



HAL
open science

Study of the benefits of tangible user interfaces applied to teaching scheduling to french nursing students using a serious game simulating a hospital ward

Benoît Landi

► To cite this version:

Benoît Landi. Study of the benefits of tangible user interfaces applied to teaching scheduling to french nursing students using a serious game simulating a hospital ward. Library and information sciences. Université Paul Sabatier - Toulouse III, 2023. English. NNT : 2023TOU30130 . tel-04356512

HAL Id: tel-04356512

<https://theses.hal.science/tel-04356512>

Submitted on 20 Dec 2023

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

En vue de l'obtention du

DOCTORAT DE L'UNIVERSITÉ DE TOULOUSE

Délivré par : *l'Université Toulouse 3 Paul Sabatier (UT3 Paul Sabatier)*

Présentée et soutenue le *11/07/2023* par :

Benoît Landi

**Study of the benefits of tangible user interfaces applied to
teaching scheduling to French nursing students using a
serious game simulating a hospital ward**

JURY

EMMANUEL DUBOIS	Professeur des universités	Président du Jury
NANCY RODRIGUEZ- DESTRUDEL	Maître de conférences	Membre du Jury
PIERRE LAGARRIGUE	Professeur des universités	Invité
MICHEL GALAUP	Maître de conférences (HDR)	Invité

École doctorale et spécialité :

MITT : Image, Information, Hypermédia

Unité de Recherche :

Institut de Recherche en Informatique de Toulouse (UMR 5505)

Directeur(s) de Thèse :

Jean-Pierre JESSEL et Catherine PONS-LELARDEUX

Rapporteurs :

Frédérique MÉRIENNE et Sébastien GEORGE

Acknowledgement

À Maurice, Jany & Michel,
À mes parents

Acknowledgements

First of all, I would like to thank the members of the jury for the time they allocated to my work and their expertise in evaluating it. I would like to first thank Frédéric Mérienne and Sébastien Georges for having accepted to be the rapporteurs of this work. I would also like to thank to Emmanuel Dubois and Nancy Rodriguez-Destruel for having accepted to be members of the jury.

I would like to thank my thesis director Jean-Pierre Jessel for having supported and guided me during this work. Thank you to Catherine Pons-Lelardeux for her supervision and advice.

Of course, I can't forget the SGRL team : Pierre Lagarrigue, David Panzoli, Hervé Pingaud, and our chief Michel Galaup. Their welcome, advice and support have been indispensable in this thesis. I want to recognize all the people who have worked with me and who have directly or indirectly participated, through our discussions, in this work, especially Jean-Benoît Culié, Hubin Zhang and Christophe Mazé. In addition, I want to thank the staff of the INU Champollion for their welcome and their help. I would like to thank my family for their patience and support during those years of work.

Last but not least, I would also like to thank the teams of the different IFSIs who helped me both during the development of the tangible version and for the experiments. In particular, the teams of Castres, including Matthieu Chincholle and Muriel Alary, those of the Toulouse University Hospital, including Sabrina Dida and Thierry REY, and those of Albi, including Nathalie Barthes and Corinne Four. Of course, I don't want to forget the students who gave me their time without who none of the experiments could have been done and for that I want to thank them warmly.

Remerciements

Je remercie tout d'abord les membres du jury pour le temps qu'ils m'ont accordé et leur expertise dans l'évaluation de mes travaux. Je voudrais remercier en premier lieu Frédéric Mérienne et Sébastien Georges d'avoir accepté d'être rapporteurs de ces travaux et bien sûr, merci à Emmanuel Dubois et Nancy Rodriguez-Destruel d'avoir accepté d'être membres du jury.

Je voudrais ensuite remercier mon directeur de thèse Jean-Pierre Jessel de m'avoir soutenu et guidé durant ces travaux. De même, je voudrais dire merci à Catherine Pons-Lelardeux pour son encadrement et ses conseils.

Bien évidemment, je ne peut pas oublier l'équipe du SGRL : Pierre Lagarrigue, David Panzoli, Hervé Pingaud, et notre chef Michel Galaup. Leur accueil, conseils et soutiens ont été indispensable dans cette thèse et je les remercie. Je remercie aussi toutes les personnes qui ont travaillé à mes côtés et qui ont, directement ou indirectement participé, par nos discussions, à ces travaux notamment Jean-Benoît Culié, Hubin Zhang et Christophe Mazé. Merci aussi au personnel de l'INU Champollion dans son ensemble pour leur accueil et leur aide.

Je voudrais remercier ma famille pour leur patience et leur support durant ces années de travail.

Enfin et pas des moindres, je remercie aussi les équipes des différents IFSIs qui m'ont aidé à la fois durant le développement de la version tangible mais aussi pour les expérimentations. Notamment les équipes de Castres dont Matthieu Chincholle et Muriel Alary, celle du CHU de Toulouse dont Sabrina Dida et Thierry REY, et d'Albi dont Nathalie Barthes et Corinne Four. Bien sûr, je n'oublie pas les étudiants qui m'ont accordé de leur temps et je les remercie tous chaleureusement.

Abstract

Abstract

Tangible User Interfaces or simply TUIs are a different type of interfaces developed in the 1990s. They aim to bring the interacted-with virtual elements into the physical environment. This is done to reduce the abstraction gap between virtual and physical. In parallel, Clone is a serious simulation game that teaches scheduling to nursing students. It was developed in response to a teaching need expressed by some Nursing Care Teaching Institutes and the Regional Health Agency of Occitania of Occitania in collaboration with our laboratory the SGRL. This is a very complex game that simulates a hospital wing and asks players to schedule a full day shift for both a nurse and a nurse assistant. This complexity lead us to look for a solution to help users. In particular, we wanted to see the potential benefits that TUIs could bring.

To measure those benefits, we developed a new tangible version of Clone's scheduling interface. This development lead us to different design choices and prototypes version presented further in this document. With a working prototype, we then tested it along with the original version with the help of 24 nursing students and measured multiple user-experience metrics. The results don't show, for most of them, a statistically significant difference between the two versions with a wide distribution of values in the metrics. While our participant pool is not large enough to make a definitive affirmation, this leads us to believe that for our very specific need, TUIs might not be the solution. However, the prototype fidelity to the original is not perfect and further studies on more simple and elemental interfaces could provide different results.

Keywords: Teaching; Tangible interfaces; Scheduling; Healthcare; Nursing;

Résumé

Les Tangible User Interfaces ou simplement TUIs sont un type différent d'interfaces développées dans les années 1990. Elles visent à intégrer les éléments virtuels avec lesquels on interagit dans l'environnement physique. Cela permet de réduire l'écart d'abstraction entre le virtuel et le physique. Parallèlement, Clone est un jeu sérieux de simulation qui enseigne la planification aux étudiants en soins infirmiers. Il a été développé en réponse à un besoin d'enseignement exprimé par les Instituts de Formation en Soins Infirmier d'Occitanie et l'Agence Régionale de Santé Occitanie en collaboration avec notre laboratoire le SGRL. Il s'agit d'un jeu très complexe qui simule une aile d'hôpital et demande aux joueurs de planifier une journée complète de travail pour une infirmière et une aide-soignante. Cette complexité nous a amenés à chercher une solution pour aider les utilisateurs. En particulier, nous voulions voir les avantages potentiels qu'TUIs pourrait apporter.

Pour mesurer ces avantages, nous avons développé une nouvelle version tangible de l'interface de planification de Clone. Ce développement nous a conduit à différents choix de conception et à des prototypes présentés dans ce document. Avec un prototype fonctionnel, nous l'avons ensuite testé avec la version originale avec l'aide de 24 étudiants en soins infirmiers et nous avons mesuré plusieurs paramètres de l'expérience utilisateur. Les résultats ne montrent pas, pour la plupart d'entre eux, de différence statistiquement significative entre les deux versions, avec une large distribution des valeurs dans les mesures. Bien que notre groupe de participants ne soit pas assez important pour nous permettre de faire une affirmation définitive, cela nous amène à penser que pour notre besoin très spécifique, TUIs ne sont peut-être pas la solution. Cependant, la fidélité du prototype à l'original n'est pas parfaite et d'autres études sur des interfaces plus simples et plus élémentaires pourraient donner des résultats différents.

Mots-Clefs : Apprentissage humain ; Interfaces tangibles ; Planification ; Santé ; Infirmier ;

Contents

Acknowledgement	2
Abstract	4
Appendices	10
Appendix A Acronyms	11
Introduction	13
1 Context	15
1.1 Clone's creation	16
1.1.1 Nursing workload in France	16
1.1.2 Project design and experimentation: design based research . .	16
1.1.3 Gameplay overview	17
1.1.4 Students' pedagogic progression in Clone	17
1.2 Game Elements	18
1.2.1 Patients' files	18
1.2.2 Steps	20
1.2.3 Workload indicators	20
1.2.4 Information	20
1.2.5 Scenarios	21
1.2.6 Conceptual model	21
1.2.7 Cognitive process	21
1.3 Clone gameplay phases	23
1.3.1 Scheduling phase	24
1.3.2 Executing phase	25
1.4 Clone's results and opportunities	26
1.4.1 Overall positive feedback	26
1.4.2 Complexity, workload and information overload	27
1.4.3 Tangible interfaces	28
2 Problematisation	30
2.1 Direction of our research	31
2.1.1 Tangibility: characteristic of interfaces	31
2.1.2 A/B testing and Multivariate test	32
2.1.3 Metric impact from tangibility	33
2.1.4 Constraint on the experimentation from the A/B testing method	33
2.1.5 Potential impacted metrics	34

2.2	Research questions	34
2.2.1	Questions of design	34
2.2.2	Questions of measures	34
2.3	Metrics measurement	35
2.3.1	Questionnaires' constraints	35
2.3.2	Questionnaires	36
2.3.3	In-Game data	39
3	State of the Art	41
3.1	Planning and Scheduling	42
3.1.1	Defining scheduling concepts	42
3.1.2	Skills set associated with scheduling	42
3.1.3	Scheduling in healthcare	44
3.2	Serious Games	45
3.2.1	From games	45
3.2.2	... to serious games	46
3.2.3	... and to learning games	46
3.2.4	Benefits of games for "serious" purposes	47
3.3	Tangible User Interfaces	48
3.3.1	Origin and early development of the concept	48
3.3.2	Defining tangible interfaces	48
3.3.3	TUI within frameworks / Paradigms	50
3.3.4	A wide array of tangible projects	53
3.3.5	Promises and drawbacks	54
3.3.6	Spatial cognition	55
3.4	Technologies	59
3.4.1	Tracking through chips	59
3.4.2	Image Recognition	63
4	Tangible schedule development	66
4.1	Prototype versions overview	67
4.1.1	Introduction	67
4.1.2	First prototype	67
4.1.3	Second and third prototypes	68
4.2	Boards	70
4.2.1	First Prototype's modular board	70
4.2.2	Second's and third's projected board	73
4.3	Patterns	77
4.3.1	Custom patterns	77
4.3.2	Pattern's border	78
4.3.3	Pattern's core: Number of squares and colours	78
4.3.4	Pattern's core: Central symmetry	79
4.3.5	Pattern's core: Colour encoding and decoding	79
4.3.6	Pattern's core: size of coloured bits	81
4.4	Detection	81
4.4.1	Methods: Image Matching vs contours and core sequencing vs Aruco	81
4.4.2	Detection loop	82
4.4.3	Connecting OpenCV and Unity	83

4.5	Activities' icon design	84
4.5.1	The need of remaking the icons	84
4.5.2	The new icons design	84
4.5.3	New icons validation	85
4.5.4	Need for categories	86
4.6	Answering our design research questions	86
4.6.1	Tangibilisation of Clone's game elements	86
4.6.2	Lesson on the tangibilisation	87
4.6.3	Criterion for our tangibilisation's quality	89
5	Experimentation and results	91
5.1	Protocol	92
5.1.1	Participants	92
5.1.2	Experimental steps	92
5.2	Data	94
5.2.1	Questionnaires: Data Clean up	94
5.2.2	Questionnaires: General Information	95
5.2.3	In-Game: Data Clean up	95
5.3	First test results	96
5.3.1	Introduction	96
5.3.2	Attrakdiff(FR) questionnaire	96
5.3.3	F-SUS questionnaire	97
5.3.4	NASA-TLX(FR) questionnaire	98
5.4	First or Second interface played difference results	98
5.4.1	Introduction	98
5.4.2	Results for all questionnaires	100
5.4.3	Graphs for all questionnaires	101
5.5	In-Game data's Results	103
5.6	Results by demographic category	105
5.6.1	Gender	105
5.6.2	Confidence on computer abilities	105
6	Interpretation and discussion	110
6.1	General interpretation of the experimental campaign's feedbacks and unfolding	111
6.1.1	Experimental campaigns	111
6.1.2	Direct feedbacks	111
6.1.3	Need of explanation and questions	111
6.2	Interpretation of the data collected	112
6.2.1	Questionnaires' feedback : first interface played	112
6.2.2	Questionnaires' feedback : first and second compared	113
6.2.3	In-Game measures	113
6.2.4	Demographics' interpretation	114
6.3	Discussion	114
6.3.1	Overall results feedbacks	114
6.3.2	Answering our metric based research questions	115
6.3.3	Other research questions	116
6.3.4	Similarity in the interfaces' versions	116
6.3.5	Reach of the results	116

6.4 Perspectives	117
6.4.1 Opportunities for Clone's future	117
6.4.2 Better bases for the analysis	117
Conclusion	119
Appendices	121
Appendix B Figures' and tables' list	122
Appendix C Questionnaires	126
C.1 AttrakDiff(FR)	126
C.2 F-SUS	127
C.3 NASA-TLX(FR)	128
Bibliography	129

Appendices

Appendix A

Acronyms

AR Augmented Reality	31, 53, 62, 63
ARS Regional Health Agency of Occitania (Agence Régionale de Santé Occitanie) 4, 5, 13–16, 21, 26, 29, 117	
AS Nurse Assistant (Aide-Soignant)	17, 20, 21, 24, 25, 77, 85, 112, 116
ASD Autism Spectrum Disorder (Troubles du spectre autistique)	53
ATT Evaluation construct (Atractivité globale)	37, 96, 112, 114
AV Augmented Virtuality	31
DLL Dynamic-link library	83
DREES Directorate for Research, Studies, Evaluation and Statistics (Direction de la recherche, des études, de l'évaluation et des statistiques)	95
EIAH Informatics environment for human learning (Environnement Informatique pour l'Apprentissage Humain)	46
F-SUS French System Usability Scale	37, 98, 102, 112, 127
GEQ Game Experience Questionnaire	38
GEQ2 Game Engagement Questionnaire	38
GUI Graphical User Interface	34, 40, 48, 49, 57, 90, 96–98, 105, 113–115, 117
HCI Human-Computer Interface (Interface Homme-Machine)	32, 36, 48
HF High Frequency	61
IDE Registered Nurse (Infirmier Diplômé d'État)	17, 20, 25, 44, 77, 85, 112, 116
IFSI Nursing Care Teaching Institute (Institut de Formation en Soins Infirmier)3, 4, 13–16, 21, 26, 29, 73, 85, 91, 92, 111, 117, 119, 120	
ISO International Standards Organisation	59
MCDM Multi-Criteria Decision-Making	88
APPENDIX A. ACRONYMS	11

MIT Massachusetts Institute of Technology	13, 14, 48, 55
MR Mixed Reality	31
MR-VC Mixed Reality Virtuality Continuum	31
NASA-TLX NASA Task Load Index	37, 112, 113, 115, 128
NASA-TLX-FR NASA Task Load Index - French Version	37, 98, 103
NFC Near-Field Communication	61, 62, 68
NRP Nurse Rostering Problem	44
NSP Nurse Scheduling Problem	44
OCIS Organising and Coordinating Health care Interventions (Organiser et coordonner des interventions soignantes)	44, 45
OpenCV Open Source Computer Vision Library	63
PVC Polyvinyl chloride (Polychlorure de vinyle)	73
QHI Hedonic quality - Identity (Qualité Hédonique - Identité)	37, 96, 112, 114
QHS Hedonic quality - Stimulation (Qualité Hédonique - Stimulation)	37, 96, 112, 113
QP Pragmatic Quality (Qualité Pragmatique)	37, 96, 97, 112, 113, 115
RFID Radio-Frequency Identification	59–62, 68
SGRL Serious Game research Lab (Serious Game research Lab)	3–5, 16, 29, 117
SUS System Usability Scale	37, 113, 115
TUI Tangible User Interface	4, 5, 13, 14, 34, 40, 41, 48–52, 54–58, 65, 90, 91, 96–98, 105, 110, 113, 114, 116–118
UES User Engagement Scale	38
UES-SF User Engagement Scale - Short Form	38
UI User Interface	16, 55, 88, 89
UX User eXperience	55, 88, 119
WIMP Windows, Icons, Menus, Pointer	13, 48, 56, 58

Introduction

Introduction in English Tangible User Interfaces or TUIs for short are a type of interfaces that were first developed in the 1990s by a team of researchers at MIT. In opposition to the more classical “Windows, Icons, Menus, Pointer” paradigm, those interfaces’ goal is to bring the virtual elements that are the target of a user’s interaction into the physical objects that the user is actually manipulating. Thereby reducing or even closing the divide between the two environments. Those ideas have been developed since then in a wide range of area and have been the subject of research from multiple perspectives. In our work, we are looking at them to try and find a possible solution for an issue with a game developed in our laboratory.

This game named Clone for “CLinical Organiser Nurse Environment”. Clone is a serious game created to answer a need expressed by “Agence Régionale de Santé Occitanie” (ARS) for Regional Health Agency of Occitania and nursing schools called “Institut de Formation en Soins Infirmier” (IFSI) for Nursing Care Teaching Institute. This need is the deficiency of students in scheduling skills. Clone simulates an hospital ward and tasks the players with scheduling the day shift of nursing staff. The player has to read the patients files and use their knowledge of clinical practices to identify the care needed. Then, he has to schedule those care acts, again using the clinical practices and the hospital internal rules, while respecting the time constraints imposed by the day shift. Finally, the player has to apply the care acts which he scheduled and deal with the unforeseen events given by the game, adapting his schedule on the fly to still be able to take care of the patient appropriately.

The game is very complex with hundreds of information about patients and the hospitals, dozens of care acts available and in consequence very dense interface. The direct feedback and the data collected during tests shows that this impact negatively the students using the application. The tangible interfaces, according to the literature available, through their different approach of interfaces could be a solution to, if not solve, at least alleviate some of those problems.

In this work, we present our work creating an alternative version of the Clone interface using tangible principles. This design and its evaluation was done throughout our time at the laboratory. Notably, we detail the decisions taken and the solutions employed throughout its development to produce a working version that would respect the constraints imposed by our intended experiments. Having tangibilised our graphical interface, we performed tests to determine the potential impact of using a tangible interface over a graphical one. Finally, we present and discuss the results of the work, offering our conclusions and potential future researches that could continue the ideas explored in this document and our project.

Introduction en français Tangible User Interfaces ou TUIs en abrégé sont un type différent d'interfaces qui ont été développées pour la première fois dans les années 1990 par une équipe de chercheurs du MIT. L'objectif de ces interfaces est d'intégrer les éléments virtuels qui sont la cible de l'interaction de l'utilisateur dans les objets physiques qui sont ce que l'utilisateur manipule réellement. Cela permet de réduire, voire de combler le fossé entre les deux espaces physiques et virtuels. Ces idées ont été développées en un large éventail de projets et ont fait l'objet de recherches sous de multiples angles. Dans le cadre de notre travail, nous les examinons pour tenter de trouver une solution possible à un problème lié à un jeu développé dans notre laboratoire.

Ce jeu est Clone pour « CLinical Organiser Nurse Environment ». Il s'agit d'un jeu sérieux conçu pour répondre à un besoin exprimé par l'Agence Régionale de Santé Occitanie (ARS) et les écoles d'infirmières appelé Institut de Formation en Soins Infirmier (IFSI). Ce besoin est lié au manque de compétences des étudiants en matière de planification. Clone simule un service hospitalier et demande aux joueurs de planifier l'équipe de jour du personnel infirmier. Le joueur doit lire les dossiers des patients et utiliser sa connaissance des pratiques cliniques pour identifier les soins nécessaires. Ensuite, il doit planifier ces soins, toujours en utilisant les pratiques cliniques ainsi que les règles internes de l'hôpital, tout en respectant les contraintes de temps. Enfin, le joueur doit appliquer les actes de soins qu'il a programmés et faire face aux événements imprévus donnés par le jeu, en adaptant son planning à la volée pour toujours pouvoir traiter le patient de manière appropriée.

Le jeu est très complexe avec des centaines d'informations sur les patients et les hôpitaux, des dizaines d'actes de soins disponibles et par conséquent une interface très dense. Les retours des étudiants qui l'utilisent ainsi que les données collectées lors des tests montrent que ça a un impact négatif durant leurs parties. Les interfaces tangibles, selon la littérature disponible, par leur approche différente pourraient être une solution pour, sinon résoudre, au moins alléger certains de ces problèmes.

Dans ce travail, nous présentons le projet de création d'une version alternative de l'interface Clone en utilisant des principes tangibles. La conception de ce projet s'est faite tout au long de notre séjour au laboratoire. Nous détaillons notamment les décisions prises et les solutions employées tout au long de son développement pour produire une version fonctionnelle qui respecterait les contraintes imposées par les expériences prévues. Après avoir tangibilisé notre interface graphique, nous avons effectué des tests pour déterminer l'impact potentiel de l'utilisation d'une interface tangible par rapport à une interface graphique. Enfin, nous présentons et discutons les résultats de ces travaux, en proposant notre conclusion et des travaux futurs potentiels qui pourraient poursuivre les idées explorées dans ce document et dans notre projet.

Chapter 1

Context

Introduction Clone is a nurse scheduling simulation serious game developed by our laboratory in collaboration with Regional Health Agency of Occitania and Nursing Care Teaching Institute of the Occitania region. Its goal is to improve the scheduling skills of nursing students and in turn upgrade their overall job performances. Being a project developed with partner schools allowed the continuous testing and improving of the design as well as collecting scientific data. The game's goals are to : first schedule the needed tasks for the day-shift depending on the patients assigned and what is written in their files, second execute those tasks and adapt the schedule to unforeseen events during play. In this section, we will go in details about Clone's creation and overall design, Clone's game elements and phases and finish on the results of the multiple questionnaires and feedbacks collected during the project.

Introduction Clone est un jeu sérieux de simulation de planning infirmier développé par notre laboratoire en collaboration avec l'Agence Régionale de Santé Occitania et des écoles en soins infirmiers de la région. Son objectif est d'améliorer les compétences de planification des étudiants en soins infirmiers afin d'améliorer leurs performances professionnelles globales. Le fait que le projet soit développé avec des écoles partenaires a permis de tester et d'améliorer continuellement le design ainsi que de collecter des données scientifiques. Les objectifs du jeu sont les suivants : premièrement, planifier les tâches nécessaires pour l'équipe de jour en fonction des patients assignés et de ce qui est écrit dans leurs dossiers, deuxièmement, exécuter ces tâches et adapter le planning aux événements imprévus pendant le jeu. Dans cette section, nous détaillerons la création et le design général de Clone, les éléments et les phases de jeu de Clone et, nous terminerons par les résultats des multiples questionnaires et retours collectés au cours du projet.

1.1 Clone’s creation

1.1.1 Nursing workload in France

Even before the pandemic, the workload of each nurse in french hospital has been steadily increasing. This phenomenon is not recent nor is it just affecting France, leading to increase job-dissatisfaction and burnouts [Aiken, 2002, McHugh et al., 2011]. Medical unit in hospitals hosts more and more patients simultaneously with less and less nurses in their care teams leading to increasing time pressure for nurses. Every profession has its specific technical skills, but also use a panel of non-technical skills to support the first ones [Fletcher et al., 2003, Flin et al., 2010]. Situations of high time pressure and tasks’ complexity [Huber et al., 2021] exacerbate the need for scheduling skills in nursing staff in order to maximise efficiency. In response, the “Agence Régionale de Santé Occitanie” (ARS) for Regional Health Agency of Occitania initiated in 2018 the project Clone, for “CLinical Organiser Nurse Environment”. This project was done through a collaboration of the ARS, the nursing schools of Occitania, called “Institut de Formation en Soins Infirmier” (IFSI) for Nursing Care Teaching Institute, and, the Serious Game research Lab (SGRL) of the University Champollion of Albi (France). The ARS and IFISIs provided : (1) expert knowledge of both the work and the curriculum of nurses and (2) the SGRL provided both the technical development and scientific foundations for analysing the final product. The challenge is to blend the complex reality of the nursing work and the game design methods to teach that reality [Landi et al., 2022c]. The teaching focuses on the skill 9 of the french nursing curriculum: “Organiser et coordonner les interventions soignantes” (“Schedule and coordinate care acts”)[France compétences, 2022]. This skill is complex, including both hard skills (scheduling, workload smoothing,...) and soft skills (intra team cooperation, inter team/care-unit cooperation,...). Before the project Clone was initiated, each IFISIs had different courses teaching the skill 9. Those were usually paper exercises that lacked the patient amount (never more than 6) and were rather hard to manage. Most notably, the exercises not being on computer meant there was no automatic workload gauges that directed players on their schedule feasibility.

1.1.2 Project design and experimentation: design based research

The project was launched in 2018 and continued development up to 2021. In term of project development, the design was conducted by the SGRL team along regular meetings with ARS and IFSI representatives to insure the veracity of the content. The IFISIs collected real patient files, which were anonymised, to build educational content dedicated to the project. They also provided data about the work in an hospital department : tasks expected of a nurse, approximate/average duration of each task, template for hospital regulations, national regulations,... [Lelardeux et al., 2020]. During those development years, different version of Clone were used in class by some of our partners in 11 IFISIs. These experiments saw 2400 students use Clone and provided us three main benefits. The first one was very practical: bug control from a wide user base which is more apt to catch elusive bugs. The second were design feedbacks. UI, game feel, challenges too difficult are very hard to test by

the development team and play-tests allowed to course-correct the design. Finally, the third was preliminary scientific data through user evaluation. During play-test, players were given questionnaires at the end of sessions to evaluate their experience with the game. In addition, the game registers a broad array of game data and send them, link to the player and the game session, to an external database for analysis later on. Intervention design [Brown, 1992], or known more as design-based research, is based around using interventions (E.I. actually applying the proposed solutions to problems) to evaluate those solutions and then exploit the feedback from the interventions. Clone's development method of using data from the in-situ uses of the game to fuel both research and its development is a good example of such method [Fallman, 2003, Obrenovic, 2011, Anderson and Shattuck, 2012]. In our case, this method is combined with a feedback-development loop that uses the research models and conclusions to improve the game's design [Pons Lelardeux et al., 2020, Galaup et al., 2021, Novoseltseva et al., 2022].

1.1.3 Gameplay overview

Clone is a scheduling game where the player takes the role of a Registered Nurse (called IDE for "Infirmier Diplômé d'État") with a Nurse Assistant (called AS for "Aide-Soignant") in charge of between 1 and 14 patients. He has to schedule and do a day shift (6:40 to 13:50). The player is given information from both the preceding night-shift and patients records (medical, nursing and administrative). With those information, his knowledge of the clinical practices and his nursing knowledge, he has to understand what tasks are needed for each patient. Then he has to schedule them according to the constraints (order, hospital deadlines, health reasons, etc). Finally, he has to perform the scheduled tasks. During this performing, the current time advances along each tasks performed and game events can give the player new information. He then has to adapt to the unforeseen by modifying his scheduling and carrying on. During the game, if the player makes a critical mistake, he loses a "life". Those mistakes are considered grave and therefore player should not schedule without taking them into account. The player loses the game if he runs out of lives and wins if he finishes the shift with some lives remaining. At the end, the player is given an auto-evaluation questionnaire and then a summary of his performances. A trailer of the game is available at the following address "<https://www.univ-jfc.fr/actu/clone-un-serious-game-dedie-lorganisation-du-travail-en-milieu-hospitalier>" on the university' website.

1.1.4 Students' pedagogic progression in Clone

The experiment were organised as lessons for students and a pedagogic progression was planned for. Across their 3 years of nursing school, students would use Clone 5 times. Each session would have different objectives. The first sessions were tutorials on how to use the software and what is expected of them while the last 3 were the actual game with an increasing difficulty across the sessions. The 5 sessions are presented in Table 1.1.

Year	Semester	Duration	Objectives	Content	Schedule	Execution
1	S2(UI 5-2)	2h	Discovery of the game - Understanding patient's files - Clinical judgement	1 patient (4,9 or 11) - Clinical rules	X	X
2	S3(UI 5-3)	3h	First scheduling without resource constraint - workload understanding	First multi-patient scenario : 3 patients (1, 8 and 14)	X	X
2	S4(UI 5-4)	3h	Scheduling with resource constraint	6 patients scenario (1,2,4,6,7,and 8)	X	X
3	S5(UI 5-5)	3h	Scheduling with resource constraint (harder)	10 patient scenario (6 patients adding (3,5,10,12))	X	X
3	S5(UI 5-5)	3h	Full scheduling without execution	14 patients scenario (10 patients adding (9,11,13,14)	X	/

Table 1.1: Clone's pedagogic progression table

1.2 Game Elements

Activities and tasks

Clone being a nursing care scheduling game, it is centred around the care acts being performed by nursing staff. In the game proper, there are however two concepts that represent those acts : activities and tasks. Activities are what is scheduled in the care plan. Tasks are the actual care acts being performed. For example, "Give oral medicine" would be the activity while "Giving Ciflox[®] and Dafalgan[®]" (both pills taken orally) would be the corresponding task for a patient needing those medications. In the interface, both activities and tasks use the icon of the activity/task category, which are shared, on a blue background but the activity UI-element is a square while the tasks is a circle.

1.2.1 Patients' files

As explained above, the player uses the patient's files to schedule the appropriate care. Like in the real nursing world, each patient has 3 files : administrative, medical and nursing (see Figure 1.1). Each contains the information a real file would. Information in patient file are covered and can be revealed by clicking on them. This information is also added to the player information "memory", a side bar that keep all information readily available and has some filter to quickly find back information.

Dossier infirmier Numéro : 76303-4

Identité du patient : Mr Adrien Pouzole
né(e) le 11/01/19xx
sexe : homme

Age : DV ATIS

Macro-cible : ΔΠΤΕΡΟΔΕΣΠΙΣ : ΗΣΡΠΠΦΣ ΦΠΓΥΦΠΛΣ, ΑΔΕΠΘΜΣ ΔΣ ΛΑ ΠΡΘΣΤΑΤΣ, ΑΔΕΠΘCΑΡCΦΠΘΜΣ ΔΥ ΡΕCΤΥΜ ΘΡΕΡΕ (+ ΡΑΔΦΘΗΘΗΡΑΡΦΣ ΣΤ CΗΦΜΦΘΗΘΗΡΑΡΦΣ),

Antécédents : Ηλεκτρική τρενάκι : Ηλεκτροεγκεφαλόγραμμα εγώ εδύ κρούση, επί χημικής οξείας φάσης για τραύμα υπερτασικό με ήπιο αρθρικό, γρήγορο έκτακτο

Historique : Αδυναμία πορείας με αποκατάσταση σε ποσοστό 60%

Motif d'hospitalisation : Αδυναμία πορείας με αποκατάσταση σε ποσοστό 60%

Transmissions

Hospitalisé depuis 1 jour.
Déorientation temporo-spatiale. Déplacement en fauteuil depuis l'hospitalisation en chirurgie orthopédique.
Incontinence urinaire et fécale.
Présence d'une escarre sur le bord externe du pied droit

(a) Administrative

Dossier administratif Numéro : 76303-4

Identité du patient : Mr Adrien Pouzole
né(e) le 11/01/19xx
sexe : homme

Age : 91 ans

Médecin référent : Dr. Dubois

Personne à prévenir : Pas de personne à prévenir

Mutuelle : PAMOUTH

Situation familiale : Il est veuf avec 2 enfants en ORFÈVRE depuis 10 ans. Il est retraité de la SNCF

Tutelle : N'ANI

(b) Medical

Dossier médical Numéro : 76303-4

Identité du patient : Mr Adrien Pouzole
né(e) le 11/01/19xx
sexe : homme

Age : 91 ans

Prescription per os
Amilor® (amlodipine) 5mg cp (1.0.0)
Aspirine kardégrip® (Acétylsalicylate de lysine) 75 mg (1.0-0)
Bactrim® forte cp (sulfaméthoxazole-triméthoprime) (1.0-0) pendant 8 jours
Dolipranel® paracétamol 500 mg 2 cp à 6h, 12h, 18h et 20h
Efferal® fentanyl 100 mg 1 cp sublingual 14 heures si douleur (maximum)
Xatral® alfuzosine LP 10 mg cp (1.0.0)

Prescription des injectables
Lantus® (insuline glargine) 11 UI SIC à 8h
Lovenox 4000 UI injection sous cutanée (SC) à 18h (Enoxaparine sodique)

Autres prescriptions
Pansement 1X sous MEOPA® (oxygène-protolyde d'azote) 6L/lin pend.
Surveillance glycémie capillaire x3.
Douleur par Echelle Visuelle Simple (EVA)
Surveillance des ingesta
Collation à 10h et 16h
Scanner du bassin prescrit hier

Compte-Rendu d'exams
Hier : CRP = 60 mg/l

(c) Nursing

Figure 1.1: The patients' files' interfaces with, from top to bottom : administrative(a), medical(b) and nursing(c). The administrative file shown has some information still “hidden” with grey block and the text in Greek alphabet.

1.2.2 Steps

Steps are the time slots that constitute the nurse shift. Most of them are an hour long.

- ▶ **6:40-8:00** *Tour du matin (prise de sang, traitement per os...)*
Morning round (blood test, oral treatment...)
- ▶ **8:00-9:00** *Petit-dejeuner (toilette, injections...)*
Breakfast (grooming, injections...)
- ▶ **9:00-10:00** *Toilette, pansements, injections*
Cleaning, dressings, injections
- ▶ **10:00-11:00** *Visite du médecin*
Visit of the doctor
- ▶ **11:00-12:00** *Relevée visite, RDV...*
Check-in, appointment...
- ▶ **12:00-13:05** *Dejeuner*
Lunch
- ▶ **13:06-13:50** *Soins*
Care

While each step is named with the usual activities being performed at those hours in a typical nurse shift, players can execute tasks more freely.

1.2.3 Workload indicators

Workload represents the percentage of time already assigned to activities for a given step. In the UI, this is usually represented with a pair of workload gauges that show both the nurse and his nurse assistant's workloads. When there is multiple steps represented in the interface (see subsection 1.3.1) there are multiple pairs of gauges. The calculation for the workload of the IDE is simply the ratio of time taken by all activities scheduled for a step within the step available time and respectively for the AS. The gauges can be seen in Figure 1.4.

1.2.4 Information

As in real life, information are central in a lot of the game. In Clone, those are split into elementary information and the player collect them during play. The information are obviously central for the players to make their decisions, but a lot of tasks can only be done if the player has unlocked the corresponding information. For example, if a patient needs some medicine like paracetamol, during the executing phase, the player will not be able to give him the medicine until he has looked at the patient chart telling him about it. Alternatively, some tests done by nurses, like "Measure of parameters" can give information; those would also unlock some tasks that may be required. With the 14 patients records, multiple scenarios (see below) with or without events, and around 320 different tasks, Clone has around 530 individual information.

1.2.5 Scenarios

While Clone has 14 different patients, the game load sets of them using scenarios (see Table 1.1). For example, while one of the most basic scenario only contains patient #1, one more advance contains the patient #1, #8 and #14 and the most advances scenarios contains all 14 patients and some events. Scenarios are mainly a list of patients but also allow designer to tweak the number of assistant and to include some specific unexpected events in addition to the patient's regular chart. This gives players an easier time to learn the game interface with a game where scheduling constraints are very low. The scheduling constraints are almost always the same as the duration of the shift is fixed across all scenarios. The scenario designer have one element which can be altered : the number of assistants. In Clone, the player can delegate some tasks to the AS and the amount of AS available depends on the scenario. This allow the player to have more available AS' workload. By default, it is at 1 but is increased to 1.5 for the 14 patients scenarios.

1.2.6 Conceptual model

As explained above in each of the elements' section, they interact with each other. Those interactions can be summed up in Figure 1.2. The centre point is the relation between activities and tasks. In the next part about phases we will see that tasks interact with the execution while activities are what is scheduled.

1.2.7 Cognitive process

Work has been done to step back and analyse the process used by the player to solve the scenario offered by Clone. One of representation of this process is the cognitive process loop available in Figure 1.3. This process has five nodes:

1. "Patient Files"
2. "Clinical Practices"
3. "Care acts"
4. "Care plan"
5. "Workload plan"

It is quite evident that those will be closely linked or even also present in the conceptual model. The conceptual model is but Clone's implementation of the cognitive process. Each of the nodes is linked with the following one in order (meaning nodes 1 and 2 are linked with the first element of the following list):

1. The patients files are understood by the player using their knowledge of the clinical practices.
2. This understanding need to be refined to become a list of necessary care acts.
3. Those care act have to be coordinated into a rough care plan.
4. This care plan need to fit the workload plan constraints.
5. Finally, despite the workload plan, unforeseen events forces the player to read-just and go back to the patient files.

While some of those links are the purview of the experts at both the IFSIs and ARS, one of the link that has been researched for Clone's design is the validation [Landi et al., 2022a]. In our work, we discussed the challenge of designing a 14 patients

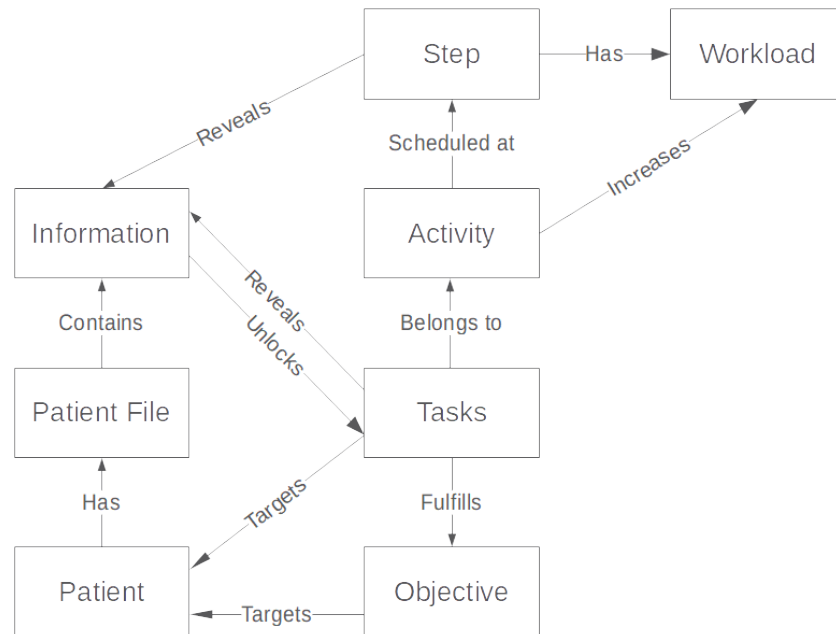


Figure 1.2: In this graph we can see some of the relationships between the different elements of Clone.

scenario and our use of an “ideal” schedule. The issue we encountered was that, for this scenario, students would take a very long time to come at the validation link of the cognitive process. One of the solution explored was to give them a starting point with an ideal schedule with all the activities that, in a perfect world, would be in the schedule at the step there were put by default. Then the player would have to correct this schedule to make it fit within the time constraints while still being within the clinical practices’ constraints. This shows that, using scenarios, we can focus on one step of the cognitive process and push students more specifically on certain skills. This model fits quite well with the the skills presented further down in this document (see subsection 3.1.2).

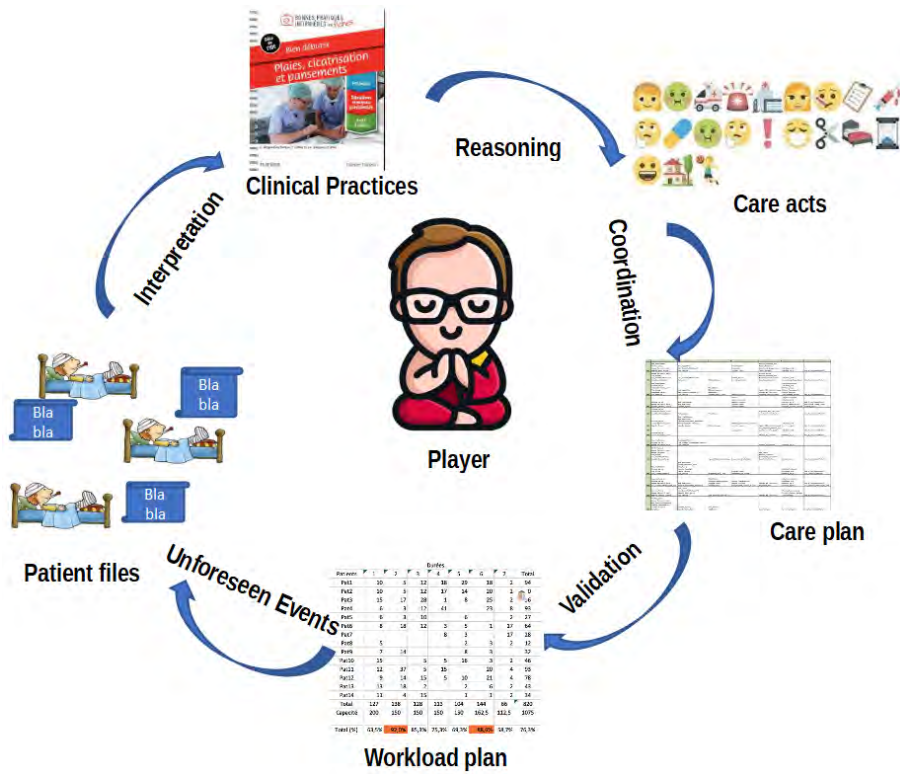


Figure 1.3: Clone’s cognitive process

1.3 Clone gameplay phases

Phases sequence

A game of Clone is structured in 5 phases with:

1. Morning communication with night-shift
2. Scheduling of the shift using patients’ records
3. Executing the tasks on patients
4. Midday communication with the next shift’s team
5. Review and end of game feedback

The phases are not equal in duration; the scheduling and executing phases constitute the bulk of the game, up to 3 hours for 14 patients scenarios, while the other three phases usually take less than a minute [Lelardeux et al., 2020]. In addition, while the phases are ordered as such and the player usually makes his schedule before executing it, he can go directly to the executing phase skipping the scheduling entirely as it is not needed for actually doing care act on patients. However, the player can go back to the schedule during the executing phase, even if players mostly only do so when an unforeseen event occurs. For example a new patient arriving might forces the player to delay some other tasks. Alternatively a new information asks for another care act to be done urgently. This means that, while the scheduling and executing phase are more or less in sequence, they can be played simultaneously if the players wants to.

1.3.1 Scheduling phase

The scheduling stage is performed using two graphical user interfaces: the global care plan (Figure 1.4) and the individual patient care plan. The global care plan is a table with all patients as lines and the shift split into seven steps/period of roughly one hour as columns. Each cell at the intersection of a patient line and a step column contains the activities scheduled for that patient for this step. This interface allows the player to easily move activities for a patient from step to step while watching the impact on the workloads. They can also delete activities but they cannot add new ones. To do that, they must click on either the “paintbrush” button in a patient-step intersection or on the “modify” next to the patient icon; this will open the personal schedule for that patient-step.

The individual patient care plan is a more focus schedule that mainly shows the scheduled activity for one patient-step. However, in this interface, the player has the list of all activities available in the game and can schedule them (by drag-and-dropping them into the central spot). They also can delegate the task to the Nurse Assistant (AS) if the task can be delegated (according to french laws). The interface also shows in smaller boxes the scheduled activities for the previous and next steps. Arrow buttons let the player navigate to the personal schedule of those step without going back to the main schedule. Finally, the interface includes the workload of the previous, the selected and the next steps so that the player can easily watch if a task can be freely added or if maybe, putting it in another steps would be better.

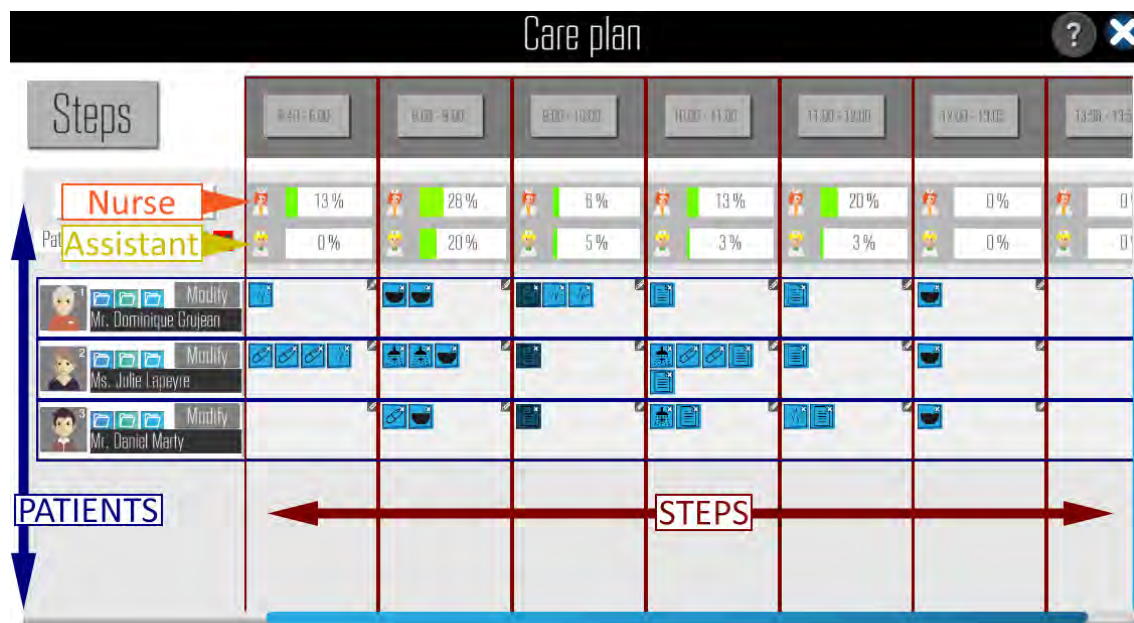


Figure 1.4: The global care plan with patient as lines (blue) and steps as columns (red). The workload gauges for the nurse (orange) and the nurse’s assistant (yellow) are displayed at the top of each step.



Figure 1.5: The patient care plan with the activities on the left (blue), the workloads above (green), the scheduled activities for the previous, current and next steps (red) and the delegation tab (yellow)

1.3.2 Executing phase

The executing phase is played on an overview of the service with patients in rooms and the nurse and his assistant moving around on the hospital floor. The player can click on patient to have the tasks performed by the IDE. Meanwhile, the AS will automatically performed its scheduled tasks when in the executing phase. The execution phase has the same steps segmentation as the schedule. Tasks' temporal objectives are aligned on those steps. For example, "Giving Inexium® 20mg(1-0-0)" on a patient charts tells the nursing student to give the medicine before 8h, as per hospital policy, which is before the end of the first step (6:40 - 8:00). Similarly when a patient charts asks for "Doliprane® paracetamol 500mg 2cp at 12h" the game will want the player to give the patient the medicine during the 12:00 to 13:00 step. We also talked about unforeseen event and new information. During the execution phase, new element can come to light requiring new care act to be done. This means that the player will have additional tasks that could mess his scheduling. When those tasks come up, he can go back to the previous phase and adapt his schedule.

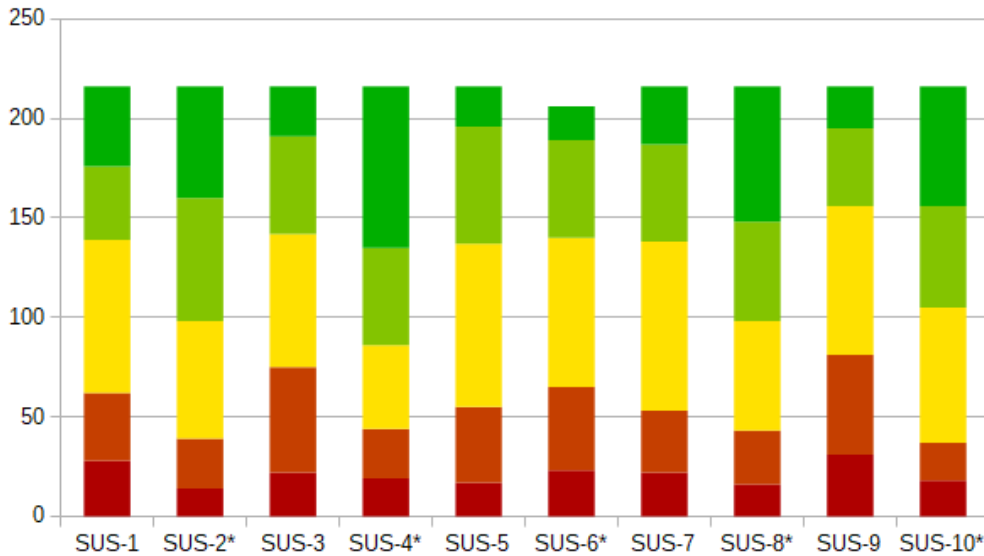


Figure 1.6: Clone’s execution interface. On the top is the current workload gauges for the current step which allow players to keep up with the impact of actions already done. The stars represent the lives of the players. Each patients has its room which has buttons to access the schedule tasks, the possible actions and his files.

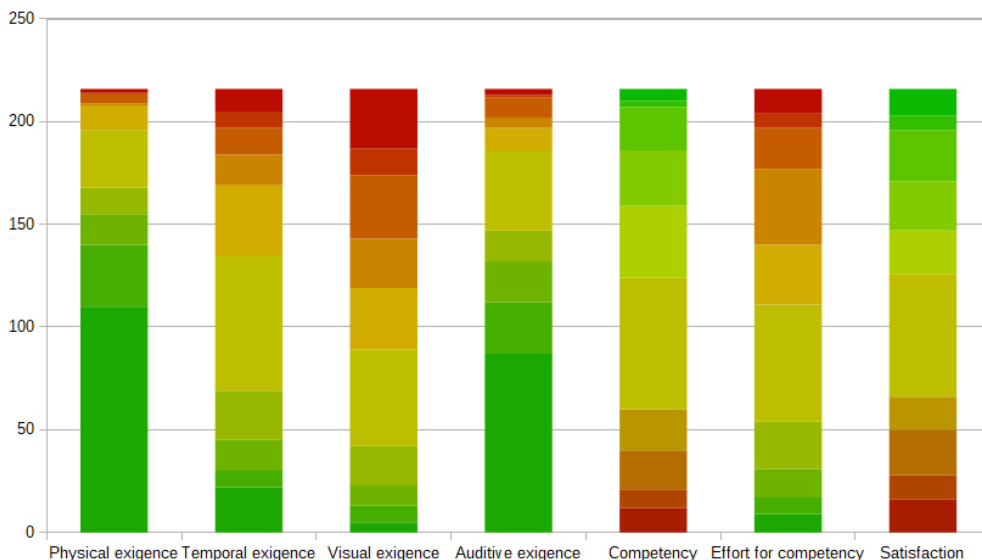
1.4 Clone’s results and opportunities

1.4.1 Overall positive feedback

If we look at the latest experimental campaign, the results are somewhat positive. For example, in the SUS part of the feedback questionnaire that students answer at the end of their session we get the results shown in Subfigure 1.7(a). While the results are not perfect, we still got some negative feedbacks about the application. Similarly, when talking with students, they seem to like using Clone in class. Some even use it again when at home. On the other project partner part, the ARS was satisfied with the project and while some IFSIs stopped using Clone, some kept it even to this day.



(a) Results from the SUS questionnaire (see section C.2 for the questions) . Items with a * have their results reversed (1-5 to 5-1 from the original questionnaire) so that green and red respectively reflect a positive and negative aspect of the software.



(b) Other metrics such as multiple workloads, competency, effort and satisfaction during the 2021 campaign.

Figure 1.7: Results from the questionnaire of the 2021 Clone's campaign

1.4.2 Complexity, workload and information overload

Still, both in some direct feedback and in questionnaires' results, we can see that Clone is a very complex game to tackle. In Subfigure 1.7(b), we see that, as can be expected, Clone is not a physically or auditively demanding game. The issues are on the temporal and most importantly visual exigence. Playing Clone is long and demand focus on a busy interface with a lot of information. In terms of perceived efficiency, seen in "Competency" and "Effort for competency" of the same figure, Clone results are very spread out. While some people perceives they have improved, some don't, or that it took a lot of effort to improve. This is echoed in the direct

feedbacks which also showed that a lot of the interfaces are unclear, users possible actions can be hard to understand and failure can come unexpectedly. All of this is despite the pedagogic progression put in place to ease people in through gradual increase of complexity.

1.4.3 Tangible interfaces

As we will see in section 2.1 and section 3.3, we look into solution to improve Clone's outcomes. One of the ways explored was tangible interfaces. Our cursory knowledge of the time directed us to think that maybe those interfaces could decrease the mental demand put on users when playing. Thus, this would increase both satisfaction, usability and outcomes. This thesis explores this direction. Our goal is to identify the impact of an tangible interface compared to a digital one on the mental demand and usability of a scheduling serious game.

Synthesis Clone is a simulation game developed to teach french nursing planning and scheduling in an clinical environment. The project was designed within the SGRL with the help of the ARS and IFSIs of Occitania from 2018 to 2022. Following the design based research method, continuous experimentation of the software on nursing students helped guide its development. Despite that and good feedback, issues still exist with Clone as the game is very taxing for users. This thesis explores one of the possible solutions, using another interface type: tangible interfaces.

Synthèse Clone est un jeu de simulation développé pour enseigner à des élèves infirmiers la planification et l'ordonnancement dans un environnement clinique. Le projet a été conçu au sein du SGRL avec l'aide de l'ARS et des IFSIs d'Occitanie de 2018 à 2022. Suivant la méthode de design based research, l'expérimentation en continu du logiciel sur des étudiants en soins infirmiers a permis de guider son développement. Malgré cela et de bons retours, des problèmes subsistent avec Clone car le jeu est très éprouvant pour les utilisateurs. Cette thèse explore une des solutions possibles à ce problème, l'utilisation d'une autre famille d'interface : les interfaces tangibles.

Chapter 2

Problematization

Introduction Tangible interfaces are a different way to conceive the interaction between human and machines. During our early works this thesis we conceptualised how to explore the impacts of using tangible interfaces. This was done both as an exploration of the tangible paradigm and as a way to solve some of Clone's issues. Among the topics explored there are: the practical constraints imposed by such endeavour, our reflection on concepts linked with tangible interfaces and solutions to measure the concepts on participants. In this section, we explore those points by presenting what was envisioned, putting forth our research questions and hypotheses, and finish by presenting with our measure methods for the concepts selected.

Introduction Les interfaces tangibles sont une façon différente de concevoir l'interaction entre l'homme et la machine. Au début de nos travaux durant cette thèse, nous avons conceptualisé la manière d'explorer les impacts de l'utilisation des interfaces tangibles. Il s'agissait à la fois d'une exploration du paradigme tangible et d'un moyen de résoudre certains problèmes de Clone. Parmi les sujets explorés, il y a : les contraintes pratiques imposées par une telle entreprise, notre réflexion sur les concepts liés aux interfaces tangibles et la solution pour mesurer les concepts sur les participants. Dans cette section, nous explorons ces points en présentant nos réflexions, en formulant nos questions de recherche et nos hypothèses et en terminant par les différents moyens de mesure mis en place.

2.1 Direction of our research

2.1.1 Tangibility: characteristic of interfaces

Tangibility continuum

The tangible aspect of an interface is about bringing the digital world into the physical. This can be achieved to a certain degree. Looking at different interfaces, we can intuitively express their rough tangibility. For example, an interface where the users never touches anything and all his interaction are voice commands isn't very tangible. On the other hand, if a very motivated air-plane hobbyist were to recreate the cockpit of his model air-plane and connected it to his model to pilot it, this would be very tangible. Those observations lead us to view the tangibility as a continuum where we can place interfaces. One force behind this idea were the design spaces presented by the MIT team in 1995 [Fitzmaurice et al., 1995], notably is the "Interaction representations" scale from "all physical artifacts" to "all virtual artifacts". Another is the mixed reality virtual continuum proposed by Milgram and Kishino [Milgram and Kishino, 1994] which focuses on the Mixed Reality Virtuality Continuum (or MR-VC) and not the tangible one but we follow a very similar logic. The MR-VC is a scale on Mixed Reality from real environment to virtual environment where Augmented Reality(AR) is nearer real environment and Augmented Virtuality(AV) is nearer virtual environment. Through this "scale" the authors shows the diversity in MR systems.

Looking at those design spaces, the MR-VC or even Fishkin's paradigm [Fishkin et al., 2004], tangibility is even more than a simple scale, but an array of properties collected into a simple word. For example, when looking at two cases that fall within tangibility like having cube of information to manage data [Patten et al., 2001] and a duplicate model of a warehouse to teach racks organisation [Schneider et al., 2011], which is more tangible? Both have small bricks that are manipulated to be inputs, one has immediate feedback while the other wait the end to give a score but the latter has tiny racks that correspond to the virtual racks while the former has knobs to represent abstracts information. Those two example have different tangible properties and it can't be fully said that one is more tangible than the other. Meaning this placement isn't precise and completely objective as some interfaces can have very similar tangibility.

Tangibilisation: analysis

As we are looking at the tangible continuum, another question comes to mind: if there is interfaces with low tangibility and interfaces with high tangibility, can the former be converted to the latter (and inversely)? For this concept, we used the word of tangibilisation of an interface (or detangibilisation if it is from high to low). The concept of tangibilisation bring further questions. Among those: can a tangibilised interface be the same as the original beside its tangible aspect? Is it not? What does "the same" means in this context? What criterion can we use to determine this? Some of those questions were the subject of an early paper [Landi et al., 2021].

Tangibilisation: change

If we look at an example of the most elemental change we can think of like for example changing a button from blue to cyan on a panel. Does this change makes the interface completely foreign? At which point does the interface becomes so foreign it is not the same anymore? Those questions are very analogous to the classical “Theseus’ ship”. The idea of Theseus ship the Argos is the following: Theseus embarks on his quest with his ship fully intact. During his journey he has to, because of unfortunate events or simply normal maintenance, replace every part of his ship at some point of another. When he comes back to Athens is the Argos still ... the Argos? And if not, at which point has is became a new ship. There is a wide range of philosophical element to this question: is the fact that the replacement are identical in form and function important? Coming back to our analysis, is it possible to tangibilised an interface while keeping it “the same”?

Tangibilisation: characteristics

To approach this issue, our idea was that interfaces have characteristics among which is tangibility. Those characteristics can be linked together. For example an interface with detection of movement only as input will necessarily require some form of computer vision and therefore be digital in some way (without considering having someone be the detection mechanism). For the tangibilisation we explored the idea of taking a graphical interface and keeping all characteristics but the tangibility as close as possible. Obviously, interfaces have a lot of different characteristics: from the actions available, their goals, their feedback provided, the symbolism and form of the interface used... Using those idea, for our tangibilisation, we would need to identify those in our original graphical interface and recreate them in our tangible version.

Interfaces’ actions and feedbacks

For the actions, we look at the concept of affordance. J. Gibson’s concept of affordances [James J. Gibson, 1977] describes what the environment offers to the individual. This term was redefined by D. Norman to better fit the HCI context. For Norman [Norman, 2008], it describes the range of possible actions a user of the object can perceive. One of the key point explored by Norman is the importance of the “suggestion” of the possible action by the object. Rather obviously, even if an object is capable of an action but does not “suggest” it, unless explained by an outside source, the user will not be able to perform it. In a less critical way, an object that misdirect its user on how to interact with it will also make the user experience worst. This is especially central with computers where the final action can be layered behind multiple intermediary action: user moves the mouse to move the cursor to move the window.

2.1.2 A/B testing and Multivariate test

Reasoning

A/B testing central principle is using two situations which are identical but for one aspect so any differences in the measure can be imputed to this aspect’s influence.

In 1747, James Lind a British doctor of the Royal navy used the 12 men that were suffering from scurvy and divided them into 6 pairs. He gave each pair a different treatment and showed that lemon and oranges would cure the ailment. He published his findings in “A Treatise on the Scurvy” [Lind, 1772]. Nowadays A/B testing is a widely used method of testing with conference around it and is a very useful methodology, especially used in UX for example [Young, 2014]. A/B testing, also called split-testing, is a simple controlled experiment where users are randomised into two groups that are given two different product/service/test (condition A or B). The conditions being identical but for one single isolated change. This allow to observe the impact of the change on the user behaviour. While the basic testing is only done with two conditions, control and variation, the number of variation can be increased. The A/B testing becoming a multivariate test. This allows us to look at more variables but also imposes bigger overall sample size so that the sample size of each test stay high enough to be statistically relevant.

2.1.3 Metric impact from tangibility

If we see tangibility as a characteristic of an interface, it is fair to assume that depending of the tangibility of an interface, it will elicit different reaction from users. Meaning that some metric would be impacted by whether or not an interface is tangible just like any other characteristic. If we combine this though with the principle of A/B testing we comes to what became the central questions of this thesis:

- ▶ Which metric are impacted by the tangible nature of an interface?
- ▶ In what ways are they impacted?

To answer those questions we would need to do the following. First we have to identify the potentially impacted metrics. We cannot be sure that all the metrics impacted were taken into account during our analysis but we must strive to be thorough in this regard. Second find or design a way to measure the metrics when using our interfaces. The best case would be to have a recognised measure methods, even better if it is also widespread as it would give us a benchmark with which to compare both A/B conditions. Then we would need to create the two conditions, test them on similar populations and measure the metrics during those tests.

2.1.4 Constraint on the experimentation from the A/B testing method

For our purposes, our two conditions will be digital/graphic and tangible. During design, we must make sure that both conditions respect the core principle of the A/B testing: conditions are identical but for one single isolated change. If only one had an additional functionality, the impact on the results could be either due to the tangibility (or absence of) or the functionality being there or not. Conclusion would be difficult to make. Therefore those two versions should have interfaces that are as similar as each other to reduce potential bias. Digital and tangible being two different ways of interacting, this constraint will not be easy to respect at all time. Especially when starting with one version already developed. We had to make compromises and introduced some biases in order to make the tangible version

works. Those biases will be talked about further in the interpretation part of this text (Appendix 6).

2.1.5 Potential impacted metrics

Metric from the state of the art

The form of the interface will, because of the inter-connectivity of the different aspects of user experience, have an impact on a wide array of possible measurable metrics. Our first instinct, as we are not specialised in psychology, was to look at which element other studies measured and ask advice from other researchers. This returned a wide array of concepts. Some names for concept appeared multiple times with significantly different definitions while some were simply a more detailed version the other definitions.

Metrics' selection criteria

First of all, as one of our goal along analysing the impact of tangibility is to see if it solves Clone's issues. We therefore need to measure the mental demand and/or overall usability of the application. This is our first criteria: measuring Clone's improvement. Secondly, while we are not certain of which metric will be impacted or not, the state of the art gives us some good insights into what we can expect. The second criteria should be: measuring concepts the state of the art tells us should be impacted. Finally, some concepts have some overlap in their definitions, given that we realistically cannot measure everything, as a third criteria we should: try spread our measures. Those practical concerns are developed more in subsection 2.3.1.

2.2 Research questions

2.2.1 Questions of design

Introduction

In 2.1.4 we explained the main constraint of similarity of the TUI version to the GUI's original design. This need is what guided all of the development of our project. The design questions are named PQ1 and PQ2 for "Prototype Question #"

Questions

PQ1 How to "tangibilise" an already existing interface?

PQ2 Which criterion can be used to judge the similarity of two interfaces?

2.2.2 Questions of measures

Usability

The usability question is named MUQ for "Measure Usability Question".

MUQ Will the tangible version be more usable for users?

Workload

The workload question is named MUQ for “Measure Workload Question”.

MWQ Will the tangible version be more demanding for users?

User Experience

The user experience questions are named MXQ1 and 2 for “Measure user-eXperience Question #”.

MXQ1 Will the tangible version offer a better user experience?

MXQ2 Will the tangible version be a more attractive version than the numeric one?

Demographic influences

It is very much possible to have demographic influences on the results. We can very easily think that people with difficulties dealing with computers will be more receptive to the tangible version (or at least more than people without such difficulties). We also chose a somewhat broad but classic array of demographic questions on population. Those questions are named MDQ1 to MDQ4 for “Measure Demographic Question #”.

MDQ1 Is there significant differences between men and women?

MDQ2 Is there significant impact from the age on the results?

MDQ3 Does people lacking confidence in their computer literacy have a preference for the tangible version?

MDQ4 Does having used Clone in class recently impact the results?

Obviously, this is the principle behind the research. The reality is that with two conditions, splitting each of them again with demographics group makes the sample size even smaller. This reduce the statistical value of our results or forces us to have even bigger groups of participants. In addition, as our test focuses on the intended target of Clone, nursing students, they are an imbalanced in certain categories which, again, forces us to use bigger and bigger groups. For example, this very apparent for men and women as women are heavily over represented in nursing students.

2.3 Metrics measurement

2.3.1 Questionnaires’ constraints

Standardisation of data

The main concern of our data collection was its interpretation reliability. While our team is made up of a broad continuum of fields’ expert, we were lacking in the creation and validation of questionnaires. More specifically, we feared that, while we would talk about measuring some aspect of our prototype, we would be criticised if we used custom-made questionnaire. This does not cover all our gathered data

but for questionnaires, the solution was to use already existing ones. Questionnaires that are, as much as possible, used extensively by the research community. We focused on the tools of the HCI community which has numerous well-researched questionnaires about broad aspects of the user experience when using an interface. It was harder to incorporate existing knowledge on our in-game analytics as they were not only designed before this work on tangible interfaces but they are also very specific to the game and its mechanics.

Public and language

The target testing group was mainly nursing students. Obviously, unless there are circumstances, studies in France will be conducted on mostly francophone participants. If we refer a 2021 report from Education First, France #31 (out of 112 countries) has a “High English” proficiency [First, 2021]. This means that most questionnaires will need to be in French unless we exclude all people not completely proficient in English. This is an issue because most standardised questionnaires are in English. Translated questionnaires cannot be directly used as replacement; they need to be validated through a specific study to make sure they are statistically similar to the original. This means that using in-house translation would block us from one of the advantage of using a standardised questionnaire which is the ability to compare our result with a broad range of data. One of the criteria for our choice of questionnaires is therefore the availability of a French translation.

Testing time

One of our main concern was the overall testing time. Clone is not a fast game as one normal game with 3 patients usually takes around 1 hour. This constraint is reflected in our choice to stay at 3 patients (instead of going for 14) and also in our selection of questionnaires. We chose to keep it at 10 minutes maximum. In the following questionnaire presentation, we also give the times, we observed, people usually took to complete them.

2.3.2 Questionnaires

General Information

The first questionnaire is the general information category. This non-standardised questionnaire gather all non-standardised questions. Those questions’ goals is to gather demographic data about participant with the goal to see if certain population have different results. It contains the following questions: id, sex, age, year of study, scenario played, interface type (tangible or numeric), user past experience with the game (if any) and computer-use ability. All questions but id, scenario and interface are the general demographic questions. Id is to identify which answers correspond to which participant and “scenario played” and “interface” are used to split the questionnaire between version and in case we needed to test different scenarios.

AttrakDiff(FR)

AttrakDiff(FR) [Lallemand et al., 2015], a translation of AttrakDiff [Hassenzahl, 2003], focuses on user experience. It contains 28 questions, grouped in 4 categories,

and usually take 3 to 5 minutes for people to fill. Questions consist of pairs of words that are somewhat opposed and participants must answer from 1 to 7 with 1 strongly linked to the first word of the pair and 7 to the other words. For example, if the pair is “Human” and “Technical”, a user thinking that the application is not at all “Technical” will answer 1 maybe 2, and if he thinks it is very “Technical” and not very “Human”, he will answer with 6 or 7. The four categories are QP (Pragmatic Quality), QHS (Hedonic quality - Stimulation), QHI (Hedonic quality - Identity) and ATT (Evaluation construct). The Attrakdiff is available at C.1 in the appendix at the end of this document.

French System Usability Scale

The F-SUS [Gronier and Baudet, 2021] for “French System Usability Scale”, a translation of the SUS the “System Usability Scale” [Brooke, 1995], focuses on usability. The french version was analysed by its authors and was determined to be a valid version of the SUS. It contains 10 questions, answered from “Strongly disagree” (1) to “Strongly agree”(5), and usually take 1 or 2 minutes for people to fill. On LimeSurvey [LimeSurvey GmbH, 2023], for the SUS there was a mistake were, while the paragraph was correct, the header of the columns was marked with “1 (Tout à fait d’accord)” (“1 (Strongly agree)”) and “5 (Pas du tout d’accord)” (“5 (Strongly disagree)”). This error was pointed out to all participant before they filled the questionnaire (as well as during just in case). The F-SUS is available at C.2 in the appendix at the end of this document.

NASA-TLX-FR questionnaire

The NASA-TLX questionnaire measure the subjective mental workload along 6 dimensions with the same number of questions. Each question is answered between 0 to 100 in increment of 5, and usually take 1 or 2 minutes for people to fill. It measures the experience of workload when performing an activity. The metrics are:

- ▶ **Mental Demand:** how much thinking, deciding, or calculating was required to perform the task.
- ▶ **Physical demand:** the amount and intensity of physical activity required to complete the task.
- ▶ **Temporal demand:** the amount of time pressure involved in completing the task.
- ▶ **Effort:** how hard does the participant have to work to maintain their level of performance?
- ▶ **Performance:** the level of success in completing the task.
- ▶ **Frustration level:** how insecure, discouraged, or secure or content the participant felt during the task.

It must be noted that the current NASA-TLX is not the earliest one. This one with 6 questions, published in 2006 is an updated form of a the scale published in 1988. The new version has a french validated translation [Cegarra and Morgado, 2009], the NASA-TLX-FR. The NASA-TLX-FR is available at C.3 in the appendix at the end of this document.

Game Experience Questionnaire

The GEQ, for Game Experience Questionnaire, is a modular questionnaire split into three sections. The sections are “Core questionnaire”, “Social Presence Module” and “Post-Game module”. The core part assesses five metrics: Immersion, Flow, Competence, Affect (negative and positive), Tension and Challenge. The social part focuses on relation with other player (e.g. “I empathised with the other(s)”). The post game enquire about how player felt right after finishing playing (e.g. “I felt energised”). The sections have, in order, 33, 17 and 17 questions for a total of 67. While multiple versions of the questionnaire exist, one of them can be found on the authors’ university website [Wijnand A. IJsselsteijn et al., 2013]. The GEQ, even if a bit long, is a very broad tool, and as we’ll see right after it has been used by numerous people, which would make it an interesting addition to our questionnaires. However, in their review of 73 publications [Law et al., 2018] are quite critical of some of the GEQ’s aspects. First they complain about the multiple versions lacking a clear final version. They also criticised the questionnaire validity and consistency even concluding “the psychometric properties of [the GEQ] are yet to be fully established”. For those reason, we have chosen to not use the GEQ for our study. The Game Experience Questionnaire, here referred as GEQ, should not be mistaken for the Game Engagement Questionnaire, here referred as GEQ2.

Game Engagement Questionnaire

The GEQ2, for Game Engagement Questionnaire, is a 19 questions questionnaire which focuses on engagement. Presented in [Brockmyer et al., 2009], it focuses on 4 metrics and uses the following definitions for them:

- ▶ **Psychological absorption:** “total engagement in the present experience”
- ▶ **Flow:** “feelings of enjoyment that occur when a balance between skill and challenge is achieved in the process of performing an intrinsically rewarding activity” [Cziksentmihalyi, 1990]
- ▶ **Presence:** “being in a normal state of consciousness and having the experience of being inside a virtual environment”
- ▶ **Immersion:** “the experience of becoming engaged in the game-playing experience while retaining some awareness of one’s surroundings”

User Engagement Scale - Short Form

The UES-SF, for User Engagement Scale - Short Form, is, as its names indicate, a short form of the User Engagement Scale or UES for short. The UES [O’Brien and Toms, 2010] was refined from the original 31 questions to 12 in the UES-SF [O’Brien et al., 2018] while keeping the validity of its measurement. User engagement was one of the element considered during our preliminary reflection about the impact of tangibility. However, while not completely redundant with the other selected questionnaire, we chose not to include it to reduce the length of the experience and focus on the other metrics.

2.3.3 In-Game data

Triggered log

Clone measure data in multiple ways. One is however the most central: triggered log. In Clone's code, when certain events happen it triggers a log send to the database. There is a lot of events tracked within Clone, basically most user inputs, from opening the care plans, adding activities to them and game events. Depending on the triggered log, different information are logged.

Saved schedules and step saves

Clone saves the game state during play. The main save simply keep the last schedule made by the player. There is also a step saves done when players change steps. This is done because most objectives can be failed when getting in a new step so if the player fails, he can reload to the nearest possible point without losing much progress.

Tangible and digital version

With the idea of the tangibilisation of the digital interface, we aim to keep game elements as similar as possible. This means that, except in rare cases, most trigger will work regardless of the version. In fact, we only added 6 logs compared to the 40+ already in place. The logs added were for 2 reasons. Either there was a new, tangible-only event that required it, like the special actions. Or, there were some difference in the behind-the-scene inner working of the code that meant that some trigger were skipped in the tangible version and had to be re-added, like schedule opening.

Synthesis The goal of our work is to measure the impact of TUIs. The central idea developed for our work is to create a pair of interfaces, one GUI and one TUI. Then, those interfaces would be tested on their intended users and we would measure a wide array of metrics. The metrics measured are selected using their expected link with tangibility. They are “Usability”, “User eXperience” and “Mental workload”. By having both those interfaces as similar as possible but for the tangibility, our goal is to see the influence that having or not a tangible interface has on the metric measured. This research direction identified, we are led to an analysis of the state of the art with the goal of better understanding the systems behind Tangible User Interfaces, scheduling and serious games.

Synthèse L’objectif de notre travail est de mesurer l’impact des TUIs. L’idée centrale de ce travail est de créer une paire d’interfaces, l’une graphique et l’autre tangible. Ensuite, ces interfaces seront testées sur les utilisateurs cible et nous mesurerons un large éventail de paramètres. Les paramètres mesurés sont sélectionnés en fonction de leur lien attendu avec la tangibilité. Il s’agit d’usabilité, d’expérience utilisateur et de charge cognitive. En ayant ces deux interfaces aussi semblables que possible, à l’exception de la tangibilité, notre objectif est de voir l’influence qu’à, sur les paramètres, le fait d’avoir ou non une interface tangible. Cette direction générale identifiée, nous sommes conduit à une analyse de l’état de l’art dans le but de mieux comprendre les systèmes qui sous-tendent les interfaces tangibles, la planification et les jeux sérieux.

Chapter 3

State of the Art

Introduction With the direction of analysing the impact of tangible interfaces we now look at the work of others on the subject to see their methods, the concepts proposed and the theoretical frameworks already developed. This state of the art is divided into the main subject that are the central themes of our research: scheduling, serious games and TUIs. In parallel of that theoretical work, we are also looking at the practical applications of those themes, and most importantly on technologies that are used in TUIs in order to inform our future tangible design.

Introduction Dans le but d'analyser l'impact des interfaces tangibles, nous nous penchons maintenant sur le travail d'autres personnes sur le sujet afin de voir leurs méthodes, les concepts proposés et les cadres théoriques déjà développés. Cet état de l'art est divisé selon les principaux sujets qui constituent les thèmes centraux de notre recherche : la planification, les jeux sérieux et les TUIs. Parallèlement à ce travail théorique, nous nous penchons également sur les applications pratiques de ces thèmes, et surtout sur les technologies utilisées dans les TUIs afin d'éclairer notre futur design tangible.

3.1 Planning and Scheduling

3.1.1 Defining scheduling concepts

Defining Scheduling

Scheduling (and sequencing) is at the centre of our industrialised world. It is used in all industries from manufacturing to services. There exist multiple definitions for scheduling (among others [Mellor, 1966], [Pinedo, 2016], [Agnētis et al., 2014] or [Blazewicz et al., 2019]). We propose to use a fusion of some of them “Scheduling is the decision making process of sequencing, usually by start-time, tasks and allocating them resources, especially processors/machines, in order to optimise certain objectives”. In this definition there are 3 central elements : the tasks, the resources and the objectives.

- ▶ **Tasks** the activities sometimes also referred as jobs. While they can be grouped or split into depending on the situation. Usually, the best is to group all task that will be done in sequence and that require the same resources. For example: make 20 machinery pieces, box them into an appropriate container, load them into a cargo truck, etc.
- ▶ **Resources** broad term for all the things needed to do the tasks. They are sometimes referred as machines but also includes personnel. For example: a machine to make a machinery piece, its operator, its input, the delivery system, etc.
- ▶ **Objectives/Constraints** rules that will need to be followed which exceed the simple task-resources dynamic. For example: security rules, shift hours, deadlines, etc.

Scheduling modeling

Like any other science, scheduling uses models. Models, being a representation of reality under a certain set of rules, will simplify what is a more varied reality. This is particularly useful for common problem that fit well within some models’ rules (e.g. job shop scheduling problems which has an extensive history of heuristics [Mellor, 1966]). However more complex situations require more complex models (e.g. taking into accounts incidents, more than 1-1 relation for machines and jobs, rolling deadlines, etc) which usually requires more complex heuristics. Those heuristics can be groups in term of skills need to be able to schedule correctly. Given that the chosen heuristics depend on the models which depend on the complexity of the subject, we will now focus on the skill set used during our study.

3.1.2 Skills set associated with scheduling

Knowledge, know-how and people-skills

Like for most skills sets in professional environment, we can categorise skills associated with scheduling along 3 groups : knowledge (“savoir”), know-how(“savoir-faire”) and people-skills(“savoir-être”). In this categorisation, knowledge contains all theoretical understandings and insight acquired through school or professional experiences. Know-how is all the practical and applied knowledge. Finally, people

skills are about the capacity of working with other people and within groups of similar or dissimilar persons.

People-skills in scheduling

People-skills includes :

1. knowing how to deal with stress
2. knowing how to be efficient and reactive
3. knowing how to discuss and persuade
4. having a broad understanding of the perimeter of the activity
5. being able to say no

Knowledge skills in scheduling

Knowledge includes :

1. knowing the state of the system being scheduled for and its workload capacities
2. knowing the tasks to schedule and their requirements
3. knowing how to evaluate the sizes of the organisation
4. knowing being able to analyse the system
5. knowing legal obligations
6. knowing being able to control the tasks' materials and tools needs
7. knowing how to model the organisation processes
8. knowing how to schedule a workload
 - ▶ design work
 - ▶ production work

Scheduling being a skill set that is applied to a specific subject, the above group is either about general scheduling (e.g. "how to schedule a workload") or about those specific to the subject (e.g. "legal obligations").

Know-how skills in scheduling

Know-how includes :

1. mastering the flexibility of scheduling to deal with the unforeseen
2. mastering big data
3. knowing how to solve acute and complex issues
 - ▶ through prioritisation
 - ▶ through identifying bottlenecks
4. knowing how to communicate with other parties involved
5. knowing how to create directives for employees
6. knowing how to plan work schedules
 - ▶ define activities
 - ▶ schedule activities
7. modeling the organisation so it can be improved

Links between the cognitive process and the scheduling skills set

As said in the subsection 1.2.7, the cognitive process is closely links with the scheduling skills set. This is quite understandable as Clone focus is scheduling, as such, students should and do need to apply scheduling to solve Clone’s scenarios. if we look at the two of them along one another we can the see the link. Knowledge of the patient files and patient current state fit mostly in “knowing the state of the system (knowledge 1)”. Being able to interpret the files using the practices and he ability to convert that understanding into care acts fit in multiple knowledge categories like “knowing how to analyse the system (knowledge 3)” and “knowing legals obligations”. Planning the care plan is first and foremost about “knowing the tasks to schedules and their requirements (knowledge 2)” and “having a broad understanding of the perimeter of the activity (people skills 4)”. The validation of the rough care plan need multiple skills with “knowing how to schedule a workload (knowledge 8a and 8b)”, “knowing how to plan work schedules (know-how 8a and 8b)” and “knowing how to solve acute and complex issues (know-how 3a and 3b)”. Finally being able to deal with the unforeseen, along with the previous skills require “mastering the flexibility of scheduling to deal with the unforeseen (know-how 1)”.

3.1.3 Scheduling in healthcare

Introduction

Scheduling in healthcare, like in a lot of industries is a critical issue that is central to the effectiveness of any organisation of the field [Abdalkareem et al., 2021]. The issues encountered in healthcare are no completely dissimilar with other industries. However, healthcare being a major industry, its issues are very much studied about. One classical scheduling problem that has had much research on it is the “Nurse Rostering Problem”(NRP) also called “Nurse Scheduling Problem”(NSP) [Cheang et al., 2003]. This problem aims at providing optimised schedules for nurses. This top down affectation, not unlike assimilating nurses and patients to workers and jobs is not what we are focusing on in this work. But the understanding that is being conveyed with Clone shares a lot with the ideas behind those problems.

Nursing scheduling skill in France

Teaching scheduling in health require, like most other teaching require to take a step back and analyse the subject [Landi et al., 2022b]. Presented further in [Galaup et al., 2021], the OCIS skill block is one of the necessary skills references of the Registered Nurse’s diploma. It contains the following skills :

1. Identifying the actors working with people (health, social, medico-social, as-sociations, etc.)
2. Organising his interventions taking into account the limits of his professional field and responsibilities, ensure continuity of care by calling on other skills
3. Choosing the tools for transmitting information adapted to the partners and situations, ensuring both implementation and effectiveness
4. Coordinating actions and care on the person being cared for with the various health, social and home care actors
5. Cooperating within a multi-professional team in order to optimise health and medico-social care

- 6. Coordinating the processing of information provided by the different actors in order to ensure the continuity and safety of care
- 7. Establishing and maintaining links with the actors, networks and structures working with people
- 8. Organising work in the different modes of nursing practice, particularly in the private sector

Comparing the OCIS skill block with the scheduling skill set

This cross analysis is the subject of a publication [Landi et al., 2023] that gives more details than what we are going to present here. In Table 3.1’s last row and last column, we have the totals for each of the respective column and row. This first thing to note is that there is no value at 0. This indicates the full coverage of both systems with each other. There is no missing element between the two. Being focus on scheduling, it is with no surprise that the OCIS’s 8th item, focus most importantly on organising work, is the one with the most representation in the scheduling skill set.

	A People-skills					B Knowledge skills								C Know-how											
	1	2	3	4	5	1	2	3	4	5	6	7	8a	8b	1	2	3a	3b	4	5	6	7a	7b	8	
1	0	0	1	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	1	0	0	0	0	1	4
2	0	0	0	1	0	1	1	0	1	1	0	1	0	1	0	0	1	1	0	0	0	1	1	1	12
3	0	1	0	0	0	0	0	0	0	0	0	1	0	0	0	1	0	0	1	1	0	0	0	0	5
4	0	0	0	1	0	1	0	0	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	17
5	0	0	0	1	1	0	1	1	1	0	0	1	0	0	0	0	0	0	1	0	0	0	0	1	8
6	1	1	0	1	0	0	0	0	1	0	0	1	0	0	0	1	1	0	1	1	0	0	1	1	11
7	0	0	1	1	0	1	0	1	1	0	0	1	0	0	0	0	0	0	1	1	1	1	1	1	12
8	0	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	21
	1	2	2	6	2	4	3	3	7	2	2	7	2	3	2	4	4	3	7	5	3	4	5	7	

Table 3.1: Table of the OCIS skills block with the scheduling skills set showing the overlap between the two. The rows are each of the OCIS’s skills. The columns are the categorised scheduling skills. The last row and columns show the total for each of their respective column and row.

3.2 Serious Games

3.2.1 From games ...

In their book “Rules of Play” [Salen and Zimmerman, 2004], the authors presents an array of definition for the concept of games. Having extracted the characteristics of each of them they offer their own definitions “A *game* is a system in which players engage in an artificial conflict defined by rules, that results in a quantifiable outcome”. This definition is very broad in scope and very dense in concepts. We note the importance of the relation between the players who interact with the system and its rules.

Video games fit this definition. They simply also require a computer (or any electronic machine) as a physical object by which the system is managed (at least in part). This definition also use the “quantifiable outcomes” concept to talk about

the results of the game within its rules like who won and lost for example. This idea of outcomes is solely within the rules of the games and does not include external benefits (or losses) like, for example, useful learning or simply entertainment.

3.2.2 ... to serious games ...

If we look at those outcomes as goals that the game aims to accomplish, we can classify games in two categories. The bulk of video-game being produced by the 117 billion of dollars (in 2019 [Palma-Ruiz et al., 2022]) that have entertainment as their primary goal and the rest.

Aristotle's view on game used for teaching in Politics are not the most generous : "Now obviously youths are not to be instructed with a view to their amusement, for learning is no amusement, but is accompanied with pain." (from Benjamin Jowett's translation Book 8 part 5 [Aristotle, 1999]). While this negative opinion on mixing the serious learning with recreational games is even today alive and well, we also see that the opposite idea is not new. The oxymoron "Serio Ludere" appears in the early Renaissance to express a similar idea of "playing seriously"[Alvarez, 2007]. In even more recent publication, Clark Abt coined the term for serious games linking serious purposes with, here tabletop games [Clark C. Abt, 1970]. Nowadays it is clear that the motivating power of games can be used to further other goal than just entertainment. The exact domain of serious games is still debated and there exist multiple definitions of it, but the main definition is just that : a video game which primary purpose isn't entertainment [Djaouti et al., 2011]. Serious games accomplished those purposes by utilising the codes and practises and exploiting the intrinsic motivational [Malone et al., 1981, Cordova and Lepper, 1996, Fenouillet et al., 2009, Habgood and Ainsworth, 2011, Robert T. Hays, 2005] of video games. As pointed in [Alvarez and Djaouti, 2011], video games can also be used for serious purposes but serious games are designed for it. Being defined by what their primary purposes is not, serious games encompass a broad range of "other purposes"; while teaching is the most common, some even equating serious games to learning games [Whyte et al., 2015], they can also be used for therapy, public campaign (activism, religion, political or advertisement) or scientific research... In terms of reason for playing, a meta-analysis [Hamari and Keronen, 2017] found that while enjoyment was strongly correlated with playing intention. The main difference between "hedonic"(here general entertainment games) and "utilitarian"(here serious games) was on perceived ease of use and perceived usefulness where hedonic games had a significantly lower correlation than utilitarian ones.

3.2.3 ... and to learning games

Children's learning

The domain of EIAH, for "Environnement Informatique pour l'Apprentissage Humain" (Informatics environment for human learning), is a very broad field that also includes serious games that aims at teaching players. In his thesis [Galaup, 2013] presents Piaget's work [Jean Piaget, 1974] on the parallel between children's (and teenager's) psychological development and the type of game played. The first are the workout games that help develop the child's sensori-motor skills. After 2 years old, games become more complex with symbolic games like imitation, acting, drawing

or imagination. Then starting around 6, the games gain rules, developing autonomy and cooperation with others. Strategies starts to appears.(Table 3.2). Games are part of learning from a very young age and this continues even after the end of children development. Learning games have since had a long history [Alvarez, 2007]. One of its early and well know title being “The Oregon trail” in 1971 by American university teachers to teach American history [Rawitsch, 2017] that even got a new version in November 2022 on PC.

Game Type	Age	Function	Examples
Workout games	0 to 2	Sensori-motor development	Object discovery
Symbolic games	2 to 6-7	Reality assimilation	Mimicking, drawing
Games with rules	6 to 12	Autonomy and cooperation development	Board games or team games

Table 3.2: From [Galaup, 2013] Children development linked to game types

3.2.4 Benefits of games for “serious” purposes

Motivation

Firstly, like we talked above, one of the main driver behind using games for non-entertainment goal is the intrinsic motivation to play that they bring. Once the player is doing the activity, here playing, we can use that interaction to also accomplish our goal. Malone in [Malone et al., 1981] present elements that motivate players in games : having a appropriate and understandable challenge, be an embodiment of a fantasy for the player, satisfy the curiosity, need for novelty of the player. Those elements are particularly well fitted for video games where there is already a need for a challenge, the possibility through a computer to give easily new experience that could create the fantasy sought by players.

Flow

Another benefit can be found with the theory of “Flow” from [Cziksentmihalyi, 1990, Sherry, 2004, Sweetser and Wyeth, 2005]. It presents a psychological state of mind where the player is absorbed by and focused on the activity with nothing around him mattering. One of the way to achieve this state in game is to balance the challenge of the tasks with the skill of the player. To much challenge leads to frustration while not enough leads to boredom. Computers can, if the software is designed for it, balance the experience to better achieve this states. In addition this can be done on a case by case basis if users have one machine each.

Scientific research

Games can also take advantage of the massive following they can gather and in our age of global communication this can be leverage in some interesting direction. Some games use games capacity to have people help scientific research in new ways. For example, Foldit is an online puzzle game about folding molecules. Described as an “interactive graphical interface of the Rosetta molecular modelling

package” [Kleffner et al., 2017]. Its aim is to find new protein designs by trying to fold them correctly. Foldit is available in multiple forms including online at Fold.it and reported 240000 registered players in 2012 [Marshall, 2012]. There is also Zooniverse[Zooniverse, 2023a] is a platform for “people-powered research” with 101 projects. Volunteers help by answering questions about images presented to them. This create data-sets for researcher. Zooniverse list 428 papers published (and data-sets) in a broad range of fields such as astronomy, physics, humanities [Zooniverse, 2023b].

Medicine

Games can also serve a a medical apparatus. Motivation is central in rehabilitation which makes serious games a good solution especially with children. For example, in this thesis [Anne-Laure Guinet, 2022] the author created a serious games to help during the rehabilitation of children with cerebral palsy.

3.3 Tangible User Interfaces

3.3.1 Origin and early development of the concept

In the field of Human-Computer Interface (HCI), the dominant way of interacting is done using a mouse and keyboard often named WIMP-GUIs for Windows, Icons, Menus, Pointer Graphical User Interfaces invented in the 70s by Xerox and now ubiquitous in most mainstream computer systems. While very useful, they do have their limitation and there exist other interface paradigms (and even a “post-WIMP” movement) [van Dam, 1997]. One of them is the Tangible User Interfaces, or TUIs. TUIs use the grasping of everyday objects and their manipulation to serve as inputs in the interface. While earlier development of such interfaces existed before, TUIs were first conceptualised in the seminal paper “Bricks: Laying the Foundations for Graspable User Interfaces” [Fitzmaurice et al., 1995]. In this paper, there were called “Graspable User Interfaces” but Graspable would be replaced by Tangible two years later at CHI 1997 [Ishii and Ullmer, 1997]. Bricks’ paper is the result of works by the Tangible Media Group at the MIT Media Lab. This team regroup a lot of the most prolific authors around the TUIs. Among them George Fitzmaurice with both the original paper for TUIs [Fitzmaurice et al., 1995] and a first experiment to test the usability interfaces with different degree of tangibility [Fitzmaurice and Buxton, 1997], Hiroshi Ishii with an exploration of the possibilities of TUIs [Ishii and Ullmer, 1997], and Brygg Ullmer with a framework for TUIs [Ullmer and Ishii, 2000], his follow up thesis furthering his framework [Ullmer, 2002] and another paper focusing on a specific part of TUIs [Ullmer et al., 2005].

3.3.2 Defining tangible interfaces

The concept of tangible interfaces has been described and/or defined in a multitude of ways :

- ▶ “We introduce the concept Graspable User Interfaces that allow direct control of electronic or virtual objects through physical handles for control.” [Fitzmaurice et al., 1995]

- ▶ “Graspable User Interfaces [2, 4] advocate providing users concurrent access to multiple, specialized input devices which can serve as dedicated physical interface widgets, affording physical manipulation and spatial arrangements. These physical artifacts are essentially a collection of input devices that can be tightly coupled or “attached” to virtual objects for manipulation or for expressing action” [Fitzmaurice and Buxton, 1997]
- ▶ “The goal of Tangible Bits is to bridge the gaps between both cyberspace and the physical environment” and “TUIs will augment the real physical world by coupling digital information to everyday physical objects and environments” [Ishii and Ullmer, 1997]
- ▶ “Tangible user interfaces are broadly concerned with giving physical form to digital information. At the highest level, there are two basic facets of this approach. First, physical objects are used as representations of digital information and computational operations. Secondly, physical manipulations of these objects are used to interactively engage with computational systems.” [Ullmer, 2002, p. 16]
- ▶ “[TUI’s] stated research goal was to extend computation beyond traditional Graphical User Interface (GUIs) to interfaces that use physical objects and the physical world as tangible embodiments of digital information.” [Pangaro and Sociology, 2003]
- ▶ In their literature review [O’Malley and Fraser, 2004], in addition to other description found here, they present tangible computing as a part of ubiquitous computing and link it to augmented reality.
 - ▷ **ubiquitous computing:** O’Malley and Fraser presents Weiser paradigm from his article “The Computer for the 21st Century” [Weiser, 1993] : “[Weiser] talked about a vision where the digital world blends into the physical world and becomes so much part of the background of our consciousness that it disappears”.
 - ▷ **augmented reality:** O’Malley and Fraser presents Milgram and Kishino’s Virtuality Continuum scale from their “A taxonomy of mixed reality visual displays” [Milgram and Kishino, 1994]. The review authors point out that “many examples of so-called tangible computing have employed AR techniques”

They conclude with “the distinction between ubiquitous computing (or ‘ubicomp’), tangible technology, augmented reality” and ‘ambient media’ has been blurred or at least the areas overlap considerably.

- ▶ In their taxonomy [Fishkin et al., 2004], Fishkin’s team use the following “broad script that characterizes TUIs”:
 1. Some input event occurs. This input event is typically a physical manipulation performed by a user with their hands on some “everyday physical object,” such as tilting, shaking, squeezing, pushing, or, most often, moving. Later, we will remove the “typically” qualifier.

2. A computer system senses this input event, and alters its state.
3. The system provides feedback. This output event is via a change in the physical nature of some object it alters its display surface, grows, shrinks, makes a sound, gives haptic feedback, etc.

Through out those definition we see that the concept is still fuzzy but with very central elements : physical objects, physical manipulation as input, embodiment of the target of the action in the object manipulated, proximity of the feedback given to the object manipulated. For this document, we will define tangible interface as “Interfaces where the user interacts with the virtual environment through the affordances of physical objects that metaphorically represent the virtual environment or its elements.”

3.3.3 TUI within frameworks / Paradigms

Multimodal human machine interaction op

The multimodal human-machine interaction loop (see Figure 3.1) is based on Nigay’s Pipe-Lines model [Laurence Nigay, 1994] itself based on Norman’s human action cycle [Norman, 2008]. In this model, the human starts with an intention or goal which are translated into actions captured by the machine through devices like keyboards, microphone, etc and transformed into data. This data is computed within the model by the software and makes the machine send back a feedback which is perceive by the human which interprets its and adjust his intention accordingly. This mirror cycle of decision - action - capture - interpretation fundamentally describes human machine interactions.

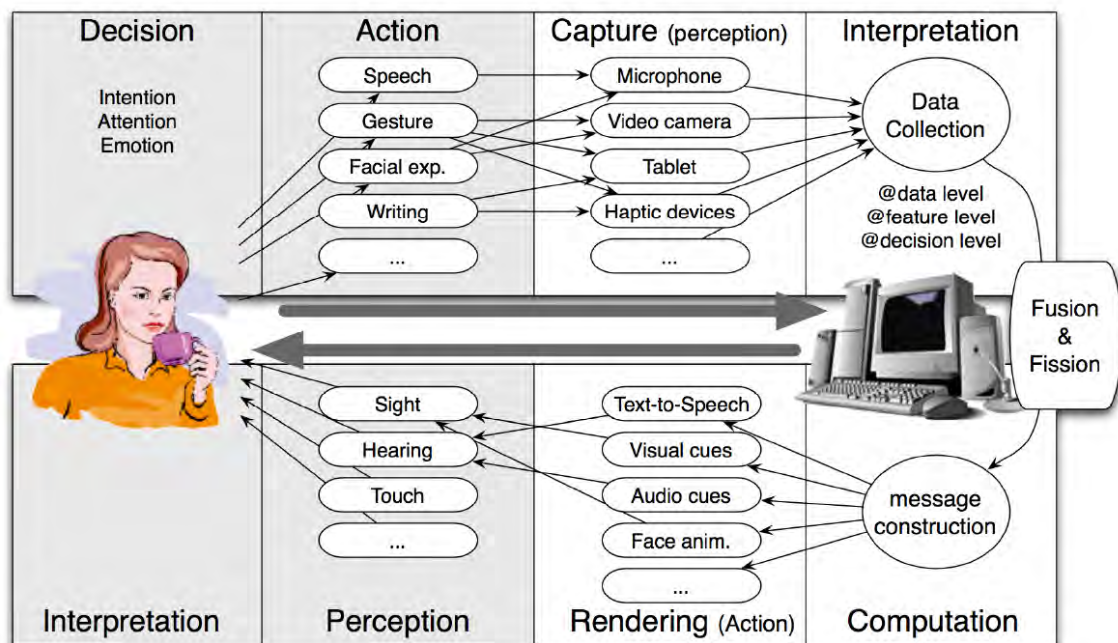


Figure 3.1: From [Dumas, 2010, p. 34], the multi modal human machine loop

Design space for Graspable User Interfaces

In their 1995 paper [Fitzmaurice et al., 1995], the MIT team offers a list of design space for what was at the time their new interface type. Each of those design space are scales on which to place the TUIs. A direct example is the number of bricks involved simultaneously. The full list can be seen in Figure 3.2.

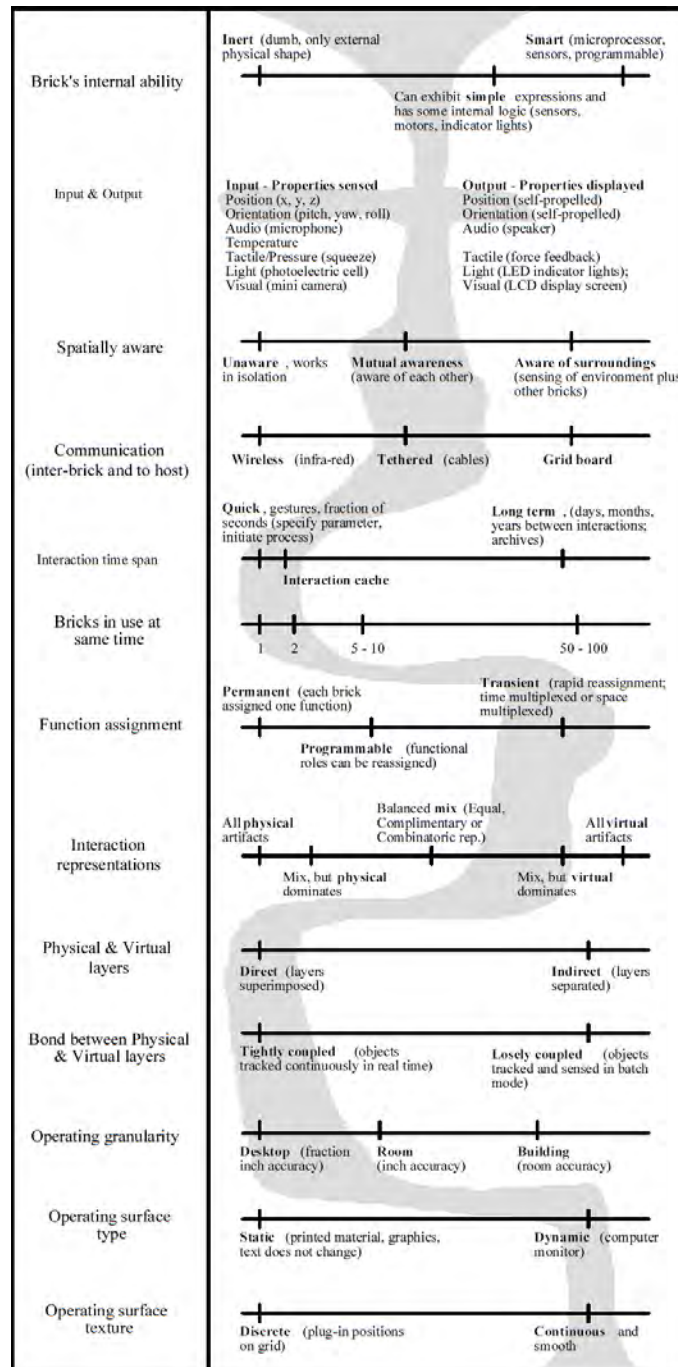


Figure 3.2: From [Fitzmaurice et al., 1995], the design spaces presented in on of the first paper written about TUIs

Fishkin's Taxonomy

Fishkin [Fishkin, 2004] also notes that definitions of tangible interface are broad. The solutions he offers two axis on which we can put tangible artifacts :

- ▶ Embodiment : represent the proximity between the output and input element
 - ▷ Full : the same object is both output and input
 - ▷ Nearby : touching or very near
 - ▷ Environmental : around the input/user, output input are tenuously linked
 - ▷ Distant : input and output are not linked but are somewhat near, like in same room/space

- ▶ Metaphor : represent the metaphorical link suggested by the physical object
 - ▷ None : self evident
 - ▷ Noun : it is the shape of object that creates the metaphor
 - ▷ Verb : the interaction with the object creates the metaphor
 - ▷ Noun and Verb : both previous category apply
 - ▷ Full : the virtual system is the physical system

In this paradigm we see that Fishkin apply the feedback loop of the general human-machine paradigm. He focuses on the link between the object through which the action is done, the input, and the system response, the output. In the tangible, the focus is on the link of the virtual with the physical. For him, the important part of this link is either spatial with the embodiment or metaphorical.

This is because of the importance for him of the cognitive dissonance in the concept of TUI.

Those categories are interesting to compare to Fitzmaurice's design spaces, specifically "Bond between Physical and Virtual layers". In their explanation, for the MIT teams this is mostly for the timely synchronisation more than the metaphorical link.

Ullmer's Token + Constraint

In his thesis titled "Tangible Interfaces for Manipulating Aggregates of Digital Information"[Ullmer, 2002], Brygg Ullmer explores the aggregation of tokens and slots; Aggregation that creates meaning. He develops a framework of "two kinds of physical/digital artifacts: tokens and constraints" thereafter named Token+Constraints frameworks (or T+C). Tokens represents their elemental digital counterparts while constraints serves as slots to put the tokens into. The framework gives us very varied tokens and constraints and tokens+constraints interactions. Tokens can be all the same, different, can be constraints (nested token), ... Constraints can have one slot, multiple slots, slots that takes specific tokens, slots where the token can be manipulated,... the interaction can simply be "Is there a token in the slot ?", their position or order in the slot can be important, the combination of tokens can be important,... This framework offer a wide scope of possibilities and some properties to classify the tokens, the constraints and their interactions.

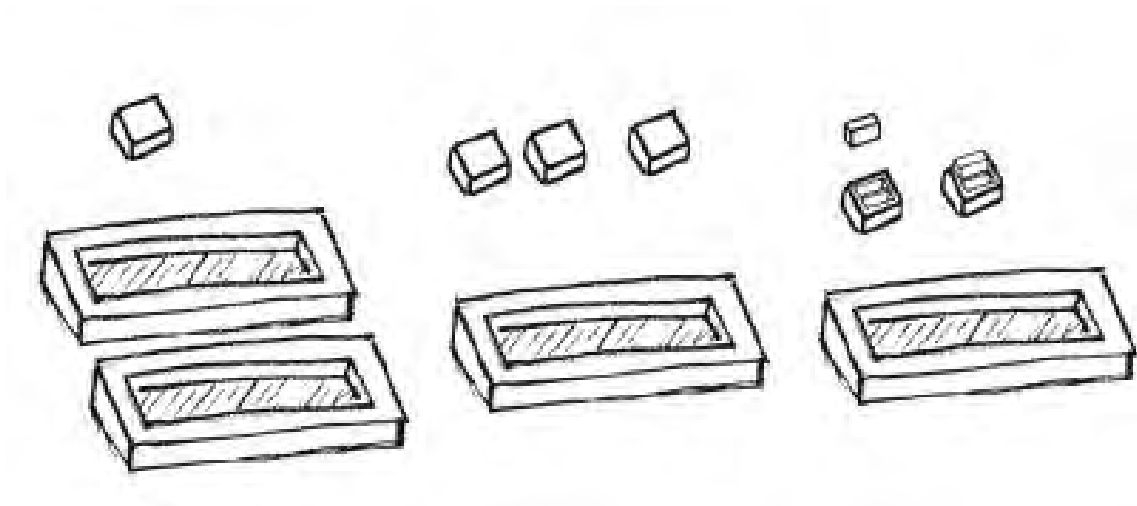
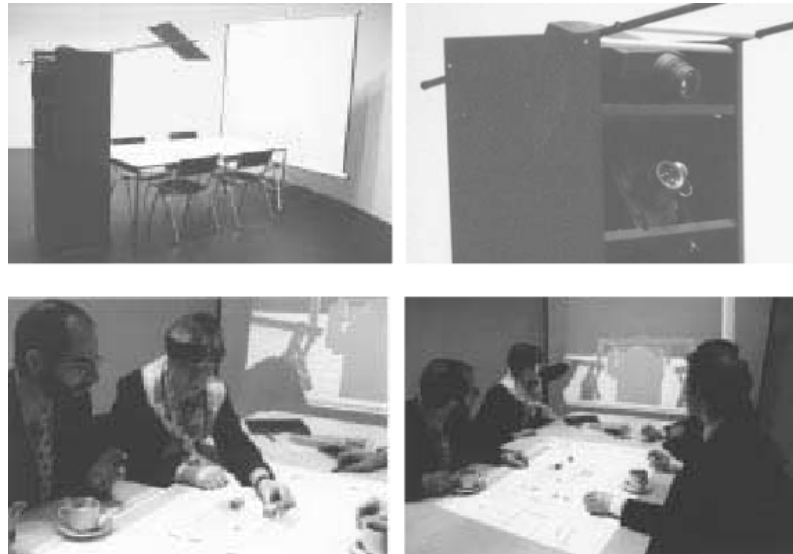


Figure 3.3: From [Ullmer, 2002], it shows 3 possible combinations among others for the T+C. (Left to right) One token with multiple separate constraints, multiple tokens in a single multi slot constraints and nested tokens/constraints

3.3.4 A wide array of tangible projects

- ▶ Sensetable [Patten et al., 2001]: a table that electromagnetically tracks tokens and gives a feedback using projected images.
- ▶ Tangible Puzzle [Xie et al., 2008]: a puzzle for 7 to 9 years old with a tracking camera and patterns on the pieces (“the markers were distributed along the edges of intersecting pieces. No one puzzle piece had an entire pattern”). The study enjoyment and engagement between multiple version of the game.
- ▶ POMA [Al Mahmud and Soysa, 2020]: an object recognition game for ASD (Autism Spectrum Disorder) 5 years old. The researchers used foam object of fruits, vegetables or animal and asked the children to recognise the object shown on an iPad and put the correct object on the screen.
- ▶ BUILD-IT system [Rauterberg et al., 1998][Fjeld et al., 1999][Fjeld et al., 2001]: a collaborative planning tool that uses computer vision and AR (see Subfigure 3.4(a)). The tool allow multiple people to work collaboratively using the bricks system for input and projection as feedbacks.
- ▶ Envisionment and Discovery Collaboratory of L3D [Arias et al., 1997][Fischer, 2001]: an application that uses tangible interface to create a context aware interface (see Subfigure 3.4(b))
- ▶ Facet-Streams [Jetter et al., 2011]: a multi-touch tabletop using tangible tokens which is compared to a web interface (see Subfigure 3.4(c)). On task performances, the results indicate a slight advantage for the hybrid interface over the standard web ones but they were not statistically significant. Usability and user experience results were also similar, the author still noting the relative easy of learning the new interface. In this study, it seems that tangible interface have relatively similar performances than multi touch one for collaborative tasks.
- ▶ Customs route for museum visitor [Rey, 2020]: a thesis work on the design and use of tangible interfaces in museums environments.

- ▶ JamaicAR [Thevin et al., 2021]: a board game using tangible technology to adapt it for the visually impaired



(a) BUILD IT [Rauterberg et al., 1998]



(b) EDC Environment



(c) Facet Streams [Jetter et al., 2011]



(d) MuseoTUI [Rey, 2020]



(e) JamaicAR [Thevin et al., 2021]

Figure 3.4: Tangible interfaces examples' images

3.3.5 Promises and drawbacks

Introduction

TUIs encompass a very wide array of interfaces as being tangible is but one of the characteristic of such interface. That diversity means that it is quite hard to precisely determines the benefits of TUIs as most studies are always done with a specific benefit in mind and on a very specific interface.

Epistemic actions

In his thesis [George W. Fitzmaurice, 1996] Fitzmaurice present the distinction between epistemic and pragmatic actions. Pragmatic actions are the regular action we think of, “whose primary function is to bring the user physically closer to the goal”. Epistemic actions are “performed to uncover information that is hidden or hard to compute mentally”. In those concept he extends the concepts of exploratory and performatory action [Gibson, 1962]. An example given of epistemic action is physically moving a chess piece without actually completing the move. The availability of epistemic actions should help users with their mental computation by reducing the memory and the steps memorise needed as well as the error from remembering. The idea, put forth by Firtzmaurice and other authors [Kim and Maher, 2008][Antle et al., 2009] [Chandrasekera and Yoon, 2015] [Kim et al., 2019] is that tangible interfaces produce more epistemic actions which in turn alleviate the cognitive load on the user improving his performance.

While the Firtzmaurice and Gibson are focused on movement and gestures, the idea of previsualisation of results and reducing the cognitive load of doing all the steps in your head before actually choosing to do the action proper is also spread in video game UX design. Online chess games like “chess.com” gives players the ability to do this move preview even in video game form. A wide array of games with a lot of value also preview the results of players choices to help reduce calculation without the players having to stop and do them or instinctively guess.

3.3.6 Spatial cognition

In their article [Kim and Maher, 2008], continuing their work available in [Maher and Kim, 2005], the authors look at the impact of TUIs on spatial cognition, defined as “reflective interaction between the external representation and the internal representation of the problem–solution processed by the perception and reasoning about visuo-spatial information”. After the introduction of other works [Ma et al., 2003][Fjeld et al., 2001][Lee et al., 2009] using TUIs in that context, the authors conclude that there is a lack of evaluation which is the goal of their enquiry. Their study shows multiple element about TUIs. Epistemic actions are dominating TUIs with a lot of focus shift observed and perceptual actions in that condition. In that condition, user put a stronger accent on the spatial relation among UI’s elements and were able to use new visio-spatial features. In terms of activity performances, the foals of producing designs was better accomplished in the TUI condition, with a note that time was more allocated to the actual design problem.

Cooperative and collaborative use

In their paper [Hornecker and Buur, 2006], Honecker and Buur present the application of tangible interfaces for cooperative designs in the context of the bricks developed by the MIT team. For her, TUIs has intuitive manipulation through the brick form which eases first access. She cite Ehn’s Scandinavian design book [Ehn, 1993] which notes the benefits of participatory design where both designers and users learns from each others and the importance of design by doing. Honecker also presents a relationship graph a key factors an characteristics that pertain to the benefits brought by TUIs during collaboration. The graph is available at Figure 3.5.

It show the two key characteristics “physical spatiality” and “haptic directness” on top, the “enabling factors” in the middle and the “concretising characteristics” below them.

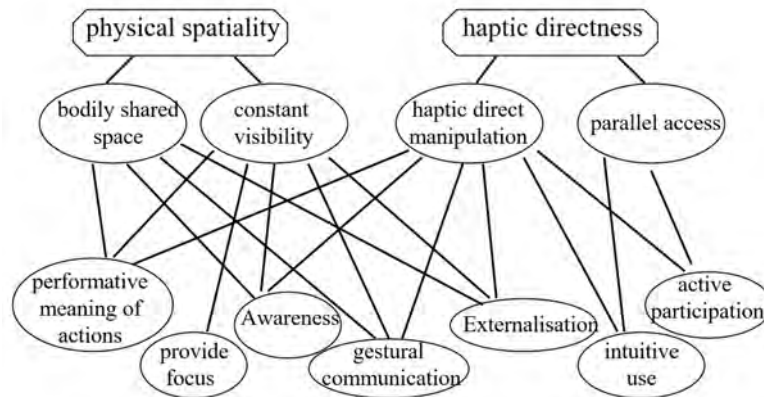


Figure 3.5: From [Hornecker and Buur, 2006] that presents the key concepts around the benefits brought by tangible interfaces

Those concepts are closely linked to the embodied actions paradigm. In their longitudinal field study and taxonomy Robertson defines embodied actions as “the publicly available, meaningful actions that people rely on to interact with others and their environment [...] These include actions like talking, touching, drawing, looking and moving around in the environment, in order to accomplish whatever the person acting wants to do.”. As noted even in his article, those actions have a particular relationship with the virtual and the physical world usually interacted with to act within the virtual. Robertson, in his taxonomy, focuses on the separation between the individual embodied actions and the group activities constituted by the individual actions of its members. The “concretising characteristics” brought force by the embodied actions performed during collaborative design are benefits of the TUIs over using multiple WIMP interfaces where such actions are not performed. It can still be noted that, those action don’t completely requires fully tangible interfaces as multi touch also benefits partially from the embodied actions. If we look at [Schneider et al., 2011] for an example of multi touch vs tangible condition it seems that participants using the tangible conditions performed better as well as collaborating better with the tangible condition.

Thinking through doing

Jean Piaget’s offers two concepts through which children learn : assimilation (adding new experiences to known concepts) and accommodation (adapting known concept to explain new experiences) [Piaget, 1997]. Those are used throughout his cognitive development work which contains 4 stages of learning for children : sensorimotor (from birth to age 2), preoperational (from 2 to 7), concrete operational (from 7 to 11) and formal operational (from 11). In the sensorimotor, the child think using his senses and actions. While the individual continue to develop from that stage the experimental nature of learning through the physical interaction with the world stays relevant. Even though it is not well considered nowadays, the Montessori method [Montessori, 1912] is a good example of experience driven learning which

emphasizes hands-on independent learning. In their paper [Hartmann et al., 2015] Hartman et al. presents Piagets work to support the benefits of the physical object during learning or thinking. The extensive work of Golding Meadow and Ping further stress the importance of the gesture during activities as gesture lighten the cognitive load [Goldin-Meadow et al., 2001][Ping and Goldin-Meadow, 2010] (even for children blind from birth).[Iverson and Goldin-Meadow, 1997]. Gestures also help speakers convey their message [Alibali et al., 1997][Holler et al., 2009][Ping and Goldin-Meadow, 2008] and speakers tend to gestures even when the audience suggesting a aid provided by the gesture [Bavelas et al., 1992]. Having gestures in your interface should therefore help the cognitive load of the action and improve the expected performance of the tasks. Another important element that Hartman and Takayama bring is motor memory, the capacity of our brain to memorise and recreate gestures using what is essentially kinesthetic memory. This is particularly central to soft-skill linked to movement like riding a bicycle or swimming. The motor memory while limited to gestures is particularly useful for situation where conscious reasoning would be slow and where the individual can rely on his past experience to do the rights actions depending on the circumstances. For example, improvising on the piano would not be possible completely consciously but instead rely on the individual past experience playing the piano to guide him to the correct next note to play [Sudnow and Dreyfus, 2001].

Previous and other studies

Studies on the differences between TUIs and GUIs have been done on numerous metrics, numerous population, about numerous themes and using numerous interfaces. Using the same format and papers compiled in [Cheng et al., 2010] and [Chen et al., 2011] with additional articles we get Table 3.3.

References	UI	Metric	Results	Users
[Wren and Reynolds, 2004]	TUI GUI	Usability and user preference	Better Learnability on TUIs	10 subjects
[Sitdhisanguan et al., 2007]	TUI GUI	Ease of use	Better ease of use for TUI	6 autistic children (3 to 7 years old)
[Xie et al., 2008]	TUI GUI Traditional Puzzle	Enjoyment, Engagement and collaboration	Similar enjoyment but better engagement for TUI and GUI. Collaboration inconclusive due to system differences	132 children (7-9 years old)
[Kim and Maher, 2008]	TUI GUI	Spatial Cognition (gestures, Perception) and performance	Significantly better for TUI with	7 designers
[Horn et al., 2009]	TUI GUI	Inviting, Apprehensible, active collaboration, Engagement and Age influence	TUIs more inviting, foster active collaboration and preferred by children but not more apprehensible and not significantly more engaging.	64 individual or group
[Marshall et al., 2009]	TUI Paper	General behaviour mainly collaboration and conflicts	Differences that might be due to changes in affordances	10 groups of 3 children (7-8 years old)
[Chen et al., 2011]	TUI GUI	Engagement	No significant differences	32 children (8 years old)
[Sylla et al., 2012]	TUI- GUI	Motivation and Learning	Better results overall for TUI	41 children(4-5 years old)
[Sei et al., 2018]	TUI- GUI	Hollnagel's Contextual Control Model	biggest variety in problem solving approach for TUI	8 participants
[Almjally et al., 2020]	TUI- GUI	Attitudinal (towards computer), Engagement and Learning	Better learning on GUI, Better Attitudinal on TUI and overall similar Enjoyment	44 students (6-7 years old)

Table 3.3: Results of paper comparing TUIs with GUIs

Model of learning through TUIs

In his paper [Marshall, 2007], the author presents an analytic framework of the concepts linked to learning with tangible technologies built using his analysis of the state of the art. He presents 6 perspectives that should guide research in that regards. The graph is available in Figure 3.6. In it, there is a lot of point we have already talked about like embodiment or the collaborative benefits of TUIs. The reflection on accessibility is also an interesting point to bring out as WIMP interface can be hard for some people with disabilities or young persons not yet fully able.

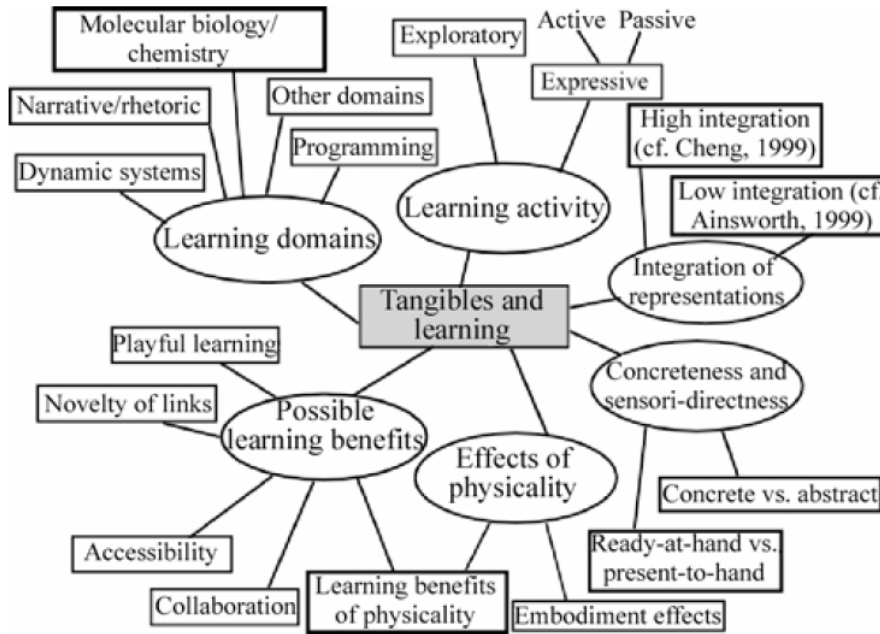


Figure 3.6: From [Marshall, 2007], his analytic framework on perspectives about benefits brought by TUIs for learning

3.4 Technologies

3.4.1 Tracking through chips

Radio-Frequency Identification

Radio-Frequency Identification or simply RFID is a wireless system with tags and readers. The tags contain information that is transmitted to the reader when it is read. RFID tags can be split into to categories : passive, semi-passive and active. Active RFID tags contains a power source and directly emit a signal using this power source. Meanwhile, passive RFID tags use a coil of wire and the radio signal of the reader to power itself using induction [Weinstein, 2005]. Having to include a power source, active tags are bigger, low power active tags being slightly larger than a smartphone. They are more expensive. Because of their power source, they can use higher frequency and be detected from 100 meters by readers. Passive tags are very inexpensive at around 20 cents per tags and quite smaller than their counterparts with the thickness of a sticker while being around 5cm per 5cm. Semi passive tags are in between the two others as they have a battery and broadcast their signal but only do so when queried by a reader. RFID has multiple standard within the International Standards Organisation (ISO) [Weinstein, 2005] and is widely used in most industries from factory's products tracking to merchandise security in stores and simple inventory management in companies. Because of this wide usage and multiple size and type available, there exists a lot of different model of RFID chips and as such very different specifications, especially in terms of data capacity. We can expect around 16 bytes for card size chips to a few kilobytes for bigger chips as it is still limited by the signal. Readers are also either passive or active, with passive readers only receiving signal from active tags. In terms of range, it depend on the

frequency of the signal from a few centimetres for low frequency to 100 meters (even up to 500 with active tags) for higher frequency [Weis, 2007] (see Table 3.4). We can also note that different frequency react differently to obstructing element and those distance could be shorter if some matters if between the tag and the reader [Deborah Dos Santos, 2021].

Frequency Range	Frequency	Range
Low Frequency (LF)	120-140 KHz	~ 10cm
High Frequency (HF)	13.56 MHz	~ 1m
Ultra-High Frequency (UHF)	868-928 MHz	~ 5m
Microwave	2.45-5.8 GHz	~ 100m
Ultra-Wide Band (UWB)	3.1-10.6 GHz	~ 100m

Table 3.4: RFID operating frequencies from [Weis, 2007]

As with all technologies used widely, some research or education projects also used it.

- ▶ Warhammer 40k tabletop RFID tracking [Hinske and Langheinrich, 2007][Hinske and Langheinrich, 2008]: a miniatures tracking system for the warhammer 40k tabletop game using RFID tags and a network of readers (see Subfigure 3.7(a)).
- ▶ Smart Playing Card [Floerkemeier and Mattern, 2006]: playing cards tracked with RFID tags that allow the tracking of the game state and offer an additional game interface on mobile phones (see Subfigure 3.7(b)).
- ▶ RFIBricks [Hsieh et al., 2018]: Interactive blocks that use RFID to track which block they are linked to and which is used in a computer game (see Subfigure 3.7(c)).
- ▶ Tangible Virtual World [Tan and Rau, 2010][Tan and Rau, 2015] A game platform that used RFID tags to track game pieces (see Subfigure 3.7(d)).
- ▶ Smart Jigsaw [Bohn, 2004] An augmented jigsaw with RFID tags (see Subfigure 3.7(e)).

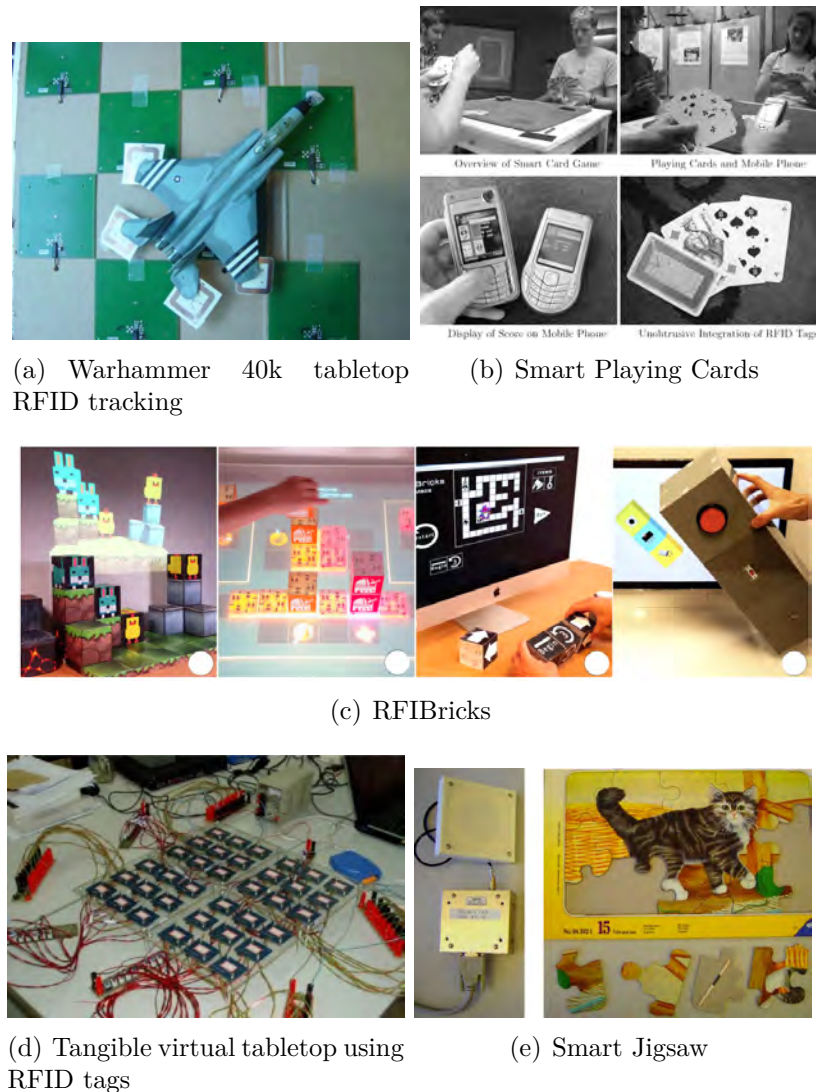


Figure 3.7: Examples of RFID applications

Near-Field Communication

More well known under the acronym NFC is a continuation of RFID. NFC also use inductive coupling between the tag's antenna and the reader's antenna. They also are active and passive NFC system as well as system like smartphones that emulate NFC cards. Like its name suggest, NFC is designed for very close ($\sim 2\text{cm}$) communication between device at the HF radio frequency (13.56 MHz). Just like for RFID, matters can impede the signal. Here there must be no metal between the tag and reader. This fact is used for security to stop tags being read when not wanted like when thief use makeshift reader to steal bank cards information. As hinted before, one of the most known use of NFC is for credit cards and contact less payment which has also been extended to phones and other electronic devices using the tag emulators. NFC are also used in gaming. Skylanders: Spyro's Adventure (see Subfigure 3.8(a)) uses it with customisable figurines and cards that are read to the main toy to play. In a similar fashion, Disney infinity also uses figurines with NFC tags to play the game. Nintendo's Amiibo are figurines of its characters and function a bit differently as they are not all developed for one, all encompassing

game. Each Amiibo is tied to usually a specific game (and at most a few), in which it gives some sort of bonus (gear for your character for example) or additional functionality (a new opponent AI that keep records of your plays for example) (see Subfigure 3.8(b)). Like for RFID, NFC has also been used in research or education projects.

- ▶ Skylanders: a commercial game where children can use collectables figurines to play using a reader hidden in a “portal of power” that read the NFC tags inside the figurine. (see Subfigure 3.8(a)).
- ▶ Amiibo: commercial figurines that unlock NFC tags inside the figurine. (see Subfigure 3.8(a)).
- ▶ Education game on Wildlife [Safrodin et al., 2019] : Indonesian board game to increase awareness about wildlife and use AR and NFC (see Subfigure 3.8(c)).
- ▶ Moodle game with NFC [Garrido et al., 2011] : Pervasive game that use NFC in mobile phones connected to a moodle platform to support learning motivation (see Subfigure 3.8(d)).



(a) Skylanders box with figurine, (b) Amiibo Figurine for the WiiU reader and game



(c) Indonesian board game with AR and NFC, (d) Skylanders box with figurine, reader and game

Figure 3.8: Examples of NFC applications

Pros and Cons of radio chips

Radio chips, both RFID and NFC, have advantages and disadvantages. The main advantages are :

- ▶ tags are relatively cheap if passive and even active are not prohibitively expensive
- ▶ readers can be connected to computers or even mini electronic cards like arduino

- ▶ connection with unity is not very hard
- ▶ their positions can be calculated using multiple readers
- ▶ they can contain quite a lot of data depending on your application but in our case the identification number can be quite large

On the other hand the disadvantages are :

- ▶ the prices are still not trivial for tags and readers can become expensive if multiple are required
- ▶ position detection does not have a very good precision for board games as it is more suitable for the position of goods in a warehouse
- ▶ readers close to one another can detect nearby tags meaning a certain distance needs to be kept between readers to avoid false positives
- ▶ common tags can be on the bigger size for our application purposes around 5x5 cm usually.

3.4.2 Image Recognition

OpenCV

OpenCV, for Open Source Computer Vision Library, is a C/C++ library for computer vision [Bradski and Kaehler, 2011]. The project is open source (Apache 2 license) the has been developed since 1999 and support other languages like Python, Ruby, etc and multiple platform from Windows, MacOS to Android. OpenCV includes a lot of computer vision functions such as faces detection, object identification movement tracking,... It is quite successful, being used by research groups and big companies among which Google, Microsoft and Intel and the organisation boasts more the 18 million downloads [OpenCV, 2023].

ARToolkit

ARToolkit is an open-source library for tracking images in video feed focused on Augmented Reality. It was first developed in 1999. While ARToolkit is now proprietary, being originally an open-source project it was forked in different versions and used with different technologies. Currently, the main successor is ARToolkitX.

Aruco

Aruco is a fiducial marker system first developed for camera pose estimation for augmented reality and robot localisation in 2013 at the university of Cordoba in Spain [Garrido-Jurado et al., 2014]. It uses markers similar to ARToolkit or ARTag. Set of markers are organised in a dictionary by their number of bits and the number of markers in the set. More bits increase the number of markers available in a set but demand a better resolution for detection. More markers in the set increase the similarity of markers which increase the error potential and reduce the available bits for error correction. Aruco has also been included in Open CV's library.

Pros and Cons of image recognition

Like radio chips, using image recognition brings some advantages and some disadvantages. The main advantages are :

- ▶ Price for scalability. The bulk of the cost in image recognition is the camera of which only one is usually needed. Patterns can be printed using common printer which camera increasing the scale very cheap.
- ▶ Encoding the pattern value is also quite easy as again, this only need a printer.
- ▶ Position tracking is very good if the camera's position can be identified.
- ▶ Having very good position allows to have almost no issue with pattern being too near each others as long as they don't obstruct one another.

However, there are some disadvantages :

- ▶ Pattern size are similarly restrictive as radio tags but for different reasons. The patterns need to be visible on the camera well enough to be recognized.
- ▶ Light condition needs to be good. Both a light strong enough and no "noise" that would mess with the patterns.
- ▶ the connection with the computer need some additional softwares that might be complicated to incorporate depending on the chosen technology.

Synthesis It is clear that there is a large variety of paradigms and methods within all the central theme of our research. The theoretical works on scheduling informed us on the key elements that must be kept at the centre of our design during our tangibilisation of Clone's graphical interface. Tangible User Interfaces, since its inception, have been developed into a broad range of applications and use cases. The technologies used for creating the link between the physical interfaces and the computer are also diverse and all have their advantages and drawbacks. This understanding allows us to carry on and start the design of the new tangible version of Clone. While still keeping in mind the constraints imposed by the goal of our study.

Synthèse Il est clair qu'il existe une grande variété de paradigmes et de méthodes au sein de tous les thèmes centraux de notre recherche. Les travaux théoriques sur la planification nous ont éclairés sur les éléments clés qui doivent rester au centre de notre conception lors de la tangibilisation de l'interface graphique de Clone. Depuis sa création, le tangible s'est développé en un large éventail d'applications et de cas d'utilisation. Les technologies utilisées pour créer le lien entre les interfaces physiques et l'ordinateur sont diverses et présentent toutes certains avantages et inconvénients. Cette compréhension nous permet de poursuivre nos travaux en commençant la conception de la nouvelle version tangible de Clone. Cela en gardant à l'esprit les contraintes imposées par l'objectif de notre étude.

Chapter 4

Tangible schedule development

Introduction Now that we have identified the available technologies and the constraints for our project, we can present its development phase which constitutes a large part of our work. This development started with a first version of the prototype tangible clone version which aimed at creating a tangible global care plan. This first prototype guided most of the subsequent design by introducing the main technology used: image recognition and patterns, the main physical artefacts: 3D printed tokens and boards and the main software: openCV. Disappointingly, due to a unforeseen constraints and issues with development, this version would be replaced by a second and then a third one that would both focus on the individual care plan. This third one would be the actual success and be the focus of our experiments.

Introduction Maintenant que nous avons identifié les technologies disponibles et les contraintes de notre projet, nous pouvons présenter sa phase de développement qui constitue la majeure partie de notre travail. Ce développement a commencé par une première version du prototype du clone tangible qui visait à créer un plan de soins global tangible. Ce premier prototype a guidé la plupart de la conception ultérieure en introduisant les principales technologies utilisées : la reconnaissance d'image et les patterns, les principaux artefacts physiques : les jetons et les planches imprimés en 3D et le principal logiciel : openCV. Malheureusement, en raison de contraintes imprévues et de problèmes de développement, cette version fut remplacée par une deuxième, puis une troisième, toutes deux axées sur le plan de soins individuel. C'est cette troisième version qui a été retenue et qui a fait l'objet de nos expérimentations.

4.1 Prototype versions overview

4.1.1 Introduction

Our project was developed through an iterative design. Each implementation was designed to answer challenges we encountered during development. The first implementation took up half the development time (from February 2020 to September 2021). The two other implementations are more similar to each other and took up the remaining time of the project.

4.1.2 First prototype

Scheduling but not executing

Our first design choice came at the start of our reflection on tangibilisation section 2.1.1. As described in subsection 1.3.1, Clone has multiple interfaces which are complex: the scheduling interface and the executing interface. To tangibilise both of them would require a lot more work than was possible during that project and we would end up with a huge physical object that would be unpractical for both the user and to transport and install for experimentation. Of the two interfaces, the executing interface is the least suited to be transcribed with fidelity into a tangible interface. This is due to a lot of information being in different sub-menus and the need for complex feedbacks like unforeseen events or post-actions information. Because the two main phases of Clone can be split more or less cleanly and scheduling is the main goal of Clone, we decided to focus on the scheduling interface for our experiment. To keep the gameplay as similar as possible during the experiment, for the digital version, we created a sub-version of Clone with only the scheduling part. This sub-version is included in the original through the scenarios mechanic. Meaning that some scenarios are defined as skipping the executing phase, going straight to the feedback at end of the game. This insured us that we would have the exact same game engine working behind the scenes for both versions with the exact same data instead of making two different applications.

Global care plan's interface

We now had to design the tangible scheduling interface of Clone. Again as presented before, Clone has two interfaces for scheduling: the global care plan and the individual care plan. The global care plan is the whole shift schedule. Meanwhile the individual care plan zoom on a specific patient at a specific step. The scheduling is mostly done on the individual care plan with the full list of activities while the global care plan only allow moving activities between steps or deleting them. We wanted to keep the physical object to a workstation. And at first we tried to design for both interfaces, but when thinking of ways to kept the global care plan as similar as possible to the original, we could not design a small physical object to not take the whole workstation. At that point, we believed that adding, on top of that, another, admittedly smaller, interface would be too much for users. We also thought of recreating the shift of care plan. In the graphical Clone, the individual care plan is not next to the global one but on top of it. But we could not think of simple ways to do so without something big like a custom desk with moving parts.

Therefore, we concluded that it would be best to focus on one of the interface. As Clone's scheduling interface is centred around the global care plan and we wanted to explore the mental workload of tasks which is more present in this interface, so we chose to focus on it. While they are done in the individual care plan, the actions of adding, removing or moving activities are simply done on the intersection which is part of the global care plan. So if we added the addition and delegation of activities to the global care plan we could get all the individual care plan's specific interactions. Our initial idea was to keep the graphical's "drag and drop" by using token for activities in constraints representing the intersection of patient and steps.

Technology choice

One choice that would impact the rest of the design is the technology chosen to connect the tangible elements to the computers. We had some solutions in mind when making that choice but the two main were using radio chips like RFID or NFC or use image recognition. Having decided to focus on the global care plan meant we would have a big board. When we looked at the characteristics of radio tags, we noticed that using readers in close proximity would create errors in the detection as they would detect tags meant for the reader next to them. If we wanted to avoid that issue, that would mean spreading the readers. Sadly, this was not realistically feasible due to the number of rows and columns needed in the global care plan. This is what prompted us to go with image recognition instead. In insight, we when scrapping the first version, this choice could have been reverted as those imperative of multiple readers did not exist anymore.

Design

The first prototype design (see Figure 4.1) used image detection coupled with plastic token with patterns on their back. A plastic board with slot serves as a play area. This board mimic the global care plan and has slots for the token that represent activities scheduled. This contraption rests on a transparent table with a camera underneath. The game, using the camera, detects which tokens are placed in the plastic board's slots. The player plays by putting the activity tokens in the slots he deems correct.

Shortcomings

The first prototype's shortcomings was the detection of the tokens. This simply did not work despite multiple variation and attempts to solve issues. As will be explained below, there were issues with detection speed (section 4.4.1) information density (section 4.3.1) and pattern size detection (section 4.4.2).

4.1.3 Second and third prototypes

Solving the first's shortcomings

Having failed to make the tangible global care plan work, the second and third prototype completely change the target interface being tangibilised. Because the main issue was the pattern size restriction, to be able to increase pattern size we went for the individual patient care's interface instead of the global care plan's one.



Figure 4.1: Design of the first prototype

The individual care plan has less pattern simultaneously: from 28x28 for the global care plan to 4x2 for the individual (not counting the rest of the interface). This meant we would have less pattern at the same time. In turn, each pattern could be bigger without getting outside of the camera fulcrum. Which meant we could use an already working pattern system. We chose Aruco (section 3.4.2) as they were available within OpenCV's library with which we were already working. While it wasn't necessary, during testing, we also worked with the camera above the area and kept it for the final version.

Projection as feedback

While the first prototype did not need a lot of feedback except for workload gauges, the change from global care plan to individual care plan increased this need. Our idea was to use a video projector onto the physical object to have information added in real time. This led to our board being projected constraint with physical tokens (see Figure 4.7).

From second to third

Like it is said in the introduction, the second was quite different to the first version. However, the third is very similar to the second. One of the features of the second version was the positional detection presented in section 4.2.2 have information displayed around the tokens. We tried to use that kept failing at implementation. Despite our effort we couldn't solve the issues and so we redesign the interface to be statically projected instead of adapting to any game area.



Figure 4.2: Third prototype being tested in the laboratory. Some of the tokens can be seen around the projected area on the bottom. The chemistry lab equipment that was used to hold the project is on the right. On the top of the image are the projector with the webcam behind it filming the game are below.

4.2 Boards

4.2.1 First Prototype's modular board

Board: 3D printing and modular design

Our prototypes use 3D printing extensively to facilitate token and board making. Because of the size of our printer, we naturally went toward making small modular elements. We decided on a puzzle-like design for our board with slots for patient, steps and activities containers which can be seen in Figure 4.3. This had several advantages. First, small identical pieces were easy to make using the 3D printers. Second, it allows for an easy transportation as all pieces could be detached and stacked in a briefcase. Third, despite having a lot of pieces, when needed, the board would be reassembled in under 10 minutes thanks to the puzzle attachments. Finally, it meant that we could use the same pieces for any scenario. Scenario with more patients would only require to add more lines for the additional patient and cell pieces. And all patients were represented using patterns that were also modular so only one patient token for each was needed to be able to play every scenarios. 3D printed plastic is also quite light with the whole boards, disassembled in the suitcase

weighing around 3 kilos. It however had one disadvantage: size once assembled. While this is a small one, the attachments between the pieces necessarily increase the size of each piece and therefore the size whole board. For our first prototype, we printed mostly in the same yellow colour which is what was used for our drawings. The patterns being printed then glued on tokens, we used what was convenient at the time. In our office, we have 2 3D printers. Each constraints object takes around 4 hours to print, each token around half an hour for the small one (activities) and 1 hour to 2 hours for the bigger ones (patients, steps and detection triangle). The 3D printers are big enough that multiple prints can be done on the same board so we could set up the machine and print 4 constraints object during the night for example. Luckily, the office did not need the printers at the time of our printing so we manage to print most of what would be needed in a month.

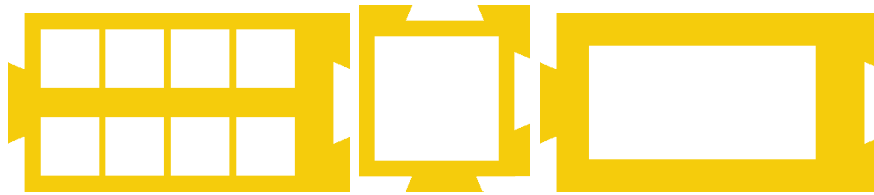


Figure 4.3: Designs of each modular board pieces. Form left to right: Activities container at the intersection of patients and steps, Patient container and Step container

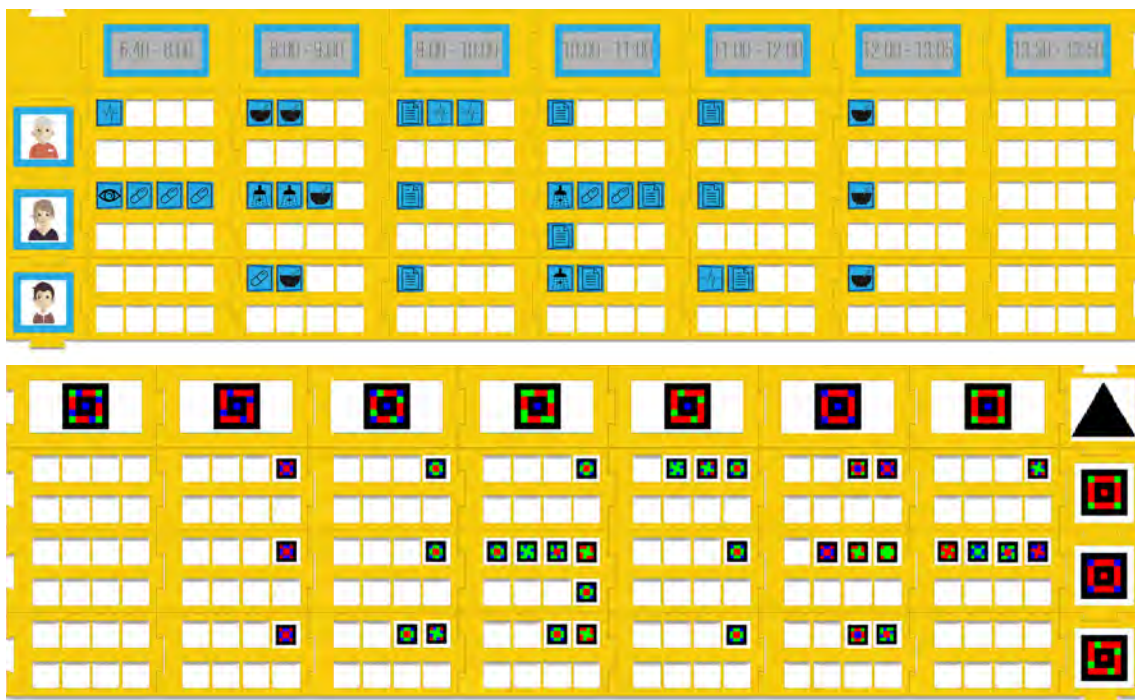


Figure 4.4: Design of the modular board with the up side (above) and the back side (below). Each token has behind their corresponding pattern. On the upper right side of the back side, there is the size calibration triangle which is used during detection.

Board Size constraint

The schedule is composed of the following parts a patient header column, a step header line, an internal (between the headers column and line) grid of patient/step intersection (cell containers) for the activities and workload meters. The number of step is fixed at 7 regardless of the scenario. The number of patient changes, going from 1 to up to 14 patients. The internal grid needs to be able to scale to 7x14. In Clone digital, cell contains up to 8 activities so Clone tangible should replicate and have at least 8 slot for tasks in cells. Finally, a space for the header should be reserved. The workload meters are not interacted with by the user, they're only a feedback. This gave us leeway on its placement because it did not need to be detected. This size of all those board elements are linked in-between themselves and with the size of the activity tokens. A cell must hold 4x2 tokens. The board has the following width and height. The later depends on the number of patient.

- ▶ $Width_{board} = Width_{PatientContainer} + 7 * Width_{Cell}$.
- ▶ $Height_{board} = Height_{StepContainer} + NumberOfPatient * Height_{Cell}$.

Board final dimensions

During preliminary testing with our camera we noticed that it would struggle to detect that patterns below 2cm and be almost useless below 1 cm. While those tests used QR and micro QR and we would chose a different pattern (see Patterns), the issue of size stayed similar. With size in mind to avoid a board too big, we choose to use 2cm patterns (1.8cm to be exact to keep a white border) for our first prototype. This cascade into the dimension of our overall board. Cells and step containers both have a size of 10.5x5.5 cm, while patients containers a size of 6x5.5 cm. This gives a 3 patients board a size of 75.5x22.5 cm and a 14 patients board a size of 75.5x83 cm. An example of a 3 patients board can be see in Figure 4.5.

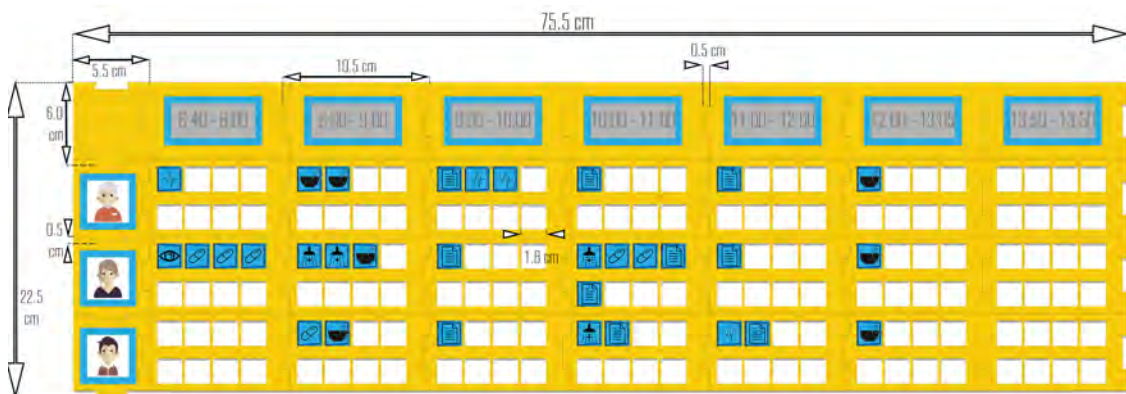


Figure 4.5: Dimensions for the first prototype's board with 3 patients

Play area

With those dimension we have quite a large assembled object that need a large table to fit. In addition, we also have to include the reserve tokens. Players need to have

access, during play, to the activity tokens he would need to place and a place where he can put the activity he wishes to remove. This obviously increase the play area needed. However, this is less of a constraint as it does not have to be visible by the camera nor has to be attached to the board. This still need to be considered but as we will see below, we had some area of the table not visible due to the trestles so we used this for our reserve token zone. We also designed some paper boxes to hold them in a ordered manner to avoid the player mixing them over their game. And making it harder and harder to find activities quickly when playing as well as making the reset between player longer.

Detection: camera and table

With image recognition on token, you need to have a camera that can see the patterns. On our flat tokens, we only had one side up and directly visible for the user. In Clone digital, because there is no pattern to worry about, the elements have icons to distinguish them from one another and to make the user experience more intuitive. With our first prototype we considered three options on where to put the patterns. First, we could put them on the up side replacing the pattern. For us this would be a huge change to the experience has player would have a hard time following the overall game. Otherwise, we could put the pattern and the icon both on the upside. This would severely reduce the pattern to token size and which in turn would force us to increase the token size to keep the pattern detectable. Finally, we could put the pattern on the backside. This would require us to have the camera below the board and with a way for it to see the pattern. We chose to go with the later option. As those experiment could be in any of the IFSIs, the "transparent table" would need to be transportable when doing the experiments. We bought transparent PVC plates and some trestles. The plates were bought quite thick to avoid bending in the middle (we also considered making a wood frame). They were further cut with round edges to reduce people bumping into them and with handles to make it easier to transport. A photo of the contraption can be seen in 4.6 showing the same distribution of lines and rows.

4.2.2 Second's and third's projected board

Projection

The second's and third's prototype use a projector as feedback. This was done to complement information feedback from the physical object. As the feedback isn't physical, this is not the most tangible but this allows us to add dynamic information directly on the playing surface. Doing so without resorting to complex electronics or object with moving parts as this would not be feasible in our time frame with our current skill set.

Projected interface

The interface was designed as a continuation of the previous ideas. By that we mean we kept a game area at the centre to place the patterns. This is because our idea for the second version was to have the tokens tracked by camera and additional information would be shown around the tokens. This meant the tokens would need

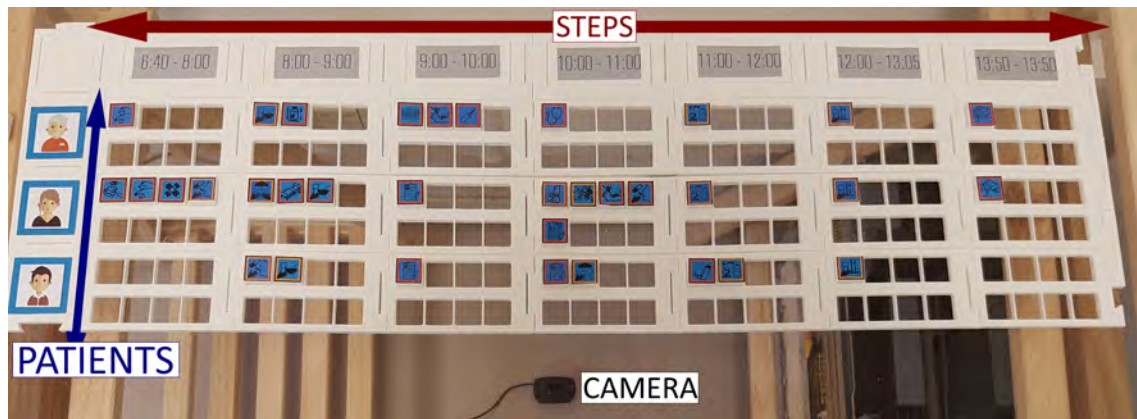
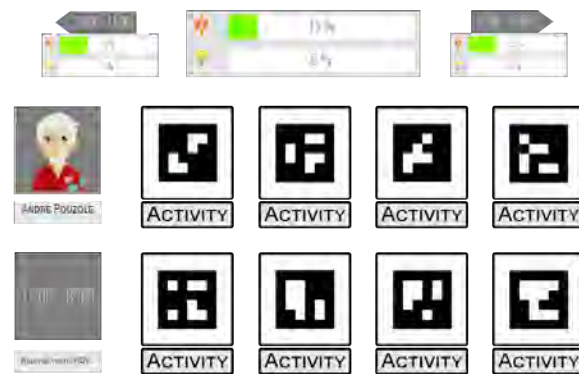
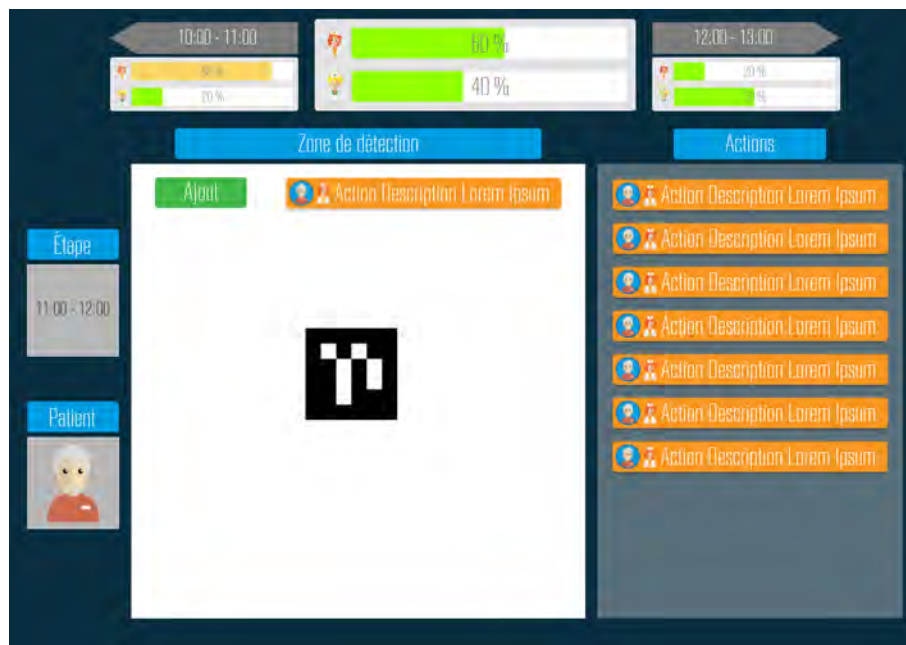


Figure 4.6: Photo of the first prototype on the transparent table and with the camera below. The same disposition of patient as row and steps as column as the designs. The activity tokens are twice as thick as the rest of the board to make them easy to grab while being very close to the dimensions of the holes so they are both easy to place in and won't move even if the table is shook slightly.

to be placed on a rather bland, if possible white surface so the projected information would be easy to see. While this wasn't necessary anymore for the third version, the third version's UI was simply the second version's UI with the tasks list displayed. Our idea was that putting an area for pattern would make the detection easier for the camera by avoiding projected clutter on the patter and make the action of placing the pattern in the white zone easy to understand and remember. For the example of the UI's mock up (see Figure 4.7 and Figure 4.8).



(a) Second Prototype



(b) Third Prototype

Figure 4.7: The two main UI mock ups made for the second (a) and third (b) prototype versions. The third mock up also uses the graphical identity of Clone.

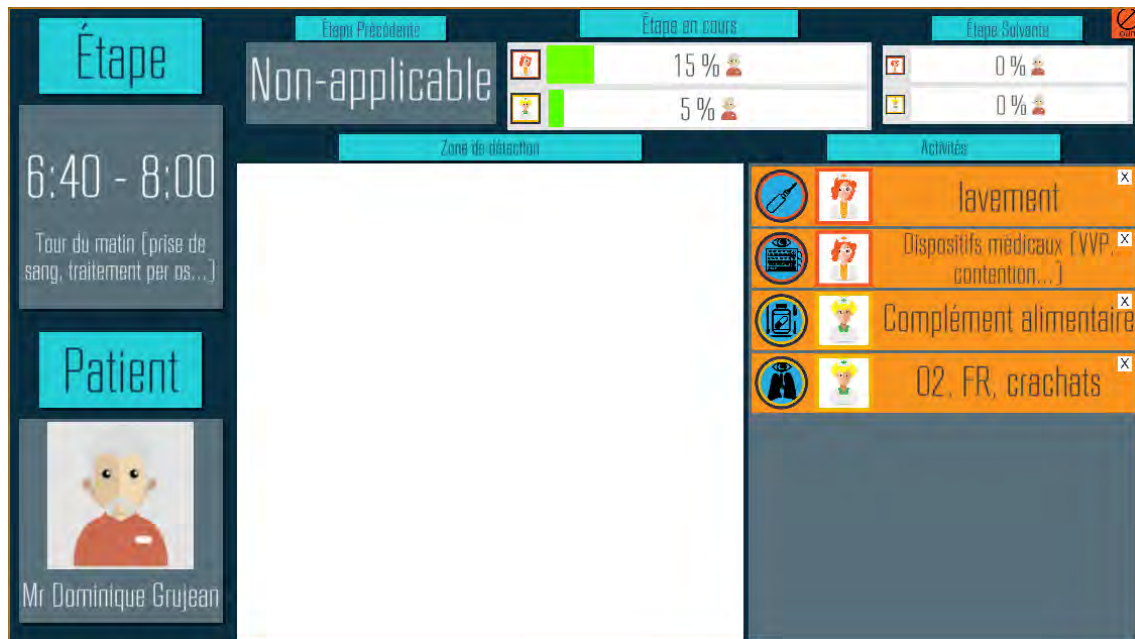


Figure 4.8: Third Prototype Main UI

Positioning and calibration for the second prototype

With the second prototype, our idea was to fully exploit the information we were getting from patterns. Meaning, we would, in addition to the presence/absence of pattern also its position. Our goal was to super impose on and around the pattern its information to serve as quality of life for the player and feedback on the detection. To do that for a pattern we had to get the position of the pattern in the camera video feed and convert it into a position in the projected space. This meant that, before anything we needed the position of the projected space within the camera fulcrum. To do that we used patterns of known position in the projected space, projected them in the camera space to get their position in the camera fulcrum and to some maths to get the conversion from camera space to projected space. This can be seen in Figure 4.9. On the left image is the camera vision with its border in red and corners as red squares with, inside the calculated projected area in blue with its corners in blue squares and inside it the 4 patterns for the calculations. On the right image is Unity's view with the projected area in blue, and same corners as before, with the 4 patterns forming rectangle with grey circle to mark their position and finally the calculated camera vision in red, with red squares as corners except for one. The one blue circle corner serves as a reminder that, due to some quirk of the detection, the horizontal axis is inverted in Unity compared to the camera video feed. As explained above, this did not work out for two reasons. The first was that the detection wasn't precise enough and would leave a skew in the detection (note how, in the camera view, the blue square is skewed compared to the actual projected area). The second is a consequence of that small skew, it is a tendency for the calculation to diverge and place the patterns completely outside of the projected area.

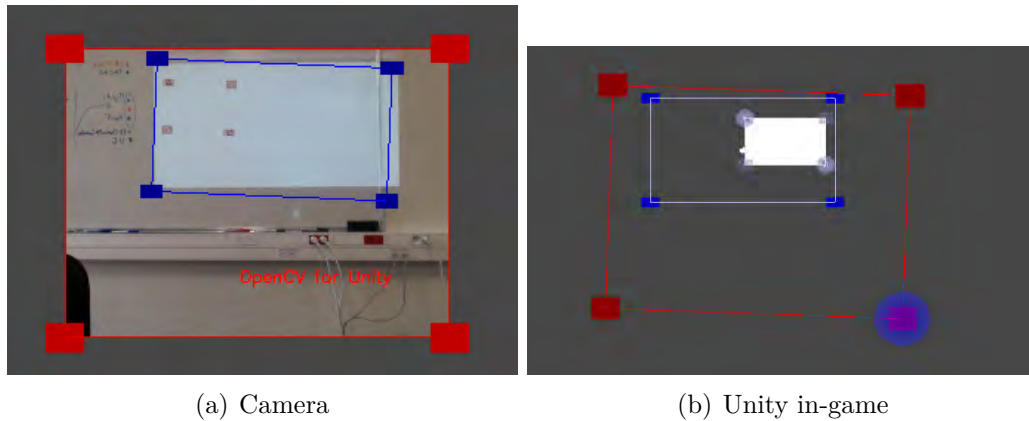


Figure 4.9: Positioning space representation. Red represents the camera fulcrum and blue the projected game area. (a) shows camera video feed. (b) shows Unity’s in-game space. Note that Unity’s in-game space is inverted horizontally.

4.3 Patterns

4.3.1 Custom patterns

Technology choice: OpenCV with custom patterns

For our first prototype’s pattern we had multiple already implemented pattern to choose from. The main issue with all of them was detecting them in a video feed that would cover the whole board. The board can have up to 28 activity/task patterns in width, a patient pattern and the structure of the board holding the tokens. It’s similar for the height but with the step pattern as there is only half the activity/task pattern per cell but there is double the number of maximum patient as steps. All of this need to fit in the camera’s fulcrum. The camera’s resolution will give us then the best possible number of pixel per pattern (and pixel per coloured square for pattern that use them to hold information).

The most compact microQR is 11x11 element wide with at least 2 dead white space for detection making it a 15x15 grid. On a 1.8 cm token this equate to each element being at best 0.12 cm wide. If we use the 14 patients board with a size of 75.5x83 cm filmed by a 1080x1920 pixels camera. The optimal ratio of pixel per centimetres we can get is either 14.3 px/cm for the width or 23.13 px/cm for the height. We have to get the smaller, most restrictive 14.3 px/cm. This means that each square of information in the microQR has, at best, 1.7 pixels in the image. This is obviously even worse for the 76x76 elements of the normal QR code which gives it 0.3 pixel for an square. This is even worse when we take into account the imperfection of the video feed with blur, non optimal camera field of view, etc. To still be able to remake the original interface, we decided to use customs patterns for our detection.

Number of variation for the first prototype’s design

In our first prototype, we incorrectly assumed we would need to use tasks directly instead of the activities (which are different as explained in subsection 1.3.2). While Clone has 65 activities (if you count them for IDE and AS), it has over 322 tasks.

And if we add those to the 14 patients and 7 steps for a total of 343 elements (tasks+patients+steps) we need to track so at least that many different patterns.

4.3.2 Pattern's border

The pattern's border has been designed to work with the detection method so the borders changed when the method changed.

Border for image matching

When we started designing patterns, it was our understanding that we would use image matching for the detection. To that end we started with a border that would be very hard to find in reality: black squares in a specific symmetrical arrangement. The point being to reduce false positives by having a complex and high contrast element.

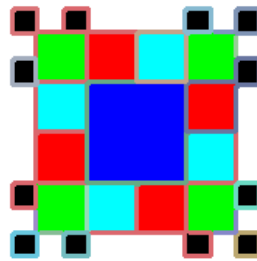


Figure 4.10: Complex borders are very unique and so easier to distinguish with image matching but they also have a lot of contours. In this image, all the numerous detected contours are drawn, with a random colour. Coloured squares are also contoured because the image wasn't "thresholded" before the contour finding.

Border for contour findings

This use of contour finding forces us to use a solid rectangle for our border so that we have one solid contour to look for. At first, we used a rectangle separated from the pattern but this duplicated our contours finding. So we made the rectangle extend to the pattern core. Those borders would be used up to our witch to Aruco.

4.3.3 Pattern's core: Number of squares and colours

All our patterns are made of two elements: the core and the border. Our main issue with patterns has been to be able to keep them as small as possible while keeping the information in their core they hold readable. To read this information, we must be able to distinguish them, so each core of the patterns has to be different. The number of variations the pattern can have a direct impact on the number of coloured square it must have. Given that the minimum size of a pattern is when the coloured squared can still be read, we also need to keep the number of squares low. If we use the number of 322 tasks and we keep the black and white squares we can quickly calculate, using $2^8 = 256 < 343 \text{ elements} < 2^9 = 512$, that we need at least 9 squares. However, we can further reduce this by using more colours, not simply binary/black and white squares. We could use three colours like red, green and blue

which allow us to use only 6 squares, using $3^5 = 243 < 343 \text{ elements} < 3^6 = 729$. We could also use six colours by adding the previous three ones yellow, magenta and cyan which would allow us to use only 4 squares, using $6^3 = 216 < 343 \text{ elements} < 6^4 = 1296$. Our pattern's colours' amount changed during development, for all tasks we started with 6 but went down to only 3 when we figured we had to use the less numerous activities.

4.3.4 Pattern's core: Central symmetry

We had an interrogation during our design. We were thinking about what could happen if player did not put the pattern in the correct orientation. If a symmetrical other pattern existed, the software would incorrectly read it and the player might not catch its mistake. We had to make sure no pattern was the rotation of another. We could either sort our pattern to remove any symmetries, we could alter our pattern to add an orientation bits to check for orientation before getting the pattern information or we could make the pattern a central symmetry. The solutions had pros and cons with the first keeping the square number low but requiring more work to implement, the last one being a bit wasteful in square number but being way easier to do and the middle one being a middle ground between the two. For our first prototype, we chose the last one at first. Our idea was that, if needed, we could swap the pattern information detection part of our software later if we needed more compact patterns.

4.3.5 Pattern's core: Colour encoding and decoding

Referencing

In Unity, activities, tasks, patients and steps all have an id that comes from the external system we use for our game data management Articy. All those id are unique which would allow us to only put the id on the pattern. However, those id are too long to fit within our squares combination limitations. So, when we were creating the pattern, we assigned each Articy id to a pattern value (simple number) and kept the reference table in a xml file. The pattern value is what was encoded/decoded in the pattern. Our pattern generation is done using python scripts and they also generate the reference file connecting the pattern to the unity object. The scripts' first step is to open the corresponding Clone's data. Clone's data is tied to each scenario so, to get as much information as possible, the scripts use one of the 14 patients scenario. Using those files, they list all objects: activities, patients and steps. For each of those, they either load the corresponding reference file or create a new empty one if none are found. Then, if they cycle through every objects and if it isn't in the reference file, add it along with an pattern value. When all object have a reference, the scripts generate the pattern's image using the encoding method. An example of such file is given in Listing 4.1

```

1 <root>
2 <reference id="0x0100000200004C2D">141</reference>
3 <reference id="0x0100000200004C35">142</reference>
4 <reference id="0x0100000200004C44">143</reference>
5 <reference id="0x0100000200004C4C">144</reference>
6 <reference id="0x0100000200004C54">145</reference>
7 <reference id="0x0100000200004CFE">146</reference>

```



```

8 <reference id="0x0100000200004C5C">147</reference>
9 <reference id="0x0100000200004C64">148</reference>
10 <reference id="0x0100000200004C6C">149</reference>
11 <reference id="0x0100000200004CD8">150</reference>
12 </root>

```

Listing 4.1: Example of reference file: Step reference xml

Encoding

To create the image we take the pattern value and convert it to either the senary numeral system or the ternary. Just like the binary system only uses 2 digits 0 and 1, the senary system uses 6, from 0 to 5 and the ternary 3, from 0, 1 and 2. We convert our pattern value from the decimal system to the other system. Each digit is linked to a colour from 0 to 5 respectively red, green, blue, cyan, magenta and yellow. Then the squares are placed in their place so they can be decoded later.

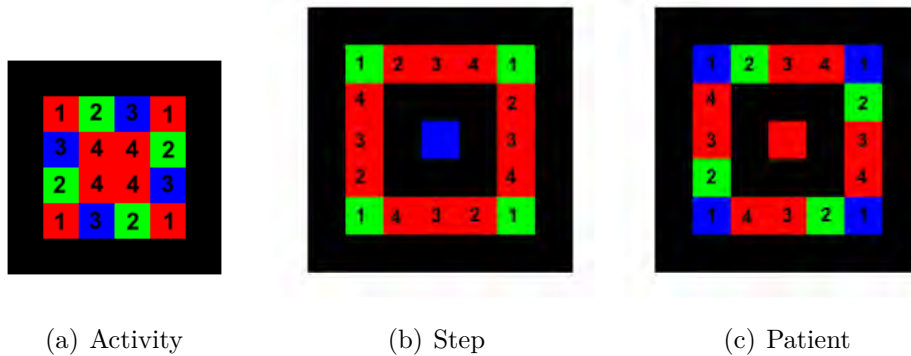


Figure 4.11: Pattern's coloured squares. They are numbered according to their corresponding digit in the pattern value. On the left (a) the task pattern, in the middle (b) a step pattern and on the right (c) a patient pattern. Step and patient pattern are similar with only the centre changing from blue (step) to red (patient).

For example, let's encode the value 42 in a ternary. 42 is equal to $1*27+1*9+2*3+0*1$. So, in ternary, 42 is 1120. Given the digit to colours table (where red is 0, green is 1 and blue is 2), we would get the colours green-green-blue-red. In the other way around, the left pattern's colours from the Figure 4.11 are red-green-blue-red so we get 0120 which is 15 in decimal.

Decoding

When we look for a method to encode and decode pattern, after trying image matching, see subsection 4.4.1 we used the same Aruco's method. By knowing the patterns' size and position and because patterns are made of coloured squares, we can cut and extract those squares. We then average the colours of pixel in the square. Each square is paired with 3 other because of the central symmetry so we can reduce error by only focusing on the most common colour among those 4 squares. With the colours identified and the senary or trinary system, we revert those colour encodings to digits and converted them to our decimal id. Using the reference files, we get us the pattern Articy/Unity id.

4.3.6 Pattern’s core: size of coloured bits

With this new pattern, we can redo our calculation from section 4.3.1 for the pixel size of smaller element needed to be detected. Still, the 14 patients board has a 14.3 px/cm. The our coloured pattern’s coloured bits of information measure 0.24 cm in both width and height. This means that each coloured bits, in the most optimal image possible, takes 3.5 pixel in the image. This is better with only 3 patients, while the height is reduced from 83 to 22.5 cm, the width won’t change. The improvement comes from rotating the camera to match the bigger resolution with the bigger size. In this case we get either 48 px/cm or 25.43 px/cm which translate to 6.10 px for a coloured bit.

4.4 Detection

4.4.1 Methods: Image Matching vs contours and core sequencing vs Aruco

During our development we used 3 different methods for our detection.

Image Matching

Our first detection method was using image matching. Image matching compares an image candidate to a frame of the video feed. To do that, it extract from the frame a sub-frame of the same size as the image candidate. It then compares, pixel at pixel, the sub-frame to the image. If the differences are below a certain threshold, it is a match at the position of the sub-frame. The image matching extract each sub-frame possible from the frame. For example, for a 1280×720p resolution in the video feed, and searching for a 100x100 pixels image, there would be $(1280 - 100) * (720 - 100) = 731600$ *subframes*. Furthermore, to search for multiple images, we would have to do this for each image candidates. Finally, this method would mostly work with exact image; this would not be the best with a real video feed.

Contour findings

After testing with image matching, we found that it was extremely slow, even with reduced resolutions which is understandable when it has to do $100 * 100 * 731600 * 3 = 2194800000$ *calulations* for each pattern searched in each video frame. To improve performances, we split the pattern detection in two: finding patterns and extracting the pattern’s information. The idea was to, at first, only do the image matching for curated spots where there should be patterns. We used contours to find pattern candidates. Open CV uses the contours algorithm described in [Suzuki and Keiichi, 1985]. In this algorithm, for binary picture (two colours), border point are any point of a colour with a 8-(or 4-) neighbouring points of the other colour and contours are set of neighbouring border points. To uses OpenCV’s “find contours” function, we slightly blur then threshold (making the image black and white with colour below a threshold black and above white) the image. Then, the “findCountours()” method returns us interesting elements in the image. By filter in those elements with characteristics expected of a pattern, we can then extract the pattern’s information.

ArUco

Open CV has a module named ArUco. ArUco has been developed by a team at the university of Cordoba in Spain. [Garrido-Jurado et al., 2016] [Romero-Ramirez et al., 2018]. Its aim is to detect the fiducial markers they have developed. The markers are black and white with a thick border and squares that identify them not unlike QR Codes. The markers can have different numbers of those squares from 4x4 to 7x7. They are grouped into a dictionary of different sizes. For example, there is a dictionary of 50 4 by 4 markers and a dictionary of 250 4 by 4 markers. The idea is that the 50 are chosen to be very different so the software can apply error correction better than with the 250 marker dictionary where the “distance” between markers is shorter. Markers with more squares can contain more information and/or the error correction is better but they require a better resolution or to be bigger for the camera to properly detect them.

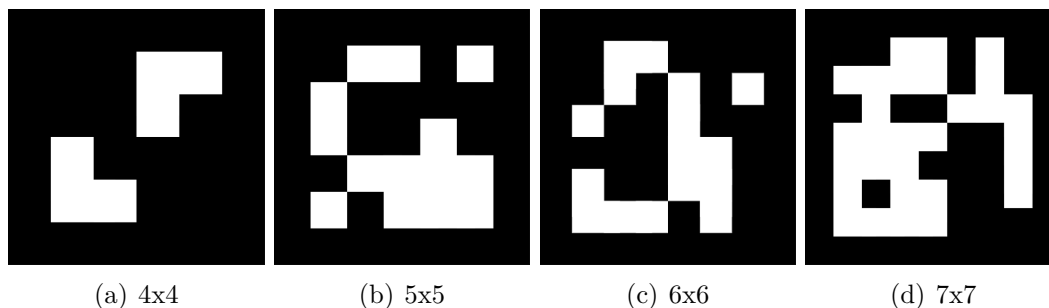


Figure 4.12: ArUco markers for the value 42 in 4x4 (a), 5x5(b), 6x6(c) and 7x7(d) squares

4.4.2 Detection loop

First prototype’s detection loop

The detection would work in steps. Before the loop started, we first initialised the variables and setup the parameters for the detection. Then the loop would go through the following steps: image modifications, contour findings, non-pattern filtering, information extraction and finally information sending. First some modifications are done so that other steps can use the video input. For contours, we blur the image slightly which reduces non interesting elements being picked up and then we threshold the image to only get black-to-white contours and canny the image (which, admittedly could be in the finding contours proper). We then do the contours finding explained above in Contour findings. With the contours, we try to filter as much as possible to avoid doing the expensive information extraction on non-pattern. To do that we find the size calibration triangle to get the expected as explained in Square pattern and the importance of size detection and use it along with their square nature to filter them. Having extracted the pattern, we cut them into their coloured squared “bits” and extract their information as explained in Pattern’s core: Colour encoding and decoding. At the end of every loop we send the information to the C# side so it update its values. As the loop is not synchronised with the C#, in the first prototype, Unity is not in the loop and work using an

event-base model using those end of loop data transfer to update as needed. This part caused a lot of memory.

Square pattern and the importance of size detection

One of the challenge we had when developing the detection was the contour filtering. One of the criteria we used was expected size of a pattern. We included a pattern that would have been easier to detect, a black equilateral triangle on a white background (see Figure 4.13), and use its size in the video feed to estimate the size a pattern should have in the feed. This allowed us to filter any contour that were too far from this size. The system was completely automatic, once the size triangle was detected, the other steps of the detection loop could carry on. The detection of said triangle was very reliable when we used a perfect image, however, this was far from the case when we used a real video feed. The criteria for selecting a contour as the size triangle were: have 3 sides and the sides are roughly the same size. The main issue was that, often, we would get false positives or false negative for the actual triangle. The false positive would be triangular shapes in view of the camera that would still be, coincidentally, equilateral. The false negative would often be when the size calibration triangle would register as 4 sided. Those issue were constant and would ruin the rest of the detection loop as we would be very likely of missing pattern that are present.

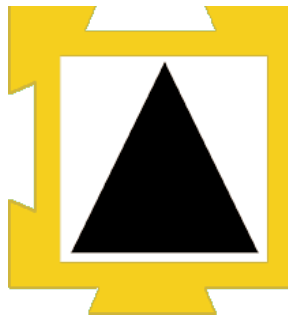


Figure 4.13: The detection triangle is on the back of upper-left side of the prototype. This stark contrast and simple geometry have been specifically chosen to ease detection.

4.4.3 Connecting OpenCV and Unity

Making a detection library using OpenCV

OpenCV is primarily a C++ library but also exist as a python package. Clone uses Unity which is, outwardly at least, in C#. While the two languages have similarities, making them work together is possible but not trivial. For the first prototype, we chose to create a DLL that uses OpenCV function and can be communicated with in Unity.

Shared objects and memory issues

One difference between C# and C++ is the way the memory is managed. While C++ have some automatic garbage collection, developers usually have to allocate

memory space to objects. C#'s management of memory is more automatic and don't really allow the developer to have the same fine control as with C++. This becomes an issue when we have to transmit data from the detecting C++ to Clone's C#. At first, for the first prototype, we started by passing primitive variable (Boolean, integer or floats). This was fine as they were the same object on each sides. To pass more complex object we used C++ pointer. All objects were stored on the C++'s side. C# had a "data" object that mirrored the C++ but only contain the memory address and methods that call the C++ to get the rest of the object's content. Consequently, for the C#, even the heaviest, in term of memory space, were only a memory address, so not very important to "garbage collect"(meaning automatically delete from memory when not used) as soon as possible. This meant that for complex and heavy object like images, the C# would not delete them causing a memory leak situation, where over time, the memory would get more and more full. There also were problems with object containing other objects that would sometimes be deleted by the C++ but still though to exist by the C#. When the C# would try to access the deleted C++ object, the application would crash. In the first prototype, this issue was sideline by only using primitives objects and only once when they are sent to the C#.

OpenCV4Unity

While we were working on the dll, we found a unity package called OpenCV4Unity. This package aim is to port OpenCV functionalities in Unity in a method similar to ours. While it meant we had to scrap our dll and remake some of our code to fit this package API, we still gain some development time by switch to it.

4.5 Activities' icon design

4.5.1 The need of remaking the icons

In the original Clone digital, activity were given a category. Those categories reflected the reality of the work and served to organised the activities in the individual patient care plan. During development, activities weren't given an icon for each of them, instead, categories were given icons and activities took their category's icon. Additionally, in the global care plan's interface, activities of the same category would combine under the same icon/category. The player could hover his mouse over the icon and get a tool tip reminding him of the scheduled activity's name. All of this would be hard to replicate in a tangible form so we opted to make each activity independent. This meant that we needed icons for the activities.

4.5.2 The new icons design

For the design, we took inspiration from the already designed icons and the overall art of Clone digital. We did not want to have icon clashing with the rest of the game. Clone had two icon type, round or squared for tasks and activities respectively. Because tasks use their activity's category's icon, our change would also impact them so we kept the round icon limits in mind during our design. The icons would be split into 3 parts: the symbol, the inner-ring and the background. The background was

shared by all icons and we kept it mostly like it was before: light blue in the back and dark blue outer-ring, we just removed the gradient to streamline the implementation in Unity. The inner-ring was an addition we made to show the delegation of the activity directly on the icon. While it is a good reminder in Clone digital, it is capital in Clone tangible as delegable activities are given two patterns, one for the IDE and one for the AS. This ring would simply be coloured using the character icons' hair colours: orange for the redhead IDE and yellow for the blonde AS. Finally the symbols were made to embodied the activity. Often using the item used like pills for the “Give oral medicine” activity or a chair for the “mobility assistance” one. We also tried to group some icons like for the “monitoring” category, we used the eye symbol, which was the category's symbol, above all of their symbol to show they belong together.

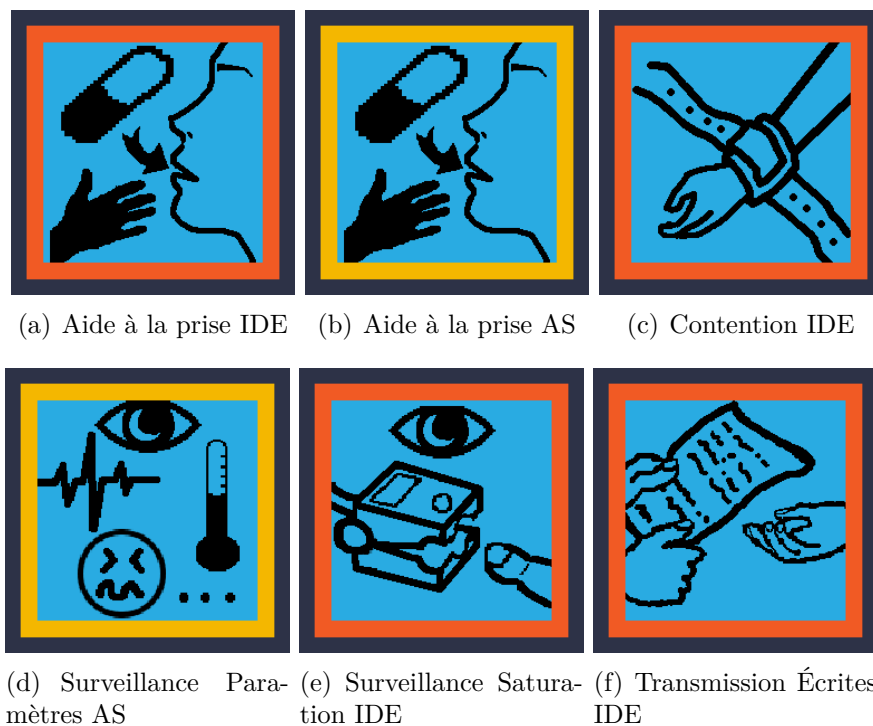


Figure 4.14: A few examples of activities icons. The first two are both version of the same activity “Help for taking pills” with (a) being the IDE version and (b) the AS's. The third one, “medical restraint” (c), is an example of an activity icon that had multiple versions after some feedback from the expert. The fourth and fifth one, “vital parameters surveillance” (d) and “saturation surveillance” (e), show the eye symbol that was the old surveillance category symbol and was used to group the surveillance activities' icons. The last one is “written transmissions” (f)

4.5.3 New icons validation

The icons were then sent to an IFSI's expert. We had multiple exchanges to tweak some of the icons. Some of the symbols had multiple versions made so the best could be chosen. We also tried them in the digital application to check the size. This last part showed us that some symbols needed simplification to be readable (some still do). Finally, we printed the icons to the size of our token and they were satisfactory.

4.5.4 Need for categories

Despite the new icons, during experimentation, some of the first participants complained of the lack of the categories for the tokens. In Clone numeric, activities are grouped in categories. We chose to organise the tangible tokens using the same categories to make the user experience more similar and easier. This arrangement was done using sheets of paper and was, admittedly, not the most efficient at keeping the token organised.

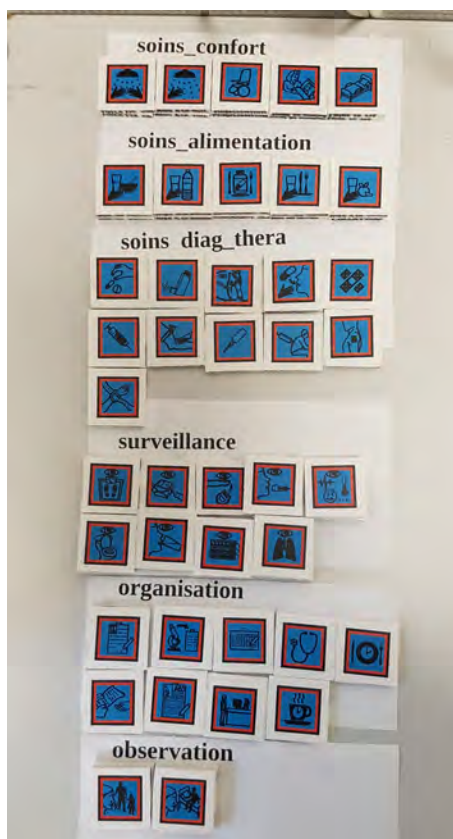


Figure 4.15: Categories for the physical ability tokens which are the same as the regular categories in the original graphical interface

4.6 Answering our design research questions

4.6.1 Tangibilisation of Clone's game elements

Patients and steps

During our initial design work, we had to think about how Clone's game elements (see section 1.2) would be transcribed in a tangible interface. In the scheduling the main action for players is scheduling (or unscheduling) activities. When doing this action, the player also has to convey which patient and step are targeted by the activity. This isn't trivial as it could be conveyed in different manners. As we explained subsection 4.1.2, in in the first prototype, the player would add a task to a patient/step pair by placing the activity in one of the pair's activity slot. This makes each patient/step pair as a constraint within the T+C paradigm section 3.3.3.



Figure 4.16: Table with the tokens' placement around the main interface that is projected on the blank paper at the centre. The paper is used to maximise light reflection as projector can be lacking in light intensity. Using it help keep the conditions similar across test sites.

On the other hand, in the second and third prototype, the patient and steps became token. The player could select the current patient and current step. When used, activities would be added/removed from those current patient and steps.

Special actions

In parallel, some players' actions don't fit in the activity/patient/step model. For example, confirming the planning at the end or getting patient files. The first prototype did not get to that point as problems accumulated before those actions were needed. However, the other prototypes used the same interaction method with token for those "special actions". On the note of patients' files, in the digital version of Clone, players have to "reveal" patient's information by clicking on them. We chose to remove this part for simplicity. In Clone tangible, opening a patient file automatically reveals all information in it.

Workload gauges

Similarly to special actions, there was some reflection on how to implement them in the first prototype (using LED lines or other external electronics) but this prototype did not get to the point where we had to actually implement it so those ideas stayed as such. For the second and third however, the solution was easy. The workload gauges are feedbacks only and can't be interacted with directly. They were therefore simply included in the projected interface.

Objectives

Clone's objectives only take into account the executing phase. As Clone tangible only contains the scheduling phase, they were not implemented during our tangibilisation.

4.6.2 Lesson on the tangibilisation

PQ1 How to "tangibilise" an already existing interface?

Tangibilisation with a goal: similarity

Not all tangibilisation are done with the same goal in mind. Necessarily, the goal will have an impact on what is important during the process. In our work, the goal was to research the impact of such changes on some key metrics and, as explained in subsection 2.1.2, we wanted to keep both versions as similar as possible. This will not be the same for other project. For example, people aiming at performances on some specific factor will only employ tangibilisation as a means to an end. Only implementing tangible principles to benefit this goal.

Method employed: analysis of affordances and feedbacks

One of the central step in our tangibilisation was the identification of interactions and feedbacks that would need to be converted. In our case, for the interactions, our tangibilisation's goal was to keep the same list and simply change the way they are performed. The feedbacks' analysis followed the same idea but, instead of changing how they are performed, the challenge was to integrate them in the physical object, and, given our limited time and expertise, do so without requiring extensive work.

Method employed: technical implementation

Once the elements to tangibilise and the goal to aim for are identified, we had to implement our ideas. The technical centre point of tangible interfaces are physical objects linked to computers. As with a lot of projects, there are multiple technologies that can make this link within our main constraints. We first list them, asking advice from other persons in the field for their opinions. Then we listed the pro and con besides our main constraints: like cost, time implementation, practicality, evolution potential, etc. With our problem, we fall into a Multi-Criteria Decision-Making (or MCDM) problem [Aruldoss et al., 2013]. This lead us towards image recognition as it was our understanding that we would be capable of implementing it using tags and could keep both the cost low and the evolution potential high with 3D printing our objects. Because of misidentification of the difficulty of image recognition this actually did not worked out as intended and other methods would have been better suited.

New tangible interaction design

With both the elements to tangibilise and the technology selected, the next step is to combine themes when designing the new tangibilised elements. The exact result will depend on both the technology and the specific interaction but the designed should keep in mind the ideas behind the tangible interfaces. To get a more tangible end interface, we can look at Fishkin's taxonomy, and try to aim for "Full embodiment" and "Full metaphor" if possible. The idea is to maximise the support from already existing knowledge for the user to create intuitive actions that need lower cognitive workload to perform and are easier to understand. The same way that semiotic and UX/UI designs link elements with similar shapes or colours, the T+C paradigm can be used to give similar information or expected interaction to user with tangible objects. Finally, the feedbacks can be design to be physical consequences (like putting a square in a slot) or be computer feedback (like a good or bad jingle when the token is in the good slot) or both.

4.6.3 Criterion for our tangibilisation's quality

PQ2 Which criterion can be used to judge the similarity of two interfaces?

Evaluating similarity

Our tangibilisation's main goal was to allow the A/B testing experiment which may not be the case for every tangibilisation. Some may only seek improvement without necessarily aiming for the measure of impact. In our case we come back to the idea of what makes an interface what it is.

We can first consider the affordances as the main point of commonality between interfaces as "ways to interact with" seems to be the central characteristic of interfaces. In that regard, we somewhat succeeded. Both version have tasks scheduling or removal, the ability to do so on specific patients and specific steps and the ability to access patients files and view workload gauges. After understanding how to use tokens with the image recognition, the interactions they allow is rather self evident and while some user did struggle to master their usage, this is not dissimilar to Clone's UI which can also take some time for user to master. For the affordance, we therefore kept the same affordance but with a different interaction method.

In terms of information on screen we also have some differences but not much. All of Clone's feedback as been recreated in one way or another. However those information are not at the same place or available at the same time. This is obviously because of the singular interface for the tangible version that can't really recreate both the global care plan and the individual care plan at the same time. Critically, not having the global care plan regularly in view can make the player miss some over charge on the gauges. There is a sub screen which show all the workload gauges to solve this but a lot of players did not use it. This is because they did not "experiment" with the token functionality and some never tried this special action to see what it would do.

For the overall appearances of both version are quite different. While the art direction was kept as similar as possible to respect our goal, there was some changes that were necessary for functionality. A quick glance at the interfaces of the version shows this.

Answering our question

The three elements that are key to judging interface similarity are, in order of importance: same affordances, same information feedback, and overall UI. In our case, except for the tangible aspect, we did keep the same affordances, the information were present but not presented in the same way but the UI is radically different. While the first prototype would have been way closer in term of overall UI we would still have a lack of the individual care plan. Regardless, we were unable to make this first prototype work so we had to make some choices on what to modify to produce a working version. It is very clear that those choices reduce the similarity, mainly regarding the overall UI, which harmed the quality of the results by introducing ever increasing biases. However, we feel those choices were necessary to complete the project with the means at our disposition and mediocre results are better than no results at all.

Synthesis Despite a long journey that forced us into multiple compromises, the third version is working and keep the Clone's core affordances for users. This version still uses image recognition and token with patterns and is linked to the Clone's software and its logging database for data collection. With the working prototype, we can now start the experiments with nursing students to compare both the original GUI and this new TUI.

Synthèse Malgré un long voyage qui nous a contraint à de multiples compromis, la troisième version fonctionne et conserve les fonctionnalités essentielles de Clone pour les utilisateurs. Cette version utilise la reconnaissance d'image et les patterns, que le première prototype avait initiée, et est liée au logiciel de Clone et à sa base de données pour la collecte de données. Avec le prototype fonctionnel, nous pouvons maintenant commencer les expériences avec les étudiants en soins infirmiers pour comparer à la fois la GUI originale et cette nouvelle TUI.

Chapter 5

Experimentation and results

Introduction The workingTUI allowed us to start the experimental campaign and recruit some volunteer nursing students to test our Clone versions. Multiple IFSIs helped us by allowing us to present our project to their students and lend us rooms to conduct the tests. During the tests, participants would be assigned to one of the interfaces, play a specific scenario design for the experiment and at the end answer the questionnaires about their experience. Then, they would do it again, but on the other interface in order to maximise the collected data. We had 24 participants which allowed us to make a statistical analysis of the data and plot graphs of the results. The interpretation of those results is available in Appendix 6.

Introduction Le TUI maintenant fonctionnel, nous avons pu lancer la campagne expérimentale et recruter des étudiants en soins infirmiers volontaires pour tester nos versions de Clone. Plusieurs IFSIs nous ont aidés en nous permettant de présenter notre projet à leurs étudiants et en nous prêtant des salles pour réaliser ces tests. Lors des tests, les participants ont été d'abord assignés à l'une des interfaces, ont joué au scénario spécifique conçu pour les tests et à la fin ont répondu aux questionnaires sur leur expérience. Ensuite, ils ont de nouveau joué au même scénario et répondu aux questionnaires mais sur l'autre interface afin de maximiser les données collectées. Nous avons eu 24 participants ce qui nous a permis de faire l'analyse statistique des données et de tracer les graphes disponibles dans ce chapitre. La discussion des résultats est quant-à elle dans l'annexe 6.

5.1 Protocol

5.1.1 Participants

Nursing students

The study focuses on the target of Clone : nursing students. To that end, we worked with some of the IFSIs that helped develop the Clone’s original digital version. This relationship between our laboratory and the IFSIs helped us when we needed students for our study. The establishments of Albi, Castres and Toulouse allowed us to ask their students for their participation. Despite the numerous students’ promotions, getting participants was somewhat difficult and would have been impossible without the help of the IFSIs’ teaching personnel helping us asking for volunteers. For example, at Albi’s campus, we send e-mails, posters and the teachers organised a presentation for us to convince students. However, despite gathering over 20 students’ information, we managed to only get 4 of them to come in the end. This is understandable as our study isn’t a priority for them and french nursing students have very dense schedules. Both campuses of Castres and Toulouse were more successful with 11 and 9 participants respectively.

ID and anonymity

To insure anonymity, once the groups were created, we assigned each participant an ID. The ID was simply “PAR” (for participant) followed by an incrementing number. We then split our data into two tables: the identity-to-ID table and the ID-to-result table. This insulate our results from the identity of the person. All data are linked to this ID has the participant connect to the game using it and answer the questionnaire with it too. This is, however, not a perfect solution. Certain demographics don’t have enough people to insure anonymity using this method.

Feedbacks to participants

All participants and their establishment will receive a result summary in accordance with the guidelines asked by the ethical committee. The idea is both to have an outlet to thanks them again for their time investment in our research and also to give them a well deserved overview of the results as they only got their own participation (or maybe the person’s they have done the experiment with).

5.1.2 Experimental steps

Introduction

The experimental steps are : waiver signature, group assignment, first play, questionnaires, second play and questionnaires. Depending on their assigned groups, participants would play either the numeric then the tangible version or the other way around. To reduce the time taken by the experiment, as much as possible, two students would participate simultaneously and swap version in the middle.

Waiver

In accordance with the recommendations of the ethic committee of Toulouse’s “Paul Sabatier” university, we gave all participant a form of informed consent. This form included the researchers’ contacts, our goals for this study, the risks and gains and, the rights of participants. When they signed the waiver, we also made the fill a form with their identity information. It is this form that would provide them an ID used during the rest of the experiment.

Group assignment

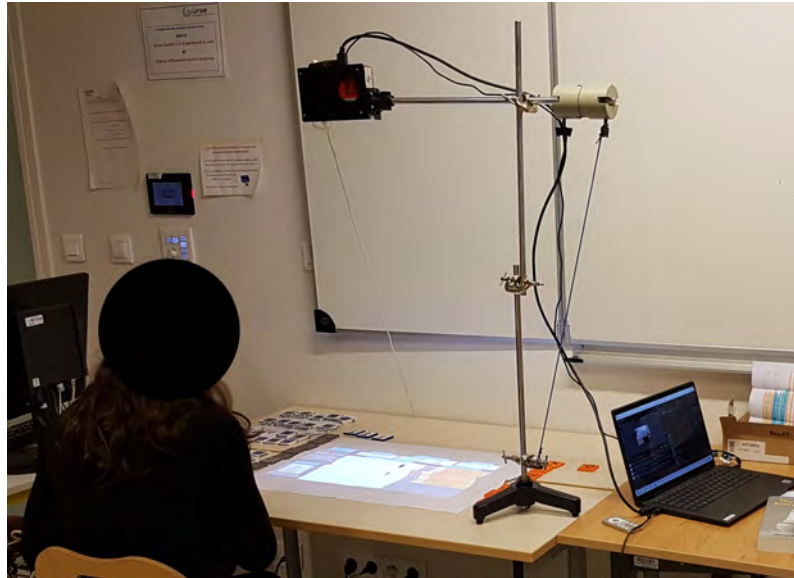
We used a pseudo-random method to create the two groups. This meant that we randomise one of the participant group and, necessarily, the second participant was assigned to the other group. The main bottleneck for this experiment is the number of participants. While we want to have randomised groups, we also need to try and have homogeneous groups. This constraint is particularly important on some demographic groups that are underrepresented in our population such as men or participant above 45 years old.

Plays

The participants were put in front of the game and told their goal (e.i. they would have to schedule the shift of a nurse and her assistant). Participants were reminded that they could ask questions and some did. Help given or comment from participants were noted. Sometimes, participants had to wait when switching version as each participant takes a different time to do the scenario. Questionnaires were a good cushion for that as it could be done while waiting to change.

Questionnaires

Questionnaires were given using a tablet. They were hosted on the Albi’s university’s server and used the web app LimeSurvey. Both versions had the same questionnaires and the questions were about the experience of the version they just played. To make sure, participants were also told that the question were about the version they just played (and for example, not about the overall experience when they would answer the questionnaire the second time at the end).



(a) Tangible Version



(b) Graphical version

Figure 5.1: The two versions in practice. During the experiments, we used a portable computer for the tangible version and a regular computer for the graphical version.

5.2 Data

The questionnaires are presented in the Questionnaires section of this document.

5.2.1 Questionnaires: Data Clean up

For the questionnaire a few clean ups were needed. Firstly, some questions were misunderstood or missing but the information could be found elsewhere. For example, the first batch of participants did not have the full “previous Clone use” questions but given that we knew their promotion classes, we were able to fill the question for them. Some also had so mistakes with wrong ID values. Because of the LimeSurvey’s system of questionnaires and the fact that we changed some of the

elements of the questionnaire our data exports were split. We then have to combine the files and do some column modification to match the versions.

5.2.2 Questionnaires: General Information

Our participant recruitment was sent broadly to the different promotion of students. Given that the nursing population is overwhelmingly female, the DREES give a 13% male ratio for french nurses [DREES, 2021], our participants were also mostly female. So this is not surprising that, in our study we got a ratio of 8% (from 2 men and 21 women).

5.2.3 In-Game: Data Clean up

Database and data amount

Firstly, during our experimentation, we used multiple PCs for our software. In total, there was 3 of them, one was used for all digital conditions and two were used for the tangible version. The in-game data were collected on a local data base for each of them to insure that they could be collected even without an internet access. This meant that the first step of the clean up was to combine those database value. This was done by copying the database, increase some of the primary key index across some data tables before the combining.

Interface condition detection

Due to the nature of the code, it would be quite precarious to change the database model. This however is not too much of an issue for the events as they are generic with 3 generic fields that simply need to be string of text. This however meant that we could not simply, for the “game” table add a Boolean to indicate whether or not it was a tangible or digital game. To get that information back, we used the both special events specific to the tangible version or modified primary key as they would give us the machine it was played on and in turn the version of interface used.

Event logs issues and games crashes

The log manager in Clone’s software is a bit unreliable. This is due to the server system it uses for its data collection. In our case this caused some of the experimental database’s data to not be recorded correctly. This forces us to do some work on the available data by removing games on an interface where the other interface for the same player was not available. In addition some of the events, has we will be discussing below, are not well monitored and seems to have been recorded multiple times for the same action which, should be taken into during the interpretation of results. Obviously, this reduce the size of our results but this solution is better than to either ignores the traces completely. In addition, the system being a bit unreliable caused some restarts at the start of games (before 5 minutes) which also had us clean the superfluous recordings.

Statistical treatment of the data

The data analysis was done using the R programming language and multiple of available libraries. The main function used throughout the analysis is “t.test”. This function is designed to compared sets of data. Either a data sample to a normal distribution or two data samples between themselves. This is done using the Student test formula. In our case, we are comparing the same data for two populations. The R function returns us a few values that can be seen in all the results tables below:

- ▶ **t** the statistic result of the test. Higher values indicate a more significant difference between the two samples provided
- ▶ **df** the degrees of freedom of the sample
- ▶ **p-value** the p-value result of the test. On the opposite of the statistic t, low p-values indicate a more significant difference between the two samples. This value is usually compared to a certain confidence threshold. We chose to use <0.05 as it is the most commonly used threshold but <0.01 is also used. We can also note that in some very high risk test, a very high confidence can be required which can mean an even lower threshold.
- ▶ **confidence interval** the mean of the sample with the variance added or subtracted. This is represented in the graphs below with the coloured box. The two values form the upper and lower side of the box.
- ▶ **mean** the means of the two sample compared. This is represented in the graphs below with a thick black line.

5.3 First test results

5.3.1 Introduction

Those results are obtained when selecting the answers from the first interface participants used. Therefore participants that did tangible TUI then digital GUI here are in the TUI group and the others are in the GUI group.

5.3.2 Attrakdiff(FR) questionnaire

As a reminder QP is for Pragmatic Quality, ATT is for Evaluation construct, QHI is for Hedonic quality - Identity and QHS is for Hedonic quality - Stimulation.

	t	df	p-value	confidence interval		Mean	
						GUI	TUI
QP	4.5219	18.512	0.0002475	1.2125	3.3091	1.4156	-0.8453
ATT	1.4372	20.085	0.1661	-0.3393	1.8436	1.4545	0.7024
QHI	0.7049	20.969	0.4886	-0.5383	1.0902	0.8831	0.6071
QHS	-1.608	18.324	0.1249	-1.6190	0.2141	1.1429	1.8452

Table 5.1: First interface played questionnaires results for the attrakDiff

The main value we are interested in is the p.value. The usual threshold to disprove the H_0 hypothesis and conclude that the two group are significantly different on this metric is $p < 0.05$. On the attrakDiff, for the first experiment, only the QP value is below 0.05. In those condition, the GUI score is higher (better) than the TUI.

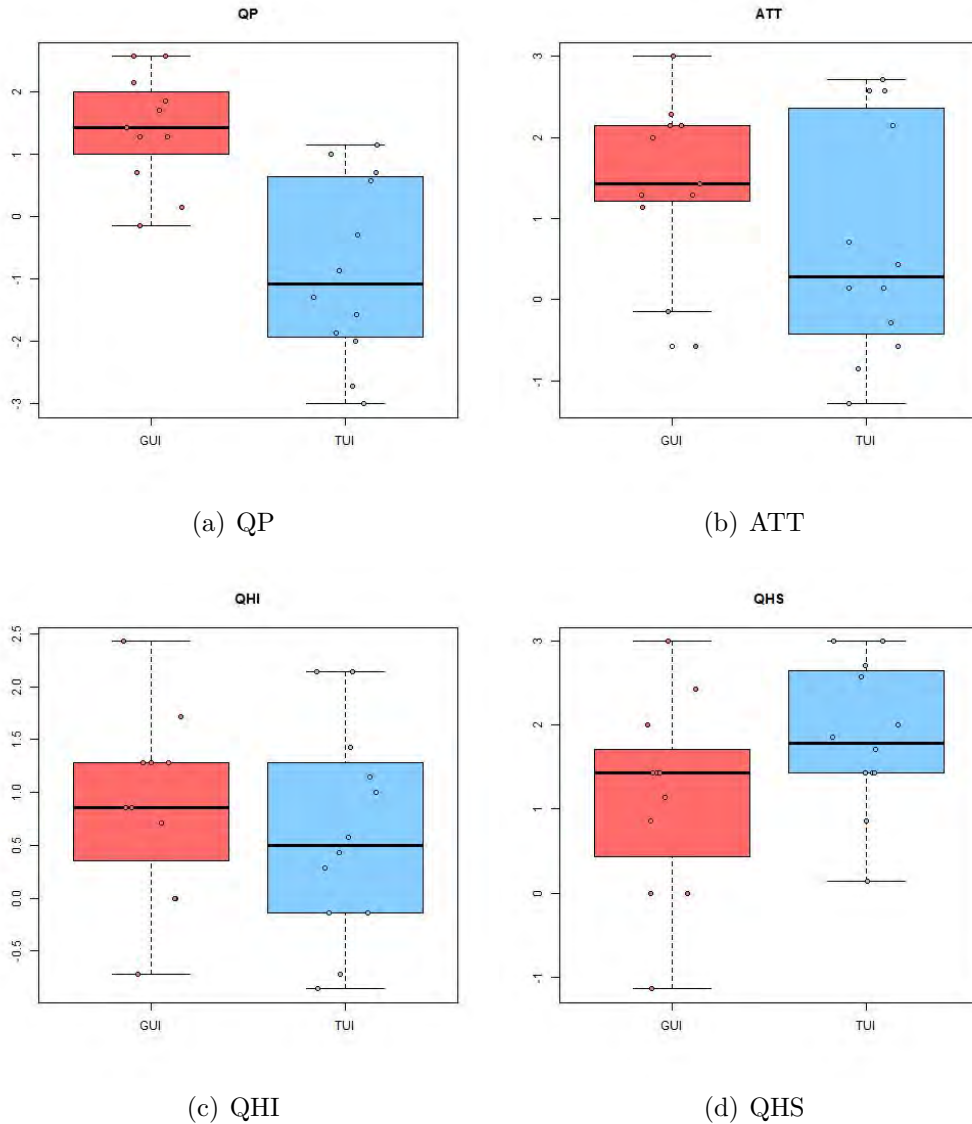


Figure 5.2: The graph for the first interface played questionnaires results for the attrakDiff questionnaire. Each results contains the boxplot of both the GUI (red) and TUI (blue) conditions as well as the values as a cloud of point overlaid.

5.3.3 F-SUS questionnaire

	t	df	p-value	confidence interval		Mean	
						GUI	TUI
F-SUS	1.4076	20.858	0.174	-0.6817	3.5340	7.1136	5.6875

Table 5.2: First interface played questionnaires result for the F-SUS

In the case of the F-SUS, the results have a p value too high so the differences are not enough to conclude of an impact with those data.

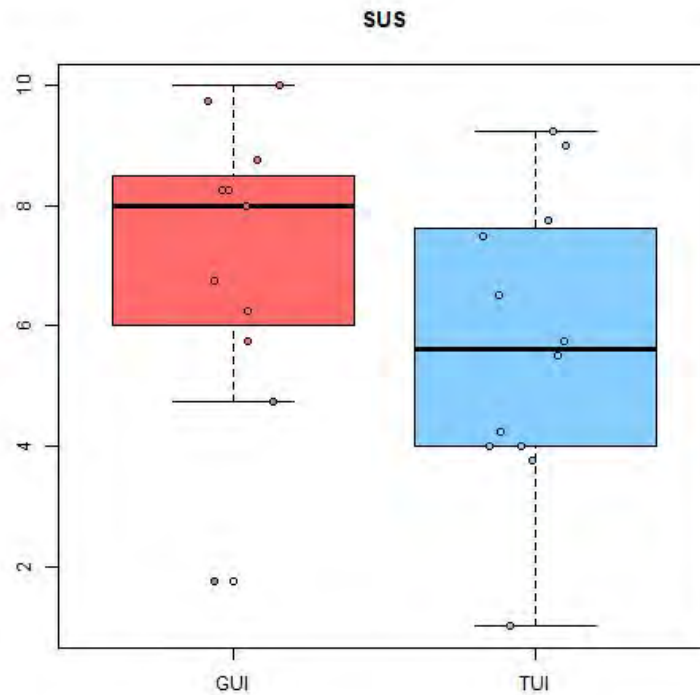


Figure 5.3: The graph for the first interface played questionnaires results for the F-SUS questionnaire. The graph contains the boxplot of both the GUI (red) and TUI (blue) conditions as well as the values as a cloud of point overlaid.

5.3.4 NASA-TLX(FR) questionnaire

	t	df	p-value	confidence interval		Mean	
						GUI	TUI
NASA-TLX-FR	-2.2275	20.287	0.03739	-29.4496	-0.9800	31.5909	46.8056

Table 5.3: First interface played questionnaires result for the NASA-TLX-FR

Similarly to the F-SUS, the NASA-TLX-FR, the results also have a p value too high.

5.4 First or Second interface played difference results

5.4.1 Introduction

Those results are obtained when we compare either the first interface use with the second for each group (TUI/GUI) respectively. Meaning we compare the tangible

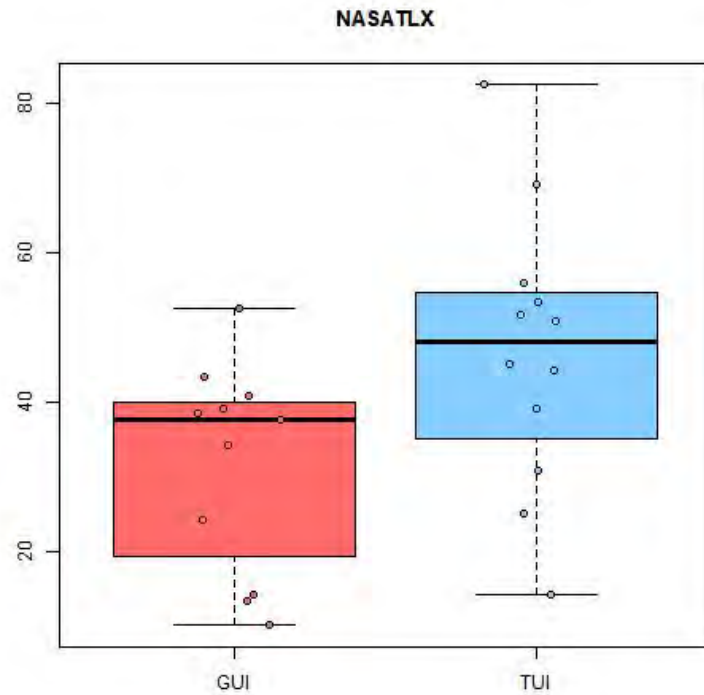


Figure 5.4: The graph for the first interface played questionnaires results for the NASA-TLX-FR questionnaire. The graph contains the boxplot of both the GUI (red) and TUI (blue) conditions as well as the values as a cloud of point overlaid.

results of both group of participants and the digital results of both group of participants. The aims is to see if there is an impact from already having played Clone right before.

5.4.2 Results for all questionnaires

	t	df	p-value	confidence interval		Mean	
						First	Second
QP GUI	0.927	20.8669	0.364528	-0.4727	1.2324	1.4156	1.0357
QP TUI	-2.5314	17.6195	0.021123	-2.2613	-0.2084	-0.8452	0.3896
ATT GUI	-1.0461	20.9958	0.307417	-1.452	0.4801	1.4545	1.9405
ATT TUI	-1.5959	17.3928	0.128512	-1.7749	0.2446	0.7024	1.4675
QHI GUI	-0.7675	20.1927	0.451652	-1.2307	0.5683	0.8831	1.2143
QHI TUI	-0.1511	20.1912	0.881386	-0.8166	0.7063	0.6071	0.6623
QHS GUI	-1.9948	14.892	0.064702	-1.6504	0.0552	1.1429	1.9405
QHS TUI	0.9537	20.7323	0.351229	-0.4159	1.1193	1.8452	1.4935
F-SUS GUI	0.3283	20.8427	0.745987	-1.9411	2.6684	7.1136	6.75
F-SUS TUI	-0.1959	20.3528	0.846621	-2.0494	1.6972	5.6875	5.8636
NASA-TLX-FR GUI	-0.4622	20.6011	0.648774	-16.8543	10.7306	31.5909	34.6528
NASA-TLX-FR TUI	1.063	20.995	0.299852	-7.5954	23.4793	46.8056	38.8636

Table 5.4: First or Second interface played questionnaires results for all the questionnaires

5.4.3 Graphs for all questionnaires

Attrakdiff(FR) questionnaire

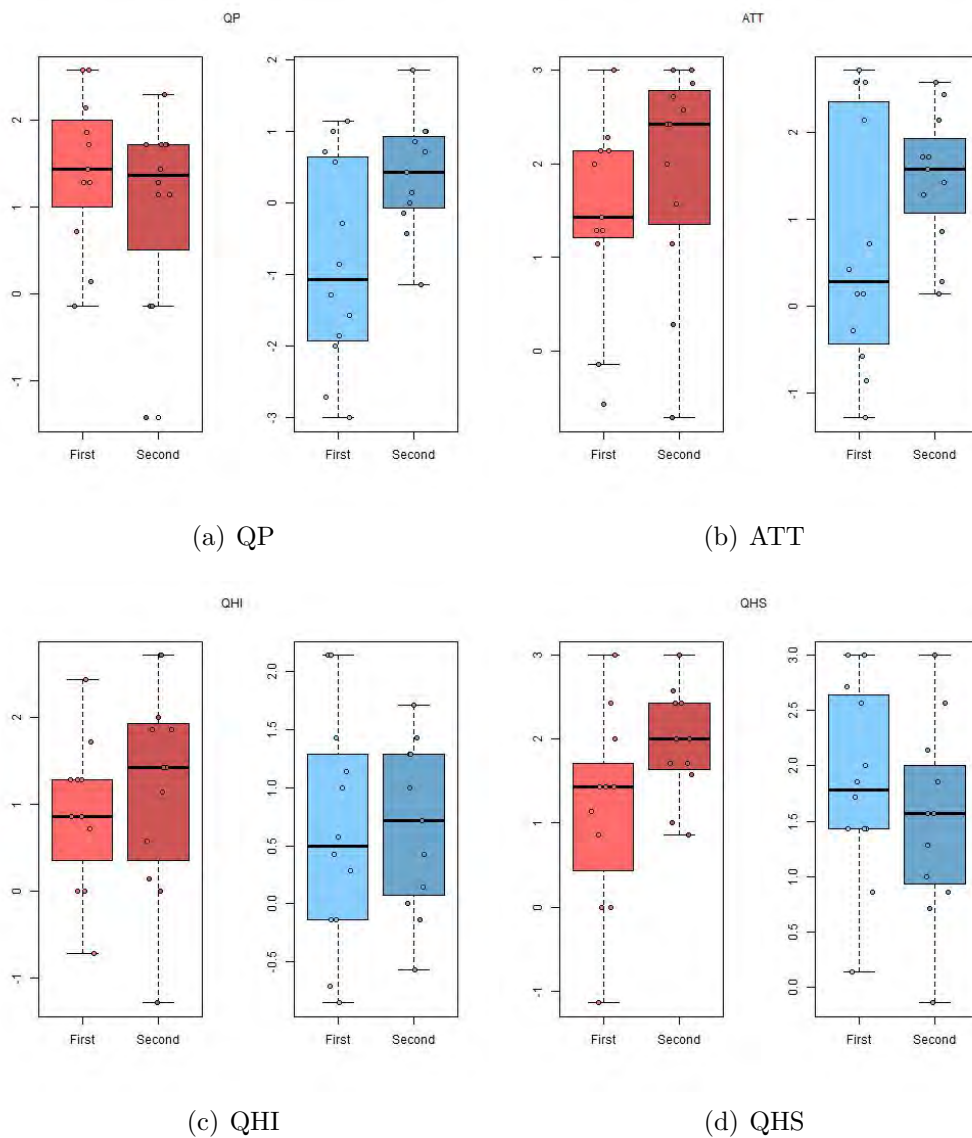


Figure 5.5: The graph that compares the results for participants playing the interface as their first interface of the test, or as their second interface for the attrakDiff questionnaire. Each results contains the boxplot of both the GUI as first (red), GUI as second (dark red), TUI as first (blue) and TUI as second (dark blue) conditions as well as the values as a cloud of point overlaid.

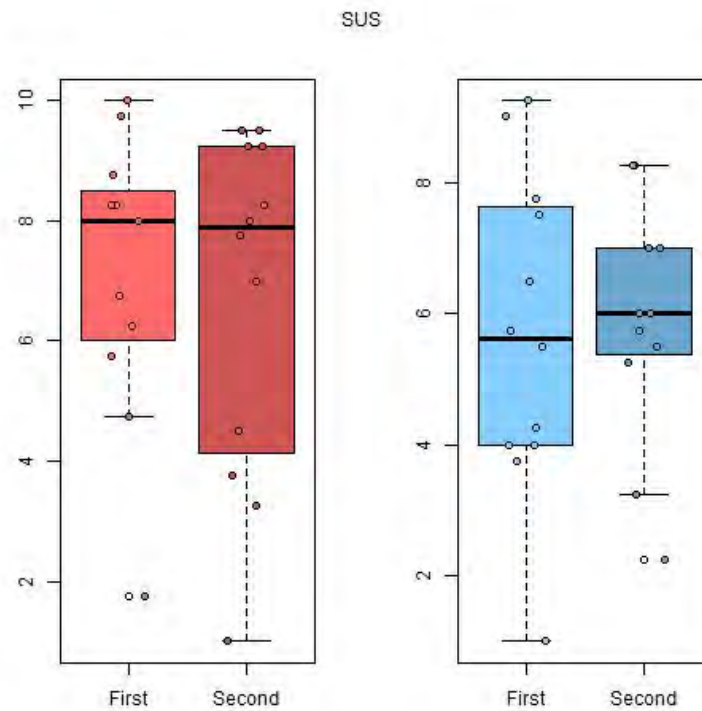
F-SUS questionnaire

Figure 5.6: The graph that compares the results for participants playing the interface as their first interface of the test, or as their second interface for the F-SUS questionnaire. Each results contains the boxplot of both the GUI as first (red), GUI as second (dark red), TUI as first (blue) and TUI as second (dark blue) conditions as well as the values as a cloud of point overlaid. questionnaire.

NASA-TLX-FR questionnaire

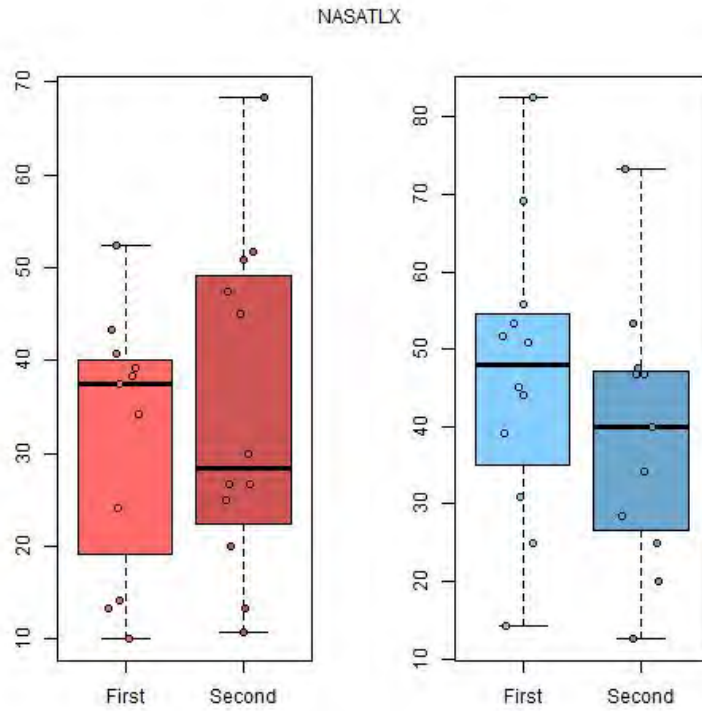


Figure 5.7: The graph that compares the results for participants playing the interface as their first interface of the test, or as their second interface for the NASA-TLX-FR questionnaire. Each results contains the boxplot of both the GUI as first (red), GUI as second (dark red), TUI as first (blue) and TUI as second (dark blue) conditions as well as the values as a cloud of point overlaid.

5.5 In-Game data's Results

Game duration

The game duration had some issues with the data collection as explained above. However, even when only taking into accounts games. For all of those results we focused on the first game played to try and reduce the variance and improve interpretability later on. Despite that, the results are not conclusive enough.

	t	df	p-value	confidence interval		Mean	
						GUI	TUI
Time Played	-0.4345	13.2645	0.670921	-969	644	2133	2295

Table 5.5: Time of the last event registered (in seconds) for each games depending on the type of interface

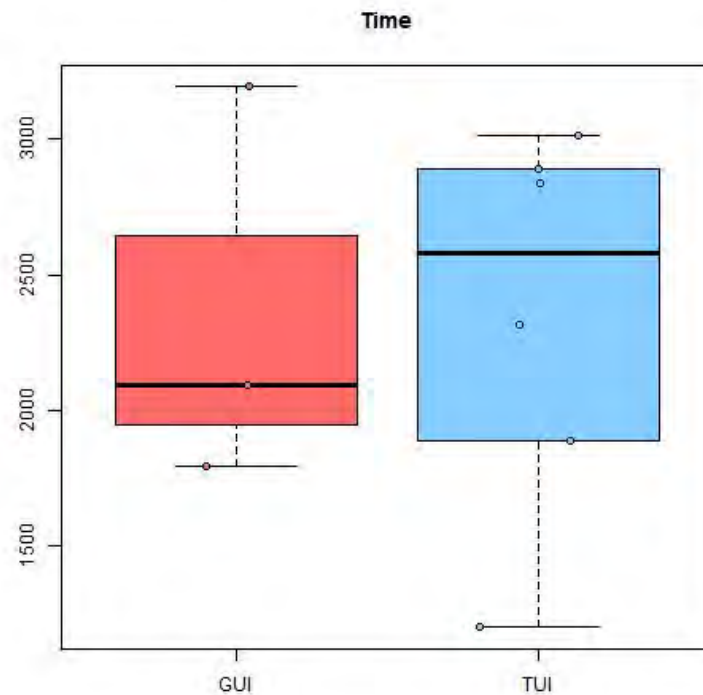


Figure 5.8: Time played for both versions. The time played is obtain by looking at the game time for the last event logged

Activities added and removed

The activities traces are the ones that caused us the most trouble. This is due to the combination of all the previously cited issues of game duplicated, missing data and, mostly for activities removed, duplicated events logged. In addition, a issue with our implementation of the image recognition in the tangible interface caused some of the patterns to be constantly detected even when absent. This forced some participants to have to removed the erroneously added pattern quite a few times which obviously should not be taken into account. All of this means that, while the data are presented, they should not be considered very robust and not much can be extracted from them.

	t	df	p-value	confidence interval		Mean	
						GUI	TUI
Added	-1.3087	6.3153	0.236235	-55.9969	16.6635	59.6667	79.3333
Removed	6.915	2.1723	0.016243	170.3031	635.3635	432.6667	29.8333

Table 5.6: Number events for activities added and removed depending on the type of interface first played.

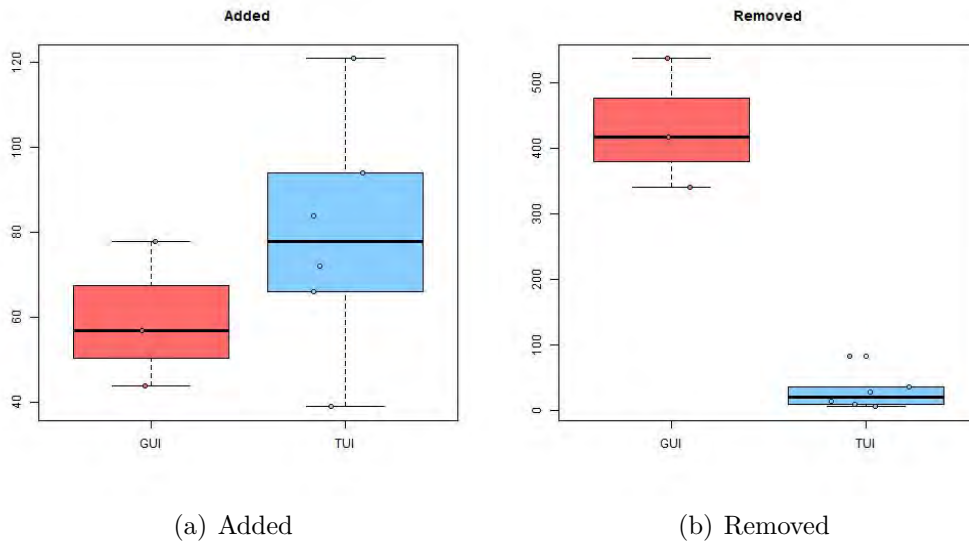


Figure 5.9: The two graph show the events logged by the system for added activities first played.

5.6 Results by demographic category

5.6.1 Gender

On our 24 participants we had 22 women and 2 men. During our experimentation, we made each of the two men do a different texting order (TUI then GUI and GUI then TUI). Despite that, we cannot realistically do a statistical analysis with only 1 value for the gender category.

5.6.2 Confidence on computer abilities

Computer Confidence Bool

We can then look at the questionnaires results differences depending on the participant confidence in their computer abilities. In the demographic questionnaire we have the 5 point scale question (see Figure 5.10)

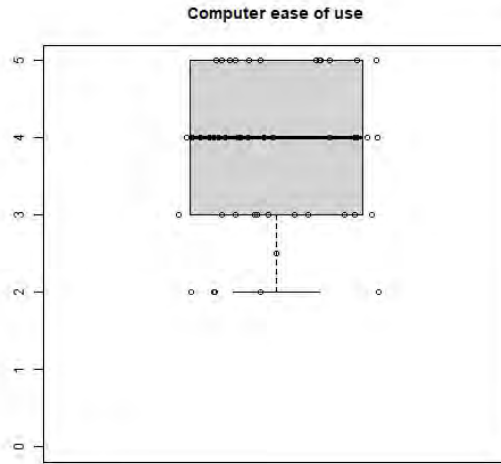


Figure 5.10: Computer ease of use boxplot

For the Student’s test, we need a 2 variables metric so we started with looking at the distribution of this question answer to see if we could observe some patterns. In the answer, the mean is on 4 with a no students putting 0 or 1 as their answer. We decided to split those answer into the “Computer Confidence Bool” where participants are score TRUE if they answer 4 or 5 in the previous question and FALSE otherwise.

Results

Using this criteria, we did the same analysis as before looking at first interface played between the tangible and graphical one.

	t	df	p-value	confidence interval		Mean	
						FALSE	TRUE
QP GUI	-0.0575	1.0981	0.962813	-12.8929	12.75	1.3571	1.4286
QP TUI	-0.0766	8.371	0.940703	-2.1409	2.0021	-0.8857	-0.8163
ATT GUI	-7.4699	3.3159	0.00349	-3.1088	-1.3198	-0.3571	1.8571
ATT TUI	0.3454	7.6204	0.73916	-1.8025	2.4311	0.8857	0.5714
QHI GUI	-3.8028	8	0.005216	-1.7339	-0.4248	0	1.0794
QHI TUI	-0.1863	9.6976	0.856035	-1.4337	1.2133	0.5429	0.6531
QHS GUI	-3.2637	1.5795	0.111768	-5.697	1.5066	-0.5714	1.5238
QHS TUI	-0.3823	5.731	0.715991	-1.6777	1.2287	1.7143	1.9388
F-SUS GUI	-0.4925	1.6043	0.68136	-10.9813	9.1758	6.375	7.2778
F-SUS TUI	-0.3508	9.9837	0.733045	-3.6242	2.6385	5.4	5.8929
NASA-TLX-FR GUI	-0.0193	1.0661	0.987568	-237.6757	236.8423	31.25	31.6667
NASA-TLX-FR TUI	-0.6813	8.5016	0.513806	-33.3518	18.0185	42.3333	50

Table 5.7: First interface played questionnaire answers depending on the participant self reported computer ease of use

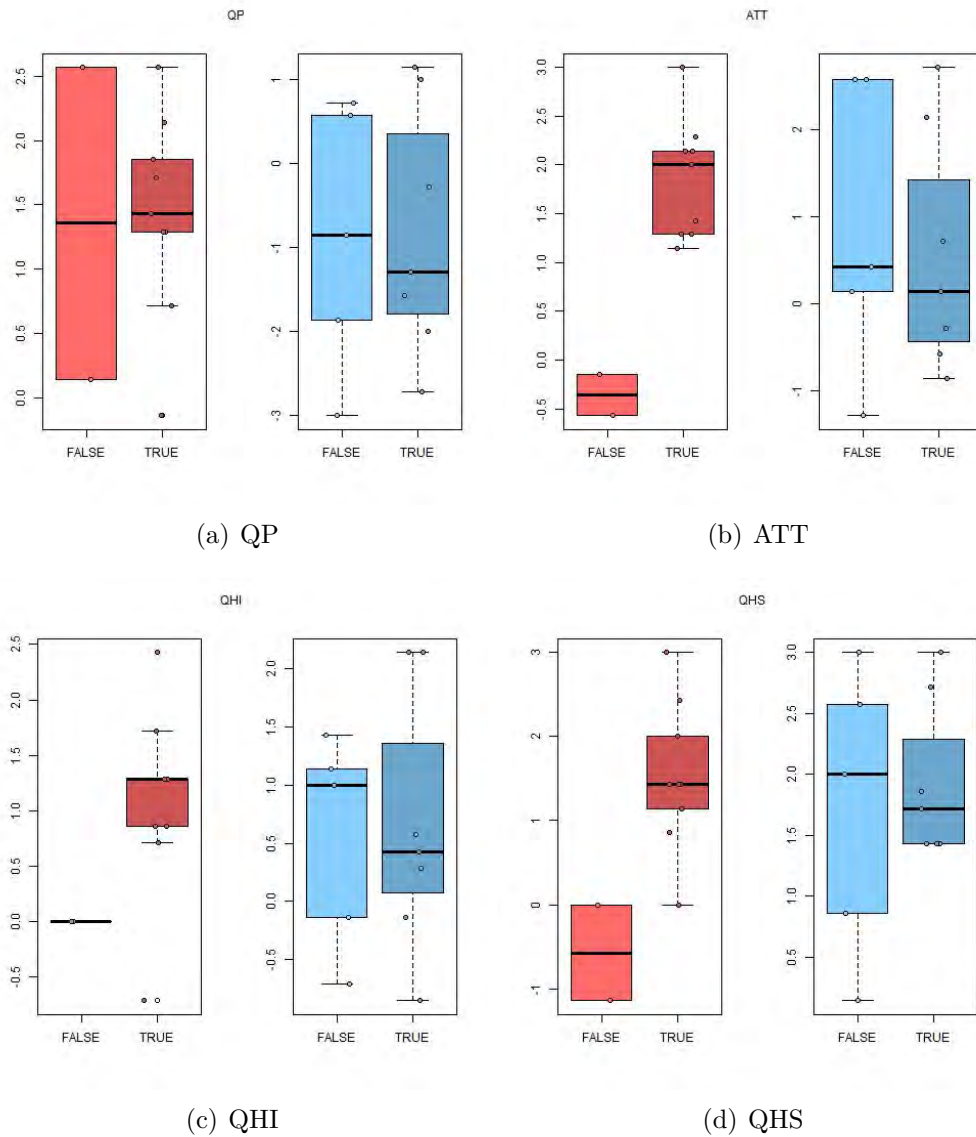


Figure 5.11: This graph compares the results for the attrakdiff questionnaire depending on the participant confidence in their computer abilities

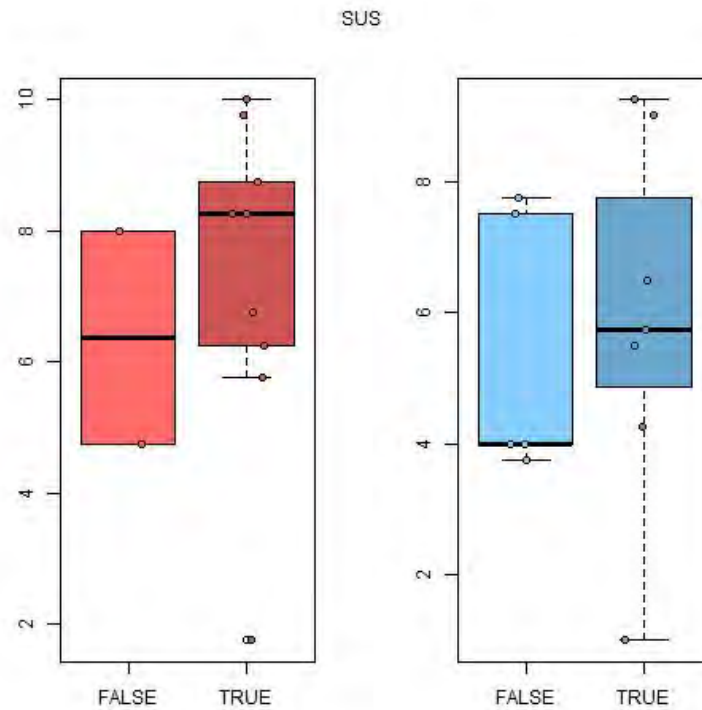


Figure 5.12: This graph compares the results for the F-SUS questionnaire depending on the participant confidence in their computer abilities

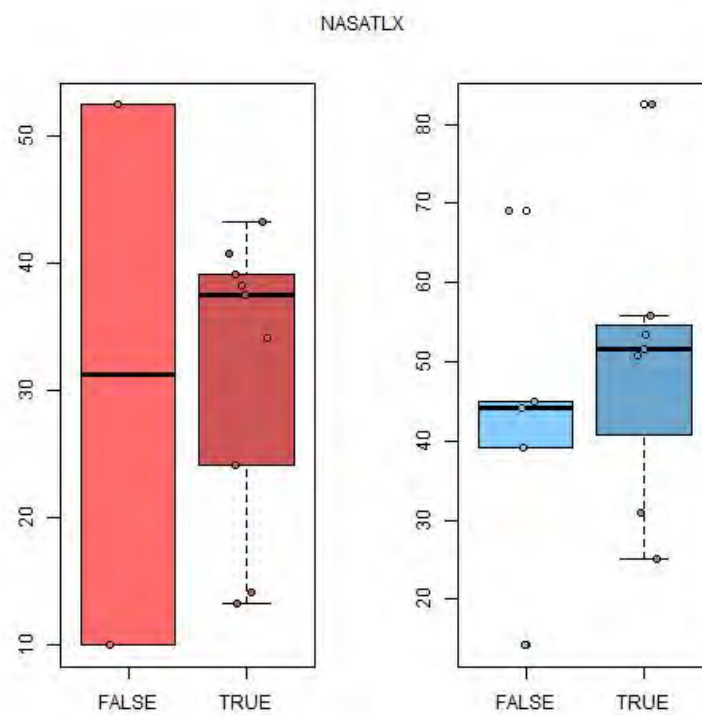


Figure 5.13: This graph compares the results for the NASA-TLX questionnaire depending on the participant confidence in their computer abilities

Synthesis The experiment gathered 24 students and 23 sets of data. We have done multiple analyses: looking at the “first interface played” results as are main focus, the differences on measures caused by either playing an interface first or second, and the demographics data that was possible to extract despite some technical issues. With a cursory view at the graphs and tables, we can see that not all results confirm our hypotheses. In the next chapter of this work, we offer an interpretation and some discussions on those results.

Synthèse Les expériences ont rassemblé 24 étudiants fournissant 23 ensembles de données. Nous avons procédé à de multiples analyses : nous nous sommes concentrés sur les résultats de « première interface jouée », sur les différences de résultats causées par le fait de jouer une interface en premier ou en second, et sur les données démographiques qu’il a été possible d’extraire en dépit de quelques problèmes techniques. En examinant les graphiques et les tableaux, nous pouvons constater que tous les résultats ne vont pas dans le sens de nos hypothèses. Dans le chapitre suivant de ce travail, nous proposons une interprétation et quelques discussions sur ces résultats.

Chapter 6

Interpretation and discussion

Introduction Having collected and the data into usable results, we can now look at them to conclude whether or not our initial hypotheses were good or misguided. Most of our results are non significant which, in the most part, seems to indicate the lack of impact from the TUIs on the metrics measured. However, that fact must be tempered by the compromises and bias introduced during the design of our prototype. The direct feedbacks that complain on the technically unreliable nature of our tangible version and the differences between the interfaces means that more work could make more apparent the possible influence of the tangibility.

Introduction Après avoir collecté et traité les données pour obtenir des résultats exploitables, nous pouvons maintenant les examiner pour déterminer si nos hypothèses initiales étaient bonnes ou erronées. La plupart de nos résultats ne sont pas significatifs, ce qui, pour l'essentiel, semble indiquer l'absence d'impact des TUIs sur les paramètres mesurés. Cependant, la valeur de ces résultats reste modérés de par les compromis et les biais introduits lors de la conception de notre prototype. Les retours d'information directs qui se plaignent de la détection ératique de notre interface tangible et des différences entre les interfaces signifient qu'un travail supplémentaire pourrait permettre de rendre plus évident un éventuel impact de la tangibilité.

6.1 General interpretation of the experimental campaign's feedbacks and unfolding

6.1.1 Experimental campaigns

The experimental campaigns were good. As alluded above, the new conditions exacerbated some of the issues with Clone's image recognition. We have not yet completely identified the full reasons but it seems that lighting play a role as well as potential damage on the paper pattern on the back of some tokens. Even more problematic was the detection issue with some specific patterns. It is our understanding that some pattern are "too geometric". They keep being detected by the camera despite not being there when said geometric forms appear in the rest of the camera's field of view. This problem plagued the tangible version during the experiments while having only occurred once or twice during testing. This could be fixed by skipping the problematic patterns but would require a bit of work that was judged risky in the middle on experimentation. This however bring the topic of more testing before hands to make sure to eliminate such issues. Our mistake was to not anticipate the issue and not test in more diverse conditions. Each of the IFSIs outside of Albi's dedicated us a room for our experiments which greatly help to keep the experiments calm and without distractions. Obviously, for Albi's, participants were brought in one of the laboratory many available room and got a similar experience. It should also be noted that some participants did discussed lightly between themselves when playing. This was kept at a minimum, only answering some questions asked to the research or some light remarks on the game overall.

6.1.2 Direct feedbacks

A lot of the participants were convinced of going by their teachers through their interest in research. Interest that was also brought by their, either future of present depending on their year at the school, course on the topic of research. Participating in our experiment was viewed as a first step or a practical experience in the research that would help them later on. It should also be noted that despite that personal benefits, during our discussion with the students, they expressed a genuine interest in our work. Their participation should not be reduced at their personal interest. All of this meant that most students were quite happy and interested in our work and, despite some issue with the software and prototype, were determined to finish the experimentation. Only one person stopped before the end and did not return afterwards, admitting to us a bit of stress during exam periods at the time. In terms of times, most students had enough time with the allocated 45 min per interface scheduled. Some did take a bit more time which was not much of an issue except when the other student had some appointments.

6.1.3 Need of explanation and questions

One thing that was not clear defined in our protocol was the amount of help the researcher was to give stuck players. As presented above, Clone is a complex game. While it most participants had already used Clone before, even they sometimes

needed a refresher when it had been a long time. To reduce bias, it was decided to give the same information at the start of the game and answer questions when prompted. Students were reminded or told the goal of using the patients files to schedule the shift of the IDE and AS and the importance of keeping the workload gauges below their maximum for each steps. For the graphical interface, they were also explained about revealing information, where and how to add and remove activities and how to delegate. For the tangible interface, they were explained how to use a token with the camera, the interaction available with the activities, steps and patients tokens and the roles of each of the special action tokens. For the tokens, they were also shown the name on the token side and given a demonstration on how to use them (or have had to use them for a first time with guidance from the researcher). When prompted the researcher also answered the participants questions on either those topics again or on other subjects.

Most questions were asked for the tangible interface. Most were about abnormal and unwanted behaviours of the application. Some were about the specific of activities. Players only have the icon and name of activities to understand what was meant and sometimes they might look for something without finding it. This happened in both version but somewhat more with the tangible. Some players did not understand the workload gauges until the end of the game. The game cannot proceed without having all gauges below their maximum which usually prompted a question from players. They, for the most part, quickly resolved the issue as the 3 patient scenario do not require a lot of work in that regard. One of the other issue was with some special activities : the breaks and meal. The graphical version of Clone show meal as already scheduled activities which can be disorienting as those cannot be scheduled manual because they are not meant for either the IDE or AS but the cook help. This lead to some question on how to proceed which were answered without issues. For the breaks, they do not really target a specific patient. In Clone graphical, this is solved by scheduling the break activity on all patient but only counting a total of one break time. In Clone tangible, this was not taken into account, and some players did give the IDE a break for each patient. This only lead to some issue with the workload gauges which was an easy fix when pointed out by the researcher.

6.2 Interpretation of the data collected

6.2.1 Questionnaires' feedback : first interface played

Results from questionnaire are the most interesting ones. Importantly, they are the most numerous as they had little issues by not being integrated in Clone's logging system. Their results are what we will be focusing on more.

Non-significant

Starting with the first interface played which is the most direct results. Most of the results are not statistically significant meaning that, in our participants experience of our prototype, those metrics were not much different between the two versions. This includes ATT, QHI, QHS for the attrakdiff and the F-SUS while the NASA-TLX was just below the 0.05 threshold and the attrakdiff's QP well below it with 0.02%. For those results we can still look at their values to informs us on the overall scores

of both version. For the attrakdiff's scores, they are all between -2 and 2. Values between 2 and 3 (or -2 and -3) are usually what is considered indicative enough. This indicates that, while on those metrics our interfaces did not fare terribly, they also did not excel either. Also, while not significant, for all 3 of the values (even the fourth significant one), the tangible interface had lower scores than the graphical one.

NASA-TLX

The NASA-TLX shows that the TUI, with an average of 46, is more demanding than the GUI, with 31. Despite that, on this scale between 30 and 49 is considered "Somewhat high" [Prabaswari et al., 2019] which means both interfaces need to be improved on that aspect. This confirmed the feedback for Clone that were some of the impetus for this work, it is very demanding for players and it seems our TUI did not solve the issue.

QP

The QP's results on usability give similar results to the SUS but with bigger mathematical validity. It suggests that, again, the GUI is preferred by users. Again, the mean scores of 1.4 and -0.8 are still near the 0 where on the scale meaning this is mediocre for both interface on this metric.

6.2.2 Questionnaires' feedback : first and second compared

Those results are less central to our analysis but can still discuss them to find some insights into our issues. All but one of the results are non significant. That indicates that there is not much impact on the answers if the player has played the other interface beforehand. Still, it should be noted that, overall most scores improved for both the graphical interface (except for QP and SUS) and the tangible one (except for QHS and NASA-TLX). If there was more significant results this would indicate that people become more tolerant in their second answers.

6.2.3 In-Game measures

Game Duration

The game duration results is not significant implying that both versions takes a similar amount of time to solve. This is actually something we observed during the experimentation with most participants finishing almost at the same time for the swap of interface at the mid-point and again at the end. It was this research's opinion that the tangible version took slightly more time which can be seen in the mean results but the differences does not seem to be enough. There is however one bias that can be discussed. The fact that, for almost all experiments, participants played side by side and with the understanding that they had to switch interface in the middle, that could have influenced their game time. This is even more of a possibility due to the end condition that only ask for the workload gauges to be under 100% which means some participants may have rushed at the end of their game to not bother their fellow participant waiting for their turn. It might have

been better to, if not having participants go one at a time, at least split them in two rooms to avoid this bias.

Action added and removed

As it was already explained, this measure cannot be used to conclude at all due to issues in the software. This is stop part of our analysis as this we could have used those results with the epistemic actions paradigm in mind. Maybe future developments will resolve this issue.

6.2.4 Demographics' interpretation

Gender

As explained when we presented our data the gender distribution of our participants does not allow us to do our analysis due to Men being to underrepresented. Some bigger experimental campaign could help solve this issue.

Self reported computer abilities

The results for the self reported computer abilities do not show a lot of significant differences in either the graphical interface or the tangible one. There is only one value that is significant "ATT TRUE" with a p-value of 3.5% (and "QHI TRUE" almost at the threshold with 5.2%). This seems to indicate the absence of impact between people confident with computer and people that are not. It should however be noted that for our analysis we had to split the results into two categories from a 5 points scale which is not always the best representation of original answer intent. In addition, the results of the original questions were very high with a mean at 4 out of 5 which would indicate that our population may lack people actually bad with computers. Finally, that question is a self report of confidence and not a actual test of abilities. This also induce a bias that should be kept in mind when reading the results.

6.3 Discussion

6.3.1 Overall results feedbacks

Overall the results are not what would have wanted. Most of the data show the TUI being inferior to the GUI in most aspects. The number of participants being 24, those results are not necessarily a perfect representation of the nursing students body general reaction but still, our new version did not solve the issues it was meant to. The overwhelming part of the feedbacks given during the tests are focused on the detection issues of the tangible version. Those issues take a tremendous place in the bad user experience for the player. This cannot be overlooked when critiquing the results of the experimentation. While this is a let down for the current results from this wave of experimentation, this also means that a more robust version could clear up the bias from the difficult detection of the TUI where pattern can take a few seconds to be detected or some may be detected erroneously. A lot of the

questionnaires also have a huge variance in their answers that could indicate some other influence from metrics not included in this study.

6.3.2 Answering our metric based research questions

Metrics

If we come back to our initial research questions on measures :

- MUQ** Will the tangible version be more usable for users ?
- ▶ The QP results with a very low p-value indicate that it is the opposite with the GUI's version being better. The SUS result do not see a significant differences. Based on our small sample size, we could be inclined answer yes to this question.
- MWQ** Will the tangible version be more demanding for users ?
- ▶ The NASA-TLX results indicates that yes our tangible version is more demanding than the graphical one.
- MXQ1** Will the tangible version offer a better user experience ?
- ▶ Except for QP, none of the values pertaining to this question are significantly different between the two conditions which leads us to answer no to this question.
- MXQ2** Will the tangible version be a more attractive version than the numeric one?
- ▶ Again, the values are not significantly different meaning than no, both version are similarly attractive.

Demographics

Similarly as for the metric research questions we can look back at our demographics questions :

- MDQ1** Is there significant differences between men and women ?
- ▶ The data collected do not allow us to conclude.
- MDQ2** Is there significant impact from the age on the results ?
- ▶ Again, the data collected do not allow us to conclude.
- MDQ3** Does people lacking confidence in their computer literacy have a preference for the tangible version ?
- ▶ There doesn't seems to be an impact of computer literacy on the participant but a wider representation of the scale, notably people worst with computer, would change this result.
- MDQ4** Does having used Clone in class recently impact the results ?
- ▶ The first and second analysis does not show an influence of having just played the game on another interface on the results of the questionnaire. This lead us to conclude the lack of impact of the last play of Clone. However, this might be obscured by the re-explanation of the gameplay at the start of the experiment. Sadly this was necessary to avoid having participant unable to play at all.

6.3.3 Other research questions

The discussion for the other questions has already been explored in section 4.6. We chose to put this analysis right at the end of the relevant chapter for better reading.

6.3.4 Similarity in the interfaces' versions

All throughout this document we have talked about the necessity of design the TUI with similarity as one of the main constraint to abide by. Despite that, we have had to compromise on that principle in order to bring this project to fruition. The two interfaces still have the same main affordances with the ability to schedule and remove activities to specific patients and steps, they also convey the workload gauges of the shift for both agents (IDE and AS) and convey the patients files necessary. In that regards, the two interfaces have been kept similar. In most other aspects they are not.

The graphical interface has two interfaces for scheduling : the global care plan and the individual care plan. This alone is one of the biggest compromise made during our tangibilisation. This first lead to the fusion of both into a tangible global care plan that did not work but would still be a very different experience for the user. The information visible during the choices, the need of swapping between care plan and other changes due to the fusion would have been bias that could not be accounted for during our analysis. This was not completed but the second and third prototypes also have the same issues. Players do spend a large amount of times within the individual care plan, but the ability (and obligation) to come back on the global care plan and see the whole picture of the schedule when changing patient can (and in my opinion does) have a impact on the student's cognitive model of the schedule. To not have the same mechanism in the tangible version certainly has an impact that muddies the results. The project Clone was maybe too ambitious to tangibilise given our skill and resources available or another approach may have yielded better results. A smaller scale analysis of the issue with simpler subject would have lead to less constraints. The idea of improving Clone using TUIs would at least require intermediates works. Using more elemental prototype more focuses on specific aspects of tangibility and then apply them with intention and knowledge on the tangibilisation of Clone might have been a better solution than to focus on similarity so much. This idea of using a simpler target subject for the impact analysis was one of the original idea during the very early discussion around the this thesis direction, but this did not end up being the chosen direction. It is very possible that other people can learn from this research on reapply the tangibilisation principles to measure tangibility's impact.

6.3.5 Reach of the results

One can question the reach of our research. As already noted in 2009 by Horn et al. [Horn et al., 2009], the reach of the works on TUIs is limited. While our original intentions were to try and offer a method of determining the impact of those interfaces, this is way beyond the possible scope of our limited resources. However, if we focus our results on what was accomplished, we can see that our work brings another point of results on tangibility. We have shown, with the value that our

results have, that in scheduling for nursing students, using tangible interface might not be the first solutions to reduce cognitive load of complex problems.

6.4 Perspectives

6.4.1 Opportunities for Clone's future

Clone's graphical version is being worked on currently by a new developer and a company outside of the original members of the project (IFSIs, ARS and SGRL). This brings new opportunities for developments and solving Clone's original issue of mental demand. While those development will not be conducted in the laboratory, this new partner is still communicating with us and using our design inputs.

6.4.2 Better bases for the analysis

Again if we look at the idea of a simpler target subject that could be easily modified. One could envisioned not just a pair of versions TUI and GUI, but multiple of those pairs each targeting an aspect of the interface that is thought to have impact on the overall experience. By again using A/B testing but on a larger scale, one could draw a clearer picture of the relationship between tangibility and its connected metrics. Given the usual participant amount for such experiment, this would need to first be an exploratory work. Followed by, once the important metrics are identified, a more focused but bigger experiment to obtains more statistically relevant data. Obviously we would once again fall into having results about the specific interface we use but by making multiple, very different, version of it, it should be possible to extract criterion of quality for TUIs.

Synthesis This project has shown its limits with results that are for the most part non significant despite the selection of metric that were supposed to be impacted by the tangibility of interfaces. This can be, most likely, attributed to both the still dissimilar nature of the interfaces, notably the detection issues in the tangible version as well as potentially other factors not taken into account in this study. While the data provided contributes to the overall understanding of TUIs, the complexity and specificity of the subject chosen is somewhat of a burden of the generalisation of the results. We can envision future works re-threading the tangibilisation and A/B testing route but focusing more on simpler and more modifiable interface to produce an array of results.

Synthèse Ce projet a montré ses limites avec des résultats qui sont pour la plupart non significatifs malgré la sélection de mesures censées être impactées par la tangibilité des interfaces. Cela peut être attribué, très probablement, à la fois à la nature encore dissemblable des interfaces, notamment les problème de détection dans la version tangible ainsi que potentiellement à d'autres facteurs qui n'ont pas été pris en compte dans cette étude. Bien que les données fournies contribuent à la compréhension globale des TUIs, la complexité et la spécificité du sujet choisi pèsent quelque peu sur la généralisation des résultats. Il n'est pas impossible d'imaginer de futurs travaux reprenant la tangibilisation et les tests A/B mais se concentrant davantage sur une interface plus simple et plus modifiable afin de produire un éventail de résultats.

Conclusion

The third iteration of our tangible version of Clone's interface has allowed us to conduct our experimental campaign on the volunteer nursing students provided by the IFSIs. Those 24 students used the two versions of the interface which provided us with a wide range of data helping us conclude on the potential benefits brought by the design of the tangible interface. Those results have been disappointing as most data do not show significant differences between the two versions. This is due to a wide variance in the answers and data collected. In addition, the final working tangible interface was shown to not be completely free of issues. During the experiments this irredeemably impacted the user experience and the others measures. Finally, the compromises made during the design of the tangible interface introduced some bias which effects on the measures can not be fully quantified and further muddy the possible conclusions. Despite all those issues some metrics show some encouraging improvements and the presence of technical problems while still having relatively similar score in other measures could indicate that removing those issues would make the tangible version the better one.

In terms of general analysis of the tangible impact on measures, this work does not reach the original aspirations. Now looking from the end of this project, the idea of building a conceptual model for the interactions between tangibility and other UX linked concepts may have been too ambitious. However, this study still provided another brick that can be added to the overall understanding of tangible interfaces. Further works that could follow this one could focus more on smaller and more elemental problems. With the idea of using a more iterative approach which is not possible with a complex and, in that specific context, constraining subject interface like Clone's. For Clone's issues, further work is needed to fully polish the tangible version which might lead to better results in the analysis and its adoption. In addition, other avenues could be explored to solve the mental demand issues facing the students.

Résumé en français La troisième itération de notre version tangible de l'interface de Clone nous a permis de mener notre campagne expérimentale sur les étudiants en soins infirmiers volontaires fournis par des IFSIs. Ces 24 étudiants ont utilisé les deux versions de l'interface, ce qui nous a permis d'obtenir un large éventail de données nous aidant à conclure sur les avantages potentiels apportés par la conception de l'interface tangible. Ces résultats ont été décevants car la plupart des données ne montrent pas de différences significatives entre les deux versions. Ceci est dû à une grande variabilité dans les réponses et les données collectées. De plus, l'interface tangible finale s'est avérée ne pas être exempte de problèmes techniques. Pendant les expériences, cela a irrémédiablement affecté l'expérience de l'utilisateur et les autres mesures. Enfin, les compromis faits durant la conception de l'interface tangible ont introduit certains biais dont les effets sur les mesures ne peuvent pas être entièrement quantifiés et brouillent encore plus les conclusions possibles. Malgré tous ces problèmes, certaines mesures montrent des améliorations encourageantes et la présence de soucis de détection tout en ayant des résultats relativement similaires dans d'autres mesures pourrait indiquer que leur élimination rendrait la version tangible significativement meilleure que l'interface originale.

En termes d'analyse générale de l'impact tangible sur les mesures, ce travail n'atteint pas les aspirations initiales. L'idée de construire un modèle conceptuel pour les interactions entre la tangibilité et d'autres concepts liés s'est avérée trop ambitieuse. Cependant, cette étude a fourni une nouvelle brique qui peut être ajoutée à l'édifice tangible. D'autres travaux qui pourraient suivre celui-ci pourraient se concentrer davantage sur des problèmes plus élémentaires. Avec l'idée d'utiliser une approche plus itérative, ce qui n'est pas possible avec une interface complexe et, dans ce contexte spécifique, contraignante comme celle de Clone. En ce qui concerne les problèmes de Clone, des travaux supplémentaires sont nécessaires pour peaufiner la version tangible, ce qui pourrait conduire à de meilleurs résultats dans l'analyse et son adoption. En outre, d'autres pistes pourraient être explorées pour résoudre les problèmes de demande mentale auxquels sont confrontés les étudiants.

Appendices

Appendix B

Figures' and tables' list

List of Figures

1.1	Patients files in Clone GUI	19
1.2	Clone’s conceptual model	22
1.3	Clone’s cognitive process	23
1.4	Clone’s global care plan	24
1.5	Clone’s individual care plan	25
1.6	Clone’s execution interface	26
1.7	Results from the questionnaire of the 2021 Clone’s campaign	27
3.1	From [Dumas, 2010, p. 34], the multi modal human machine loop	50
3.2	TUI’s design spaces from [Fitzmaurice et al., 1995]	51
3.3	Token and constraints example from [Ullmer, 2002]	53
3.4	Tangible interfaces examples’ images	54
3.5	Tangible benefits key concepts, from [Hornecker and Buur, 2006]	56
3.6	Analytic framework from [Marshall, 2007]	59
3.7	Examples of RFID applications	61
3.8	Examples of NFC applications	62
4.1	Design of the first prototype	69
4.2	Photo of third prototype in laboratory	70
4.3	Designs of each modular board piece	71
4.4	Design of the modular board	71
4.5	Dimensions for the first prototype’s board with 3 patients	72
4.6	Photo of the first prototype	74
4.7	Second and third versions’ main UI mock ups	75
4.8	Third Prototype Main UI	76
4.9	Second prototype positioning example	77
4.10	Early pattern with complex border	78
4.11	Pattern squares numbered for decoding	80
4.12	ArUco markers	82
4.13	The detection triangle	83
4.14	Examples of activities icons	85
4.15	Categories for the physical ability tokens	86
4.16	Table with the tokens’ placement	87
5.1	Photos of the two version during experiments	94
5.2	First interface played attrakdiff graphs	97
5.3	First interface played F-SUS graphs	98
5.4	First interface played NASA-TLX-FR graphs	99
5.5	First or second differences attrakdiff graphs	101

5.6	First or second differences F-SUS graphs	102
5.7	First or second differences NASA-TLX-FR graphs	103
5.8	Time played in-game graph	104
5.9	Activities added or removed in-game graph	105
5.10	Computer ease of use boxplot	106
5.11	First interface played attrakdiff graphs with self reported ease of use .	107
5.12	First interface played F-SUS graphs with self reported ease of use . .	108
5.13	First interface played NASA-TLX-FR graphs with self reported ease of use	108

List of Tables

1.1	Clone's pedagogic progression table	18
3.1	OCIS to scheduling skills comparison table	45
3.2	From [Galaup, 2013] Children development linked to game types . . .	47
3.3	Results of paper comparing TUIs with GUIs	58
3.4	RFID operating frequencies from [Weis, 2007]	60
5.1	First interface played attrakdiff results	96
5.2	First interface played F-SUS results	97
5.3	First interface played NASA-TLX-FR results	98
5.4	First or second differences all results	100
5.5	Time played in-game results	103
5.6	Activities added or removed in-game results	104
5.7	First interface played all results with self reported ease of use	106

Appendix C

Questionnaires

C.1 AttrakDiff(FR)

ID		1	2	3	4	5	6	7	
QP 2	Simple								Complicqué
QP 3	Pratique								Pas pratique
QP 4	Fastidieux								Efficace
QP 5	Prévisible								Imprévisible
QP 6	Confus								Clair
QP 7	Incontrôlable								Maîtrisable
ATT 1	Plaisant								Déplaisant
ATT 2	Laid								Beau
ATT 3	Agréable								Désagréable
ATT 4	Rebutant								Attirant
ATT 5	Bon								Mauvais
ATT 6	Repoussant								Attrayant
ATT 7	Motivant								Décourageant
QHI 1	M'isole								Me sociabilise
QHI 2	Professionnel								Amateur
QHI 3	De bon goût								De mauvais goût
QHI 4	Bas de gamme								Haut de gamme
QHI 5	M'exclut								M'intègre
QHI 6	Me rapproche des autres								Me sépare des autres
QHI 7	Non présentable								Présentable
QHS 1	Original								Conventionnel
QHS 2	Sans imagination								Créatif
QHS 3	Audacieux								Prudent
QHS 4	Novateur								Conservateur
QHS 5	Ennuyeux								Captivant
QHS 6	Peu exigeant								Challenging
QHS 7	Nouveau								Commun

C.2 F-SUS

	1 (Pas du tout d'accord)	2	3	4	5 (Tout à fait d'accord)
1. Je voudrais utiliser ce système fréquemment					
2. Ce système est inutilement complexe					
3. Ce système est facile à utiliser					
4. J'aurais besoin du soutien d'un technicien pour être capable d'utiliser ce système					
5. Les différentes fonctionnalités de ce système sont bien intégrées					
6. Il y a trop d'incohérences dans ce système					
7. La plupart des gens apprendront à utiliser ce système très rapidement					
8. Ce système est très lourd à utiliser					
9. Je me suis sentie très en confiance en utilisant ce système					
10. J'ai eu besoin d'apprendre beaucoup de choses avant de pouvoir utiliser ce système					

Bibliography

- [Abdalkareem et al., 2021] Abdalkareem, Z. A., Amir, A., Al-Betar, M. A., Ekhan, P., and Hammouri, A. I. (2021). Healthcare scheduling in optimization context: a review. *Health and Technology*, 11(3):445–469.
- [Agnētis et al., 2014] Agnētis, A., Billaut, J.-C., Gawiejnowicz, S., Pacciarelli, D., and Soukhal, A. (2014). *Multiagent Scheduling*. Springer Berlin Heidelberg, Berlin, Heidelberg.
- [Aiken, 2002] Aiken, L. H. (2002). Hospital Nurse Staffing and Patient Mortality, Nurse Burnout, and Job Dissatisfaction. *JAMA*, 288(16):1987.
- [Al Mahmud and Soysa, 2020] Al Mahmud, A. and Soysa, A. I. (2020). POMA: A tangible user interface to improve social and cognitive skills of Sri Lankan children with ASD. *International Journal of Human-Computer Studies*, 144:102486.
- [Alibali et al., 1997] Alibali, M. W., Flevares, L. M., and Goldin-Meadow, S. (1997). Assessing Knowledge Conveyed in Gesture: Do Teachers Have the Upper Hand?
- [Almjally et al., 2020] Almjally, A., Howland, K., and Good, J. (2020). Comparing TUIs and GUIs for Primary School Programming. In *Proceedings of the 51st ACM Technical Symposium on Computer Science Education*, pages 521–527, Portland OR USA. ACM.
- [Alvarez, 2007] Alvarez, J. (2007). DU JEU VIDÉO AU SERIOUS GAME Approches culturelle, pragmatique et formelle. Publisher: Unpublished.
- [Alvarez and Djaouti, 2011] Alvarez, J. and Djaouti, D. (2011). An introduction to Serious game Definitions and concepts. *Proceedings of the Serious Games & Simulation Workshop*, pages 10–15.
- [Anderson and Shattuck, 2012] Anderson, T. and Shattuck, J. (2012). Design-Based Research: A Decade of Progress in Education Research? *Educational Researcher*, 41(1):16–25.
- [Anne-Laure Guinet, 2022] Anne-Laure Guinet (2022). *Multimodal sensory feedback in augmented reality for gait rehabilitation in children with cerebral palsy*. PhD thesis, Université Paris-Saclay, Paris.
- [Antle et al., 2009] Antle, A. N., Droumeva, M., and Ha, D. (2009). Hands on what?: comparing children’s mouse-based and tangible-based interaction. In *Proceedings of the 8th International Conference on Interaction Design and Children*, pages 80–88, Como Italy. ACM.

- [Arias et al., 1997] Arias, E., Eden, H., and Fisher, G. (1997). Enhancing communication, facilitating shared understanding, and creating better artifacts by integrating physical and computational media for design. In *Proceedings of the conference on Designing interactive systems processes, practices, methods, and techniques - DIS '97*, pages 1–12, Amsterdam, The Netherlands. ACM Press.
- [Aristotle, 1999] Aristotle (1999). Politics. In *Politics*.
- [Aruldoss et al., 2013] Aruldoss, M., Lakshmi, T. M., and Venkatesan, V. P. (2013). A Survey on Multi Criteria Decision Making Methods and Its Applications. *American Journal of Information Systems*, 1(1):31–43.
- [Bavelas et al., 1992] Bavelas, J. B., Chovil, N., Lawrie, D. A., and Wade, A. (1992). Interactive gestures. *Discourse Processes*, 15(4):469–489.
- [Blazewicz et al., 2019] Blazewicz, J., Ecker, K. H., Pesch, E., Schmidt, G., Sterna, M., and Weglarz, J. (2019). *Handbook on Scheduling: From Theory to Practice*. Springer International Publishing, Cham.
- [Bohn, 2004] Bohn, J. (2004). The Smart Jigsaw Puzzle Assistant: Using RFID Technology for Building Augmented Real-World Games. page 6.
- [Bradski and Kaehler, 2011] Bradski, G. R. and Kaehler, A. (2011). *Learning OpenCV: computer vision with the OpenCV library*. Software that sees. O'Reilly, Beijing, 1. ed., [nachdr.] edition.
- [Brockmyer et al., 2009] Brockmyer, J. H., Fox, C. M., Curtiss, K. A., McBroom, E., Burkhart, K. M., and Pidruzny, J. N. (2009). The development of the Game Engagement Questionnaire: A measure of engagement in video game-playing. *Journal of Experimental Social Psychology*, 45(4):624–634.
- [Brooke, 1995] Brooke, J. (1995). SUS - A quick and dirty usability scale. *Usability Evaluation In Industry*, page 8.
- [Brown, 1992] Brown, A. L. (1992). Design Experiments: Theoretical and Methodological Challenges in Creating Complex Interventions in Classroom Settings. *The Journal of the Learning Sciences*, 2(2):141–178.
- [Cegarra and Morgado, 2009] Cegarra, J. and Morgado, N. (2009). Étude des propriétés de la version francophone du NASA-TLX. *EPIQUE 2009: 5ème Colloque de Psychologie Ergonomique*, page 6.
- [Chandrasekera and Yoon, 2015] Chandrasekera, T. and Yoon, S.-Y. (2015). The Effect of Tangible User Interfaces on Cognitive Load in the Creative Design Process. In *2015 IEEE International Symposium on Mixed and Augmented Reality - Media, Art, Social Science, Humanities and Design*, pages 6–8, Fukuoka, Japan. IEEE.
- [Cheang et al., 2003] Cheang, B., Li, H., Lim, A., and Rodrigues, B. (2003). Nurse rostering problems—a bibliographic survey. *European Journal of Operational Research*, 151(3):447–460.

- [Chen et al., 2011] Chen, M., Kolko, B. E., Cuddihy, E., and Medina, E. (2011). Modeling but NOT measuring engagement in computer games. *Conference: Games Learning Society 2011*, page 9.
- [Cheng et al., 2010] Cheng, L. K., Der, C. S., Omar, R., and Sidhu, M. S. (2010). Learnability Measures in GUI Versus TUI.
- [Clark C. Abt, 1970] Clark C. Abt (1970). *Serious games*. Viking Press, New York.
- [Cordova and Lepper, 1996] Cordova, D. I. and Lepper, M. R. (1996). Intrinsic motivation and the process of learning: Beneficial effects of contextualization, personalization, and choice. *Journal of Educational Psychology*, 88(4):715–730.
- [Cziksentmihalyi, 1990] Cziksentmihalyi, M. (1990). *Flow – The Psychology of optimal experience*.
- [Deborah Dos Santos, 2021] Deborah Dos Santos (2021). RFID technology, how does it work?
- [Djaouti et al., 2011] Djaouti, D., Alvarez, J., and Jessel, J.-P. (2011). Classifying Serious Games: the G/P/S model. *Handbook of Research on Improving Learning and Motivation through Educational Games: Multidisciplinary Approaches*.
- [DREES, 2021] DREES (2021). Démographie des professionnels de santé - DREES.
- [Dumas, 2010] Dumas, B. (2010). *Frameworks, description languages and fusion engines for multimodal interactive systems*. PhD thesis, Université de Fribourg, Fribourg.
- [Ehn, 1993] Ehn, P. (1993). Scandinavian Design: On Participation and Skill. In *Usability: Turning Technologies into Tools*. Oxford University Press.
- [Fallman, 2003] Fallman, D. (2003). Design-oriented Human—Computer Interaction. *NEW HORIZONS*, (5).
- [Fenouillet et al., 2009] Fenouillet, F., Kaplan, J., and Yennek, N. (2009). Serious games et motivation. IMT Atlantique, Brest.
- [First, 2021] First, E. (2021). *EF EPI EF English Proficiency Index A Ranking of 112 Countries and Regions by English Skills*. Education First.
- [Fischer, 2001] Fischer, G. (2001). Articulating the Task at Hand and Making Information Relevant to It. *Human–Computer Interaction*, 16(2-4):243–256.
- [Fishkin, 2004] Fishkin, K. (2004). A taxonomy for and analysis of tangible interfaces. *Personal and Ubiquitous Computing*, 8(5).
- [Fishkin et al., 2004] Fishkin, K. P., Jiang, B., Philipose, M., and Roy, S. (2004). I Sense a Disturbance in the Force: Unobtrusive Detection of Interactions with RFID-tagged Objects. In Hutchison, D., Kanade, T., Kittler, J., Kleinberg, J. M., Mattern, F., Mitchell, J. C., Naor, M., Nierstrasz, O., Pandu Rangan, C., Steffen, B., Sudan, M., Terzopoulos, D., Tygar, D., Vardi, M. Y., Weikum, G., Davies, N., Mynatt, E. D., and Siio, I., editors, *UbiComp 2004: Ubiquitous Computing*, volume 3205, pages 268–282. Springer Berlin Heidelberg, Berlin, Heidelberg.

- [Fitzmaurice and Buxton, 1997] Fitzmaurice, G. W. and Buxton, W. (1997). An empirical evaluation of graspable user interfaces: towards specialized, space-multiplexed input. In *Proceedings of the ACM SIGCHI Conference on Human factors in computing systems*, volume CHI '97: Proceedings of the ACM SIGCHI Conference on Human factors in computing systems, pages 43–50, Atlanta Georgia USA. ACM.
- [Fitzmaurice et al., 1995] Fitzmaurice, G. W., Ishii, H., and Buxton, W. A. S. (1995). Bricks: laying the foundations for graspable user interfaces. In *Proceedings of the SIGCHI conference on Human factors in computing systems - CHI '95*, pages 442–449, Denver, Colorado, United States. ACM Press.
- [Fjeld et al., 2001] Fjeld, M., Ironmonger, N., Guttormsen Schär, S., and Krueger, H. (2001). Design and Evaluation of Four AR Navigation Tools Using Scene and Viewpoint Handling. Medium: application/pdf, Online-Ressource Publisher: ETH Zurich.
- [Fjeld et al., 1999] Fjeld, M., Voorhorst, F., Bichsel, M., Lauche, K., Rauterberg, M., and Krueger, H. (1999). Exploring Brick-Based Navigation and Composition in an Augmented Reality. In Goos, G., Hartmanis, J., van Leeuwen, J., and Gellersen, H.-W., editors, *Handheld and Ubiquitous Computing*, volume 1707, pages 102–116. Springer Berlin Heidelberg, Berlin, Heidelberg. Series Title: Lecture Notes in Computer Science.
- [Fletcher et al., 2003] Fletcher, G., Flin, R., McGeorge, P., Glavin, R., Maran, N., and Patey, R. (2003). Anaesthetists' Non-Technical Skills (ANTS): evaluation of a behavioural marker system † †Declaration of interest: The ANTS system was developed under research funding from the Scottish Council for Postgraduate Medical and Dental Education, now part of NHS Education for Scotland, through grants to the University of Aberdeen from September 1999 to August 2003. The views presented in this paper are those of the authors and should not be taken to represent the position or policy of the funding body. *British Journal of Anaesthesia*, 90(5):580–588.
- [Flin et al., 2010] Flin, R., Patey, R., Glavin, R., and Maran, N. (2010). Anaesthetists' non-technical skills. *British Journal of Anaesthesia*, 105(1):38–44.
- [Floerkemeier and Mattern, 2006] Floerkemeier, C. and Mattern, F. (2006). Smart Playing Cards – Enhancing the Gaming Experience with RFID. page 10.
- [France compétences, 2022] France compétences (2022). RNCP8940 - DE - Diplôme d'Etat d'infirmier ou d'infirmière, grade de licence - France Compétences.
- [Galaup, 2013] Galaup, M. (2013). *De la conception à l'usage d'un jeu sérieux de génie mécanique : phénomènes de transposition didactique dans l'enseignement secondaire et universitaire. Le cas de Mecagenius®*. PhD thesis, Université Toulouse 2 Le Mirail (UT2 Le Mirail), Toulouse.
- [Galaup et al., 2021] Galaup, M., Pingaud, H., Pons-Lelardeux, C., and Lagarrigue, P. (2021). Vers une plénitude du potentiel positif des infirmiers(e)-élèves : une étude de cas avec le jeu sérieux CLONE. Medium: pdf Publisher: STICEF Version Number: 1.

- [Garrido et al., 2011] Garrido, P. C., Miraz, G. M., Ruiz, I. L., and Gomez-Nieto, M. A. (2011). Use of NFC-based Pervasive Games for Encouraging Learning and Student Motivation. In *2011 Third International Workshop on Near Field Communication*, pages 32–37, Hagenberg, Austria. IEEE.
- [Garrido-Jurado et al., 2014] Garrido-Jurado, S., Muñoz-Salinas, R., Madrid-Cuevas, F., and Marín-Jiménez, M. (2014). Automatic generation and detection of highly reliable fiducial markers under occlusion. *Pattern Recognition*, 47(6):2280–2292.
- [Garrido-Jurado et al., 2016] Garrido-Jurado, S., Muñoz-Salinas, R., Madrid-Cuevas, F., and Medina-Carnicer, R. (2016). Generation of fiducial marker dictionaries using Mixed Integer Linear Programming. *Pattern Recognition*, 51:481–491.
- [George W. Fitzmaurice, 1996] George W. Fitzmaurice (1996). *Graspable User Interfaces*. PhD thesis, University of Toronto, Toronto Ontario Canada.
- [Gibson, 1962] Gibson, J. J. (1962). Observations on active touch. *Psychological Review*, 69(6):477–491.
- [Goldin-Meadow et al., 2001] Goldin-Meadow, S., Nusbaum, H., Kelly, S. D., and Wagner, S. (2001). Explaining Math: Gesturing Lightens the Load. *Psychological Science*, 12(6):516–522.
- [Gronier and Baudet, 2021] Gronier, G. and Baudet, A. (2021). Psychometric Evaluation of the F-SUS: Creation and Validation of the French Version of the System Usability Scale. *International Journal of Human-Computer Interaction*, pages 1–12.
- [Habgood and Ainsworth, 2011] Habgood, M. P. J. and Ainsworth, S. E. (2011). Motivating Children to Learn Effectively: Exploring the Value of Intrinsic Integration in Educational Games. *Journal of the Learning Sciences*, 20(2):169–206.
- [Hamari and Keronen, 2017] Hamari, J. and Keronen, L. (2017). Why do people play games? A meta-analysis. *International Journal of Information Management*, 37(3):125–141.
- [Hartmann et al., 2015] Hartmann, T., Wirth, W., Schramm, H., Klimmt, C., Vorderer, P., Gysbers, A., Böcking, S., Ravaja, N., Laarni, J., Saari, T., Gouveia, F., and Sacau, A. (2015). The Spatial Presence Experience Scale (SPES). *Journal of Media Psychology: Theories, Methods, and Applications*, 1:1–15.
- [Hassenzahl, 2003] Hassenzahl, M. (2003). The thing and I: Understanding the relationship between user and product. In *Funology: From usability to enjoyment*, volume 3 of *Human-Computer Interaction Series*, pages 31–42. Kluwer Academic, blyth, m. a., monk, a. f., overbeeke, k. & wright, p. c. edition.
- [Hinske and Langheinrich, 2007] Hinske, S. and Langheinrich, M. (2007). An RFID-based Infrastructure for Automatically Determining the Position and Orientation of Game Objects in Tabletop Games. page 21.

- [Hinske and Langheinrich, 2008] Hinske, S. and Langheinrich, M. (2008). Using a Movable RFID Antenna to Automatically Determine the Position and Orientation of Objects on a Tabletop. In Roggen, D., Lombriser, C., Tröster, G., Kortuem, G., and Havinga, P., editors, *Smart Sensing and Context*, volume 5279, pages 14–26. Springer Berlin Heidelberg, Berlin, Heidelberg.
- [Holler et al., 2009] Holler, J., Shovelton, H., and Beattie, G. (2009). Do Iconic Hand Gestures Really Contribute to the Communication of Semantic Information in a Face-to-Face Context? *Journal of Nonverbal Behavior*, 33(2):73–88.
- [Horn et al., 2009] Horn, M. S., Solovey, E. T., Crouser, R. J., and Jacob, R. J. (2009). Comparing the use of tangible and graphical programming languages for informal science education. In *Proceedings of the 27th international conference on Human factors in computing systems - CHI 09*, page 975, Boston, MA, USA. ACM Press.
- [Hornecker and Buur, 2006] Hornecker, E. and Buur, J. (2006). Getting a Grip on Tangible Interaction: A Framework on Physical Space and Social Interaction. *CHI '06: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, page 10.
- [Hsieh et al., 2018] Hsieh, M.-J., Liang, R.-H., Huang, D.-Y., Ke, J.-Y., and Chen, B.-Y. (2018). RFIBricks: Interactive Building Blocks Based on RFID. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems - CHI '18*, pages 1–10, Montreal QC, Canada. ACM Press.
- [Huber et al., 2021] Huber, E., Kleinknecht-Dolf, M., Kugler, C., and Spirig, R. (2021). Patient-related complexity of nursing care in acute care hospitals – an updated concept. *Scandinavian Journal of Caring Sciences*, 35(1):178–195.
- [Ishii and Ullmer, 1997] Ishii, H. and Ullmer, B. (1997). Tangible bits: towards seamless interfaces between people, bits and atoms. In *Proceedings of the SIGCHI conference on Human factors in computing systems - CHI '97*, pages 234–241, Atlanta, Georgia, United States. ACM Press.
- [Iverson and Goldin-Meadow, 1997] Iverson, J. M. and Goldin-Meadow, S. (1997). What’s communication got to do with it? Gesture in children blind from birth. *Developmental Psychology*, 33(3):453–467.
- [James J. Gibson, 1977] James J. Gibson (1977). Theory of Affordances. In *Theory of Affordances*, pages 67–82. Lawrence Erlbaum Associates, Inc. Publishers, Hillsdale, USA.
- [Jean Piaget, 1974] Jean Piaget (1974). *La prise de conscience*. PUF, Paris.
- [Jetter et al., 2011] Jetter, H.-C., Gerken, J., Zöllner, M., Reiterer, H., and Milic-Frayling, N. (2011). Materializing the query with facet-streams: a hybrid surface for collaborative search on tabletops. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 3013–3022, Vancouver BC Canada. ACM.

- [Kim et al., 2019] Kim, J., Maher, M. L., and Lee, L. (2019). The Effect of Tangible Interaction on Spatial Design Tasks. In Gero, J. S., editor, *Design Computing and Cognition '18*, pages 135–153. Springer International Publishing, Cham.
- [Kim and Maher, 2008] Kim, M. J. and Maher, M. L. (2008). The Impact of Tangible User Interfaces on Designers' Spatial Cognition. *Human-Computer Interaction*, 23(2):101–137.
- [Kleffner et al., 2017] Kleffner, R., Flatten, J., Leaver-Fay, A., Baker, D., Siegel, J. B., Khatib, F., and Cooper, S. (2017). Foldit Standalone: a video game-derived protein structure manipulation interface using Rosetta. *Bioinformatics*, 33(17):2765–2767.
- [Lallemand et al., 2015] Lallemand, C., Koenig, V., Gronier, G., and Martin, R. (2015). Création et validation d'une version française du questionnaire AttrakDiff pour l'évaluation de l'expérience utilisateur des systèmes interactifs. *European Review of Applied Psychology*, 65(5):239–252.
- [Landi et al., 2022a] Landi, B., Pingaud, H., Culié, J.-B., and Galaup, M. (2022a). A serious game to upgrade the learning of organization skills in nurses schools. Tokyo, Japan.
- [Landi et al., 2022b] Landi, B., Pingaud, H., and Galaup, M. (2022b). Apprentissage par le numérique de la compétence organisationnelle des étudiants en soins infirmiers. In *L'impact des nouvelles technologies et de l'Intelligence Artificielle sur les organisations de santé*, Saint-Etienne, France.
- [Landi et al., 2022c] Landi, B., Pingaud, H., and Galaup, M. (2022c). Compromis de ludicisation dans un jeu de planification sérieux. Polytech Lille, France.
- [Landi et al., 2023] Landi, B., Pingaud, H., and Galaup, M. (2023). Montée en compétences d'organisation des interventions soignantes à l'hôpital. In *Propulser la performance Interconnectivité et collaboration dans un contexte d'intelligence artificielle.*, Québec, Canada.
- [Landi et al., 2021] Landi, B., Pons Lelardeux, C. P., Lagarrigue, P., and Jessel, J.-P. (2021). Tangibilisation of an educational scheduling game's graphical interface: Tangibilisation de l'interface d'un jeu de formation à la planification. In *32e Conférence Francophone sur l'Interaction Homme-Machine*, pages 1–5, Virtual Event France. ACM.
- [Laurence Nigay, 1994] Laurence Nigay (1994). *Conception et modelisation logicielles des systemes interactifs : application aux interfaces multimodales. (Software design and implementation of interactive systems: A case study of multimodal interfaces)*. PhD thesis, UNIVERSITÉ JOSEPH FOURIER, GRENOBLE.
- [Law et al., 2018] Law, E. L.-C., Brühlmann, F., and Mekler, E. D. (2018). Systematic Review and Validation of the Game Experience Questionnaire (GEQ) - Implications for Citation and Reporting Practice. In *Proceedings of the 2018 Annual Symposium on Computer-Human Interaction in Play*, pages 257–270, Melbourne VIC Australia. ACM.

- [Lee et al., 2009] Lee, S., Choi, J., and Park, J.-i. (2009). Interactive e-learning system using pattern recognition and augmented reality. *IEEE Transactions on Consumer Electronics*, 55(2):883–890.
- [Lelardeux et al., 2020] Lelardeux, C. P., Pingaud, H., Galaup, M., Ramolet, A., and Lagarrigue, P. (2020). The Challenge of Designing Interactive Scenarios to Train Nurses on Rostering Problems in a Virtual Clinical Unit. In Auer, M. E. and Tsiatsos, T., editors, *The Challenges of the Digital Transformation in Education*, volume 916, pages 589–601. Springer International Publishing, Cham.
- [LimeSurvey GmbH, 2023] LimeSurvey GmbH (2023). LimeSurvey — Free Online Survey Tool.
- [Lind, 1772] Lind, J. (1772). *A Treatise on the Scurvy. In Three Parts. Containing An Inquiry into the Nature, Causes, and Cure, of that Disease. Together with A Critical and Chronological View of what has been published on the Subject.* Printed for S. Crowder, D. Wilson and G. Nicholls, T. Cadell, T. Becket and Co., G. Pearch, and W. Woodfall. Printed for S. Crowder, D. Wilson and G. Nicholls, T. Cadell, T. Becket and Co., G. Pearch, and W. Woodfall., Edimbourg, third edition.
- [Ma et al., 2003] Ma, Y.-P., Lee, C.-H., and Jeng, T. (2003). Inavigator: A Spatially-Aware Tangible Interface for Interactive 3D Visualization. pages 963–974, Bangkok. Thailand.
- [Maher and Kim, 2005] Maher, M. L. and Kim, M. J. (2005). Do Tangible User Interfaces Impact Spatial Cognition in Collaborative Design? In Hutchison, D., Kanade, T., Kittler, J., Kleinberg, J. M., Mattern, F., Mitchell, J. C., Naor, M., Nierstrasz, O., Pandu Rangan, C., Steffen, B., Sudan, M., Terzopoulos, D., Tygar, D., Vardi, M. Y., Weikum, G., and Luo, Y., editors, *Cooperative Design, Visualization, and Engineering*, volume 3675, pages 30–41. Springer Berlin Heidelberg, Berlin, Heidelberg. Series Title: Lecture Notes in Computer Science.
- [Malone et al., 1981] Malone, T. W., Calfee, R., Clancey, B., Collins, A., Friend, J., Furnas, G., and Gould, L. (1981). *Toward a Theory of Intrinsically Instruction Motivating*, volume Cognitive Science. Wadsworth Publishing Co.
- [Marshall, 2012] Marshall, J. (2012). Victory for crowdsourced biomolecule design. *Nature*, page nature.2012.9872.
- [Marshall, 2007] Marshall, P. (2007). Do tangible interfaces enhance learning? In *Proceedings of the 1st international conference on Tangible and embedded interaction - TEI '07*, page 163, Baton Rouge, Louisiana. ACM Press.
- [Marshall et al., 2009] Marshall, P., Fleck, R., Harris, A., Rick, J., Hornecker, E., Rogers, Y., Yuill, N., and Dalton, N. S. (2009). Fighting for control: children’s embodied interactions when using physical and digital representations. In *Proceedings of the 27th international conference on Human factors in computing systems - CHI 09*, page 2149, Boston, MA, USA. ACM Press.

- [McHugh et al., 2011] McHugh, M. D., Kutney-Lee, A., Cimiotti, J. P., Sloane, D. M., and Aiken, L. H. (2011). Nurses’ Widespread Job Dissatisfaction, Burnout, And Frustration With Health Benefits Signal Problems For Patient Care. *Health Affairs*, 30(2):202–210.
- [Mellor, 1966] Mellor, P. (1966). A Review of Job Shop Scheduling. *Journal of the Operational Research Society*, 17(2):161–171.
- [Milgram and Kishino, 1994] Milgram, P. and Kishino, F. (1994). A taxonomy of mixed reality visual displays. *IEICE Transactions on Information Systems*, E77-D(12):15.
- [Montessori, 1912] Montessori, M. (1912). *The Montessori method*. Routledge, Taylor & Francis Group, London. OCLC: 1003855476.
- [Norman, 2008] Norman, D. A. (2008). *The design of everyday things*. Basic Books, New York, first basic paperback, [nachdr.] edition.
- [Novoseltseva et al., 2022] Novoseltseva, D., Pons, C., and Jessel, N. (2022). Examining Students’ Behavior in a Digital Simulation Game for Nurse Training. *International Journal of Serious Games*, 9(4):3–24.
- [Obrenovic, 2011] Obrenovic, Z. (2011). Design-based research: what we learn when we engage in design of interactive systems. *Interactions*, 18(5):56–59.
- [O’Brien and Toms, 2010] O’Brien, H. L. and Toms, E. G. (2010). The development and evaluation of a survey to measure user engagement. *Journal of the American Society for Information Science and Technology*, 61(1):50–69.
- [O’Malley and Fraser, 2004] O’Malley, C. and Fraser, D. (2004). Literature Review in Learning with Tangible Technologies. *NESTA Futurelab Rep.*, 12.
- [OpenCV, 2023] OpenCV (2023). OpenCV About.
- [O’Brien et al., 2018] O’Brien, H. L., Cairns, P., and Hall, M. (2018). A practical approach to measuring user engagement with the refined user engagement scale (UES) and new UES short form. *International Journal of Human-Computer Studies*, 112:28–39.
- [Palma-Ruiz et al., 2022] Palma-Ruiz, J. M., Torres-Toukoumidis, A., González-Moreno, S. E., and Valles-Baca, H. G. (2022). An overview of the gaming industry across nations: using analytics with power BI to forecast and identify key influencers. *Heliyon*, 8(2):e08959.
- [Pangaro and Sociology, 2003] Pangaro, G. A. and Sociology, B. A. (2003). *The Actuated Workbench: 2D Actuation in Tabletop Tangible Interfaces*. PhD thesis, Massachusetts Institute of Technology.
- [Patten et al., 2001] Patten, J., Ishii, H., Hines, J., and Pangaro, G. (2001). Sensetable: A Wireless Object Tracking Platform for Tangible User Interfaces. In *Proceedings of the {CHI} 2001 Conference on Human Factors in Computing Systems, Seattle, WA, USA, March 31 - April 5, 2001*, pages 253–260, Seattle, WA, USA. ACM.

- [Piaget, 1997] Piaget, J. (1997). THE ORIGINS OF INTELLIGENCE IN CHILDREN. 3.
- [Pinedo, 2016] Pinedo, M. L. (2016). *Scheduling*. Springer International Publishing, Cham.
- [Ping and Goldin-Meadow, 2010] Ping, R. and Goldin-Meadow, S. (2010). Gesturing Saves Cognitive Resources When Talking About Nonpresent Objects. *Cognitive Science*, 34(4):602–619.
- [Ping and Goldin-Meadow, 2008] Ping, R. M. and Goldin-Meadow, S. (2008). Hands in the air: Using ungrounded iconic gestures to teach children conservation of quantity. *Developmental Psychology*, 44(5):1277–1287.
- [Pons Lelardeux et al., 2020] Pons Lelardeux, C., Galaup, M., Pingaud, H., Mercadier, C., and Lagarrigue, P. (2020). A Method to Balance Educational Game Content and Lesson Duration: The Case of a Digital Simulation Game for Nurse Training. In Auer, M. E., Hortsch, H., and Sethakul, P., editors, *The Impact of the 4th Industrial Revolution on Engineering Education*, volume 1134, pages 125–136. Springer International Publishing, Cham.
- [Prabaswari et al., 2019] Prabaswari, A. D., Basumerda, C., and Utomo, B. W. (2019). The Mental Workload Analysis of Staff in Study Program of Private Educational Organization. *IOP Conference Series: Materials Science and Engineering*, 528(1):012018.
- [Rauterberg et al., 1998] Rauterberg, M., Fjeld, M., Krueger, H., Bichsel, M., Leonhardt, U., and Meier, M. (1998). BUILD-IT: a planning tool for construction and design. In *CHI 98 Conference Summary on Human Factors in Computing Systems*, pages 177–178, Los Angeles California USA. ACM.
- [Rawitsch, 2017] Rawitsch, D. (2017). Classic Game Postmortem: Oregon Trail.
- [Rey, 2020] Rey, S. (2020). *Apports des Interactions Tangibles pour la Création, le Choix et le Suivi de Parcours de Visite Personnalisés dans les Musées*. PhD thesis, Université de Bordeaux.
- [Robert T. Hays, 2005] Robert T. Hays (2005). The effectiveness of instructional games: A literature review and discussion. Technical report, NAVAL AIR WARFARE CENTER TRAINING SYSTEMS DIVISION ORLANDO, NAVAL AIR WARFARE CENTER TRAINING SYSTEMS DIVISION ORLANDO, FL 32826-3275.
- [Romero-Ramirez et al., 2018] Romero-Ramirez, F. J., Muñoz-Salinas, R., and Medina-Carnicer, R. (2018). Speeded up detection of squared fiducial markers. *Image and Vision Computing*, 76:38–47.
- [Safrodin et al., 2019] Safrodin, M., Cahya Bagar, F. N., and Yoga Pralista, F. (2019). The Development of Digital Board Game to Introduce Indonesian Wildlife using AR Technology and NFC. In *2019 International Electronics Symposium (IES)*, pages 277–282, Surabaya, Indonesia. IEEE.

- [Salen and Zimmerman, 2004] Salen, K. and Zimmerman, E. (2004). *Rules of Play - Game Design Fundamentals*.
- [Schneider et al., 2011] Schneider, B., Jermann, P., Zufferey, G., and Dillenbourg, P. (2011). Benefits of a Tangible Interface for Collaborative Learning and Interaction. *IEEE Transactions on Learning Technologies*, 4(3):222–232.
- [Sei et al., 2018] Sei, N., Oka, M., and Mori, H. (2018). Investigation of Learning Process with TUI and GUI Based on COCOM. In Yamamoto, S. and Mori, H., editors, *Human Interface and the Management of Information. Information in Applications and Services*, volume 10905, pages 331–339. Springer International Publishing, Cham. Series Title: Lecture Notes in Computer Science.
- [Sherry, 2004] Sherry, J. L. (2004). Flow and Media Enjoyment. *Communication Theory*, 14(4):328–347.
- [Sitdhisanguan et al., 2007] Sitdhisanguan, K., Dechaboon, A., Chotikakamthorn, N., and Out, P. (2007). Comparative study of WIMP and tangible user interfaces in training shape matching skill for autistic children. In *TENCON 2007 - 2007 IEEE Region 10 Conference*, pages 1–4, Taipei, Taiwan. IEEE.
- [Sudnow and Dreyfus, 2001] Sudnow, D. and Dreyfus, H. L. (2001). *Ways of the hand: a rewritten account*. MIT Press, Cambridge, Mass.
- [Suzuki and Keiichi, 1985] Suzuki, S. and Keiichi, A. (1985). Topological Structural Analysis of Digitized Binary Images by Border Following. *Comput. Vis. Graph. Image Process.*, 30:15.
- [Sweetser and Wyeth, 2005] Sweetser, P. and Wyeth, P. (2005). GameFlow: a model for evaluating player enjoyment in games. *Computers in Entertainment*, 3(3):3–3.
- [Sylla et al., 2012] Sylla, C., Branco, P., Coutinho, C., and Coquet, E. (2012). TUIs vs. GUIs: comparing the learning potential with preschoolers. *Personal and Ubiquitous Computing*, 16(4):421–432.
- [Tan and Rau, 2015] Tan, J. and Rau, P.-L. P. (2015). A Design of Augmented Tabletop Game Based on RFID Technology. *Procedia Manufacturing*, 3:2142–2148.
- [Tan and Rau, 2010] Tan, J. and Rau, P. P. (2010). Building Your Own Tangible Virtual World: A Design of an RFID-Based Tabletop Game Platform. In *2010 Third IEEE International Conference on Digital Game and Intelligent Toy Enhanced Learning*, pages 151–155, Kaohsiung, Taiwan. IEEE.
- [Thevin et al., 2021] Thevin, L., Rodier, N., Oriola, B., Hachet, M., Jouffrais, C., and Brock, A. M. (2021). Inclusive Adaptation of Existing Board Games for Gamers with and without Visual Impairments using a Spatial Augmented Reality Framework for Touch Detection and Audio Feedback. *Proceedings of the ACM on Human-Computer Interaction*, 5(ISS):1–33.
- [Ullmer and Ishii, 2000] Ullmer, B. and Ishii, H. (2000). Emerging frameworks for tangible user interfaces. *IBM Systems Journal*, 39(3.4):915–931.

- [Ullmer et al., 2005] Ullmer, B., Ishii, H., and Jacob, R. J. K. (2005). Token+constraint systems for tangible interaction with digital information. *ACM Transactions on Computer-Human Interaction (TOCHI)*, 12(1):81–118.
- [Ullmer, 2002] Ullmer, B. A. (2002). *Tangible Interfaces for Manipulating Aggregates of Digital Information*. PhD thesis, Massachusetts Institute of Technology.
- [van Dam, 1997] van Dam, A. (1997). Post-WIMP user interfaces. *Communications of the ACM*, 40(2):63–67.
- [Weinstein, 2005] Weinstein, R. (2005). RFID: a technical overview and its application to the enterprise. *IT Professional*, 7(3):27–33.
- [Weis, 2007] Weis, S. A. (2007). RFID (Radio Frequency Identification): Principles and Applications.
- [Weiser, 1993] Weiser, M. (1993). Some computer science issues in ubiquitous computing. *Communications of the ACM*, 36(7):75–84.
- [Whyte et al., 2015] Whyte, E. M., Smyth, J. M., and Scherf, K. S. (2015). Designing Serious Game Interventions for Individuals with Autism. *Journal of Autism and Developmental Disorders*, 45(12):3820–3831.
- [Wijnand A. IJsselsteijn et al., 2013] Wijnand A. IJsselsteijn, Yvonne A.W. de Kort, and K. Poels (2013). Game Experience Questionnaire.
- [Wren and Reynolds, 2004] Wren, C. R. and Reynolds, C. J. (2004). Parsimony & Transparency in Ubiquitous Interface Design.
- [Xie et al., 2008] Xie, Z., Antle, A., and Motamedi, N. (2008). Are tangibles more fun?: Comparing children’s enjoyment and engagement using physical, graphical and tangible user interfaces. In *Proceedings of the 2nd International Conference on Tangible and Embedded Interaction 2008, Bonn, Germany, February 18-20, 2008*, pages 191–198, Bonn, Germany. ACM.
- [Young, 2014] Young, S. W. H. (2014). Improving Library User Experience with A/B Testing: Principles and Process. *Weave: Journal of Library User Experience*, 1(1).
- [Zooniverse, 2023a] Zooniverse (2023a). Zooniverse.
- [Zooniverse, 2023b] Zooniverse (2023b). Zooniverse - Publications.