

UNIVERSITÉ DE MONTRÉAL

ÉTUDE COMPARATIVE DE CINQ DISPOSITIFS D'INTERACTION POUR LES GRANDS  
ÉCRANS TACTILES VERTICAUX : TACTILE DIRECT, PAVÉ TACTILE, TABLETTE,  
SOURIS, BOULES DE COMMANDE

ETIENNE VIALE

DÉPARTEMENT DE MATHÉMATIQUES ET DE GÉNIE INDUSTRIEL

ÉCOLE POLYTECHNIQUE DE MONTRÉAL

MÉMOIRE PRÉSENTÉ EN VUE DE L'OBTENTION  
DU DIPLÔME DE MAÎTRISE ÈS SCIENCES APPLIQUÉES

(GÉNIE INDUSTRIEL)

MARS 2014

UNIVERSITÉ DE MONTRÉAL

ÉCOLE POLYTECHNIQUE DE MONTRÉAL

Ce mémoire intitulé :

ÉTUDE COMPARATIVE DE CINQ DISPOSITIFS D'INTERACTION POUR LES GRANDS  
ÉCRANS TACTILES VERTICAUX : TACTILE DIRECT, PAVÉ TACTILE, TABLETTE,  
SOURIS, BOULES DE COMMANDE

présenté par : VIALE Etienne

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

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

M. FRAYRET Jean-Marc, Ph.D., président

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

M. CYBIS Walter, Doctorat, membre

# DÉDICACE

*À mes parents*

## REMERCIEMENTS

Je souhaite tout d'abord remercier L-3 MAPPS pour les ressources accordées, et en particulier M. Yvan Lamontagne, Mme Isabelle Saint-Louis, et M. Stéphane Léger, ainsi que toute la division Marine Software Engineering, pour leur aide et leur disponibilité.

Je remercie également M. Jean-Marc Robert, professeur titulaire à l'École Polytechnique de Montréal, qui a guidé mes efforts pendant ce travail de longue haleine.

J'exprime toute ma gratitude envers M. Philippe Doyon-Poulin, étudiant au doctorat en Interactions humain-machine à l'École Polytechnique de Montréal, qui a contribué à combler certaines de mes lacunes techniques, permettant la complétion du code de test. M. Doyon-Poulin a également proposé la lecture de plusieurs articles scientifiques qui ont orienté ce travail.

Je remercie aussi M. Maxim Morin, doctorant à l'Université de Montréal, chargé de cours en analyse statistique, pour son expertise concernant le traitement des données issues de l'étude.

Enfin, je souhaite remercier ma famille et mes amis, sans le support desquels rien n'aurait pu être accompli.

Cette recherche a été financée par une Bourse en Milieu Pratique BMP Innovation : bourse conjointe du Conseil de Recherches en Sciences Naturelles et en Génie du Canada (CRSNG), du Fonds Québécois de la Recherche sur la Nature et les Technologies (FQRNT) et de L-3 MAPPS.

## RÉSUMÉ

Cette recherche a été réalisée au sein de L-3 MAPPS, une entreprise montréalaise qui conçoit, développe et commercialise des systèmes de simulation et d'entraînement pour les domaines maritime, spatial et énergétique. Certains de ces systèmes font usage d'écrans tactiles de grandes dimensions devant lesquels les opérateurs doivent se déplacer pour les utiliser. L'objectif de cette recherche était d'évaluer la performance et la satisfaction relatives à l'utilisation de six dispositifs d'interaction à l'écran. À noter : il y a cinq types de dispositifs, mais l'un d'eux comprenant deux modèles, six dispositifs ont finalement été évalués. C'est pourquoi dans le reste du document, il est question de six dispositifs.

L'hypothèse formulée est que l'utilisation d'un dispositif d'interaction déporté (c'est-à-dire placé à distance du Grand Écran Vertical Tactile (GETV)) comparativement à une commande tactile directe améliore significativement la performance à la tâche de pilotage d'un curseur sur un écran tactile de grandes dimensions disposé verticalement.

Pour tester cette hypothèse, on a comparé la performance et la satisfaction humaine relatives à l'utilisation de six dispositifs de pointage suivant la méthodologie proposée par ISO 9241-9 (International Standards Organization [ISO], 2000). La partie 9 de cette norme (Exigences relatives aux dispositifs d'entrée autres que les claviers) définit un essai d'utilisabilité de façon à permettre d'évaluer la précision atteinte pour le test de frappe par unité de temps (débit), le taux de cibles atteintes par rapport au total de cibles (taux d'erreur), et la satisfaction liée à l'utilisation d'un dispositif de pointage.

Les six dispositifs retenus pour l'étude sont : un GETV de diagonale 46 pouces, une tablette tactile de diagonale 12 pouces, un pavé tactile (plus communément appelé trackpad), une souris, et deux modèles de boule de commande : un modèle militarisé, et un modèle commercial. Chaque dispositif a été testé suivant six conditions expérimentales différentes : deux types de tâches (multidirectionnelle et unidirectionnelle) et trois niveaux de difficultés (par variation de la taille des cibles).

La tâche multidirectionnelle consistait à frapper 24 cibles apparaissant successivement sur un cercle. La tâche unidirectionnelle consistait à frapper 24 cibles apparaissant successivement de part et d'autre de l'écran.

Vingt sujets travaillant dans l'entreprise d'accueil ont participé à cette étude. Chaque sujet a effectué le test de frappe pour les six conditions expérimentales avec les six dispositifs, pour un total de trente-six "séries" de 24 frappes par sujet. Pour chaque sujet, on a recueilli le débit et le taux d'erreur pour chaque série, ainsi qu'une note de satisfaction globale pour chaque dispositif.

Les résultats montrent que tous les sujets ont été capables d'effectuer la tâche expérimentale avec chacun des dispositifs d'interaction. Ils montrent aussi que comparativement au reste de la sélection, le GETV se classe troisième sur six en termes de débit (2,31 bits / s contre 4,66 pour la tablette tactile, le dispositif le plus rapide, et 3,60 bits / s pour la souris, second dispositif le plus rapide), cinquième en termes de taux d'erreurs (17,62 % contre 8,73 pour la souris, dispositif le plus fiable), et sixième en termes de satisfaction moyenne. Il est donc légitime et même souhaitable de chercher à améliorer les performances des GETV. D'autre part, parmi les autres dispositifs tactiles de la sélection, on constate un avantage en vitesse pour la tablette (fort débit avec 4,66 bits / s contre 2,31 pour le GETV, second dispositif tactile en vitesse), et pour le pavé tactile en fiabilité (faible taux d'erreur avec 10,51 % contre 17,62 pour le GETV, second dispositif tactile en fiabilité). Enfin, la souris reste le dispositif le plus performant, que ce soit en termes de débit ou de taux d'erreurs.

## ABSTRACT

This study is based off a research carried out at L-3 MAPPs, Montréal. L-3 MAPPs is a company that designs, develop and sells simulation and training systems for the the naval, spatial and energy industries. Some of those systems use large touch screens requiring the user to move during use. The initial goal of this work was to evaluate alternative interaction methods for the pointing devices currently employed by the company.

The starting hypothesis states that the use of a remote pointing device (as in placed at some distance from the Large Vertical Touch Screen (LVTS)), when compared with a touch command significantly improves the performances for a pointing task on a large vertical screen.

To test this hypothesis, a comparative experimental study following the methodology prescribed by ISO 9241-9 standard (International Standards Organization [ISO], 2000) was conducted. Section 9 of this standard (Requirements for non-keyboard input devices) defines a usability test allowing the evaluation of a pointing device in terms of speed, error rate, and user satisfaction (measured as throughput, error rate, and satisfaction rate, respectively).

The six pointing devices selected for this study are : a LVTS, a touch screen tablet, a touchpad, a mouse, and two trackball models. The six experimental conditions were defined by two task types (multi and unidirectional) and three difficulty levels (defined by target size). For the multidirectional task the subjects hit 24 successive targets distributed on a circle. For the unidirectional task the subjects hit 24 successive targets on the left and right of the screen. Twenty subjects participated in this study. All subjects tested each device in each of the six conditions, resulting in 36 series for each subject.

The results of this study show that, comparatively to the other devices, the LVTS comes 3rd in throughput (with 2,31 bits / s against 4,66 for the tablet, fastest device), 5th in error rate (with 17,62 % against 8,73 for the mouse, most reliable device), and 6th in satisfaction rate. We conclude that seeking better interaction methods for the LVTS is not only legitimate, but desirable. Furthermore, amongst the other touch pointing devices evaluated, we notice a significant speed advantage for the tablet (high throughput with 4,66 bits / s against 2,31 for the LVTS, second fastest touch device), and a significant reliability advantage for the touchpad (low error rate with 10,51 % against 17,62 for the LVTS, second most reliable touch device). Finally, the mouse remains the overall best performing device, whether in throughput or error rate.

# TABLE DES MATIÈRES

DÉDICACE.....	iii
REMERCIEMENTS.....	iv
RÉSUMÉ.....	v
ABSTRACT.....	vii
TABLE DES MATIÈRES.....	viii
LISTE DES TABLEAUX.....	x
LISTE DES FIGURES.....	xi
LISTE DE LA TERMINOLOGIE.....	xiii
LISTE DES ANNEXES.....	xiv
INTRODUCTION.....	1
CHAPITRE 1 : REVUE DE LITTÉRATURE SUR LA PERFORMANCE ET LA SATISFACTION HUMAINE LIÉES À L'UTILISATION DE DISPOSITIFS D'INTERACTION À L'ÉCRAN.....	3
1.1 Introduction.....	3
1.2 Impact.....	5
1.3 Historique.....	6
1.4 Études comparables.....	9
1.5 Prédiction des performances.....	14
CHAPITRE 2 : PROBLÉMATIQUE, OBJECTIFS, HYPOTHÈSE.....	17
2.1 Problématique.....	17
2.2 Objectifs de recherche.....	18
2.3 Hypothèse.....	18
2.4 Définition du projet de recherche.....	19
2.5 Propositions.....	21
2.6 Utilisation des résultats de la recherche.....	23
CHAPITRE 3 : MÉTHODOLOGIE DE L'ÉTUDE COMPARATIVE DES DISPOSITIFS D'INTERACTION AVEC UN ÉCRAN.....	24
3.1 Configuration.....	25
3.2 Sujets.....	26
3.3 Procédure expérimentale.....	31
3.4 Procédure.....	39
3.5 Déroulement des pré-tests.....	42



CHAPITRE 4 : PRÉSENTATION DES RÉSULTATS.....	43
4.1 Débit et taux d'erreurs pour l'ensemble des dispositifs.....	43
4.2 Données de débit et de taux d'erreur par dispositif.....	45
4.3 Niveau de satisfaction de l'utilisateur.....	50
4.4 Analyses effectuées.....	53
4.5 Discussion.....	70
CONCLUSION.....	74
Limites de l'expérimentation.....	77
Proposition de travaux futurs.....	78
RÉFÉRENCES.....	79
ANNEXES.....	84

## LISTE DES TABLEAUX

Tableau 1.1 : Coefficients de corrélation entre le temps de mouvement et l'indice de difficulté.....	14
Tableau 2.1 : Dimension des GETVs disponibles sur le site de test.....	17
Tableau 2.2 : Types de tests disponibles en fonction du dispositif testé.....	20
Tableau 3.1 : Critères de recrutement de l'échantillon.....	28
Tableau 3.2 : Niveaux d'expérience par dispositif.....	30
Tableau 3.3 : Valeurs de $d$ et $w$ en fonction des indices de difficulté retenus.....	34
Tableau 3.4 : Valeurs de l' $ID_e$ en fonction de l'équation utilisée.....	37
Tableau 3.5 : Composition des carrés latins.....	41
Tableau 4.1 : Résultats de débit et de taux d'erreur par dispositif.....	45
Tableau 4.2 : Performances par dispositif.....	53
Tableau 4.3 : Analyse de variance pour le débit des dispositifs, niveau de difficulté 1.....	56
Tableau 4.4 : Analyse de variance pour le débit des dispositifs, niveau de difficulté 2.....	57
Tableau 4.5 : Analyse de variance pour le débit des dispositifs, niveau de difficulté 3.....	59
Tableau 4.6 : Analyse de variance pour le taux d'erreur des dispositifs.....	62
Tableau 4.7 : Synthèse des performances, par dispositif.....	70
Tableau 4.8 : Comparaison avec les résultats issus de la littérature.....	71
Tableau 4.9 : Comparaison de nos résultats à ceux de l'étude MacKenzie et Jusoh (2001).....	73

## LISTE DES FIGURES

Figure 3.1 : Age des participants.....	28
Figure 3.2 : Taille des participants.....	29
Figure 3.3 : Distance entre l'œil du participant et le GETV (dispositifs D3 à D6).....	32
Figure 3.4 : Placement des cibles pour le test de frappe multidirectionnel.....	33
Figure 3.5 : Placement des cibles pour le test de frappe unidirectionnel.....	34
Figure 3.6 : Répartition des cibles en fonction du niveau de difficulté ; test multidirectionnel.....	34
Figure 3.7 : Répartition des cibles en fonction du niveau de difficulté ; test unidirectionnel.....	35
Figure 4.1 : Répartition des données de débit.....	43
Figure 4.2 : Répartition des données de taux d'erreurs.....	44
Figure 4.3 : Répartition des données de débit par dispositif.....	46
Figure 4.4 : Répartition des valeurs de débit par dispositif.....	47
Figure 4.5 : Répartition des données de taux d'erreurs par dispositif.....	48
Figure 4.6 : Répartition des valeurs de taux d'erreurs par dispositif.....	49
Figure 4.7 : Données de satisfaction utilisateur.....	50
Figure 4.8 : Moyennes de débit par dispositif, niveau de difficulté 1.....	56
Figure 4.9 : Moyennes de débit par dispositif, niveau de difficulté 2.....	58
Figure 4.10 : Moyennes de débit par dispositif, niveau de difficulté 3.....	59
Figure 4.11 : Moyennes de taux d'erreurs par dispositif.....	63
Figure 4.12 : Moyennes de taux d'erreurs par tâche.....	64
Figure 4.13 : Moyennes de taux d'erreurs par niveau.....	64
Figure 4.14 : Moyennes de taux d'erreurs par dispositif et par tâche.....	65
Figure 4.15 : Moyennes de taux d'erreurs par dispositif et par niveau.....	66
Figure 4.16 : Moyennes de taux d'erreurs par tâche et par niveau.....	66
Figure 4.17 : Moyennes de taux d'erreurs par dispositif et par tâches, niveau 1.....	67
Figure 4.18 : Moyennes de taux d'erreurs par dispositif et par tâches, niveau 2.....	68
Figure 4.19 : Moyennes de taux d'erreurs par dispositif et par tâches, niveau 3.....	68
Figure F.1 : Expérience avec les GETV.....	95

Figure F.2 : Expérience avec les petits écrans tactiles..... 95

Figure F.3 : Expérience avec les pavés tactiles..... 96

Figure F.4 : Expérience avec la souris..... 96

Figure F.5 : Expérience avec les boules de commande militarisées..... 97

Figure F.6 : Expérience avec les boules de commande commerciales..... 97

## LISTE DE LA TERMINOLOGIE

- GETV : Grand Écran Tactile Vertical.
- Système C4ISR : système de Commandement, Contrôle, Communications, Informatique, Renseignement, Surveillance et Reconnaissance.
- DCS : Système de contrôle des dommages (Damage Control System).
- Frappe : ce terme désigne une action unique sur une cible à l'écran par un participant. C'est l'équivalent tactile d'un clic de souris. Par souci de simplicité et d'uniformité, on parlera de frappe pour tous les dispositifs, y compris la souris.
- Série : ce terme désigne l'exécution par un participant d'un test de frappe standardisé tel que défini par ISO 9241-9. Une série est constituée de 24 frappes. Chaque série évalue une condition expérimentale pour un dispositif donné.
- Test : ce terme désigne l'exécution par un participant de l'ensemble des six séries de frappes pour un dispositif donné. Un participant exécutera six tests (un par dispositif) au cours de l'expérimentation.
- Campagne : la période d'expérimentation a été divisée en deux campagnes successives.
- Définition (d'un écran) : densité de pixels affichés à l'écran. S'exprime en pixels par pouce linéaire (ppi).
- Résolution : quantité de pixels affichés sur un écran. Exemple : 1680 \* 1050.

## **LISTE DES ANNEXES**

ANNEXE A - RÉPARTITION DES DONNÉES PAR DISPOSITIF.....	84
ANNEXE B - DONNÉES BRUTES RECUEILLIES PAR LE CODE.....	86
ANNEXE C - PAGES DU SITE DE TEST.....	90
ANNEXE D - FORMULAIRE D'APPROBATION SOUMIS AU CER.....	93
ANNEXE E - QUESTIONNAIRE DE SATISFACTION.....	94
ANNEXE F - RÉPARTITION DE L'EXPÉRIENCE.....	95
ANNEXE G - MATRICE DE PASSAGE.....	98

# INTRODUCTION

L-3 MAPPS conçoit, développe et commercialise des systèmes de simulation et d'entraînement pour la marine, l'énergie et l'espace. Certains de ces systèmes font usage de grands écrans tactiles verticaux (GETV). Or l'utilisation prolongée de tels écrans, en position debout et nécessitant des déplacements, crée des problèmes de fatigue pour l'opérateur humain et ne semble pas très efficace. L'objectif de cette étude était d'évaluer des dispositifs d'interaction alternatifs par rapport à ces GETV.

À bord des navires militaires et civils, la gestion des incidents est une fonction essentielle à la survie de l'équipage et à la bonne conduite de la mission. Cette activité, longtemps effectuée sur des supports papiers et des consoles câblées, a graduellement évolué vers des systèmes informatisés. Cette gestion est de plus en plus intégrée aux organes et capteurs du navire, et peut être effectuée par n'importe quel membre d'équipage que ce soit pour le signalement d'un incident, la planification spatiale des éléments d'un plan de contingence, ou la communication entre les équipes d'intervention. L'étude présentée ici vise à contribuer à améliorer le travail à bord des navires faisant usage de GETV.

Les progrès récents en matière de dispositifs tactiles font progressivement disparaître les périphériques classiques (clavier / boule de commande) des compartiments d'opérations au profit d'écrans tactiles de très grandes dimensions. Ce changement, ainsi que l'apparition dans le domaine naval de nouvelles technologies telles que les tablettes sans fil pour la prise d'information ou les écrans tactiles multi-points, imposent de repenser le paradigme d'interaction afin de mieux tirer parti de ces nouveaux éléments.

Le pilotage de grands écrans tactiles (de 22 à 65 pouces) positionnés verticalement (GETV) engendre actuellement plusieurs problèmes, amplifiés par la longueur des périodes de service (entre deux et quatre heures suivant les habitudes de travail), l'urgence de la situation, et surtout les contraintes du service à la mer (tangage, roulis, secousses, chocs).

Ces problèmes peuvent être listés tels que suit :

1. Dans une configuration tactile classique (tablette ou téléphone intelligent), l'écran tactile est tenu à la main, voire posé à plat sur une surface horizontale. Le guidage du geste s'effectue typiquement depuis le coude, le poignet ou la base du pouce. Le positionnement vertical des écrans muraux demande un guidage du bras depuis l'épaule, ce qui entraîne un geste plus ample et moins précis.
2. Cette même position du bras en extension amène une rapide augmentation de la fatigue physique.
3. La distance d'action pour une surface tactile murale est celle du bras humain. À cette distance, l'angle visuel humain n'observe avec clarté qu'une faible partie de la surface d'affichage. La qualité des sollicitations issues de la périphérie de la surface est fortement réduite.
4. Enfin, des observations que nous avons faites sur le terrain révèlent que le grand écran est perçu comme envahissant par l'utilisateur, impactant négativement sur le stress et l'attention.

Cette recherche se propose de tester et de comparer la performance et la satisfaction humaine liées à l'utilisation de solutions de pilotage alternatives pour le GETV, permettant d'atténuer ou d'éliminer les problèmes relevés.

Structure du mémoire :

Le premier chapitre présente une revue de littérature qui permet de mettre en lumière les problèmes rencontrés dans l'usage courant d'un GETV, et les résultats de plusieurs études de comparaison de la performance des dispositifs d'interaction menées au cours des dernières décennies. Le second chapitre présentera la problématique, les objectifs et l'hypothèse de notre projet de recherche. Le troisième chapitre exposera la méthodologie employée pour comparer la performance des différents dispositifs d'interaction. Le quatrième chapitre détaillera les résultats bruts des données collectées. Le cinquième chapitre présentera les analyses statistiques qui ont été menées sur les données brutes et une discussion des résultats. Nous concluons en soulignant les principaux résultats et en proposant quelques pistes de recherche pour la suite.



# **CHAPITRE 1 : REVUE DE LITTÉRATURE SUR LA PERFORMANCE ET LA SATISFACTION HUMAINE LIÉES À L'UTILISATION DE DISPOSITIFS D'INTERACTION À L'ÉCRAN**

Ce chapitre présente l'état de l'art de la recherche sur la performance humaine et la satisfaction liées à l'utilisation de différents dispositifs d'interaction à l'écran. Il comprend six sections. La première section décrit les problèmes rencontrés actuellement dans l'usage courant d'un GETV. La deuxième section délimite l'impact potentiel de l'amélioration des modes d'interaction avec un GETV. La troisième section présente les différentes études qui ont conduit à l'établissement du modèle de performance utilisé dans la nôtre. La quatrième section présente les études dont les données peuvent être comparées à la nôtre. La cinquième section cherche à prédire les performances des dispositifs mis à l'étude sur la base des modèles trouvés dans la littérature. La sixième section présente les questions auxquelles cette étude tentera de répondre.

## **1.1 Introduction**

Ce projet de recherche a été initialement orienté par l'opportunité d'utiliser ces périphériques peu communs que sont les GETV. Un ensemble d'observations sur le terrain au début de la recherche nous a permis de mettre en évidence certaines difficultés des utilisateurs lors de l'utilisation des GETV, notamment lors de l'atteinte de cibles de petite taille. Nous avons pu assister aux séances d'essai du Damage Control System par les personnels de la Marine royale canadienne sur le site de l'entreprise d'accueil. On a ainsi pu noter les doléances des utilisateurs envers le système. On a constaté que les utilisateurs professionnels, quelque soit leur niveau d'expérience avec le dispositif, sont frustrés par plusieurs aspects du fonctionnement.

### **FONCTIONNEMENT DU SYSTÈME DE CONTRÔLE DES DOMMAGES**

Les marins que nous avons eu l'occasion d'observer faisaient partie de l'encadrement des équipages des frégates de classe Halifax de la Marine royale canadienne. La classe de bâtiments militaires Halifax désigne douze frégates multirôles conçues et construites au Canada. La configuration permet d'engager des menaces sous-marines, notamment par la présence de l'hélicoptère de bord, aériennes, et le soutien des opérations terrestres.

La classe est mise en service de 1992 à 1996, et se voit attribuer des missions de souveraineté et d'accompagnement des forces de l'OTAN.

En octobre 2010 a débuté une campagne de modernisation de la classe, incluant un nouveau système de commandement et de contrôle, de nouveaux radars, un nouveau système de guerre électronique ainsi que des systèmes de communication et des missiles améliorés. La modernisation du système de contrôle des dommages fait partie de ce programme de modernisation.

Le système de contrôle des dommages (Damage Control System, DCS) remplace les outils traditionnels (sur papier) pour la gestion des incidents à bord de navires. L'outil évolue depuis 5 ans chez L-3. C'est le premier DCS informatisé de la Marine canadienne et le premier DCS à grands écrans tactiles de L-3.

Le système de contrôle des dommages représente une partie du système de Commandement, Contrôle, Communications, Informatique, Renseignement, Surveillance et Reconnaissance (C4ISR) du bâtiment. Ils permettent la prise en compte des avaries, quels que soient leur origine, et la planification des opérations de lutte contre l'avarie. Un bâtiment de surface à propulsion non nucléaire doit préparer des plans de contingence contre les risques d'incendie, de voie d'eau, d'incident chimique, électrique, de perforation de la coque, de perte d'intégrité structurelle, de percement d'une canalisation, d'endommagement d'un sous-système mécanique, ou encore de radiations (subséquentes à une attaque nucléaire).

Dans le navire, le DCS complet est constitué d'un ensemble de GETVs (7 pour la classe Halifax) reliés entre eux par un réseau redondant. Le système est relié aux différents capteurs (fumées, chaleurs, gaz, etc) et au réseau de caméras de surveillance. En plus du GETV contrôlé par l'officier de contrôle des dommages, les autres écrans sont disposés dans les centres locaux de contrôle des sections du navire, ainsi que dans certains compartiments critiques : salle machine, passerelle, central opérations, etc.

Par exemple, en cas de feu dans un compartiment du bâtiment, l'équipe en charge du système (ou à défaut de la section du navire) touché va signaler l'incident à l'officier de contrôle des dommages. en créant un "drapeau d'avarie" sur un plan du navire affiché sur leur GETV local. Ce drapeau d'avarie caractérise le type d'incident en cours. Pour un feu, on signale le type de carburant, l'heure de détection de l'avarie, l'état de l'incident, et les actions prises pour le résoudre. Depuis un poste centralisé, l'officier de contrôle des dommages va coordonner les actions de l'équipe de système et des équipes d'intervention dans tout le navire. Il va demander via les GETV d'exécuter un script

d'arrêt et mises en route de systèmes, et tracer sur le plan du navire les instructions (personnels, matériels, axes d'approches et actions à exécuter) pour l'intervention des équipes. Pour un feu, l'officier pourra par exemple demander la mise en route des pompes de la section et la coupure de la ventilation dans les compartiments affectés. Simultanément, il pourra demander à une équipe d'intervention d'arroser les parois du ou des compartiments affectés pendant qu'une autre équipe s'équipe de bouteilles d'air comprimé en vue de pénétrer dans le compartiment en feu. Les chefs d'équipes de sections vont rendre compte de leurs progrès ou difficultés dans l'exécution de ces actions. Les plans, les ordres et les comptes rendus sont mis à jours en temps réel sur tout le réseau.

Les communications verbales viennent compléter la transmission d'information, et sont à ce stade mieux perçues par les opérateurs que les mises à jour purement visuelles, du fait de l'effet de diminution de la conscience de la situation qui se crée naturellement en période de gestion de crise.

Auparavant, cette gestion se faisait par de grands plans sur papier représentant le navire, et les ordres et comptes rendus étaient transmis par téléphone. Les systèmes modernes de gestions des risques informatisent le processus en l'intégrant au réseau global du navire.

Ces systèmes sont apparus il y a environ vingt-cinq ans et se sont répandus vers des bâtiments dont la criticité va décroissant. On voit aujourd'hui apparaître, dans une forme réduite, des systèmes de contrôle des dommages à bord de navire de croisière civils.

## MODE D'INTERACTION

Le mode d'interaction courant avec un GETV implique que l'utilisateur se tienne debout, face à l'écran, et se déplace pour pouvoir toucher les cibles à l'écran avec le doigt (ou avec un pointeur ad-hoc tel qu'un stylo), habituellement à bout de bras. Les tâches observées à l'écran nécessitent généralement des actions de frappe, de traçage, ou d'écriture (utilisation d'un clavier virtuel ; hors du cadre de cette étude) ponctuelles, de l'ordre de 10 à 20 secondes par minute, mais l'observation du GETV est quasi permanente.

Les tendances lourdes issues de ces séances d'observation ont fait ressortir notamment les limitations du périphérique GETV ainsi que de l'interface, conçue pour la souris. Spécifiquement, les limitations du dispositif concernent :

## LA PRÉCISION DU GESTE DE FRAPPE

Hammer (2007) montre que le taux d'erreurs pour une tâche de commande de repas sur un écran vertical est jusqu'à 15 % plus élevé que ceux correspondant à des positions de travail avec un écran incliné ou disposé à plat. Chourasia et al. (2013) indiquent que, dans une étude où des sujets en

bonne santé exécutent une tâche d'entrée de données au moyen d'un clavier numérique sur un écran quasi-vertical tactile (écran d'angle et de hauteur fixe à 70° par rapport à l'horizontale et 103 cm par rapport au sol), en position assise vs. debout, les cibles (touches du clavier numérique) de faible dimension (10 et 15 mm) présentent un taux d'erreur supérieur de 20 % pour la position debout. Cet effet disparaît pour les cibles de taille supérieure (20, 25, et 30 mm). Cette même étude indique également un temps de complétion plus long pour les tâches effectuées debout.

## LE CONFORT DE L'OPÉRATEUR

Hammer (2007) indique que la position de travail debout face à un écran vertical entraîne de hauts niveaux d'inconfort pour le bas du dos, le cou et l'épaule. La même étude classe cette position neuvième en terme de préférence utilisateur sur 12 postures examinées.

Chourasia et al. (2013) montrent que sur un groupe de 15 personnes interrogées au sujet d'une tâche d'entrée de données, 13 préfèrent la position de travail assise. Au test de perception de la charge mentale (NASA-TLX, Hart & Staveland, 1988), la même étude affiche un indice de charge mentale de 18,3 pour la position assise contre 23,5 pour la position debout. Mais certaines pratiques de l'industrie semblent indiquer que cet inconfort provoqué par les écrans tactiles verticaux se retrouve quand l'utilisateur est en position assise :

"Apple has filed for a number of patents related to touchscreen iMacs, fueling speculation that the company would soon introduce a touchscreen desktop or laptop computer.

Jobs [Steve, co-fondateur et ancien président du conseil d'administration d'Apple] put a damper on the speculation today saying, "We've done tons of user testing on this, and it turns out it doesn't work. [...]"

It gives great demo but after a short period of time, you start to fatigue and after an extended period of time, your arm wants to fall off. It doesn't work, it's ergonomically terrible." (Jay Yarow, 2010)"

Il semble donc que les écrans tactiles verticaux engendrent un inconfort chez leurs utilisateurs quelle que soit la position de travail.

## LA CRÉATION D'UN FORT EFFET TUNNEL

D'après Bartram, Ware, & Calvert (2003), dans une situation où l'utilisateur complète une tâche au centre de son champ visuel, les sollicitations affichées à la périphérie de ce champ sont souvent ignorées. En effet, des sollicitations statiques (changement de couleur ou de forme d'une icône) à la

périphérie de la zone d'affichage (de 40 à 52 degrés d'angle horizontal par rapport à la ligne du regard de l'humain au repos) ne sont pas détectées dans 17 à 28 % des cas, contre 0 à 8 % de non-détection pour le champ visuel proche (7 à 20 degrés). De plus, dans le cas où elles sont détectées, ces sollicitations statiques périphériques ont un temps de détection de 3200 à 4300 millisecondes, contre 1500 à 3000 ms pour des sollicitations proches de la ligne du regard de l'humain au repos. Considérant une distance moyenne d'interaction face au GETV de 711 mm (Tilley, 1959), 40 degrés d'angle représentent un rayon de 597 mm.

En situation défavorable, si la tâche sur laquelle l'attention de l'opérateur est concentrée est située dans le coin d'un GETV de 46 pouces (identique à celui utilisé dans notre expérience), alors il existe une probabilité de non détection des sollicitations supérieure ou égale à 17 % pour 52 % de la surface d'affichage. Dans le cas d'un GETV de diagonale 65 pouces, cette proportion passe à 76 % de la surface.

## 1.2 Impact

Sans pouvoir établir un compte précis, les utilisateurs professionnels de dispositifs GETV représentent un nombre de l'ordre de plusieurs milliers d'individus dans le monde entier. Ce nombre peut aujourd'hui sembler limité, mais on doit garder à l'esprit que les dispositifs tactiles, quelles que soient leurs dimensions, sont amenés à se répandre. Il est plausible d'imaginer avec la croissance fulgurante du nombre de téléphones intelligents et de tablettes interactives, que dans un futur proche l'interaction tactile constituera le mode primaire d'interaction pour une majorité des utilisateurs. Par exemple, le pourcentage de trafic internet mobile — quasi exclusivement effectué via plateforme tactile — représentait 0,9 % du trafic total en mai 2009 contre 15 % en mai 2013, avec une prédiction de plus de 25 % pour le début de 2015 (Meeker & Wu, 2013).

Dans cette optique, l'étude présentée ici tente de se positionner en précurseur d'une utilisation des GETV de plus en plus répandue.

## 1.3 Historique

On définit les paramètres usuels d'une expérience de comparaison des performances humaines avec des dispositifs d'interaction comme suit :

- L'indice de difficulté (ID) mesure (en bits) la précision que l'on requiert de l'utilisateur pour une tâche. Les formules de calcul de l'indice de difficulté sont fournies au chapitre 3.3. Plus les cibles sont petites et éloignées, plus l'ID est grand.

- Le débit mesure (en bits par seconde) la performance de l'utilisateur en quantité d'information par unité de temps. Il exprime la quantité de temps nécessaire pour compléter une tâche (dans notre cas la frappe de 24 cibles dans le temps le plus court possible) d'une difficulté (ID) donnée. Les formules de calcul du débit sont fournies au chapitre 3.3.
- Le temps d'atteinte de la cible (en secondes) est défini comme l'indice de difficulté divisé par le débit. Il était initialement utilisé pour mesurer la performance de l'utilisateur à la place du débit, ce dernier étant exprimé comme une constante déterminée lors de l'expérimentation.
- Les tâches proposées peuvent être unidirectionnelles (cibles de part et d'autre de l'écran), ou bien multi-directionnelles (cibles disposées en cercle).

L'étude de la comparaison des performances des dispositifs de pointage remonte au début de l'ère informatique. En 1963, Douglas Engelbart (décédé le 2 juillet 2013), du Stanford Research Institute [SRI], inventa la souris.

Rapidement, English, Engelbart, et Melvin (1967) publient un article constituant la toute première étude comparative des performances pour des dispositifs d'interaction à l'écran. Les cinq dispositifs comparés sont : un bâton de commande, un "Grafacon" (équivalent informatique d'un pantographe de dessinateur), un crayon optique, un système de contrôle activé par le genou de l'utilisateur, et la souris nouvellement développée par le SRI. Dans l'article, la tâche examinée est une tâche d'édition de texte (deux tâches de pointage, une taille des cibles fixe, et la distance pointeur / cible variant aléatoirement), avec la volonté de généraliser ces résultats pour d'autres types de tâches d'interface par ligne de commande. L'échantillon était constitué de huit sujets expérimentés connaissant les interfaces de commande par ligne de texte, et de trois sujets inexpérimentés. Les performances étaient déterminées simplement par mesure du temps de complétion de la tâche.

Les résultats de cette étude indiquent que les sujets expérimentés obtiennent de meilleures performances avec la souris (perçue comme satisfaisante et peu fatigante par les deux groupes), alors que les sujets inexpérimentés obtiennent de meilleures performances avec le stylo optique et le contrôle par le genou. Le bâton de commande obtient les performances les plus basses dans les deux groupes.

Outre les résultats scientifiques de cette étude, celle-ci a posé les bases pour un demi-siècle d'études comparatives en facteurs humains. Cependant, il existe une différence majeure entre cette étude et celles qui la suivront : en effet, elle ne mesure pas l'effet de la distance entre le pointeur et la cible.

Il faut attendre l'étude publiée par Card, English, et Burr (1978) pour que soit prise en compte la relation établie par Fitts (1954) entre distance, taille des cibles, et durée du mouvement. Cette relation, dite "loi de Fitts", prédit le temps requis pour parcourir la distance entre un point de départ et une cible, en fonction de la taille et de la distance de cette cible. Elle s'applique aux interfaces informatiques aussi bien qu'au monde physique. De manière simplifiée, la loi de Fitts implique que plus une cible est éloignée du point de départ, et plus sa taille est petite, plus le temps d'atteinte augmente.

Dans l'étude de Card et al., à la différence de la précédente, les auteurs comparent quatre dispositifs : une souris, un bâton de commande, les quatre touches directionnelles du clavier, et un groupe de quatre touches d'édition de texte ayant disparu des claviers modernes. Cette étude contrôle les effets induits par la distance et la taille de la cible, et cherche pour la première fois à comparer les performances obtenues aux prédictions de la loi de Fitts. L'échantillon est composé de cinq sujets inexpérimentés. La tâche, exécutée sur une interface par ligne de commande, consiste à déplacer un curseur vers une cible (mot ou phrase au sein d'un texte), puis à rendre compte de l'atteinte de la cible par une touche (au clavier ou à la souris). La tâche présente cinq distances, quatre tailles de cibles, et 10 positions de la cible différentes, définissant 200 conditions expérimentales différentes. Chaque sujet effectue un total compris entre 1200 et 1800 passages sur chaque dispositif (sur plusieurs jours, apprentissage inclus).

Les résultats montrent que la souris constitue le dispositif à la fois le plus rapide et le plus précis, avec un temps moyen d'atteinte de la cible égal à 1,66 secondes, et un taux d'erreur moyen de 5 %. Le bâton de commande se classe second (1,83 s ; 11 %), les touches d'édition troisièmes (2,26 s ; 9 %), et les touches directionnelles dernières (2,51 s ; 13 %).

L'article indique que la performance (définie en terme de temps de mouvement) des dispositifs continus (la souris et le bâton de commande) est prédite par la loi de Fitts, alors que pour les dispositifs à base de touches (directionnelles ou d'édition de texte), le temps de mouvement est proportionnel au nombre de frappes au clavier.

Concernant la prédiction des performances en utilisant la loi de Fitts, et à la différence des travaux qui succéderont, les auteurs choisissent de caractériser les performances en termes de temps de mouvement, et non de débit. D'autre part, l'étude utilise la formulation de l'indice de difficulté établie par Welford (1960). Cette formulation, largement employée jusque dans les années 1990, a progressivement été remplacée par celle dite de Shannon. On présente ici les différentes formulations exprimées en fonction du temps de mouvement  $t_m$  pour faciliter la comparaison :

Formulation de Fitts :

$$t_m = a + b \cdot \log_2\left(\frac{2w}{w}\right)_{(1)}$$

Où :

- $t_m$  temps de mouvement.
- $a$  et  $b$  sont des constantes déterminées empiriquement,
- $w$  est la largeur de la cible affichée le long de l'axe d'approche.

Formulation de Welford :

$$t_m = a + b \cdot \log_2\left(\frac{d + 0,5w}{w}\right)_{(2)}$$

Où :

- $d$  est la distance entre le point de départ et la cible.

Formulation de Shannon :

$$t_m = a + b \cdot \log_2\left(\frac{d + w}{w}\right)_{(3)}$$

Toutefois, l'étude de Card et al. est une contribution fondamentale aux sciences de l'interaction humain-machine. Il s'agit en effet de la première instance de l'utilisation d'un modèle qui fait encore référence plus de 40 ans plus tard, dans un domaine pourtant extrêmement dynamique.

Les études se basant sur la loi de Fitts pour établir un comparatif de dispositifs en termes de débit et de taux d'erreur sont graduellement devenues plus nombreuses avec l'abandon des interfaces par ligne de commande et la popularisation des interfaces graphiques.

On peut citer entre autres l'étude de Sperling et Tullis (1988) qui se concentre sur une comparaison des performances des dispositifs possédant a priori de meilleures performances : la souris et la boule de commande (trackball). L'article compare ces deux dispositifs sur trois tâches de sélection (clic), déplacement (glisser-déposer) et traçage. L'étude est divisée en deux parties. La première partie évalue les performances de 12 sujets, dont six expérimentés à la souris, et six expérimentés à la boule de commande, uniquement sur la tâche de sélection. La seconde partie évalue les performances de 24 sujets sur l'ensemble des trois tâches.



Pour la première partie, les résultats ne montrent pas de différences significative entre la boule de commande et la souris en termes de taux d'erreurs, mais la souris affiche un temps d'exécution des tâches inférieur de 15 % en moyenne. Pour la seconde partie, il n'existe pas de différence significative de taux d'erreurs entre les deux dispositifs pour les tâches de sélection et de déplacement, avec une valeur d'environ 3 % dans les deux cas. En revanche, pour la tâche de traçage, la souris et la boule de commande affichent un taux d'erreurs de 19 % et 38 %, respectivement. Le temps d'exécution ne fait pas apparaître de relation significative entre les dispositifs d'interaction et les tâches. Cependant, il existe des différences significatives entre les dispositifs, avec une souris plus rapide, notamment pour la tâche de traçage, ainsi qu'entre les tâches, la tâche de sélection étant plus rapide que les tâches de déplacement et de traçage (environ 20, 55 et 70 secondes en moyenne, respectivement). L'étude conclut sur une supériorité apparente de la souris pour les tâches examinées.

Bien que ces travaux ne puissent être mis directement en comparaison avec ceux qui leur succéderont, du fait de leurs objets et méthodologies très dissemblables, ils constituent les éléments fondateurs de plus de quatre décennies de recherche en interaction humain-machine. Ils ont permis de poser les bases de la réflexion qui conduiront à l'établissement d'une méthodologie unifiée pour les tests d'évaluation de dispositifs de pointage.

## **1.4 Études comparables**

Après de nombreuses années de tests conçus ad-hoc, une norme (ISO 9241-9) paraît qui unifie la méthodologie employée lors des tests. Dès 1999, Douglas, Kirkpatrick & MacKenzie publient en précurseurs une étude faisant usage du brouillon de la norme paru en 1998.

Cette étude compare deux dispositifs d'interaction pour ordinateur portable, à savoir un bâton de commande isométrique et un pavé tactile (trackpad), au regard de deux des trois dimensions nouvellement prescrites par ISO 9241 : la performance mesurée en débit, et le confort de l'utilisateur. Le confort est mesuré par un questionnaire a posteriori demandant au sujet d'évaluer ce qu'il ressent en termes de fatigue physique et mentale, ainsi que par rapport à la vitesse, la souplesse et la précision du dispositif. La mesure de l'effort (troisième dimension) n'a pas été effectuée dans cette étude par manque d'un matériel approprié. L'étude se concentre sur une tâche de pointage dans deux conditions de test : uni-directionnel (dit Fitts 1D), et multi-directionnel (Fitts 2D). Pour le test uni-directionnel, les cibles sont disposées de part et d'autre de la surface d'affichage, et le sujet frappe la cible affichée alternativement d'un côté à l'autre. Pour le test multi-directionnel, les cibles

sont disposées sur un cercle, et le sujet frappe la cible affichée à  $+\pi$  (cibles paires) ou  $+\pi - 2\pi/n$  (cibles impaires) par rapport à la position initiale,  $n$  étant le nombre total de cibles sur le cercle. L'échantillon est constitué de 24 sujets inexpérimentés (un groupe de 12 par dispositifs).

Pour le test multi-directionnel, le débit du bâton de commande est en moyenne 27 % plus élevé que celui du pavé tactile, cette différence est significative. De même, le bâton de commande présente un taux d'erreurs inférieur de 61 % à celui du pavé tactile. Pour le test uni-directionnel, ni le débit ni le taux d'erreurs ne présentent de différence significative. Enfin, l'étude ne trouve pas de différence significative en termes de confort entre les deux dispositifs examinés.

L'article s'interroge ensuite sur la validité scientifique et pratique des prescriptions d'ISO 9241. Les auteurs concluent sur trois points importants (on ne liste ici que ceux qui n'ont pas été corrigés dans la version finale) : la norme ne fixe pas de modalités pour la prise en compte de l'apprentissage ; elle ne fixe pas une formulation et une étendue des valeurs pour l'indice de difficulté ; enfin, aucune précaution n'est prise pour assurer la comparabilité des résultats entre plusieurs études. Si l'établissement de cette norme constitue un effort louable, elle laisse trop de zones floues pour pouvoir assurer une reproductibilité des résultats.

On s'est limité par la suite à l'examen des études évaluant au moins un des dispositifs retenus dans notre étude, dans des conditions expérimentales identiques à celles que nous avons retenues.

On peut citer notamment l'étude de MacKenzie, Kauppinen & Silfverberg (2001), qui compare les performances d'une souris, d'une boule de commande, d'un bâton de commande, et d'un pavé tactile. Dans cet article, les 12 participants sont expérimentés à la souris et à la boule de commande, mais inexpérimentés au bâton de commande (sauf pour un participant) et au pavé tactile. La tâche examinée est un essai de frappe multidirectionnel tel que décrit par ISO 9241-9. La variable de satisfaction utilisateur n'est pas mesurée. Les résultats en débit montrent des différences significatives, la souris se classant en première position, avec un débit de 4,9 bits / s. On obtient pour la boule de commande, le pavé tactile, et le bâton de commande des valeurs de débit de 3 bits / s, 2,9 bits / s, et 1,8 bits / s, respectivement. Les résultats en termes de taux d'erreurs ne sont pas significativement différents. On trouve un taux d'erreurs de 7 % pour le pavé tactile, 8,6 % pour la boule de commande, 9 % pour le bâton de commande, et 9,4 % pour la souris. La souris est à nouveau le dispositif le mieux classé globalement.

Parmi les études passées en revue, nous en citerons une qui présente un intérêt tout particulier à nos yeux. Dans un article datant de 2002, MacKenzie & Jusoh (2002) s'intéressent aux systèmes de

pointage à distance (mais pour des écrans de petites dimensions). Dans cette étude, les participants sont placés à une distance de 1,5 m, face à un écran de 15 pouces. Les auteurs comparent les performances de trois dispositifs de pointage à distance. Les dispositifs sont : un pointeur portatif "GyroPoint" (ce pointeur peut opérer suivant deux modes : un mode "aérien", où un gyroscope détecte les variations d'angle en lacet et en tangage pour diriger le curseur ; et un mode "bureau" où le pointeur opère comme une souris), un pointeur portatif "RemotePoint", où l'utilisateur pilote un petit bâton de commande au bout du pouce, et une souris. L'échantillon comprenait 12 participants, tous expérimentés à la souris, mais aucun n'ayant d'expérience avec un dispositif de pointage à distance. La tâche est un essai de frappe unidirectionnel tel que décrit par ISO 9241-9. Le questionnaire de satisfaction est également tel que décrit par ISO 9241-9, avec une notation du dispositif sur 12 dimensions (force requise pour l'activation, souplesse d'utilisation, efforts requis par l'utilisation, précision, vitesse d'utilisation, confort global, fonctionnement global, fatigue des doigts, du poignet, du bras, de l'épaule, du cou), mais s'effectue sur une échelle de Likert de 1 à 5 (au lieu de 1 à 7 conformément à la norme).

Les résultats montrent que le dispositif générant le moins d'erreurs de frappe est le RemotePoint, avec un taux d'erreur de 1,6 %. Il est suivi par la souris, le GyroPoint "bureau" et le GyroPoint "aérien", avec respectivement des taux d'erreurs 2,4 %, 2,7 %, et 3,5 %. Pour les performances en débit, on obtient un débit de 4,1 bits / s pour le GyroPoint "bureau". Les autres dispositifs obtiennent 3,7 bits / s pour la souris, 2,8 bits / s pour le GyroPoint "aérien" et 1,4 bits / s pour le RemotePoint. En termes de satisfaction utilisateur, la souris obtient un score moyen de 5/5, le GyroPoint "bureau" 5/5, le GyroPoint "aérien" 3,1/5, et le RemotePoint 1,9/5. L'étude conclut sur une supériorité sans conteste de la souris, et les mauvaises performances des deux dispositifs de pointage à distance.

Une comparaison détaillée des résultats de cette étude avec la nôtre est faite au chapitre 6.3. On aura pris soin de ne comparer les résultats que pour des conditions similaires (tâche unidirectionnelle, étendue de l'indice de difficulté comparable).

Outre la méthodologie tirée d'ISO 9241, beaucoup des choix de paramétrage qui apparaissent dans les propositions de recherche et dans la méthodologie de notre étude découlent des recommandations exprimées par Soukoreff & MacKenzie (2004), dans leur revue exhaustive des études d'évaluation des dispositifs de pointage. Ce travail très important explore les résultats et les méthodologies de 34 études, de 1978 à 2004. Les auteurs établissent notamment sept recommandations qui contribuent à restreindre le cadre expérimental dans des limites utiles. Elles

comblent ainsi certains des flous apparaissant dans ISO 9241. Au vu de leur importance, elles sont ici citées dans leur intégralité :

- I. Use the Shannon formulation of ID.
- II. Use a wide and representative range of ID values (ID values ranging from 2 to 8 bits should apply to most situations).
- III. Measure the scatter of movement end-point positions, using error rates and/or end-point data.
- IV. Perform the adjustment for accuracy to convert the ID values to  $ID_e$  (la formule de calcul de l'indice effectif de difficulté  $ID_e$  est donnée au chapitre 3.2) values. Note any large discrepancy between the ID and  $ID_e$  values.
- V. Use linear regression to measure the goodness of fit (to decide whether Fitts' law indeed applies) and to verify that the intercept is small. Investigate and explain a large intercept value.
- VI. Limit predictions made from the resulting Fitts' law linear regression model to the range of  $ID_e$  values used (and specifically, the intercept should not be misconstrued as the time to make a zero-distance movement).
- VII. If devices or experiment conditions are to be compared, calculate the dependent measure throughput via the mean of means.

(Soukoreff & MacKenzie, 2004)

Sans évoquer en détail les variations de chacun de ces paramètres au cours des différentes études, on s'attachera au plus proéminent d'entre eux : la formulation de l'indice de difficulté.

MacKenzie (1989) prouve les bénéfices issus de l'utilisation de la formulation de Shannon & Weaver (1949) pour l'expression de l'indice de difficulté, par opposition aux formulations de Fitts (1954) et Welford (1960), largement utilisées jusque là. Dans cette étude, l'auteur reprend les données de l'expérimentation originale de Fitts (1954), et examine les coefficients de corrélation entre le temps de mouvement et l'indice de difficulté (ID) pour les trois formulations de l'indice de difficulté (exprimées dans l'article en termes de temps de mouvement  $t_m$ , elles sont présentées plus haut dans ce chapitre).

Tableau 1.1 : Coefficients de corrélation entre le temps de mouvement et l'indice de difficulté

Modèle	Frappe (28 g)	Frappe (454 g)
Fitts	0,9907	0,9885
Welford	0,9927	0,9911
Shannon	0,9938	0,9927

La formulation de Shannon exprime au mieux la relation existant entre le temps de mouvement et l'indice de difficulté. D'autre part, elle permet dans certains cas d'éviter d'obtenir des valeurs de débit négatives (MacKenzie & Buxton, 1993). On constate aussi que l'emploi d'une formulation particulière par un auteur rend invalide la comparaison directe des résultats avec ceux d'un auteur employant une autre formulation ; à moins de disposer des résultats complets, à l'image du travail effectué par MacKenzie en 1989.

Dans notre étude, on s'est attaché non seulement à respecter la méthodologie établie par ISO 9241-9, mais également à se conformer aux recommandations exprimées ci-dessus (Soukoreff & MacKenzie, 2004).

## 1.5 Prédiction des performances

La tendance générale qui apparaît à l'issue de cette revue place clairement la souris en tête de classement dans les différentes études, quelles que soient les tâches observées. Ce résultat n'est pas surprenant : la souris est un dispositif commercialisé depuis 1968 ([www.computerhistory.org/revolution/input-output/14/346/1874](http://www.computerhistory.org/revolution/input-output/14/346/1874)), et elle domine le marché des dispositifs de pointage grand public depuis les années 1980. Toute personne susceptible de participer à une étude sur les dispositifs de pointage pourra donc être considérée comme experte à la souris. Même si les temps d'apprentissages des différents dispositifs comparés dans notre étude sont relativement courts, comme on le verra plus loin, on s'attend donc à des performances de la souris dépassant celles des autres dispositifs, y compris en termes de satisfaction.

La prédiction des performances humaines pour les autres dispositifs s'avère plus malaisée. Si le comportement de certains d'entre eux est connu, la boule de commande étant globalement plus performante que le pavé tactile mais moins que la souris, leurs performances relatives dans une configuration nouvelle (pilotage à distance) n'est pas aisément prévisible. Le niveau de performance avec le dispositif de référence, le GETV, est quant à lui complètement inconnu.

Les problèmes identifiés avec les GETV dans la première partie de la revue de littérature nous amènent à nous demander si le GETV est un dispositif d'interaction efficient, en terme de performance humaine, pour accomplir une tâche donnée. Pour y répondre, on se propose de comparer ce dispositif à d'autres, dans une configuration tentant de résoudre ces problèmes (en plaçant le sujet à distance de l'écran). Les critères de comparaison qui doivent être retenus font l'objet d'un consensus dans la communauté scientifique. La revue aura cependant montré l'importance qu'il convient d'accorder à la formulation des critères et aux plages de valeurs pour lesquelles ces critères doivent être comparés.

Dans ce chapitre, nous avons dressé une liste non exhaustive des travaux menés ces quarante dernières années comparant des dispositifs d'interaction entre eux. On retiendra notamment l'importance de l'établissement d'une méthodologie commune afin d'assurer la comparabilité des résultats. Au chapitre suivant on précisera la problématique, les objectifs et l'hypothèse de recherche qui ont gouverné ce travail.

# CHAPITRE 2 : PROBLÉMATIQUE, OBJECTIFS, HYPOTHÈSE

Ce chapitre comprend six sections. La première pose la problématique ainsi que la question de recherche. Les deuxième et troisième sections présentent respectivement les objectifs et l'hypothèse de recherche qui ont été fixés initialement. La quatrième section présente les variables indépendantes et dépendantes retenues pour l'expérimentation. La cinquième section présente les choix qui ont été faits durant la création de l'expérience. La sixième section présente le processus d'établissement des conditions expérimentales.

## 2.1 Problématique

Les observations de terrain que nous avons effectuées auprès de plusieurs marins canadiens utilisant le GETV montrent que, comparativement à des dispositifs d'interaction traditionnels (ex. souris, stylet) utilisés sur de petits écrans, un GETV diminue le confort de l'utilisateur dans des proportions qui pourraient s'avérer dangereuses dans certains contextes professionnels.

En particulier, la taille de ces dispositifs (voir Tableau 2.1) impose à l'utilisateur des mouvements importants du bras, des rotations du torse, voire des déplacements du corps pour les individus de moindre stature, pour atteindre des cibles situées de part et d'autre de l'écran. D'où des temps d'exécution de tâche plus lents. Il existe aussi une possibilité de non-détection ou de détection lente des nouvelles informations affichées aux extrémités d'un écran de grande taille.

Tableau 2.1 : Dimension des GETVs disponibles sur le site de test

Diagonale 65 pouces	Diagonale 165 cm	Largeur 143 cm	Hauteur 80 cm
Diagonale 46 pouces	Diagonale 117 cm	Largeur 102 cm	Hauteur 57 cm

Afin de tenter de remédier au problème d'accès rapide de l'utilisateur à la zone d'interaction, on imagine l'utilisation de dispositifs déportés, servant de périphériques d'entrée principaux. Le GETV serait alors utilisé prioritairement comme dispositif d'affichage, et verrait son usage tactile se spécialiser dans le travail ponctuel ou de détail, ou le travail collaboratif à proximité de l'écran.

## 2.2 Objectifs de recherche

Notre étude poursuit les trois objectifs suivants :

1. Tester et comparer la performance et la satisfaction humaine liées à l'utilisation de six dispositifs d'interaction (GETV, pavé tactile, tablette, souris, boules de commandes militarisée et commerciale) avec un grand écran vertical en vue de choisir le dispositif d'interaction optimal avec ce type d'écran.

Ce faisant, notre étude permet aussi de :

2. Démontrer la possibilité d'amélioration du mode d'interaction actuel pour les GETVs en termes de performance et de satisfaction humaine .
3. Tester la viabilité et la performance de différentes commandes tactiles déportées pour les grands écrans verticaux.

## 2.3 Hypothèse

$H_1$  : L'utilisation d'un dispositif d'interaction déporté (c'est-à-dire placé à distance de l'écran vertical) comparativement à une commande tactile directe améliore significativement la performance à la tâche de pilotage d'un curseur sur un écran tactile de grandes dimensions disposé verticalement.

Ces dispositifs déportés peuvent également améliorer la satisfaction utilisateur par un meilleur confort d'utilisation.

Le projet de recherche réalise la comparaison de six dispositifs déportés avec un grand écran vertical. Chaque dispositif est évalué par rapport à la complétion d'une tâche standardisée sur les critères usuels de l'utilisabilité.

### ORIGINALITÉ DE L'HYPOTHÈSE

L'utilisation classique d'un GETV (D1) entraîne l'augmentation du temps d'exécution des tâches, du taux d'erreurs, et de l'inconfort par rapport aux dispositifs d'interaction classiques (écran tactile de petites dimensions, pavé tactile, clavier / souris (Bartram et al., 2003 ; Chourasia et al., 2013 ; Hammer, 2007 ; Yarow, 2010).

À la connaissance de l'auteur, il n'existe pas de données chiffrées concernant l'usage de GETV en tant que dispositifs d'interaction, ou caractérisant le pilotage à distance des écrans de grandes dimensions via une surface tactile.



## 2.4 Définition du projet de recherche

### CONCEPTION DU TEST

Il arrive souvent que les paramètres des études de comparaison des dispositifs d'interaction avec un écran (indice de difficulté, type de tâches) soient fixés ad-hoc par les auteurs des études. Cette liberté présente l'inconvénient d'empêcher une comparaison rigoureuse entre des données issues de différentes études. Afin d'éviter cet état de fait et de pouvoir effectuer des comparaisons de performances entre les résultats obtenus, nous avons utilisé une méthodologie commune à plusieurs auteurs. Cette volonté d'obtenir des résultats qui puissent se comparer à ceux de la littérature nous vient de la lecture de l'article de Soukoreff & MacKenzie (2004), qui ont recensé et confronté les résultats issus de nombreuses études entre 1978 et 2004. La conclusion de cette étude a mis en évidence la nécessité d'utiliser une méthodologie commune, afin de pouvoir comparer les résultats sur une base commune, et également de pouvoir détecter d'éventuels biais expérimentaux. La recherche de comparabilité des résultats est cruciale pour atteindre les deux premiers objectifs de recherche définis ci-dessus.

La plupart des choix méthodologiques qui découlent de cette intention sont directement issus de la norme ISO 9241-9 qui définit les conditions expérimentales dans lesquelles le test dit de la loi de Fitts doit s'effectuer (test standardisé "Essai d'efficience et d'efficacité"). Cette norme permet d'évaluer quantitativement l'utilisabilité d'un dispositif au regard des tâches élémentaires suivantes : pointage, sélection, déplacement ("glisser- déposer"), entrée à main levée, traçage. Elle décrit un ensemble de conditions de test standardisées bâties autour de la loi de Fitts. Ces conditions sont catégorisées en fonction du dispositif et du type de tâche qu'elles évaluent : test de frappe (pointage, sélection), unidirectionnel ou multidirectionnel ; essai de déplacement (déplacement, traçage), unidirectionnel ou multidirectionnel ; essai d'entrée à main levée ; tâche de préhension et de remise en place du dispositif (synthèse des tâches). Le tableau 2.2 page suivante présente les types de tâches qu'il est, selon la norme, possible d'exécuter avec différents dispositifs d'interaction avec un ordinateur en vue de comparer les performances de ces derniers.

Tableau 2.2 : Types de tests disponibles en fonction du dispositif testé.

Dispositifs	Tâches élémentaires				
	Pointage	Sélection	Déplacement	Traçage	Entrée
Souris	✓	✓	✓	✓	
Boule de com.	✓				
Tablette	✓	✓	✓	✓	✓
Styler	✓	✓		✓	
Photostyle	✓	✓		✓	
Ecran tactile	✓	✓			

Le couple Tâches élémentaires / dispositifs correspondant à l'usage pratique des dispositifs retenus (ici souris, boule de commande, tablette et écran tactile) autorise l'utilisation des tests de pointage et de sélection.

#### ESSAIS DE FRAPPE UNI- ET MULTI-DIRECTIONNELS

Le choix entre le test de frappe uni- ou multidirectionnel ne devrait se faire que sur la base de la tâche réelle testée. La majorité des tâches à l'écran étant en deux dimensions, le test unidirectionnel ne devrait donc se rencontrer que très rarement. Il est cependant sur-représenté dans la littérature, et en particulier depuis 1999, date d'introduction de la norme. Sur 34 études totalisant 57 tâches relevées (Soukoreff & MacKenzie, 2004), la procédure unidirectionnelle représente 11 tâches depuis 1999 (36 %), contre six tâches entre 1978 et 1996 (23 %). Il est permis de penser que le test unidirectionnel, défini par Fitts en 1954, a été redécouvert et formalisé à l'occasion de la publication de la norme. On a décidé d'inclure les tests unidirectionnels afin d'obliger les participants à utiliser toute la largeur du GETV.

#### QUESTIONNAIRES DE SATISFACTION

Les questionnaires de satisfaction des utilisateurs proviennent de la norme ISO 9241. Un spécimen du questionnaire est fourni en annexe E de ce document.

L'utilisation du test de la charge mentale NASA-TLX (Hart, Staveland, 1988) avait initialement été envisagée en complément du questionnaire de satisfaction. Cependant, au vu des délais accordés à l'expérimentation, cette partie du test a été supprimée.

## 2.5 Propositions

### DÉTERMINATION DES CONDITIONS EXPÉRIMENTALES

Tenter de reproduire les tâches quotidiennes des utilisateurs finaux du GETV (tâches complexes faisant apparaître des cibles sur toute la surface de l'écran) exigerait la sélection multidirectionnelle de cibles, sur la pleine largeur de l'écran, en faisant usage des deux mains, voire parfois de plusieurs doigts par main. Or il n'existe pas dans la norme de condition permettant de tester cette exacte situation. Notamment, la norme ne prévoit que l'utilisation de pointeurs mono-point, excluant de fait les écrans multi-points modernes qui représentent aujourd'hui l'immense majorité du marché des écrans tactiles. On avait prévu au début de la recherche de tester suivant une seule condition expérimentale :

- C0 : Permettre la sélection des cibles avec une seule main en utilisant toute la largeur de l'écran. Ceci implique l'exécution d'un test de Fitts unidirectionnel. Il n'est pas nécessaire d'effectuer un test unidirectionnel sur la hauteur de l'écran. Celle-ci étant inférieure d'un rapport 1,778 à la largeur, il est raisonnable de penser qu'elle est susceptible de générer des résultats moins significatifs.

Cette condition avait été définie alors que le GETV disponible pour le test était un matériel d'ancienne génération, à détection infrarouge. L'obtention de performances correctes ne pouvait se faire qu'en prenant un certain nombre de précautions (sélection des cibles au moyen d'un stylet, problèmes de parallaxe). Quand un GETV à détection optique (triangulation de la position des doigts sur l'écran par des caméras noyées dans la bordure) est devenu disponible, cette condition a été retirée.

Plusieurs couples de conditions expérimentales possibles ont par la suite été proposées (notées C1x pour la condition multidirectionnelle et C2x pour la condition unidirectionnelle) :

- C1 : permettre la sélection des cibles avec une seule main en utilisant la moitié de la largeur de l'écran : prise en compte de l'aspect sélection multidirectionnelle. Il faut noter que pour le test multidirectionnel, les cibles s'affichant en cercle, il est impossible d'utiliser un rayon plus grand.
- C2 : permettre la sélection des cibles avec deux mains (suivant la préférence de l'utilisateur) en utilisant la pleine largeur de l'écran : prise en compte des problèmes posés par la taille de l'écran du GETV (perception visuelle sur une surface d'affichage plus large que l'angle visuel au repos, éventuels déplacements).

Ou bien :

- C1 : idem ; et,
- C21 : permettre la sélection des cibles avec deux mains, au choix, en utilisant la pleine largeur de l'écran. Cette condition permettrait de s'approcher un peu plus des conditions de l'emploi réel (sélection multidirectionnelle, pleine largeur de l'écran, deux mains, multi-point possible). Compter au cours de l'exécution des tâches le nombre de frappes effectuées avec la main droite, afin de déterminer si l'utilisation du GETV "force" un comportement non-naturel (par rapport à D2 ou D3) et si ce comportement contribue à une diminution des performances.

#### CONDITIONS RETENUES

- L'utilisation de la définition ISO du test de la loi de Fitts a cependant imposé un cadre strict, délimitant fortement les conditions expérimentales retenues :
- Il n'a finalement pas été permis aux sujets de faire usage des deux mains pour pointer à l'écran, la norme ne définissant pas de condition multi-point. On discutera en conclusion des problèmes que pose cette restriction dans un environnement où les périphériques tactiles multi-points sont la norme.
- Autoriser C21 aurait de plus conduit à biaiser inconsidérément les résultats en faveur du GETV et de la tablette (seuls capables de prendre en compte deux pointeurs ou plus) par rapport aux autres dispositifs.
- Les déplacements du corps (y compris les mouvements de jambes) n'ont pas non plus été inclus lors de la définition du projet : si l'on prend en compte l'utilisation des deux mains (hypothèse de l'écran multi-point : sélection d'une seule cible à la fois avec une ou deux mains), l'utilisateur n'a plus à se déplacer pour atteindre une cible, où qu'elle soit. La plus grande distance parcourue correspond alors à la moitié de la largeur de l'écran.

On a finalement retenu les conditions expérimentales suivantes :

- C12 : permettre la sélection des cibles avec un seul doigt en affichant celles-ci sur la moitié de la largeur de l'écran (multidirectionnel).
- C22 : permettre la sélection des cibles avec un seul doigt en affichant celles-ci sur la pleine largeur de l'écran (unidirectionnel).

Cette notation pour les deux conditions étant répétée dans tout le texte, on la simplifie en les renommant C1 pour la condition multidirectionnelle, et C2 pour la condition unidirectionnelle.

## 2.6 Utilisation des résultats de la recherche

Les utilisateurs potentiels des résultats de la recherche incluent :

- Les constructeurs des grands écrans tactiles,
- Les acheteurs institutionnels de grands écrans tactiles (n'utilisant pas directement les dispositifs),
- Les utilisateurs directs et indirects du dispositif :
  - Utilisateurs directs connus : marins militaires et civils, dans plusieurs pays,
  - Utilisateurs directs potentiels : toute personne amenée à utiliser un GETV au quotidien, notamment les personnels des services de l'éducation (tableau blanc interactif),
  - Utilisateurs indirects : cadres responsables de la doctrine d'emploi du dispositif, constructeurs, personnels d'entretien et de commercialisation.

Dans ce chapitre, on a établi les objectifs de cette recherche : on souhaite démontrer que l'usage d'un GETV peut faire l'objet d'améliorations en termes de performances humaines et de confort d'utilisation. Pour ce faire, on se propose de comparer le GETV à cinq autres dispositifs d'interaction en faisant exécuter un test standardisé à un groupe d'utilisateurs. Au chapitre suivant, on précisera la méthodologie employée pour parvenir à ces objectifs.

# **CHAPITRE 3 : MÉTHODOLOGIE DE L'ÉTUDE COMPARATIVE DES DISPOSITIFS D'INTERACTION AVEC UN ÉCRAN**

Ce chapitre comprend six sections. La première présente les caractéristiques biographiques des sujets ayant participé à l'étude. La deuxième section explique le choix des différents paramètres physiques de l'expérimentation pas clair. La troisième section présente les caractéristiques des dispositifs d'interaction testés, ainsi que la description de la procédure de test employée. La quatrième section spécifie la configuration des différents matériels et logiciels employés. La cinquième section fixe la procédure employée lors des tests. La sixième section indique quelles modifications ont été apportées au protocole à l'issue des pré-tests. On s'attache à ce que la recherche soit menée dans un environnement reconstituant de manière simplifiée les conditions d'utilisation opérationnelle du dispositif ; il était impossible d'effectuer les tests directement à bord d'un navire.

## **DISPOSITIFS DE POINTAGE À L'ÉCRAN**

Initialement, la comparaison ne devait s'effectuer qu'entre le GETV et deux autres dispositifs tactiles :

D2 : tablette répétant l'affichage du GETV.

D3 : pavé tactile (trackpad), ne répétant pas l'affichage du GETV.

Au cours de la revue, il est apparu que l'objectif interne de comparaison des performances avec celles issues de la littérature ne pourrait être atteint qu'en incluant des dispositifs pour lesquels des résultats étaient déjà publiés. Il a donc été décidé d'inclure dans la sélection une souris, qui est un périphérique pour lequel un grand nombre de données est déjà disponible. Les quatre dispositifs initiaux ont été testés entre le 30 juillet et le 15 août 2012.

Une fois la campagne de test initiale clôturée, l'entreprise d'accueil a émis le souhait de voir confronté aux dispositifs de la sélection un dispositif usuellement utilisé à bord des navires militaires, à savoir une boule de commande militarisée ("durcie"). Ces périphériques intégralement métalliques sont d'une dimension supérieure (63 mm de diamètre) aux dispositifs commerciaux et peuvent s'avérer d'emploi malaisé après plusieurs heures d'utilisation ininterrompue (inertie due au

poids, angle de la main peu naturel). La question à résoudre pour l'entreprise d'accueil étant de proposer aux industriels réalisant les consoles une alternative plus aisée à employer et à configurer que la boule de commande militarisée, on a également inclus dans la sélection une boule de commande commerciale de diamètre 38 mm en plastique plus légère, dont l'inertie est moindre que la boule militarisée. Suite aux demandes de l'entreprise, deux dispositifs supplémentaires ont été évalués du 17 septembre au 5 octobre 2012. Une campagne de pré-test avait pris place au préalable du 5 au 11 juillet 2012. Cette sélection imposant un temps de test par participant d'environ une heure trente au total, il a été décidé de limiter la liste des dispositifs testés à six.

### 3.1 Configuration

Cette section présente la configuration matérielle, logicielle et physique de notre étude.

#### CONFIGURATION MATÉRIELLE

Aucune modification n'a été apportée aux valeurs de vitesse de pointeur par rapport aux valeurs par défaut données par les différents constructeurs. Pour les dispositifs tactiles, la fonction de gain vaut un, car un déplacement du doigt d'une distance  $d$  correspond mécaniquement à un déplacement du curseur sur la même distance.

- D1 : GETV :
  - Ecran du GETV (toutes configurations sauf D2) : NEC LCD 4610 ; diagonale : 46 pouces ; ratio : 16:9, soit 102 \* 45 cm ; résolution : 1366 \* 768 pixels.
  - Surface de détection du GETV (D1 uniquement) : Smart Board Interactive flat panel ; détection : caméras optiques ; fonction de gain : facteur d'échelle = 1 (écran tactile), c'est à dire qu'un mouvement du bras d'une distance  $a$  correspond à un déplacement du curseur sur l'écran d'une même distance  $a$ .
- D2 : Tablette tactile (la tablette exécute également le code du test pour les configurations D1, D2, D4, D5, D6) : Asus Eee Slate. Ce modèle particulier à été choisi car il possède un ratio d'écran, ainsi qu'une résolution proche de celle du GETV. Diagonale : 12,1 pouces ; ratio : 16:10 ; résolution : 1280 \* 800 pixels ; détection : écran capacitif ; fonction de gain : facteur d'échelle = 1 (écran tactile).
- D3 : pavé tactile Apple Magic trackpad. Ce dispositif particulier a été choisi du fait de la grande surface disponible pour le guidage et la frappe, ainsi que de la possibilité de l'utiliser à grande

distance de l'ordinateur exécutant le code ; fonction de gain : facteur d'échelle = 0,6875 (valeur par défaut sous Mac OS X).

- Ordinateur (configuration D3 uniquement) : Apple MacBook Pro 15" (2008).
- D4 : souris laser Dell trois boutons ; fonction de gain : facteur d'échelle = 1 (valeur par défaut sous Windows 7).
- D5 : boule de commande militarisée (marque et modèle non accessibles) ; diamètre de la boule : 63 mm ; fonction de gain : facteur d'échelle = 1 (valeur par défaut sous Windows 7).
- D6 : boule de commande Logitech Trackman Marble ; diamètre de la boule : 38,1 mm ; fonction de gain : fonction de gain : facteur d'échelle = 1 (valeur par défaut sous Windows 7).

## CONFIGURATION LOGICIELLE

Originellement, la définition du test requérait une exécution sur plusieurs configurations matérielles différentes. Le code a donc été écrit en PHP / JavaScript exécutable dans n'importe quel navigateur. Compte tenu des exigences de sécurité de l'entreprise d'accueil, les stations de travail exécutant le code du test ne peuvent communiquer par réseau (exception faite du réseau wifi local liant le GETV à la tablette). Ceci impose le stockage en local des données issues du test dans une base de données locale au format SQL, ainsi que de la ressource jQuery animant le mouvement des cibles.

Originellement, la répétition de l'image du GETV sur la tablette se faisait par protocole Remote Desktop (RDP) tierce partie. Cependant, les performances obtenues par cette configuration au cours des pré-tests étaient très en deçà des attentes (débit entre 0 et 1 bit, alors que les valeurs issues de la littérature nous permettaient d'attendre un débit entre 2 et 6 bits/s). On a donc choisi d'exécuter le test directement sur la tablette pour la configuration D2.

Pour toutes les configurations :

- OS : Microsoft Windows 7 (sauf D3) ; Mac OS 10.7 (D3 uniquement),
- Stack logicielle : Apache, MySQL, PHP, sous stack XAMPP 1.7.3,
- Navigateur : Google Chrome 19.

## CONFIGURATION PHYSIQUE

Pour toute la durée de l'expérimentation, les éclairages du site de test ont été masqués afin d'éviter les reflets. Les distances d'interactions proposées pour les dispositifs D2 à D6 (rappel du chapitre 3.2 Détermination du paramétrage) sont marquées et numérotées au sol :



- 1,7 m si on considère  $\alpha = 15^\circ$ , soit l'angle de rotation facile de l'œil d'après Tilley (1959) (compromis "réaliste" compte tenu de la taille des compartiments d'opération),
- 2,6 m si on considère  $\alpha = 10^\circ$ , soit l'angle limite supérieure pour la reconnaissance des mots (Tilley, 1959).

Hormis les distances d'interaction, le sujet a la possibilité de repositionner les dispositifs (sauf le GETV), la table et le fauteuil en fonction de sa morphologie et de ses préférences.

Par dispositif, on a :

- D1 : le sujet est placé debout face au GETV, à la distance d'action de son bras. Il effectue les deux tâches (C1 et C2 telles que définies dans le chapitre 4.2 Détermination du paramétrage) en adaptant librement sa posture.
- D2 : le sujet est assis (chaise ajustable en hauteur) de manière à ce que son œil soit à la distance choisie de l'écran. La tablette est posée sur une table devant le sujet. La tablette fait un angle de 15 degrés par rapport à une ligne horizontale afin de réduire les reflets lumineux autant que possible. La tablette répète en miniature l'affichage de l'écran, sans aucune modification hormis la mise à l'échelle.
- D3 : le sujet est assis (chaise ajustable en hauteur) de manière à ce que son œil soit à la distance choisie de l'écran. Le pavé tactile est posé à plat sur une table devant le sujet.
- D4 : idem D3, pilotage par souris, à plat sur la table.
- D5 : idem D3, pilotage par boule de commande militarisée, fixe et à plat sur la table. Un support en styro-mousse a été taillé pour éviter au dispositif de basculer.
- D6 : idem D3, pilotage par boule de commande commerciale, fixe et à plat sur la table.

## 3.2 Sujets

L'étude s'applique à une population d'utilisateurs faisant un usage quotidien, allant de régulier à intensif d'interfaces graphiques et de dispositifs de pointage. L'idéal aurait été de constituer un échantillon à partir des marins canadiens utilisant le GETV au quotidien. Au vu du faible temps passé par ces marins dans les locaux de l'entreprise d'accueil, cela s'est avéré impossible. Compte tenu des ressources et du temps alloué à la recherche, les sujets constituant l'échantillon ont donc été choisis par volontariat au sein de la population de l'entreprise d'accueil susceptible de répondre aux critères d'inclusion et d'exclusion présentés dans le tableau 3.1 page suivante.

Tableau 3.1 : Critères de recrutement des sujets de l'échantillon.

Critères d'inclusion	Critères d'exclusion
<ul style="list-style-type: none"> <li>• Travailler sur le site de l'entreprise d'accueil.</li> <li>• Etre volontaire pour participer à l'expérience.</li> <li>• Avoir déjà utilisé des systèmes informatiques munis d'interfaces graphiques.</li> </ul>	<ul style="list-style-type: none"> <li>• Handicaps aux membres supérieurs.</li> <li>• Incapacité visuelle empêchant l'emploi d'une interface graphique.</li> <li>• Incapacité de maintenir la position debout ou de se déplacer rapidement en position debout.</li> </ul>

Les sujets n'ont reçu aucun paiement pour leur participation à l'étude. L'expérimentation (environ 1h30 sur le total des deux campagnes) s'est tenue pendant les heures de travail.

### RÉSULTATS DE L'ENQUÊTE BIOGRAPHIQUE

L'échantillon initialement prévu était de 24 sujets. Compte tenu des contraintes liées au milieu de travail, 20 sujets ont participé à la campagne 1 (excepté D3, 19 sujets seulement suite à une défaillance technique) et 18 sujets ont participé à la campagne 2. Les deux campagnes ont fait appel aux mêmes sujets. Tous les sujets à l'étude étaient droitiers.

Le questionnaire biographique distribué au début de l'expérimentation demandait à chaque participant d'identifier son âge, son sexe, sa taille, sa main dominante, ainsi que son expérience avec divers dispositifs d'interaction (voir figures F.1 à F.6 en annexe F).

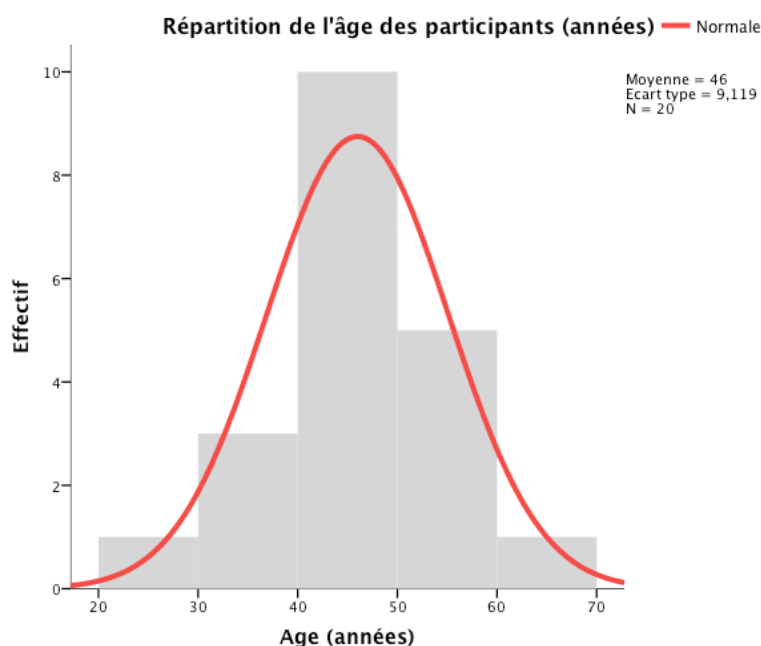


Figure 3.1 : Répartition de l'âge des participants

On observe sur la figure 3.1 page précédente une forte prédominance des sujets âgés de 40 à 60 ans (75 % du total).

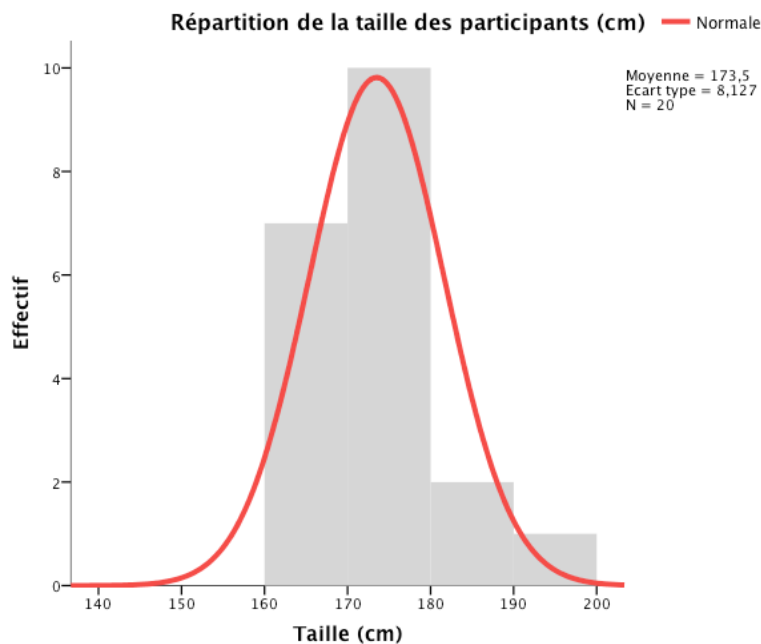


Figure 3.2 : Répartition de la taille des sujets

On observe que 85 % des sujets ont une taille comprise entre 160 et 180 cm.

Sur 20 sujets, seulement trois étaient de sexe féminin (15 % du total). Ce déséquilibre s'explique par la faible proportion de femmes dans un milieu de travail très technique.

#### RÉPARTITION DE L'EXPÉRIENCE EN FONCTION DU DISPOSITIF

Le tableau 3.2 page suivante (graphiques dans l'annexe F) montre les niveaux d'expérience des sujets en fonction des différents dispositifs utilisés.

Tableau 3.2 : Niveaux d'expérience par dispositif.

	GETV	Petits écrans tactiles	Pavés tactiles	Souris	Boules mil.	Boules com.
Alpha	< 10 ans	< 10 ans	< 10 ans	> 10 ans	< 1 ans	aucune
Bravo	< 10 heures	< 10 ans	< 10 ans	> 10 ans	< 10 ans	< 10 heures
Charlie	< 10 ans	< 10 ans	< 10 ans	> 10 ans	< 10 ans	< 10 ans
Delta	< 10 ans	< 10 ans	> 10 ans	> 10 ans	n/a	n/a
Echo	> 10 ans	< 10 ans	> 10 ans	> 10 ans	> 10 ans	> 10 ans
Foxtrot	< 1 ans	< 1 ans	< 10 ans	> 10 ans	n/a	n/a
Golf	< 100 heures	< 1 ans	< 10 heures	> 10 ans	< 1 ans	< 1 ans
Hotel	< 10 heures	< 10 heures	< 100 heures	> 10 ans	< 100 heures	< 10 heures
India	< 100 heures	< 10 heures	< 10 ans	> 10 ans	> 10 ans	> 10 ans
Juliet	< 10 heures	< 10 heures	aucune	< 10 ans	< 1 heures	< 1 heures
Kilo	aucune	< 10 ans	< 10 ans	< 10 ans	< 100 heures	< 10 ans
Lima	< 10 ans	< 10 ans	< 10 ans	> 10 ans	> 10 ans	> 10 ans
Mike	< 10 ans	< 10 ans	< 10 ans	> 10 ans	> 10 ans	> 10 ans
November	< 10 ans	< 1 ans	> 10 ans	> 10 ans	> 10 ans	aucune
Oscar	< 100 heures	< 1 ans	< 10 ans	< 10 ans	< 100 heures	< 1 ans
Papa	< 1 heures	< 1 heures	< 100 heures	> 10 ans	aucune	aucune
Quebec	< 1 ans	< 10 ans	< 10 ans	> 10 ans	< 1 heures	< 1 heures
Romeo	< 10 heures	< 100 heures	< 10 heures	> 10 ans	< 100 heures	< 1 heures
Sierra	< 100 heures	< 1 ans	< 10 heures	> 10 ans	< 100 heures	< 100 heures
Tango	< 10 heures	< 1 heures	< 1 heures	> 10 ans	< 100 heures	< 1 heures

Deux dispositifs se distinguent dans ce comparatif : tous les sujets ont au moins un an d'expérience avec la souris. Ceci confirme la non-pertinence d'une série d'essais pour la souris. Par ailleurs, les petits écrans tactiles ayant été introduits au grand public depuis moins de 10 ans, aucun participant ne possède plus de 10 d'expérience pour ce genre de dispositif. Ce fait doit être nuancé par l'utilisation intensive qui peut en être faite par certains.

### 3.3 Procédure expérimentale

#### DISTANCES D'INTERACTION

Les distances d'interactions proposées ont été choisies en fonction de la largeur maximale des objets affichés à l'écran (1200 pixels, soit 90 cm de largeur pour l'affichage des cibles unidimensionnelles droite et gauche) et d'une série d'angles visuels prédéterminés (voir ci- dessous).

Dans la figure 3.3, on représente la distance de l'œil du participant à l'écran. Les distances sont indiquées en m. Pour tous les dispositifs sauf la tablette, les participants regardent l'écran, quel que soit la distance où ils sont placés. Pour la tablette, les participants regardent directement cette dernière. Pour le GETV (D1, pour dispositif 1), Les participants se placent debout directement devant l'écran, à une distance qu'ils jugent confortable pour atteindre toutes les zones de l'écran. Ils sont invités à faire quelques essais de positionnement afin de vérifier que l'atteinte est possible sans efforts. Pour la tablette (D2) les participants sont assis à une table sur laquelle on a posé la tablette, faisant un angle de  $15^\circ$  avec l'horizontale. Les sujets regardent la tablette uniquement.

Pour tous les autres dispositifs (D3 à D6), les participants sont assis à une table sur laquelle sont posés les dispositifs (voir figure 3.3 page suivante). Le protocole leur permet de choisir entre quatre distances déterminées. On souhaitait ainsi déterminer l'influence de l'éloignement (et donc la taille des cibles) sur les performances, sans toutefois nuire au confort en imposant un placement. Ces distances correspondent à des valeurs remarquables d'angle visuel, compte tenu de la largeur de la surface d'affichage des cibles (Tilley, 1959) :

- 1,7 m si on considère  $\alpha = 15^\circ$ , soit l'angle de rotation facile de l'œil,
- 2,6 m si on considère  $\alpha = 10^\circ$ , soit l'angle limite supérieure pour la reconnaissance des mots,

Des distances de 0,8 et 5,1 m, (correspondant à des angles visuels de  $\alpha = 30^\circ$  et  $\alpha = 5^\circ$ , respectivement) étaient aussi proposés, mais en pratique seules les distances de 1,7 et 2,6 m auront été choisies par les participants.

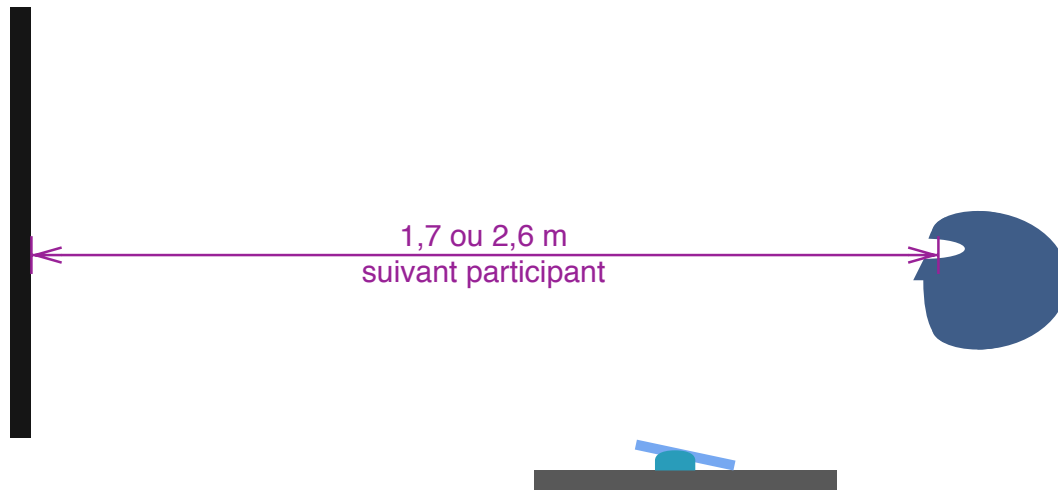


Figure 3.3 : Distance entre l'œil du participant et le GETV (dispositifs D3 à D6).

La hauteur du siège ainsi que la disposition (main dominante, positionnement des éléments, angle de la tablette) des périphériques sur la table sont laissés au choix du participant. Les participants sont invités à effectuer des essais de positionnement au début de chacune de leurs séries d'entraînement.

#### DISTANCES À L'ÉCRAN - TÂCHES & DISPOSITION DES CIBLES

Dans la suite du chapitre,  $d$  représente la distance entre les centres des cibles, et  $w$  la largeur des cibles dans le sens du mouvement.

Tous les éléments visuels du test (contrôles, textes d'instruction, cibles) sont affichés en noir et blanc, afin d'accommoder d'éventuels sujets daltoniens.

Déclenchement du test : le sujet frappe un bouton situé au centre de l'écran (rectangle gris au centre des figures 3.4 et 3.5 pages suivantes). Le curseur n'est pas repositionné d'une cible à l'autre : le point de départ pour la cible suivante est le point de frappe de la cible précédente.

#### TEST DE FRAPPE MULTIDIRECTIONNEL - C1

On choisit des cibles circulaires (et non carrées comme le définit la norme) pour présenter la même largeur  $w$  (ici, le diamètre) indépendamment de l'angle d'approche (Isokoski & Raisamo, 2002).

Le sujet doit frapper 24 cibles circulaires s'affichant une par une sur un cercle : la première cible s'affiche à  $0^\circ$  (haut du cercle), la seconde à  $180^\circ$ , la troisième à  $15^\circ$ , la quatrième à  $195^\circ$ , etc. À tout moment, une seule cible est présente à l'écran, la cible frappée disparaissant immédiatement, et la cible à frapper apparaissant immédiatement.

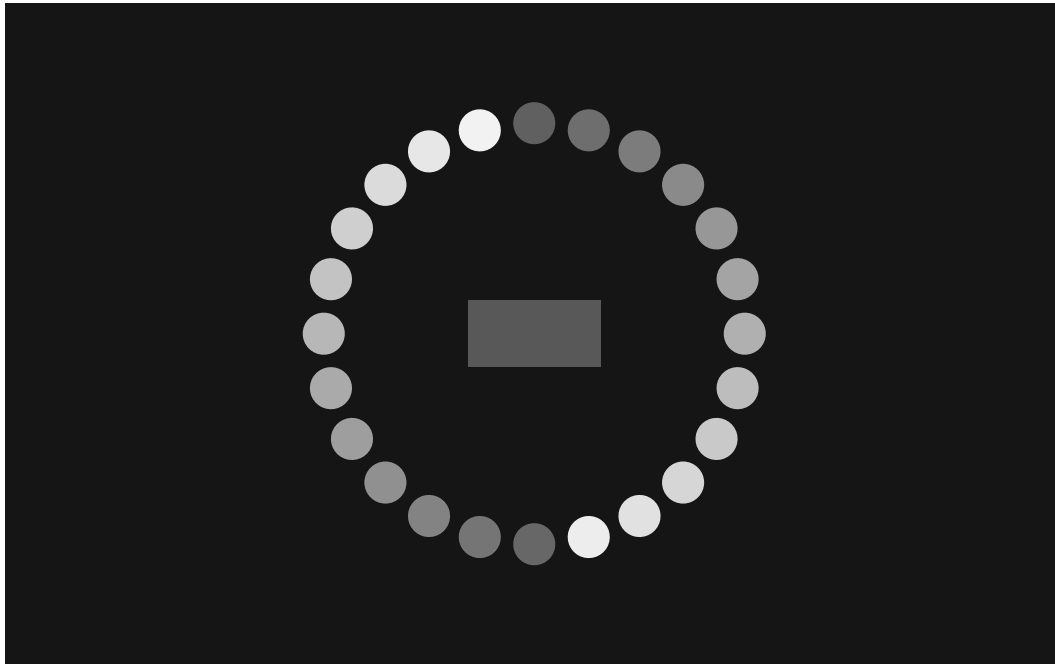


Figure 3.4 : Placement des cibles pour le test de frappe multidirectionnel

À la fin, quelles qu'aient été ses performances, une page confirme au sujet la bonne réussite du test et l'invite à attendre la suite des évènements.

#### TEST DE FRAPPE UNIDIRECTIONNEL - C2

Le sujet doit frapper 24 cibles rectangulaires s'affichant alternativement à droite et à gauche de l'écran. À tout moment, une seule cible est présente à l'écran (voir figure 3.5).

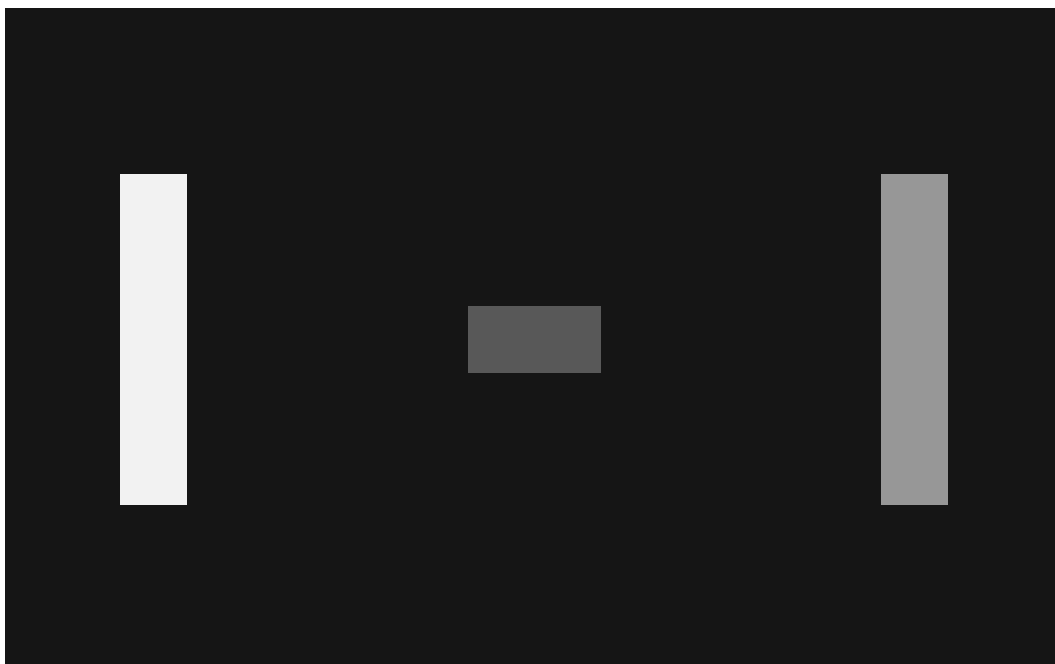


Figure 3.5 : Placement des cibles pour le test de frappe unidirectionnel

À la fin, quelles qu'aient été ses performances une page confirme au sujet la bonne réussite du test et l'invite à attendre la suite des évènements.

### POSITIONNEMENT ET TAILLE RELATIVE DES CIBLES

Les figures 3.6 et 3.7 ci-dessous présentent la répartition spatiale de toutes les cibles juxtaposées pour les deux types de tâches : figure 3.6 le test multidirectionnel, et figure 3.7 le test unidirectionnel. Toutes les dimensions sont homothétiques par rapport à celles du test.

On a représenté les cibles pour les trois niveaux de difficulté, gris clair pour le niveau 1, gris moyen pour le niveau 2, gris foncé pour le niveau 3. Le choix des indices de difficulté dicte les dimensions expérimentales telles que présentées dans le tableau 3.3.

Tableau 3.3 : Valeurs de  $d$  et  $w$  en fonction des indices de difficulté retenus ;  $d$  et  $w$  en pixels.

Condition expérimentale	$d$	$w$	Largeur d'affichage $d+w$	Indice de difficulté ( $ID$ )
Multi N1	500	167	667	1,998
Multi N2	500	33	533	4,014
Multi N3	500	8	508	5,989
Uni N1	900	300	1200	2,000
Uni N2	900	60	960	4,000
Uni N3	900	14	914	6,029

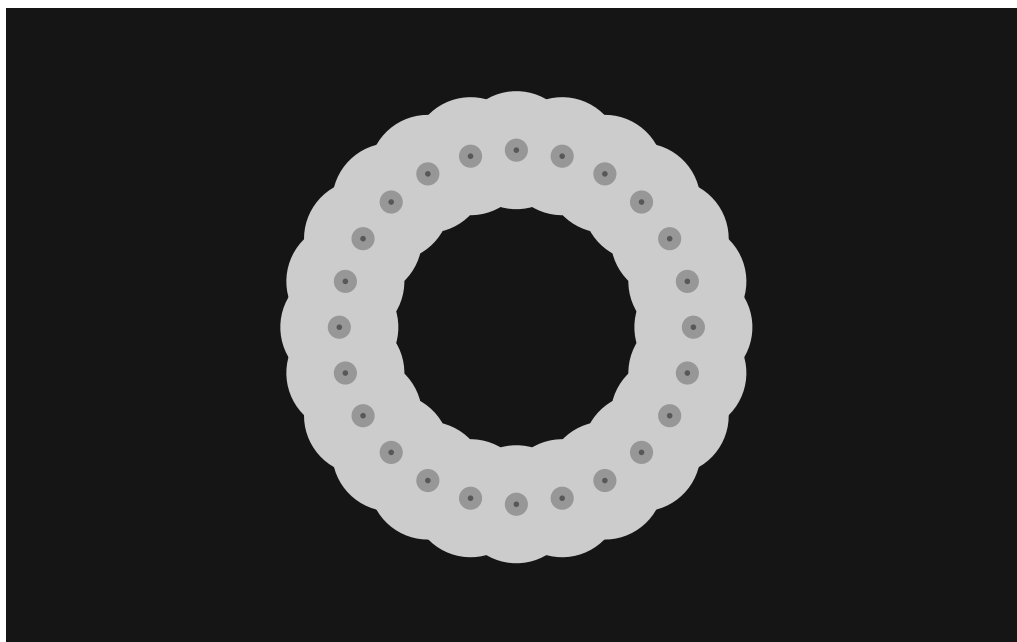


Figure 3.6 : Répartition des cibles en fonction du niveau de difficulté ; test multidirectionnel.



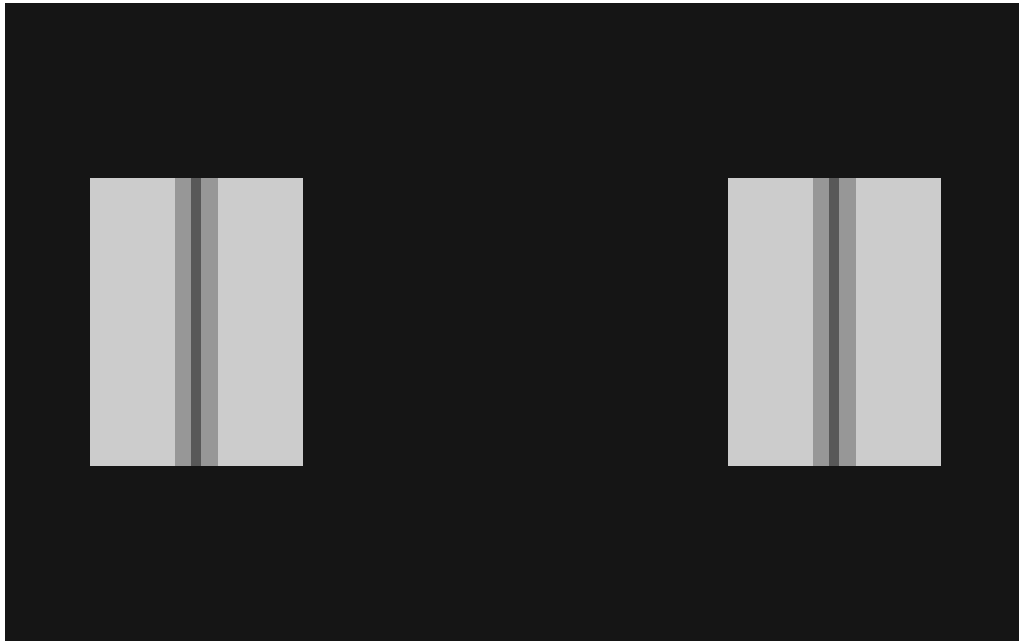


Figure 3.7 : Répartition des cibles en fonction du niveau de difficulté ; test unidirectionnel.

#### ORDRE DES SÉRIES

L'expérimentation a été menée en deux campagnes : campagne 1 : Écran mural, Tablette, Pavé tactile, Souris ; campagne 2 : Boule de commande militarisée, Boule de commande commerciale.

Total de séries par sujet : 6 dispositifs  $\times$  2 tâches  $\times$  3 niveaux de difficulté = 36 passages.

Afin de mitiger les effets de l'apprentissage, de la fatigue et du stress, l'ordre de passage des dispositifs, des tâches et des niveaux de difficulté a été alterné entre chaque série suivant trois carrés latins (voir tableau 3.4 : Composition des carrés latins).

#### CALCUL DES PERFORMANCES

Les formules utilisées pour le calcul des performances des participants ont été adaptées d'ISO 9241-9 : 2000. Il n'a pas été possible de respecter la norme à la lettre, car elle comporte des inconsistances présentées ci-dessous (équations 3, 3a, 3b).

Les unités de distances sont arbitraires, le débit ne dépendant que de dimensions relatives.

##### Largeur effective de la cible

Elle est calculée comme l'étendue de la distribution des coordonnées de frappe :

$$w_e = 4,133 s_x (1)$$

Où :

- $w_e$  représente la largeur effective de la cible,
- $s_x$  représente l'écart type des coordonnées de frappe, dans la direction du mouvement.

#### Indice de difficulté

Il mesure (en bits) la précision de l'utilisateur requise pour une tâche :

$$ID = \log_2\left(\frac{d+w}{w}\right) \quad (2)$$

Où :

- $d$  est la distance entre le point de départ et la cible,
- $w$  est la largeur de la cible affichée le long de l'axe d'approche.

#### Indice effectif de difficulté

Il mesure (en bits) la précision atteinte par l'utilisateur dans la réalisation d'une tâche. On utilise ici la formulation dite de Shannon (Shannon & Weaver, 1949), qui donne une meilleure corrélation des données (MacKenzie, 1989) et permet d'éviter les valeurs négatives du débit (MacKenzie & Buxton, 1993).

C'est pour le calcul de cette valeur que les valeurs de la littérature diffèrent de celles fournies par la norme. ISO 9241 fournit pour le calcul de l'indice de difficulté multidirectionnel :

$$ID_e = \log_2\left(\frac{d+w_e}{w_e}\right) \quad (3)$$

Parfois retrouvée dans la littérature (Murata, 1991), cette équation donne un  $ID_e$  (et par suite un débit) légèrement surévalué ; une rapide estimation à partir des résultats de pré-test a donné en moyenne :

$$\frac{ID_{e\text{ Norme}}}{ID_{e\text{ Fitts}}} = 1,025 \quad (3a)$$

D'autre part, la norme donne pour le calcul de l'indice de difficulté unidirectionnel :

$$ID_e = \frac{\log_2(d+w_e)}{w_e} \quad (3b)$$

Dans certaines configurations, cette équation peut donner des valeurs aberrantes pour  $ID_e$  (voir tableau 3.3) :

Tableau 3.4 : Valeurs de l' $ID_e$  en fonction de l'équation utilisée.

$d$	$w$	$ID_e$ Fitts (3d)	$ID_e$ Norme (3b)
200	30	2,939	0,262
400	30	3,841	0,292

Nous n'avons donc pas tenu compte de ces formules dans le calcul de l'indice de difficulté. On a à la place utilisé l'équation (3d) — expression originale de Shannon & Weaver, (1949) — à la fois pour les essais uni- et multidirectionnel :

$$ID_e = \log_2 \left( \frac{d_e + w_e}{w_e} \right) \quad (3d)$$

Où :

- $d_e$  est la distance effective parcourue par l'utilisateur d'une frappe à l'autre,
- $w_e$  est la largeur effective de la cible.

### Débit

Le débit mesure (en bits par seconde, unité arbitraire) la quantité d'informations qu'un dispositif peut passer par unité de temps. La mesure du débit de chaque dispositif permet de le comparer aux autres en terme de rapidité. Le débit est indépendant de toute considération de taille ou de résolution de l'écran.

$$Débit = \frac{ID_e}{t_m} \quad (4)$$

Où :

- $ID_e$  est l'indice effectif de difficulté, (combinant la distance entre les deux cibles et la taille de la cible),
- $t_m$  est la durée du mouvement entre le point de départ et la cible.

### Taux d'erreurs

Cette mesure ne fait pas partie de la définition proposée par la norme. Elle est toutefois nécessaire pour comparer les dispositifs entre eux en terme d'efficacité. Par série, le taux d'erreurs est calculé comme suit :

$$e\% = 100 \cdot \frac{\text{Nombre de frappes manquées}}{\text{Nombre total de frappes}} \quad (5)$$

L'efficacité d'un dispositif est définie comme l'inverse du taux d'erreurs moyen pour toutes les séries.

### Taux de satisfaction

La satisfaction de l'utilisateur a été évaluée par une série de 12 variables pouvant prendre 7 valeurs entre 1 et 7 (échelle de Likert). Ces variables sont les suivantes : force requise pour l'activation, souplesse d'utilisation (c'est à dire la capacité à s'adapter à différentes tâches), efforts requis pour l'utilisation, précision, vitesse d'utilisation, confort global, fonctionnement global du dispositif d'entrée, fatigue des doigts, fatigue du poignet, fatigue du bras, fatigue de l'épaule, fatigue du cou. Pour chacune des 12 variables, on identifie 1 comme note la plus basse et 7 comme note la plus haute. Des adjectifs qualitatifs adaptés à chaque variable sont utilisés sur le questionnaire de satisfaction.

### CHOIX DES INDICES DE DIFFICULTÉ

La largeur et l'éloignement relatif des cibles sont choisis afin de présenter trois niveaux de difficulté (*ID*) échelonnés à 2, 4, et 6 bits (équation (3d), voir tableau 3.3 pour les distances et tailles de cibles correspondantes), dans l'intervalle prescrit par la norme ( $1 < ID < 6$ ). Afin de ne pas introduire de variations liées à une amplitude de mouvements variable, le degré de difficulté est varié uniquement en jouant sur la taille des cibles  $w$ , la distance  $d$  entre les cibles restant fixe. L'indice de difficulté étant fonction de la taille des cibles ainsi que de leur éloignement, on a joué sur ce premier paramètre pour conserver un *ID* compris entre 1 et 6. Les résultats sont donc comparables avec ceux des études se situant dans le même intervalle de l'*ID*.

On doit ici noter que ce choix particulier pour  $d$  ( $d$  fixe pour tous les niveaux de difficulté) va entraîner — à taux d'erreur et durée de mouvement égaux — une augmentation significative du débit relevé au fur et à mesure que l'indice de difficulté augmente. Ce résultat contre-intuitif s'explique par le fait que la cible se faisant de plus en plus petite, on a tendance à frapper de plus en plus proche du centre réel.

Cette amélioration artificielle du débit avec la difficulté sera négligée. En effet, l'expérience portant sur la comparaison de deux dispositifs et non sur l'impact de la variation de l'indice de difficulté, on aura soin de comparer les résultats uniquement entre dispositifs à indice de difficulté constant.

L'indice de difficulté est identique quel que soit le dispositif d'affichage, car à  $d$  et  $w$  constants,  $ID$  et  $IDe$  ne dépendent pas de la définition de l'écran (équations 2 et 4).

D'après Gillan et al. (1990), pour  $w = 1$  constant, une hauteur  $h$  des cibles supérieure à 2 n'a pas d'influence notable sur le temps de pointage. Le temps de pointage augmente pour  $h < 2$ . On choisit pour le test unidirectionnel une hauteur de cible constante égale à 400 pixels.

## VARIABLES

La variable indépendante est le type de dispositif d'interaction utilisé pour agir sur l'écran. Six dispositifs d'interaction ont été comparés:

1. D1 : Commande tactile directe sur le GETV,
2. D2 : Tablette répétant l'affichage du GETV,
3. D3 : Pavé tactile (trackpad),
4. D4 : Souris,
5. D5 : Boule de commande militarisée ("durcie") (trackball durcie),
6. D6 : Boule de commande commerciale (trackball commerciale).

Les trois variables dépendantes sont les suivantes :

- Efficience mesurée en débit, (tel que défini plus haut),
- Efficacité mesurée en pourcentage de frappes réussies (comptage du nombre de frappes dans la cible),

Ces deux variables seront relevées directement par le code exécutant le test.

- Satisfaction de l'utilisateur évaluée par questionnaire a posteriori.

## 3.4 Procédure

Les phrases entre guillemets sont données telles quelles à l'oral.

### ACCUEIL

L'accueil dure environ 10 minutes.

Accueil du sujet, présentation du projet de recherche. On fait lire et signer au sujet le formulaire d'information et de consentement. Ensuite, il complète un bref questionnaire de données biographiques : âge, sexe, taille, main dominante, expérience avec le GETV, expérience avec les interfaces tactiles en général, les pavés tactiles, la souris. On répond aux éventuelles questions.

### INSTRUCTIONS GÉNÉRALES

"Vous pouvez choisir le pseudonyme qui vous convient, mais il est important de conserver le même surnom au cours de l'expérience. Ce pseudonyme sera utilisé plusieurs fois, il est donc judicieux de rester concis."

"Les cibles s'affichent en clair sur un fond sombre. Utilisez un seul doigt pour atteindre chaque cible. Ne changez pas de main au cours de l'expérience."

"Vous devrez exécuter le test aussi rapidement que possible, mais en faisant le moins d'erreurs possible."

"Ne tentez pas de corriger les erreurs : si vous avez manqué votre cible, frappez immédiatement la cible suivante."

L'expérimentateur vérifie la compréhension et répond aux éventuelles questions.

### DISPOSITIFS

"Placez-vous debout devant l'écran / la tablette / le pavé tactile / la souris / la boule de commande, dans une position où vous jugez confortable l'atteinte de toutes les zones de l'écran. Vous êtes libre de vous déplacer au cours du test."

L'expérimentateur vérifie la compréhension et répond aux éventuelles questions.

### ENTRAÎNEMENT

L'entraînement ne dure pas plus de 10 minutes pour l'ensemble des dispositifs.

Il s'effectue au cours du test, à mesure que le sujet se voit présenter les différents dispositifs. Il s'effectue toujours dans la condition C2 N2 (tâche 'C' unidirectionnelle, niveau de difficulté 'N' intermédiaire, pour ne pas perturber le participant) pour le dispositif concerné. On ne propose pas d'entraînement pour la souris. Chaque participant choisit le nombre d'entraînements qu'il décide d'effectuer avant de passer aux séries réelles. Pour chaque dispositif, la série réelle est exécutée immédiatement après la fin de l'entraînement.

L'expérimentateur vérifie la compréhension et répond aux éventuelles questions.

## EXÉCUTION DE LA TÂCHE

La durée est d'environ 30 minutes.

Le sujet exécute les tâches C1 et C2 sur les dispositifs D1 et D2, ventilé par niveaux de difficulté N1, N2, N3, dans l'ordre déterminé par trois carrés latins (voir tableau 3.4).

Tableau 3.5 : Composition des carrés latins

Dispositifs				Tâches		Niveau de difficulté		
D1	D2	D3	D4	C1	C2	N1	N2	N3
D2	D3	D4	D1	C2	C1	N2	N3	N1
D3	D4	D1	D2			N3	N1	N2
D4	D1	D2	D3					

Le premier participant exécute : D1 C1 N1 ; D1 C1 N2 ; D1 C1 N3 ; D1 C2 N1 ; etc.

Le second participant exécute : D1 C1 N2 ; D1 C1 N3 ; D1 C1 N1 ; D1 C2 N2 ; etc.

Le quatrième participant exécute : D1 C2 N1 ; D1 C2 N2 ; D1 C2 N3 ; D1 C1 N1 ; etc.

Le septième participant exécute : D2 C1 N1 ; D2 C1 N2 ; D2 C1 N3 ; D2 C2 N1 ; etc.

Pour la première campagne (dispositifs D1 à D4), chaque participant effectue un total de : 4 dispositifs \* 2 tâches \* 3 niveaux de difficulté = 24 passages. Pour la deuxième campagne (dispositifs D5 et D6), cette matrice a été réutilisée en remplaçant les dispositifs 1 et 2 par les dispositifs 5 et 6. On a donc 2 dispositifs \* 2 tâches \* 3 niveaux de difficulté = 12 passages par participant.

La matrice de passage pour chaque participant (première campagne uniquement) est donnée à l'annexe G de ce document.

## QUESTIONNAIRES

La complétion des questionnaires dure environ 10 minutes.

A l'issue de chaque groupe de six séries, le sujet complète le questionnaire de satisfaction relatif au dispositif de pointage qu'il vient de tester. Un spécimen de questionnaire de satisfaction est fourni à l'annexe E. À la fin du test, on demande au sujet d'établir deux classements des dispositifs, le premier différencié par tâches, et le second global. La durée totale du test est de 60 minutes maximum pour la première campagne, 30 minutes maximum pour la seconde.

### 3.5 Déroutement des pré-tests

Une campagne de pré-tests préliminaire à la première campagne s'est tenue dès que le site de test a permis l'accueil des sujets. Cette campagne n'a concerné que trois sujets (qui n'ont pas participé aux campagnes réelles). N'ont été évalués que les dispositifs D1, D2 et D3 (D3 testé seulement avec le pavé tactile intégré de l'ordinateur). Le but principal des pré-tests de cette préparation a été de valider les aspects techniques et organisationnels de l'expérimentation, ainsi que de préparer au mieux l'expérimentateur pour une exécution fluide des séquences de test. Une seconde campagne préliminaire a été menée avant la seconde campagne réelle. Elle a été limitée aux dispositifs D5 et D6, et n'a concerné qu'un seul participant.

Ces campagnes de pré-test ont permis d'intégrer deux changements au protocole :

- Pour le dispositif tablette tactile (D2), il était originellement prévu de donner le choix suivant au participant : soit utiliser la tablette comme simple dispositif de pointage et regarder le grand écran, soit utiliser seulement la tablette. Les participants ayant choisi la première option se sont trouvés trop perturbés par ce mode de fonctionnement. Les participants reportaient constamment leur vision sur la tablette afin de vérifier la position de leur doigt par rapport à la cible, conduisant à des temps de complétion de tâche exagérément longs. Pour l'expérimentation réelle, on a donc retiré le choix et demandé aux participants d'utiliser la tablette seule. Plutôt que de retirer complètement la tablette de l'expérimentation, on a choisi de la conserver parmi les dispositifs évalués, en tant que tablette simple.
- On a également pris la décision de remplacer le pavé tactile intégré (à l'ordinateur de test) de D3 par un pavé tactile sans fil, la solution "intégrée" amenant instinctivement les participants à chercher l'image sur l'écran de l'ordinateur, celui-ci étant pourtant désactivé.

Dans ce chapitre, on a présenté les données de l'enquête biographique réalisée, ainsi que le choix de la configuration et de la procédure retenus pour le test, notamment l'adaptation qui a pu être faite de la norme ISO 9241 aux particularités de notre hypothèse. Au chapitre suivant, on a présenté les résultats détaillés du test, ainsi que les analyses qui ont été menées afin de déterminer leur signification. Ces résultats sont discutés au regard de ceux obtenus par les chercheurs ayant précédé ce travail.



## CHAPITRE 4 : PRÉSENTATION DES RÉSULTATS

Ce chapitre présente les résultats de performance pour les six dispositifs. Il comprend six sections. La première présente les données pour les variables de débit et de taux d'erreurs pour tous les dispositifs. La deuxième détaille les données de débit et de taux d'erreur pour chaque dispositif. La troisième présente les données de satisfaction utilisateur pour tous les dispositifs. La quatrième présente les résultats des analyses statistiques effectuées sur les données de débit et de taux d'erreurs. La cinquième discute les résultats de l'expérimentation au regard d'expériences comparables. Toutes les analyses statistiques effectuées ont été réalisées avec SPSS Statistics.

### 4.1 Débit et taux d'erreurs pour l'ensemble des dispositifs

Afin de faciliter la lecture, les données relatives au taux de satisfaction sont regroupées à la fin de document et traitées à part. La partie automatisée du test comporte un total de 807 séries observées. En excluant les entraînements, on obtient un total de 691 séries réelles. On a choisi pour ces deux variables une mesure de la tendance centrale par la moyenne, plus adaptée à la comparaison de performances entre les dispositifs. Les figures 4.1 et 4.2 présentent la répartition des données issues de la partie automatisée du test :

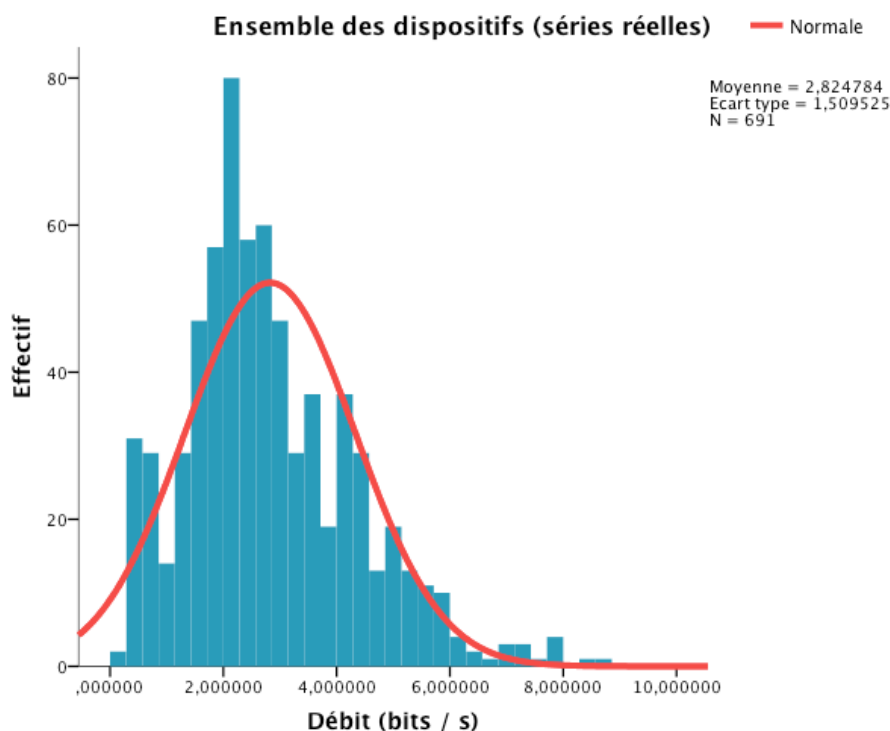


Figure 4.1 : Répartition des données de débit.

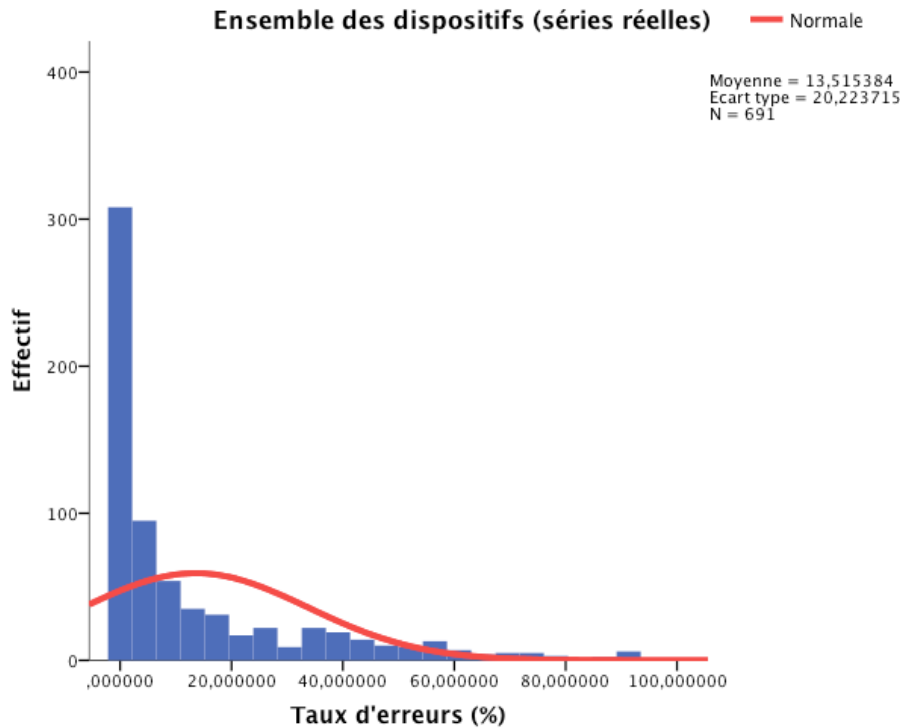


Figure 4.2 : Répartition des données de taux d'erreurs.

Le débit est une variable continue de type échelle de rapport, généralement comprise entre 0 et 8 bits par seconde. Elle est donc représentée sous forme d'histogramme. Pour le débit, on cherche à obtenir une moyenne la plus élevée possible (cela signifie que le dispositif est rapide). On observe sur la figure 4.1 un groupement des données entre 0 et 6 bits / s.

Le taux d'erreurs est une variable discrète considérée comme une variable de type intervalle. Une série étant constituée de 24 frappes de cibles, le taux d'erreur d'une série peut prendre 24 valeurs entre 0 et 100 %. On a volontairement choisi de représenter le taux d'erreur sous forme d'histogramme comportant 24 intervalles de classe d'étendue égale. Pour le taux d'erreurs, on cherche à obtenir une moyenne la plus basse possible (cela signifie que le dispositif est précis). On observe naturellement un fort aplatissement et une forte asymétrie (groupement des valeurs vers 0 %). Ne cherchant pas à utiliser nos résultats pour établir des prédictions, on ne réalisera pas de standardisation vers la loi normale.

Pour la suite du document, on abrège les termes Dispositif, Type de tâche et Niveau de difficulté par D, T et N, respectivement.

## 4.2 Données de débit et de taux d'erreur par dispositif

Dans le tableau 4.1 ci-dessous, pour le débit, une différence négative par rapport à la moyenne de l'ensemble des séries indique une performance inférieure à la moyenne. Pour le taux d'erreurs, une différence négative par rapport à la moyenne de l'ensemble des séries signale une performance supérieure à la moyenne.

Tableau 4.1 : Résultats de débit et de taux d'erreur par dispositif.

Dispositif	Données de débit			Données de taux d'erreur		
	Débit moy. (Ecart-type)	Différence avec la moy. des séries	Rang	Taux d'erreur moy. (Ecart-type)	Différence avec la moy. des séries	Rang
D1 : GETV	2,31 bits/s (1,38)	- 0,52 bits/s	3ème	17,62 % (21,98)	+ 4,10 %	5ème
D2 : Tablette	4,66 bits/s (1,59)	+ 1,83 bits/s	1er	23,25 % (30,80)	+ 9,73 %	6ème
D3 : Pavé tactile	2,01 bits/s (0,87)	- 0,82 bits/s	5ème	10,51 % (13,74)	- 3,01 %	4ème
D4 : Souris	3,60 bits/s (0,96)	+ 0,77 bits/s	2ème	8,73 % (14,27)	- 4,79 %	1er
D5 : Boule mil.	1,97 bits/s (0,77)	- 0,86 bits/s	6ème	10,20 % (15,04)	- 3,32 %	3ème
D6 : Boule com.	2,14 bits/s (0,73)	- 0,69 bits/s	4ème	9,84 % (13,75)	- 3,68 %	2ème

### DÉBITS

Un dispositif qui obtient un débit faible est perçu comme lent, car il possède une inertie qui est généralement due au principe d'action (friction du doigt avec D3, inertie de la boule avec D5 et D6, bras en extension avec D1).

La tablette se classe clairement en première place en terme de débit moyen (voir figures 4.3 et 4.4). Cette performance est probablement due au fait que ce dispositif ne requérait pas la consultation de l'écran à distance. Cependant, l'écart-type le plus élevé de la sélection indique une hétérogénéité des résultats. Le débit élevé de la souris peut s'expliquer par le fait que les sujets possèdent une longue expérience avec ce pointeur. Les dispositifs tactiles D1 et D2 présentent tous deux une forte dispersion des résultats (voir figure 4.3). Ceci s'explique en partie par des niveaux d'expérience très différents entre les sujets.

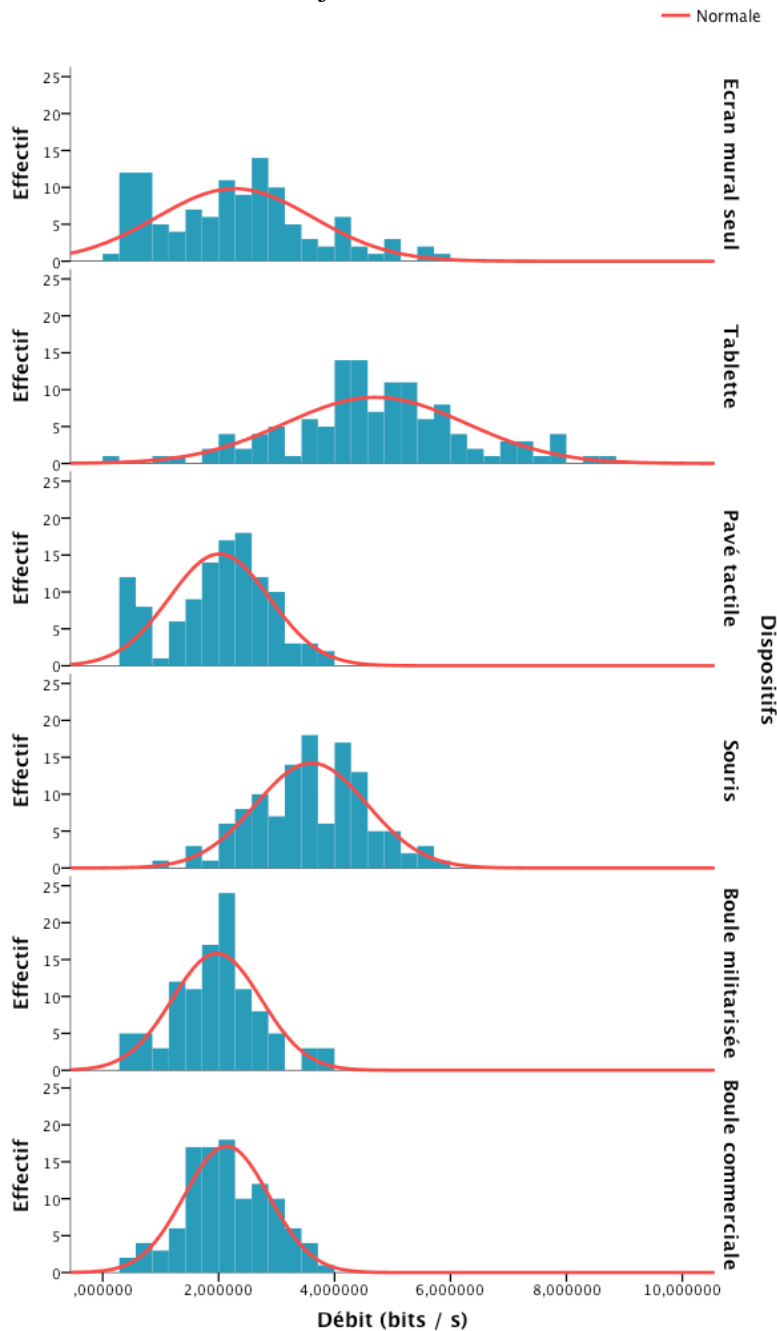


Figure 4.3 : Répartition des données de débit par dispositif

Les dispositifs de type boule de commande (D5 et D6) se caractérisent par un débit moyen peu élevé et une dispersion faible. On peut expliquer ceci par deux phénomènes :

- un manque d'expérience plus marqué avec ce type de dispositif,
- une inertie physique plus importante, imposant un geste plus lent.

On constate sur la figure 4.3 que les dispositifs D1, D3 et D5 présentent tous un groupement de données important entre 0 et 1 bit. Cette répartition des points de donnée permet de penser que ces dispositifs imposent un apprentissage plus long, car de nombreux participants inexpérimentés obtiennent des résultats très faibles.

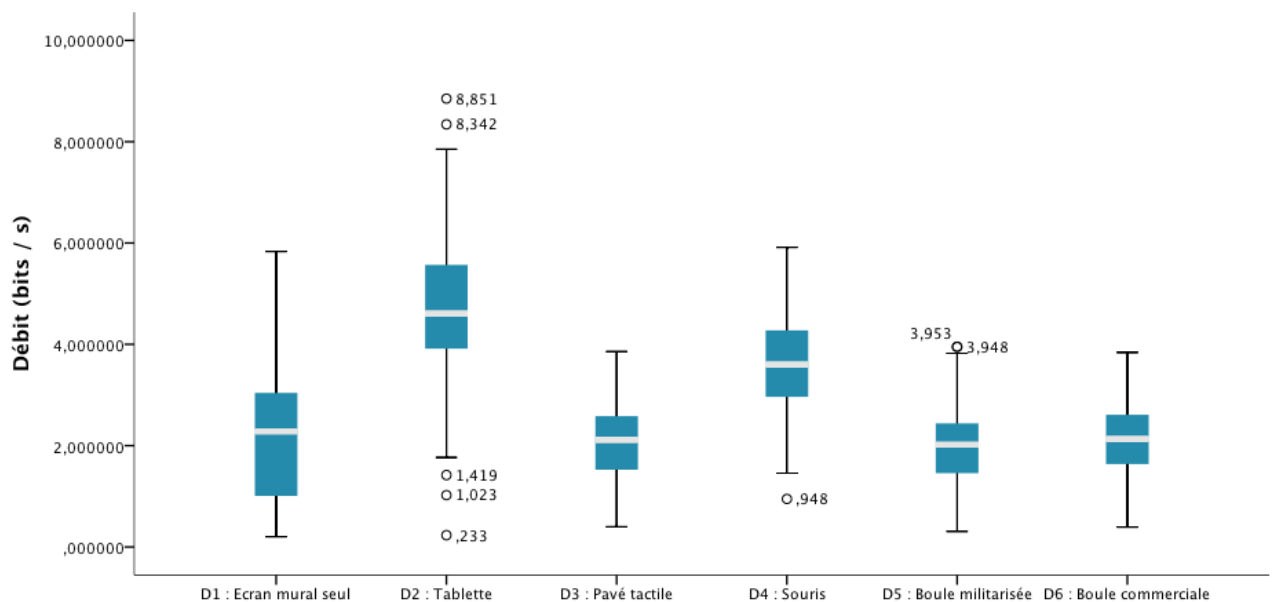


Figure 4.4 : Répartition des valeurs de débit par dispositif.

On observe dans la figure 4.4 la présence de valeurs de débit dispersées (précédées par un cercle) plus nombreuses pour D2, ce qui est cohérent avec l'écart type relevé pour ce dispositif. D3, D4, D5, et D6 présentent des étendues plus resserrées, d'où des valeurs de débit plus homogènes. Du point de vue de l'utilisateur, ceci implique un comportement du dispositif plus facilement prévisible.

#### TAUX D'ERREURS

Un dispositif qui obtient un fort taux d'erreur est considéré peu fiable et peut être dangereux dans les environnements critiques (applications industrielles, militaires, dispositifs d'urgence). De façon cohérente avec l'écart-type de la variable débit, la tablette présente le plus fort taux d'erreurs de la sélection (voir figures 4.5 et 4.6). Le faible taux d'erreur de la souris peut s'expliquer par le fait que

tous les sujets possèdent une longue expérience avec ce pointeur. Les dispositifs tactiles (D1 et D2) présentent tous deux une forte dispersion des résultats ( $30,80 > S > 21,98$ ). Ceci peut être dû à de mauvaises performances pour la sélection de cibles de petite taille. En effet, avec une interface tactile, la taille du pointeur est celle du doigt de l'opérateur. Les autres dispositifs (D3, D4, D5, et D6) voient s'afficher à l'écran un curseur d'une résolution égale à un pixel.

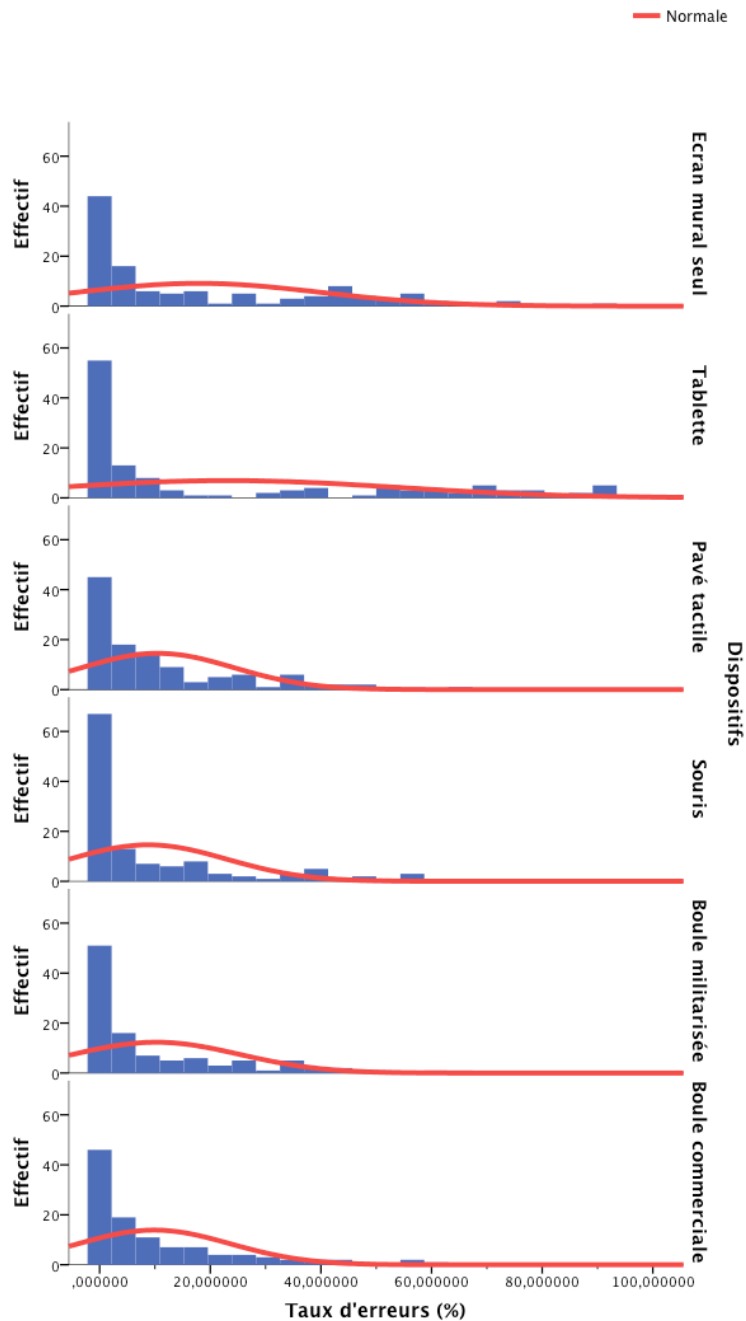


Figure 4.5 : Répartition des données de taux d'erreurs par dispositif.

Certains de ces dispositifs cumulent d'autres avantages :

- La souris (D4) est le dispositif auquel les sujets sont très habitués ;
- Les dispositifs de type boule de commande (D5 et D6) se caractérisent par une inertie physique plus importante, imposant un geste plus lent, ce qui conduit à avoir moins de frappes hors cibles.

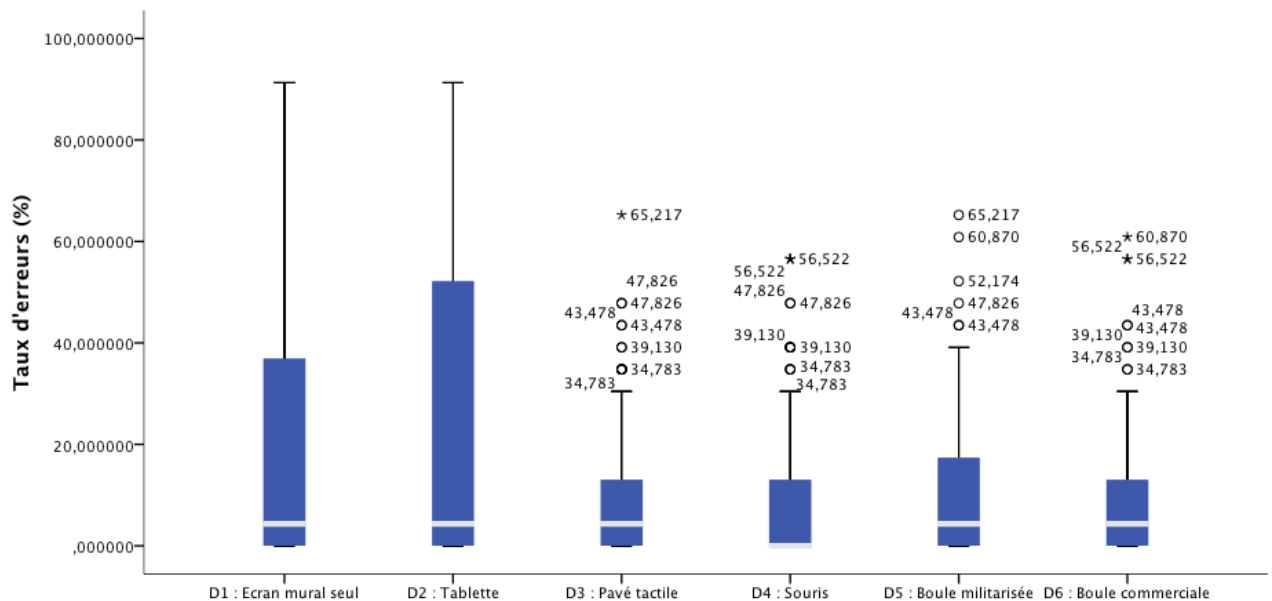


Figure 4.6 : Répartition des valeurs de taux d'erreurs par dispositif.

On constate ici que la majorité des dispositifs présentent des valeurs très dispersées (précédées par un cercle) voire extrêmes (précédées par une astérisque), sauf D1 et D2. Ces valeurs proviennent majoritairement des séries réalisées avec une difficulté élevée.

Ce qui dans l'absolu pourrait être interprété positivement signifie en fait que D1 et D2 présentent un grand nombre de séries avec un très fort taux d'erreurs (au-dessus de 40 % de cibles manquées dans la série).

La médiane de D4 est égale à 0, ce qui nous indique que plus de 50 % des réponses ont un taux d'erreurs nul, un excellent résultat.

### 4.3 Niveau de satisfaction de l'utilisateur

La satisfaction de l'utilisateur a été évaluée à travers une série de 12 variables discrètes considérées comme des variables de type intervalle, pouvant prendre sept valeurs entre 1 et 7 (échelle de Likert). Ces variables sont les suivantes : force requise pour l'activation, souplesse d'utilisation, efforts requis pour l'utilisation, précision, vitesse d'utilisation, confort global, fonctionnement global du dispositif d'entrée, fatigue des doigts, fatigue du poignet, fatigue du bras, fatigue de l'épaule, fatigue du cou. On a volontairement choisi de représenter les niveaux de satisfaction sous forme de diagramme bâton. Compte tenu du nombre de réponses pour chaque dispositif (N = 18, suite à l'exclusion des deux participants n'ayant pu être disponibles pour la seconde campagne), on a limité l'analyse aux moyennes obtenues par dispositif et par dimension.

Pour toutes les dimensions mesurées, 1 indique toujours la plus mauvaise note, et 7 la meilleure note. Notamment, pour les résultats de fatigue, une note proche de 1 indique que le dispositif entraîne une grande fatigue dans l'articulation concernée. De même pour les échelles de force ou d'effort, 1 indique qu'une grande force doit être employée, et 7 une force faible.

La figure 4.7 ci dessous montre les scores relatifs des différents dispositifs.

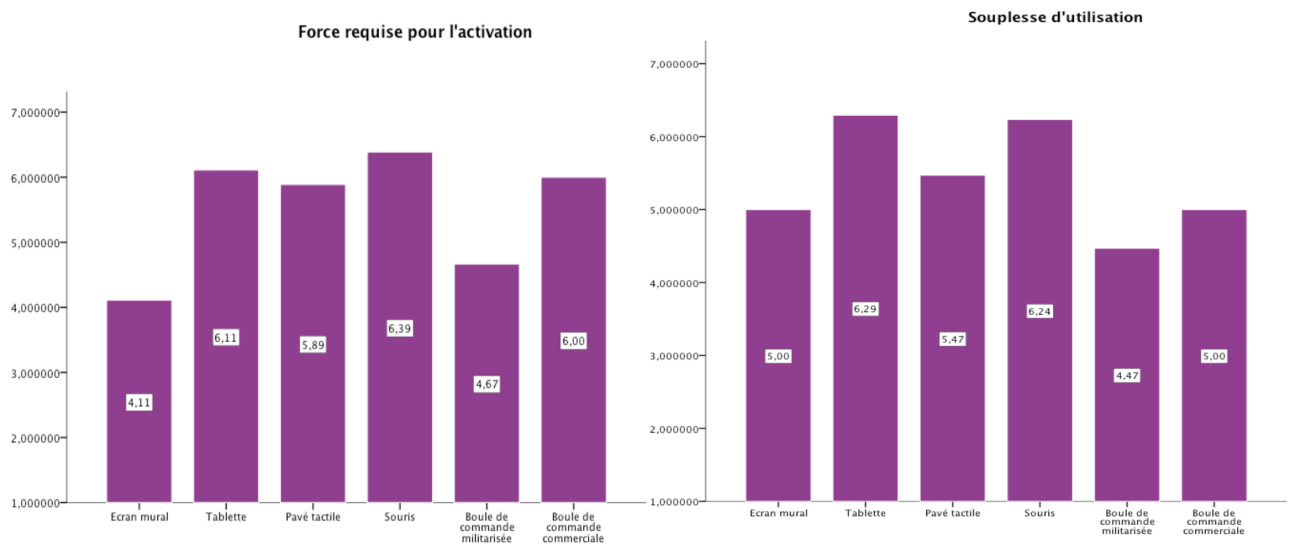


Figure 4.7 : Données de satisfaction utilisateur.



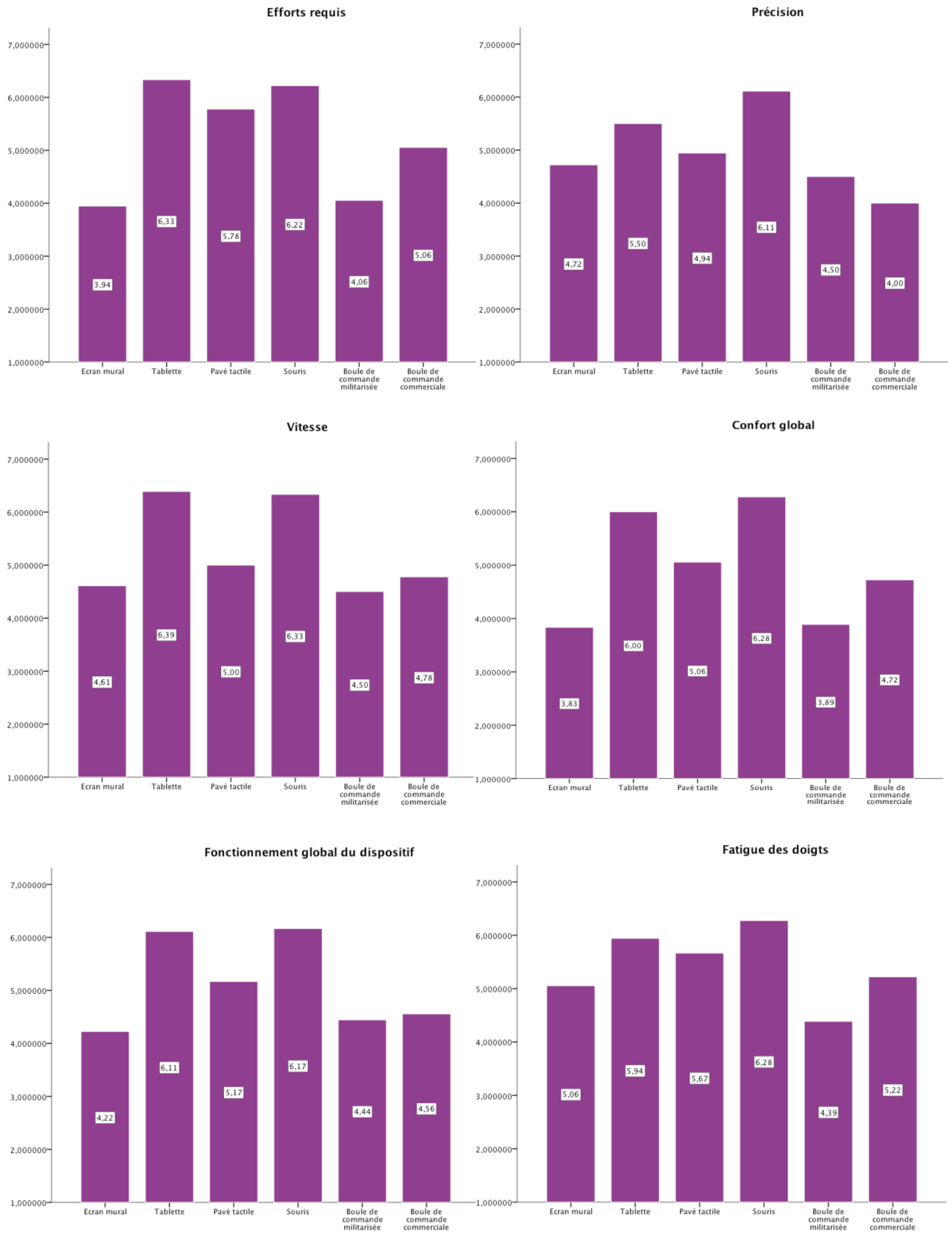


Figure 4.7 : Données de satisfaction utilisateur (suite).

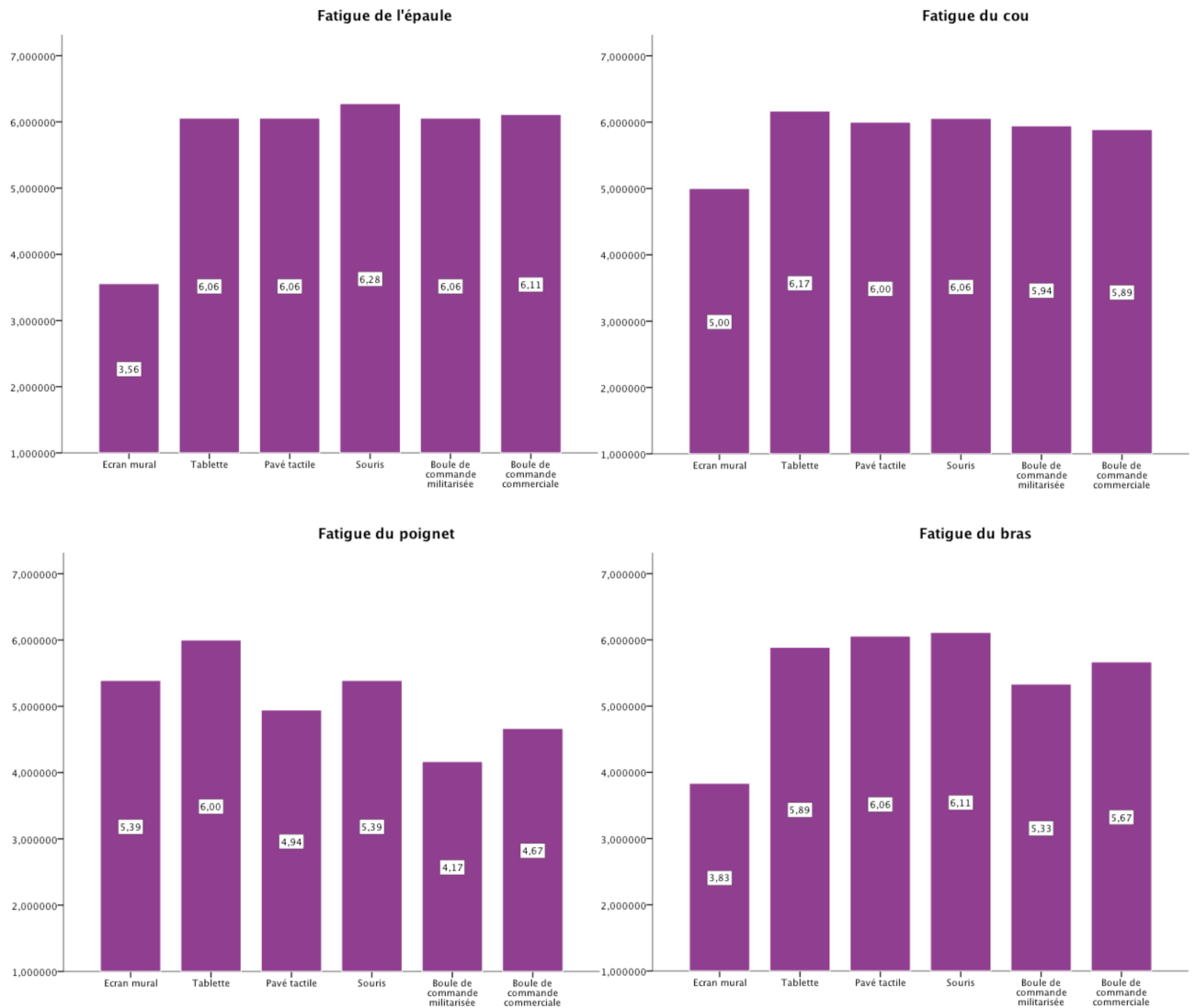


Figure 4.7 : Données de satisfaction utilisateur (suite et fin).

On observe que le GETV est pratiquement toujours classé dernier pour ce qui est du niveau de satisfaction de l'utilisateur, notamment pour ce qui concerne les dimensions de fatigue. La fatigue provoquée par ce dispositif dans le bras, l'épaule et le cou fait ressortir la supériorité des dispositifs de pilotage en position assise en terme de fatigue perçue par l'utilisateur. À l'inverse, les dispositifs Souris (D4) et Tablette (D2) sont pratiquement toujours classés premiers pour la satisfaction de l'utilisateur. Étonnamment, la tablette (D2) se classe seconde en précision perçue, avec une bonne note de 5,5 / 7, alors que ce dispositif obtient le taux d'erreur le plus élevé de la sélection. Ce résultat soulève la question d'un manque de retour sensitif (feedback) approprié sur la tablette tactile. En revanche, les résultats en vitesse perçue recourent ceux relevés par le code, pour l'ensemble des dispositifs.

## SYNTHÈSE

On rassemble ici les moyennes des performances par dispositif :

Tableau 4.2 : Performances par dispositif.

	Débit moyen (bits / s)	Taux d'erreurs moyen (%)	Satisfaction utilisateur amalgamée (Lickert de 1 à 7)
D1 : GETV	2,31	17,62	4,44
D2 : Tablette	4,66	23,25	6,07
D3 : Pavé tactile	2,01	10,51	5,50
D4 : Souris	3,60	8,73	6,16
D5 : Boule mil.	1,97	10,20	4,70
D6 : Boule com.	2,14	9,84	5,14

### 4.4 Analyses effectuées

Afin de déterminer la pertinence de ces résultats et l'existence de différences significatives, nous avons procédé à une série d'analyses de variance pour mesures répétées :

1. Trois analyses ventilées par niveau de difficulté pour la variable débit,
2. Une analyse unique pour la variable taux d'erreurs.

#### OBJECTIF

L'objectif est de comparer les six dispositifs en termes de débit et de taux d'erreurs au regard des six conditions expérimentales testées, afin de déterminer s'il existe des différences significatives entre les dispositifs pour ces deux variables.

#### TYPE DE TEST

Les deux variables examinées sont considérées comme continues. On retient pour l'analyse un échantillon de 108 séries, correspondant aux combinaisons des six conditions expérimentales et des 18 sujets, répétées pour chacun des six dispositifs, pour un total théorique de 648 séries (on rappelle qu'une série est la complétion du test par un participant sous une condition expérimentale définie par le triplet dispositif / type de test / niveau de difficulté). Afin d'atteindre l'objectif visé, on choisit

donc de pratiquer pour chaque variable un test d'analyse de variance pour mesures répétées. Le niveau de signification retenu est le niveau standard avec  $\alpha = 0,05$ .

Le test de sphéricité de Mauchly s'avérant négatif pour certains tests de l'analyse de variance, on a retenu pour chaque analyse le critère de la borne inférieure, donnant les valeurs de p les plus conservatrices. Ce choix explique que la somme des carrés et la moyenne des carrés ont des valeurs égales.

## PRÉ-TRAITEMENT DES DONNÉES

On exclut les données d'entraînement de l'analyse.

On exclut les données d'entraînement de l'analyse.

Un total de 20 personnes ont participé à la campagne 1 (excepté D3, 19 participants seulement suite à une défaillance technique) et 18 personnes ont participé à la campagne 2. Les deux campagnes ont fait appel aux mêmes sujets. L'analyse de variance pour mesures répétées ne tolérant pas les données manquantes, on procède à une réorganisation des données :

- On considère seulement les données issues des 18 sujets ayant participé aux deux campagnes. Le nombre total de séries considérées passe de 691 à 648, soit une suppression de 6,223 % des séries.
- Les données manquantes restantes ( $N = 8$ , soit 1,24 %) sont remplacées par la moyenne des valeurs pour la condition expérimentale correspondante. On choisit le remplacement par la moyenne des valeurs car c'est celle qui est la plus "neutre" vis-à-vis de la signification des données.

Le remplacement par la moyenne ou la médiane des points voisins, ou bien l'interpolation linéaire suppose toutes deux que la performance d'un sujet dans une condition expérimentale soit liée à la performance des autres sujets dans la même condition, ce qui n'est pas le cas. En effet, l'ordre de passage des sujets sur chaque dispositif a été organisé suivant un carré latin.

Afin de vérifier l'innocuité du pré-traitement sur la distribution des données, on réalise deux analyses descriptives (une pour chaque variable) comparant la distribution des données avant suppression des données manquantes à la distribution des données après suppression. La comparaison des deux ensembles de données par tests de Student n'est ici pas justifiée, du fait du faible impact de la modification :

- Les sujets n'ayant pas complété les 36 séries et le reste de la population ne présentent pas de différence majeure (au vu de l'enquête biographique ou de leurs autres résultats),

- Retrait d'une faible proportion du total des séries.

#### 4.4.1 Analyse de variance pour mesures répétées : débits

Des restrictions liées à la méthodologie s'appliquent à cette analyse de variance : en raison des choix faits lors de la conception du test, on doit procéder à trois analyses en fonction des niveaux de difficulté. En effet, la formule de calcul du débit prend en compte la distance de la frappe par rapport au centre de la cible. Or la distance inter-cibles a été volontairement gardée constante pour imposer un mouvement de grande ampleur entre chaque frappe. On va donc procéder à trois analyses, soit une par niveau de difficulté (Niveau 1 : 2 bits, niveau 2 : 4 bits, niveau 3 : 6 bits).

##### HYPOTHÈSES STATISTIQUES

Pour chacun des trois niveaux de difficulté  $n$  ( $1 \leq n \leq 3$ ), on établit les hypothèses statistiques suivantes :

$H_{01-n}$  : Il n'existe pas de différence significative entre les moyennes de débits de chacun des six dispositifs.

$H_{02-n}$  : Il n'existe pas de différence significative entre les moyennes de débits de chacune des deux tâches.

$H_{03-n}$  : Il n'existe pas de différence significative entre les moyennes de débits de chacun des 12 couples dispositif / tâche.

$H_{11-n}$  : Au moins une des moyennes de débits par dispositifs est significativement différente des cinq autres.

$H_{12-n}$  : Les deux moyennes de débit par tâches sont significativement différentes.

$H_{13-n}$  : Au moins une des moyennes de débits par couples dispositif / tâches est significativement différente des 11 autres.

On présente les résultats ci-dessous, ventilés par niveaux de difficulté.

## RÉSULTATS NIVEAU 1

Tableau 4.3 : Analyse de variance pour le débit des dispositifs, niveau de difficulté 1.

Tests des effets intra-sujets (borne inférieure)	Somme des carrés	Moyenne des carrés	F	p
Dispositifs	217,123	217,123	67,115	< 0,000
Erreur(Dispositifs)	54,997	3,235		
Tâches	1,510E-06	1,510E-06	0,000	0,998
Erreur(Tâches)	5,716	0,336		
Dispositifs * Tâches	3,742	3,742	2,155	0,160
Erreur(Dispositifs * Tâches)	29,518	1,736		

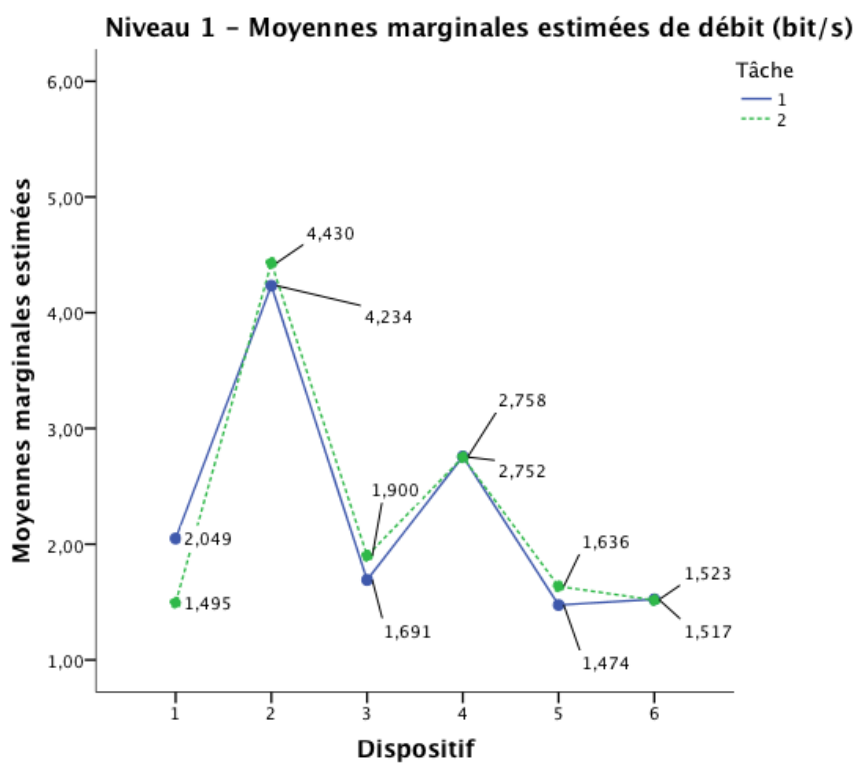


Figure 4.8 : Moyennes de débit par dispositif, niveau de difficulté 1.

On observe que la probabilité statistique (signification) associée à l'analyse de variance est inférieure à 0,05 seulement pour l'effet induit par les dispositifs.

## INTERPRÉTATION

Pour le niveau de difficulté 1 (2 bits) :

H<sub>01-1</sub> "Il n'existe pas de différence significative entre les moyennes de débits de chacun des six dispositifs." est rejetée. Au moins un des dispositifs obtient une moyenne combinée significativement différente des cinq autres.

H<sub>02-1</sub> "Il n'existe pas de différence significative entre les moyennes de débits de chacune des deux tâches." n'est pas rejetée. Le type de tâche (unidirectionnel ou bidirectionnel) n'a pas d'influence significative sur les résultats moyens d'un dispositif.

H<sub>03-1</sub> "Il n'existe pas de différence significative entre les moyennes de débits de chacun des douze couples dispositif / tâche." n'est pas rejetée. Pour un dispositif donné, le type de tâche n'a pas d'influence significative sur les résultats de ce dispositif.

Cependant, compte tenu des faibles puissances observées pour les tests Tâche et Dispositif \* Tâche (0,05 et 0,283 respectivement), on ne peut retenir les conclusions issues de ces tests. En effet, la probabilité de commettre une erreur de type II (faux négatif) est de 0,95 et 0,717, respectivement. On peut seulement conclure sur le fait qu'un ou plusieurs des dispositifs présentent une moyenne significativement différente des autres.

## RÉSULTATS NIVEAU 2

Tableau 4.4 : Analyse de variance pour le débit des dispositifs, niveau de difficulté 2.

Tests des effets intra-sujets (borne inférieure)	Somme des carrés	Moyenne des carrés	F	p
Dispositifs	272,306	272,306	47,835	< 0,000
Erreur(Dispositifs)	96,774	5,693		
Tâches	24,238	24,238	64,752	< 0,000
Erreur(Tâches)	6,363	0,374		
Dispositifs * Tâches	21,788	21,788	21,788	0,003
Erreur(Dispositifs * Tâches)	29,782	1,752		

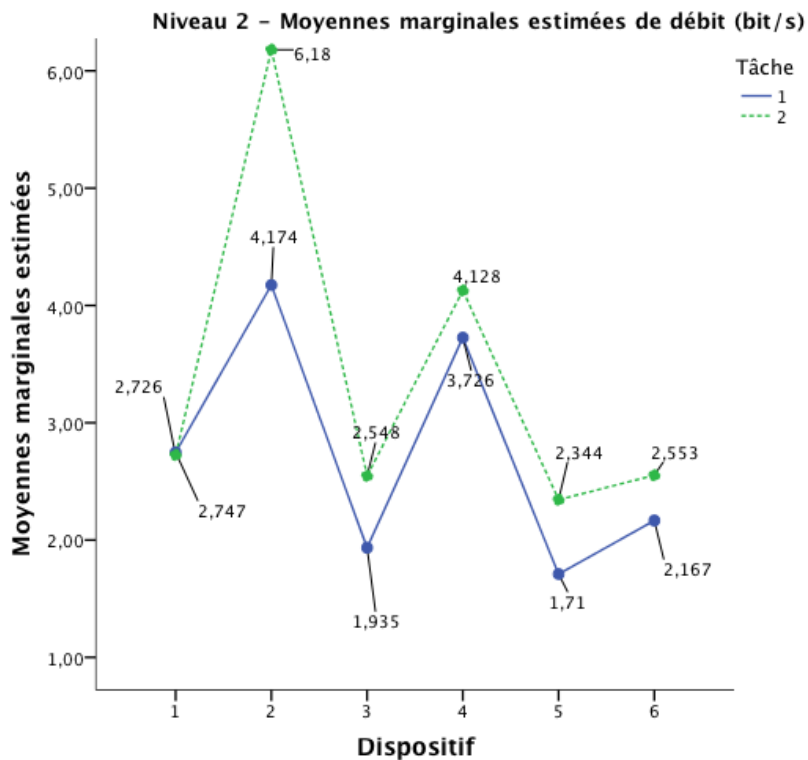


Figure 4.9 : Moyennes de débit par dispositif, niveau de difficulté 2.

On observe que la probabilité statistique (signification) associée à l'analyse de variance est inférieure à 0,05 pour l'effet induit par les dispositifs, l'effet induit par les tâches, et la combinaison de ces deux effets.

### INTERPRÉTATION

Pour le niveau de difficulté 2 (4 bits) :

$H_{01-2}$  "Il n'existe pas de différence significative entre les moyennes de débits de chacun des six dispositifs." est rejetée. Au moins un des dispositifs obtient une moyenne combinée significativement différente des cinq autres.

$H_{02-2}$  "Il n'existe pas de différence significative entre les moyennes de débits de chacune des deux tâches." est rejetée. Le type de tâche a une influence significative sur les résultats moyens d'un dispositif.

$H_{03-2}$  "Il n'existe pas de différence significative entre les moyennes de débits de chacun des douze couples dispositif / tâche." est rejetée. Pour un dispositif donné, le type de tâche a une influence significative sur les résultats de ce dispositif.

Pour chacun des trois tests, les puissances observées sont suffisantes pour valider l'interprétation.



## RÉSULTATS NIVEAU 3

Tableau 4.5 : Analyse de variance pour le débit des dispositifs, niveau de difficulté 3.

Tests des effets intra-sujets (borne inférieure)	Somme des carrés	Moyenne des carrés	F	p
Dispositifs	214,412	214,412	35,929	< 0,000
Erreur(Dispositifs)	101,449	5,968		
Tâches	33,188	33,188	95,259	< 0,000
Erreur(Tâches)	6,363	0,374		
Dispositifs * Tâches	13,991	13,991	5,980	0,026
Erreur(Dispositifs * Tâches)	39,772	2,340		

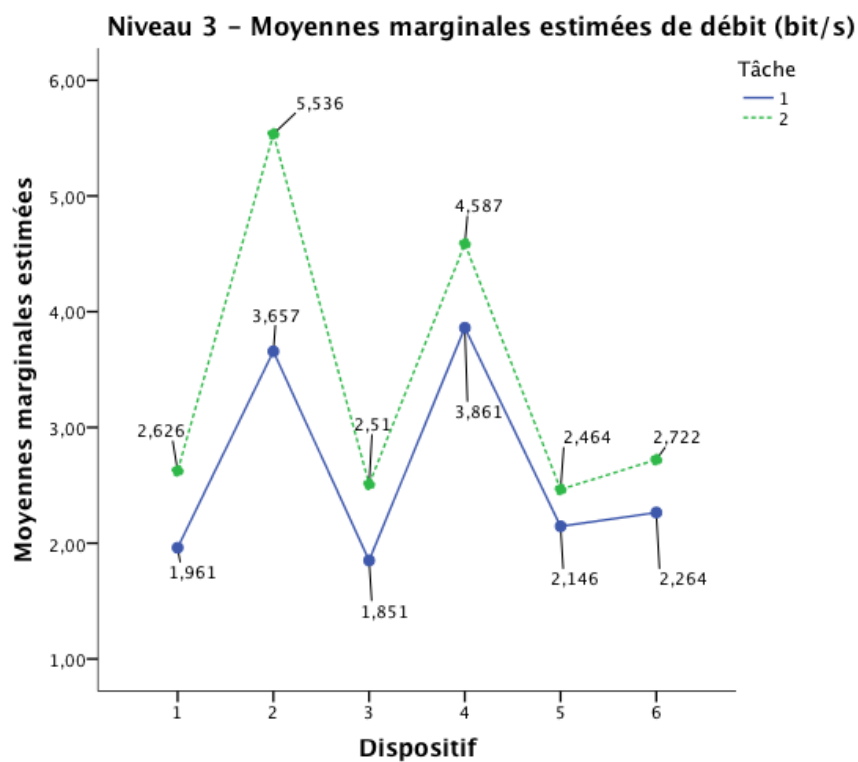


Figure 4.10 : Moyennes de débit par dispositif, niveau de difficulté 3.

On observe que la probabilité statistique (signification) associée à l'analyse de variance est inférieure à 0,05 pour l'effet induit par les dispositifs, ainsi que l'effet induit par les tâches.

## INTERPRÉTATION

Pour le niveau de difficulté 3 (6 bits) :

H<sub>01-3</sub> "Il n'existe pas de différence significative entre les moyennes de débits de chacun des six dispositifs." est rejetée. Au moins un des dispositifs obtient une moyenne combinée significativement différente des cinq autres.

H<sub>02-3</sub> "Il n'existe pas de différence significative entre les moyennes de débits de chacune des deux tâches." est rejetée. Le type de tâche a une influence significative sur les résultats moyens d'un dispositif.

H<sub>03-3</sub> "Il n'existe pas de différence significative entre les moyennes de débits de chacun des douze couples dispositif / tâche." n'est pas rejetée. Pour un dispositif donné, le type de tâche n'a pas d'influence significative sur les résultats de ce dispositif.

Pour chacun des trois tests, les puissances observées sont suffisantes pour valider l'interprétation.

#### OBSERVATIONS POUR L'ENSEMBLE DES RÉSULTATS DE DÉBIT

Le résultat principal apparaissant à l'issue de cette analyse montre que le classement des dispositifs en termes de débit moyen ne varie pas en fonction de la difficulté du test ou du type de tâche effectué. Dans l'ordre de débit moyen décroissant, on retrouve toujours :

- Première place : Tablette tactile,
- Deuxième place : Souris,
- Fin de classement : GETV, pavé tactile, boule de commande militaire & commerciale.

Ce résultat peut s'expliquer par deux facteurs : la proximité de la tablette (posée sur la table devant les sujets) par rapport aux autres dispositifs (déportés sur le grand écran) d'une part, une grande expertise du pointeur souris chez tous les participants d'autre part.

Autres tendances relevées :

- La différence de performance entre les deux tâches augmente avec l'augmentation du niveau de difficulté, à l'avantage de la tâche unidirectionnelle. Cette tendance peut être attribuée à une plus grande prévisibilité du test unidirectionnel.
- Le dispositif 1 (GETV) présente un comportement différent de celui des cinq autres dispositifs : pour le niveau 1, les moyennes de débit sont partout équivalentes pour les deux tâches, sauf pour le GETV. Quand on augmente le niveau de difficulté (diminution de la taille des cibles), cet effet s'estompe. Le GETV égalise les résultats des deux tâches au niveau 2 et rattrape le comportement des autres dispositifs au niveau 3.

La distance parcourue par le bras du participant semble donc exister une influence sur les résultats: pour les tâches à faible difficulté (grandes cibles) effectuées avec le GETV, l'inertie venant de l'amplitude du mouvement du bras sur le GETV annule les bénéfices issus de la prévisibilité de la tâche unidirectionnelle.

#### **4.4.2 Analyse de variance pour mesures répétées : taux d'erreurs**

Les restrictions présentes au cours dans la section 5.4.1 n'existent pas pour l'analyse des résultats des taux d'erreurs. En effet, le calcul du taux d'erreur ne dépend pas de la distance entre la frappe et le centre de la cible. Il s'agit pour chaque cible d'un simple test binaire. On procède à une seule analyse pour l'ensemble des données.

##### **HYPOTHÈSES STATISTIQUES**

$H_{0-D}$  : Il n'existe pas de différence significative entre les moyennes de taux d'erreurs de chacun des six dispositifs.

$H_{0-T}$  : Il n'existe pas de différence significative entre les moyennes de taux d'erreurs de chacune des deux tâches.

$H_{0-N}$  : Il n'existe pas de différence significative entre les moyennes de taux d'erreurs de chacune des trois niveaux de difficulté.

$H_{0-DT}$  : Il n'existe pas de différence significative entre les moyennes de taux d'erreurs de chacun des 12 couples dispositif / tâche.

$H_{0-DN}$  : Il n'existe pas de différence significative entre les moyennes de taux d'erreurs de chacun des 18 couples dispositif / niveau.

$H_{0-TN}$  : Il n'existe pas de différence significative entre les moyennes de taux d'erreurs de chacun des six couples tâche / niveau.

$H_{0-DTN}$  : Il n'existe pas de différence significative entre les moyennes de taux d'erreurs de chacun des 36 triplets dispositif / tâche / niveau.

$H_{1-D}$  : Au moins une des moyennes de taux d'erreurs par dispositifs est significativement différente des cinq autres.

$H_{1-T}$  : Les deux moyennes de taux d'erreurs par tâches sont significativement différentes.

$H_{1-N}$  : Au moins une des moyennes de taux d'erreurs par niveaux est significativement différente des deux autres.

$H_{1-DT}$  : Au moins une des moyennes de taux d'erreurs par couples dispositif / tâche est significativement différente des 11 autres.

$H_{1-DN}$  : Au moins une des moyennes de taux d'erreurs par couples dispositif / niveau est significativement différente des 17 autres.

$H_{1-TN}$  : Au moins une des moyennes de taux d'erreurs par couples tâche / niveau est significativement différente des cinq autres.

$H_{1-DTN}$  : Au moins une des moyennes de taux d'erreurs par triplets dispositif / tâche / niveau est significativement différente des 35 autres.

## RÉSULTATS

Tableau 4.6 : Analyse de variance pour le taux d'erreur des dispositifs.

Tests des effets intra-sujets (borne inférieure)	Somme des carrés	Moyenne des carrés	F	p
Dispositif	16523,844	16523,844	25,764	< 0,000
Erreur(Dispositif)	10903,033	641,355		
Tâche	8165,395	8165,395	101,332	< 0,000
Erreur(Tâche)	1369,875	80,581		
Niveau	110170,228	110170,228	425,511	< 0,000
Erreur(Niveau)	4401,517	258,913		
Dispositif * Tâche	7524,772	7524,772	23,735	< 0,000
Erreur(Dispositif*Tâche)	5389,596	317,035		
Dispositif * Niveau	37405,999	37405,999	44,888	< 0,000
Erreur(Dispositif*Niveau)	14166,530	833,325		
Tâche * Niveau	4242,263	4242,263	64,758	< 0,000
Erreur(Tâche*Niveau)	1113,667	65,510		
Dispositif * Tâche * Niveau	9538,187	9538,187	18,927	< 0,000
Erreur(Dispositif*Tâche*Niveau)	8567,246	503,956		

On observe que la probabilité statistique (signification) associée à l'analyse de variance est

inférieure à 0,005 pour tous les effets induits : par les dispositifs, les tâches, les niveaux, et leurs l'ensemble de leurs combinaisons respectives.

D'autre part, les puissances observées pour ces tests ne remettent pas en cause la validité des probabilités statistiques relevées.

## INTERPRÉTATION

$H_{0-D}$  est rejetée. Au moins un des dispositifs obtient une moyenne combinée significativement différente des cinq autres. On constate (figure 4.11) que les dispositifs 2 et 1 (tablette tactile et GETV) présentent un taux d'erreurs plus élevé par rapport à l'ensemble de la sélection.

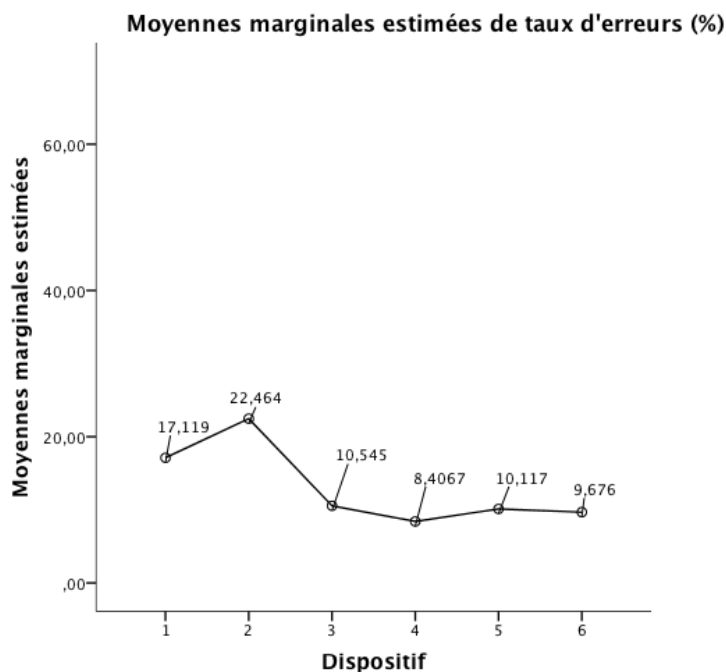


Figure 4.11 : Moyennes de taux d'erreurs par dispositif.

$H_{0-T}$  est rejetée. Le type de tâche a une influence significative sur les résultats moyens d'un dispositif. On constate (figure 4.12 page suivante) que la tâche 2 (tâche unidirectionnelle) présente un taux d'erreurs plus élevé par rapport à la tâche 1 (tâche unidirectionnelle).

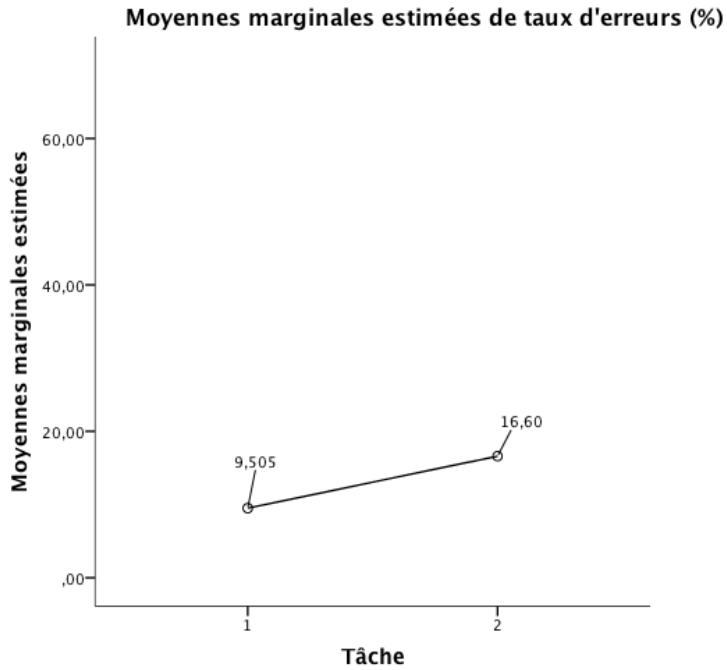


Figure 4.12 : Moyennes de taux d'erreurs par tâche.

$H_{0-N}$  est rejetée. Au moins un des niveaux de difficulté obtient une moyenne combinée significativement différente des trois autres. On constate (figure 4.13) que le niveau de difficulté 3 (6 bits) présente un taux d'erreurs significativement plus élevé par rapport autres niveaux (2 et 4 bits).

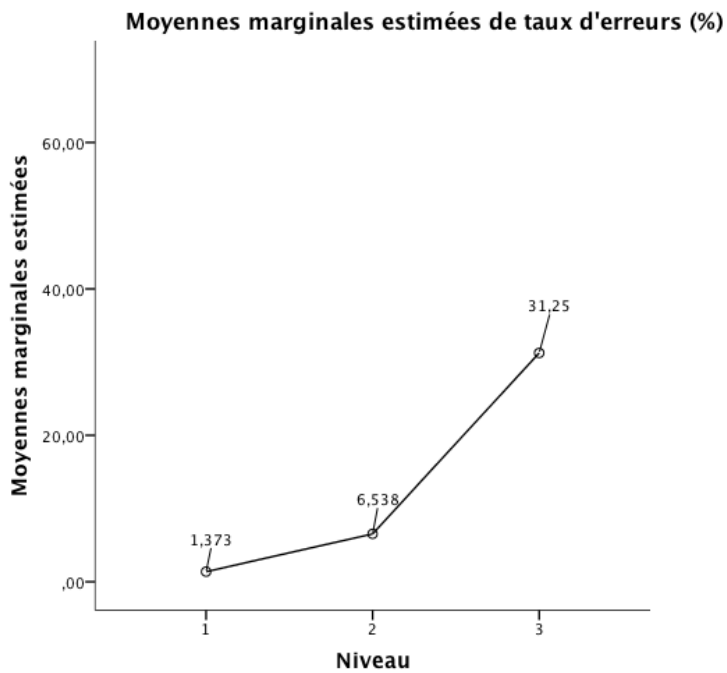


Figure 4.13 : Moyennes de taux d'erreurs par niveau.

$H_{0-DT}$  est rejetée. Au moins une des moyennes de taux d'erreurs par couples dispositif / tâche est significativement différente des 11 autres. Alors que le taux d'erreurs pour la tâche 2 (unidirectionnelle) varie peu entre chaque dispositif (Figure 4.14), la tâche bidirectionnelle présente de fortes variations. En particulier, les dispositifs 2 et 1 (tablette tactile et GETV) présentent un taux d'erreurs plus élevé par rapport aux autres dispositifs de la sélection.

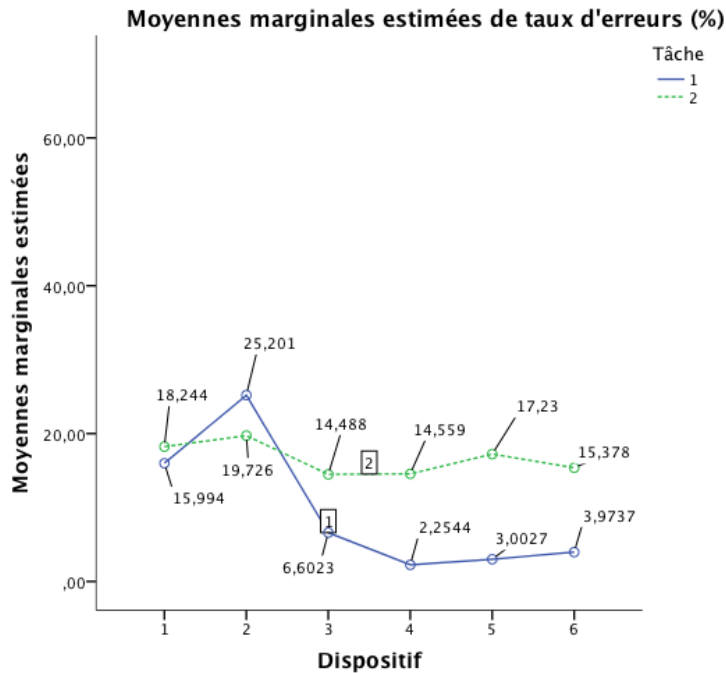


Figure 4.14 : Moyennes de taux d'erreurs par dispositif et par tâche.

$H_{0-DN}$  est rejetée. Au moins une des moyennes de taux d'erreurs par couples dispositif / niveau est significativement différente des 17 autres (Figure 4.15 page suivante). Les moyennes de taux d'erreurs varient peu pour les niveaux de difficulté 1 et 2 (2 et 4 bits respectivement). Le niveau 3 (6 bits) montre de fortes variations, avec les dispositifs 2 et 1 (tablette tactile et GETV) qui présentent un taux d'erreurs plus élevé par rapport aux quatre autres dispositifs.

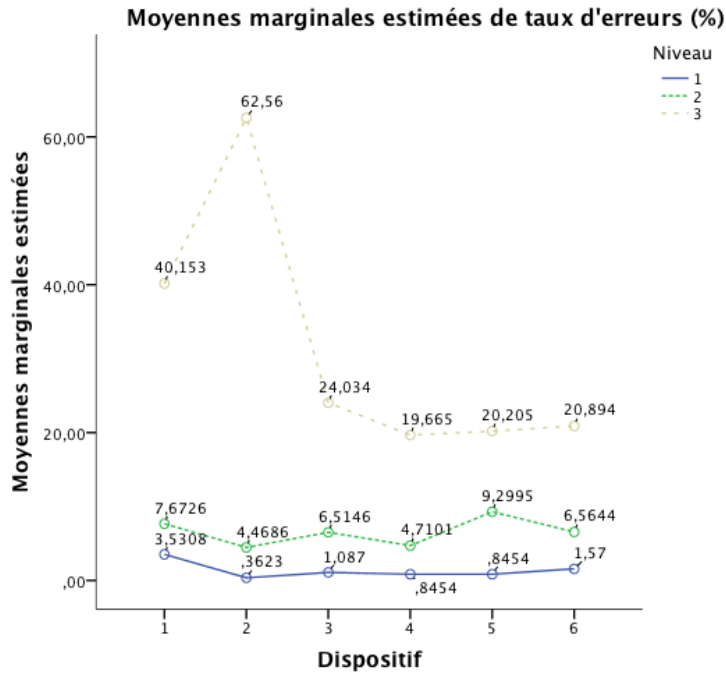


Figure 4.15 : Moyennes de taux d'erreurs par dispositif et par niveau.

$H_{0-TN}$  est rejetée. Au moins une des moyennes de taux d'erreurs par couples tâche / niveau est significativement différente des cinq autres (Figure 4.16). Le niveau de difficulté 3 présente un taux d'erreur plus élevé que les niveaux 1 et 2. De plus, pour les niveaux de difficulté 2 et 3, on constate un taux d'erreur plus élevé pour la tâche unidirectionnelle par rapport à la tâche bidirectionnelle.

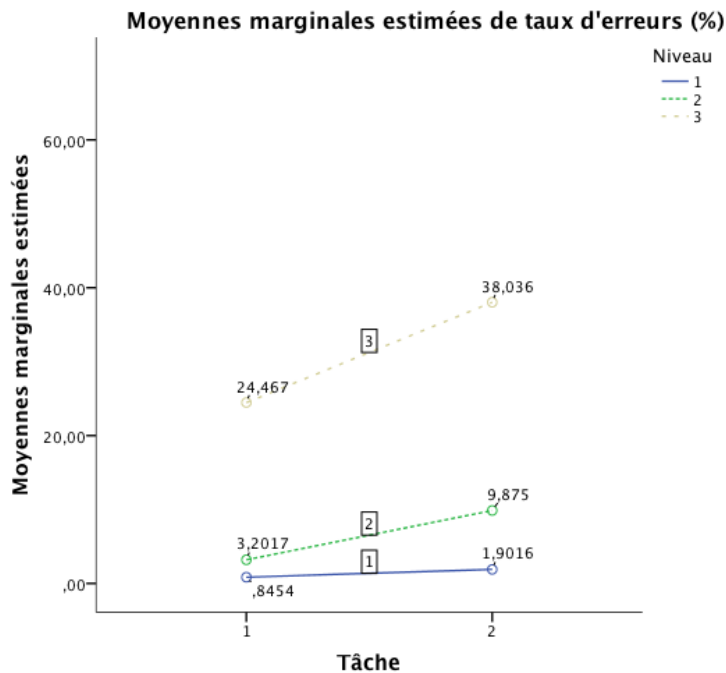


Figure 4.16 : Moyennes de taux d'erreurs par tâche et par niveau.



Par souci de clarté, on présente les résultats de taux d'erreurs par triplets dispositif / tâche / niveau découpés par niveaux de difficulté.

$H_{0-DTN}$  est rejetée. Au moins une des moyennes de taux d'erreurs par triplets dispositif / tâche / niveau est significativement différente des trente-cinq autres. Pour le niveau de difficulté 1 : 2 bits (Figure 4.17), les taux d'erreurs sont très proches les uns des autres et peu différents de zéro. On s'abstiendra donc d'effectuer une comparaison pour ces douze triplets.

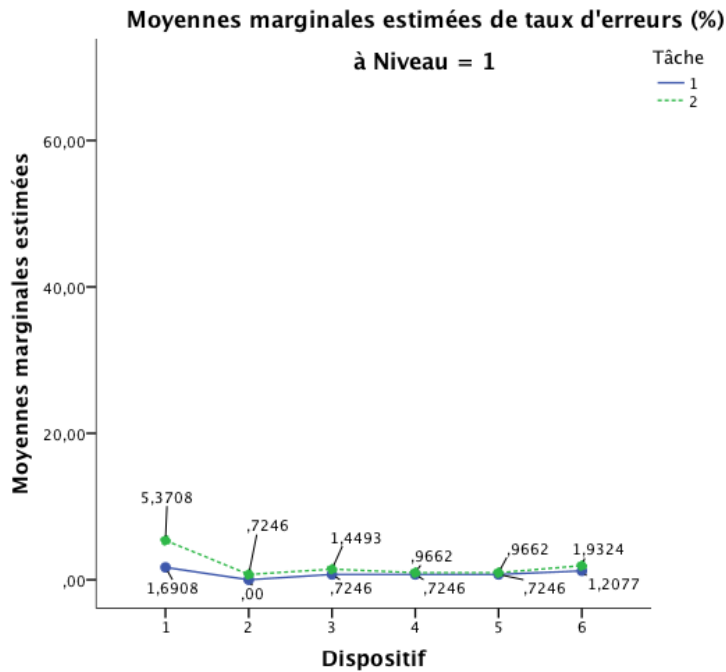


Figure 4.17 : Moyennes de taux d'erreurs par dispositif et par tâches, niveau 1.

Pour le niveau 2 : 4 bits (Figure 4.18 page suivante), on observe un taux d'erreurs plus élevé pour la tâche unidirectionnelle, sauf pour le dispositif tablette. Il est intéressant de noter que les dispositifs présentant les taux d'erreurs les plus faibles pour la tâche bidirectionnelle (dispositifs 4, 5, et 6, respectivement pavé tactile, boule de commande militaire & commerciale) sont également ceux qui montrent le plus fort taux d'erreurs pour la tâche unidirectionnelle.

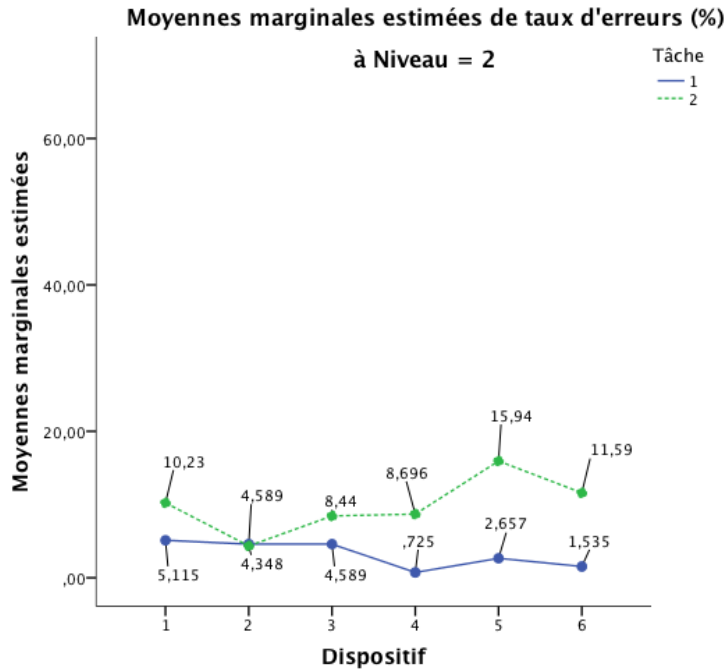


Figure 4.18 : Moyennes de taux d'erreurs par dispositif et par tâches, niveau 2.

Pour le niveau de difficulté 3 : 6 bits (Figure 4.19), on observe ici encore une performance meilleure pour la tâche bidirectionnelle, excepté pour les dispositifs tactiles (1 et 2, respectivement GETV et tablette). Ces deux dispositifs présentent également un taux d'erreurs significativement plus élevé que les dispositifs à curseur.

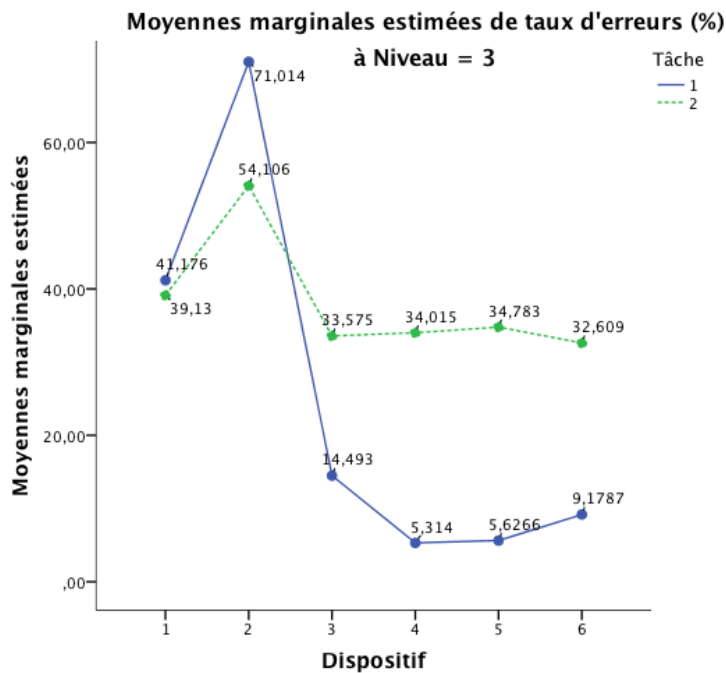


Figure 4.19 : Moyennes de taux d'erreurs par dispositif et par tâches, niveau 3.

## OBSERVATIONS POUR L'ENSEMBLE DES RÉSULTATS DE DÉBIT ET DE TAUX D'ERREURS

Les résultats principaux apparaissant à l'issue de cette analyse montrent que :

- Les dispositifs tactiles (D1 : GETV et D2 : Tablette) présentent un taux d'erreurs significativement plus élevé que les dispositifs à curseur (D3, D4, D5 et D6).
- Le taux d'erreurs augmente au fur et à mesure que la taille des cibles diminue.
- La tâche unidirectionnelle, pourtant plus prévisible, présente un taux d'erreur systématiquement plus élevé que la tâche multidirectionnelle (sauf pour la tablette). Ce résultat en apparence contradictoire peut être expliqué par le fait qu'il existe un effet d'inertie non négligeable pour les dispositifs D1, D3, D4, D5 et D6. Par conception, la distance entre les cibles est en effet plus grande pour la tâche unidirectionnelle que pour la tâche bidirectionnelle. Les participants ont tendance à dépasser les cibles lors du mouvement d'approche à cause des effets de l'inertie.

Le GETV est affecté également, car il s'agit cette fois-ci de déplacer son bras sur une distance très longue.

La tablette, dispositif tactile mais à déplacement court, bénéficie au contraire de la prévisibilité de la tâche unidirectionnelle.

- Les dispositifs à curseurs sont affectés d'une manière similaire par l'augmentation du niveau de difficulté. Les dispositifs tactiles voient eux leurs performances se dégrader très rapidement avec l'augmentation de la difficulté.
- Pour les tâches de faible difficulté, le taux d'erreurs est similaire entre les tâches unidirectionnelles et les tâches bidirectionnelles. Pour les difficultés supérieures, le taux d'erreur est plus fort pour les tâches unidirectionnelles.

## 4.5 Discussion

En préambule, on rappelle les moyennes des performances par dispositif dans le tableau Tableau 4.7 ci-dessous.

Tableau 4.7 : Synthèse des performances, par dispositif.

	Débit moyen (bits / s)	Taux d'erreurs moyen (%)	Satisfaction utilisateur amalgamée (Lickert de 1 à 7)
D1 : GETV	2,31	17,62	4,44
D2 : Tablette	4,66	23,25	6,07
D3 : Pavé tactile	2,01	10,51	5,50
D4 : Souris	3,60	8,73	6,16
D5 : Boule mil.	1,97	10,20	4,70
D6 : Boule com.	2,14	9,84	5,14

### 4.5.1 Comparaison avec les résultats issus de la littérature

Soukoreff & MacKenzie (2004) ont réalisé une compilation des tests de frappe suivant ISO 9241-9 menés entre 1998 et 2002. Dans le tableau 4.8 (page suivante), on présente comment les résultats de notre étude se comparent à ceux des études effectuées dans les mêmes conditions.

Tableau 4.8 : Comparaison avec les résultats issus de la littérature.

PAVÉ TACTILE					
	Type de tâche	Apprentissage	Etendue de l'ID (bits)	Débit (bits / s)	Taux d'erreurs (%)
MacKenzie & Oniszcak, 1998	Uni	Oui	1,0 - 4,1	1,43	9,92
Nos résultats	Uni	Oui (hors étude)	2,0 - 6,0	2,32	14,25
SOURIS					
	Type de tâche	Apprentissage	Etendue de l'ID (bits)	Débit (bits / s)	Taux d'erreurs (%)
MacKenzie & Jusoh, 2001	Uni	Non	1,0 - 4,1	3,7	2,4
Isokoski & Raisamo, 2002	Uni	Oui (hors étude)	1,58 - 8,00	4,40 - 4,60	6,90 - 9,40
Viale, 2013	Uni	Non	2,0 - 6,0	3,82	14,56
MacKenzie et al. (2001)	Multi	Oui	3,8	4,9	9,4
Keates et al. (2002)	Multi	Non	–	4,88	11
O h & Stuerzlinger, 2002	Multi	Oui	2,58 - 3,75	4,09	20
Nos résultats	Multi	Non	2,0 - 6,0	3,39	6,76
BOULE DE COMMANDE					
	Type de tâche	Apprentissage	Etendue de l'ID (bits)	Débit (bits / s)	Taux d'erreurs (%)
MacKenzie et al. (2001)	Multi	Oui	3,8	4,9	9,4
Nos résultats (militarisée)	Multi	Oui (hors étude)	2,0 - 6,0	1,78	3
Nos résultats (commerciale)	Multi	Oui (hors étude)	2,0 - 6,0	1,99	3,97

Pour la souris, on constate que les résultats que nous avons obtenus sont très comparables aux résultats issus de la littérature, et ce quel que soit le type de tâche ou l'étendue de l'indice de difficulté considéré. En particulier, les résultats sont très similaires à l'étude de MacKenzie et Jusoh (2001), effectuée dans des conditions comparables : contrôle du curseur sur un grand écran situé à distance.

Pour le pavé tactile, les valeurs que nous avons obtenues à la fois pour le débit et pour le taux d'erreur sont légèrement supérieures à celles obtenues par MacKenzie & Oniszczak (1998), mais restent comparables. Ce positionnement des valeurs pourrait indiquer que les participants à cette première étude ont effectué leurs séries de frappes moins rapidement. Il est donc permis de penser que l'utilisation d'un écran à distance avec ces deux dispositifs n'amène pas chez l'utilisateur de comportement nouveau susceptible de modifier ses habitudes de travail.

Les deux modèles de boule de commande obtiennent des résultats très proches, malgré de fortes différences d'inertie entre ces deux dispositifs. Les résultats que nous avons obtenus pour ces deux dispositifs sont très différents de ceux issus de la littérature. Le modèle utilisé dans l'expérimentation réalisée par MacKenzie et al. (2001) n'étant pas précisé, il est impossible d'en expliquer ces différences.

Outre les résultats de la boule de commande, on retient que le protocole et la méthodologie retenus pour cette étude n'introduisent pas de biais important rendant les résultats invalides.

Comment les dispositifs de contrôle à distance de notre étude se comparent-ils à ceux de l'étude de MacKenzie et Jusoh (2001) réalisée sur des dispositifs de commande à distance (1,5 m, écran de 15 pouces) ? On restreint la comparaison aux conditions de notre étude correspondantes (tâche unidirectionnelle, étendue du niveau de difficulté de 2 - 4 bits, à comparer à l'étendue de 1 - 4,1 bits pour MacKenzie & Jusoh). On a rassemblé dans le tableau 4.9 les performances de ces dispositifs. Les résultats pour la tâche unidirectionnelle avec une étendue de 2 - 6 bits y sont affichés entre parenthèses.

Tableau 4.9 : Comparaison de nos résultats à ceux de l'étude MacKenzie et Jusoh (2001)

Étude	Dispositif	Débit (bits / s)	Taux d'erreurs (%)
MacKenzie & Jusoh, 2001	GyroPoint-air	2,8	3,5
	GyroPoint-desk	4,1	2,7
	Souris	3,7	2,4
	RemotePoint	1,4	1,6
Nos résultats	Tablette	5,31 (5,38)	2,54 (19,73)
	Pavé tactile	2,22 (2,32)	4,95 (14,49)
	Souris	3,44 (3,69)	4,83 (14,56)
	Boule mil.	1,99 (2,15)	2,05 (17,23)
	Boule com.	2,04 (2,27)	1,37 (15,38)

On constate que les dispositifs de notre étude présentent globalement un taux d'erreur plus élevé, qui peut s'expliquer par le choix d'une étendue de l'ID légèrement plus large (les valeurs de l'ID de 4 à 6 bits présentant une difficulté plus élevée). En revanche, les boules de commande présentent un taux d'erreur relativement bas. Les performances en débit ne présentent pas de différence marquée avec ceux de l'étude de 2001.

Il serait malvenu de conclure que la tablette et le GyroPoint-desk sortent gagnants incontestés de la comparaison, car un choix d'étendue de l'ID différent réduit considérablement les performances de la tablette (cf. résultats entre parenthèses) : il se pourrait qu'il ait des effets similaires sur les dispositifs de l'étude de MacKenzie et Jusoh.

Ce chapitre a présenté les résultats détaillés de notre recherche. On retiendra notamment que les performances du GETV sont faibles par rapport aux autres dispositifs de la sélection. Il est donc légitime de chercher des dispositifs alternatifs ou complémentaires pour certains contextes d'emploi. Dans la conclusion, on évalue la réponse aux objectifs initiaux. On déterminera aussi les limites et les possibles suites à donner à ce travail.

[1] Le GyroPoint est un pointeur portatif qui peut opérer suivant deux modes : un mode "aérien" ("GyroPoint-air"), où un gyroscope détecte les variations d'angle en lacet et en tangage pour diriger le curseur ; et un mode "bureau" ("GyroPoint-desk") où le pointeur opère comme une souris.

[2] Le RemotePoint est un pointeur portatif, où l'utilisateur pilote un petit bâton de commande au bout du pouce.

## CONCLUSION

L'hypothèse de ce projet de recherche était la suivante :

H<sub>1</sub> : L'utilisation d'un dispositif d'interaction déporté (c'est-à-dire placé à distance de l'écran vertical) comparativement à une commande tactile directe améliore significativement la performance à la tâche de pilotage d'un curseur sur un écran tactile de grandes dimensions disposé verticalement.

Ces dispositifs déportés peuvent également améliorer la satisfaction utilisateur par un meilleur confort d'utilisation.

En effectuant la comparaison de performances du GETV avec plusieurs dispositifs déportés sur une tâche standardisée, on a découvert que le GETV se classe troisième sur six en termes de débit (2,31 bits / s contre 4,66 pour la tablette tactile, le dispositif le plus rapide, et 3,60 bits / s pour la souris, second dispositif le plus rapide), cinquième en termes de taux d'erreurs (17,62 % contre 8,73 pour la souris, dispositif le plus fiable), et sixième (dernier) en termes de satisfaction moyenne. Les différences de performances en termes de débit et de taux d'erreur sont statistiquement significatives.

L'hypothèse de départ de cette étude est donc prouvée : l'utilisation d'un dispositif d'interaction déporté améliore significativement la performance comparativement à l'utilisation du tactile direct pour la tâche de pilotage d'un curseur sur grand écran.

Les objectifs détaillés de ce projet de recherche étaient les suivants :

1. Démontrer la possibilité d'amélioration du mode d'interaction actuel pour les GETVs en termes de performance et de satisfaction humaine.
2. Démontrer la viabilité de la commande tactile déportée pour les GETVs et sa supériorité en termes de performance et de confort pour l'utilisateur.
3. Déterminer lequel des deux dispositifs tactiles utilisés à distance (tablette, pavé tactile) permet d'obtenir les meilleures performances et le meilleur confort.

Pour atteindre ces objectifs, on a comparé la performance et la satisfaction humaine relatives à l'utilisation de six dispositifs de pointage suivant la méthodologie proposée par ISO 9241-9 évaluant le débit, le taux d'erreur, et la satisfaction pour chaque dispositif. Les six dispositifs comparés étaient : un GETV de diagonale 46 pouces, une tablette tactile de diagonale 12 pouces, un



pavé tactile, une souris, et deux modèles de boule de commande (un modèle militarisé, et un modèle commercial). Chaque dispositif a été testé suivant six conditions expérimentales différentes (deux types de tâches (multidirectionnelle et unidirectionnelle) et trois niveaux de difficultés (par variation de la taille des cibles)).

Les principaux résultats ont montré que :

1. le GETV se classe troisième sur six en terme de débit, cinquième en terme de taux d'erreurs, et sixième en terme de satisfaction de l'utilisateur. Il est donc légitime et même souhaitable de chercher à améliorer les performances des GETVs, que ce soit par la recherche de dispositifs d'interactions alternatifs, mais aussi par la conception d'interfaces adaptées à ce dispositif.
2. Les dispositifs D2 (Tablette) et D3 (Pavé tactile) se classent dans l'ensemble positivement par rapport au GETV, avec un point fort pour la tablette en vitesse (4,66 bit / s de débit contre 2,31 bit / s pour le GETV) et en satisfaction (6,07 / 7 en satisfaction contre 4,44 / 7 pour le GETV), et pour le pavé tactile en fiabilité (10,51 % de taux d'erreur contre 17,62 % pour le GETV).
3. Les résultats pour les deux dispositifs de commande tactile (D2 : tablette et D3 : pavé tactile) sont plutôt contrastés :
  1. La tablette tactile (dispositif 2) obtient d'excellents résultats en termes de débits et de satisfaction utilisateurs (vitesse et aspect ludique du dispositif), mais obtient également les plus mauvaises performances en termes de taux d'erreurs. D'autre part, tous les participants ont fait le choix de regarder seulement la tablette au cours de ce test, transformant de fait le GETV en simple grand écran, non utilisé dans le cadre du test. Cet écran pourrait devenir un périphérique "pilote" dans un contexte avec plusieurs utilisateurs. Ces trois caractéristiques amènent à penser que cette configuration serait plutôt adaptée pour une utilisation domestique sans applications critiques, la tablette commandant un écran géant consacré par exemple au divertissement, à la surveillance, etc.
  2. Le pavé tactile a présenté des performances décevantes par rapport à ce qui était attendu : si le pavé est bien classé en termes de taux d'erreurs (10,51 %, meilleur dispositif tactile) et de la satisfaction (5,50 / 7), le débit est en revanche inférieur à celui obtenu par le GETV (2,01 bit / s de débit contre 2,31 bit / s pour le GETV). Ceci malgré une expérience à priori plus grande avec le pavé tactile qu'avec le GETV. Toute éventuelle implantation de la configuration dans un milieu critique devra donc s'accompagner d'une période d'apprentissage afin d'augmenter la rapidité de l'opérateur sans pour autant dégrader son taux

d'erreurs. Entre ces deux dispositifs tactiles, aucun ne se détache clairement quel que soit le cadre d'emploi.

Il faut souligner quelques autres résultats intéressants :

- La souris reste globalement le dispositif le plus performant (débit 3,60 bits / s, deuxième meilleur dispositif, taux d'erreur 8,73 %, premier dispositif, satisfaction 6,16, meilleur dispositif), quel que soit la tâche ou le niveau de difficulté.
- Malgré la différence d'inertie importante – qui s'accompagne également d'une différence marquée en terme de satisfaction, toutes deux à l'avantage de la boule commerciale – entre la boule de commande militarisée (plus lourde) et sa version commerciale (plus légère), la différence de performance est minime (2,14 bit / s de débit pour la boule commerciale contre 1,97 bit / s pour boule militarisée, et 9,84 % de taux d'erreur pour la boule commerciale contre 10,20 % pour boule militarisée).

Se pose maintenant pour le choix d'un dispositif de remplacement du GETV la question de la pondération relative des variables de l'étude, en fonction du contexte d'emploi du système.

En effet, pour les objectifs 2 et 3, la réponse peut varier en fonction de l'environnement : on peut présumer qu'une tablette sera destinée à une application domestique ou de loisir (rapidité, aspect ludique, erreurs non critiques), alors qu'une boule de commande sera plutôt destinée à un contexte professionnel : industrie, domaine militaire (aucune tolérance aux erreurs).

Concernant l'adéquation des dispositifs étudiés avec la tâche réelle de contrôle des dommages effectuée à bord de navires militaires ou civils, on constate que le GETV et la boule de commande militarisée (actuellement utilisés par la Marine royale canadienne) présentent de sérieux désavantages par rapport à d'autres dispositifs de cette recherche. Sans supprimer totalement la fonctionnalité tactile des GETVs, une grande partie des tâches gagneront à être effectuées en utilisant un dispositif déporté complémentaire.

Dans un contexte militaire, on préconise de compléter l'emploi du GETV par un pavé tactile "militarisé", de type résistif analogique ou bien à capteurs infrarouges. Certaines tablettes tactiles militarisées font déjà usage de ces technologies. Dans un contexte civil, il convient d'étudier la possibilité de compléter le GETV par une souris, périphérique le plus économique. Si les contraintes opérationnelles rendent l'emploi d'une souris impossible, on préconise alors l'utilisation d'un pavé tactile.

## Limites de l'expérimentation

Certains choix faits lors de cette expérimentation imposent d'exploiter et d'interpréter les résultats avec prudence. Ces éléments sont détaillés ci-dessous.

- Approche choisie pour la distance inter-cible fixe : on a choisi de garder la distance inter-cible fixe, en faisant varier le niveau de difficulté uniquement en jouant sur la taille des cibles. Ce choix introduit un biais négatif qui joue en défaveur du GETV comparativement aux autres dispositifs, faisant de ce dernier le seul dispositif imposant de large mouvement du corps pour frapper les cibles d'une extrémité à l'autre de l'écran. Il est cependant justifié : en effet, l'usage courant de ce dispositif impose à l'opérateur de frapper des cibles sur toute la surface de l'écran. La distance inter-cibles fixe tente d'approcher ce comportement. Ne pas imposer une distance inter-cible fixe aurait légèrement amélioré les résultats du GETV pour la durée de l'expérimentation, mais cet avantage n'existe pas en dehors d'un milieu expérimental : une distance  $d$  existant entre deux cibles augmente proportionnellement avec la diagonale de l'écran, et vient à imposer des geste d'une ampleur peu confortable quand on atteint les dimensions d'un GETV.
- Biais en faveur de la souris : toute étude comparative de dispositifs de pointage incluant une souris est biaisée en faveur de ce dispositif parce que l'immense majorité de la population – en particulier lorsque l'échantillon est composé de travailleurs provenant d'une entreprise de haute technologie – peut être considérée comme experte à la souris. On a tout de même choisi d'inclure un pointeur souris dans l'étude afin de pouvoir comparer les résultats de l'étude à ceux issus de la littérature. Pour s'affranchir de ce biais positif envers la souris, il pourrait s'avérer intéressant de reprendre l'expérimentation auprès d'une population non entraînée (en effet, la génération qui accède aujourd'hui à l'informatique est beaucoup plus susceptible de devenir experte avec une interface tactile durant l'enfance, bien avant de découvrir la souris).
- Le trop faible nombre de participants ( $N=18$ ) empêche d'effectuer une analyse de variance des performances en fonction des niveaux d'expérience des participants avec les différents dispositifs, le nombre de participants par niveau d'expérience étant trop faible pour faire des tests de comparaison. Un nombre de participants plus élevé aurait pu permettre de caractériser mieux les effets de l'apprentissage.
- Nous n'avons pas effectué de tests statistiques pour la satisfaction à cause du faible nombre de réponses par dispositif ( $N = 18$ ), à la différence des variables de performance où il a été possible de faire des regroupements par dispositifs, tâches, niveaux de difficultés ( $N = 18 * 2 * 3 ; 18 * 6 *$

3 ; 18 \* 6 \* 2, respectivement). On s'est donc limité à une analyse descriptive commentée présentant les niveaux de satisfaction envers chaque dispositif.

## **Proposition de travaux futurs**

Il existe des aberrations à corriger (ou plus probablement des erreurs de typographie, étant donné que la formule est correcte ailleurs dans les autres tâches) dans la formulation de l'indice de difficulté pour la tâche unidirectionnelle dans la norme ISO 9241-9. La dernière version parue date de 2000. À la connaissance de l'auteur, aucun travail n'est en cours pour la réviser.

De plus, dans sa formulation actuelle, la norme n'est pas adaptée aux interfaces tactiles multi-points modernes. ISO 9241 est initialement prévue pour tester les performances d'un seul pointeur à l'écran, elle est donc peu adaptée à l'utilisation de deux ou plusieurs pointeurs utilisés simultanément dans les configurations tactiles multi-points qui sont la norme actuellement (jusqu'à 10 en fonction du constructeur). De plus, les tâches élémentaires que la norme permet de tester (Pointage, Sélection, Déplacement, Entrée à main levée, Traçage) constituent seulement une sous-partie des tâches élémentaires d'une interface tactile moderne (les gestes tactiles de zoom, de rotation, etc n'étant pas pris en compte).

Avec la complexité nouvelle des tâches que l'informatique actuelle permet, Glessner, Bérard et Cooperstock (2013) évaluent par exemple la performance de divers dispositifs : souris, surface tactile multipoint, tactile multipoint sur deux plans perpendiculaires ("Cubtile"), et stylet de pointage dans l'espace ("Phantom") pour une tâche de translation et rotation d'un objet en trois dimensions. Glessner et al. concluent sur la supériorité des dispositifs Cubtile et Phantom pour la manipulation d'objets dans les trois dimensions.

Leurs travaux, à l'instar de ceux présentés dans ce mémoire, nous amènent à nous interroger sur l'évolution des tâches qui seront demain accomplies au travers d'interfaces graphiques, sur la nécessité de créer des dispositifs d'interaction adaptés à ces tâches, et enfin sur celle de pouvoir mesurer ces dispositifs entre eux.

## RÉFÉRENCES

1. Bartram, L., Ware, C., & Calvert, T. (2003). Moticons: detection, distraction and task. *International Journal of Human-Computer Studies*, 58(5), 515-545. Tiré de <http://www.interruptions.net/literature/Bartram-IJHCS03-BW.pdf>
2. Bied Sperling, B., & Tullis, T. S. (1988). Are you a better Mouser or Trackballer? A comparison of cursor-positioning performance. *ACM SIGCHI Bulletin*, 19(3), 77-81.
3. Chourasia, A. O., Wiegmann, D. A., Chen, K. B., Irwin, C. B., & Sesto, M. E. (2013). Effect of Sitting or Standing on Touch Screen Performance and Touch Characteristics. *Human Factors: The Journal of the Human Factors and Ergonomics Society*. Tiré de <http://hfs.sagepub.com/content/early/2013/01/08/0018720812470843.full.pdf+html>
4. Card, S. K., English, W. K., & Burr, B. J. (1978). Evaluation of mouse, rate-controlled isometric joystick, step keys, and text keys for text selection on a CRT. *Ergonomics*, 21(8), 601-613.
5. Douglas, S. A., Kirkpatrick, A. E., & MacKenzie, I. S. (1999). Testing pointing device performance and user assessment with the ISO 9241, Part 9 standard. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, 215-222. ACM. Tiré de <http://csi.ufs.ac.za/resres/files/Douglas.pdf>
6. English, W. K., Engelbart, D. C., & Berman, M. L. (1967) Display-Selection Techniques for Text Manipulation. *IEEE Transactions on Human Factors in Electronics*, March, Vol. HFE-8, No. 1, 5-15. Tiré de <http://sloan.stanford.edu/MouseSite/Archive/AugmentingHumanIntellect62/Display1967.html>

7. Fitts, P. M. (1954). The information capacity of the human motor system in controlling the amplitude of movement. *Journal of Experimental Psychology*, 47, 381-391.
8. Gillan, D. J., Holden, K., Adam, S., Rudisill, M., & Magee, L. (1990). How does Fitts' law fit pointing and dragging?. In *Proceedings of the SIGCHI conference on Human factors in computing systems: Empowering people* (pp. 227-234). ACM. Consulté le 10 mai 2012, tiré de <http://dl.acm.org/citation.cfm?id=97278>
9. Glesser, D., Bérard, F., & Cooperstock, J. R. (2013). Overcoming limitations of the trackpad for 3d docking operations. In *CHI'13 Extended Abstracts on Human Factors in Computing Systems*, 1239-1244. ACM. Tiré de <http://www.cim.mcgill.ca/sre/publications/2013-CHI-WIP.pdf>
10. Hammer, M. (2007). [Ergonomic Comparison of Keyboard and Touch Screen Data Entry While Standing and Sitting](#) (Mémoire de maîtrise). Université de Cincinnati.
11. Hart, S. G., & Staveland, L. E. (1988). Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. *Human mental workload*, 1(3), 139-183. Tiré de <http://humanfactors.arc.nasa.gov/groups/TLX/downloads/NASA-TLXChapter.pdf>
12. Ishiyama, K., & Yano, S. (2000). A study of characteristics of pointing devices for television operation. In *Systems, Man, and Cybernetics, 2000 IEEE International Conference on Systems, Man, and Cybernetics* (Vol. 2, 1307-1312). IEEE. Tiré de : <http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=886034>
13. Isokoski, P., & Raisamo, R. (2002). Speed-accuracy measures in a population of six mice. *Proceedings of APCHI 2002 (5th Asia Pacific Conference on Computer Human Interaction)* (Science Press, Beijing, 2002), 765-777. Consulté le 3 mars 2012, tiré de [http://www.sis.uta.fi/~pi52316/apchi\\_2002\\_paper/apchi\\_paper.html](http://www.sis.uta.fi/~pi52316/apchi_2002_paper/apchi_paper.html)

14. Keates, S., Hwang, F., Langdon, P., Clarkson, P. J., & Robinson, P. (2002). Cursor measures for motion-impaired computer users. In *Proceedings of the fifth international ACM conference on Assistive technologies* (pp. 135-142). ACM. Tiré de <http://link.springer.com/content/pdf/10.1007%2Fs10209-002-0033-7.pdf>
15. MacKenzie, I. S. (1989). A note on the information-theoretic basis for Fitts' law. *Journal of Motor Behavior*, 21, 323-330. Tiré de : <http://www.yorku.ca/mack/JMB89.html>
16. MacKenzie, I. S., & Buxton, W. (1993). [A tool for the rapid evaluation of input devices using Fitts' law models](#). *SIGCHI Bulletin*, 25(3), 58-63. Consulté le 10 mai 2012.
17. MacKenzie, I. S., & Jusoh, S. (2001). [An evaluation of two input devices for remote pointing](#). *Proceedings of the Eighth IFIP International Conference on Engineering for Human-Computer Interaction – EHCI 2001*, (pp. 235-249) Heidelberg, Germany: Springer-Verlag. Consulté le 3 mars 2012.
18. MacKenzie, I. S., Kauppinen, T., & Silfverberg, M. (2001). [Accuracy measures for evaluating computer pointing devices](#). *Proceedings of the ACM Conference on Human Factors in Computing Systems – CHI 2001*, (pp. 9-16). New York: ACM. Consulté le 3 mars 2013.
19. MacKenzie, I. S., & Oniszczak, A. (1998). [A comparison of three selection techniques for touchpads](#). *Proceedings of the ACM Conference on Human Factors in Computing Systems – CHI '98*, (pp. 336-343). New York: ACM.
20. Meeker, M., Wu, L., (2013, mai). *2013 Internet Trends*. Communication présentée à la conférence D11, Rancho Palos Verdes. Tiré de : <http://www.kpcb.com/insights/2013-internet-trends>

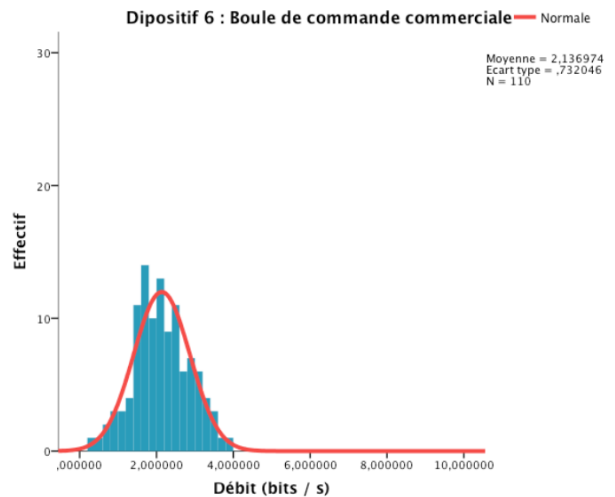
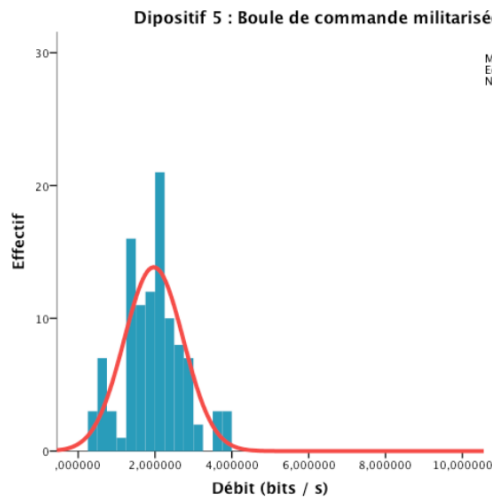
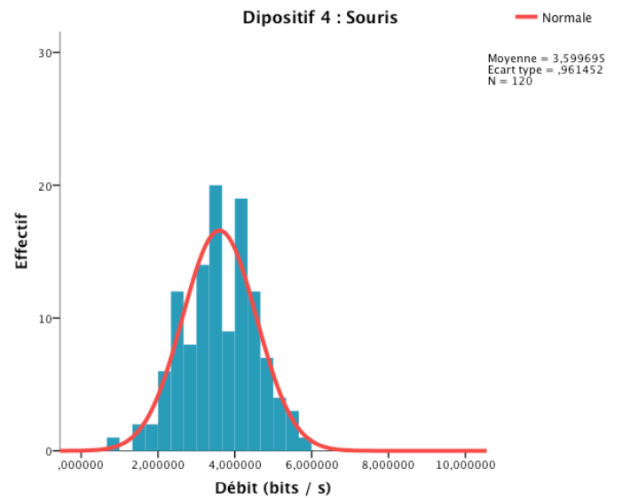
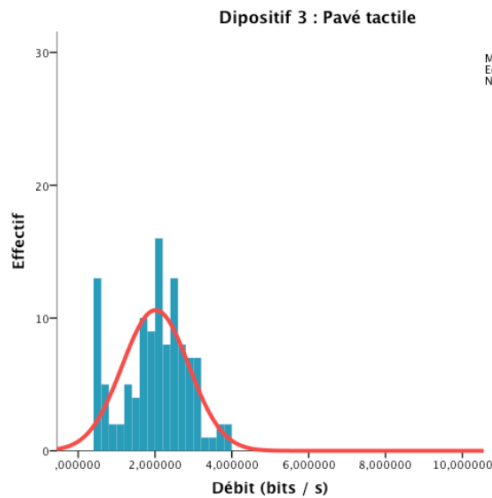
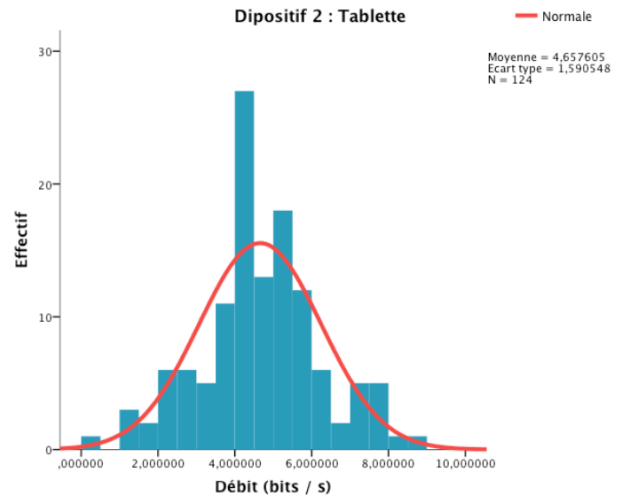
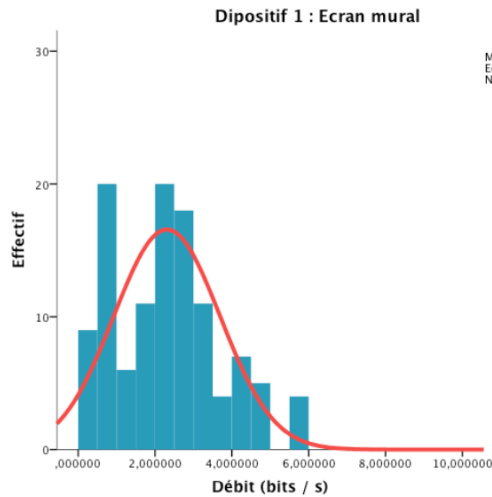
21. Murata, A. (1991). An experimental evaluation of mouse, joystick, joycard, lightpen, trackball, and touchscreen for pointing: Basic study on human interface design. Human aspect in computing: Design and Use of Interactive Systems and Work with Terminal, H.J. Bullinger (ed), pp 123 -127.
22. Oh, J. Y., & Stuerzlinger, W. (2002). Laser pointers as collaborative pointing devices. In *Graphics Interface* (pp. 141-150).
23. Organisation internationale de normalisation. (2000). 9241-9 Exigences ergonomiques pour travail de bureau avec terminaux à écrans de visualisation (TEV) - Partie 9 : Exigences relatives aux dispositifs d'entrée autres que les claviers.
24. National Aeronautics and Space Administration. (1981). Preferred Height and Angle Settings of CRT and Keyboard for a Display Station Input Task, Proceedings of the Human Factors and Ergonomics Society Annual Meeting, October 1981 25: 492-496.
25. Shannon, C. E., & Weaver, W. (1949). The mathematical theory of communication (Urbana, IL. *University of Illinois Press*, 19(7), 1.
26. Soukoreff, R. W., & MacKenzie, I. S. (2004). Towards a standard for pointing device evaluation: Perspectives on 27 years of Fitts' law research in HCI. *International Journal of Human-Computer Studies*, 61, 751-789. Tiré de <http://www.yorku.ca/mack/ijhcs2004.html>
27. Tilley, A. R. (1959). *The Measure of Man and Woman : Human Factors in Design* (éd. 2002). New York : Henry Dreyfuss and Associates.
28. Welford, A. T. (1960). The measurement of sensory-motor performance: survey and reappraisal of twelve years' progress. *Ergonomics*, 3(3), 189-230.



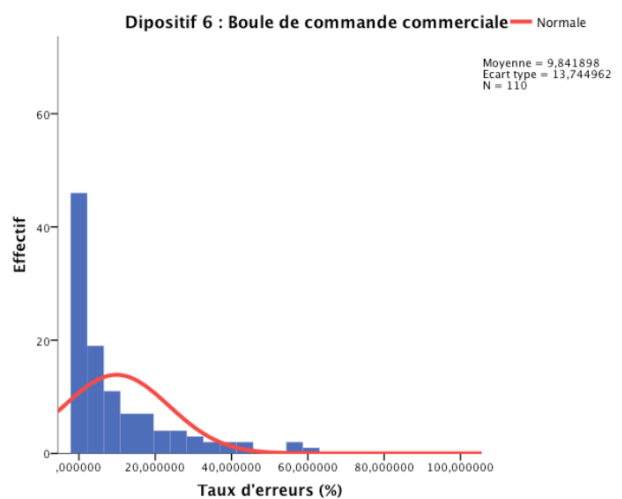
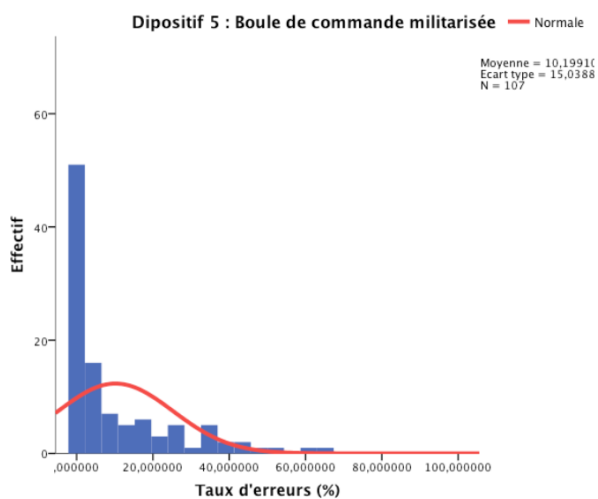
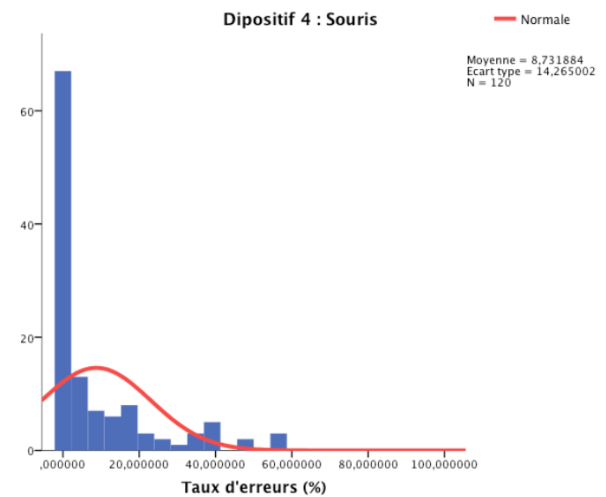
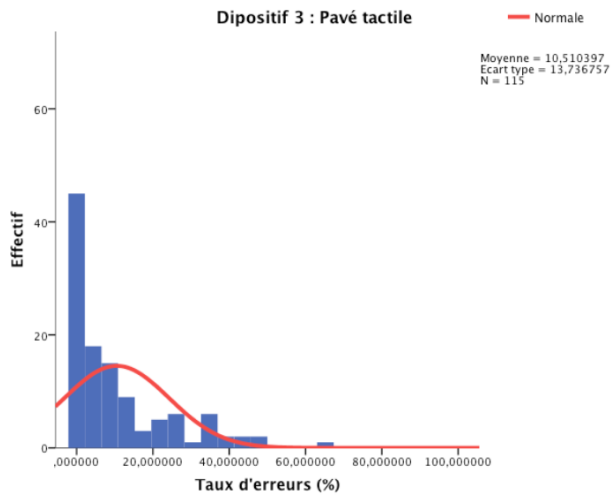
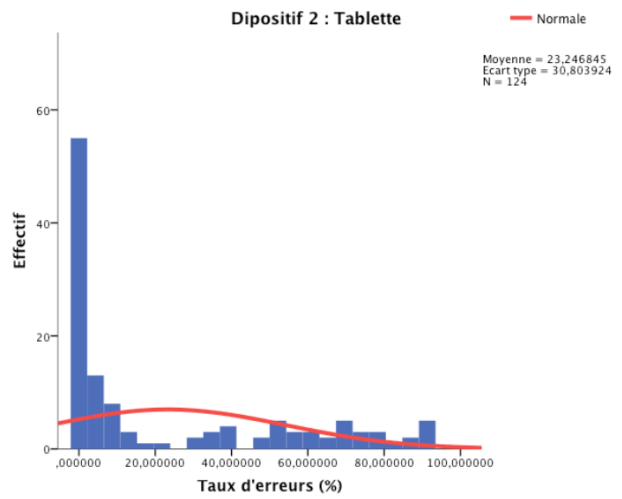
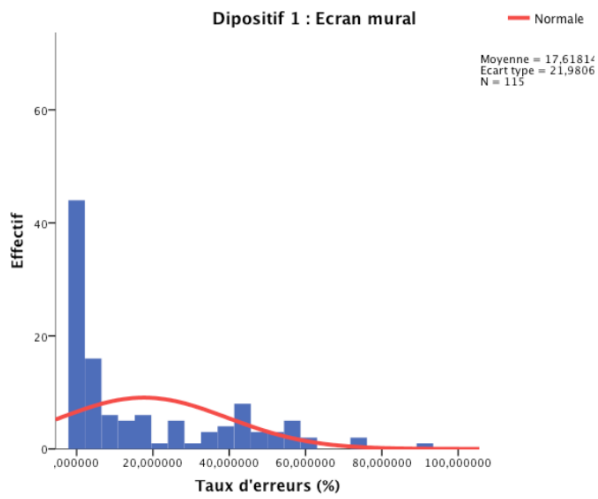
29. Yarow, J. (2010). Steve Jobs: Touchscreen Laptops Don't Work. In *Business Insider*. Tiré de [http://articles.businessinsider.com/2010-10-20/tech/29974850\\_1\\_multitouch-trackpad-notebook#ixzz2Sv1i30mj](http://articles.businessinsider.com/2010-10-20/tech/29974850_1_multitouch-trackpad-notebook#ixzz2Sv1i30mj)

# ANNEXE A - RÉPARTITION DES DONNÉES PAR DISPOSITIF

## DÉBITS PAR DISPOSITIF



## TAUX D'ERREURS PAR DISPOSITIF



# ANNEXE B - DONNÉES RECUEILLIES PAR LE CODE

DÉBIT (BIT/S) PAR PARTICIPANT ET PAR SÉRIE

Condition	ex	Sujet	Alpha	Bravo	Charli	Delta	Echo	Foxtr	Golf	Hotel	India	Juliet	Kilo	Lima	Mike	Nover	Oscar	Papa	Quebr	Rome	Sierra	Tango
D1	C1	N1	2,353	2,373	1,863	1,934	2,072	2,771	0,559	0,764	2,847	1,641	2,347	1,679	1,519	1,683	4,265	0,543	2,648	2,794	2,224	2,054
D1	C1	N2	2,866	3,804	1,463	2,126	0,684	2,694	2,168	2,153	4,067	4,221	5,689	1,936	2,592	1,385	4,876	0,415	2,409	3,928	2,406	
D1	C1	N3	0,714	2,865	1,401	0,375	0,202	2,601	2,622		3,211	4,620	3,261	0,363	0,822	2,289	3,109	0,301	0,599	4,176	2,151	2,797
D1	C2	N1	1,091	2,443		0,519	0,896	2,367	1,417		3,093	2,268	0,886	0,892	0,627	0,851	2,603	2,435	1,430	0,853	1,740	0,788
D1	C2	N2	2,900	3,171		0,548	2,216	2,619	3,001	1,796	4,158	4,013	3,563	2,831	1,724	1,318	4,967	2,186	3,167	2,156	2,814	1,668
D1	C2	N3	3,424	4,550	0,384	3,112	0,305	0,584	0,925	0,636	4,966	5,641	3,083	0,601	3,478	2,658	5,833	0,462	3,000	4,390	2,880	3,556
D2	C1	N1	3,644	4,878	3,689	2,646	5,234	4,181	4,148	4,380	5,486	4,365	3,775	4,914	4,520	3,076	5,427	4,410	5,249	1,764	3,297	5,071
D2	C1	N2	4,820	5,320	4,990	2,651	4,226	5,361	4,239	4,366	5,022	5,058	4,464	4,262	4,291	4,034	1,419	4,155	5,042	1,843	2,977	5,938
D2	C1	N3	5,831	3,943	4,314	2,419	2,751	4,089	4,402	3,712	4,177	4,095	4,608	3,039	3,768	4,122	3,136	1,022	4,043	2,237	2,645	5,530
D2	C2	N1	2,026	7,500	6,361	3,726	5,351	4,359	3,606	4,822	3,884	5,731	5,047	5,693	2,191	2,400	7,756	4,323	6,129	2,914	3,675	2,134
D2	C2	N2	8,850	7,852	7,737	5,740	5,966	5,284	7,272	6,769	6,448	7,822	5,215	7,378	4,885	4,776	5,059	5,791	6,239	3,554	4,131	7,094
D2	C2	N3	8,342	6,208	7,244	4,490	5,396	5,524	5,981	5,603	5,373	5,632	4,606	5,727	5,251	6,937	4,398	4,623	6,120	4,309	4,004	7,009
D3	C1	N1	2,496	0,831	2,067	0,701	0,662	1,486	1,347	1,788	1,795	1,738	2,186	1,429	2,125		2,559	1,563	1,580	1,687	1,189	2,033
D3	C1	N2	2,815	1,704	2,466	0,472	0,695	2,006	1,858	2,124	2,163	2,653	0,834	2,063	2,155		2,765	0,525	2,564	2,345	1,776	2,534
D3	C1	N3	2,907	0,471	2,493	2,113	0,435	0,548	1,891	0,542	2,610	2,703	1,067	2,054	0,493		2,490	1,843	2,368	2,333	1,988	2,426
D3	C2	N1	2,570	1,295	1,614	0,595	1,372	1,935	1,648	1,244	2,304	1,883	2,326	1,692	2,076		2,704	1,915	1,787	2,399	1,889	2,238
D3	C2	N2	3,155	1,939	3,100	0,532	2,146	2,563	2,092	2,681	2,860	2,854	3,165	2,119	2,919		3,668	0,764	3,319	3,017	2,351	
D3	C2	N3	3,827	0,541	3,032	2,589	0,400	0,617	2,740	2,455	2,923	3,139	3,675	2,711	0,529		3,856	0,482	2,925	3,578	2,575	3,110
D4	C1	N1	3,981	2,570	2,792	2,947	2,364	3,029	1,669	2,975	2,511	2,140	2,622	3,161	2,741	2,498	4,134	2,249	2,265	2,568	2,810	2,671
D4	C1	N2	4,453	3,633	4,055	4,592	3,468	4,280	3,087	3,540	3,335	3,498	4,205	3,145	3,575	3,675	4,937	2,615	3,152	3,310	3,354	4,451
D4	C1	N3	4,946	3,233	4,295	4,162	3,564	4,025	3,381	3,525	4,107	3,627	4,47	3,606	3,614	3,686	3,984	3,236	3,232	4,192	3,011	4,510
D4	C2	N1	3,285	2,665	3,340	3,283	2,161	2,767	1,798	3,461	1,456	2,705	2,650	3,596	1,460	3,480	4,344	2,081	2,272	2,510	2,488	3,865
D4	C2	N2	5,086	3,784	4,182	5,915	3,016	4,452	3,050	4,492	4,419	3,636	4,698	4,051	0,947	4,409	5,578	3,407	4,086	4,344	4,167	4,766
D4	C2	N3	5,656	3,897	4,382	4,937	4,118	5,150	3,507	4,333	4,703	4,038	5,599	4,101	4,270	3,890	5,275	3,701	4,840		4,242	5,114
D5	C1	N1	1,716		1,480	1,326	0,957	1,309	1,404	1,401	1,428	0,700	1,400	1,158	1,341		2,807	0,967	2,146	1,559	1,519	1,910
D5	C1	N2	0,654		0,728	2,093	1,494	1,751	2,078	1,847	1,887	1,864	1,840	1,742	0,561		3,519	1,325	1,983	2,232	0,646	2,528
D5	C1	N3	3,078		2,433	2,264	2,068	2,248	2,471	2,262	2,452	2,174	2,239	2,074	2,029		3,953	0,404	1,585	0,307	2,433	
D5	C2	N1	2,143		1,431	1,390	1,445	1,740	1,570	1,561	1,546	1,268	1,976	1,340	0,664		2,886	0,987	2,445	1,838	1,467	1,734
D5	C2	N2	2,768		2,587	2,061	2,019	2,200	2,273	1,945	2,176	1,813	2,523	1,851	2,117		3,947	1,741	2,679	2,196	2,445	2,843
D5	C2	N3	3,613		3,007	2,526	0,493	2,312	2,527	2,682	2,842	2,075	2,713	2,189	0,507		3,823	2,037	1,545	2,949	2,905	3,604
D6	C1	N1	1,760		1,673	1,429	1,294	1,613	1,522	1,354	1,667	1,478	1,685	1,308	1,638		2,014	0,600	1,489	1,901	1,463	1,526
D6	C1	N2	2,663		2,599	2,158	1,725	2,099	2,170	1,977	2,330	1,745	2,140	1,773	2,446		2,599	1,583	1,898	2,396	2,127	2,562
D6	C1	N3	3,005		2,251	2,365	1,568	2,194	2,150	1,973	2,615	1,856	2,609	2,167	2,779		2,927	0,425	2,330	2,860	2,090	2,584
D6	C2	N1	2,275		1,673	1,290	1,492	0,818	1,053	2,026	1,755	0,826	1,770	1,117	1,829		1,753	1,197	0,916	2,060	1,806	1,634
D6	C2	N2	2,992		2,705	2,558	2,241	2,468	1,999	2,923	2,572	2,150	2,901	2,252	2,594		3,186	1,845	1,949	2,517	3,160	2,924
D6	C2	N3	3,466		3,097	3,473	0,728	2,800	2,520	3,296	3,014	2,160	3,838	2,257	3,292		3,288	0,391	1,402	3,008	3,259	3,689

## TAUX D'ERREURS (%) PAR PARTICIPANT ET PAR SÉRIE

Condition	ex	Sujet	Alpha	Bravo	Charli	Delta	Echo	Foxtrot	Golf	Hotel	India	Juliet	Kilo	Lima	Mike	Nover	Oscar	Papa	Quebec	Rome	Sierra	Tango
D1	C1	N1	0	0	0	0	0	0	4,347	4,347	0	0	0	8,695	0	0	0	13,04	0	0	0	0
D1	C1	N2	0	0	26,08	0	17,39	4,347	8,695	4,347	0	0	0	4,347	4,347	8,695	0	17,39	0	0	0	
D1	C1	N3	17,39	34,78	47,82	56,52	43,47	17,39	56,52		26,08	47,82	43,47	60,86	39,13	73,91	26,08	73,91	60,86	34,78	8,695	39,13
D1	C2	N1	4,347	0	43,47	17,39	0	0	0		0	0	8,695	0	4,347	4,347	0	0	0	8,695	0	4,347
D1	C2	N2	4,347	0		21,73	17,39	13,04	0	43,47	4,347	4,347	4,347	0	39,13	13,04	0	13,04	0	4,347	0	4,347
D1	C2	N3	26,08	43,47		43,47	39,13	52,17	47,82	43,47	30,43	52,17	56,52	56,52	26,08	56,52	0	52,17	34,78	47,82	13,04	43,47
D2	C1	N1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
D2	C1	N2	4,347	13,04	4,347	0	8,695	0	17,39	0	13,04	0	8,695	4,347	4,347	8,695	4,347	0	4,347	8,695	0	0
D2	C1	N3	39,13	91,30	73,91	91,30	91,30	73,91	86,95	52,17	69,56	60,86	91,30	65,21	86,95	69,56	56,52	52,17	78,26	73,91	82,60	52,17
D2	C2	N1	4,347	0	0	0	0	0	0	0	0	0	0	0	4,347	0	0	0	0	0	0	4,347
D2	C2	N2	0	0	8,695	4,347	0	21,73	8,695	0	8,695	0	0	0	4,347	30,43	4,347	0	13,04	4,347	0	0
D2	C2	N3	30,43	56,52	47,82	52,17	39,13	69,56	78,26	60,86	52,17	65,21	60,86	34,78	69,56	39,13	56,52	34,78	78,26	39,13	69,56	34,78
D3	C1	N1	0	4,347	0	4,347	4,347	0	0	0	0	0	4,347	0	0		0	0	0	0	0	0
D3	C1	N2	4,347	0	8,695	8,695	8,695	0	13,04	13,04	4,347	8,695	4,347	0	0		0	8,695	0	0	0	0
D3	C1	N3	8,695	4,347	13,04	4,347	17,39	13,04	26,08	21,73	13,04	21,73	26,08	4,347	21,73		4,347	0	4,347	43,47	0	17,39
D3	C2	N1	0	0	0	8,695	0	0	4,347	0	4,347	4,347	4,347	0	0		0	0	0	0	0	0
D3	C2	N2	8,695	13,04	8,695	34,78	8,695	8,695	8,695	13,04	4,347	17,39	0	0	0		0	8,695	0	13,04	8,695	
D3	C2	N3	34,78	26,08	34,78	26,08	39,13	30,43	39,13	43,47	65,21	47,82	26,08	4,347	47,82		13,04	34,78	21,73	26,08	34,78	34,78
D4	C1	N1	0	0	8,695	0	0	0	0	0	0	0	4,347	0	0	0	0	0	0	0	0	0
D4	C1	N2	0	0	4,347	0	4,347	0	0	0	4,347	0	0	0	0	0	0	0	0	0	0	0
D4	C1	N3	0	13,04	8,695	0	8,695	13,04	17,39	8,695	8,695	4,347	4,347	0	13,04	34,78	0	4,347	0	4,347	0	0
D4	C2	N1	0	0	0	0	0	0	0	0	8,695	0	0	0	4,347	0	0	0	0	0	0	4,347
D4	C2	N2	8,695	26,08	26,08	0	13,04	0	17,39	13,04	17,39	17,39	4,347	0	30,43	21,73	0	0	0	4,347	0	4,347
D4	C2	N3	17,39	39,13	39,13	34,78	39,13	34,78	56,52	56,52	56,52	39,13	13,04	21,73	47,82	39,13	21,73	17,39	17,39		17,39	47,82
D5	C1	N1	0		0	0	0	0	0	0	0	8,695	0	0	0		0	0	0	4,347	0	0
D5	C1	N2	4,347		4,347	0	0	4,347	0	4,347	0	17,39	4,347	0	4,347		0	0	0	0	4,347	0
D5	C1	N3	0		17,39	0	4,347	4,347	0	0	0	21,73	26,08	0	0		0	8,695	0	13,04	0	
D5	C2	N1	0		4,347	0	0	0	0	0	0	4,347	0	0	4,347		0	0	0	0	0	4,347
D5	C2	N2	13,04		30,43	17,39	13,04	26,08	8,695	26,08	8,695	21,73	13,04	21,73	17,39		0	17,39	8,695	26,08	8,695	8,695
D5	C2	N3	26,08		52,17	39,13	34,78	39,13	34,78	34,78	13,04	65,21	60,86	34,78	43,47		4,347	34,78	0	47,82	17,39	43,47
D6	C1	N1	0		0	4,347	4,347	0	0	4,347	0	0	0	0	4,347		0	4,347	0	0	0	0
D6	C1	N2	0		0	0	0	0		0	0	4,347	4,347	0	4,347		4,347	0	0	8,695	0	0
D6	C1	N3	0		30,43	4,347	13,04	17,39	17,39	8,695	0	21,73	4,347	4,347	4,347		0	17,39	0	4,347	8,695	8,695
D6	C2	N1	0		0	0	0	13,04	8,695	0	0	4,347	0	0	0		4,347	0	4,347	0	0	0
D6	C2	N2	8,695		17,39	8,695	8,695	21,73	21,73	4,347	13,04	13,04	8,695	13,04	4,347		0	17,39	8,695	30,43	0	8,695
D6	C2	N3	26,08		60,86	13,04	43,47	34,78	34,78	13,04	39,13	17,39	21,73	43,47	26,08		26,08	56,52	30,43	56,52	26,08	17,39

## SCORE DE SATISFACTION (LICKERT DE 1 À 7) PAR DISPOSITIF ET PAR PARTICIPANT

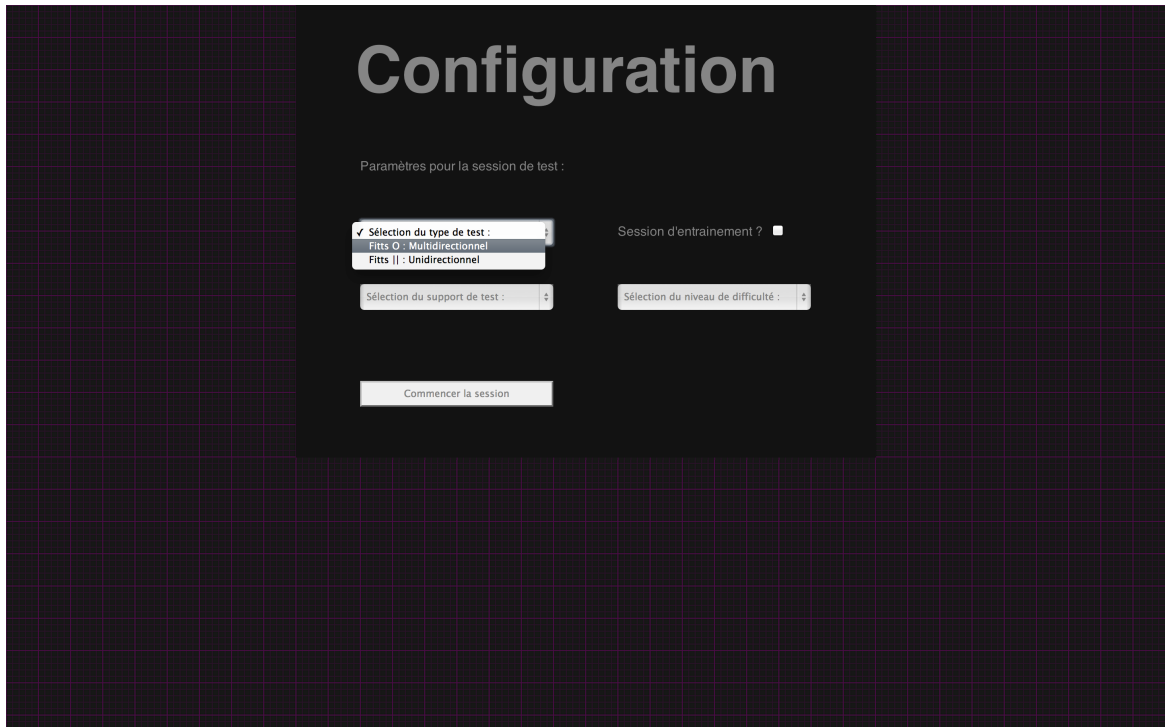
	Alpha						Bravo						Charlie						Delta					
	Ecrar	Table	Track	Souri	TBM	TBC	Ecrar	Table	Track	Souri	TBM	TBC	Ecrar	Table	Track	Souri	TBM	TBC	Ecrar	Table	Track	Souri	TBM	TBC
Force requise pour l'ac	4	6	6	7	4	7	6	6	6	7	3	5	3	6	7	7	6	6	6	7	X	6		
Souplesse d'utilisation	6	7	5	7	5	6	6	6	5	6	3	3	X	5	6	7	5	4	6	7	X	5		
Efforts requis	6	6	6	7	5	7	4	5	6	7	3	5	3	6	6	7	6	5	6	7	X	6		
Précision	6	7	5	6	5	6	6	6	4	4	2	2	5	4	7	7	4	5	5	6	X	5		
Vitesse d'utilisation	7	7	5	7	5	7	3	5	3	4	3	3	5	6	7	7	6	6	6	6	X	4		
Confort global	5	7	5	7	4	6	4	4	4	6	2	5	4	6	5	7	6	6	6	6	X	5		
Fonctionnement global	6	6	5	7	5	6	6	5	4	6	4	4	5	6	6	6	6	5	6	6	X	5		
Fatigue des doigts	7	6	6	7	6	4	6	6	6	7	6	6	7	7	7	6	7	6	6	6	X	3		
Fatigue du poignet	7	7	7	5	6	4	6	6	5	6	2	3	7	7	5	6	5	6	7	6	X	6		
Fatigue du bras	5	7	7	7	5	4	2	3	6	7	6	6	5	7	7	6	7	7	5	6	X	6		
Fatigue de l'épaule	7	7	7	7	7	7	4	5	6	7	6	6	2	7	7	7	7	7	5	5	X	5		
Fatigue du cou	7	7	7	7	7	7	3	3	6	6	6	6	7	7	7	7	7	6	6	X	6			
	Echo						Foxtrot						Golf						Hotel					
	Ecrar	Table	Track	Souri	TBM	TBC	Ecrar	Table	Track	Souri	TBM	TBC	Ecrar	Table	Track	Souri	TBM	TBC	Ecrar	Table	Track	Souri	TBM	TBC
Force requise pour l'ac	4	7	6	6	6	6	6	7	7	7			6	5	7	6	7	5	3	5	5	7	6	6
Souplesse d'utilisation	6	7	4	6	2	4	5	6	5	5			7	5	7	7	7	4	4	6	5	7	4	6
Efforts requis	5	7	5	6	1	6	3	7	5	6			4	5	6	6	6	4	3	7	5	7	4	5
Précision	6	7	4	6	3	5	6	6	4	5			6	6	6	7	6	5	4	5	4	7	5	5
Vitesse d'utilisation	6	7	4	6	1	5	5	7	5	5			5	5	6	7	6	4	3	6	5	7	5	5
Confort global	6	7	4	6	1	4	4	6	5	4			5	6	6	7	6	5	4	7	6	7	5	6
Fonctionnement global	6	7	4	6	2	4	4	7	4	6			6	5	7	7	6	5	4	6	6	7	5	5
Fatigue des doigts	5	7	4	6	5	6	7	7	3	7			3	4	6	7	6	5	6	7	6	7	5	6
Fatigue du poignet	5	7	4	6	1	6	7	7	4	2			3	4	6	6	5	4	6	7	5	6	4	6
Fatigue du bras	5	7	4	6	4	6	1	3	7	7			1	5	6	6	5	5	3	7	7	7	6	7
Fatigue de l'épaule	5	7	4	6	7	6	5	6	7	7			1	4	6	6	7	6	3	7	7	7	6	7
Fatigue du cou	5	7	4	6	6	6	7	7	6	7			1	6	6	6	6	5	5	7	7	7	7	7
	India						Juliet						Kilo						Lima					
	Ecrar	Table	Track	Souri	TBM	TBC	Ecrar	Table	Track	Souri	TBM	TBC	Ecrar	Table	Track	Souri	TBM	TBC	Ecrar	Table	Track	Souri	TBM	TBC
Force requise pour l'ac	7	7	7	7	6	7	2	3	6	7	3	2	5	7	5	5	5	6	1	7	6	6	3	7
Souplesse d'utilisation	7	7	7	7	5	7	3	4	6	7	3	4	3	7	4	4	4	4	7	7	6	6	5	7
Efforts requis	7	7	6	7	7	7	3	5	6	7	3	3	6	4	4	3	5	3	7	6	5	3	4	
Précision	4	5	5	7	7	6	3	4	5	6	4	3	2	5	5	4	4	5	5	5	6	6	4	
Vitesse d'utilisation	5	6	6	7	7	7	5	6	5	7	4	3	2	7	5	4	4	5	7	6	6	6	5	6
Confort global	5	7	7	7	7	7	4	5	6	7	4	4	1	5	4	4	4	5	4	7	6	5	3	6
Fonctionnement global	6	7	7	7	7	7	3	5	6	7	4	4	1	5	4	4	4	5	5	7	6	6	4	6
Fatigue des doigts	7	7	6	6	6	7	2	5	5	7	3	4	3	5	6	5	3	5	5	7	7	6	3	6
Fatigue du poignet	7	7	7	5	7	5	4	5	5	4	4	3	3	4	3	3	4	7	7	5	5	6	7	
Fatigue du bras	5	7	7	7	7	7	3	5	5	5	3	4	2	6	6	3	5	5	5	7	7	6	6	7
Fatigue de l'épaule	7	7	7	7	7	7	3	5	4	5	3	4	2	6	6	5	5	5	4	7	7	6	6	7
Fatigue du cou	7	7	7	7	7	7	3	5	4	5	3	4	1	5	6	6	5	5	6	7	7	6	6	7
	Mike						November						Oscar						Papa					
	Ecrar	Table	Track	Souri	TBM	TBC	Ecrar	Table	Track	Souri	TBM	TBC	Ecrar	Table	Track	Souri	TBM	TBC	Ecrar	Table	Track	Souri	TBM	TBC
Force requise pour l'ac	6	7	5	7	5	6	3	6	6	6	5	5	6	4	6	7	5	7	6	7	5	4	5	6
Souplesse d'utilisation	7	7	5	7	5	6	4	6	6	5	5	5	5	6	6	7	4	6	6	7	6	7	5	6
Efforts requis	7	7	6	7	5	6	4	7	7	6	5	5	2	6	6	7	5	4	3	7	6	5	4	6
Précision	7	7	6	7	6	5	7	6	6	7	4	4	4	5	6	6	5	6	4	6	6	7	6	4
Vitesse d'utilisation	7	7	7	7	5	6	4	7	6	6	4	4	6	6	7	7	6	5	4	7	6	7	6	6
Confort global	6	7	6	7	4	6	3	5	5	6	4	4	3	6	6	6	6	4	4	7	6	7	4	5
Fonctionnement global	6	7	5	7	5	6	4	6	7	6	5	4	4	6	6	6	7	5	4	6	5	6	5	5
Fatigue des doigts	6	7	6	6	4	5	4	6	4	6	4	5	6	5	6	6	3	3	7	5	7	7	7	7
Fatigue du poignet	7	7	4	6	4	5	2	6	7	6	5	5	6	5	6	5	5	3	7	7	4	6	3	5
Fatigue du bras	6	6	7	7	6	6	5	6	7	7	5	5	3	6	6	6	6	6	5	7	7	7	5	7
Fatigue de l'épaule	5	7	7	7	6	6	4	6	7	7	5	5	3	6	6	6	6	6	2	5	7	6	7	7
Fatigue du cou	7	7	7	7	6	6	7	7	4	4	5	5	6	5	6	5	6	6	6	7	6	6	7	7

	Quebec						Romeo						Sierra						Tango					
	Ecrar	Table	Track	Souri	TBM	TBC	Ecrar	Table	Track	Souri	TBM	TBC	Ecrar	Table	Track	Souri	TBM	TBC	Ecrar	Table	Track	Souri	TBM	TBC
<b>Force requise pour l'ac</b>	2	7	5	7	2	7	2	6	6	6	3	6	3	7	7	7	6	7	5	7	5	6	4	7
<b>Souplesse d'utilisation</b>	3	7	6	7	5	4	3	5	6	5	4	6	4	6	6	6	5	5	4	7	3	5	5	2
<b>Efforts requis</b>	2	7	6	6	3	6	3	6	6	5	3	7	3	6	6	6	4	4	6	7	5	7	3	2
<b>Précision</b>	6	5	3	6	2	1	3	4	5	6	5	3	2	6	2	6	2	2	5	6	4	5	5	1
<b>Vitesse d'utilisation</b>	3	7	4	7	5	6	5	7	4	6	6	5	2	6	2	6	1	1	4	7	2	6	2	2
<b>Confort global</b>	3	7	4	6	2	2	4	5	6	6	4	6	1	4	2	6	3	2	3	6	3	6	1	2
<b>Fonctionnement global</b>	3	6	4	6	2	2	3	7	6	6	5	5	1	6	2	7	2	2	3	7	3	4	2	2
<b>Fatigue des doigts</b>	6	6	6	5	2	2	2	5	4	7	3	4	2	5	4	5	5	6	7	7	6	6	2	6
<b>Fatigue du poignet</b>	6	6	6	5	6	6	6	6	5	6	6	4	1	4	1	4	2	1	7	7	3	6	1	6
<b>Fatigue du bras</b>	4	6	6	5	6	6	3	4	5	5	6	5	1	3	2	6	4	4	6	7	7	7	4	5
<b>Fatigue de l'épaule</b>	2	7	6	6	6	6	6	6	6	6	6	6	1	3	2	5	7	6	3	7	7	7	5	6
<b>Fatigue du cou</b>	3	6	5	7	6	6	6	6	7	7	6	6	4	6	6	5	7	7	6	6	6	5	4	2

## ANNEXE C - PAGES DU SITE DE TEST

On montre seulement les captures dans l'ordre de déroulement normal du test français. Le site en anglais, identique, n'est pas présenté.

### PAGE DE CONFIGURATION



The screenshot shows a dark-themed configuration interface for a test session. The title "Configuration" is prominently displayed at the top. Below it, the text "Paramètres pour la session de test :" indicates the settings for the test session. The interface includes several interactive elements: a dropdown menu for "Sélection du type de test :" with options "Fitts O : Multidirectionnel" and "Fitts IJ : Unidirectionnel"; a checkbox for "Session d'entraînement ?"; two more dropdown menus for "Sélection du support de test :" and "Sélection du niveau de difficulté :"; and a "Commencer la session" button at the bottom.

Configuration

Paramètres pour la session de test :

✓ Sélection du type de test :  
Fitts O : Multidirectionnel  
Fitts IJ : Unidirectionnel

Session d'entraînement ?

Sélection du support de test :  
Sélection du niveau de difficulté :

Commencer la session



## PAGE D'ACCUEIL

**Bienvenue !**

Vous allez bientôt débiter une série. Veuillez renseigner les champs suivants :

Entrez votre surnom

Le support utilisé durant la série sera *un écran mural avec tablette aveugle*.

Indiquez votre expérience avec ce support :

Accéder à la série

## PAGE D'INITIALISATION

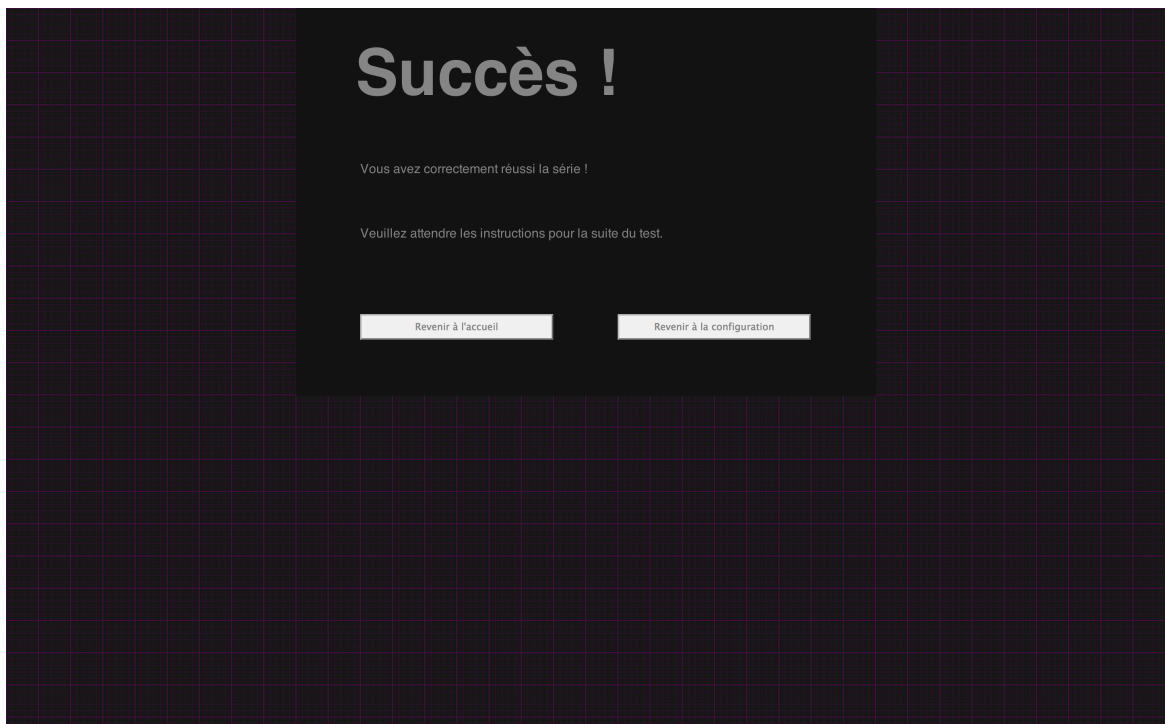
Quand vous êtes prêt(e), pressez le bouton au centre de l'écran.

Commencer la série

## PAGE DE TEST - TEST UNIDIRECTIONNEL, DIFFICULTÉ MOYENNE



## PAGE DE FIN DE SÉRIE



# ANNEXE D - CERTIFICAT D'ACCEPTATION PAR LE CER



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

Montréal, le 6 décembre 2012

Comité d'éthique de la  
recherche avec des  
êtres humains

M. Étienne Viale  
M. Jean-Marc Robert  
Département de mathématiques et génie industriel  
École Polytechnique de Montréal

Adresse civique :  
Campus de l'Université de Montréal  
900, boul. Édouard-Montpetit  
École Polytechnique  
500, chemin de Polytechnique  
H3T 1J4

N/Réf : Dossier CÉR-11/12-17

Adresse postale :  
C.P. 6079, succursale Centre-ville  
Montréal (Québec) Canada  
H3C 3A7

Messieurs,

téléphone : (514) 340-4990  
télécopieur : (514) 340-4992

École affiliée à  
Université de Montréal

Membres réguliers du comité :

Marie-Josée Bernardi, juriste et éthicienne  
Mario Bourgault, mathématiques et génie  
industriel  
Farida Cheriet, génie Informatique et  
génie logiciel\*  
Sophie De Serres, IRSST  
Delphine Périé-Curnier, génie mécanique  
Lodje Petit, juriste et éthicienne  
Jean-Marc Robert, mathématiques et  
génie industriel

Céline Roehrig, secrétaire

Présidente du Comité

J'ai le plaisir de vous informer que les membres du Comité d'éthique de la recherche ont procédé à l'évaluation en comité restreint de votre projet de recherche intitulé « *Optimisation et adaptation d'une interface tactile pour le système de gestion intégrée d'une plateforme de commande et de contrôle dans le domaine naval* » et en ont recommandé l'approbation sur la base des modifications apportées à votre documentation transmises par courriel à Mme Roehrig.

Veillez noter que le présent certificat est valable pour le projet tel que soumis au Comité d'éthique de la recherche avec des sujets humains. La secrétaire du Comité d'éthique de la recherche avec des sujets humains devra immédiatement être informée de toute modification qui pourrait être apportée ultérieurement au protocole expérimental, de même que de tout problème imprévu pouvant avoir une incidence sur la santé et la sécurité des personnes impliquées dans le projet de recherche (sujets, professionnels de recherche ou chercheurs).

Nous vous prions également de nous faire parvenir un bref **rapport annuel** ainsi qu'un avis à la fin de vos travaux.

Je vous souhaite bonne chance dans vos travaux de recherche,

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

c.c.: Céline Roehrig, DRI

## ANNEXE E - QUESTIONNAIRE DE SATISFACTION

### Questionnaire de satisfaction utilisateur

Souris

Code :

Ce questionnaire a pour but d'évaluer votre satisfaction vis-à-vis du dispositif d'entrée que vous venez d'utiliser.

Veillez noter le dispositif au regard des douze aspects suivants, avec 1 note la plus basse et 7 meilleure note :

*Force requise pour l'activation*

<b>Grande</b>				<b>Faible</b>		
1	2	3	4	5	6	7

*Souplesse d'utilisation*

<b>Peu souple</b>				<b>Très souple</b>		
1	2	3	4	5	6	7

*Efforts requis pour l'utilisation*

<b>Grands</b>				<b>Faibles</b>		
1	2	3	4	5	6	7

*Précision*

<b>Peu précis</b>				<b>Très précis</b>		
1	2	3	4	5	6	7

*Vitesse d'utilisation*

<b>Lent</b>				<b>Rapide</b>		
1	2	3	4	5	6	7

*Confort global*

<b>Inconfortable</b>				<b>Confortable</b>		
1	2	3	4	5	6	7

*Fonctionnement global du dispositif*

<b>Insatisfaisant</b>				<b>Satisfaisant</b>		
1	2	3	4	5	6	7

*Fatigue des doigts*

<b>Grande</b>				<b>Faible</b>		
1	2	3	4	5	6	7

*Fatigue du poignet*

<b>Grande</b>				<b>Faible</b>		
1	2	3	4	5	6	7

*Fatigue du bras*

<b>Grande</b>				<b>Faible</b>		
1	2	3	4	5	6	7

*Fatigue de l'épaule*

<b>Grande</b>				<b>Faible</b>		
1	2	3	4	5	6	7

*Fatigue du cou*

<b>Grande</b>				<b>Faible</b>		
1	2	3	4	5	6	7

Vous pouvez également noter toute autre remarque dans la partie ci-dessous :

## ANNEXE F - RÉPARTITION DE L'EXPÉRIENCE

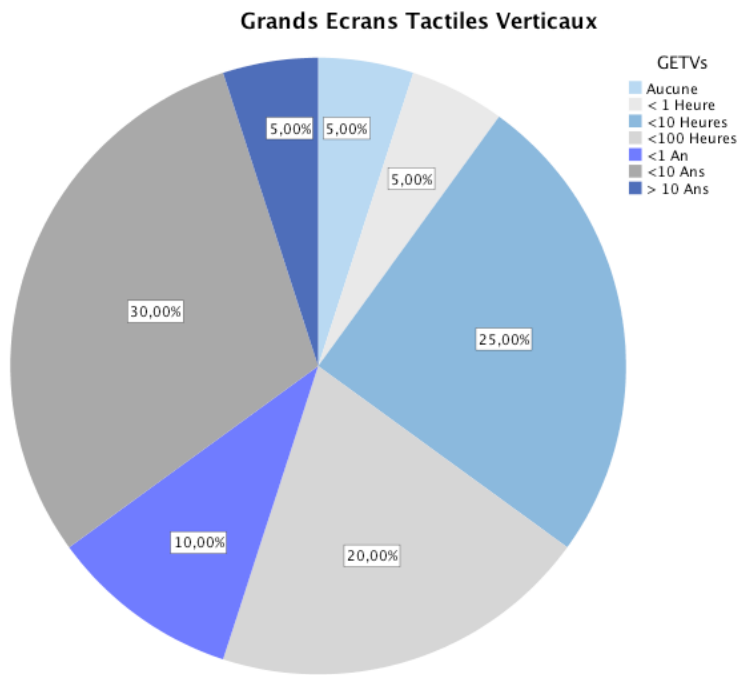


Figure F.1 : Expérience avec les GETV

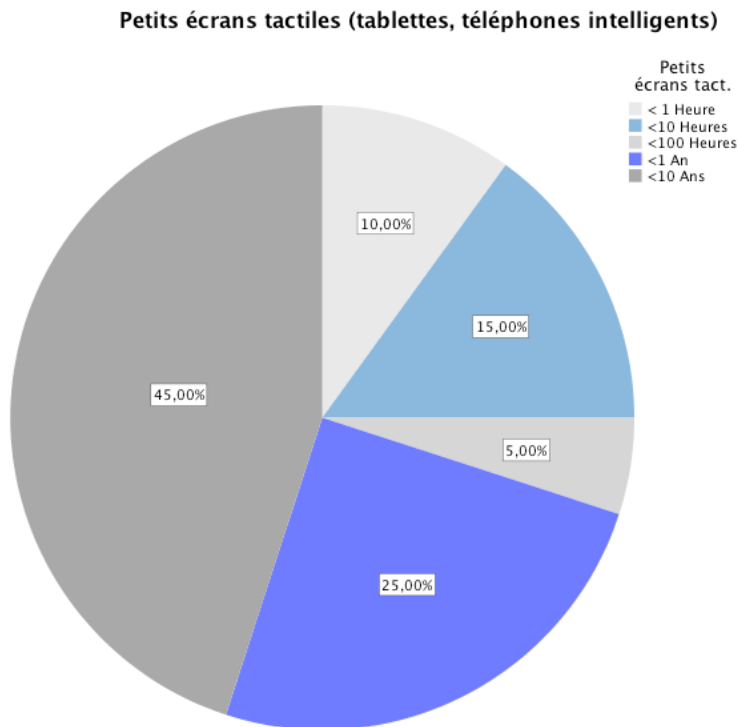


Figure F.2 : Expérience avec les petits écrans tactiles

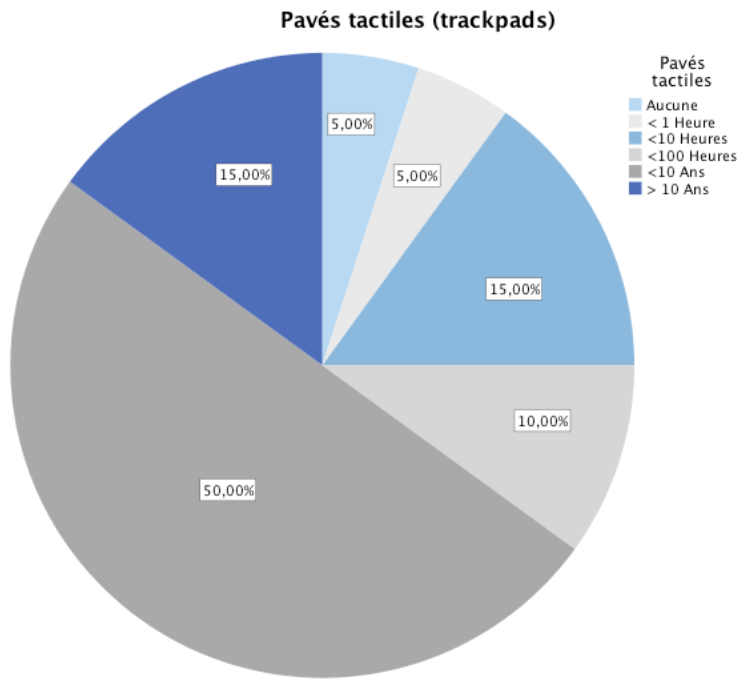


Figure F.3 : Expérience avec les pavés tactiles

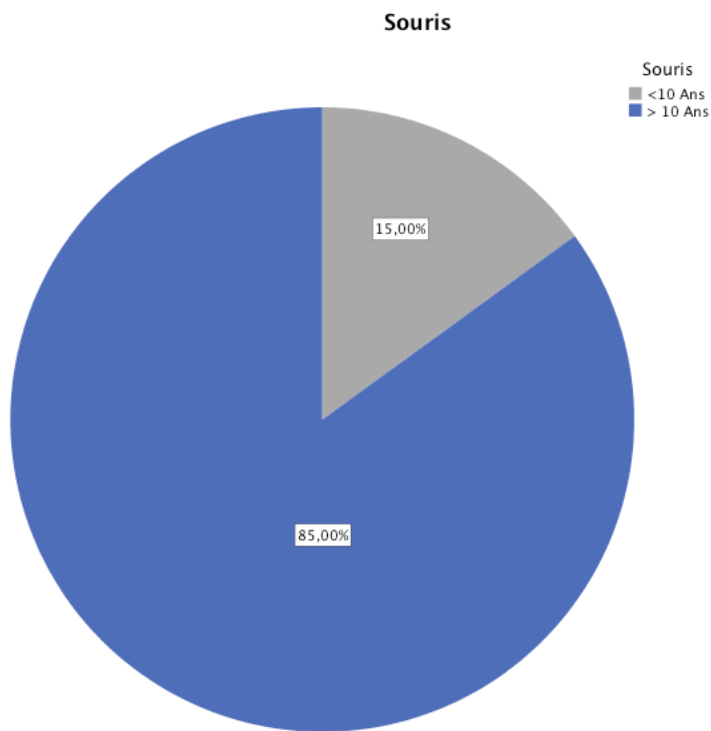


Figure F.4 : Expérience avec la souris

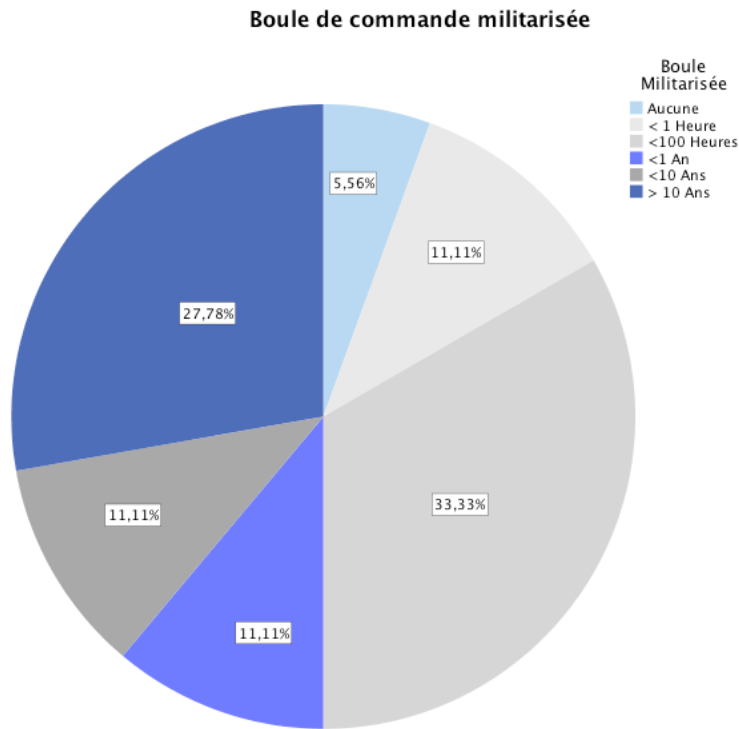


Figure F.5 : Expérience avec les boules de commande militarisées

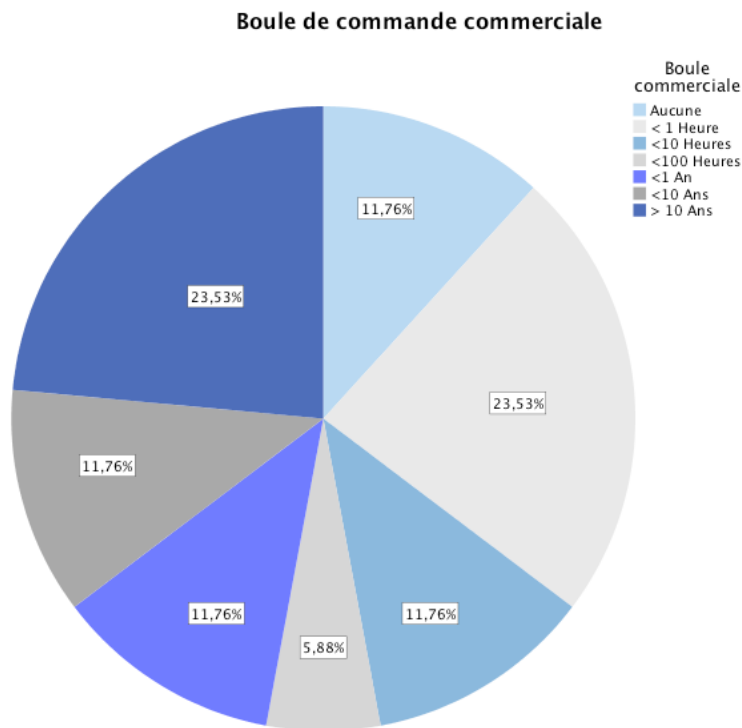


Figure F.6 : Expérience avec les boules de commande commerciales

