned. *International Journal of Human-Computer Interaction*, 30(9), 663-684. <https://doi.org/10.1080/10447318.2014.930311>
- Lewis, J. R., & Sauro, J. (2009). The Factor Structure of the System Usability Scale. *Human Centered Design*, 94-103. https://doi.org/10.1007/978-3-642-02806-9_12
- Liang, J., Xian, D., Liu, X., Fu, J., Zhang, X., Tang, B., & Lei, J. (2018). Usability Study of Mainstream Wearable Fitness Devices: Feature Analysis and System Usability Scale Evaluation. *JMIR MHealth and UHealth*, 6(11), e11066. <https://doi.org/10.2196/11066>
- Lindgaard, G., & Dudek, C. (2003). What is this evasive beast we call user satisfaction? *Interacting with Computers*, 15, 429–452. [doi:10.1016/S0953-5438\(02\)00063-2](https://doi.org/10.1016/S0953-5438(02)00063-2)
- Lindwall, O., & Ekström, A. (2012). Instruction-in-Interaction: The Teaching and Learning of a Manual Skill. *Human Studies*, 35(1), 27-49. <https://doi.org/10.1007/s10746-012->

- Lippeveld, T., Sauerborn, R., Bodart, C., & Organization, W. H. (2000). *Design and implementation of health information systems*. World Health Organization. <https://apps.who.int/iris/handle/10665/42289>
- Liu, H.-Y. T., Chia, R.-M., Setiawan, I. M. A., Crytzer, T. M., & Ding, D. (2019). Development of « My Wheelchair Guide » app: a qualitative study. *Disability and Rehabilitation. Assistive Technology*, 14(8), 839-848. <https://doi.org/10.1080/17483107.2018.1499140>
- Loorbach, N., Karreman, J., & Steehouder, M. (2007, August). *Adding Motivational Elements to an Instruction Manual for Seniors: Effects on Usability and Motivation* [Text]. <https://www.ingentaconnect.com/content/stc/tc/2007/00000054/00000003/art00007#expand/collapse>
- Loup-Escande, É., Burkhardt, J.-M., & Richir, S. (2013). Anticiper et évaluer l'utilité dans la conception ergonomique des technologies émergentes : une revue. *Le travail humain*, Vol. 76(1), 27-55. <https://www.cairn.info/revue-le-travail-humain-2013-1-page-27.htm>
- Lyles, C. R., Sarkar, U., & Osborn, C. Y. (2014). Getting a technology-based diabetes intervention ready for prime time: a review of usability testing studies. *Current Diabetes Reports*, 14(10), 534. <https://doi.org/10.1007/s11892-014-0534-9>
- Loup-Escande, E., Valentin, A., Metayer, N., Ricquebourd, V., Bidard, A., Damay, P., Boisson, C., Cuvillon, P., & Chaniaud, N. (2019). Adaptation de la démarche ergonomique en phase amont de la conception de technologies e-santé : exemple du projet Smart Angel. *EPIQUE 2019*.
- Macefield, R. (2007). Usability studies and the Hawthorne effect. *Psychology*, 2(3), 145-154. <https://doi.org/13964590>
- Mackert, M., Mabry-Flynn, A., Champlin, S., Donovan, E. E., & Pounders, K. (2016). Health Literacy and Health Information Technology Adoption: The Potential for a New Digital Divide. *Journal of Medical Internet Research*, 18(10), e264. <https://doi.org/10.2196/jmir.6349>
- Mahieu, A., & Raffy-Pihan, N. (1997). *La chirurgie ambulatoire en France, bilan et perspectives* (p. 73).
- Mahlke, S. (2008). *User Experience of Interaction with Technical Systems*. <http://dx.doi.org/10.14279/depositonce-1793>
- Mansfield, E. D., Wahba, R., Gillis, D. E., Weiss, B. D., & L'Abbé, M. (2018). Canadian adaptation of the Newest Vital Sign©, a health literacy assessment tool. *Public Health Nutrition*, 21(11), 2038-2045. <https://doi.org/10.1017/S1368980018000253>
- Marcilly, R., Peute, L., & Beuscart-Zephir, M.-C. (2016). From Usability Engineering to Evidence-based Usability in Health IT. *Studies in Health Technology and Informatics*,

- Marcus, A. (2002). Global and intercultural user-interface design. In *The human-computer interaction handbook: fundamentals, evolving technologies and emerging applications* (p. 441–463). L. Erlbaum Associates Inc.
- Margat, A., Gagnayre, R., Lombrail, P., Andrade, V. de, & Azogui-Levy, S. (2017). Interventions en littératie en santé et éducation thérapeutique : une revue de la littérature. *Sante Publique, Vol. 29(6)*, 811-820. <https://www.cairn.info/revue-sante-publique-2017-6-page-811.htm>
- Martínez-Pérez, B., de la Torre-Díez, I., & López-Coronado, M. (2015). Experiences and Results of Applying Tools for Assessing the Quality of a mHealth App Named Heartkeeper. *Journal of Medical Systems, 39(11)*, 142. <https://doi.org/10.1007/s10916-015-0303-6>
- Masny, D., & Dufresne, T. (2007). Lire au 21e siècle : la perspective des littératies multiples. In A.-M. Dionne & M.-J. Berger, *Les littératies: perspectives linguistique, familiale et culturelle* (p. 209-224). Presses de l'Université d'Ottawa. <https://www.icrml.ca/fr/recherches-et-publications/references/item/67113-lire-au-21e-siecle-la-perspective-des-litteraties-multiples>
- Matthew-Maich, N., Harris, L., Ploeg, J., Markle-Reid, M., Valaitis, R., Ibrahim, S., Gafni, A., & Isaacs, S. (2016). Designing, Implementing, and Evaluating Mobile Health Technologies for Managing Chronic Conditions in Older Adults: A Scoping Review. *JMIR MHealth and UHealth, 4(2)*, e29. <https://doi.org/10.2196/mhealth.5127>
- Mayberry, L. S., Berg, C. A., Harper, K. J., & Osborn, C. Y. (2016). The Design, Usability, and Feasibility of a Family-Focused Diabetes Self-Care Support mHealth Intervention for Diverse, Low-Income Adults with Type 2 Diabetes. *Journal of Diabetes Research, 2016*, 7586385. <https://doi.org/10.1155/2016/7586385>
- Mayer, R. E. (2001). *Multimedia Learning*. Cambridge University Press. <https://hal.archives-ouvertes.fr/hal-00702927>
- Mayer, R. E. (2003). Elements of a Science of E-Learning. *Journal of Educational Computing Research, 29(3)*, 297–313. <https://doi.org/10.2190/YJLG-09F9-XKAX-753D>
- Mayer, R. E., & Moreno, R. (1998). A split-attention effect in multimedia learning: Evidence for dual processing systems in working memory. *Journal of Educational Psychology, 90(2)*, 312–320. <https://doi.org/10.1037/0022-0663.90.2.312>
- McLellan, S., Muddimer, A., & Peres, S. C. (2012). The effect of experience on System Usability Scale ratings. *Journal of Usability Studies, 7*, 56–67.
- Mennecier, D. (2014). *De la e-Santé à la m-Santé : innovation mais attention aux données* [Académie nationale de médecine].
- Metelmann, B., Metelmann, C., Schuffert, L., Hahnenkamp, K., & Brinkrolf, P. (2018). Medical Correctness and User Friendliness of Available Apps for Cardiopulmonary

- Resuscitation: Systematic Search Combined With Guideline Adherence and Usability Evaluation. *JMIR MHealth and UHealth*, 6(11), e190. <https://doi.org/10.2196/mhealth.9651>
- Lallemand, C., Gronier, G. (2018). *Méthodes de design UX*. Éditions Eyrolles. <https://www.editions-eyrolles.com/Livre/9782212141436/methodes-de-design-ux>
- Mich, L., Franch, M., & Cilione, G. (2003). The 2QCV3Q Quality Model for the Analysis of Web Site Requirements. *J. Web Eng.*, 2(1), 105–127. <http://dl.acm.org/citation.cfm?id=2011119.2011126>
- Midler, C., Garel, G., & Kessler, A. (1997). Le co-développement, définitions, enjeux et problèmes. *Éducation permanente*, 131, 95-108. <https://hal.archives-ouvertes.fr/hal-00263243>
- Midler, C., Maniak, R., & Beaume, R. (2007). *Du co-développement à la co-innovation - Analyse empirique des coopérations verticales en conception innovante*. 15th GERPISA International Colloquium. <https://hal.archives-ouvertes.fr/hal-00263194>
- Mirkovic, J., Kaufman, D. R., & Ruland, C. M. (2014). Supporting Cancer Patients in Illness Management: Usability Evaluation of a Mobile App. *JMIR MHealth and UHealth*, 2(3), e33. <https://doi.org/10.2196/mhealth.3359>
- Monkman, H., Fellow, A. W. K. (ACMI, & Fellow), C. (2015). eHealth literacy issues, constructs, models, and methods for health information technology design and evaluation. *Knowledge Management & E-Learning: An International Journal*, 7(4), 541-549. <http://www.kmel-journal.org/ojs/index.php/online-publication/article/view/305>
- Monkman, H., & Kushniruk, A. (2013). A health literacy and usability heuristic evaluation of a mobile consumer health application. *Studies in Health Technology and Informatics*, 192, 724-728.
- Monkman, H., & Kushniruk, A. W. (2015). The Consumer Health Information System Adoption Model. *Studies in Health Technology and Informatics*, 218, 26-31.
- Moore, M. (1999). The evolution of telemedicine. *Future Generation Computer Systems*, 15(2), 245-254. [https://doi.org/10.1016/S0167-739X\(98\)00067-3](https://doi.org/10.1016/S0167-739X(98)00067-3)
- Moradian, S., Krzyzanowska, M. K., Maguire, R., Morita, P. P., Kukreti, V., Avery, J., Liu, G., Cafazzo, J., & Howell, D. (2018). Usability Evaluation of a Mobile Phone–Based System for Remote Monitoring and Management of Chemotherapy-Related Side Effects in Cancer Patients: Mixed-Methods Study. *JMIR Cancer*, 4(2), e10932. <https://doi.org/10.2196/10932>
- Morrell, R. W., Park, D. C., & Poon, L. W. (1990). Effects of Labeling Techniques on Memory and Comprehension of Prescription Information in Young and Old Adults. *Journal of Gerontology*, 45(4), P166–P172. <https://doi.org/10.1093/geronj/45.4.P166>

- Mousavi, S. Y., Low, R., & Sweller, J. (1995). Reducing cognitive load by mixing auditory and visual presentation modes. *Journal of Educational Psychology*, 87(2), 319–334. <https://doi.org/10.1037/0022-0663.87.2.319>
- Mõttus, R., Johnson, W., Murray, C., Wolf, M. S., Starr, J. M., & Deary, I. J. (2014). Towards understanding the links between health literacy and physical health. *Health Psychology: Official Journal of the Division of Health Psychology, American Psychological Association*, 33(2), 164-173. <https://doi.org/10.1037/a0031439>
- Murphy, P. W., Davis, T. C., Long, S. W., Jackson, R. H., & Decker, B. C. (1993). Rapid Estimate of Adult Literacy in Medicine (REALM): A Quick Reading Test for Patients. *Journal of Reading*, 37(2), 124-130. JSTOR. www.jstor.org/stable/40033408
- Murray, T., Clermont, Y., & Binkley, M. (2005). *Mesurer la litt ratie et les comp tences des adultes: des nouveaux cadres d' valuation* [Statistique Canada].
- Mykityshyn, A. L., Fisk, A. D., & Rogers, W. A. (2002). Learning to use a home medical device: mediating age-related differences with training. *Human Factors*, 44(3), 354-364. <https://doi.org/10.1518/0018720024497727>
- Nagamachi, M. (1995). Kansei Engineering: A new ergonomic consumer-oriented technology for product development. *International Journal of Industrial Ergonomics*. [https://doi.org/10.1016/0169-8141\(94\)00052-5](https://doi.org/10.1016/0169-8141(94)00052-5)
- Narv ez, S., Tobar,  . M., L pez, D. M., & Blobel, B. (2016). Human-Centered Design of an mHealth App for the Prevention of Burnout Syndrome. *Studies in Health Technology and Informatics*, 228, 215-219.
- Nelson, L. A., Mulvaney, S. A., Johnson, K. B., & Osborn, C. Y. (2017). mHealth Intervention Elements and User Characteristics Determine Utility: A Mixed-Methods Analysis. *Diabetes Technology & Therapeutics*, 19(1), 9-17. <https://doi.org/10.1089/dia.2016.0294>
- Neubert, T.-A., Dusch, M., Karst, M., & Beissner, F. (2018). Designing a Tablet-Based Software App for Mapping Bodily Symptoms: Usability Evaluation and Reproducibility Analysis. *JMIR MHealth and UHealth*, 6(5), e127. <https://doi.org/10.2196/mhealth.8409>
- Ng, A. H., Crowe, T. C., Ball, K., & Rasmussen, B. (2019). A mHealth Support Program for Australian Young Adults with Type 1 Diabetes: A Mixed Methods Study. *Digital Health*, 5, 2055207619882179. <https://doi.org/10.1177/2055207619882179>
- Nguyen, M. H., van Weert, J. C. M., Bol, N., Loos, E. F., Tytgat, K. M. A. J., van de Ven, A. W. H., & Smets, E. M. A. (2017). Tailoring the Mode of Information Presentation: Effects on Younger and Older Adults' Attention and Recall of Online Information. *Human Communication Research*, 43(1), 102-126. <https://doi.org/10.1111/hcre.12097>
- Nielsen, C. (1998). *Testing in the Field | Proceedings of the Third Asian Pacific Computer and Human Interaction*. <https://dl.acm.org/doi/abs/10.5555/786112.786315>

- Nielsen, J. (1993). *Usability Engineering*. Morgan Kaufmann Publishers Inc.
- Nielsen, J. (1994a). *Usability Engineering*. Elsevier.
- Nielsen, J. (1994b). Usability Inspection Methods. *Conference Companion on Human Factors in Computing Systems*, 413–414. <https://doi.org/10.1145/259963.260531>
- Nielsen, J. (2005). Medical Usability: How to Kill Patients Through Bad Design. *Nielsen Norman Group*. <https://www.nngroup.com/articles/medical-usability/>
- Nielsen, J., & Molich, R. (1990). Heuristic Evaluation of User Interfaces. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 249–256. <https://doi.org/10.1145/97243.97281>
- Nishiuchi, N., & Takahashi, Y. (2015). Objective Evaluation Method of Usability Using Parameters of User's Fingertip Movement. In M. L. Gavrilova, C. J. K. Tan, K. Saeed, N. Chaki, & S. H. Shaikh (Éds.), *Transactions on Computational Science XXV* (p. 77-89). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-662-47074-9_5
- Nitsch, M., Adamcik, T., Kuso, S., Zeiler, M., & Waldherr, K. (2019). Usability and Engagement Evaluation of an Unguided Online Program for Promoting a Healthy Lifestyle and Reducing the Risk for Eating Disorders and Obesity in the School Setting. *Nutrients*, *11*(4). <https://doi.org/10.3390/nu11040713>
- Nitsch, M., Dimopoulos, C. N., Flaschberger, E., Saffran, K., Kruger, J. F., Garlock, L., Wilfley, D. E., Taylor, C. B., & Jones, M. (2016). A guided online and mobile self-help program for individuals with eating disorders: An iterative engagement and usability study. *Journal of Medical Internet Research*, *18*(1).
- Norman, D. A. (1987). *Interfacing Thought: Cognitive Aspects of Human-computer Interaction* (J. M. Carroll (Éd.); p. 325–336). MIT Press. <http://dl.acm.org/citation.cfm?id=28446.28458>
- Nunnally, J. C. (1978). *Psychometric theory* (McGraw Hill).
- O'Brien, H. L., & Toms, E. G. (2008). What is user engagement? A conceptual framework for defining user engagement with technology. *Journal of the American Society for Information Science and Technology*, *59*(6), 938-955. <https://doi.org/10.1002/asi.20801>
- Olsina, L., Lafuente, G., Lafuente, G., & Ross, G. (s. d.). *Using Web-site QEM to Measure the Quality of Sites*. 3.
- Ose, D., Kunz, A., Pohlmann, S., Hofmann, H., Qreini, M., Krisam, J., Uhlmann, L., Jacke, C., Winkler, E. C., Salize, H.-J., & Szecsenyi, J. (2017). A Personal Electronic Health Record: Study Protocol of a Feasibility Study on Implementation in a Real-World Health Care Setting. *JMIR Research Protocols*, *6*(3), e33. <https://doi.org/10.2196/resprot.6314>

- Paas, F., Van Gerven, P., & Tabbers, H. (2005). The cognitive aging principle in multimedia learning. In R. Mayer E., *The Cambridge Handbook of multimedia learning* (Cambridge University Press, pp. 339–351).
- Paasche-Orlow, M. K., & Wolf, M. S. (2007). The causal pathways linking health literacy to health outcomes. *American Journal of Health Behavior*, *31 Suppl 1*, S19-26. <https://doi.org/10.5555/ajhb.2007.31.supp.S19>
- Papalambros, P. Y. (2015). Design Science: Why, What and How. *Design Science*, *1*. <https://doi.org/10.1017/dsj.2015.1>
- Paré, G., Jaana, M., & Sicotte, C. (2007). Systematic Review of Home Telemonitoring for Chronic Diseases: The Evidence Base. *Journal of the American Medical Informatics Association : JAMIA*, *14*(3), 269-277. <https://doi.org/10.1197/jamia.M2270>
- Parker, R. (2000). Health literacy: a challenge for American patients and their health care providers. *Health Promotion International*, *15*(4), 277–283. <https://doi.org/10.1093/heapro/15.4.277>
- Parker, R. M., Baker, D. W., Williams, M. V., & Nurss, J. R. (1995). The test of functional health literacy in adults: a new instrument for measuring patients' literacy skills. *Journal of General Internal Medicine*, *10*(10), 537-541. <https://doi.org/10.1007/bf02640361>
- Parmanto, B., Pramana, G., Yu, D. X., Fairman, A. D., Dicianno, B. E., & McCue, M. P. (2013). iMHere: A Novel mHealth System for Supporting Self-Care in Management of Complex and Chronic Conditions. *JMIR MHealth and UHealth*, *1*(2), e10. <https://doi.org/10.2196/mhealth.2391>
- Patel, V. L., Eisemon, T. O., & Arocha, J. F. (1990). Comprehending instructions for using pharmaceutical products in rural Kenya. *Instructional Science*, *19*(1), 71–84. <https://doi.org/10.1007/BF00377986>
- Pelayo, S., Costa, S. B. D., & Beuscart-Zéphir, M.-C. (2012). Intégration des Facteurs Humains dans les démarches de conception et d'implémentation des NTIC en Santé. *IRBM*, *33*(3), 236-239. <https://doi.org/10.1016/j.irbm.2012.04.002>
- Pelayo, Sylvia, Watbled, L., Guerlinger, S., & Beuscart-Zéphir, M.-C. (2012). Prise en compte des facteurs humains pour une implémentation sûre et efficace des nouvelles technologies de l'information en santé. In P. M. Staccini, A. Harmel, S. J. Darmoni, & R. Gouider (Éds.), *Systèmes d'information pour l'amélioration de la qualité en santé: Comptes rendus des quatorzièmes Journées francophones d'informatique médicale, Tunis, 23 et 24 septembre 2011* (p. 321-330). Springer. https://doi.org/10.1007/978-2-8178-0285-5_28
- Perrot, C., Meillon, B., Mandran, N., & Jambon, F. (2014). Evaluation des systèmes mobiles et ubiquitaires: proposition de méthodologie et retours d'expérience. *Journal d'Interaction Personne-Système, Volume 1*. <https://jips.episciences.org/61/pdf>

- Petrie, H., & Bevan, N. (2009). The Evaluation of Accessibility, Usability, and User Experience. *The Universal Access Handbook*.
<https://doi.org/10.1201/9781420064995-c20>
- Peute, L. W., Driest, K. F., Marcilly, R., Bras Da Costa, S., Beuscart-Zephir, M.-C., & Jaspers, M. W. M. (2013). A framework for reporting on human factor/usability studies of health information technologies. *Studies in Health Technology and Informatics*, 194, 54-60.
- Peute, L. W., Lichtner, V., Baysari, M. T., Hägglund, M., Homco, J., Jansen-Kosterink, S., Jauregui, I., Kaipio, J., Kuziemy, C. E., Lehnbohm, E. C., Leite, F., Lesselroth, B., Luna, D., Otero, C., Pedersen, R., Pelayo, S., Santos, R., Silva, N.-A., Tyllinen, M., ... Marcilly, R. (2020). Challenges and Best Practices in Ethical Review of Human and Organizational Factors Studies in Health Technology: a Synthesis of Testimonies. *Yearbook of Medical Informatics*. <https://doi.org/10.1055/s-0040-1701979>
- Peute, L. W. P., Spithoven, R., Bakker, P. J. M., & Jaspers, M. W. M. (2008). Usability studies on interactive health information systems; where do we stand? *Studies in Health Technology and Informatics*, 136, 327-332.
- Piau, A., Campo, E., Rumeau, P., Vellas, B., & Nourhashemi, F. (2014). Aging society and gerontechnology: A solution for an independent living? *The Journal of Nutrition, Health & Aging*, 18(1), 97–112. <https://doi.org/10.1007/s12603-013-0356-5>
- Powell, A. C., Torous, J., Chan, S., Raynor, G. S., Shwartz, E., Shanahan, M., & Landman, A. B. (2016). Interrater Reliability of mHealth App Rating Measures: Analysis of Top Depression and Smoking Cessation Apps. *JMIR MHealth and UHealth*, 4(1), e15. <https://doi.org/10.2196/mhealth.5176>
- Prado-Jimenez, D., Rosales-Cinco, R. A., Aceves-Gonzalez, C., & Lang, A. (2018). *Self-monitoring of blood glucose: The user experience and device interactions of older adults with type 2 diabetes mellitus | Avances en Interacción Humano-Computadora*. 1(3), 65-68. <http://publicaciones.amexihc.org/index.php/aihc/article/view/49>
- Preece, J., Rogers, Y., Sharp, H., Benyon, D., Holland, S., & Carey, T. (1994). *Human-Computer Interaction*. Addison-Wesley Longman Ltd.
- Prentice, R. (2007). Drilling Surgeons: The Social Lessons of Embodied Surgical Learning. *Science, Technology, & Human Values*, 32(5), 534-553. <https://doi.org/10.1177/0895904805303201>
- Price, M., Sawyer, T., Harris, M., & Skalka, C. (2016). Usability Evaluation of a Mobile Monitoring System to Assess Symptoms After a Traumatic Injury: A Mixed-Methods Study. *JMIR Mental Health*, 3(1), e3. <https://doi.org/10.2196/mental.5023>
- Pruitt, J., & Grudin, J. (2003). Personas: practice and theory. *DUX '03*. <https://doi.org/10.1145/997078.997089>
- Pulantara, I. W., Parmanto, B., & Germain, A. (2018). Development of a Just-in-Time Adaptive mHealth Intervention for Insomnia: Usability Study. *JMIR Human Factors*,

5(2), e21. <https://doi.org/10.2196/humanfactors.8905>

- Quaedackers, L., De Wit, J., Pillen, S., Van Gilst, M., Batalas, N., Lammers, G. J., Markopoulos, P., & Overeem, S. (2020). A Mobile App for Longterm Monitoring of Narcolepsy Symptoms: Design, Development, and Evaluation. *JMIR MHealth and UHealth*, 8(1), e14939. <https://doi.org/10.2196/14939>
- Rahimi, B., & Vimarlund, V. (2007). Methods to evaluate health information systems in healthcare settings: a literature review. *Journal of Medical Systems*, 31(5), 397-432. <https://doi.org/10.1007/s10916-007-9082-z>
- Rahmat, H., Zulzalil, H., Ghani, A. A. A., & Kamaruddin, A. (2018, mars). *A Comprehensive Usability Model for Evaluating Smartphone Apps* [Text]. <https://doi.org/info:doi/10.1166/asl.2018.11125>
- Rajput, Z. A., Mbugua, S., Amadi, D., Chepngeno, V., Saleem, J. J., Anokwa, Y., Hartung, C., Borriello, G., Mamlin, B. W., Ndege, S. K., & Were, M. C. (2012). Evaluation of an Android-based mHealth system for population surveillance in developing countries. *Journal of the American Medical Informatics Association*, 19(4), 655-659. <https://doi.org/10.1136/amiajnl-2011-000476>
- Rantanen, P., Parkkari, T., Leikola, S., Airaksinen, M., & Lyles, A. (2017). An In-home Advanced Robotic System to Manage Elderly Home-care Patients' Medications: A Pilot Safety and Usability Study. *Clinical Therapeutics*, 39(5), 1054-1061. <https://doi.org/10.1016/j.clinthera.2017.03.020>
- Rasmussen, J. (1983). Skills, rules, and knowledge; signals, signs, and symbols, and other distinctions in human performance models. *IEEE Transactions on Systems, Man, and Cybernetics*, SMC-13(3), 257-266. <https://doi.org/10.1109/TSMC.1983.6313160>
- Reason, J. (2013). *L'erreur humaine* (Mines ParisTech). <https://www.pressesdesmines.com/produit/l-erreur-humaine/>
- Reolon, M., Lacerda, T. C., Krone, C., Wangenheim, C. G. von, Xafranski, J., Nunes, J. V., & Wangenheim, A. von. (2016). *Usability Heuristics for Evaluating Health-care Applications for Smartphones : A Systematic Literature*.
- Rettig, M. (1991). Nobody reads documentation. *Communications of the ACM*, 34(7), 19-24. <https://doi.org/10.1145/105783.105788>
- Rico-Olarte, C., López, D. M., & Kepplinger, S. (2018). Towards a Conceptual Framework for the Objective Evaluation of User Experience. In A. Marcus & W. Wang (Éds.), *Design, User Experience, and Usability: Theory and Practice* (p. 546-559). Springer International Publishing.
- Robinson, M. D., Branham, A. R., Locklear, A., Robertson, S., & Gridley, T. (2016). Measuring satisfaction and usability of facetime for virtual visits in patients with uncontrolled diabetes. *Telemedicine and e-Health*, 22(2), 138-143.

- Rochat, J., Weinhold, T., Lovis, C., Schiro, J., Marcilly, R., & Pelayo, S. (2017). *A Literature Review to Define Concepts and Dimensions of Ecological Validity/Fidelity for Usability Validation*. Informatics for Health, Manchester, UK.
<https://doi.org/10.13140/RG.2.2.31190.19525>
- Rogers, C. R. (1942). *Counseling and psychotherapy: newer concepts in practice*. Houghton Mifflin company.
- Root, R. W., & Draper, S. (1983). Questionnaires As a Software Evaluation Tool. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 83–87. <https://doi.org/10.1145/800045.801586>
- Rootman, I., & Ronson, B. (2005). Literacy and health research in Canada: where have we been and where should we go? *Canadian Journal of Public Health = Revue Canadienne De Sante Publique*, 96 Suppl 2, S62-77.
- Rouquette, A., Nadot, T., Labitrie, P., Van den Broucke, S., Mancini, J., Rigal, L., & Ringa, V. (2018). Validity and measurement invariance across sex, age, and education level of the French short versions of the European Health Literacy Survey Questionnaire. *PLoS ONE*, 13(12). <https://doi.org/10.1371/journal.pone.0208091>
- Rubin, J. (1994). *Handbook of Usability Testing: How to Plan, Design, and Conduct Effective Tests* (1st éd.). John Wiley & Sons, Inc.
- Rubin, J. et Chisnell, D. (2008). *Handbook of Usability Testing: How to Plan, Design, and Conduct Effective Tests*. Second Edition, Wiley.
- Rudd, R. E., Anderson, J. E., Oppenheimer, S., Nath, C., Comings, J. P., Garner, B., Smith, C. (2007). Health literacy: An update of patient and public health literature, *Review of Adult learning and Literacy*, 7.
- Ruel, J., Alaire, C., Moreau, A., Kassi, B., Brumagne, A., Delample, A., Grisard, C., & Pinto da Silva, F. (2018). Communiquer pour tous - Guide pour une information accessible. *Santé publique France*, 116.
<http://inpes.santepubliquefrance.fr/accessible/pdf/communiquerpourtous.pdf>
- Sadegh, S. S., Khakshour Saadat, P., Sepehri, M. M., & Assadi, V. (2018). A framework for m-health service development and success evaluation. *International Journal of Medical Informatics*, 112, 123-130. <https://doi.org/10.1016/j.ijmedinf.2018.01.003>
- Sagnier, C., Loup-Escande, É., & Valléry, G. (2019). Acceptabilité de la réalité virtuelle : une revue de la littérature. *Le travail humain*, Vol. 82(3), 183-212.
<https://www.cairn.info/revue-le-travail-humain-2019-3-page-183.htm>
- Sagot, J.-C., Gouin, V., & Gomes, S. (2003). Ergonomics in product design: safety factor. *Safety Science*, 41(2), 137-154. [https://doi.org/10.1016/S0925-7535\(02\)00038-3](https://doi.org/10.1016/S0925-7535(02)00038-3)
- Sarkar, U., Gourley, G. I., Lyles, C. R., Tieu, L., Clarity, C., Newmark, L., Singh, K., & Bates, D. W. (2016). Usability of Commercially Available Mobile Applications for Diverse Patients. *Journal of General Internal Medicine*, 31(12), 1417-1426.

<https://doi.org/10.1007/s11606-016-3771-6>

- Sauro, J. (2011). *A practical guide to the system usability scale: Background, benchmarks & best practices*. Measuring Usability LLC Denver, CO.
- Savage, P. A., Kemp, D. G., & Payne, D. E. (1991). Auditory versus Visual Presentation of Help Messages. *Proceedings of the Human Factors Society Annual Meeting*, 35(4), 244–248. <https://doi.org/10.1177/154193129103500408>
- Schmidt, J. D. E., & De Marchi, A. C. B. (2017). Usability evaluation methods for mobile serious games applied to health: a systematic review. *Universal Access in the Information Society*, 16(4), 921-928. <https://doi.org/10.1007/s10209-016-0511-y>
- Schmuckler, M. A. (2001). What Is Ecological Validity? A Dimensional Analysis. *Infancy*, 2(4), 419-436. https://doi.org/10.1207/S15327078IN0204_02
- Schnall, R., Bakken, S., Brown Iii, W., Carballo-Diequez, A., & Iribarren, S. (2016). Usability Evaluation of a Prototype Mobile App for Health Management for Persons Living with HIV. *Studies in Health Technology and Informatics*, 225, 481-485.
- Schnall, R., Rojas, M., Bakken, S., Brown, W., Carballo-Diequez, A., Carry, M., Gelaude, D., Mosley, J. P., & Travers, J. (2016a). A user-centered model for designing consumer mobile health (mHealth) applications (apps). *Journal of Biomedical Informatics*, 60, 243-251. <https://doi.org/10.1016/j.jbi.2016.02.002>*
- Schnall, R., Rojas, M., Bakken, S., Brown, W., Carballo-Diequez, A., Carry, M., Gelaude, D., Mosley, J. P., & Travers, J. (2016b). A user-centered model for designing consumer mobile health (mHealth) applications (apps). *Journal of Biomedical Informatics*, 60, 243-251. <https://doi.org/10.1016/j.jbi.2016.02.002>
- Schoffman, D. E., Turner-McGrievy, G., Jones, S. J., & Wilcox, S. (2013). Mobile apps for pediatric obesity prevention and treatment, healthy eating, and physical activity promotion: just fun and games? *Translational Behavioral Medicine*, 3(3), 320-325. <https://doi.org/10.1007/s13142-013-0206-3>
- Scholtz, J. (2004). Usability evaluation. *National Institute of Standards and Technology*, 1.
- Scott, A. R., Alore, E. A., Naik, A. D., Berger, D. H., & Suliburk, J. W. (2017). Mixed-Methods Analysis of Factors Impacting Use of a Postoperative mHealth App. *JMIR MHealth and UHealth*, 5(2), e11. <https://doi.org/10.2196/mhealth.6728>
- Seffah, A., Donyaee, M., Kline, R. B., & Padda, H. K. (2006). Usability measurement and metrics: A consolidated model. *Software Quality Journal*, 14(2), 159-178. <https://doi.org/10.1007/s11219-006-7600-8>
- Seffah, A., Kececi, N., & Donyaee, M. (2001). QUIM: a framework for quantifying usability metrics in software quality models. *Proceedings Second Asia-Pacific Conference on Quality Software*, 311-318. <https://doi.org/10.1109/APAQS.2001.990036>

- Shackel, B., & Richardson, S. J. (1991). *Human Factors for Informatics Usability*. Cambridge University Press.
- Shackel, B. (1981). *The concept of usability Poughkeepsie* (Proceedings of IBM Software and Information Usability Symposium).
- Shackel, B. (1986). Ergonomics in design for usability York, UK: Proceedings of the Second Conference of the British Computer Society. *Human Computer Interaction Specialist Group on People and Computers: designing for usability*, 44-64.
- Shackel, B. (1990). Human factors and usability. In J. Preece & L. Keller, *Human-computer interaction: Selected readings* (Prentice Hall International, p. 27-41).
- Shackel, B. (1991). Usability--context, Framework, Definition, Design and Evaluation. Shackel, B. and Richardson, S., Ed. *Human Factors for Informatics Usability*. pp. 21-37. Cambridge, UK, Cambridge University Press.
- Shackel, B. (2009). Usability – Context, framework, definition, design and evaluation. *Interacting with Computers*, 21(5-6), 339-346. <https://doi.org/10.1016/j.intcom.2009.04.007>
- Shawgi, E., & Noureldien, N. A. (2015). Usability Measurement Model (UMM): A New Model for Measuring Websites Usability. *International Journal of Information Science*, 5(1), 5-13.
- Sheehan, B., Lee, Y., Rodriguez, M., Tiase, V., & Schnall, R. (2012). A Comparison of Usability Factors of Four Mobile Devices for Accessing Healthcare Information by Adolescents. *Applied Clinical Informatics*, 03(4), 356-366. <https://doi.org/10.4338/ACI-2012-06-RA-0021>
- Sheehan, B., & Lucero, R. J. (2015). Initial Usability and Feasibility Evaluation of a Personal Health Record-Based Self-Management System for Older Adults. *EGEMS (Washington, DC)*, 3(2), 1152. <https://doi.org/10.13063/2327-9214.1152>
- Shneiderman, B. (1997). *Designing the User Interface: Strategies for Effective Human-Computer Interaction* (3rd ed.). Addison-Wesley Longman Publishing Co., Inc.
- Shneiderman, B., & Plaisant, C. (2010). *Designing the user interface: strategies for effective human-computer interaction*. Pearson Education India.
- Shrout, P. E., & Fleiss, J. L. (1979). Intraclass correlations: Uses in assessing rater reliability. *Psychological Bulletin*, 86(2), 420-428. <https://doi.org/10.1037/0033-2909.86.2.420>
- Silver, N. C., & Wogalter, M. S. (1991). Strength and Understanding of Signal Words by Elementary and Middle School Students. *Proceedings of the Human Factors Society Annual Meeting*, 35(9), 590-594. <https://doi.org/10.1518/107118191786754671>
- Simon, P. (2017). Bénéfices et risques des objets connectés en télémédecine et santé mobile. *European Research in Telemedicine / La Recherche Européenne en Télémédecine*,

6(2), 47-57. <https://doi.org/10.1016/j.eurtel.2017.05.002>

Simon, P., Hansske, A., Boiron, P., & al. (2011). HIT Paris 2011. Systèmes d'information en santé et TIC. *HIT Paris 2011. Systèmes d'information en santé et TIC*, 539, 64-93.

Simon, P. (2015). *Télémédecine - Enjeux et pratiques*. Le Coudrier.

Sinatra, G. M., Heddy, B. C., & Lombardi, D. (2015). The Challenges of Defining and Measuring Student Engagement in Science. *Educational Psychologist*, 50(1), 1–13. <https://doi.org/10.1080/00461520.2014.1002924>

Skorin-Kapov, L., Dobrijevic, O., & Piplica, D. (2014). Towards Evaluating the Quality of Experience of Remote Patient Monitoring Services: A Study Considering Usability Aspects. *International Journal of Mobile Human Computer Interaction (IJMHCI)*, 6(4), 59-89. <https://doi.org/10.4018/ijmhci.2014100104>

Smaradottir, B., Gerdes, M., Fensli, R., & Martinez, S. (2015). Usability Evaluation of a COPD Remote Monitoring Application. *Studies in Health Technology and Informatics*, 210, 845-849.

Sonderegger, A., Schmutz, S., & Sauer, J. (2016). The influence of age in usability testing. *Applied Ergonomics*, 52, 291-300. <https://doi.org/10.1016/j.apergo.2015.06.012>

Sørensen, K., Broucke, S. V. den, Fullam, J., Doyle, G., Pelikan, J., Slonska, Z., & Brand, H. (2012). Health literacy and public health: A systematic review and integration of definitions and models. *BMC Public Health*, 12(1), 80. <https://doi.org/10.1186/1471-2458-12-80>

Sørensen, K., Pelikan, J. M., Röthlin, F., Ganahl, K., Slonska, Z., Doyle, G., Fullam, J., Kondilis, B., Agrafiotis, D., Uiters, E., Falcon, M., Mensing, M., Tchamov, K., van den Broucke, S., Brand, H., & HLS-EU Consortium. (2015). Health literacy in Europe: comparative results of the European health literacy survey (HLS-EU). *European Journal of Public Health*, 25(6), 1053-1058. <https://doi.org/10.1093/eurpub/ckv043>

Sørensen, K., Van den Broucke, S., Fullam, J., Doyle, G., Pelikan, J., Slonska, Z., & Brand, H. (2012). Health literacy and public health: A systematic review and integration of definitions and models. *BMC Public Health*, 12, 80. <https://doi.org/10.1186/1471-2458-12-80>

Sørensen, K., Van den Broucke, S., Fullam, J., Doyle, G., Pelikan, J., Slonska, Z., Brand, H., & (HLS-EU) Consortium Health Literacy Project European. (2012). Health literacy and public health: A systematic review and integration of definitions and models. *BMC Public Health*, 12(1), 80. <https://doi.org/10.1186/1471-2458-12-80>

Sørensen, K., Van den Broucke, S., Pelikan, J. M., Fullam, J., Doyle, G., Slonska, Z., Kondilis, B., Stoffels, V., Osborne, R. H., Brand, H., & HLS-EU Consortium. (2013). Measuring health literacy in populations: illuminating the design and development process of the European Health Literacy Survey Questionnaire (HLS-EU-Q). *BMC*

Public Health, 13, 948. <https://doi.org/10.1186/1471-2458-13-948>

- Sparkes, J., Valaitis, R., & McKibbin, A. (2012). A usability study of patients setting up a cardiac event loop recorder and BlackBerry gateway for remote monitoring at home. *Telemedicine Journal and E-Health: The Official Journal of the American Telemedicine Association*, 18(6), 484-490. <https://doi.org/10.1089/tmj.2011.0230>
- Sperandio, J.-C. (1980). *La Psychologie en ergonomie*. FeniXX.
- Sperandio, J.-C. (2001). *Critères ergonomiques de l'assistance technologique aux opérateurs*. Communication au congrès JIM'2001 : Interaction Homme – Machine & Assistance, Metz, France.
- Stead, W. W., Haynes, R. B., Fuller, S., Friedman, C. P., Travis, L. E., Beck, J. R., Fenichel, C. H., Chandrasekaran, B., Buchanan, B. G., & Abola, E. E. (1994). Designing medical informatics research and library--resource projects to increase what is learned. *Journal of the American Medical Informatics Association: JAMIA*, 1(1), 28-33. <https://doi.org/10.1136/jamia.1994.95236134>
- Stiles-Shields, C., Montague, E., Lattie, E. G., Schueller, S. M., Kwasny, M. J., & Mohr, D. C. (2017). Exploring User Learnability and Learning Performance in an App for Depression: Usability Study. *JMIR Human Factors*, 4(3), e18. <https://doi.org/10.2196/humanfactors.7951>
- Stoll, R. D., Pina, A. A., Gary, K., & Amresh, A. (2017). Usability of a Smartphone Application to Support the Prevention and Early Intervention of Anxiety in Youth. *Cognitive and Behavioral Practice*, 24(4), 393-404. <https://doi.org/10.1016/j.cbpra.2016.11.002>
- Stone, D. E., & Glock, M. D. (1981). How do young adults read directions with and without pictures? *Journal of Educational Psychology*, 73(3), 419-426. <https://doi.org/10.1037/0022-0663.73.3.419>
- Strich, T. G. (1985). Understanding Readers and their Uses of Texts. In T. M. Duffy & R. Willaer (Eds.), *Designing Usable Texts* (pp. 31-340). London: Academic Press Inc.
- Svensson, M. S., Luff, P., & Heath, C. (2009). Embedding instruction in practice: contingency and collaboration during surgical training. *Sociology of Health & Illness*, 31(6), 889-906. <https://doi.org/10.1111/j.1467-9566.2009.01195.x>
- Sweller, J., & Chandler, P. (1994). Why Some Material Is Difficult to Learn. *Cognition and Instruction*, 12(3), 185-233. https://doi.org/10.1207/s1532690xci1203_1
- Szlichcinski, K. P. (1979). Telling people how things work. *Applied Ergonomics*, 10(1), 2-8. [https://doi.org/10.1016/0003-6870\(79\)90002-4](https://doi.org/10.1016/0003-6870(79)90002-4)
- Tabbers, H. K., Martens, R. L., & van Merriënboer, J. J. G. (2004). Multimedia instructions and cognitive load theory: effects of modality and cueing. *The British Journal of Educational Psychology*, 74(Pt 1), 71-81. <https://doi.org/10.1348/000709904322848824>

- Tarricone, R., Torbica, A., & Drummond, M. (2017). Challenges in the Assessment of Medical Devices: The MedtechHTA Project: Challenges in the assessment of medical devices. *Health Economics*, 26, 5-12. <https://doi.org/10.1002/hec.3469>
- Tearl, D. K., Cox, T. J., & Hertzog, J. H. (2006). Hospital discharge of respiratory-technology-dependent children: role of a dedicated respiratory care discharge coordinator. *Respiratory Care*, 51(7), 744-749.
- Ten Haken, I., Ben Allouch, S., & van Harten, W. H. (2018). The use of advanced medical technologies at home: a systematic review of the literature. *BMC Public Health*, 18(1), 284. <https://doi.org/10.1186/s12889-018-5123-4>
- Teo, C. H., Ng, C. J., Lo, S. K., Lim, C. D., & White, A. (2019). A Mobile Web App to Improve Health Screening Uptake in Men (ScreenMen): Utility and Usability Evaluation Study. *JMIR MHealth and UHealth*, 7(4), e10216. <https://doi.org/10.2196/10216>
- Theureau, J., & Jeffroy, F. (Éds.). (1994). *Ergonomie des situations informatisées: la conception centrée sur le cours d'action des utilisateurs*. Octares éd.
- Tissier, I., & Gronier, G. (2014). Impact of usability, user experience and motivation on the engagement to using a mobile application. *Proceedings of the 26th Conference on l'Interaction Homme-Machine - IHM '14*, 195-200. <https://doi.org/10.1145/2670444.2670470>
- Tractinsky, N. (2018). The Usability Construct: A Dead End? *Human-Computer Interaction*, 33(2), 131-177. <https://doi.org/10.1080/07370024.2017.1298038>
- Triantafyllidis, A., Velardo, C., Chantler, T., Shah, S. A., Paton, C., Khorshidi, R., Tarassenko, L., & Rahimi, K. (2015). A personalised mobile-based home monitoring system for heart failure: The SUPPORT-HF Study. *International Journal of Medical Informatics*, 84(10), 743-753. <https://doi.org/10.1016/j.ijmedinf.2015.05.003>
- Tricot, A., Plé gat-Soutjis, F., Camps, J.-F., Amiel, A., Lutz, G., & Morcillo, A. (2003). *Utilité, utilisabilité, acceptabilité : interpréter les relations entre trois dimensions de l'évaluation des EIAH*. 391-402. <https://edutice.archives-ouvertes.fr/edutice-00000154>
- Van der Vaart, R., Driel, D. van, Pronk, K., Paulussen, S., Boekhorst, S. te, Rosmalen, J. G. M., & Evers, A. W. M. (2019). The Role of Age, Education, and Digital Health Literacy in the Usability of Internet-Based Cognitive Behavioral Therapy for Chronic Pain: Mixed Methods Study. *JMIR Formative Research*, 3(4), e12883. <https://doi.org/10.2196/12883>
- Van Hooijdonk, C., & Krahmer, E. (2008). Information modalities for procedural instructions: The influence of text, pictures, and film clips on learning and executing RSI exercises. *IEEE Transactions on Professional Communication*, 51(1), 50-62. <https://doi.org/10.1109/TPC.2007.2000054>

- Vermersch, P. (1985). Données d'observation sur l'utilisation d'une consigne écrite : l'atomisation de l'action. *Le Travail Humain*, 48(2), 161–172. JSTOR. <https://www.jstor.org/stable/40657326>
- Vermersch, P. (2019). *L'entretien d'explicitation*. ESF Sciences humaines.
- Vial, F., Bouaziz, H., Mekler, G., Cornet, C., Merle, M., & Laxenaire, M. C. (2000). [Postoperative pain and surgical treatment of trapeziometacarpal osteoarthritis of the thumb after ambulatory surgery]. *Annales Francaises D'anesthesie Et De Reanimation*, 19(9), 643-648.
- Voncken-Brewster, V., Moser, A., Weijden, T. van der, Nagykaldi, Z., Vries, H. de, & Tange, H. (2013). Usability Evaluation of an Online, Tailored Self-Management Intervention for Chronic Obstructive Pulmonary Disease Patients Incorporating Behavior Change Techniques. *JMIR Research Protocols*, 2(1), e3. <https://doi.org/10.2196/resprot.2246>
- Vooijs, M., Alpay, L. L., Snoeck-Stroband, J. B., Beerthuisen, T., Siemonsma, P. C., Abbink, J. J., Sont, J. K., & Rövekamp, T. A. (2014). Validity and Usability of Low-Cost Accelerometers for Internet-Based Self-Monitoring of Physical Activity in Patients With Chronic Obstructive Pulmonary Disease. *Interactive Journal of Medical Research*, 3(4), e14. <https://doi.org/10.2196/ijmr.3056>
- Weichbroth, P. (2018). Usability Attributes Revisited: A Time-Framed Knowledge Map. *2018 Federated Conference on Computer Science and Information Systems (FedCSIS)*, 1005-1008.
- Weick-Brady, M., & Lazerow, R. (2006). Medical devices: Promoting a safe Migration Into the Home. *Home Healthcare Nurse: The Journal for the Home Care and Hospice Professional*, 24(5), 298-304. [insights.ovid.com](https://www.insights.ovid.com)
- Weiss, B. D., Mays, M. Z., Martz, W., Castro, K. M., DeWalt, D. A., Pignone, M. P., Mockbee, J., & Hale, F. A. (2005). Quick assessment of literacy in primary care: the newest vital sign. *Annals of Family Medicine*, 3(6), 514-522. <https://doi.org/10.1370/afm.405>
- Werner, N. E., Jolliff, A. F., Casper, G., Martell, T., & Ponto, K. (2018). Home is where the head is: a distributed cognition account of personal health information management in the home among those with chronic illness. *Ergonomics*, 61(8), 1065-1078. <https://doi.org/10.1080/00140139.2018.1435910>
- WHO. (2016). *WHO Global Model Regulatory Framework for Medical Devices including in vitro diagnostic medical devices* (World Health Organization). http://www.who.int/medical_devices/publications/global_model_regulatory_framework_meddev/en/
- Wildenbos, G. A., Peute, L., & Jaspers, M. (2018). Aging barriers influencing mobile health usability for older adults: A literature based framework (MOLD-US). *International Journal of Medical Informatics*, 114, 66-75. <https://doi.org/10.1016/j.ijmedinf.2018.03.012>

- Wildenbos, G. A., Peute, L. W., & Jaspers, M. W. M. (2015). A framework for evaluating mHealth tools for Older Patients on Usability. *Studies in Health Technology and Informatics*, 210, 783-787.
- Williams, M. V., Baker, D. W., Parker, R. M., & Nurss, J. R. (1998). Relationship of Functional Health Literacy to Patients' Knowledge of Their Chronic Disease: A Study of Patients With Hypertension and Diabetes. *Archives of Internal Medicine*, 158(2), 166-172. <https://doi.org/10.1001/archinte.158.2.166>
- Wogalter, M. S., & Barlow, T. (1990). *Injury Severity and Likelihood in Warnings*. 580-583. <https://doi.org/10.1177/154193129003400801>
- Wogalter, M. S., Desaulniers, D. R., & John W. Brelsford, J. (2016). Consumer Products: How are the Hazards Perceived?: *Proceedings of the Human Factors Society Annual Meeting*. <https://doi.org/10.1177/154193128703100530>
- Wong, J., Eakin, J., Migram, P., Cafazzo, J. A., Halifax, N. V. D., & Chan, C. T. (2009). Patients' experiences with learning a complex medical device for the self-administration of nocturnal home hemodialysis. *Nephrology Nursing Journal: Journal of the American Nephrology Nurses' Association*, 36(1), 27-32.
- Woot, P. (1970). *Stratégie et management*. Dunod.
- Wozney, L. M., Baxter, P., Fast, H., Cleghorn, L., Hundert, A. S., & Newton, A. S. (2016). Sociotechnical Human Factors Involved in Remote Online Usability Testing of Two eHealth Interventions. *JMIR Human Factors*, 3(1), e6. <https://doi.org/10.2196/humanfactors.4602>
- Wright, P., Creighton, P., & Threlfall, S. M. (1982). Some factors determining when instructions will be read. *Ergonomics*, 25(3), 225-237. <https://doi.org/10.1080/00140138208924943>
- Wright, P., Soroka, A. J., Belt, S., Pham, D. T., Dimov, S., DeRoure, D. C., & Petrie, H. (2008). Modality preference and performance when seniors consult online information. *Gerontechnology*, 7(3). <https://doi.org/10.4017/gt.2008.07.03.004.00>
- Wyatt, J., & Spiegelhalter, D. (1991). Field trials of medical decision-aids: potential problems and solutions. *Proceedings. Symposium on Computer Applications in Medical Care*, 3-7.
- Yang, F., & Mahmud, A. A. (2018). Issues in evaluating the user performance of new home medical devices: a field test of two devices produced by a leading provider. *Journal of Design Research*, 16(3-4), 196-213. <https://doi.org/10.1504/JDR.2018.099537>
- Yen, P.-Y. (2010). *Health information technology usability evaluation: methods, models, and measures*. <https://search.proquest.com/openview/0df85be25e4479c55d4b72b070dc034c/1?pq-origsite=gscholar&cbl=18750&diss=y>

- Yen, P.-Y., & Bakken, S. (2012). Review of health information technology usability study methodologies. *Journal of the American Medical Informatics Association*, 19(3), 413-422. <https://doi.org/10.1136/amiajnl-2010-000020>
- Yen, P.-Y., Lara, B., Lopetegui, M., Bharat, A., Ardoin, S., Johnson, B., Mathur, P., Embi, P. J., & Curtis, J. R. (2016). Usability and Workflow Evaluation of « RhEumAtic Disease activitY » (READY). A Mobile Application for Rheumatology Patients and Providers. *Applied Clinical Informatics*, 7(4), 1007-1024. <https://doi.org/10.4338/ACI-2016-03-RA-0036>
- Yu, D. X., Parmanto, B., Dicianno, B. E., Watzlaf, V. J., & Seelman, K. D. (2017). Accessibility needs and challenges of a mHealth system for patients with dexterity impairments. *Disability and Rehabilitation: Assistive Technology*, 12(1), 56-64. <https://doi.org/10.3109/17483107.2015.1063171>
- Zapata, B. C., Fernández-Alemán, J. L., Idri, A., & Toval, A. (2015). Empirical Studies on Usability of mHealth Apps: A Systematic Literature Review. *Journal of Medical Systems*, 39(2), 1. <https://doi.org/10.1007/s10916-014-0182-2>
- Zenklusen, S., Bischoff, T., Bodenmann, P., Panese, F., Bodenmann, P., Madrid, C., & Wolff, H. (2012). Compétences en santé déficientes : obstacle à une prise en charge optimale. *Revue Médicale Suisse*, 8, 1016-1021. <https://www.revmed.ch/RMS/2012/RMS-340/Competences-en-sante-deficientes-obstacle-a-une-prise-en-charge-optimale>
- Zhang, D., & Adipat, B. (2005). Challenges, Methodologies, and Issues in the Usability Testing of Mobile Applications. *International Journal of Human-Computer Interaction*, 18(3), 293-308. https://doi.org/10.1207/s15327590ijhc1803_3

ANNEXES

Annexe 1 : Revue des modèles d'utilisabilité de 1991 à 2018

Auteurs	Les métriques des modèles d'utilisabilité								
McCall (1971)	Opérabilité	Training	communicativeness						
Shackel (1991, p.24-25)	Efficacité	Facilité d'apprentissage	Flexibilité	Attitude					
Schneiderman (1992)	Efficacité	Efficiency							
Nielsen (1993, p.26)	Efficiency	Facilité d'apprentissage	Apprenabilité	Erreurs	satisfaction				
Dix et al. (1993)	Efficacité	Efficiency	Satisfaction	Apprenabilité					
Vanderdonckt (1994)	Compatibilité	Cohérence	Charge de travail	Gestion d'erreur	Contrôle de dialogue	de	représentativité	Guidage	Adaptabilité
Preece et al. (1994)	Efficacité	Efficiency	Apprenabilité	Attitude					
Scapin et Bastien (1997)	Compatibilité	homogénéité/Cohérence	Charge de travail	Gestion d'erreur	Contrôle explicite		Signifiante des codes et dénominations	Guidage	Adaptabilité
(ISO 9241-11, 1998 ; 2018)	Efficacité	Efficiency	Satisfaction	Liberté de risques	Couverture de contexte				
ISO 9126-1 (1991)	Compréhensibilité	Apprenabilité	Productivité	Attractivité	Conformité à l'utilisabilité				
Norman (1998)	Familiarité	Standardisation	Charge cognitive	Prévention des erreurs	des	visibilité	Guidage		
Dee et Allen (2001)	Efficacité	Efficiency	Satisfaction	Production	Confiance		Justesse		
Rengger (1991)	Atteinte des objectifs (exactitude et efficacité)	Rythme de travail (productivité et efficacité)	Acquisition des connaissances (apprenabilité)	des	Opérabilité (taux d'erreur et utilisation des fonctions)				

Auteurs	Les métriques des modèles d'utilisabilité						
Bevan et al. (1991)	Type de produit	Type d'utilisateur	La facilité d'utilisation	Acceptabilité			
(Mich et al., 2003) : 2QCV3Q quality model	Accessibilité	Navigabilité	compréhensibilité				
(Dix, Finlay, Abowd, & Russell, 2003)	Flexibilité	Apprenabilité	Robustesse				
Fisk, Rogers, Charners, Czaja, & Sharit (2004, pp. 30-31)	Efficience	Apprenabilité	Mémorisation	Erreurs	Satisfaction		
Wang et Huang (2009)	Efficience	Apprenabilité	Taux d'erreurs	Satisfaction			
PACMAD (Harrison, Flood, & Duce, 2013, p.4)	Efficacité	Efficience	Satisfaction	Apprenabilité	Mémorisation	Erreurs	Charge cognitive
(Abran, Khelifi, Suryan, & Seffah, 2003, p.14)	Efficacité	Efficience	Satisfaction	Apprenabilité	Sécurité		
Bass et al (Bass & John, 2003)	Modifiabilité	Évolutivité	Réutilisabilité	Performance	Sécurité		
(Shneiderman & Plaisant, 2010)	Temps d'apprentissage	Rapidité d'exécution,	taux d'erreurs	Rétention dans le temps	Satisfaction		
Dubey et al (Alonso-Ríos, Vázquez-García, Mosqueira-Rey, & Moret-Bonillo, 2009, p.5)	Efficacité	Efficience	Satisfaction	Apprenabilité			
Ham et al. 2009	Efficacité	Efficience	Satisfaction	Apprentissage	Customization		

health-ITUEM (Yen, 2010)	Efficacité	Efficiency	Satisfaction (utilité perçue/ d'utilisation perçue)	utilité facilité		
(Dubey & Gulati, 2012, p. 431)	Efficacité	Efficiency	Satisfaction	Compréhensibilité	Sécurité	

Annexe 2 : Critères d'inclusion et d'exclusion de la population pouvant utiliser le dispositif *Smart Angel*

Les critères présentés ci-dessous sont tirées du CPP pour l'étude clinique intitulée « Evaluation de l'impact du dispositif *Smart Angel* sur le suivi du patient à domicile après une chirurgie ambulatoire majeure ou intermédiaire : étude contrôlée randomisée en ouvert.

Description de la population à étudier

La population de l'étude est constituée de patients majeurs opérés de chirurgie réglée ambulatoire majeure ou intermédiaire :

- en orthopédie : épaule (arthroscopie, butée, prothèse), hanche (prothèse, arthrolyse), genou (ostéotomie, ligamentoplastie), cheville (prothèse, ligamentoplastie) pied (hallux valgus) ; ablation matériel majeur ou multisite
- en digestif (cholécystectomie, cure de hernie, gastrectomie partielle, colectomie) ;
- en gynécologie (hystérectomie, ovariectomie, mastectomie, quadrantectomie) ;
- en urologie (résection totale ou partielle de prostate).
- ORL (thyroïdectomie, amygdalectomie)
- Neurochirurgie (hernie)
- Vasculaire (stripping varice, création fistule)

Les patients seront présélectionnés dans chaque centre investigateur au cours de la visite d'annonce de l'acte chirurgical et inclus lors de la consultation d'anesthésie préopératoire qui est réalisée à distance de l'intervention chirurgicale.

Critères d'inclusion pour le choix des sujets

Critères d'inclusion généraux :

- Le patient doit avoir donné son consentement libre et éclairé et signé le consentement
- Le patient doit être affilié ou bénéficiaire d'un régime d'assurance maladie
- Le patient est âgé d'au moins 18 ans (\geq) et de moins de 80 ans ($<$)

Critères d'inclusion concernant la population cible :

- Le patient est opéré en chirurgie ambulatoire réglée :
 - en orthopédie : épaule (arthroscopie, butée, prothèse), hanche (prothèse, arthrolyse), genou (ostéotomie, ligamentoplastie), cheville (prothèse, ligamentoplastie) pied (hallux valgus complexe) ;

- en digestif (cholécystectomie, cure de hernie avec plaque) ;
- en gynécologie (hystérectomie, ovariectomie, mastectomie, quadrantectomie avec curage ou picking ganglionnaire) ;
- en urologie (résection totale ou partielle de prostate).
- Le patient a les capacités cognitive et intellectuelle nécessaires pour pouvoir interagir avec des objets connectés.
- Le patient doit réussir le test d'apprentissage qui sera réalisé à la consultation d'anesthésie, à savoir : ouvrir et connecter la tablette, activer la mesure de la pression artérielle.
- Le patient doit avoir accès à la 4G à son domicile.

Critères de non inclusion pour le choix des sujets

Critères de non-inclusion généraux :

- La patiente est enceinte.
- Le sujet est en période d'exclusion déterminée par une étude précédente
- Le sujet est sous sauvegarde de justice, sous tutelle ou sous curatelle
- Le sujet refuse de signer le consentement
- Il s'avère impossible de donner au sujet des informations éclairées

Critères de non-inclusion concernant les maladies ou conditions associé(e)s interférent(e)s :

- Le patient nécessite une chirurgie en urgence.
- Patient de classe physiologique ASA 5
- Patient non éligible à la chirurgie ambulatoire pour raison médicale (pathologie médicale décompensée) ou sociale selon les critères définis par la SFAR (société française anesthésie réanimation). Personne seule au domicile, éloignement géographique d'un centre hospitalier > 30 km (ou > 45 min en voiture).

5.1.Critères d'exclusion

- Patient ne respectant pas les critères de sortabilité de la chirurgie ambulatoire : patient en échec de sortie.

Annexe 3 : Mapping des produits de la m-santé et patterns des usages



Figure 67. Mapping des produits de la m-santé et patterns des usages suivant deux axes (curatif-préventif *versus* applications-objets connectés) faisant apparaître 5 patterns (l'urgence, le pathologique, l'autonomie, le sport, la beauté/bien-être).

Annexe 4 : Tableau d'analyse systématique des méthodes et outils utilisés dans l'évaluation de l'utilisabilité de dispositif en santé

Articles	Revue	Matériel(s) et intérêt : nom du produit	Tâche(s)	Environnement(s)	Métrique(s)	Quantitative/Qualitatif	Subjectif/Objectif	Outils/méthodes	Population	Analyse(s) de données	Résultats
1. Brown et al., 2013	Journal of Biomedical Informatics	Application: MyFitnessPal, SparkPeople, NIH obesity, et Asthma Check	Utilisation des applications libres	Laboratoire (1 ^{er} session) <i>In situ</i> (2 ^{ème} session)	Uniquement satisfaction de ITUEM-health	qualitatif	Subjectif	Focus group pour explorer l'utilisation Focus group après utilisation en milieu écologique	N = 60 adolescents	Descriptive / comparaison	Mémorisation et évaluation de la technologie sont des métriques peu pertinentes
2. Burnay et al., 2012)	Telemedecine and e-Health	App pour rhinite allergique : m.Carat	Tests utilisateurs : Manipulation par scénario (10 tâches)	Laboratoire	efficacité	Quantitative	Subjectif	Report d'erreur	N = 8 (caractéristiques non spécifié)	Descriptive	Aucun problème d'utilisabilité grave trouvé
3. Sheehan et al., 2012	Applied Clinical Informatics	4 téléphones mobiles différents	Tests utilisateurs : Manipulation par scénario (5 tâches)	Laboratoire	ISO/IEC 25010.3/ efficacité, satisfaction /apprentissage	Quantitative/Qualitative	Objectif/Subjectif	Think aloud / questionnaire (Post-Study System Usability Questionnaire (PSSUQ) / observation (enregistrement vidéo)	N = 20 adolescent urbain (avec détails des caractéristiques démographique et expérience en informatique)	ANOVA	Tous les utilisateurs ont réalisé les tâches sans problèmes, seuls l'efficacité était significativement différence en fonction de le l'appareil et de l'app
4. Sparkes et al., 2012	Telemedecine and e-health	Test cardiaque à domicile pendant 1 semaine pour remplacer le Holter : Vitaphone 3100BT sur Blackberry	Tests utilisateurs : Manipulation par scénario (20 tâches)	Terrain (le participant est interrogé à son domicile et filmé par le chercheur qui l'aide à poursuivre la manipulation en cas de problème)	Métrique de Nielsen : efficacité, efficacité, satisfaction, facilité d'apprentissage, mémorisation	Quantitative/Qualitative	Objectif/Subjectif	Enregistrement vidéo / think aloud / entretien	N = 5 participants ciblés par une arythmie cardiaque (sexe, âge, expérience en IT)	Descriptive	le sexe, l'âge et la familiarité avec la technologie semblaient influencer la capacité des participants à installer ces dispositifs avec succès
5. Brown-Johnson et al., 2015	Patient Education and Counseling	Jeu RV pour le cancer de la langue : TLC mhealth	Tests utilisateurs : Manipulation pendant 30min	Laboratoire	Satisfaction	Qualitative	Subjectif	Entretiens semi-structurés/ think aloud	N = 8 professionnels de santé (20-50ans) lié au cancer de	Descriptive	Le TLC mhealth relève d'une forte charge émotionnelle et peut promouvoir

6. Georgsson & Staggers, 2016a et Georgsson & Staggers, 2016b	<i>Journal of the American Medical Informatics Association</i> Journal of Biomedical Informatics	Messagerie pour patient atteint de diabète : Care4Life	Tests utilisateurs : Manipulation par scénario (8 tâches)	Laboratoire	ISO 9241-11 : efficacité, efficacité, satisfaction	Quantitative Qualitative	Objectif Subjectif	Questionnaire (SUS), observation direct, think aloud, entretien semi-structuré	la langue par expérience N = 10 (caractéristiques utilisateurs :	Descriptive	une autogestion intéressante Dans l'ensemble : tâche difficile à réaliser. les hommes réussissaient mieux à accomplir les tâches et que les jeunes et les expérimentés en IT sont plus efficient + dépend du diagnostic de diabète
7. Mirkovic et al., 2014	JMIR mHealth and uHealth	App pour gérer les problèmes de santé liés au cancer : Connect Mobile	Tests utilisateurs Manipulation par scénario (8 tâches) / heuristiques (Nielsen)	Laboratoire	Efficacité, efficacité, satisfaction	Quantitative Qualitative	Objectif Subjectif	Questions semi-structuré après chaque tâche. Questionnaire (SUS) Entretien / Analyse vidéo / Observation direct	N = 7 patients atteints de cancer (caractéristique démographique : âge, genre, education, experience IT)	Descriptive	27 problèmes d'utilisabilité détectés.
8. Ayubi et al., 2014	JMIR mHealth and uHealth	App de podomètre	Test utilisateurs puis déploiement sur 4 semaines	Laboratoire (1 ^{er} partie), <i>in situ</i> (2 ^{ème} partie)	Métrique Nielsen : apprentissage, mémorisation, Efficacité, efficacité, satisfaction (motivation)	Quantitative Qualitative	Objectif Subjectif	Think aloud, questionnaire (PSSUQ ; TAM), entretien semi-directif, traceur d'activité	N = 13 (âge, genre, habitude de marche, IMC, Experience IT)	Descriptive	Les résultats relatifs à l'utilisabilité suggèrent que le système est utilisable et que la satisfaction des utilisateurs était élevée.
9. Isaković et al., 2016	Journal of Diabetes Reserach	Application de surveillance du diabète, DeStress Assistant (DeSA)	Tests utilisateurs Manipulation par scénario (5 tâches)	Laboratoire	Compréhensibilité, présentation (image et text), utilisabilité, caractéristique général	Qualitatif Quantitatif	Subjectif	Think aloud, entretien, questionnaire (MARS ; SUS)	N = 10 personnes âgées diabétiques	Descriptive	avec un petit nombre de changements, il est possible d'augmenter l'utilisabilité de l'application dans un certain nombre de catégories
10. Schnall, Bakken, et al., 2016; Schnall, Rojas, et al., 2016	Journal of Biomedical Informatics / Studies in Health Technology and Informatics	Application pour la prévention, le traitement et les soins du VIH : MSM app	Évaluation heuristique de Nielsen par expert puis par les utilisateurs	Laboratoire	Satisfaction	Qualitatif Quantitatif	Subjectif	Questionnaire (PSSUQ), enregistrement des sessions d'utilisation	N = 10 (caractéristique : expérience en IT)	Descriptive	Les scores PSSUQ ont eu tendance à diminuer avec le temps ce qui indique que les utilisateurs finaux ont perçu

11. Househ et al., 2015	Acta informatica Medica	principales applications mHealth pour le diabète (4 application)	Aucune information	Laboratoire	ITUEM-health = satisfaction (Besoins d'informations, Flexibilité/ personnalisation, Facilité d'apprentissage, rapidité)	Qualitative	Subjectif	Aucune information : analyse du discours	N = 405	Descriptive	l'application comme étant de plus en plus utile Les consommateurs sont plus susceptibles d'utiliser des applications de mSanté liées au diabète qui sont performantes et qui répondent à leurs besoins d'information.
12. Jibb et al., 2017	Journal of Pediatric Oncology Nursing	Gestion de la douleur pour adolescent atteint de cancer : Pain Squad+	Test utilisateur sur 3 cycles itératifs : consignes précises	Laboratoire	Efficience, , Difficultés technique rencontrée, facilité d'utilisation et acceptabilité	Qualitatif Quantitatif	Objectif Subjectif	Think aloud, chronomètre, entretien semi-structuré	N = 16 adolescents atteints de cancer ayant déclaré une douleur forte intensité la semaine précédente (âge, sexe, diagnostic, modalité de traitement)	Descriptive	Dans l'ensemble, l'application était acceptable. Les problèmes étaient liés au mauvais fonctionnement du logiciel, aux défauts de conception de l'interface et à la confusion du texte
13. Stoll et al., 2017	Cognitive and Behavioral Practice	App prévention de l'anxiété des jeunes : REACH	Tests utilisateurs Manipulation par scénario (7 tâches)	Laboratoire	Facilité d'utilisation, facilité d'apprentissage, qualité de l'information, satisfaction et stigmatisation, erreurs	Quantitative	Subjectif	Questionnaires (USE : usefulness, satisfaction and ease of Use + RPS : Reaction Program Scale)	N = 177 (132 enfants et 45 fournisseurs de services)	ANOVA, test de Student	Une application bien apprécié sauf lors d'erreurs survenues / la compréhension du système mène à une meilleure satisfaction
14. Alnosayan et al., 2017	JMIR Human Factors	Dispositif de prévention de l'insuffisance cardiaque (balance, tensiomètre, glucomètre, téléphone avec application MyHeart)	Formation puis utilisation libre à leur domicile pendant 6 mois (bilan tous les trois mois)	In situ retrospective	Satisfaction	Quantitative Qualitative	Subjectif	Questionnaire : SUS (system usability scale) Focus group (tous les 3 mois)	N = 8 patients (sur 6 mois)	Descriptive	Recommandations : la nécessité de communiquer par d'autres moyens, de personnaliser les questions et les messages sur les symptômes, d'intégrer d'autres technologies de suivi de la santé et de fournir des méthodes supplémentaires

15. Gordon et al., 2017	Patient Education and Counseling	Application pour arrêter de fumer : RxCoach	Focus group Tests utilisateurs (2 itérations)	Laboratoire	Satisfaction	Quantitative Qualitative	Subjectif	Focus group avec questionnaire puis observation directe de manipulation Test U : Think aloud Questionnaire SUS	N = 4 (*2) groupes de discussion N =10 les deux étaient fumeurs utilisant une autre app	Descriptive	pour analyser et visualiser leurs données nous avons observé que les participants ne lisaient pas tout le contenu lors de l'installation et dans les messages de notification.
16. Mayberry et al., 2016	Journal of Diabetes Research	Application de gestion et soutien des « nuisances » de la famille sur le diabète : FAMs	3 séries de tests itératifs	<i>In situ</i>	Satisfaction, facilité de compréhension	Qualitative	Subjectif	Entretien téléphonique	N = 19 diabétique (âge, sexe, auto-gestion du diabète, littératie en santé, participation de la famille, expérience IT)	Descriptive	Le FAMS est utilisable et réalisable et semble aider les patients à gérer le soutien aux soins auto-administrés
17. Fu et al., 2016	Mhealth	Application de Gestion et des symptômes post-cancer du sein : TOLF	Évaluation heuristique et tests utilisateur Manipulation par scénario (X tâches)	Laboratoire	perception, qualité du système, intention comportementale, facilité d'utilisation et utilité perçues	Qualitative	Subjectif	Think aloud, questionnaires : Perceived Ease of Use and Usefulness Questionnaire et SUS	N = 30 femmes en rémission d'un cancer du sein (âge, niveau d'éducation, littératie, et expérience IT)	Descriptive	
18. Grindrod et al., 2014	JMIR mHealth and uHealth	Ipad + Applications d'observance médicamenteuse : MyMedRec, DrugHub, Pillboxie et PocketPharmacist	tests utilisateur Manipulation par scénario (X tâches) * 4 applications	Laboratoire	Satisfaction (facilité d'apprentissage, expérience utilisateur, facilité d'utilisation, risque erreur, qualité perçue, préférences)	Quantitative Qualitative	Subjectif	Questionnaire SUS, entretiens focus group	N = 35 participants sous traitement chroniques (âge>50ans, expérience IT, état de santé)	Descriptive	Lors de l'évaluation des applications mobiles de gestion des médicaments, les participants ont eu du mal à penser à la nécessité de ces applications dans leur propre vie
19. Rajput et al., 2012	<i>Journal of the American Medical Informatics Association</i>	Application de conseil, dépistage, recueil d'information sanitaire pour le VIH et tuberculose.	Test utilisateur et enquête	Terrain	Efficacité et satisfaction (Test U et enquête)	Qualitative	Subjectif	Observation 24h au domicile patient, think aloud, entretien semi-structurés et focus group, questionnaire	Test U : N = ? (aucune information) Enquête : N = 18 850 ménages	Descriptive	

20. Gerdes et al., 2014	Computer Science		tele-health and tele-care services : United4Health	tests utilisateur Manipulation par scénario	Laboratoire (simulation dans un service de soin de santé réel)	Efficacité	Quantitative	Objectif	(uniquement pour l'enquête) Enregistrement audio et vidéo, observation direct	N = ? patients et infirmiers	Pas de résultat présenté	
21. Kriek et al., 2012	Journal of medical Internet research	of	système de soutien pour personne atteinte de schizophrénie : ROM	Évaluation heuristique et test utilisateur manipulation par scénario (même tâche que pour l'évaluation heuristique)	Laboratoire (2/3 des participants) Terrain (tâche précise à réaliser) (1/3 au domicile patient)	Critère de Nielsen : apprenabilité, efficacité, mémorisation, erreurs, satisfaction	Quantitative	Subjectif	Retranscription vidéo de l'écran à distance en direct, observation direct, questionnaire	N = 4 experts pour l'évaluation heuristique et N = 15 patients schizophrénique (expérience IT)	Descriptive	il a été démontré que le partage de la prise de décision améliore l'alliance thérapeutique et mène à de meilleurs soins. Il est possible de faire participer des patients schizophréniques à la conception malgré leurs troubles cognitifs.
22. Voncken-Brewster et al., 2013	Journal of medical Internet research : Research Protocols	of	Application pour soutenir l'autogestion des patients atteints de maladies pulmonaires obstructives chroniques	3 séries de tests itératifs : manipulation par scénario	Laboratoire	Efficacité et satisfaction	Quantitative	Objectif Subjectif	Think aloud, entretien semi-directif, enregistrement vidéo	N = 8 (âge 51-70 ans) patients atteints de maladie pulmonaire obstructive chronique (expérience IT, caractéristiques démographiques)	Descriptive	La plupart des problèmes ont pu être résolus en trois itérations d'amélioration
23. Carmody et al., 2019	Journal of Pediatric Psychology	of	4 Appli pour l'observance médicamenteuse : Medisafe, MyTherapy, MemoHealth, Care4Today	Test utilisateur (4 tâches pour chaque app : entrer un de leurs médicaments, à définir un rappel, à activer la fonction de partage des soins et à examiner les autres composants de l'application)	Laboratoire	l'engagement, de la fonctionnalité, de l'esthétique et de l'information	Quantitatif	Subjectif	Questionnaire (uMARS Mobile Application Rating Scale)	N = 20 jeunes adultes avec maladie inflammatoire de l'intestin	ANOVA & Descriptif	

24. Parmanto et al., 2013	JMIR mHealth and uHealth	Système de soutien, d'autosoins dans la gestion des conditions complexes et chroniques : iMHere	3 série de tests itératifs : phase 1 et 3 manipulation pendant plusieurs semaines, phase 2 : manipulation par scénario	Phase 1 et 3 : <i>in situ</i> Phase 2 : laboratoire	Non précisé	Quantitative Qualitative	Objectif	Non précisé	N = 14 utilisateurs potentiels de iMHere (8 phase 1, 3 phase 2, et 3 phase 3. (âge : 18-40ans atteints de SB et hydrocéphalie)	Descriptive	Les modèles d'utilisation ont montré une augmentation de l'utilisation au cours du premier mois, suivie d'un plateau
25. Narváez et al., 2016	Studies health technologie informatics	Application de prévention du syndrome du Burnout	3 série de tests itératifs sur 3 prototype	Laboratoire	ISO 9126 Fonctions évidentes, fonction compréhensible et Satisfaction	Quantitative Qualitative	Objectif Subjectif	Questionnaire (SUS), think aloud	N = 6 travailleurs *3 prototype (âge moy = 41ans)	Descriptive	(score SUS : 89.36)
26. Metelmann et al., 2018	JMIR mHealth and uHealth	Applications pour la réanimation cardiopulmonaire	Évaluation de 3 applications sur leur propre téléphone portable	Laboratoire	Satisfaction	Quantitatif	Subjectif	Questionnaire (SUS),	N = 9 urgentistes et N = 10 participants (âge moy = 23ans) avec forte expérience IT)	Test paramétrique et non paramétrique Descriptive	Sur 34 app, 3 correspondant aux critères médicaux, Seulement 1 des 3 applications testées avait une mesure du SUS correct. La mise en place d'un contrôle de qualité systématique pour les applications liées à la santé devrait être appliquée pour garantir une exactitude médicale et une utilisabilité suffisante
27. Yu et al., 2017	Disability and rehabilitation : assistive Technology	système pour les patients ayant des troubles de la dextérité : iMHere	Utilisation du système pendant 1 semaine puis tests utilisateurs Manipulation par scénario (tâche)	Laboratoire	Efficacité : Temps de manipulation chaque tâche, Efficacité : Le nombre d'erreur Satisfaction	Quantitative Qualitative	Objectif Subjectif	Observation par enregistrement vidéo, questionnaire	N = 9 patients avec divers niveaux de dextérité	Corrélation descriptive	Une corrélation négative statistiquement significative a été identifiée entre les niveaux de dextérité des participants et leurs taux d'erreur (r = -0,434, n =

												36, p = 0,004). a conception personnalisée peut être la clé pour relever ces défis en matière d'amélioration de l'accessibilité.
28. Pulantara et al., 2018	JMIR Human Factors	Plateforme d'intervention adaptative pour les insomniaque : iREST	Formation en laboratoire puis essai pendant une semaine puis retour au laboratoire	<i>in situ retrospectif</i>	Efficacité, fiabilité, satisfaction	Quantitative	Objectif Subjectif	Questionnaire (SUS) et Usability Questionnaire (TUQ), question ouverte	N = 19 vétérans insomniaque (tactiques de sommeil militaires)	Test t Descriptif	L'utilisation de vêtements pour le suivi du sommeil est prometteuse, mais la précision de cette technologie doit encore être améliorée	
29. Cho, Powell, et al., 2018	JMIR mHealth and uHealth	programme de prévention du VIH : MyPEEPS Mobile.	Test d'heuristique (expert) et test utilisateur avec Manipulation par 2 scénario (2 groupe)	laboratoire	Satisfaction	Quantitative	Subjectif	Enregistrement vidéo de l'écran, think aloud retrospectif, questionnaire (PSSUQ)	N = 5 experts (tests heuristiques) N = 20 jeunes homme (moy 17.4ans ; niveau d'étude)	Descriptif	Les commentaires faits par les experts et les utilisateurs finaux serviront à raffiner MyPEEPS Mobile avant la tenue d'une étude pilote évaluant l'acceptabilité de l'application	
30. Gunter et al., 2016	JMIR mHealth and uHealth	Application de surveillance (prise de photo de plaie + questionnaires) post-chirurgical : WoundCheck	Test utilisateur avec Manipulation par scénario (7 tâche)	Laboratoire (à l'hôpital)	(à ISO 9241-11 Efficacité, efficacité, satisfaction)	Quantitative	Objectif Subjectif	Focus group (2), Questionnaire (SUS), observation direct, codage vidéo, enregistrement vidéo de l'écran	N = 9 patients sortant de chirurgie vasculaire et générale	Descriptif	Les patients chirurgicaux peuvent apprendre à utiliser une application pour smartphone pour la surveillance postopératoire des plaies avec une grande satisfaction des utilisateurs	
31. Neubert et al., 2018	JMIR mHealth and uHealth	Application de dessin de symptômes pour sujets âgés ou handicapés	Test-retest Test utilisateur avec manipulation par scénario (1 tâche)	Terrain (vaguement précisé) (avant le rdv avec le médecin)	Non précisé Satisfaction (patient)	Quantitative (patient)	Subjectif (patient)	Patient : Questionnaires fait maison, Médecin : Observation, question ouverte, Questionnaire SUS, Attrakdiff,	N = 52 patients (moy = 56.2ans, sexe, IMC, niveau d'étude, niveau de douleur, expérience IT) N = 6 médecins	Test t descriptif	L'évaluation de l'utilisabilité par les patients est un facteur important qui ne doit pas être négligé lors de la conception d'applications mobiles ou	

32. Batsis et al., 2018	Gerontechnology	Dispositif de port d'amulette chez les personnes âgées en milieu rural	Groupe de discussion avec manipulation libre et jugement de l'esthétique	Laboratoire	No précisé Satisfaction	Qualitative	Subjectif	et score ISONORM Entretiens semi-directifs, focus group, questionnaire	N = 29 adultes âgé + N = 4 leaders communautaires et N = 7 cliniciens ruraux	Descriptive	d'applications eHealth. Les préoccupations de la mSanté comprennent l'exactitude du dispositif, son intrusion dans la vie quotidienne et la protection de la vie privée. Les évaluations formatives sont d'une importance critique avant de déployer des interventions à grande échelle.
33. Hussain et al., 2015	Jurnal Teknologi	portail Web de sensibilisation à la santé : e-Ebola (sur Samsung et Lenovo)	Test utilisateur avec manipulation par scénario (4 tâche)	Laboratoire	ISO 9241-11 Efficience (temps de travaux) Efficacité (taux d'achèvement des tâches et erreurs) Satisfaction	Quantitative Qualitative	Objectif Subjectif	Enregistrement vidéo, observation, questionnaire (item conçu par les auteurs) + SUS, debriefing	N = 9 (sexe, âge, état matrimonial, expérience IT)	Descriptif	il existe un certain nombre de normes/modèles internationaux sur l'utilisabilité, mais ils sont rarement utilisés pour l'évaluation de la convivialité pratique.
34. Alnasser et al., 2018	JMIR Human Factors	application de perte de poids : Twazon Arabic	4 mois d'utilisation puis entretien	<i>In situ</i> <i>retrospectif</i>	satisfaction	Qualitative Quantitative	Subjectif	Entretiens, questionnaire SUS	N = 26	Descriptif	Score au SUS = 69.3 aucune relation n'a été trouvée entre les attentes de l'utilisateur et les pratiques de perte de poids fondées sur des données probantes.
35. Alanzi et al., 2016	JMIR Research Protocols	système pour le changement de comportement social et la gestion adapté aux patients saoudiens atteints de diabète, : Saudi Arabia Networking for	Test utilisateur avec manipulation par scénario (8 tâche)	Non précisé (Laboratoire)	Non précisé Satisfaction	Quantitative	Subjectif	questionnaire (QUIS)	N = 33 diabétiques (âge 10- 30 ; niveau d'éducation, statut marital, diagnostic diabète)	Descriptif	l'acceptabilité du système était élevée chez les patients saoudiens atteints de diabète, et des travaux sont en cours dans ce domaine de recherche pour

36. Price et al., 2016	JMIR Mental Health	Aiding Diabetes (SANAD) Système de surveillance mobile pour évaluer les symptômes après une blessure traumatique	5 Tests utilisateur avec manipulation par scénario (tâche)	Laboratoire (mise en situation par un scénario d'un accident)	TAM (satisfaction, efficacité)	Qualitative Quantitative	Subjectif Objectif	Questionnaire (PUEU) et Entretiens (enregistrement audio), Enregistrement vidéo	N = 21 patients ayant un vécu de trauma (moy = 19 ans, expérience IT)	Descriptif	mener une étude pilote clinique Une rétroaction personnalisée à l'utilisateur a été jugée essentielle
37. Skorin-Kapov et al., 2014	International Journal of Mobile Human Computer Interaction	mesure des signes vitaux via des capteurs médicaux (tensio mètre balance, ECG, spiromètre, glucomètre, oxymètre): Ericsson Mobile Health	3 itérations : Test utilisateur avec manipulation par scénario (8 tâches)	Laboratoire	QoE métriques : Efficacité, efficacité, apprenabilité, satisfaction, utilité, intégrité, fiabilité, privé, réponse, confort et esthétique	Qualitative Quantitative	Subjectif Objectif	Journal de bord, think aloud, hiérarchisation de la tâche, questionnaire QoE, AttrakDiff	N = 26 (âge, motivation, expérience IT)	Correlation, descriptif	Très complet, mais très long une forte corrélation entre l'efficacité perçue de l'interface mobile, la facilité perçue à effectuer une tâche de mesure de la tension artérielle et la motivation de l'utilisateur pour l'utilisation du service.
38. Constantinescu et al., 2019	Health informatics journal	Thérapie de déglutition à domicile pour les survivants du cancer du cou : Mobili-T®	Démonstration du Test utilisateur avec manipulation par scénario (5 tâches)	Laboratoire	ISO 9241-11 Efficacité, efficacité et satisfaction	Qualitative Quantitative	Subjectif Objectif	Observation (enregistrement vidéo), think aloud, entretien, questionnaire (ASQ)	N = 5 (sexe, âge, littératie en santé)	Descriptif	Parmi les méthodes d'utilisation utilisées, l'observation vidéo (capture d'écran synchronisée avec des gestes vidéo) a révélé le plus d'informations
39. Ben-Zeev et al., 2013	Psychiatric Rehabilitation Journal	Auto-gestion de la schizophrénie : FOCUS	2 cycles itératifs : Test utilisateur avec manipulation par scénario (tâche ?)	Laboratoire	Satisfaction, utilité perçue	Qualitative Quantitative	Subjectif	Think aloud, questionnaire (non précisé)	N = 12 (moy = 45 ans ; niveau d'étude ; expérience IT) Schizophrène : non précisé	Descriptif	Les auteurs déclarent avoir produit une appli qui sera utilisée avec succès, et qui est prête à être déployée et évaluée de façon systémique dans des conditions réelles.
40. Beauchemin et al., 2019	International Journal of	Autogestion du VIH : WiseApp	3 cycles : 2 tests utilisateurs	Laboratoire	Health-ITUES Efficacité, satisfaction	Qualitative Quantitative	Subjectif Objectif	Think aloud, enregistrement de l'écran,	N = 20 (moy 514.3 ans) utilisateur	Descriptif	Les résultats de la démarche cognitive avec les

	Medical Informatics		avec manipulation (26 tâches) et 1 tests heuristiques avec experts					observation direct, cognitive walkthrough, questionnaire (health-ITUES)	N = 4 experts		experts et les utilisateurs finaux ont permis d'apporter des améliorations itératives à WiseApp
41. Daugherty et al., 2012	Journal of Medical Internet Research	le relevé alimentaire du téléphone portable	Tests utilisateurs sans expérimentateur sur deux repas + instruction les tâches à réaliser	Laboratoire	Satisfaction	Qualitative Quantitative	Subjectif Objectif	Questionnaires ouverts, séance rétroaction, analyse des photos des repas	N = 135 volontaire (78 ado, 57 adultes, âge)	Test paramétrique (MnNemar)	Les adultes ont été moins efficaces lorsqu'ils ont utilisé le relevé alimentaire par téléphone mobile. Une formation sera probablement nécessaire
42. Barbara et al., 2016	JMIR Human Factors	Plateforme pour information sur le vieillissement : McMaster	Tests utilisateurs avec manipulation	Laboratoire (ou domicile patient par skype)	Non précisé Efficacité	Qualitative Quantitative	Subjectif Objectif	Entretien, observation direct (partage d'écran sur skype), think aloud,	N = 37 (sexe, moy = 69 ans, état de santé, niveau d'éducation, salaire, expérience IT)	Descriptif	La rétroaction a été appliquée pour améliorer de façon itérative l'utilisabilité du site Web.
43. Moradian et al., 2018	JMIR Cancer	surveillance et la gestion à distance de la toxicité liée à la chimiothérapie : ASyMS	Tests utilisateurs avec manipulation	Laboratoire	Non précisé Satisfaction, efficacité, efficacité	Qualitative Quantitative	Subjectif Objectif	Entretiens semi-structurés, questionnaire (TAM), enregistrement vidéo, think aloud	N = 10 patients cancéreux (âge, niveau d'éducation, expérience IT)	Descriptif	Une évaluation des effets de l'ASyMS sur les résultats des symptômes dans le cadre d'un essai clinique est nécessaire.
44. Harte et al., 2017	JMIR Human Factors	Evaluer en permanence le risque de chute système WIISEL (Wireless Insole for Independent and Safe Elderly Living)	3 phase de conception : 1er phase besoin utilisateur (entretien) 2 ^{ème} phase : heuristic evaluations et cognitive walkthroughs 3 ^{ème} : tests utilisateurs	Laboratoire	Efficacité, (le taux d'erreur et d'achèvement) efficacité, (temps nécessaire pour accomplir la tâche) satisfaction (questionnaire)	Qualitative Quantitative	Subjectif Objectif	Entretien, heuristique + cognitives walkthrough avec une petite équipe, think aloud, enregistrement vidéo, prise de note, questionnaire (SUS, ASQ) Entretiens post-test	Petit échantillon	Pas de résultat	Exemple de méthode exemplaire d'après les auteurs « Cette surveillance peut être effectuée dans différents environnements, avec des sessions de laboratoire permettant un meilleur contrôle de l'expérience et des données plus robustes, mais avec une perte de

45. Van de Vaart et al., 2019	JMIR Formative Research	Thérapie cognitive du comportement sur internet (iCBT) pour la douleur chronique : Master Your Pain	Tests utilisateurs avec manipulation (6 compétences numériques)	Laboratoire	Non précisé Efficacité, efficacité, satisfaction	Qualitative Quantitative	Objectif Subjectif	Observation direction (avec tâche de performance), enregistrement video, think aloud, questionnaire	N = 32 (moy = 49.9 ans, niveau d'éducation, en esanté)	Corrélation, descriptif	fidélité au monde réel » L'utilisabilité a été corrélée de manière modéré avec l'âge et positivement de manière modéré avec les compétences en littératie en esanté mais pas avec le niveau d'éducation.
46. Quaedacker s et al., 2020	JMIR mHealth and uHealth	Surveillance de symptômes des narcoleptiques (trouble chronique du sommeil)	Utilisation de l'application pendant 30jours puis entretien	In situ rétrospectif (manuels envoyé par la poste)	Efficacité et facilité d'utilisation	Qualitative Quantitative	Subjectif	Entretiens et questionnaires (UEQ)	N = 7 patients narcoleptiques (moy = 44 ans)	Descriptif	Les entretiens personnels approfondis ont révélé 3 thèmes majeurs : (1) les raisons d'utiliser, (2) l'utilisabilité, et (3) les caractéristiques.
47. Teo et al., 2019	JMIR mHealth and uHealth	Application pour améliorer le taux de participation au dépistage chez l'homme : ScreenMen	Utilisation libre puis réflexion rétrospective	Laboratoire	Non précisé Satisfaction	Qualitative Quantitative	Subjectif	Enregistrement de l'écran, think aloud rétrospectif, questionnaire (SUS)	N = 24 hommes d'une institution bancaire (poste, âge ,niveau d'éducation, statut dépistage)	Descriptif	Ce document présentait également les principales leçons tirées du test bêta, qui est utile aux experts et aux chercheurs en santé publique lorsqu'ils élaborent une application Web mobile axée sur l'utilisateur.
48. Georgsson et al., 2019	Journal of Biomedical Informatics	Autogestion de maladie chronique (diabète)	Tests utilisateurs avec manipulation	Laboratoire	Efficacité (violation d'heuristiques) , efficacité (temps), acceptation	Qualitative Quantitative	Objectif Subjectif	Enhanced cognitive walkthrough (UC-CW) (groupe 1) , think aloud (groupe 2), questionnaire (RTLX, NASA), questions ouvertes, heuristiques	N = 12 diabétiques (caractéristiques démographique similaires)	Descriptif	Le UC-CW a été efficace pour trouver des problèmes d'utilisation graves et récurrents et il a mis en évidence le besoin de fonctionnalités personnalisées pour l'utilisateur.

49. Liu et al., 2019	Disability and Rehabilitation : Assistive Technology	Appli guide pour fauteuil roulant : «My Wheelchair Guide »	Sondage Manipulation libre	Laboratoire	Non précisé Satisfaction	Qualitatif	Subjectif	Sondage, entretiens semi-directifs	N = 10 utilisateurs N = 4 professionnels de la mobilité	Descriptif	La méthode était également efficace et a été très bien acceptée par les utilisateurs. D'autres études pourraient être menées pour évaluer l'efficacité de l'application sur l'amélioration des connaissances des utilisateurs de fauteuils roulants et la facilitation de l'autodéfense des fauteuils roulants appropriés
50. Ng et al., 2019	Digital Health	System de gestion du diabète de type 1 pour les jeunes : Diabetes YES	Participation au programme pendant 12 semaines	In situ (sauf qu'il y a un dédommagement de 50\$ AUD)	STROBE guidelines, Non spécifié facilité d'utilisation	Qualitative Quantitative	Objectif Subjectif	Google analytics (traceur d'activité), sondage, questionnaire et questions ouvertes	N = 30 jeunes diabétiques (engagement des pairs sur le soutien, zone d'habitation)	Descriptif	Le nombre de visites de la page Web a fortement diminué, tandis que l'engagement des groupes de soutien par les pairs par l'entremise de Facebook est demeuré constant tout au long de la période d'intervention. Le soutien émotionnel des pairs a été l'avantage le plus apprécié par les participants.
51. Garcia, 2019	PLoS ONE	Télémonitoring de patients atteints de problèmes reins chronique	Test utilisateur, puis utilisation libre pendant 5 mois	Laboratoire (test utilisateur) In situ (manipulation libre, mais aucune information sur l'utilisabilité présenté	Critères Nielsen : congruence Liberté Consistences Prévention des erreurs recognitions, flexibilité et effcience, esthetic, help,	Quantitative	Subjectif	questionnaire	N = 24 (moy = 53,4 ans, niveau d'étude)	Descriptif	La méthode de l'article se trouve dans l'intro

							objectif, navigabilité, interactivité, activité, feedback, motivation, satisfaction, applicabilité						
52. Cho, Yen, et al., 2018	Journal of Biomedical Informatics	Application mobile d'auto-gestion des symptômes du VIH : mVIP	Niveau 1 : tâche-utilisateur Niveau 2 : user-task-system : test utilisateur + évaluation heuristique Niveau 3 (environnement-utilisateur-tâche-système) : essai de 3 mois	Laboratoire In situ et	Health-ITUEM Erreur, complétion, mémorabilité, besoin d'information, flexibilité, apprenabilité, efficacité, compétence, autres	Qualitative Quantitative	Objectif Subjectif	Niveau 1 : tri de carte Niveau 2 : think aloud retrospectif, eye-tracking heuristique (expert) Niveau 3 : entretien + enquête	N = 20 (N1)+ 20 (N2)+ adults diagnostiqué HIV	Descriptif			
53. Stiles-Shields et al., 2017	JMIR Human Factors	application de renforcement des compétences qui engage les utilisateurs dans une restructuration cognitive : CBITs, Chicago	Test utilisateur	Laboratoire	Facilité d'apprentissage : taux d'erreur (efficacité) et temps d'exécution (efficacité), Performance d'apprentissage (connaissances + compétences)	Qualitative Quantitative	Objectif Subjectif	Enregistrement vidéo, entretien directif, questionnaire, temps d'interaction, enregistrement des actions des utilisateurs	N = 20 adults (historique des dépressions)	Test non paramétrique (U de Mann-Withney) Descriptif	Les objectifs d'apprentissage du Thought Challenger lors de son utilisation initiale ont été atteints lors d'une évaluation auprès d'utilisateurs finaux probables. Les résultats suggèrent donc que les applications sont capables de fournir aux utilisateurs des opportunités d'apprentissage des compétences d'intervention.		
54. Fuller-Tyszkiewicz et al., 2018	JMIR Mental Health	intervention personnalisée, autoguidée et basée sur une application pour la	Utilisation de l'application pendant 12 semaines	In situ	Non spécifié Satisfaction	Qualitative Quantitative	Subjectif	Questionnaire (SUS), entretiens semi-directifs	N = 5 ends users (18-25 ans), N = 5 pro de santé et N = 5 chercheurs	Descriptif	Les groupes d'experts ont nettement moins bien noté l'application que		

			dépression : BlueWatch									le groupe d'utilisateurs finaux
55. Jeon & Park, 2018	Healthcare Informatics Research	Application d'autogestion du diabète basée sur le modèle information-motivation-comportement (IMB)	Algorithme, heuristiques, test utilisateur (utilisation sur 2 semaines)	In situ retrospectif	Satisfaction	Qualitative	Subjectif	Entretiens semi-directifs	N = 14 patients diabétiques (test U) (moy = 43 ans)	Descriptif	Il reste nécessaire d'évaluer pleinement l'effet de l'application DSM sur les changements de comportement des patients diabétiques. / En gros il reste à faire des évaluations objectives	
56. Krishnamurti et al., 2017	JMIR mHealth and uHealth	Application afin d'évaluer et de communiquer les risques de grossesse liés à la naissance prématurée. + une balance	Utilisation libre sur 3 mois	In situ	Satisfaction	Qualitative Quantitative	Subjectif	Entretiens semi-directifs mensuels téléphonique Questionnaire quotidien	N = 16 patiente (éducation, salaire, nombre de grossesses)	Descriptif Test paramétrique	L'utilisation de l'application était plus élevée chez les participantes les plus à risque,	
57. Ose et al., 2017	JMIR Research protocol	Dossier de santé électronique personnel (projet INFOPAT) : PEPA	Prise en compte des conditions organisationnelles + simulation de l'activité (tests utilisateurs)	Laboratoire	Non précisé Satisfaction	Qualitative Quantitative	Objectif Subjectif	Think aloud, questionnaire (SUS), entretiens semi-directifs, enregistrement vidéo + entrevu avec les patients, médecins pour l'organisation	N = 15 patients Cancer colorectal	Pas de résultat / uniquement présentation de la méthode	Les patients ont accès à leurs DSP et ils peuvent donner accès à d'autres personnes (par exemple, leur médecin généraliste). Avec cette approche, de nouvelles possibilités de collaboration professionnelle et d'engagement des patients peuvent se présenter.	
58. Yen et al., 2016	Applied Clinical Informatics	plateforme partagée intégrant les données des patients et des médecins : RhEumAtic Disease activitY (READY)	Tests utilisateurs avec manipulation (tâches) (laboratoire) TMS (In situ)	Laboratoire In situ	Non spécifié	Qualitative Quantitative	Objectif Subjectif	Think aloud, Time motion studies (TMS)	N = 33 patients et N = 15 médecins N = 89 observation TMS	Descriptif Test t	Les patients ont préféré le système READY aux documents papier. Beaucoup ont trouvé qu'il était plus facile d'entrer l'information en raison de la taille plus grande de la	

police et de la facilité de " taper " plutôt que d'écrire ou d'encrer les réponses.

59. Wozney et al., 2016	JMIR Human Factors	Plateforme collaborative pour intervention à distance : Blackboard Collaborate	Cycle itératif de test. test utilisateur suivi par un modérateur	In situ/terrain (présence du modérateur, mais consigne plutôt libre)	Contenu et esthétique (satisfaction)	Qualitatif	Subjectif	Entretiens téléconférence, notes de synthèse,	N = 4 modérateurs N = utilisateur non spécifié	Descriptif (verbatim)	La plupart des recherches comparatives se sont concentrées sur l'impact de l'environnement physique sur les résultats, et rarement sur la façon dont les éléments sociotechniques de l'environnement de test affectent les performances des modérateurs et des utilisateurs. Une meilleure compréhension des rôles des modérateurs et des expériences de conduite de tests d'utilisabilité "en situation réelle" peut conduire à l'amélioration des techniques et des stratégies
60. Berry et al., 2015	BMC Medical Informatics Decision Making	Logiciel d'adaptation du profil personnel du patient de la santé en ligne - aide à la décision pour les hommes hispanophones d'origine latino : the Personal Patient Profile-Prostate (P3P)	Test utilisateur	Laboratoire	Non spécifié (les auteurs citent Nieslen sans citer les métriques)	Qualitative	Subjectif	Entretien, think aloud, enregistrement des écran, questionnaire (Acceptability e-Scale)	N = 7 patients (54 à 67 ans), niveau d'étude avec cancer de la prostate	Description	

Articles	Revue	Matériel(s) et intérêt : nom du produit	Tâche(s)	Environnement(s)	Métrique(s)	Quantitative/Qualitatif	Subjectif/Objectif	Outils/méthodes	Population	Analyse(s) de données	Résultats
61. Smaradottir et al., 2015	Studies in Health Technology and Informatics	Surveillance à distance des patients atteints de bronchopneumopathie chronique obstructive (COPD) : United4Health L'outil comprend un oxymètre et questionnaire	Formation, manipulation libre (3 tâches quotidiennes pendant 7 jours) et tests utilisateurs (devant infirmier et chercheurs)	Terrain + labo (formation puis utilisation de l'application pendant 7 jours)	Non spécifié	Qualitative	Subjectif	Entretien au domicile, think aloud, manipulation libre	N = 6 patients (moy = 72 ans) avec COPD (expérience IT,	Analyse qualitative de contenu	
62. Davidson et al., 2014	International Journal Of Psychiatry Medecine	Module de dépression pour adolescent	Test utilisateur	Laboratoire	satisfaction	Qualitative Quantitative	Objectif Subjectif	Enregistrement video, observation direct, think aloud, enregistrement de l'écran, questionnaire	N = 24 adolescents (moy = 14,12 ans)	Descriptif	Détail du protocole de l'évaluation de l'utilisabilité sur un autre papier (en cours de soumission, mais introuvable). Réalisation d'une second étude « clinique » sur 2000 ado
63. Sheehan & Lucero, 2015	eGEMS : The Journal for Electronic Health Data and Methods	SAPHeR (Self-management system-self-Assessment)	20 tasks for 4 different module	Laboratoire + in situ (etude de faisabilité en 8 semaines)	TURF Apprenabilité (nombre d'erreurs) efficience (nombre d'étapes requis) satisfaction	Qualitative Quantitative	Objectif Subjectif	Direct observation (error rate) Think aloud enregistrement de l'écran And PSHUQ (Post-Study E-health Usability Questionnaire) questionnaire	N= 10 participants (moy = 72.7ans)	Descriptif	
64. Fung et al., 2015	Disabilities and health Journal	Equipement medical : pression positif des voies aériennes (PAP)	entretiens	Laboratoire	Non précisé	Qualitatif	Subjectif	Entretiens	N= 19 (9 patients atteints d'une déficiences physique/sensorielles et 10 pro de santé)	Descriptif	Les dispositifs médicaux à domicile tels que le PAP peuvent ne pas être actuellement conçus pour répondre aux besoins de certains utilisateurs

											souffrant de handicaps physiques/sensoriels.
65. Abts et al., 2014	Proceedings of the International Symposium on Human Factors and Ergonomics in Health Care	Pompe à perfusion (dispositif médical), mais surtout évaluation des instructions	Heuristiques puis 2 cycles de tests utilisateurs (scénario d'exécution de tâche)	Laboratoire	Non précisé Satisfaction	Qualitatif Quantitatif	Subjectif	Think aloud, entretiens, questionnaire	1 ^{er} session N = 5 (4 hommes) récemment utilisent une pompe à perfusion (49-69 ans) 2 ^{ème} session N = 6 novices (25-45ans, expérience IT,	Descriptif	Une tendance dans les réponses était le manque de ressources d'aide supplémentaires que les participants ont déclaré avoir reçu en dehors de la formation pratique initiale sur la pompe. En gros, les instructions ne suffisent pas. L'étude démontre la valeur des guides visuels pour les patients utilisant des dispositifs médicaux
66. Yang & Mahmud, 2018	Journal of Design Research	1 système de télésurveillance. 2 idem sauf conçu par le partenaire Pas d'identification des dispositifs (HMD) pour des raisons de confidentialité. Mais ce sont des dispositifs qui existent sur le marché	Test utilisateur	Terrain (50€ de dédommagement)	Non spécifié	Quantitatif Qualitatif	Objectif Subjectif	Entretien avec les partenaires industriels, entretien, observation, questionnaire, notes et photographies	N = 82 au domicile (N=41 par devices)	Descriptif	Les résultats ont révélé trois problèmes concernant l'approche du fournisseur : les caractéristiques d'utilisation n'avaient pas grand-chose à voir avec les réactions positives des utilisateurs ; le manque de compréhension des caractéristiques du public cible ; et les résultats contradictoires liés aux performances et aux préférences des utilisateurs.

											Les causes étaient liées à l'objectif du test utilisateur, à l'expertise du personnel qui l'a réalisé et à la nature tacite et non articulée de l'expérience et des attentes des utilisateurs.
67. Chong et al., 2019	Proceedings of the International Symposium on Human Factors and Ergonomics in Health Care	Tensiomètre : ReliOn™ BP200 Blood + évaluation des instructions	hierarchical task analysis, analyse des besoins utilisateurs, et tests heuristique et test utilisateurs	Laboratoire	Non précisé Efficacité Satisfaction	Quantitatif	Objectif Subjectif	Observation direct, questionnaire (SUS), tests de compréhension	N = 19 (moy = 25ans), pas d'expérience en tensiomètre	Descriptif	Notre évaluation du GSG (getting Started Guide) et de l'IFU (instruction for use) s'est avérée critique, car environ 10 % des utilisateurs testés n'ont pas été en mesure de trouver des informations sur la manière d'interpréter leurs relevés de pression artérielle
68. Jones & Caird, 2017	Proceeding of the Human Factors and Ergonomics Society Annual Meeting	Deux glucomètres de marché : Accu-Check Compact Plus® and 2) the One Touch Ultra 2®	Test utilisateurs	Laboratoire (Healthcare Human factors Simulation Laboratory (HHFSL) avec une chambre de simulation	Non précisé Succès et échec et nombre d'erreurs (efficacité) Temps de complétion de la tâche (efficience)	Quantitatif	Objectif	Enregistrement vidéo	N = 45 (N = 16 jeunes diabétiques (18-27) et N = 29 personnes âgées (65-85)); démographique, et habilité cognitive	Descriptif , Khi-2	Les personnes âgées font significativement plus d'erreur que les jeunes
69. Prado-Jimenez et al., 2018	Avances en Interaccion Humano-Computadora	Glucomètre	Entretiens avec clinicien, tests utilisateurs avec end-users (5 tâches)	Laboratoire	Efficience, efficacité, satisfaction	Qualitatif	Subjectif	Entretien avec clinicien, entretiens avec patients, Verbal Protocol Analysis (VPA), think aloud	N = 20 adults (65ans diabétique)	Non spécifié	

Articles	Revue	Matériel(s) et intérêt : nom du produit	Tâche(s)	Environnement(s)	Métrique(s)	Quantitatif/Qualitatif	Subjectif/Objectif	Outils/méthodes	Population	Analyse(s) de données	Résultats
70. Werner et al., 2018	Ergonomics	PHIM : Gestion des informations personnelles sur la santé : VizHOME project	Méthodes ethnographique et cognitive task analysis / exploration libre	Laboratoire (Réalité virtuel) : récréations virtuelles d'environnements domestiques réels (LIDAR) ; 3D replica virtuel + 60\$ de dédommagement	Non spécifié	Qualitatif	Subjectif	Verbatim concurrent et retrospective (think aloud), entretien	N = 20 diabétiques (moy = 54 ans), utilisation d'un glucomètre, diet particulière, observance médicamenteuse	Descriptif	Un modèle conceptuel de la distribution de la cognition par le système sociotechnique est essentielle à la conception des processus et des technologies qui soutiennent la PHIM des individus à domicile. Les approches dans lesquelles l'individu (sans contexte) est censée être l'unité d'analyse pertinente ; négligez les interactions essentielles entre le l'individu et son environnement, ses outils et son entourage
71. Robinson et al., 2016	Telemedicine Journal and e-Health	Visioconf sur Ipad2 FaceTime® (Apple)	Méthode longitudinale (30, 90 et 180 jours)	In situ	Satisfaction (et utilisabilité)	Quantitatif	Subjectif	questionnaire	N = 34 diabétique (moy = 51 ans)	Descriptif	Soixante-seize pour cent des patients ont convenu que FaceTime était efficace pour améliorer le diabète. Des recherches supplémentaires sont nécessaires pour déterminer les résultats économiques et cliniques de la technologie

											visuelle bidirectionnelle
72. Nitsch et al., 2016 & Nitsch et al., 2019	Journal of Medical Internet Research & Nutrients	le programme "Image corporelle saine"	Test utilisateur itératif (2 tests)	Laboratoire	Non spécifié Satisfaction	Quantitative Qualitative	Subjectif	Entretien cognitive et semi-directif , think aloud et questionnaire (SUS)	N = 9 personnes ayant des problèmes d'image corporelle ou des troubles alimentaires	(Analyse thématique) Descriptif	La plupart des études sur les interventions de santé numérique manquent de conception de recherche et de méthodes appropriées pour étudier les questions d'engagement individuel. L'engagement des utilisateurs et l'utilisabilité des programmes sont inextricablement liés. Le soutien du coach, l'investissement en temps et la gravité des symptômes des utilisateurs, les caractéristiques et l'efficacité du programme, la confiance, l'anonymat et l'accessibilité financière sont autant d'éléments pertinents pour l'engagement.

73. Albrecht et al., 2018	JMIR Aging	Tensiomètre avec télémonitoring	Utilisation du dispositif pendant 1 semaine (4 mesures par jours)	In situ rétrospectif (utilisation du device à la maison puis retour au lab pour entretien	Non spécifié	Qualitative	Subjectif	entretiens	N=7 seniors (mean age = 86, entre 70-95 ans, statut marital)	Descriptif	
74. Scott et al., 2017	JMIR Mhealth Uhealth	Application pour la chirurgie ambulatoire colorectal	Utilisation pendant 14 jours après leur sortie à l'hôpital	In situ rétrospectif (utilisation du device à la maison puis retour au lab pour entretien / Étude de faisabilité prospective	Non spécifié (Satisfaction) + (Nombre de connexion = efficacité)	Quantitative Qualitative	Subjectif Objectif	Questionnaire (SUS), entretiens semi-directifs (avec le patient et les membres de la famille), Traceur d'activité	N = 17 patients	Descriptif	Cette étude montre qu'une grande utilisabilité ne suffit pas à elle seule à motiver les patients à utiliser les applications des smartphones dans la période post-opératoire. Les schémas d'utilisation et les taux de réponse après avoir été invités à contacter l'équipe chirurgicale suggèrent que les applications doivent être très attrayantes pour être adoptées par les patients.
75. Vooijs et al., 2014	Interactive Journal of Medical Research	Accéléromètre bon marché : Fitbit & Physical Activity Monitor (PAM) Un accéléromètre multicapteur sophistiqué (SensWearArband) qui a servi de critère de référence	Utilisation pendant 7 jours	In situ rétrospectif (utilisation du device à la maison puis retour au lab pour entretien	Non spécifié	Quantitatif	Subjectif	Questionnaire	N = 16 patients	Modèle de régression, test t	Le Fitbit et le PAM, peu coûteux, sont des appareils valides et utilisables pour mesurer l'activité physique chez les patients atteints de BPCO.

76. Rantanen et al., 2017	Clinical Therapeutics	Dispositif robotique et système de téléassistance pour l'observance du traitement médicamenteux chez les patients âgés à domicile	Phase 1 : condition contrôlée maison de soins pendant 26.9 jours Phase 2 : patients à domicile	In situ	Satisfaction	Quantitative Qualitative	Subjectif Objectif	Entretien (patient et infirmière) et enregistrement des données du robot, nombre d'alerte du robot	N = 17 maison de repos N = 27 domiciliés (78% seuls)	Descriptif	This pilot clinical safety and usability study has limitations: (1) small sample size, (2) short duration, (3) no direct observation of medicine consumption (patients were selected by nurses based on their assessment that the potential participant was motivated to take their medicines), and (4) data were missing on the frequency of home visits by nurses at baseline in Phase II
77. Arora et al., 2018	The Journal of Sexual Medicine	auto-injecteur sous-cutané de testostérone à usage unique, pré-rempli et jetable : SCTE-AI incluant l'emballage et les instructions d'utilisation	Tests utilisateurs à 4 groupes : avec ou sans expérience d'injection & Formation avec manuel d'utilisation ou seulement manuel d'utilisation	Laboratoire : Simulation à domicile	Compréhensibilité, l'intuitivité/facilité d'utilisation + non précisé : efficacité	Quantitatif Qualitatif	Subjectif Objectif	Succès/echec de la tâche, quasi-erreur. enregistrement vidéo, entretiens semi-directifs et observation direct	N = 65 patients (1 dose) et N = 59 patient (2 doses)	Descriptif	La principale limite de l'étude est la présence d'observateurs/caméras qui peuvent avoir distrait ou créé une anxiété de performance, contribuant potentiellement aux erreurs. Les résultats de cette étude de validation de la convivialité indiquent que le dispositif SCTE-AI est sûr et intuitif à utiliser, qu'il présente un faible potentiel de nuisance et qu'il est associé à un taux élevé de réussite des

											injections, indépendamment de la formation ou de l'expérience antérieures.
78. Agnisarman et al., 2017	Applied Ergonomics	4 plateformes de télémédecine pour le domicile patient : Doxy.me, Vidyo, VSee, and Polycom	Tests utilisateurs avec scénario à suivre	Laboratoire (les participants jouent le rôle de diabétique et les chercheurs le rôle de docteur)	Le nombre d'erreur, le temps de complétion de la tâche, NASA : Charge de travail, effort, frustration, satisfaction, utilité, qualité de l'information, la facilité d'utilisation, préférence	Quantitatif Qualitatif	Subjectif Objectif	Observation direct (temps d'utilisation), retrospective think aloud, questionnaire (CSUQ, NASA, questionnaire adhoc sur l'utilité et la facilité d'utilisation dérivé du TAM)	N = 19 participants (moy = 25ans, niveau d'éducation, utilisation de video chat)	ANOVA, Test non paramétrique (friedman test et Wilcoxon Signed-Rank)	Technological systems need to be designed to serve the needs of people across different population and socioeconomic background to promote the adoption
79. Liang et al., 2018	JMIR mhealth & uhealth	7 dispositifs portables : apple Watch, Samsung Gear S, Fitbit Surge, Jawbone Up3, Mi band, Huawei Honor B2 et Misfit Shine. Fonction de podomètre, minteur d'activité et enregistrement à distance	Questionnaire en ligne : Enquête	In situ retrospectif	Efficacité, efficacité, satisfaction (mais dans les faits, uniquement satisfaction)	Quantitatif	Subjectif	Méthode de comparaison, questionnaire (SUS)	N = 388 participants (utilisateurs d'une application évaluée) (âge, fréquence d'utilisation, éducation, profession, revenu)	ANOVA. Corrélation (entre le SUS et les caractéristiques du produit)	une très faible corrélation négative entre les scores SUS et le prix (r=-.10, P=.03) et faible corrélation positive entre les scores SUS et les appareils qui prennent en charge la fonction d'identification du type d'activité (r=.11, P=.02). Aucune différence significative n'a été observée entre les marques. Le score SUS avait une corrélation positive modérée avec l'expérience de l'utilisateur (durée d'utilisation

du dispositif)
($r=0,32$, $P<0,001$).
les participants
des secteurs de la
médecine et des
soins de santé ont
donné un score
significativement
plus élevé
(moyenne 61,1,
ET 17,9 contre
moyenne 68,7, ET
14,5, $P=0,03$). pas
de corrélation
entre la profession
et le SUS, mais
corrélation avec le
revenu mensuel.

Annexe 5 : Fiche produit iHealth Sense Tensiomètre poignet connecté BP7

iHealth Sense

Tensiomètre poignet connecté BP7

Compatible avec iPhone, iPad et smartphones Android



Suivi de l'évolution de la tension sur smartphone et iPad

Pour iOS et Android version 3.0 et +.
Voir la liste des appareils compatibles sur ihealthlabs.eu/support

Fonctionnalités iHealth Sense

- Mesure de la tension artérielle (systole, diastole) et de la fréquence cardiaque (nombre de battements cardiaque par minute)
- Détection d'arythmie
- Détection de la position du poignet pour une mesure optimale
- Suivi des données, graphiques et historiques sur l'application
- Synchronisation sans fil Bluetooth
- Évaluation des mesures selon les recommandations de l'OMS*
- Facile à utiliser, léger et compact
- Partage possible des résultats avec des professionnels de santé et/ou sa famille
- Rechargement par câble USB
- Stockage des données gratuit et sécurisé

*Organisation Mondiale de la Santé

Suivi des données de santé avec l'application gratuite iHealth MyVitals



Compatible avec la gamme de produits de santé iHealth :
les trackers d'activité, balances et oxymètres de pouls

www.iHealthLabs.eu



CE0197



iHealth
Garantir votre santé



Spécifications techniques :

- Connectivité : Bluetooth 3.0 + EDR
- Taille : 72mm x 74mm x 17.6mm
- Circonférence brassard : 13.5cm-22cm
- Poids : 105g
- Méthode de mesure : technique oscillométrique avec inflation et mesure automatique
- Alimentation : DC : 5V, 1.0A, 1 x 3.7 V, Li-ion 400mAh
- Plage de mesure :
Sys : 60-260mmHg
Dia : 40-199mmHg
- Plage de mesure du pouls : 40-180bts/min,
- Conditions d'utilisation :
Temp : 5°C-35°C, Humidité : <90%RH
- Conditions de transport et de stockage :
Temp : -20°C-55°C, Humidité : <95%RH



Dans la boîte :

- 1 tensiomètre connecté iHealth Sense
- 1 guide d'installation rapide
- 1 manuel d'utilisation
- 1 câble de recharge USB
- 1 coffret de transport

Ce dispositif médical de classe IIa est un produit de santé réglementé qui porte, au titre de cette réglementation le marquage CE. Lire attentivement la notice.

Annexe 6 : Fiche produit Oxymètre de pouls connecté iHealth Air P03M

iHealth Air

Oxymètre de pouls connecté P03M

Compatible avec iPhone, iPad et smartphones Android



Suivi de la saturation en oxygène et du pouls sur iPhone, iPad et smartphones Android

Pour iOS et Android version 4.0 et +.
Voir la liste des appareils compatibles sur ihealthlabs.eu/support

Fonctionnalités iHealth Air

- Mesure et affiche instantanément sur l'oxymètre et sur le smartphone/tablette :
 - La SpO2 : Saturation pulsée en oxygène dans le sang
 - Le Pouls : Nombre de battements de coeur par minute
 - IP : Indice de perfusion sur l'application
- Suivi des données, graphiques et historiques sur l'application
- Synchronisation sans fil Bluetooth 4.0
- Partage possible des résultats avec des professionnels de santé et/ou famille
- Rechargement par câble USB
- Stockage des données gratuit et sécurisé



Suivi des données avec l'application gratuite iHealth MyVitals



Compatible avec la gamme de produits de santé iHealth : les trackers d'activité, tensiomètres et balances

www.iHealthLabs.eu



CE 0197 FDA



Spécifications techniques :

- Alimentation : Bluetooth 4.0 BLE
- Système d'affichage : LED
- Taille : 62mm x 33mm x 28mm
- Poids : 42g
- Alimentation : batterie, 3.7V Li-ion, 300mAh
- Plage de mesure SpO2 : 70-99%
- Plage de mesure du pouls : 30-250bpm
- Précision SpO2 : 70 - 99%, ± 2%; < 70%
- Précision de la fréquence du pouls : ± 2bpm
- Arrêt automatique après 8 secondes
- Environnement de fonctionnement : Temp : 5 - 40°C; Humidité < 80%RH
- Environnement de stockage : Temp : -20 - 55°C; Humidité < 95%RH



Dans la boîte :

- 1 oxymètre de pouls iHealth Air
- 1 dragonne
- 1 câble de recharge USB
- 1 guide d'installation rapide
- 1 manuel d'utilisation

iHealth
Quantitative & Qualitative Health

Ce dispositif médical de classe IIa est un produit de santé réglementé qui porte, au titre de cette réglementation le marquage CE. Lire attentivement la notice

Annexe 7 : Prospection sur le nombre d'interventions et l'âge des patients en chirurgie ambulatoire

Racine	Libellé	Nombre d'intervention par année (pourcentage en chirurgie ambulatoire en 2018)	Nombre d'intervention en chirurgie ambulatoire (proportion sur l'ensemble des actes de chirurgie ambulatoire)
02C05	Interventions sur le cristallin avec ou sans vitrectomie	877 859 (95.3 %)	836 769 (22.40 %)
03C02	Affections de la bouche et des dents avec certaines extractions, réparations et prothèses dentaires	322 443 (97.1%)	313 003 (8.37 %)
08C44	Autres interventions sur la main	146 690 (90,4 %)	132 553 (3,54 %)
01C15	Libérations du médian au canal carpien	133 893 (96.7%)	129 504 (3.46 %)
05C17	Ligatures de veines et éveinages	104 928 (90.3%)	94 738 (2.53 %)
08C14	Résections osseuses localisées et/ou ablation de matériel de fixation interne au niveau d'une localisation autre que la hanche et le fémur	110 466 (84.5%)	93 305 (2.49 %)
09C03	Greffes de peau et/ou parages de plaie à l'exception des ulcères cutanés et cellulites	111 026 (79,5 %)	91 831 (2.45 %)
12C08	Circoncision	91 357 (96,7 %)	88 897 (2.38 %)
06C12	Interventions réparatrices pour hernies inguinales et crurales, âge supérieur à 17 ans	125 258 (70.0%)	87 648 (2.34 %)
08C45	Ménisectomie sous arthroscopie	88 647 (94.3%)	83 579 (2.23 %)
07C14	Cholécystectomies sans exploration de la voie biliaire principale à l'exception des affections aiguës	86 454 (52.9%)	45 713 (1.22 %)
08C48	Prothèses de hanche pour des affections autres que des traumatismes récents	108 757 (2.1%)	2 292 (0.06 %)
08C37	Interventions sur le pied, âge supérieur à 17 ans	107 163 (66.3%)	70 999 (0.02 %)
08C39	Interventions sur l'avant-bras	69 896 (30.6%)	21 377 (0.57 %)

(source [Scansanté](#))

Annexe 8 : Présentation des personas et leurs scénarios

Damien

a. Personas Damiens

Damien





- **Technologies**
 - Grand technophile : Téléphone, PC, tablette...
 - Utilisation courante et réseaux sociaux
 - Grande confiance
 - *Si j'ai une question, je regarde sur internet*
- **Santé**
 - Grande forme sauf accident ponctuel
 - Bien être, forme et état général mais pas médical
 - Il prend ses rendez-vous médicaux sur le net
 - *« J'aimerais mieux faire les consultations à distance plutôt que de perdre mon temps dans les salles d'attente »*
 - Fait attention à une alimentation stricte
 - Adaptée à ses sports
- **Plaisir/ émotions, comportement**
 - Fait de la musculation et du cross fit
 - Il aime sortir et voyager
 - Autonome, il n'aime pas attendre, les contraintes
 - Dynamique, toujours pressé, tendance à aller vite
 - Risque d'oublier (géolocaliser la tablette)

Opération Ligaments croisés

Figure 68. Personas Damien

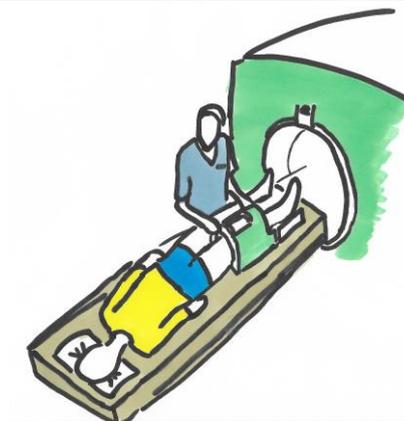
b. Scénario Damien

Texte audio (voix d'homme)	Extrait des images
<p>Damien a 23 ans et habite à Nice. C'est un étudiant qui aime vivre sans contrainte et passe beaucoup de temps à pratiquer ses activités sportives comme le fitcross.</p>	
<p>Damien surveille son activité physique grâce à ses objets connectés comme sa montre connectée. Il peut ensuite partager ses performances avec ses amis.</p>	

Malheureusement, ce jour là, Damien n'a pas fait attention à cette branche. Il est mal retombé et c'est visiblement blessé au genou.



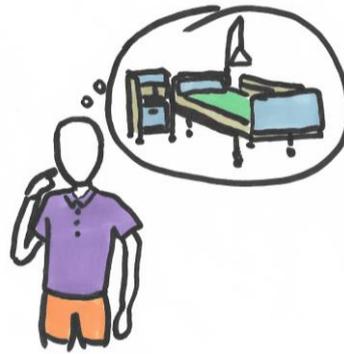
Comme Damien avait extrêmement mal après sa chute il a préféré faire des analyses. Le personnel soignant trouve une rupture des ligaments croisés. Il faut alors prendre rendez-vous chez un chirurgien pour une opération.



Trois semaines plus tard, Damien voit son chirurgien. Il lui explique comment va se dérouler l'opération et la rééducation.



Damien n'est vraiment pas rassuré. Il a peur d'avoir mal et de ne pas pouvoir bouger après l'opération.
Damien s'inquiète pour ses cours et pour ses compétitions. La saison est terminée pour lui...



Quelques jours plus tard, Damien voit l'anesthésiste pour obtenir des informations sur son état de santé. C'est la première fois qu'il est opéré, il ne sait pas s'il fera une réaction à l'anesthésie. L'anesthésiste lui demande s'il veut avoir un dispositif médical pour prendre soin de lui lorsqu'il sera chez lui. Damien accepte sans trop savoir de quoi il s'agit.



Il voit ensuite une infirmière qui lui explique comment il doit se laver et se raser avant l'opération.



Le jour j, Damien se rend à l'hôpital dans le service de chirurgie ambulatoire la boule au ventre. Les infirmiers lui demandent de se déshabiller pour partir au bloc rapidement.

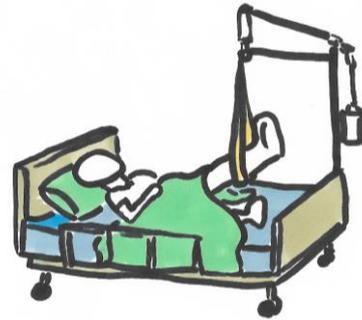


Pendant l'opération tout se passe bien. Damien est toujours sous anesthésie lorsqu'il ressort.



Après son réveil, il reste dans son lit, nauséux à attendre qu'on lui autorise de sortir.

La douleur commence à se réveiller, lorsqu'une infirmière entre pour lui présenter un dispositif *Smart Angel* qu'il doit utiliser lorsqu'il rentrera chez lui. L'infirmière lui montre comment on utilise la tablette et les deux objets connectés (tensiomètre et oxymètre de pouls). Elle lui explique que grâce à ces outils, l'hôpital a des nouvelles du patient et qu'ils n'auront pas besoin de l'appeler sauf s'il ne se sent pas bien.



Damien rentre accompagné de ses parents chez lui avec la mallette *Smart Angel*.



Une fois chez lui, la douleur s'est réveillée, et Damien reste dans son lit à ne plus pouvoir bouger. Sa mère vient alors l'aider à répondre aux questions de la tablette.



Toute la semaine, Damien mesure sa tension et son taux d'oxygène grâce au dispositif. Il n'a pas eu de problème post-chirurgical. Il reprend les cours le lundi prochain mais devra attendre la fin de la rééducation avant de refaire du sport.

L'hôpital était informé en temps réel des besoins de Damien. Mais les résultats de ses mesures étant satisfaisants, le personnel hospitalier n'a pas eu besoin de joindre Damien.



Marianna

a. Personnas Marianna

Marianna





- **Profil**
 - Jeune, 37 ans, 2 enfants, cadre marketing
 - Vit en banlieue de Lyon
 - Concilier sa vie professionnelle et familiale
 - Pas de temps pour elle
 - Une phrase clé : « *Chaque minute compte* »
- **Plaisir/ émotions, comportement**
 - J'ai jamais de temps mais je me réserve 1h de Yoga
 - Elle n'aime pas les embouteillages, faire les courses...
 - Elle est stressée, organisée, limite psycho rigide
 - Elle sait anticiper (utilisation de l'agenda)
 - Souvent inquiète surtout pour les enfants

Opération Ligature des trompes

- **Technologies**
 - Elle utilise beaucoup l'informatique pour son travail
 - Elle n'aime pas quand sa connexion déraile
 - Utilisation courante
 - d'un smart phone, d'une tablette, d'un PC
 - Un peu sur les réseaux mais n'a pas beaucoup de temps
- **Santé**
 - Fatiguée, manque de sommeil, limite dépressive
 - Elle pense à ses enfants avant de penser à elle
 - Elle prend des médicaments
 - elle fait confiance à son médecin
 - L'opération prévue : ligature des trompes
 - Connaît des appareils de domotique
 - Mais rien lié à la santé
 - Connaissances/ médicaments
 - J'ai toujours un doliprane et du sédatif PC
 - Motivations/ service proposé
 - C'est très rassurant et on peut prévoir les RdV
 - C'est pratique parce qu'on peut rester à la maison

Figure 69. Personnas Marianna

b. Scénario Marianna

Texte audio (voix de femme)	Extrait des images
<p>Marianna est très occupée. Entre son travail et ses trois enfants elle a peu de temps à se consacrer à elle. D'ailleurs, elle est sûre de ne plus vouloir d'enfants et aimerait un moyen de contraception adapté afin d'éviter les surprises.</p>	
<p>Un jour, une amie lui parle de l'opération de la ligature des trompes. Une convalescence courte et pas de prise de contraceptif, le tout en chirurgie ambulatoire.</p>	

Marianna s'est laissée séduire. Elle prend donc rendez-vous chez un chirurgien.



Son mari et elle rencontrent le chirurgien qui leur a expliqué l'intervention, les risques mais aussi le caractère irréversible de l'opération.



Après la consultation, Marianna est décidée, elle veut se faire opérer. Elle craint juste que sa convalescence dure trop longtemps.



Peu de temps avant l'intervention, elle rencontre l'anesthésiste qui prend soin de lui poser des questions sur sa santé et ses antécédents. Marianna semble pressée. Il ne faudrait pas qu'elle rate ses enfants à la sortie de l'école. Elle semble peu préoccupée par ce que le médecin ou l'infirmière éducative lui expliquent.



Elle voit ensuite une infirmière qui lui explique comment elle doit se laver et se raser avant l'opération.



Trois semaines plus tard, c'est le jour de l'opération et Marianna se sent de plus en plus anxieuse. Elle a peur d'avoir oublié quelques choses dans son organisation d'habitude sans failles.

Marianna avance dans les couloirs de l'hôpital la boule au ventre. Les infirmières lui demandent de se déshabiller pour partir au bloc rapidement.



L'opération se passe rapidement. En salle de réveil, elle se réveille avec de fortes douleurs au bas ventre.



Après avoir pris les premiers calmants, la douleur est plus faible. Elle est surtout pressée de rentrer pour reprendre le cours de sa vie. Son mari est venu la chercher. Mais avant de partir, un infirmier vient lui présenter le dispositif *Smart Angel*. Marianna comprend alors ce que voulait dire l'anesthésiste. Elle écoute avec son mari les instructions de l'infirmier et prend la mallette avec elle.



Marianna a eu son autorisation de sortie. Elle rentre chez elle accompagnée de son mari.



Une fois au domicile, Marianna suit les recommandations du médecin et utilise le dispositif. Cependant, fidèle à elle-même, elle reprend le travail avant la fin de son arrêt maladie.



Toute la semaine, Marianna a suivi à la lettre les demandes de la tablette. Elle a été très satisfaite d'avoir ce petit appareil qui la raccroche à l'hôpital. Elle déclare : « au moins, ils savent que je vais bien, et ça, cela me rassure ». L'hôpital n'a pas eu besoin de contacter Marianna au vu des résultats satisfaisants transmis par le dispositif.



Pierre

a. Personnas Pierre

Pierre





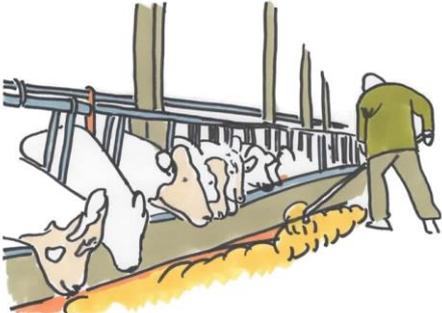
- **Profil**
 - 45 ans, marié, 3 enfants en internat
 - Eleveur en Normandie (lait), Maire de sa commune
 - Habitat rural; Développer l'exportation
 - Contacts en Chine; Il travaille 7jours/ 7
 - « *Je peux pas m'arrêter* »
- **Plaisir/ émotions, comportement**
 - Il défend la cause agricole et son village en particulier
 - Il valorise ses produits au salon de l'agriculture
 - Il aime la bonne nourriture, les choses saines, le bon vin
 - Il n'aime pas les grandes villes, le monde, les hôpitaux
 - Il est bon vivant, très sociable, joyeux, un peu impulsif
 - I prend souvent des risques dans son métier

Opération hernie inguinale

- **Technologies**
 - Il a un PC pour la compta et les contacts
 - Mais c'est sa femme qui fait la gestion
 - Il a un smart phone mais pas de tablette, ça fait gadget
 - Il n'est pas sur les réseaux sociaux, il n'a pas de page web
 - Il est relativement confiant dans la technologie,
 - voire plutôt ignorant d'internet et des risques
- **Santé**
 - Un peu de surpoids, un peu d'hypertension
 - « *Je peux pas être malade* »
 - Le médecin du village est son ami (adjoint au maire)
 - Il n'aime pas l'hôpital: se déplacer, attendre, impersonnel
 - Il ne fait pas attention à sa santé, il manque de sommeil
 - Opération prévue: hernie inguinale
 - Il ne connaît pas les objets connectés mais ça l'intrigue
 - Il ne se rend pas compte de toutes les possibilités
 - Mais il est ouvert
 - Il s'embête pas avec les questions de santé
 - Il demande à son ami
 - *Il a des médicaments pour l'hypertension mais il oublie*
 - Motivations/ Smart
 - Il est curieux; ça peut être pratique pour rentrer + vite
 - Il n'a pas toujours beaucoup de réseau
 - Il ne fait pas toujours attention aux objets
 - Il va sauter des Rdv à cause de son travail

Figure 70. Personnas Pierre

b. Scénario Pierre

Texte audio (voix d'homme)	Extrait des images
<p>Pierre est agriculteur et prend très peu de vacances. Il aime son métier et ses animaux. C'est aussi quelqu'un de très engagé, il est maire de sa commune en Normandie.</p>	
<p>A force de porter de lourdes charges, Pierre a de fortes douleurs abdominales et une bosse s'est formée. Il est parti consulter son médecin traitant qui lui a rapidement diagnostiqué une hernie inguinale. Pierre doit subir une petite opération qui se pratique généralement en chirurgie ambulatoire.</p>	

Pierre se rend donc chez un chirurgien qui lui explique l'opération et ses suites.



Pierre est très angoissé. L'idée même de se faire opérer lui fait terriblement peur. Il est aussi inquiet pour son exploitation, et du temps de sa convalescence.



Quelques jours plus tard, il rencontre son anesthésiste qui lui propose de faire l'opération en ambulatoire mais qu'il serait aidé par un dispositif du nom de *Smart Angel* à la maison. Pierre n'est pas très rassuré de ce système connecté mais si le médecin le dit, c'est que cela doit être une bonne chose. Il accepte sans se poser de questions. Ce qu'il veut, c'est rentrer chez lui rapidement.



Il voit ensuite une infirmière qui lui explique comment il doit se laver et se raser avant l'opération.



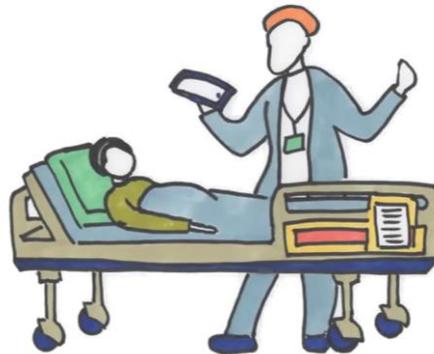
Le jour de l'opération, Pierre est très stressé. Il tremble pour enlever ses vêtements et se prépare à l'opération. Il a vraiment hâte d'en finir.



L'opération se passe bien. A son réveil Pierre est un peu nauséux à cause de l'anesthésie mais la douleur est supportable.



Après son opération, Pierre n'a qu'une envie, rentrer chez lui au plus vite. Peu de temps après sa sortie du bloc, un infirmier vient lui présenter le dispositif. Pierre essaye 3-4 fois avant de comprendre comment tout fonctionne.



Pierre a eu son autorisation de sortie. Sa femme est venue le chercher.



Une fois dans sa maison, Pierre s'allonge dans son canapé et allume la tablette *Smart Angel*. Il prend ses mesures seul.



Toute la semaine, Pierre a suivi à la lettre les demandes de la tablette. Il a été très satisfait d'avoir ce petit appareil qui le raccroche à l'hôpital. Ça le rassure de se savoir toujours en lien avec le personnel hospitalier. D'ailleurs, ses mesures étant satisfaisante, l'hôpital n'a pas eu besoin de joindre Pierre. Il a pu reprendre le travail rapidement.



Fatima

a. Personnas Fatima

Fatima





- Profil

- 55ans, vie seule avec son chien, tient un salon de coiffure
 - Sa fille vient souvent, elle sera là pour l'opération
- Elle travaille beaucoup; Elle est bénévole aux restos du cœur
- Une phrase clé : « *Jamais à la maison, toujours être actif* »

- Plaisir/ émotions, comportement

- *Le dimanche c'est sacré; c'est pour le bébé de ma fille*
- Elle aime, les échanges, les gens, les potins, les promenades
- Elle n'aime pas le vide, le silence, les clients exigeants
- Elle est prudente, organisée

Opération canal carpien

- Technologies

- Elle a un ordinateur au salon, mais c'est pas pratique
 - Elle préfère téléphoner qu'envoyer des textos
 - Elle s'est achetée une tablette pour les photos
- Elle a un site Web pour le salon
- Elle suit les blogueuses de la mode et de la coiffure

- Santé

- Elle a des rhumatismes mais elle prend soin d'elle
 - Elle fait très attention à son bien être
 - Crème, maquillage, sauna, esthétique...
- Elle fait attention à son régime
 - Mais elle aime bien le sucré
- Elle a besoin de comprendre tout de suite
 - Ex les analyses médicales
- L'opération prévue : canal carpien
- Elle ne s'intéresse pas à la santé connectée
- Elle a des connaissances faibles en santé
 - *Si on me dit de prendre des médicaments, je les prends*
- Motivations/ Smart Angel
 - C'est rassurant pour le suivi et les explications
 - Mais est-ce que je pourrais poser des questions?
 - Elle a confiance dans le corps médical

Figure 71. Personnas Fatima

b. Scénario Fatima

Texte audio (voix de femme)	Extrait des images
<p>Cela fait 35ans que FATIMA est coiffeuse. Elle a sa propre équipe et son propre salon.</p>	
<p>Malheureusement, comme beaucoup de coiffeurs, Fatima souffre du syndrome du canal carpien. Ce syndrome est fortement lié à son activité professionnelle. Elle a déjà subi plusieurs infiltrations mais ces traitements ne sont plus efficaces.</p>	
<p>La nuit, la douleur est devenue tellement forte qu'elle ne pouvait plus dormir.</p>	

Elle s'est donc rendue chez un médecin qui a poussé les analyses. Fatima n'a plus d'autres choix que de se faire opérer.



Elle se rend donc chez un chirurgien qui lui explique l'opération et ses suites.



Fatima est très angoissée. L'idée même de se faire opérer de la main lui fait terriblement peur.



Quelques jours plus tard, elle rencontre son anesthésiste qui lui propose de faire l'opération en ambulatoire mais qu'elle serait aidée par un dispositif du nom de *Smart Angel* à la maison. Fatima n'est pas très rassurée de ce système connecté mais si le médecin le dit, c'est que cela doit être une bonne chose. Elle accepte sans se poser de questions.



Elle voit ensuite une infirmière qui lui explique comment elle doit se laver et se raser avant l'opération.



Le jour de l'opération, Fatima est très stressée. Elle tremble pour enlever ses vêtements et se préparer à l'opération. Elle a vraiment hâte d'en finir.



L'opération se passe bien. Fatima est sous anesthésie locale.



A sa sortie au bloc, Fatima se sent mieux. L'anesthésie ne l'a pas rendu nauséuse. Elle n'a qu'une envie, rentrer chez elle au plus vite.

Peu de temps après sa sortie, un infirmier vient lui présenter le dispositif *Smart Angel*. Fatima essaye 3-4 fois avant de comprendre comment tout fonctionne. La tablette lui rappelle de prendre ses médicaments, alors elle se sent rassurer d'avoir cet outil.



Fatima a eu son autorisation de sortie. Elle rentre chez elle accompagnée de sa fille.



Une fois dans sa maison, Fatima s'allonge dans son lit et allume la tablette *Smart Angel*. Elle essaye de prendre ses mesures seule, mais sa fille vient l'aider pour enfiler le brassard du tensiomètre.



Toute la semaine, Fatima a suivi à la lettre les demandes de la tablette. Elle a été très satisfaite d'avoir ce petit appareil qui la raccroche à l'hôpital. Elle déclare : « au moins, ils savent que je vais bien, et ça, cela me rassure ». L'hôpital quant à lui, n'a pas eu besoin de contacter Fatima au vu des résultats satisfaisants transmis par le dispositif.



Patrick

a. Personnas Patrick

<p>Patrick</p>  <ul style="list-style-type: none"> - Profil <ul style="list-style-type: none"> - 70 ans, vit seul, à la retraite, diabétique - Il vit en appartement, en ville avec un petit chien - Une phrase clé : « <i>On ne dit pas tout</i> » - Plaisir/ émotions, comportement <ul style="list-style-type: none"> - Il aime être seul devant son ordinateur <ul style="list-style-type: none"> - Il aime bien recevoir quand même car il n'aime pas la solitude - Il croit au complotisme (surveillance par internet) - Il n'aime pas la contradiction, il ne fait pas de compromis - Il n'aime pas le sport; Il maniaque mais peu organisé <ul style="list-style-type: none"> - Je range mais à ma façon - Il est très exigeant avec les autres - Il dit qu'il ne prend pas de risque mais il en génère <div style="border: 1px solid black; padding: 5px; text-align: center; margin-top: 10px;">Opération Varice</div>	 <ul style="list-style-type: none"> - Technologies <ul style="list-style-type: none"> - Il est souvent sur son PC, il achète tout sur internet - il joue en réseau - Utilisation courante <ul style="list-style-type: none"> - un smart phone, un PC, des objets connectés - Tensiomètre, tablette, oxymètre... - Domotique, balance connectée, - Capteur de sommeil... - Mais il ne fait confiance à aucun appareil <ul style="list-style-type: none"> - Il croise les données lui-même - Santé <ul style="list-style-type: none"> - Il est diabétique, allergique à tout (chats, pollens) - Il est hypocondriaque; il s'inquiète vite <ul style="list-style-type: none"> - Il appelle sans arrêt son médecin - Il est souvent sur Doctissimo - Il aime « les médicaments qui guérissent tout » - Il n'aime pas les piqûres d'insuline, le sport - Il prend des risques : diabétique mange gras et sucré - L'opération prévue : varices - Il connaît très bien les objets connectés - Il est très informé sur le suivi théorique de sa santé <ul style="list-style-type: none"> - <i>Je sais ce qu'il faut faire, c'est moi qui décide</i> - Motivations/ Smart <ul style="list-style-type: none"> - Motivé mais il a déjà vu des systèmes similaires - Il sait comment sa marche - Il pourra comparer avec ce qu'il a déjà - Il risque de répondre faux à certaines questions
---	--

Figure 72. Personnas Patrick

b. Scénario Patrick

Texte audio (voix d'homme)	Extrait des images
<p>Patrick a arrêté de travailler il y a maintenant 7 ans. Il jouit d'une retraite tranquille. De nature gourmande il a beaucoup de mal à suivre les recommandations de son médecin et Patrick souffre de plus en plus de son diabète. Mais il a su s'équiper de technologie pour suivre sa santé. Malheureusement, son embonpoint lui a provoqué des douleurs liées à des varices.</p>	
<p>Patrick s'est alors rendu chez un généraliste qui l'a envoyé chez un chirurgien. Le chirurgien lui a donné les détails de cette petite opération et ses suites.</p>	

Patrick est très angoissée. L'idée même de se faire opérer lui fait terriblement peur. Il n'est pourtant pas à sa première opération, mais l'idée de retourner sur un lit d'hôpital l'effraye.



Quelques jours plus tard, il rencontre son anesthésiste qui lui propose de faire l'opération en ambulatoire mais qu'il serait aidé par un dispositif du nom de *Smart Angel* à la maison. Patrick n'est pas très rassuré de ce système connecté mais si le médecin le dit, c'est que cela doit être une bonne chose. Il accepte sans se poser de questions. Ce qu'il veut, c'est rentrer chez lui rapidement.



Il voit ensuite une infirmière qui lui explique comment il doit se laver et se raser avant l'opération.



Le jour de l'opération, Patrick est très stressé. Il tremble pour enlever ses vêtements et se prépare à l'opération. Il a vraiment hâte d'en finir.



L'opération s'est bien passée. A sa sortie au bloc, Patrick se sent mieux. L'anesthésie locale ne l'a pas rendu que faiblement nauséeux.



Peu de temps après sa sortie, un infirmier vient lui présenter le dispositif *Smart Angel*. Patrick utilise déjà des objets similaires pour son diabète, il comprend rapidement comment tout fonctionne.



Patrick a eu son autorisation de sortie. Sa nièce est venue le chercher.



Une fois dans sa maison, Patrick s'allonge dans son canapé et allume la tablette *Smart Angel*. Il prend ses mesures seul.



Toute la semaine, Patrick a suivi à la lettre les demandes de la tablette. Il a été très satisfait d'avoir ce petit appareil qui le raccroche à l'hôpital. Il espère ne pas remettre les pieds à l'hôpital avant un bon bout de temps. D'ailleurs, les résultats de ses mesures étant satisfaisants, le personnel hospitalier n'a pas eu besoin de l'appeler.



Annexe 9 : Questionnaire sur les connaissances antérieures en santé en français (40 items)

Vrai Faux Ne sais pas

1. Le cycle cardiaque

- a. Le cycle cardiaque comprend une contraction suivie d'un relâchement
- b. Le cycle cardiaque comprend une inspiration suivie d'une expiration
- c. La contraction cardiaque est appelée diastole
- d. Le cœur fonctionne comme une pompe qui propulse le sang vers les tissus

2. Pression artérielle

- a. La pression artérielle peut être définie comme la force exercée par le sang sur la paroi des artères.
- b. La pression artérielle peut être définie comme le volume de sang éjecté par le cœur.
- c. La pression artérielle peut être définie comme le nombre de battements du cœur par minute.
- d. La pression artérielle est synonyme de la tension artérielle.

3. Les valeurs de référence ou valeurs « normales »

- a. Les deux chiffres obtenus lors de la mesure de la tension correspondent à la valeur de référence suivie de la valeur obtenue par le patient
- b. Les valeurs de référence chez les sujets jeunes sont 140 / 90 mmHg
- c. Avec l'âge, la tension artérielle baisse.
- d. Plus l'artère est rigide et plus les deux chiffres sont élevés.

4. Variation de tension

- a. On parle d'hypertension quand les deux chiffres sont supérieurs aux valeurs de référence
- b. Le cerveau est le seul organe épargné dans le cas d'une hypertension de longue durée.
- c. Une hémorragie peut entraîner une hypotension
- d. L'hypotension peut entraîner la perte de connaissance

5. L'utilisation du tensiomètre

- a. Le tensiomètre se met généralement au bras
- b. Le tensiomètre nécessite l'application d'un gel
- c. L'utilisation du tensiomètre mesure l'activité électrique des artères
- d. Seul un personnel médical peut utiliser un tensiomètre.

9 questions faciles, 9 questions moyennes, 2 questions difficiles / 20 questions totales

Questions saturometre :

6. Le rôle des poumons :

- a. L'oxygène traverse les parois des alvéoles dans les poumons
 - b. Le temps de contact entre les alvéoles et l'hémoglobine permet la saturation des hématies en oxygène.
 - c. C'est le cœur qui permet la répartition de l'oxygène dans les tissus (organes, muscles...)
 - d. L'oxygène est transporté dans le sang grâce aux globules blancs
-

7. Transport des gaz respiratoires dans le corps

- a. Le dioxyde de carbone est éliminé par les reins
- b. Le sang riche en dioxyde de carbone et le sang riche en oxygène ne se rencontrent pas.
- c. Le sang circule en sens unique dans un système clos.
- d. Le dioxyde de carbone est produit par les tissus (organes, muscles...).

8. Saturomètre

- a. Le saturomètre mesure le taux d'oxygène dans le sang
- b. Permet de repérer les carences en oxygène dans les tissus
- c. Permet de repérer les anémies
- d. L'utilisation du saturomètre nécessite une prise de sang.

9. Le taux d'oxygène

- a. Le taux d'oxygène dans le sang a tendance à baisser avec l'âge
- b. Le taux d'oxygène dans le sang diffère selon le sexe
- c. Les fumeurs ont tendance à avoir une saturation en oxygène plus importante que celle des non-fumeurs.
- d. Une valeur inférieure de saturation en oxygène à 60% est signe d'hypoxémie.

10. l'utilisation du saturomètre

- a. Le saturomètre peut être positionné sur le lobe de l'oreille
 - b. L'utilisation du saturomètre nécessite l'application d'un gel
 - c. Seul un personnel médical peut utiliser un saturomètre.
 - d. Le vernis à ongle peut fausser les valeurs données par le saturomètre
-

Annexe 10 : Questionnaire sur les connaissances antérieures en santé en anglais (40 items)

		Don't
Blood pressure questions :	True	False know
<hr/>		
1. The heart cycle		
a. The cardiac cycle consists of a contraction followed by a relaxation.		
b. The cardiac cycle consists of one inhalation followed by one exhalation.		
c. The cardiac contraction is called diastole.		
d. The heart acts as a pump that propels blood to the tissues and tissues of the body.		
<hr/>		
2. Blood pressure		
a. Blood pressure can be defined as the force exerted by the blood on the walls of the arteries.		
b. Blood pressure can be defined as the volume of blood ejected by the heart.		
c. Blood pressure can be defined as the number of heart beats per minute.		
d. Blood pressure is synonymous with "arterial pressure".		
<hr/>		
3. The reference or "normal" values		
a. The two values obtained from the blood pressure measurement are the reference value followed by the value obtained by the patient.		
b. The reference values in young subjects are 140 / 90 mmHg		
c. With age, blood pressure drops.		
d. The more inflexible the artery, the higher the two numbers.		
<hr/>		
4. Blood pressure variation		
a. Hypertension is defined as when both numbers are above the reference values.		
b. The brain is the only organ spared in long-term hypertension		
c. A hemorrhage can lead to hypotension.		
d. Hypotension can lead to unconsciousness.		
<hr/>		
5. The use of the blood pressure monitor		
a. The monitor is usually placed on the arm		
b. The blood pressure monitor requires the application of a gel.		
c. The use of the blood pressure monitor measures the electrical activity of the arteries.		
d. Only medical personnel may use a blood pressure monitor.		
<hr/>		
Pulse Oximeter questions:		
<hr/>		
6. The role of the lungs :		
a. Oxygen passes through the walls of the alveoli in the lungs		
b. The contact time between the alveoli and the haemoglobin allows the oxygen saturation of the red blood cells.		
c. It is the heart that allows the distribution of oxygen in the tissues (organs, muscles...)		
d. Oxygen is transported in the blood by white blood cell		
<hr/>		
7. Transport of respiratory gases in the body		
a. Carbon dioxide is eliminated by the kidneys.		
b. Carbon dioxide-rich blood and oxygen-rich blood do not meet.		
c. Blood flows in a closed system in one direction only.		
<hr/>		

-
- d. Carbon dioxide is produced by tissues (organs, muscles, etc.).
-

8. Pulse oximeter

- a. The pulse oximeter measures the level of oxygen in the blood.
b. The pulse oximeter identifies oxygen deficiencies in tissues.
c. The pulse oximeter identifies anemias.
d. The use of the pulse oximeter requires a blood test.
-

9. Oxygen level

- a. Oxygen levels in the blood tend to decrease with age.
b. The level of oxygen in the blood differs according to gender
c. Smokers tend to have higher oxygen saturation than non-smokers.
d. A lower oxygen saturation value of 60% is a sign of hypoxemia.
-

10. the use of the pulse oximeter

- a. The pulse oximeter can be positioned on the earlobe.
b. The use of the hydrometer requires the application of a gel
c. Only medical personnel may use a pulse oximeter.
d. Nail polish can distort the values given by the pulse oximeter.
-

Annexe 11 : Détails des caractéristiques de l'utilisateur sur la base des mesures d'utilisabilité du tensiomètre et de l'oxymètre de pouls

Tableau 27. Détails des caractéristiques de l'utilisateur sur la base des mesures d'utilisabilité du tensiomètre.

Variable (N=137)	Efficacité (erreurs)	Efficience (secondes)	Satisfaction (SUS)
Caractéristiques sociodémographiques			
	Moyenne (ET)	,77 (1,41)	260,91 (107,12)
	N = 149		72,48 (16,23)
Age en années (Moyenne (ET) ;)	Spearman unilateral correlation	$r = -,054,$ $p = ,53$	$r = -,065, P = ,45$ $p = ,97$
	Homme (n = 37)	0,7 (1,56)	266,11 (106,64)
Genre	Femme (n = 100)	0,8 (1,36)	258,99 (109,72)
	Mann-Whitney	U = 1646 $p = ,26$	U = 1892 $p = ,84$
	1 ^{ère} année (n = 2)	0 (0)	156,14 (1,19)
	2 ^{ème} année (n = 80)	0,87 (1,55)	270,41 (112,2)
Niveau d'éducation	3 ^{ème} année (n = 55)	0,67 (1,21)	250,63 (100,1)
	Kruskal-Wallis	$\chi^2 = 2,43$ $p = ,49$	$\chi^2 = 3,16$ $p = ,37$
	Elevé (n = 47 ; 29.51 (2.46))	0,85 (1,61)	266,3 (122,3)
	Modéré (n = 59 ; 23.63 (1.74))	0,59 (0,97)	258,45 (112,68)
Expérience en technologie de l'information (TI)	Faible (n = 31; 24.39 (4.83))	1 (1,75)	257,43 (66,83)
	Kruskal-Wallis	$\chi^2 = 0,51$ $p = ,77$	$\chi^2 = 0,64$ $p = ,73$
			73,39 (16,93)
			$\chi^2 = 0,24$ $p = ,88$

Expérience dans l'utilisation de dispositifs médicaux				
Tensiomètre	Oui (n = 61)	0,66 (1,26)	250,81 (100,4)	72,54 (16,03)
	Non (n = 76)	0,87 (1,52)	269,02 (112,22)	72,43 (16,5)
	Mann-Whitney	U = 2490,5, <i>p</i> = ,39	U = 2529, <i>p</i> = ,36	U = 2364, <i>p</i> = ,84
Oxymètre de pouls	Oui (n = 19)	0,68 (0,75)	264,61 (96,73)	75,79 (15,34)
	Non (n = 118)	0,79 (1,49)	260,32 (109,07)	71,94 (16,37)
	Mann-Whitney	U = 1247,5, <i>p</i> = ,37	U = 1142, <i>p</i> = ,89	U = 1286, <i>p</i> = ,3

Tableau 28. Détails des caractéristiques de l'utilisateur sur la base des mesures d'utilisabilité de l'oxymètre de pouls.

Variable (N=147)	Efficacité (erreurs)	Effizienz (secondes)	Satisfaction (SUS)
Caractéristiques sociodémographiques			
Moyenne (ET)	0,99 (0,921)	158,42 (75,75)	75,75 (17,29)
Age en années (Moyenne (ET) ;)	N = 147		
Spearman unilateral correlation	<i>r</i> = ,068, <i>p</i> = ,42	<i>r</i> = - ,012, <i>p</i> = ,89	<i>r</i> = -,027, <i>p</i> = ,75
Homme (n = 41)	0,73 (0,77)	139,7 (62,5)	77,68 (14,65)
Femme (n = 106)	1,08 (0,96)	165,67 (79,38)	69,27 (17,72)
Mann-Whitney	U = 1686,5 <i>p</i> = ,02*	U = 1768 <i>p</i> = ,08	U = 2756 <i>p</i> = ,012*
1 ^{ère} année (n = 2)	2 (0)	144,2 (116,75)	67 (7,071)
Niveau d'éducation	2 ^{ème} année (n = 86)	1 (0,98)	149 (68,72) 73,62 (15,8)

	3 ^{ème} année (n = 59)	,93 (0,83)	172,22 (84)	68,47 (19,2)
	Kruskal-Wallis	$\chi^2= 4,37$ $p = ,22$	$\chi^2= 3,07$ $p = ,38$	$\chi^2= 4,61$ $p = ,2$
Expérience en technologie de l'information (TI)	Elevé (n = 46; 29,52 (2,49))	0,89 (0,8)	164,97 (85,83)	76,04 (17,2)
	Modéré (n =58; 23,59 (1,73))	,95 (1,02)	157,87 (71,2)	71,02 (16,02)
	Faible (n = 31; 18,1 (2,72))	2.16 (0,97)	174,28 (69,81)	69,1 (16,01)
	Kruskal-Wallis	$\chi^2= 1,83$ $p = ,4$	$\chi^2= 1,68$ $p = ,43$	$\chi^2= 6,036$ $p = ,049^*$
	Expérience dans l'utilisation de dispositifs médicaux			
Tensiomètre	Oui (n =62)	0,97 (0,77)	158,65 (86,93)	71,29 (18,17)
	Non (n = 85)	1 (1,02)	158,26 (66,98)	71,86 (16,73)
Oxymètre de pouls	Mann-Whitney	U = 2565,5, $p = ,76$	U = 2779, $p = ,57$	U = 2682, $p = ,85$
	Oui (n = 19)	0,95 (0,7)	179,01 (57,31)	77,53 (15,43)
	Non (n = 128)	0,99 (0,95)	155,37 (77,84)	70,74 (17,43)
	Mann-Whitney	U = 1255,5, $p = ,8$	U = 1526, $p = ,07$	U = 1501, $p = ,09$

Annexe 12 : Questionnaire de littératie en santé – le Newest Vital Sign

À LIRE AU PARTICIPANT : Ces renseignements figurent sur un contenant de crème glacée de 500ml.

1. Si vous mangez toute la crème glacée dans le contenant, combien de calories aurez-vous mangées?

Réponse: 1000 est la seule bonne réponse.

2. S'il vous est permis de manger 60 grammes de glucides comme collation, quelle portion de crème glacée pouvez-vous vous servir?

Réponse: Toutes les réponses suivantes sont correctes: 250ml ou 1 tasse (n'importe quelle quantité jusqu'à concurrence de 250ml ou 1 tasse), la moitié du contenant.

Remarque: Si le participant répond «deux portions», posez-lui la question suivante: «Combien de crème glacée cela représente-t-il si vous devez en mesurer la quantité dans un bol?»

3. Votre médecin vous recommande de réduire la quantité de gras saturés dans votre alimentation. Habituellement, vous mangez 42g de gras saturés chaque jour, ce qui comprend une portion de crème glacée. Si vous arrêtez de manger de la crème glacée, combien de grammes de gras saturés mangeriez-vous chaque jour?

Réponse: 33 est la seule bonne réponse.

4. Si vous mangez habituellement 2500 calories par jour, quel pourcentage de la valeur de vos calories quotidiennes mangerez-vous en en prenant une portion de crème glacée?

Réponse: 10% est la seule bonne réponse.

À LIRE AU PARTICIPANT : Imaginez que vous êtes allergique aux substances suivantes : à la pénicilline, aux arachides, aux gants de latex et aux piqûres d'abeilles.

5. Pouvez-vous manger cette crème glacée sans danger?

Réponse: Non

6. (Posez la question seulement si le participant a répondu Non à la question 5) Pourquoi pas?

Réponse: Parce qu'elle contient de l'huile d'arachide et des arachides.

Valeur nutritive	
Portion 1/2 tasse (125 mL) Portions par contenant 4	
Teneur par portion	% valeur quotidienne*
Calories 250	
Lipides 13 g	20 %
saturés 9.0 g	45 %
+ trans 0 g	
Cholestérol 30 mg	
Sodium 55 mg	2 %
Glucides 30 g	10 %
Fibres 0 g	0 %
Sucres 23 g	
Protéines 4 g	
Vitamine A 10 %	Vitamine C 0 %
Calcium 15 %	Fer 4 %

* Pourcentage de la valeur quotidienne selon un régime alimentaire de 2 000 Calories. Vos valeurs quotidiennes personnelles peuvent être plus ou moins élevées selon vos besoins énergétiques.

INGRÉDIENTS : Crème, lait écrémé, sucre liquide, eau, jaunes d'œuf, sucre brun, matière grasse du lait, huile d'arachide, sucre, beurre, sel, carraghénane, extrait de vanille

Figure 73. Capture d'écran de l'étiquette présenté aux participants

Annexe 13 : Questionnaire de littératie en santé – Le HLS-EU-Q16

Indiquez, sur une échelle de très facile à très difficile, dans quelle mesure il est facile pour vous de...

	Très facile	Facile	Difficile	Très difficile
... trouver des informations sur les traitements des maladies qui vous concernent ?				
... savoir où obtenir l'aide d'un professionnel quand vous êtes malade ?				
... comprendre ce que vous dit votre médecin ?				
... comprendre les consignes de votre médecin ou de votre pharmacien sur la manière de prendre vos médicaments ?				
... savoir quand il serait utile d'avoir l'avis d'un autre médecin ?				
... utiliser les informations que le médecin vous donne pour prendre des décisions concernant votre maladie ?				
... suivre les consignes de votre médecin ou de votre pharmacien ?				
... trouver des informations sur comment faire en cas de problèmes psychologiques comme le stress ou la dépression ?				
... comprendre les mises en gardes concernant l'impact sur la santé de certains comportements comme fumer, ne pas faire assez d'exercices et boire trop ?				
... comprendre l'intérêt des dépistages ?				
... évaluer la fiabilité des informations disponibles dans les médias sur ce qui est dangereux pour la santé ?				
... savoir comment vous préserver des maladies à partir des informations disponibles dans les médias ?				
... vous renseigner sur les activités bénéfiques pour votre bien être psychologique ?				
... comprendre les conseils de votre famille ou de vos amis en matière de santé ?				
... comprendre les informations disponibles dans les médias pour être en meilleure santé ?				
... identifier quels sont les comportements de votre vie de tous les jours qui ont un impact sur votre santé ?				

Annexe 14 : Questionnaire caractéristiques sociodémographiques études 2 et 3

14.1. Caractéristiques individuelles

1. Age :ans

2. Sexe :
 - Homme
 - Femme
 - Autre

3. Niveau d'étude :
 - Primaire
 - Collège (6^{ème}, 5^{ème}, 4^{ème}, 3^{ème})
 - Technique court (CAP, BEP)
 - Secondaire (seconde, 1^{ère}, Terminale), niveau BAC ou Brevet professionnel
 - Techniques supérieur (IUT, BTS)
 - Supérieur 1^{er} cycle (Bac +1 ou 2)
 - Supérieur 2^{ème} cycle (Bac +3 ou 4)
 - Supérieur 3^{ème} cycle (Bac +5 et plus)

4. Profession :
 - Agriculteurs
 - Artisans, commerçants, chefs d'entreprise
 - Cadres, professions libérales
 - Professions intermédiaires
 - Employés
 - Ouvriers
 - Retraités
 - Sans emploi
 - Etudiant
 - Autre :

5. Vous vivez en zone :
 - Rural
 - Semi-urbain
 - Urbain

6. Avez-vous déjà été opéré dans votre vie ? (Dent de sagesse, coloscopie, implantation de broche...)
- Oui
 - Non
 - Je ne sais pas

7. Avez-vous déjà subi une opération en chirurgie ambulatoire (sans hospitalisation)?
- Oui
 - Non
 - Je ne sais pas

14.2. Utilisation des outils numériques

1. Avez-vous des dispositifs médicaux à votre domicile ? (Tensiomètre, glucomètre...)
- Oui
 - Non
2. Avez-vous déjà pris votre tension vous-même ?
- Oui
 - Non
3. Avez-vous déjà pris votre taux d'oxygène vous-même ?
- Oui
 - Non

4. Lesquelles de ces technologies utilisez-vous et à quelle fréquence ?

	Jamais	Peu	Parfois	Souvent	Très souvent
Téléphone	<input type="radio"/>				
Tablette	<input type="radio"/>				
Ordinateur	<input type="radio"/>				
Objet(s) connecté(s) (balance, montre, bracelet....)	<input type="radio"/>				
Dispositifs médicaux (tensiomètre, glucomètre)	<input type="radio"/>				

5. Dans quelle mesure êtes-vous à l'aise quant à l'utilisation des appareils précités

	Pas du tout à l'aise	Un peu à l'aise	Moyennement à l'aise	Assez à l'aise	Tout à fait à l'aise
Téléphone	<input type="radio"/>				
Tablette	<input type="radio"/>				
Ordinateur	<input type="radio"/>				
Objet(s) connecté(s) (balance, montre, bracelet....)	<input type="radio"/>				
Dispositifs médicaux (tensiomètre, glucomètre)	<input type="radio"/>				

Annexe 15 : Guide d'entretien pour l'étude 2

Tâches	Retranscription <i>Think aloud</i>	Annotations gestuelles (e.g., montre le tensiomètre)
Ouverture du matériel et lancement de l'application		
Lancement du rdv		
Utilisation du tensiomètre		
Utilisation du saturomètre		
Réponse aux questions		
Reprise de la saturometre		
Remarque		
Remarque au SUS :		
Thématique	Guide d'Entretien	
<i>Perception dispositif</i>	du	<u><i>Perception positive</i></u> Qu'avez-vous apprécié dans le matériel ? Dans son utilisation ? Dans son utilité ? <u><i>Perception négative</i></u> Il y a-t-il des éléments qui vous freinent à utiliser ce matériel ? Éléments rédhibitoires ? Peu confortable ?
Acceptation dispositif	du	Si vous vous retrouviez dans une réelle situation de chirurgie ambulatoire, comme le PERSONA X, utiliseriez-vous le dispositif ? accepteriez-vous le matériel chez vous ? Pourquoi ? Utilisation 3 fois par jour après opération, l'utiliseriez-vous ?
erreurs		Avez-vous eu l'impression d'avoir commis des erreurs de manipulation. Pourquoi ? Avez-vous eu un doute lors de l'utilisation du matériel ? Quand et pourquoi ? Avez-vous compris vos mesures prises ? Pourquoi ?
Recommandation de conception		Est-ce qu'il y a des choses à améliorer dans le dispositif ? IL y a-t-il des choses qui vous ont manqué ? Que vous n'avez pas compris ?
Remarque autres :		

Annexe 16 : Scénario Anesthésiste et rappel du déroulement d'une consultation pré-opératoire

FICHE ANESTHESISTE

Matériel : tensiomètre avec stéthoscope

Bonjour,

Pour commencer, toute l'équipe vous remercie pour votre participation à cette étude.

Nous allons vous demander de vous mettre dans le rôle d'un(e) anesthésiste. Vous allez devoir suivre le déroulement d'une consultation préopératoire. Tous les patients que vous devrez examiner souffrent d'une hernie inguinale et devront subir une opération.

A l'issue de votre consultation, vous devrez

- Déterminer quel type d'anesthésie vous pourrez réaliser sur votre patient
- Convaincre le patient de l'utilisation d'une aide post-chirurgicale : le dispositif *Smart Angel*.

Afin de réaliser votre consultation anesthésique, prenez connaissance des questions et manipulations que vous allez devoir réaliser.

Bonnes consultations

Rappel : déroulement d'une consultation d'anesthésie Pré-opératoire

La **consultation pré-anesthésique** (ou **consultation d'anesthésie**) est obligatoire avant toute opération ou examen nécessitant une **anesthésie** quel que soit son type. Elle permet d'évaluer les risques, de définir la technique d'**anesthésie** qui sera mise en place et de vous donner toutes les informations nécessaires.

Voici le déroulement de la consultation anesthésique :

- Prenez la feuille « *questionnaire et document pré-anesthésique* » apportée par le patient
- Complétez avec le patient « *le questionnaire et document pré-anesthésique* » :
 - o Vérifiez les antécédents chirurgicaux du patient
 - o La santé générale du patient (allergies, problème cardiaque ou respiratoire...)
 - Prenez la tension du patient et vérifiez le cœur
 - Son IMC
 - Sa prise de médicament
- Définissez le score ASA du patient (à l'aide du schéma ci-dessous) et **notez-le sur la feuille questionnaire du patient.**
- Expliquez comment va se passer l'anesthésie
- Expliquez ce que doit faire le patient avant l'anesthésie
 - o À jeun depuis au moins 6 heures (nourriture et eau)
 - o Ne pas fumer le jour de l'intervention
 - o Venez sans maquillage, sans vernis, sans bijoux, sans piercing et sans objets de valeur
- Expliquez ce que doit faire le patient après l'anesthésie
 - o Ne pas consommer d'alcool
 - o Ne pas faire d'activité sportive
- Proposez le dispositif *Smart Angel* à votre patient et expliquez-lui les intérêts pour son suivi post-chirurgical ; essayer de le convaincre d'accepter le dispositif.

Classification de l'état de santé du patient	
Score	État de santé du patient
1	Patient sain, en bonne santé, c'est-à-dire sans atteinte organique, physiologique, biochimique ou psychique.
2	Maladie systémique légère, patient présentant une atteinte modérée d'une grande fonction, par exemple : légère hypertension, anémie, bronchite chronique légère.
3	Maladie systémique sévère ou invalidante, patient présentant une atteinte sévère d'une grande fonction qui n'entraîne pas d'incapacité, par exemple : angine de poitrine modérée, diabète, hypertension grave, décompensation cardiaque débutante.
4	Patient présentant une atteinte sévère d'une grande fonction, invalidante, et qui met en jeu le pronostic vital, par exemple : angine de poitrine au repos, insuffisance systémique prononcée (pulmonaire, rénale, hépatique, cardiaque...)
5	Patient moribond dont l'espérance de vie ne dépasse pas 24 heures, sans intervention chirurgicale. (« Patient moribond dont la survie est improbable sans l'intervention » : source SFAR)
6	Patient en état de mort cérébrale, candidat potentiel au don d'organes (selon critères spécifiques)

Annexe 17 : Scénario Infirmier formateur du dispositif *Smart Angel*

Bonjour,

Pour commencer tout l'équipe vous remercie pour votre participation à cette étude.

Vous allez devoir jouer le rôle d'infirmier/ère formateur dans un service de chirurgie ambulatoire. Vos patients se seront tous fait opérer d'une hernie inguinale. Vous les retrouverez après leur opération pour leur présenter le dispositif *Smart Angel* et les éduquer à leur sortie post-ambulatoire.

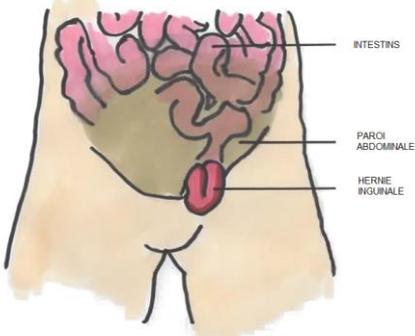
Nous comptons sur vous pour leur diffuser les informations que l'équipe technique vous aura présenté au préalable.

Bon suivi !

Annexe 18 : Scénario Patient/Participant

Tout d'abord, toute l'équipe vous remercie de votre participation pour cette étude. Nous tenons à préciser qu'aucun traitement invasif n'est réalisé pendant l'étude. Les données récoltées sont anonymes et utilisées uniquement dans le cadre de la recherche.

Nous nous intéressons à la relation patient-soignant. Laissez-vous porter par le personnel hospitalier comme si vous vous rendiez à une chirurgie.

<p>³⁹ Cela fait plusieurs jours que vous vous plaignez d'une douleur dans le bas ventre, proche de la cuisse gauche. Lorsque vous passez votre main dans cette zone, vous pouvez même sentir une grosseur qui prend le plus en plus de place.</p> 	<p>Vous décidez de prendre rendez-vous chez votre généraliste qui vous diagnostique une hernie inguinale.</p> 
<p>La hernie inguinale est un déplacement des viscères qui traversent la paroi abdominale au niveau de l'aîne et plus précisément au niveau du canal inguinal.</p> 	<p>Votre médecin traitant vous oriente vers un chirurgien digestif et celui-ci décide de vous opérer.</p> <p>Il vous explique le déroulement de l'opération.</p> <p>Le traitement de la hernie inguinale est chirurgical et consiste à replacer son contenu dans l'abdomen et à réparer la paroi abdominale. Vous fixez la date de votre chirurgie avec votre chirurgien qui se pratiquera sous anesthésie générale.</p>
<p>Les complications post-chirurgical de la hernie inguinale sont rares, mais des infections de la cicatrice ou une hémorragie abdominale peuvent survenir et être extrêmement grave. Il est donc important de signaler tous symptômes anormaux.</p>	<p>Aujourd'hui, vous allez vous rendre chez votre anesthésiste pour vérifier votre état de santé.</p>

³⁹ Il s'agit de la version pour les participantes. Nous proposons le même scénario avec des dessins d'hommes pour les participants masculins.

Annexe 19 : Guide d'entretien et focus group pour l'étude 3

Les questions sont à poser aux deux interviewés. Veillez au bon partage de la parole.

Perception du dispositif	Qu'avez-vous pensé du dispositif ? Avez-vous eu des difficultés dans son utilisation ? Avec le tensiomètre ? Avec le saturomètre ? Avec la tablette ?
Acceptation du dispositif	Trouvez-vous que les explications que vous a données l'anesthésiste étaient suffisantes ? Avez-vous accepté spontanément le dispositif ou a-t-il eu besoin de vous convaincre ? Au final, avez-vous accepté le dispositif ? Seriez-vous prêt à l'utiliser chez vous pendant une semaine, à raison de 3 fois par jour si un jour vous vous faites opérer en chirurgie ambulatoire ? Il y a-t-il des éléments du dispositif qui pourraient être un frein pour vous à son utilisation ?
Formation	Avez-vous trouvé que la formation proposée était suffisante ? <i>Du point de vue de l'infirmier</i> <i>Du point de vue du patient</i> Comment pourriez-vous améliorer cette formation ? Avez-vous manqué de temps ? de matériel ? d'explication ? de formation ?
Recommandation de conception	Que pensez-vous des explications de la tablette ? Est-elle suffisante ? Que pensez-vous des objets connectés ? Pensez-vous que vous puissiez un jour utiliser des objets comme cela à l'avenir ? Avez-vous d'autres remarques à ajouter ?
Mise en situation	Qu'avez-vous pensé de la mise en situation ? Avez-vous eu l'impression de vous retrouver dans un vrai hôpital ? Avez-vous eu peur, étiez-vous stressé ?

Annexe 20 : Grille de notation des erreurs de manipulation du dispositif

Système	Erreurs	Type d'erreurs / conséquences	Variation de la mesure ⁴⁰	Description
Tensiomètre	Le tensiomètre est à l'envers	Erreurs non-bloquantes / mesures erronées	Modérée/Importante	Le bouton Start-Stop doit se trouver sur le dessus de l'appareil en direction de la main ;
	Le tensiomètre est dans le mauvais sens par rapport à la paume de la main	Erreurs non-bloquantes / mesures erronées	Importante	Les capteurs du tensiomètre doivent être positionnés sur l'intérieur du poignet et non sur l'extérieur. En cas de mauvaise position, le tensiomètre peut continuer à tourner longtemps avant de trouver une tension et peut blesser le patient lié à un serrage trop important.
	Ne positionne pas l'avant-bras correctement – mauvaise inclinaison	Erreurs non-bloquantes / mesures erronées	Modérée	Pendant la prise de mesure, le tensiomètre doit être positionner exactement à la hauteur du cœur du participant. Pour cela, le patient doit positionner son coude sur une table et incliner son avant-bras de manière à surélevé suffisamment son poignet à la hauteur de son cœur.
	Bouge pendant la prise de mesure	Erreurs non-bloquantes / mesures erronées	Modérée/importante	L'utilisateur fait des mouvements avec son bras et avant-bras (i.e., inclinaison de l'avant-bras abaissée ou surélevée) pendant toute la durée de la prise de mesure
	Ne connecte pas le Bluetooth du tensiomètre à la tablette	Erreurs bloquantes / pas de mesure, abandons	-	L'utilisateur passe l'étape sur la tablette où il doit connecter le matériel ; Et/ou l'utilisateur n'allume pas le tensiomètre mais appuie sur le bouton « connecté » de la tablette ; Et/ou l'utilisateur n'attend pas la fin du chargement du Bluetooth.
	Positionne le brassard sur le poignet droit	Erreurs non-bloquantes / mesures erronées	Modérée	Et/ou l'utilisateur positionne le tensiomètre sur la main droite et non sur la gauche ;

⁴⁰ Une variation « importante » de la mesure entraîne une mesure de tension improbable (à partir de +/- 5 points de tension par rapport au référentiel de base) ;
 Une variation « modérée » de la mesure entraîne une mesure proche de la mesure originelle (entre 3 à 4 points de +/- de tension par rapport au référentiel de base) ;
 Une variation « faible » de la mesure entraîne une mesure très proche de la mesure originelle (1 à 2 points de +/- de tension par rapport au référentiel de base).

	Mauvais serrage – ne sert pas suffisamment le brassard à son poignet	Erreurs non-bloquantes / mesures erronées voire Erreurs bloquantes / pas de mesure, abandons	Importante		Et/ou l'utilisateur serre mal le brassard du tensiomètre à son bras. Lors d'une utilisation correcte, il n'est pas possible d'insérer un doigt entre le brassage et le poignet de l'utilisateur. En cas de mauvais serrage, le tensiomètre peut gonfler dans le vide, et/ou se décrocher de lui-même. Afin de garantir une bonne inclinaison et de ne pas biaiser la prise de tension, il est fortement recommandé de positionner son coude gauche (sur lequel est positionné le brassard) sur un support plat (e.g., table, bureau...)
	Ne positionne pas le coude sur un support	Erreurs non-bloquantes / mesures erronées	Faible		L'utilisateur doit décontracter sa main sur lequel est posé le brassard pendant la prise de mesure. Le poignet du participant doit être dénudé
	Sers le poing pendant la prise de mesure	Erreurs non-bloquantes / mesures erronées	Modéré		L'utilisateur doit décontracter sa main sur lequel est posé le brassard pendant la prise de mesure.
	Pose le tensiomètre par-dessus son vêtement	Erreurs non-bloquantes / mesures erronées	Importante		Le poignet du participant doit être dénudé
	Reste trop longtemps appuyé sur le bouton start/stop pour prendre la mesure	Erreurs bloquantes / pas de mesure, abandons	-		Afin de lancer la mesure, l'utilisateur doit appuyer une fois sur le bouton Start/stop. Puis la mesure doit se lancer. Si ce dernier reste appuyé, le tensiomètre s'éteint. Il faut alors recommencer l'opération.
	Arrêt subitement pendant la prise de mesure	Erreurs bloquantes / pas de mesure, abandons	-		L'utilisateur appuie sur le bouton Start/stop pendant la prise de mesure. Cela stop immédiatement le gonflement du tensiomètre. Et/ ou L'utilisateur n'attend pas la fin du chargement de la prise de mesure. Il retire le brassard pendant le chargement.
	Ne positionne pas l'oxymètre dans le bon sens	Erreurs non-bloquantes / mesures erronées	Faible		L'utilisateur ne positionne pas le capteur du tensiomètre (comprenant le bouton d'allumage) dans le sens de la paume de la main.
	N'appuie pas sur le bouton pour lancer la mesure	Erreurs bloquantes / pas de mesure, abandons	-		L'utilisateur ne réussit pas à appuyer sur le bouton situé sur le cadran de l'appareil pour lancer la prise de mesure.
	N'insert pas son doigt jusqu'au capteur ;	Erreurs bloquantes / pas de mesure, abandons	-		L'utilisateur ne met pas son doigt sur le capteur. Cela nécessite de mettre le doigt le plus au fond possible de l'appareil. Si ce n'est pas le cas, la phalange du doigt peut être plus visible.
Oxymètre de pouls	Retire l'oxymètre trop tôt pendant la prise de mesure ;	Erreurs bloquantes / pas de mesure, abandons	-		L'utilisateur n'attend pas la fin du chargement de l'oxymètre en cours de mesure. Ce chargement est visible sur la tablette. Si ce dernier retire de son doigt l'oxymètre afin que la tablette n'annonce par elle-même les résultats, alors le participant a retiré l'oxymètre trop tôt.
	Ne connecte pas le Bluetooth de l'oxymètre à la tablette.	Erreurs bloquantes / pas de mesure, abandons	-		L'utilisateur passe l'étape sur la tablette où il doit connecter l'oxymètre ; Et/ou l'utilisateur n'allume pas l'oxymètre de pouls mais appuie sur le bouton « connecté » de la tablette ;

						Et/ou l'utilisateur n'attend pas la fin du chargement du Bluetooth.
Tablette	Ne rentre pas dans le « rendez-vous » de l'application.		-		-	L'utilisateur réalise des mesures spontanées sans passer par le rendez-vous demandé.
	Déverrouillage de la tablette	la	Erreurs bloquantes / pas de mesure, abandons		-	L'utilisateur ne fait pas glisser son doigt sur la tablette pour la déverrouiller et chercher pendant plus de 3min un moyen de déverrouiller la tablette.

Annexe 21 : Nombre d'erreurs de manipulation des patients/participants lors de l'étude 3 en fonction de la grille de codage des erreurs (cf. annexe 20) et des trois essais demandés

Systeme	Erreurs de manipulation réalisées par les patients/participants	Type d'erreurs / conséquences	Variation de la mesure	Nombre d'utilisateur réalisant l'erreurs	1 ^{ER} essai	2 ^{EME} essai	3 ^{EME} essai
Tensiomètre	Mauvaise inclinaison	Erreurs non-bloquantes / mesures erronées	Modérée	6	5	2	6
	Bouge pendant la prise de mesure	Erreurs non-bloquantes / mesures erronées	Modérée/importante	7	4	4	4
	Positionne le brassard sur le poignet droit	Erreurs non-bloquantes / mesures erronées	Modérée	11	10	10	8
	Ne positionne pas le coude sur un support	Erreurs non-bloquantes / mesures erronées	Faible	9	5	4	3
	Mauvais serrage (erreur bloquante)	Erreurs bloquantes / pas de mesure, abandons	-	2	1	1	1
	Mauvais sens par rapport à la paume de la main	Erreurs non-bloquantes / mesures erronées	Important	1	1	1	0
	Pose le tensiomètre par-dessus son vêtement	Erreurs non-bloquantes / mesures erronées	Important	1	1	1	1
	Reste trop longtemps appuyé sur le bouton start/stop pour prendre la mesure	Erreurs bloquantes / pas de mesure, abandons	-	1	1	0	0
Oxymètre de pouls	Arrêt subitement pendant la prise de mesure	Erreurs bloquantes / pas de mesure, abandons	-	3	3	0	1
	Ne positionne pas l'oxymètre dans le bon sens	Erreurs non-bloquantes / mesures erronées	Faible	25	24	22	23
Tablette	N'insert pas son doigt jusqu'au capteur	Erreurs bloquantes / pas de mesure, abandons	-	1	1	0	1
	Le déverrouillage de la tablette	Erreurs bloquantes / pas de mesure, abandons	-	4	4	0	0
	Ne rentre pas dans le rendez-vous	-	-	3	3	3	3

Annexe 22 : Restitution des recommandations avec l'entreprise Evolucare

Recommandations de conception

Étu des	Recommandations	Réponses Evolucare	
Étude 1 –tensiomètre et oxymètre	Améliorer la compréhension	1.L'implantation de rappel de connaissances sur le fonctionnement de la tension et de l'oxygénation (idem étude 3)	Partie commerciale Sera présente dans la FAQ avec une notice dans la mallette et numérique
	Aide à la navigation dans la tablette	2.L'utilisation de format audio – accompagnement oral (notice obligatoire pour manipulation du matériel)	Petit tutoriel Vidéo
	Choix du matériel	3.Tensiomètre particulièrement complexe à utiliser – cependant, celui sans affichage montre davantage de mesures erronées. Nous recommandons donc celui avec affichage (flèches lumineuses sur le boîtier pour indiquer l'inclinaison de l'avant-bras)	Étude clinique -> CHU Nîmes Sinon objets connectés variés pour la version commerciale
Étude 2 – prototype VI	Aide à la navigation dans la tablette	4.Limiter la précipitation ⁴¹ : aborder les consignes point par point (page par page) – griser les étapes réalisées et les futures, mais mettre en couleur celles en cours	+ Fragmentation + pause pour le participant
		5.Schéma trop petit (pouvoir agrandir les schémas et ajouter des légendes)	Agrandi pour l'étude au simusanté – impact sur le design visuel Agrandir pour version commerciale « <i>on va voir ce que ça va donner</i> »
		6.Manque d'indications : - Serrage - Éteindre le matériel	Serrage : Intégré – chose importante Le matériel s'éteint seul.
		7.Ajout d'un lien hypertexte pour chaque étape	Coupler à d'autres fonctionnalités (inclus sur la version commerciale)
		8.Le participant s'attend à un feedback de l'appareil sur son utilisation incorrecte – prévention des dangers.	Seuil de mesure permet de vérifier la bonne utilisation Prévention -> non prise en compte
		9.Manque de feedback sur le Bluetooth – Bluetooth non visible	Pris en compte pour l'étude 3 -> création de pauses – bien intégrée
	Améliorer la compréhension	10. Notice papier glisser dans la mallette : - Explications sur le déverrouillage de la tablette - Explications sur le fait que les notices et les prises de commandes des objets proviennent de la tablette.	Prise en compte pour l'étude clinique : explication sur le déverrouillage
Limiter la frustration	11.Relance pour la recharge de la batterie	(important)	

⁴¹ Rappel : La première utilisation est cruciale pour la suite

			À faire pendant le questionnaire ou mesures spontanées
	Aide à la navigation dans la tablette	12. Changement de la taille de la police	À faire pour plus tard. Proposer un mode de changement manuel de la taille de la police.
	Limiter la frustration	13. Respect de l'hygiène	Procédure de nettoyage Étude Clinique -> CHU Pour version commerciale – important pour CE
		14. Présence d'un service technique. Exemple : que faire en cas de panne ?	Adapté à l'EC (idem pour perte de mdp)
	Aide à la navigation	15. Retour en arrière (idem étude 3)	Non Contrainte technique – refaire le formulaire
	Améliorer la compréhension	16. Terme « rdv » (idem étude 3)	« <i>on sait que c'est un problème</i> » et on veut le prendre en compte
	Aide à la navigation	17. Lors du rdv, griser les mesures spontanées (idem étude 3)	Déjà pris en compte
	Formation	18. Laisser le patient s'entraîner avant un retour au domicile	rendez-vous de démonstration – checklist pour le pro
		19. Penser à former les participants sur les autres onglets	Pas d'info dessus.
	Informations supplémentaires - tensiomètre	20. Précision notice : Quelles mains pour le tensiomètre ?	GIF – main rose et on fonce les lignes de la main (couleur chair et objet en blanc)
	Améliorer la compréhension	21. Dépliant pour présenter le dispositif (avec photos et explications d'à quoi servent les objets connectés) lors de la consultation anesthésique	Vidéo Infirmière de recherche fait la première démonstration Triptyque ?
	Informations supplémentaires - tensiomètre	22. Hiérarchisation des informations. Donner plus d'importance à certains éléments (exemple : mettre le tensiomètre au niveau du cœur)	il faut faire un atelier tensiomètre.
		23. Manque des informations : ne pas serrer le poing	Non – mettre avec le détail (consigne plus précise pour la première fois)
	Améliorer la compréhension Tablette vs objet connecté	24. On ne traite pas de la même manière la tablette (influencé par le niveau de technophilie et l'âge) et les DM (objet connecté) (non influencé par le niveau de technophilie et l'âge) – notice sur l'utilisation de la tablette (idem étude 2)	Notice papier dans la mallette
	Apprentissage du matériel	25. Un programme spécifique au démarrage	Petit tutoriel. « <i>on sait qu'il y en aura un, mais on ne sait pas quelle forme il aura</i> » Bcp de temps à mettre en place donc pas prioritaire. Première étape : ton premier questionnaire qui sera plus complet. et sur la démo ?
	Limiter la frustration	26. Explications sur la charge des objets connectés	Idem étude 2 Information du quotidien – visuel en recto verso qui se trouve dans la mallette
	Informations supplémentaires	27. L'ajout d'un historique	Version bonus Récap à la fin.
		28. Aide sur la navigation dans la liste des contacts : qui appeler ?	Complicé ! « <i>On sait que c'est important les contacts</i> » qui je dois contacter ? Écran de paramétrage et affichage des horaires par rapport aux services

Étude 3 - prototype V2

			Sinon une possible fiche contact dynamique
		29. Feedback sur l'envoi des données	Pas encore de feedback , « on va l'intégrer, c'est important » et à qui c'est envoyé
	Recommandations secondaires	30. Messagerie instantanée	Casse-tête (équipe derrière qui gère ça) ça ne sera pas mis en place de suite. on va vers ça, mais pas
		31. Application sur téléphone	1 ^{er} version pour téléphone. Mais pas de certification. Paramétrage complexe avec le Bluetooth. freins légaux et
		32. L'ajout de la température	Plus haut
		33. Mode daltonisme	« il faudrait faire du design accessible » Priorité : pas forcément
		34. Fonction zoom pour les légendes et le texte	Légende : non (moins de texte avec plus d'illustration)

Annexe A : Liste des publications

Articles Scientifiques à comité de lecture

Chaniaud, N., Metayer, N., Megalakaki, O., Loup-Escande, E. (2020). Effect of prior health knowledge on Usability: Usability Study. *JMIR Mhealth Uhealth*, 8(9). [10.2196/17983](https://doi.org/10.2196/17983)

Chaniaud, N., Metayer, N., Loup-Escande, E., & Megalakaki, O. (2021). Effects of mode presentation on the execution of Home Medical Device instructions: the case of a blood pressure monitor. *Cognitive Processing*. [10.1007/s10339-020-01006-1](https://doi.org/10.1007/s10339-020-01006-1)

Chaniaud, N., Megalakaki, O., Capo, S., & Loup-Escande, E. (Under review). Effects of user characteristics on the usability of a home connected medical device named Smart Angel: Usability Study (*Revue JMIR Human Factors*)

Loup-Escande, E., Burkhardt, J.-M., Christmann, O., **Chaniaud, N** & Richir, S. (2020). Usefulness and needs construction process in innovative software: an exploratory study of designers' viewpoints. *International Journal of Virtual Reality*, 20(1), 48-71. <https://doi.org/10.20870/IJVR.2020.20.1.2924>

Chevallier, T., Buzancais, G., Ocean, B., Rataboul, P., Boisson, C., Simon, N., Maseguin, C., **Chaniaud, N.**, Vialles, N., Lefrant, J.-L., Cuvillon, P. (2020). Feasibility of remote digital monitoring using wireless blue-tooth monitors, app and original web platform in patient following outpatient surgery: a prospective observational study - The SMART ANGEL 1 Study. *BMC Anesthesiology*, 20:259. <https://doi.org/10.1186/s12871-020-01178-5>

Conférences Internationales

Chaniaud, N., Loup-Escande, E., Megalakaki, O. (2018). Design, understanding and usability evaluation of connected devices in the field of health: contribution of cognitive psychology and ergonomics. IEA 2018.

Chaniaud, N., Megalakaki, O, Loup-Escande, E., Marty, J.-L., Cuvillon, P., Boisson, C., Ricquebourg, V. (2018). Dispositifs technologiques en santé : quelle contribution de la psychologie cognitive ergonomique ? Journée d'étude AFPSA « Communication psychologie de la santé et outils numériques : Applications, perspectives et éthique » (Liège).

Conférences Nationales

Chaniaud, N., Loup-Escande, E., & Megalakaki, O. (2020). Effets de la littératie en santé sur l'efficacité et l'efficacité de deux dispositifs médicaux connectés. AFPSA 2020. (Présentation orale)

Chaniaud, N., Loup-Escande, E., Metayer, N., Megalakaki, O. (2019). Evaluation de l'utilisabilité des dispositifs médicaux à domicile : Exemple d'un tensiomètre connecté. EPIQUE 2019 (Communication affichée).

Loup-Escande, E., Valentin, A., Metayer, N., Ricquebourg, V., Bidart, A., Damay, P., Boisson, C., Cuvillon, P. & **Chaniaud, N.** (2019, juillet). Adaptation de la démarche ergonomique en phase amont de la conception de technologies e-santé : un retour d'expérience. EPIQUE, 10^{ème} conférence de psychologie ergonomique (Lyon, France), 9-12 Juillet.

Chaniaud, N., Megalakaki, O., Metayer, N., Loup-Escande, E. (2018). Compréhension de dispositifs médicaux dans le cas de la chirurgie ambulatoire. SFP 2018. (Présentation orale)

Chaniaud, N., Loup-Escande, E., Megalakaki, O. (2018). Conception, compréhension et évaluation d'un dispositif technologique en santé : quelques pistes de réflexion. Doctoriales ARPEGE.

Chaniaud, N., Jean, C. (2017) Open Innovation : Une nouvelle méthode pour sélectionner de jeunes partenaires innovants. CONFER 2017 (Séville).

Comité de protection des personnes (CPP) pour lancement d'étude clinique multicentrique

Cuvillon, P., Boisson, C., Buzancais, G., Lefrant, J.-Y, Chaumeron, A., Candela, D., Gache, A., Nucci, B., Aya, G., Morau, D., Thibault, S., Toumi, M., Chanques, G., Jaber, S., Capdevila, X., Diaz, J., Zoric, L., Lorne, E., Raux, M., Reubrecht, V., Lesache, F., Langeron, O., Nouette, K., Lebuffé, G., Garrot, M., Ruinet, O., Molliex, S., Piriou, V., Beaussier, M., Leone, M., Boyadjiev, L., Lesache, F., Loup-Escande, E., & **Chaniaud, N.** (2018) Smart Angel 2 : Evaluation de l'impact du dispositif SMART ANGEL sur le suivi du patient à domicile après une chirurgie ambulatoire majeure ou intermédiaire : étude contrôlée randomisée en ouvert. 2018-A03212-53

Reviewer

Reviewer pour *Journal of Medical Internet Research* (JMIR)

Reviewer pour la *Société d'Ergonomie de Langue Française* (SELF)

Vulgarisation

Lauréate (1^{er} prix du jury) à MT180 Régionale des Hauts-de-France (2020)

Pitch Thesis – Forum des sciences cognitives (2020)

Ma thèse en récit – BD – finaliste pour le festival de Bande Dessin d'Angoulême (2019)

C'est quoi cette thèse ? – RJCE (2019)

mme michu

sort de chirurgie ambulatoire



