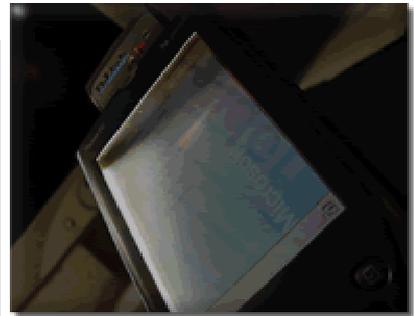
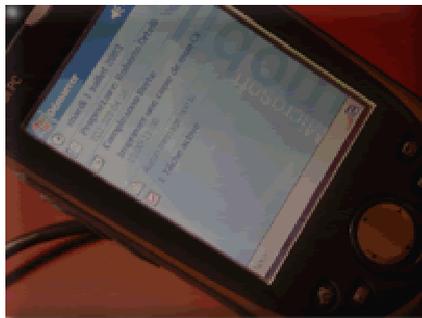


## Styles d'interaction dans les PocketPC: analyses et comparaisons



Roberto Ortelli  
Juillet 2003  
rortelli@ticino.com

### **Jury**

Mireille Bétrancourt, Directeur (TECFA, UniGE)  
Nicolas Nova (CRAFT, EPFL)  
Daniel K. Schneider (TECFA, UniGE)

# Table de matières

<b>REMERCIEMENTS</b>	<b>4</b>
<b>RESUME</b>	<b>5</b>
<b>INTRODUCTION</b>	<b>5</b>
<b>1. CADRE THEORIQUE</b>	<b>8</b>
1.1 LA MANIPULATION DIRECTE	8
1.1.1 LE POINT DE DEPART : LA MANIPULATION DIRECTE SELON BEN SHNEIDERMAN	8
1.2 MODÈLE SSOA D'INTERACTION HOMME-MACHINE	11
1.3 LE MODÈLE DE L'INTERACTION	14
1.3.1 LA DIRECTNESS	14
1.3.2 LA DISTANCE	15
1.4 LE MODÈLE DES NIVEAUX D'ACTIVITÉ DE RASMUSSEN	18
<b>2. LE POCKETPC</b>	<b>21</b>
2.1 L'INTERFACE	21
2.2 LES STYLES D'INTERACTION (SI)	21
2.3 ÉTUDES ET RECHERCHES	26
<b>3. OBJECTIFS DE RECHERCHE</b>	<b>28</b>
3.1 QUESTION DE RECHERCHE ET HYPOTHÈSES	28
3.2 ANALYSE DES PROPRIÉTÉS DES STYLES D'INTERACTION	29
3.3 PARTIE EXPÉRIMENTALE	34
3.3.1 VARIABLES	34
3.3.2 HYPOTHÈSES OPERATIONNELLES	35
3.4 PARTIE QUALITATIVE	37
3.5 MÉTHODE	38
3.5.1 SUJETS	38
3.5.2 LE QUESTIONNAIRE DE RECRUTEMENT	38
3.5.3 MATÉRIEL	38
3.5.4 PROCÉDURE	39
<b>4. RESULTATS</b>	<b>44</b>
4.1 QUESTIONNAIRE EXPERTS	44
4.1.1. GENRE ET ÂGE :	44
4.1.2. TEMPS D'UTILISATION DU PPC	44
4.1.3. DEGRÉ DE COMPÉTENCE SUBJECTIF PPC	44
4.1.4. UTILISATION DU PPC	44
4.1.5. LES STYLES D'INTERACTION	45
4.2 RESULTATS PARTIE QUANTITATIVE	47
4.2.1 TEMPS MOYEN D'EXECUTION DES TÂCHES	47
4.2.2 TÂCHE FERMÉE 1 : ADRESSE	50
4.2.3 TÂCHE FERMÉE 2 : ADRESSE WEB	58
4.2.4 TÂCHE FERMÉE 3 : RENDEZ-VOUS	65

4.2.5 TACHE FERMEE 4 : EMAIL	72
4.2.6 NOMBRE MOYEN D'ERREURS	78
4.2.7 TACHES LIBRES	80
4.2.8 TACHE LIBRE 1 : ADRESSE	81
4.2.9 TACHE LIBRE 2 : RENDEZ-VOUS	86
<b>4.3 REPONSES AU QUESTIONNAIRE POST-SERIE</b>	<b>89</b>
4.3.1 SATISFACTION DU SI UTILISE	90
4.3.2 PREFERENCE D'UN AUTRE SI	91
<b>4.4 RESULTATS PARTIE QUALITATIVE</b>	<b>93</b>
4.4.1 EVALUATION DES PERFORMANCES	93
4.4.2 ANALYSE DES SI	95
<b>5. DISCUSSION DES RESULTATS</b>	<b>98</b>
<b>5.1 H1 : ROLE DE L'EXPERTISE DANS LE CHOIX DES SI</b>	<b>98</b>
<b>5.2 H2 : ROLE DES ERREURS</b>	<b>99</b>
<b>5.3 H3 : RECONNAISSANCE DU SYSTEME</b>	<b>100</b>
<b>5.4 H4, H5 &amp; H6 : TACHES FERMEES VS. TACHES LIBRES</b>	<b>101</b>
<b>5.5 H7 : LES ERREURS</b>	<b>102</b>
<b>5.6 H8 : VITESSE DES PERFORMANCES</b>	<b>103</b>
<b>5.7 H9 : SATISFACTION SUBJECTIVE</b>	<b>104</b>
<b>6. CONCLUSION</b>	<b>105</b>
<b>7. BIBLIOGRAPHIE</b>	<b>107</b>
<b>ANNEXES</b>	<b>109</b>
<b>ANNEXE 1 : QUESTIONNAIRE D'EVALUATION DES SUJETS EXPERTS</b>	<b>110</b>
<b>ANNEXE 2 : QUESTIONNAIRE POST-TACHES FERMEES</b>	<b>116</b>
<b>ANNEXE 3 : FICHES SI</b>	<b>118</b>
<b>ANNEXE 4 : TEXTES DE FAMILIARISATION</b>	<b>123</b>
<b>ANNEXE 5 : FICHE DE BRIEFING</b>	<b>124</b>
<b>ANNEXE 6 : ITEMS TACHES FERMEES</b>	<b>125</b>
<b>ANNEXE 7 : ITEMS DES TACHES LIBRES</b>	<b>141</b>
<b>ANNEXE 8 : ENTRETIEN NOVICES</b>	<b>143</b>
<b>ANNEXE 9 : ENTRETIEN EXPERTS</b>	<b>144</b>
<b>ANNEXE 10 : ANALYSE DES ENTRETIENS</b>	<b>145</b>

## Remerciements

Merci à tous ceux qui m'ont apporté leur soutien durant la préparation et la rédaction de ce mémoire. J'aimerais remercier tout particulièrement Mireille Bétrancourt pour son support, ses conseils et sa patience.

Mes remerciements vont aussi à toute l'équipe de Logitech à Romanel-sur-Morges en particulier à Alexis Richard, Jean-Marc Flueckiger et à John Skeeahan qui m'ont incité à faire une étude sur les PocketPC.

Merci à Daniel.

Merci à mes amis et collègues de TECFA, en particulier à Nicolas (qui m'a toujours stimulé avec ses links), David, Lio, Yves, René, Paul et Elia, sans oublier tout le staff de TECFA et ceux qui sont partis au CRAFT.

Merci aussi à tous ceux qui ont participé gratuitement aux tests ;)

Merci également à ma famille qui m'a soutenu et encouragé au cours de tous mes études et qui a toujours cru en moi.

Merci à ceux qui, pendant mon adolescence, n'imaginaient pas que j'aurais pu arriver là ;)

Encore un grand merci aux administrateurs des forums de discussion de MobileMag, Bpit et PocketPCParadise.

Merci enfin à Nicole, qui a rempli mes journées de travail de joie, tendresse et amour, c'est surtout grâce à toi si je suis là. Merci pour tout ce que tu m'apportes.

## Résumé

Ce mémoire présente les résultats d'une étude quantitative et qualitative se focalisant sur les styles d'interaction (SI) présent par défaut dans les PocketPC. Nous voulons vérifier l'hypothèse générale selon laquelle lorsque l'utilisateur doit choisir le style d'interaction le plus approprié, son choix est guidé par plusieurs facteurs distincts mais reliés entre eux comme le niveau d'expertise avec le PocketPC et la maîtrise des styles d'interaction. Ces deux facteurs sont sensés modifier positivement ou négativement la satisfaction subjective ressentie par l'utilisateur qui en fonction de celle-ci guidera son choix présent et futur. Les résultats obtenus infirment en partie les hypothèses formulées. De manière générale les experts ont de meilleures performances que les novices et connaissent mieux les propriétés de chaque SI ; seulement la moitié des experts n'utilise que des SI de haut niveau, l'autre moitié se fie au clavier virtuel. Quant aux performances, les experts comme les novices ont de meilleurs résultats avec le clavier virtuel suivi par la reconnaissance des blocs. Les deux groupes n'apprécient pas le Trnscriber.

## Introduction

Ces derniers années la technologie a énormément évoluée cherchant de plus en plus à offrir au public toute sorte d'outils compacts offrant un maximum de mobilité et de services. Pensons aux ordinateurs portables ou encore tout récemment les TablePC ou les SmartPhones ainsi que le succès des téléphones portables. Dans ce contexte, les PDA ou *Personal Digital Assistant* (agenda électronique de poche) ont eux aussi évolué ouvrant le marché à de nombreuses marques et systèmes opératifs (Palm, Microsoft CE, Microsoft PocketPC, Linux, ...). Nous avons identifiés plusieurs intérêts aux PDA et tout particulièrement aux PocketPC (PPC) – hébergeant le dernier système opératif de Microsoft. – qui seraient sensés être approfondis.

D'un côté ce système opératif offre à ses utilisateurs plusieurs façons d'interaction, à l'aide d'un stylet comme support intermédiaire. Par conséquent, nous avons estimé qu'il serait intéressant de savoir comment les utilisateurs choisissent une modalité d'interaction plutôt qu'une autre (nous emploierons tout au cours de notre récit le terme de *styles d'interaction* (SI) lorsque nous parlerons des modalités impliqués dans le PPC). Suivant cette idée de départ, nous nous somme aperçus donc que le PCP permet une grande liberté et flexibilité à l'utilisateur pour choisir le SI ou la combinaison de SI afin d'achever différentes tâches. Cette liberté dans le choix des SI à utiliser détermine aussi les préférences et les acceptations du système (Calvet, Kahn, Zouinar, Salembier, Briois, Nigay, Rey & Pasqualetti,

2001). Le PPC, comme d'ailleurs les ordinateurs classiques, permettent instantanément le repérage des erreurs et en situations conflictuelles ou d'incompréhension entre la relation homme-machine avec une modalité, l'utilisateur a, à tout moment, la possibilité de la changer, influençant, nous pensons, son attitude future vis-à-vis du SI utilisé, ainsi que la situation contextuelle présente lors du changement d'attitude intervenu.

La particularité de notre étude est d'analyser les 4 styles d'interaction du PocketPC (PPC) par le biais d'une procédure expérimentale visant à évaluer d'un côté les performances et de l'autre côté la satisfaction subjective des utilisateurs ayant ou pas une expertise du PPC; de plus, notre méthodologie comporte aussi une partie d'entretiens semi-directifs nous aidant à mieux cerner des aspects que des tests quantitatifs ne le permettent pas. En effet, plusieurs études (MacKenzie et Zhang, 1997 ; Sears et Agora, 2002 ; Frankish, Hull, Morgan, 1995) ont été menées comparant seulement deux SI à la fois, ou bien des SI qui ne sont pas présents dans les PocketPC ; nous nous sommes donc inspirés de ces études pour développer nos hypothèses et attentes. Bien entendu nous nous sommes aussi appuyés sur d'autres concepts et modèles clés comme la manipulation directe (Shneiderman, 1992, Hutchins, Holland, Norman, 1986, Norman, 1986, Frohlich, 1997), les modèles des niveaux d'activité de Rasmussen (1986) et partiellement par la notion de multimodalité (Vigouroux, N., Poirier, F., Nigay, L., 1996, Coutaz, 1996, Calvet *et al.*, 2001).

Un but sous-jacent de notre travail est celui de rendre l'interaction homme-machine plus efficace, robuste et plus flexible comme le souligne le travail de Oviatt (1999) en apportant des éléments supplémentaires quant à l'usage réel d'utilisateurs des PPC, sans oublier le souci de comprendre et d'observer le comportement de sujets novices. Pour ce faire nous nous sommes posés les questions suivantes: Qu'est ce qui guide le choix de l'usage d'un SI ? Qu'est ce qui guide à utiliser une modalité plutôt qu'une autre ? Quel est le rôle de l'expérience ? Quel est le rôle joué par les erreurs ? Qu'est ce qui guide l'interaction et les intentions de l'utilisateur ? Quel est le rôle des propriétés des modalités, de l'utilisateur, de la tâche..., des choix? Est-ce que des SI sont mieux adaptés selon le niveau de compétence de l'utilisateur ? ...

Nous pensons que les réponses à ces questions pourraient nous permettre de mieux encadrer des études futures dans le domaine des « Embodied User Interfaces » ou « Embedded Physically ». Ces études (Calvet *et al.* ; 2001, Fishkin, Moran, and Beverly & Harrison , 1998 ; Harrison, Fishkin, Gujar, Mochon, & Want, R , 1998 ; Hynckley, Pierce, Sinclair & Horvitz, 2000) tentent d'analyser l'interaction alternée ou simultanées de plusieurs SI. De nouvelles modalités, expérimentales, ont été aussi testées, comme par exemple certaines qui permettent d'affecter la réaction du système en fonction de la

position ou de l'inclinaison du PocketPC. Cette nouvelle modalité qui pourrait être implémentée dans le futur permettrait par exemple de visualiser sur le PDA un texte en fonction de l'inclinaison donnée au PDA. Si par exemple l'utilisateur bascule le PDA avec un geste décidé du bas vers le haut, le système pourrait l'interpréter comme une demande d'indiquer au système de tourner la page du document ouvert.

La décision de ne pas aborder une étude dans ce sens, réside dans le fait que nous pensons qu'avant d'étudier ces nouvelles formes de modalités il faut d'abord connaître plus en profondeur l'interaction homme-machine réelle avec les modalités présentes actuellement sur la marché. Par la suite, ces études, dont la nôtre pourrait faire partie, pourraient permettre justement d'apporter des éléments supplémentaires à ces nouvelles fonctionnalités améliorant l'expérience homme-machine au niveau de leur flexibilité et robustesse.

Dans la partie théorique, nous allons d'abord expliquer trois modèles généraux d'interaction homme-machine : le modèle SSOA de Shneiderman, le modèle des niveaux d'activité de Hutchins *et al.*, et le modèle des niveaux d'activité de Rasmussen. Avant de présenter ce modèles, nous nous arrêterons quelques instants dans la définition de la manipulation directe. Ensuite nous présenterons plus en détail le PocketPC et les styles d'interaction étudiés dans ce mémoire. Dans la troisième partie nous présentons les objectifs de l'étude ainsi qu'une analyse détaillée des 4 SI, pour ensuite expliquer la méthode utilisée pour la partie quantitative et la partie qualitative. Nous présentons les résultats et nous les commentons dans les partie 4 et 5 pour ensuite dresser les conclusions.

## 1. Cadre théorique

Dans le cadre de notre étude, nous voulons analyser le comportement de l'utilisateur confronté à un choix entre 4 styles d'interaction<sup>1</sup> (SI). Dans des situations réelles des utilisateurs expérimentés ou novices sont confrontés au choix entre ces 4 SI, présents dans le PPC, pour interagir avec le système. Or, nous devons d'abord connaître quelques modèles nous permettant de mieux comprendre les processus déterminant ce choix. Pour ce faire, nous nous appuyons sur trois modèles : le modèle d'interaction développé par Hutchins *et al.* (1986), le modèle SSOA de Shneiderman (1992) et le modèle des niveau d'activité de Rasmussen (1986).

Avant d'entrer dans le vif du sujet nous devons faire un détour en introduisant d'abord la notion de la manipulation directe introduite par Ben Shneiderman (1992). L'éclairage de cette notion nous permettra par la suite de mieux cerner les concepts et les modèles qui suivront.

### 1.1 La manipulation directe

#### 1.1.1 Le point de départ : la manipulation directe selon Ben Shneiderman

Pour montrer ce qu'est la manipulation directe, nous pouvons reprendre l'exemple préféré de Shneiderman (1992, p.183) : la conduite automobile. Lorsqu'on conduit une voiture on a directement devant nous la scène visible. Le fait d'accélérer ou de freiner est une connaissance acquise dans notre culture et retenue par l'expérience sensori-motrice. Pour virer à gauche ou à droite, il suffit de tourner le volant dans la direction souhaitée. La réponse est immédiate et on voit devant nous le résultat de l'action entreprise par l'acteur. L'acteur a ainsi un feedback immédiat lui permettant d'ajuster l'action.

Dans l'exemple ci-dessus, nous pouvons déjà observer quelques propriétés de la manipulation directe que nous détaillerons ci-après.

Les concepts de la manipulation directe sont nés suite au changement de paradigme dans les interfaces informatiques. Si au début on manipulait les interfaces par l'introduction d'instructions écrites (comme on le fait encore en partie maintenant sous les systèmes basés sur Unix par exemple), l'avènement de l'interface de Xerox Star a impliqué un changement radical du paradigme donnant ainsi la naissance des

---

<sup>1</sup> La reconnaissance de blocs (RB), la reconnaissance des lettres (RL), le clavier virtuel (CV) et le Transcriber (TR)

théories liées à la manipulation directe. Avec ces nouvelles interfaces, l'utilisateur ne commande plus les systèmes avec des commandes écrites mais agit directement sur les éléments présents à l'écran (avec une souris par exemple). Le paradigme a donc changé en passant du « dialogue » à la « manipulation », des commandes écrites aux actions manuelles.

A ce propos Hutchins *et al.* (1986) introduisent la notion d'*engagement*. L'engagement est une sensation qualitative que l'on perçoit lors de la manipulation directe dans une interface. Ce concept se réfère à la perception du contrôle par rapport au système, c'est-à-dire si l'utilisateur estime être l'acteur principal ou pas. Dans les systèmes de commandes, il y a un intermédiaire caché qui effectue la tâche demandée sous forme linguistique (Figure 1). On appelle cette interaction « engagement indirect » (Frohlich, 1996, p.4). Par contre dans des systèmes basés sur des interactions « graphiques », on observe un « engagement direct » (Figure 2). C'est l'utilisateur qui manipule directement des objets.

Par exemple, si on veut effacer un document, dans le premier cas de figure on tape quelque chose qui peut ressembler à « efface mon fichier », dans le deuxième cas de figure on sélectionne le fichier et on le transporte jusqu'à la poubelle.

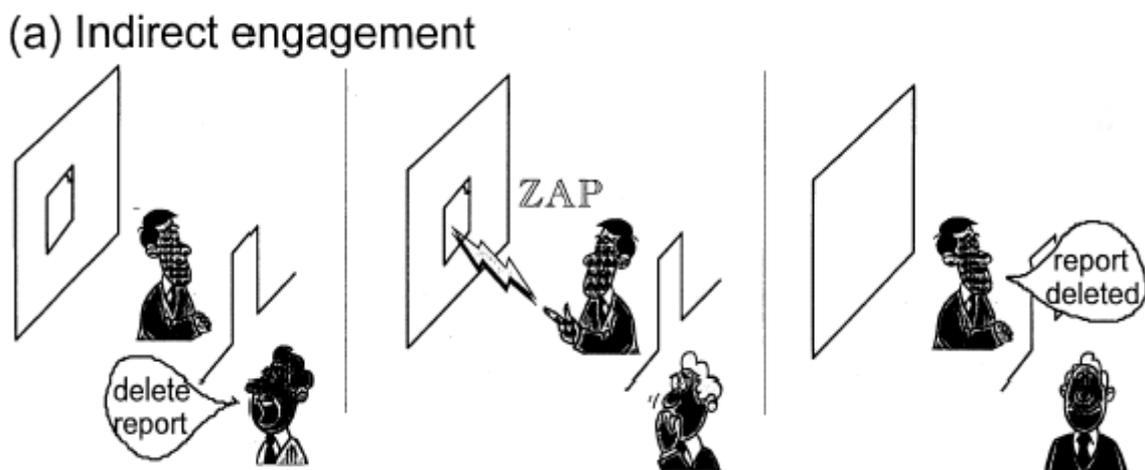


Figure 1: Engagement indirect d'après Frohlich (1996)

## (b) Direct engagement



Figure 2: Engagement direct d'après Frohlich (1996)

La première chose qu'on pense d'une interface est sa facilité d'usage mais nous ne devons pas oublier un élément central pour la manipulation directe. L'analyse du processus d'exécution d'une tâche et l'évaluation du processus permettent d'expliquer les raisons des difficultés dans l'usage d'un système. Il indique aussi sur quels éléments il faut agir pour réduire les efforts cognitifs nécessaires pour interagir avec le système.

Dans son livre, Shneiderman (1992) indique trois propriétés communes à la manipulation directe (p. 205) :

1. Représentation permanente des objets et des actions concernés :

Par exemple, dans les traitements de textes ont les propriétés dites WYSIWYG<sup>2</sup>. Les modifications sur la page peuvent être effectuées maintenant avec différents moyens et ces modifications sont directement affichées à l'écran.

2. Utilisation d'actions directes ou de pressage de boutons étiquetés au lieu d'une syntaxe compliquée :

Si on veut modifier le style d'un mot il suffit d'appuyer sur un bouton et on voit directement le résultat (par exemple le bouton gras ou l'on peut utiliser un raccourci clavier ou encore passer par la fenêtre du format du caractère). L'utilisateur peut maintenant ouvrir un fichier, l'imprimer, l'effacer simplement en agissant manuellement sur des icônes.

---

<sup>2</sup> WYSIWYG signifie en anglais « what you see is what you get ». Signifiant que ce que l'on fait on le voit directement à l'écran ou inversement.

3. Opérations réversibles et incrémentales dont l'impact sur les objets concernés est immédiatement visible :

L'utilisateur a toujours une représentation visuelle de ce qu'il est en train d'effectuer. Par conséquent il peut aussi faire « machine arrière » pour revenir à un état précédant.

A partir de ces propriétés, Shneiderman (1992) indique qu'en utilisant ces trois principes, il est possible de concevoir des systèmes qui ont ces attributs bénéfiques :

- Les personnes inexpérimentées à une interface peuvent l'apprendre vite, simplement par le biais d'une démonstration d'une personne expérimentée.
- Les personnes expérimentées peuvent travailler vite en menant à bien différentes tâches ; ils arrivent aussi à définir de nouvelles fonctions et de nouveaux outils.
- Les utilisateurs intermédiaires peuvent retenir des concepts opérationnels.
- Les messages d'erreur sont rarement utiles.
- Les utilisateurs peuvent voir immédiatement si leur action a atteint le but fixé ou pas ; si non, alors ils peuvent modifier facilement la direction de leur activité.
- Les utilisateurs sont moins anxieux car le système est compréhensible et leurs actions sont facilement réversibles.
- Les utilisateurs acquièrent de la confiance et un savoir-faire car ils sont les initiateurs de l'action, ils perçoivent le contrôle et les réponses du système sont prévisibles.

Notons que Frohlic (1996) indique que ces éléments doivent être retenus avec prudence car il s'agit d'une représentation idéalisée du potentiel offert par ces interfaces manipulables.

## **1.2 Modèle SSOA d'interaction homme-machine**

Shneiderman différencie le niveau syntaxique du niveau sémantique afin de décrire la manipulation directe. Parmi le niveau sémantique, il distingue les concepts propres à la tâche et les concepts liés au traitement informatique (Figure 3).

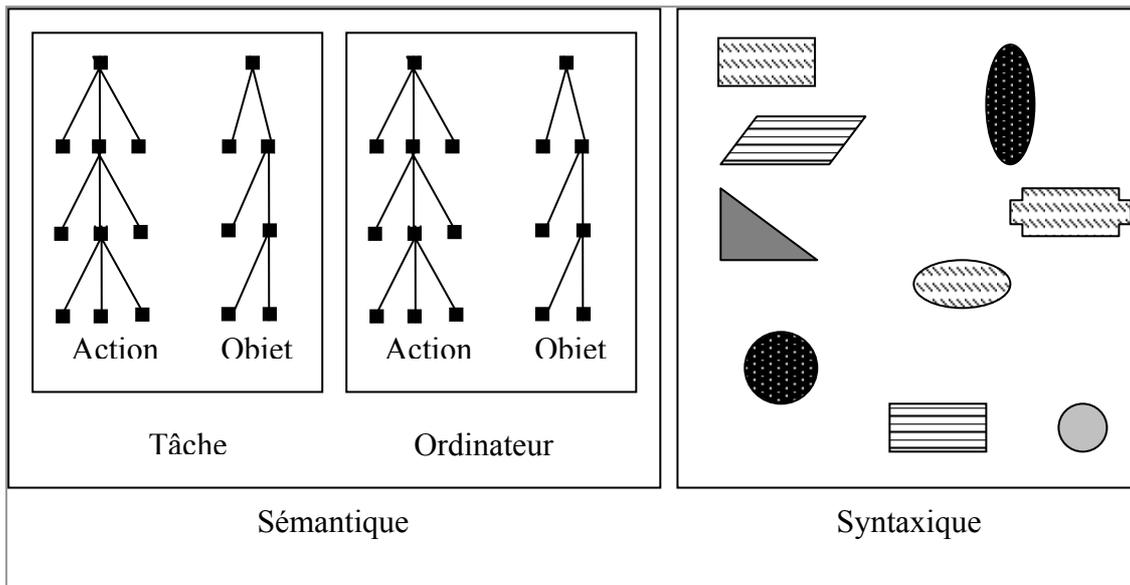


Figure 3: Le modèle SSOA de Shneiderman (1992)

L'exemple du traitement de texte permet de mieux cadrer le modèle SSOA (*Syntactic-Semantic object-action*):

Le sémantique se réfère premièrement aux concepts liés à la tâche, comme par exemple convaincre, résumer, orthographe, paragraphe..., et aux concepts liés à l'ordinateur, fichier, style, paragraphe, glossaire, remplacement... et ces connaissances requièrent un apprentissage important. Ces connaissances sont généralement stables et leur apprentissage permet de construire un réseau de concepts hiérarchisés.

Observons qu'à ce niveau, le concept de paragraphe se situe aux deux niveaux, de la tâche et de l'ordinateur. En effet, ce concept diffère si il est traité sur papier ou sur un traitement texte. Pour le traitement texte la notion de paragraphe engendre plusieurs propriétés comme le style et l'espacement par rapport aux autres paragraphes. Dans la modalité « papier », le paragraphe n'est qu'une succession de lignes vides. Les utilisateurs qui utilisent la méthode « papier » dans un traitement texte (comme séparer deux paragraphes par des lignes vides, plutôt que de paramétrer correctement la distance entre les paragraphes), ont acquis sémantiquement les concepts propres à la tâche mais ont des difficultés de représentation au niveau de l'ordinateur.

Le syntaxique, par contre, est caractérisé par des connaissances arbitraires qui sont spécifiques au système. Si ces connaissances ne sont pas utilisées régulièrement elles seront vite oubliées. Par exemple à ce niveau nous pouvons faire référence à des concepts comme effacer, taper DELETE ou utiliser le

raccourci clavier Ctrl-X, ou justifier à droite, cliquer sur l'icône correspondante ou utiliser son raccourci clavier.

Un autre exemple, celui des « contacts », dans l'environnement PocketPC:

Imaginons qu'un utilisateur doit rajouter un contact dans son PDA. Au niveau sémantique et syntaxique nous aurons les éléments suivants :

- Niveau sémantique – Tâche : connaissances des éléments propres à un contact (par exemple les éléments composant une carte de visite ;
- Niveau sémantique – Ordinateur : insérer, éditer, effacer un contact ; utilisation du vocabulaire du PocketPC, de signets, de notes... ;
- Niveau syntaxique : connaissances des propriétés des styles d'interaction et des aides pour effectuer la tâche.

Dans notre étude nous nous concentrerons sur ce dernier niveau afin de mieux comprendre d'une part quel sont les styles d'interaction les plus adaptés et performants pour les novices et les expérimentés et d'un autre côté mieux comprendre le comportement des utilisateurs lorsqu'ils manipulent ces styles d'interaction. Le modèle ajoute d'autres éléments à la manipulations directe ; afin de rendre encore plus directe une interface, il faut que l'utilisateur développe aussi des connaissances au niveau sémantique et syntaxique.

Notons aussi que Shneiderman distingue trois types d'utilisateurs confrontés à une interface:

1. le novice : étant celui qui utilise pour la première fois le logiciel. Il a des concepts de la tâche, normalement pas de concepts liés à l'ordinateur et aux connaissances syntaxiques ;
2. l'intermédiaire : il a acquis les connaissances sémantiques mais il n'a pas de connaissances syntaxiques ;
3. l'expert : il manie les connaissances sémantiques et syntaxiques.

Le prochain modèle nous permet de mieux comprendre l'interaction entre l'homme et la machine, les processus qui entrent en jeu durant la planification, l'exécution et l'évaluation d'une tâche.

### 1.3 Le modèle de l'interaction

Avant de décrire le modèle développé par Hutchins *et al.* (1986) nous devons encore une fois introduire d'autres concepts qui lui sont intimement liés.

Hutchins *et al.* (1986) reprennent les caractéristiques de la manipulation directe définie par Shneiderman (1982) pour la compléter et introduire d'autres concepts clé. Mais qu'est ce qui fait qu'une interface soit *directe* ? La réponse à cette question cruciale n'est pas aussi simple, car elle demeure plutôt abstraite. Les auteurs se concentrent premièrement sur comment l'utilisateur réfléchi par rapport au rôle de l'ordinateur dans l'interaction et deuxièmement à l'effort nécessaire pour exécuter un but. Les auteurs nomment respectivement ces deux éléments *engagement* (que nous avons décrit plus haut) et *distance* dont la résultante est la *directness* (Frohlich, 1996, p. 4).

#### 1.3.1 La *directness*

La manipulation directe n'est pas un concept unitaire mais au contraire il engendre d'autres éléments tels que les concepts de *directness*<sup>3</sup>. La *directness* est « une impression ou une sensation » (Hutchins *et al.*, 1986, p.93) que l'on perçoit lorsqu'on est confronté à une interface. La sensation de *directness* peut aussi être perçue lorsqu'il y a le « besoin d'allouer d'autres ressources cognitives dans l'usage d'une interface » (Hutchins *et al.*, 1989, p.93). La *directness* est dégagée lorsque plusieurs ressources cognitives sont impliquées lors de l'interaction de l'utilisateur avec une interface donnée. En d'autres termes, la *directness* est le sentiment de manipuler directement les objets du domaine concerné et non pas des concepts abstrait ou symboliques.

La production de cette sensation dépend de l'adaptation de l'utilisateur. Plus l'utilisateur s'adapte et s'expérimente sur l'interface et plus il pourra percevoir cette sensation.

La sensation de *directness* est toujours relative. On ne peut pas dire qu'une interface est plus directe qu'une autre ou qui est directe tout court. C'est souvent l'interaction de plusieurs facteurs (Hutchins *et al.*, 1986, p. 93). Tous ces facteurs augmentent la sensation de *directness* impliquant des coûts au niveau cognitif.

---

<sup>3</sup> Nous n'avons pas trouvé une traduction cohérente pour ce terme.

Comme nous l'avons souligné en précédente la *directness* est le produit de deux facteurs: la distance et l'engagement. Le prochain sous-chapitre introduira le concept manquant pour la compréhension de leur modèle : la distance.

### 1.3.2 La distance

La notion de distance se réfère au efforts mentaux qui sont nécessaires pour traduire en action les buts que l'utilisateur s'est fixé et à l'évaluation des effets de l'action. Plus cette distance sera courte et plus les idées et les intentions seront facilement traduites en actions physiques ; en outre, plus la distance sera courte et plus le résultat de l'action sera facilement interprétable.

Une interface introduit de la distance dans la mesure où il y a un écart entre les buts et la connaissance d'une personne et le niveau de description fourni par le système avec lequel cette personne a à faire.

Les auteurs définissent deux types de fossés créés par la distance (Figure 4) : le fossé d'exécution et le fossé d'évaluation. Le fossé d'exécution est franchi par le biais de commandes et des mécanismes du système qui correspondent avec les buts et les intentions de l'utilisateur. Par exemple, dans le cas du fossé d'évaluation, le pont est franchi par l'affichage du résultat ce qui demeure être un bon modèle conceptuel du système étant facilement interprétable et évaluable.

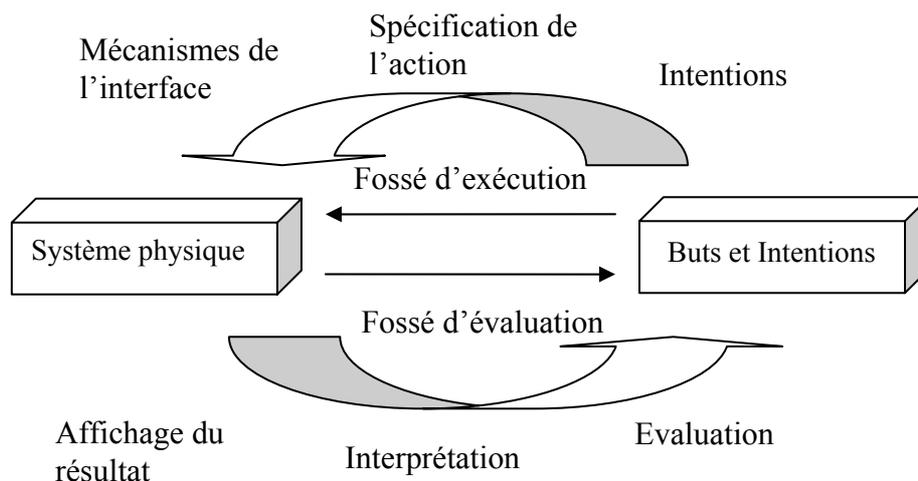


Figure 4: Les fossés d'exécution et d'évaluation

Plus les écarts à l'exécution et à l'évaluation seront grands, plus la charge cognitive impliquée sera forte et par conséquent moins le système sera direct. Nous pouvons donc affirmer que :

- la sensation de *directness* est inversement proportionnelle à la totalité des efforts cognitifs qui sont nécessaires pour manipuler et évaluer le système ;
- l'effort cognitif étant le résultat direct des fossés d'exécution et d'évaluation ;
- les meilleures interfaces permettent de faire le pont le plus direct avec le moindre effort cognitif, ayant comme résultat une sensation d'interaction.

Le modèle de Hutchins *et al.* (1986)

Dans la figure ci-dessous (Figure 5), nous retrouvons les éléments que nous venons de décrire. Dans ce modèle nous retrouvons aussi les sept étapes de l'activité de l'utilisateur :

1. Etablir un but
2. Formaliser une intention
3. Spécifier l'action en différentes séquences
4. Exécuter l'action planifiée
5. Percevoir l'état du système
6. Interpréter le résultat du système
7. Evaluer le résultat du système avec les buts et les intentions préalables

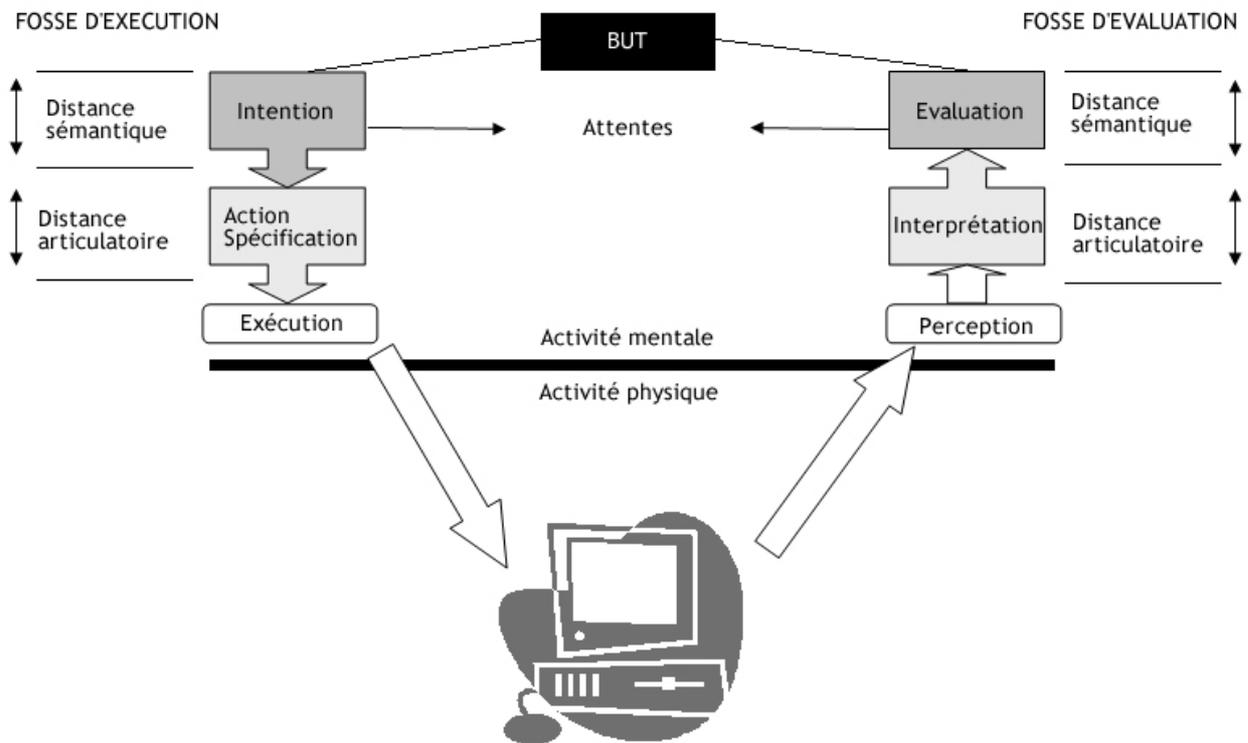


Figure 5: le modèle d'interaction de Hutchins *et al.* (1986)

Nous pouvons aussi observer que dans le fossé d'évaluation et dans le fossé d'exécution, les auteurs distinguent deux types de distances : sémantique et articulatoire. Le premier concernant la mise en place des intentions pour affronter un problème ou concernant l'évaluation du résultat de l'action avec les intentions initiales ; le deuxième affectant la mise en place « théorique » des séquence d'actions à effectuer ou affectant l'interprétation du résultat affiché par le système.

L'intérêt de l'utilisation de ce modèle réside dans le fait qu'il montre de manière concrète l'action de l'utilisateur ainsi que toutes les phase précédant et suivant l'action même.

Le modèle suivant, celui des niveaux d'activité de Rasmussen (1986), nous permettra de le mettre en relation avec les modèles de Hutchins *et al.* (1986) et de Shneiderman (1992) pour les compléter et pour expliquer le changement de comportement et d'attitude des utilisateurs lorsqu'ils ont à disposition plusieurs méthode de résolution de problèmes.

## **1.4 Le modèle des niveaux d'activité de Rasmussen**

Rasmussen (1986) est l'auteur d'un modèle qui nous permettra par la suite d'expliquer d'un côté le comportement de nos utilisateurs lors de l'expérimentation et de l'autre côté de compléter le modèle de l'interaction de Hutchins *et al.* (1986) et de Shneiderman (1992).

Le modèle de Rasmussen définit trois niveaux :

1. Le comportement basé sur les habilités (« skill-based behavior ») : est aussi appelé niveau des réflexes. Il représente les actions sensori-motrices effectuées par un opérateur suite à une intention mais qui se déroulent sans réel contrôle conscient. C'est le niveau le plus bas, engendrant le moindre effort cognitif. C'est aussi un niveau qui peut engendrer des comportements extrêmement rapides, efficaces et automatiques.
2. Le comportement basé sur les règles (« rule-based behavior ») : est aussi appelé niveau des procédés. Selon Rasmussen, « L'activité est dirigée par un but, mais elle est structurée par un contrôle proactif, grâce à une règle stockée. Très souvent, le but n'est même pas formulé explicitement, mais on le trouve implicitement dans la situation, et il libère les règles stockées. Le contrôle est téléologique, en ce sens que la règle ou le contrôle sont sélectionnés sur la base d'expériences antérieurement réussies. Le contrôle évolue selon le principe de la priorité à la règle la plus adaptée ». Par conséquent, les comportements de ce niveau fonctionnent sur des règles pré-stockées par la situation présente. A différence du niveau précédent, ce type de comportement est conscient, et l'opérateur comme l'environnement fournissent les paramètres de la règle. L'opérateur n'a pas besoin d'avoir déjà vécu une situation identique, mais il utilise une règle (ou schéma en termes Piagetiens) qui est valable dans une situation analogue. Il peut avoir vécu cette situation, l'avoir apprise, ou encore l'avoir planifiée au niveau le plus haut. Le choix de la règle à appliquer utilise le principe de la priorité à la règle la plus adaptée selon les données sur la situation. La limite entre les niveaux des réflexes et des procédés n'est pas nettement définie, et dépend de l'entraînement de l'opérateur tout comme de son attention. En règle générale, il est possible de faire la distinction selon que l'opérateur peut ou non rapporter sans difficulté l'ensemble des actions qu'il a effectuées.
3. Le comportement basé sur les connaissances (« knowledge-based behavior ») : est aussi appelé niveau du savoir. Ce niveau de comportement, le plus haut cognitivement, est utilisé lorsque qu'aucun réflexe ou aucune règle ne convient à la situation rencontrée par l'utilisateur (deux

niveaux précédents). Le sujet se trouve confronté à un processus de résolution de problème. Ce comportement est utilisé lorsque l'opérateur est confronté à des situations non familières ou bien totalement nouvelles (aucun schéma familier ne peut être appliqué). Pour ces situations, le but est clairement exprimé par l'opérateur en fonction de son analyse de l'environnement et par ses objectifs. Les traitements effectués à ce niveau sont conscients, lents et séquentiels. À ce niveau de raisonnement, la structure du comportement de l'opérateur s'appuie sur un modèle mental de la situation. Selon Rasmussen, ce type de comportement pourrait aussi être appelé "model-based behavior", c'est-à-dire comportement basé sur les modèles. Le sujet peut définir son plan d'action soit en utilisant une méthode d'essai-erreur, soit en utilisant une méthode plus conceptuelle par la compréhension des propriétés fonctionnelles de l'environnement.

Selon le modèle de Rasmussen (Figure 6), le comportement humain utilise de manière simultanée ces trois niveaux de comportement. Le niveau du savoir, le plus haut, définit les buts tandis que les deux autres se chargent de les satisfaire en fonction de leurs compétences. Par exemple, lorsque l'on cherche son chemin en voiture dans une ville inconnue, le niveau du savoir lit la carte et essaye de planifier une route. S'il faut faire demi-tour, c'est le niveau des procédés qui déterminera la procédure à utiliser en fonction de l'environnement, tandis que le niveau réflexes s'occupera de changer les vitesses, de contrôler la trajectoire du véhicule, et de surveiller le trafic.

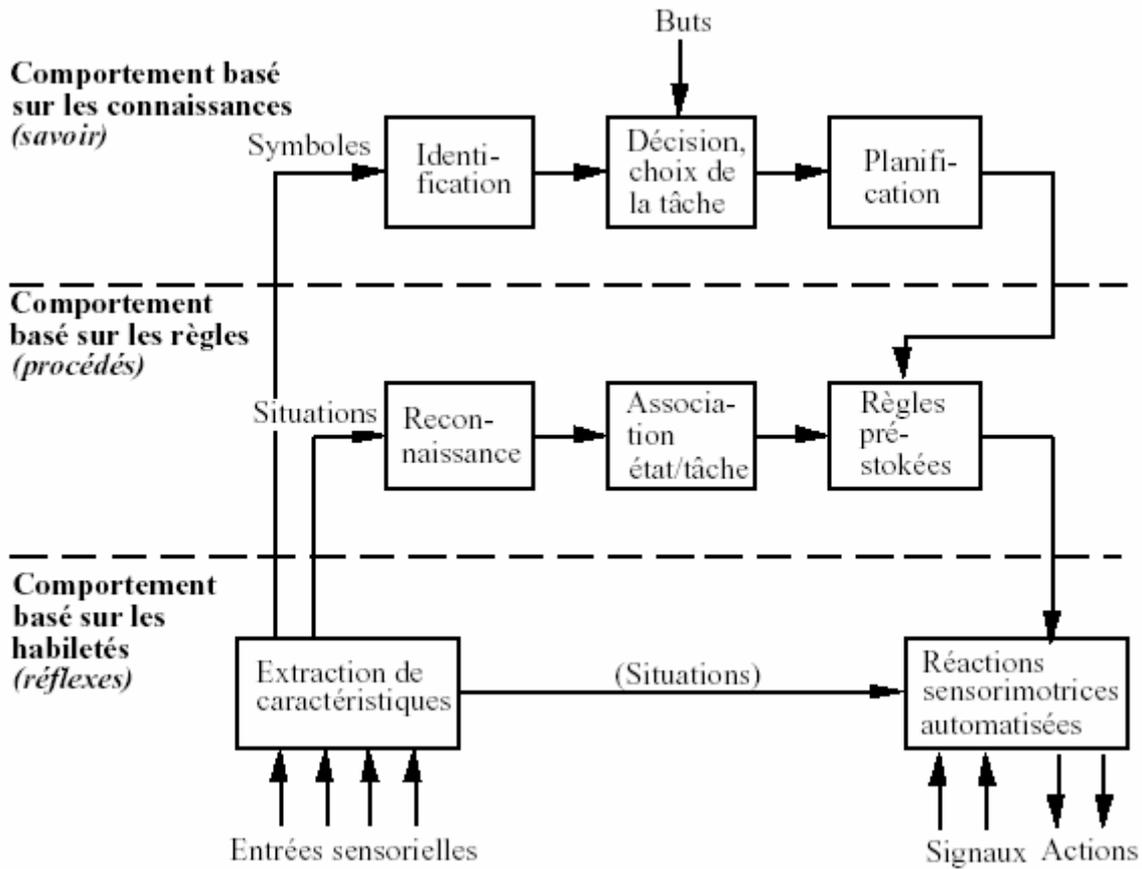


Figure 6: Modèle simplifié de Rasmussen

Pour terminer avec les concepts et les modèles théoriques utilisés dans le développement de notre travail, nous souhaitons aborder brièvement le sujet de la multimodalité qui pourrait sembler être primordiale dans notre étude mais que nous ne le considérons autant que tel pour des raisons que nous expliciterons par la suite.

Nous avons choisi de décrire ces trois modèles car nous estimons qu'ils sont complémentaires. Si nous reprenons la figure représentant le modèle de l'interaction (page 17), nous pouvons considérer que les connaissances sémantiques et syntaxiques décrites par Shneiderman entrent en jeu pendant la phase de la planification de l'action et pendant sa mise en œuvre. En effet, plus l'utilisateur a appris ou découvert ces connaissances, et moins la distance sémantique et articulatoire sera grande. Le modèle de Rasmussen, par contre, se situe dans le modèle d'interaction au niveau de l'évaluation de la tâche effectuée. En fonction du résultat de la tâche, l'utilisateur changera de niveau de comportement.

## 2. Le PocketPC

Le PocketPC rentre dans la famille des PDA, *Personal Digital Assistant*, intégrant un système opératif (OS) développé par Microsoft. De manière générale, il y a deux grands système opératifs qui dominent le marché des ordinateurs de poche : PALM et PocketPC. D'autres OS sont présents dans les magasins mais leur succès est inférieur aux deux cités ci-dessus.

Des statistiques récentes nous montrent que le marché des ordinateurs de poche voit une concurrence entre les différents OS. Il y a trois ans, Palm détenait le 70% du marché alors qu'aujourd'hui Palm ne contrôle plus que le 50,7% tandis que PocketPC a grimpé de 20% alors qu'Epoc en compte 13,5%. D'autres systèmes opératifs voient le jour, comme Linux Light.

### 2.1 L'interface

L'interface des PocketPC se base sur la métaphore du bureau comme celle qu'on retrouve dans les systèmes Windows (ou Mac). Les PocketPC se basent donc sur le concept du WIMP (*Windows, Icons, Menus and Pointing*) qui avait été développé au en 1982, dont le nom était Xerox Star. Cette interface était une nouveauté car elle permettait, par le biais d'une interface graphique et à l'application des principes de la manipulation directe, de rendre l'ordinateur accessible facilement à un plus grand public en offrant ainsi nombreuses applications (Smith, D.C., Irby, C., Kimball, R., Verplank, B., Harslem, E., 1982).

### 2.2 Les styles d'interaction (SI)

Pour écrire sur le PocketPC, le système met à disposition quatre outils :

1. Clavier virtuel (CV) : Il s'agit d'un clavier virtuel alphanumérique. Avec le stylet on choisit les lettres, les chiffres ou les symboles désirés.

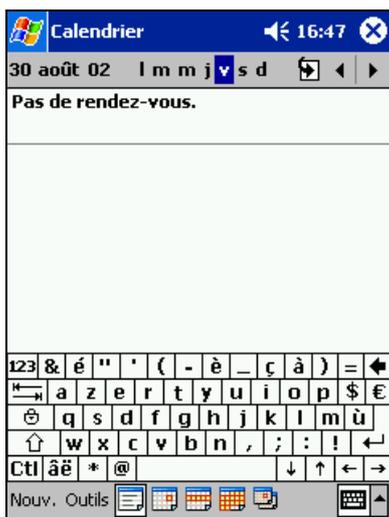


Figure 7: Clavier virtuel alphanumérique

2. Reconnaissance des blocs (RB) : La reconnaissance des blocs se base sur une écriture prédéfinie (la même que « Graffiti<sup>4</sup> » pour Palm) que l'utilisateur doit apprendre pour l'utiliser. L'insertion des symboles s'effectue dans une zone spéciale divisée principalement en deux parties : la zone de gauche permet de rentrer des lettres, celle de droite des chiffres.

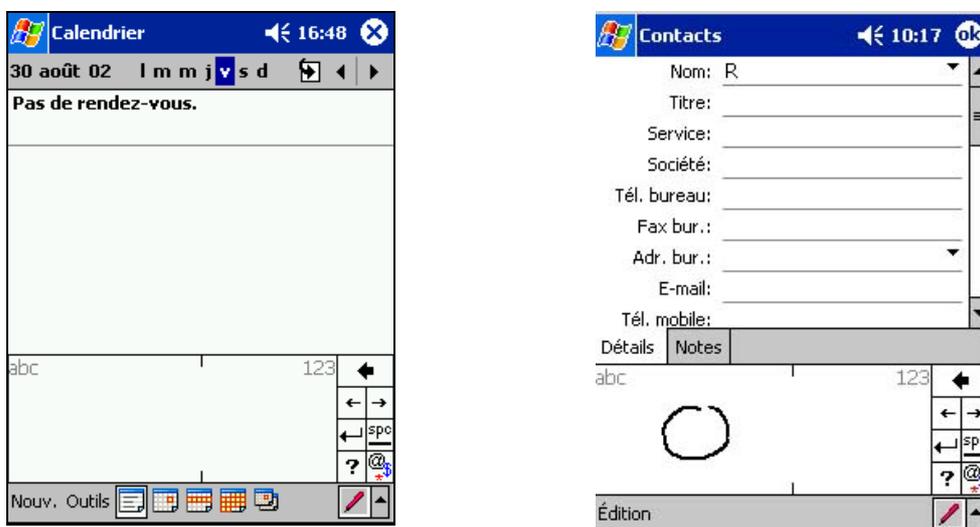


Figure 8: Reconnaissance des blocs

Voici quelques exemples d'écriture avec la reconnaissance des blocs qui utilise un alphabet dont l'écriture est simplifiée :

<sup>4</sup> Dans le système PocketPC, la reconnaissance des blocs est en fait un clone de « Graffiti » utilisé dans le système Palm. Précisons que l'alphabet « Graffiti » appartient au domaine public alors les droits sur ce langage se basent seulement sur le moteur qui encode l'écriture (Mitrani, 2002). Pour cette raison nous allons parler de Graffiti entre guillemets.

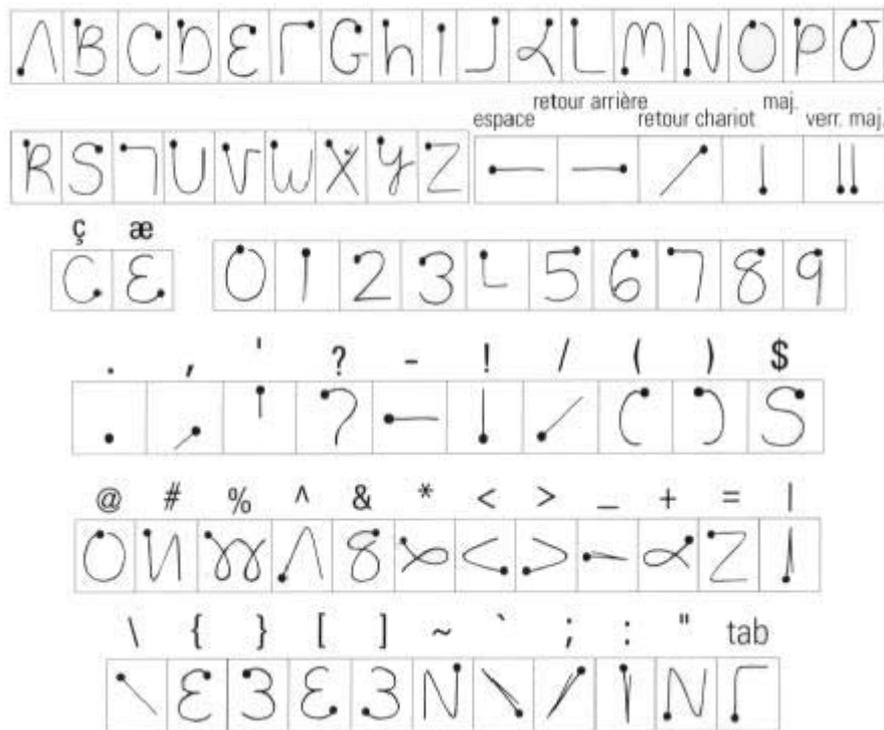


Figure 9: La Reconnaissance des blocs par lettre, chiffres et ponctuation

La reconnaissance des blocs ne se limite pas à la simple écriture alphanumérique mais au contraire elle permet d'effectuer aussi des commandes avancées comme copier, coller, déplacer le curseur à gauche ou à droite, aller à la ligne ... avec des commandes spécifiques et communes à tous les PDA (Palm, PocketPC, Epos, ...) muni de « Graffiti ».

3. Reconnaissance des lettres (RL): La reconnaissance des lettres permet à l'utilisateur d'insérer des lettres et des chiffres sans suivre nécessairement un standard prédéfini comme dans le cas de la reconnaissance de blocs. Le système a une série de lettres prédéfinies mais a aussi une capacité d'adaptation à l'écriture humaine. Notons aussi que ce style d'interaction est compatible avec « Jot<sup>5</sup> » qui lui est le système de « reconnaissance des lettres » sous Palm OS. Précisons enfin qu'aucune phase d'acquisition d'écriture n'est prévue.

<sup>5</sup> Pour plus d'infos : [https://secure.cic.com/product\\_details/jot\\_details.asp](https://secure.cic.com/product_details/jot_details.asp)

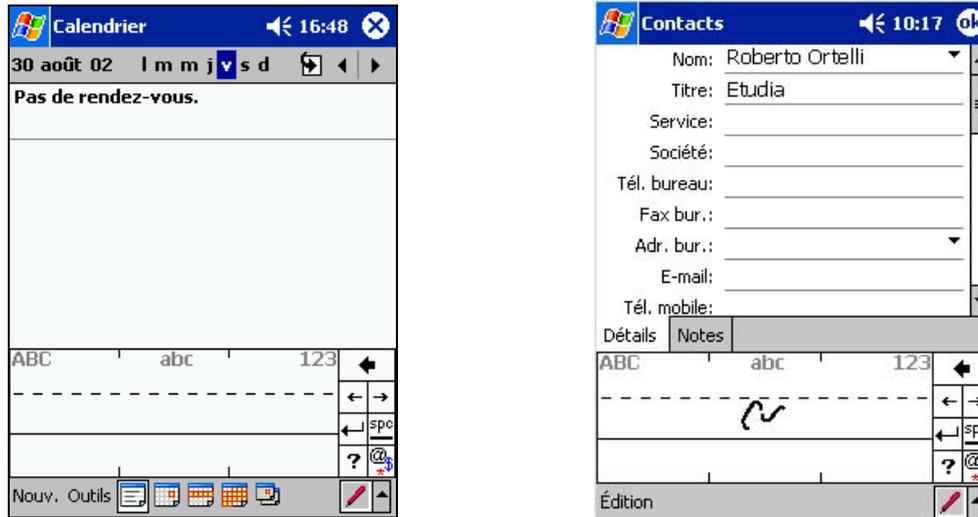


Figure 10: Reconnnaissance des lettres

4. Transcrire : Permet à l'utilisateur d'écrire directement à l'écran (dans ce cas « ciao », voir Figure 11). Après un laps de temps, le système décode l'écriture et la copie dans la cellule où le curseur est actif. L'écran entier joue ici le rôle de zone sensible pour introduire le texte.

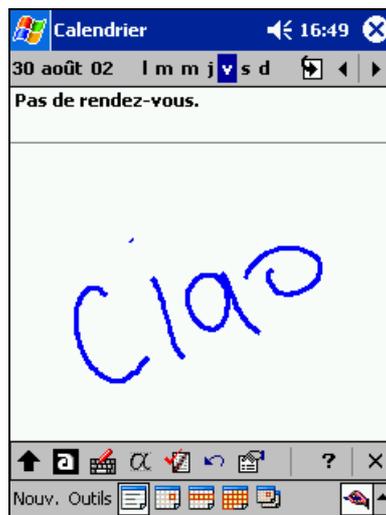


Figure 11: Transcrire

5. Vocabulaire : le système PocketPC 2002 offre à l'utilisateur un système de vocabulaire lui permettant d'avoir un choix pour compléter le mot dans un laps de temps inférieur (Figure 12) :



Figure 12: Vocabulaire intégré dans PocketPC 2002

Les styles d'interaction sans l'appui du stylet :

6. Boutons : tous les PPC sont aussi dotés de boutons « physiques » situés généralement dans la partie inférieure de l'organisateur.

Généralement ces boutons sont utilisés pour lancer des applications pré-installées dans le système comme par exemple la liste des contacts, le calendrier et la liste des tâches. Il existe aussi un quatrième bouton qui permet à l'utilisateur de revenir à l'écran de départ (**Erreur. L'origine riferimento non è stata trovata.**).

Il y a aussi d'autres types de boutons qui permettent à l'utilisateur de déplacer le curseur dans les quatre axes et un autre bouton qui permet de valider un choix. Soulignons que ces boutons ne sont pas toujours activés et disponibles pour toutes les situations. Par exemple, dans l'écran de démarrage ils n'ont aucune fonction y associée, par contre ils sont activés lorsque l'on veut créer un nouveau rendez-vous ou bien lancer une application.

Avant d'aborder la partie de la méthodologie nous devons nous arrêter à analyser plus proprement les quatre SI. Nous nous appuyerons encore une fois sur les théories de Shneiderman et par des études effectués dans le même domaine.

## 2.3 Etudes et recherches

La recherche de littérature dans le domaine des PocketPC s'est avérée plutôt difficile. Si d'un côté il existe une liste importante d'article concernant la multimodalité dans des PDA (MacKenzie et Zhang, 1997 ; Sears et Agora, 2002 ; Chang et MacKenzie, 1994 ; Calvet *et al.*, 2001 ; Frankish, Hull, Morgan, 1995 ; Fishkin, Moran et Harrison , 1998 ; Harrison, Fishkin, , Gujar, Mochon, et Want, 1998 ; Hynckley, Pierce, Sinclair et Horvitz, 2000), des articles sur les styles d'interaction ont été plus difficiles à trouver.

Pourtant, notre attention a été focalisé sur trois articles qui abordaient de plus près les styles d'interaction. Le premier (Sears, Arora, 2002) concerne la comparaison dans le système Palm entre Jot et Graffiti, respectivement reconnaissance des lettres et reconnaissance des blocs dans PocketPC, sur un public naïf, qui ne connaît pas les PDA. 21 sujets novices ont dû effectuer 6 tâches de longueur et contenu différents. Chaque sujet passe les 6 tâches avec l'un des deux styles d'interaction après une phase de familiarisation d'une durée de 10 minutes avec des tâches comparables à celles soumises lors du test. Les auteurs montrent que du point de vue de la satisfaction subjective, Jot est meilleur que Graffiti. Du point de vue des erreurs, les deux styles d'interaction sont comparables et ils relèvent aussi que Jot est plus facile à apprendre que Graffiti lorsqu'on est débutant.

La deuxième recherche que nous avons retenu porte sur Graffiti (ou reconnaissance des blocs dans le PocketPC). MacKenzie et Zhang (1997) ont passé un test sur 25 étudiants universitaire qui n'ont aucune connaissance dans la manipulation de systèmes dotés d'un stylet comme médium d'insertion de texte et d'interaction. La méthode consiste à faire effectuer à leur sujets une tâche comportant l'introduction répétée (5 fois) avec le stylet de l'alphabet de A à Z après avoir étudié pendant la durée d'une minute l'orthographe de Graffiti. Une fois cette tâche effectuée, ils ont 5 minutes pour apprendre librement l'orthographe de Graffiti pour ensuite effectuer la même tâche que l'étape précédente, mais cette fois sans l'aide d'une carte avec les symboles de Graffiti. Après 7 jours ils reviennent effectuer la même tâche sans aides ni d'étape de familiarisation. Les résultats montrent que pour la première étape (après une minute de familiarisation) le pourcentage de précision s'élève à 86% ; après 5 minutes de familiarisation la précision augmente à 97% ; après une semaine, cette précision stagne au 97%. Notons que lors de ces tests, ils n'ont pas demandés aux sujets d'introduire les lettres *X*, *K*, *U*, *V* et *F* en raison du fait qu'elles posent des problèmes particulier.

La troisième recherche qui nous intéresse porte sur l'utilisation de logiciels de reconnaissance de l'écriture et de l'acceptation des utilisateurs pour des interfaces fonctionnant avec un stylet (Frankish *et*

*al.*, 1995). Les sujets (24) doivent passer une série de tâches sur l'un des trois logiciels de reconnaissance de l'écriture. Les résultats de leur étude montrent que la satisfaction de l'utilisateur dépend de la capacité de reconnaissance du système, mais cet impact dépend de la tâche ; pour des tâches de routine, la reconnaissance de l'écriture doit être hautement structurée (définir des zones bien précises et délimitées pour rentrer du texte plutôt que de laisser libre l'utilisateur d'écrire n'importe où) ; la précision de la reconnaissance ne s'améliore pas avec l'exercice (des sujets refusaient ou étaient incapables de changer leur style d'écriture). De plus, si l'utilisateur a des difficultés à s'adapter à l'outil de reconnaissance, il faut adapter ce dernier à l'utilisateur.

Dans le prochain chapitre nous allons nous concentrer sur l'explication des objectifs de notre recherche et sur la méthodologie utilisée.

### 3. Objectifs de recherche

Cette partie présente nos motivations, le plan de recherche et les variables utilisées.

#### 3.1 Question de recherche et hypothèses

Dans notre étude nous voulons comprendre par quels biais (SI) les utilisateurs sont plus à l'aise lorsqu'ils doivent insérer du texte dans leur agenda de poche. Plus précisément nous allons essayer d'observer les utilisateurs en train d'effectuer des tâches fréquemment utilisées avec leur PDA afin de pouvoir extraire des invariants quant aux spécificités liées aux caractéristiques de l'utilisateur (expertise du PPC et maîtrise des SI), aux types de tâches (utilisation simple ou avancée), à la complexité de ces dernières et aux propriétés de chaque style d'interaction.

Nous allons utiliser une méthode mixte pour répondre à nos questions :

- Méthode expérimentale : passation de 4 séries de tâches dans lesquelles les sujets doivent effectuer pour chaque série 4 tâches. Pour chaque série de tâches fermées, l'utilisateur est obligé d'interagir avec un SI différent. Par la suite, nous présentons 2 tâches libres à effectuer ; dans ce cas, le sujet n'est plus contraint d'utiliser un SI imposé par l'expérimentateur.
- Entretiens semi-directifs : l'expérimentateur demandera aux sujets des questions concernant le déroulement du test et principalement d'identifier et d'évaluer les avantages et les inconvénients de chaque SI.

Nous essayons donc de décortiquer les éléments permettant à l'utilisateur d'effectuer le choix d'un SI parmi les quatre à disposition lors de l'exécution d'une tâche. Nous voulons aussi comprendre les raisons et les motivations de ces choix.

D'après la littérature et les recherches menées dans le domaine des PDA nous avons pu extraire quelques hypothèses et principalement la suivante:

Lorsque l'utilisateur doit choisir le style d'interaction le plus approprié, son choix est guidé par plusieurs facteurs distincts mais reliés entre eux comme : le niveau d'expertise avec le PocketPC et la maîtrise des styles d'interaction. Ces deux facteurs sont sensés modifier positivement ou négativement la satisfaction subjective ressentie par l'utilisateur qui en fonction de celle-ci guidera son choix présent et futur.

Nous essayerons aussi à répondre à des questions d'ordre plus général telles que:

- Est-ce que les utilisateurs adaptent-ils le style d'interaction en fonction de leur niveau d'expertise ? Est-ce que l'utilisateur expérimenté a un degré d'adaptation plus haut par rapport aux non expérimenté ?
- Est-ce qu'il existent des facteurs outre l'expertise du PocketPC et la maîtrise des SI qui guident le choix de l'usage du SI ?
- Quel est le rôle joué par les erreurs ? Qu'est ce qui guide l'interaction et les intentions de l'utilisateur ?
- Quel est le rôle des propriétés des SI, de l'utilisateur, de la tâche..., par rapport aux choix des utilisateurs ?
- Est-ce que des SI sont mieux adaptés selon le niveau de compétence de l'utilisateur ?

### **3.2 Analyse des propriétés des styles d'interaction**

Avant d'expliquer les hypothèses valables pour la partie expérimentale, nous allons analyser les quatre SI présents dans le PocketPC. Cette analyse nous permettra de formuler certaines hypothèses. Lorsque l'on veut comparer différents styles d'interaction afin de déterminer le meilleur il faut retenir au moins cinq facteurs (Shneiderman, 1992, p. 82):

1. *Le temps d'apprentissage* : combien de temps faut-il aux utilisateurs pour apprendre à utiliser les commandes pertinentes à la tâche.

De manière générale, plus l'apprentissage sera approfondi, meilleure sera la performance.

2. *Vitesse des performances* : combien de temps ça prend pour effectuer une série de tâches ?
3. *Fréquence d'erreurs* : Combien et de quel type d'erreurs sont faits lors de l'exécution d'une série de tâches ?

Si l'usager commet souvent beaucoup d'erreurs, il oublie des paramètres, ..., on peut dire que le logiciel est moins convivial. L'usage occasionnel tend à faire augmenter le taux d'erreurs. La gravité des erreurs devra aussi être considérée, car elles engendrent une diminution de l'efficacité globale (ça prends du temps pour réparer les erreurs), elles diminuent la fiabilité du

système (pensons au cockpit des avions, une erreur peut s'avérer impardonnable) et surtout elles diminuent la satisfaction globale perçue par les utilisateurs.

4. *Rétention dans le temps* : comment les utilisateurs se souviennent-ils de leurs connaissances après un certain laps de temps (par rapport à leur fréquence d'usage) ?

A long terme se manifeste comme la propriété intuitive qui fait qu'un utilisateur occasionnel s'y retrouve facilement dans le logiciel sans grande perte de performance.

5. *Satisfaction subjective* : les utilisateurs aiment-ils différents aspects du système ?

Pour obtenir une bonne satisfaction subjective, il faut obtenir des résultats rapidement. Parfois on peut trouver des cas où des compromis de facilité d'apprentissage sont acceptables. On pense alors à des logiciels dont l'incidence peut être critique pour la vie comme pour le contrôle aérien. Dans les cas de cette espèce, on accepte mieux un long apprentissage pour que, par la suite, les actions soient rapides et sans erreurs.

Nous allons maintenant analyser chaque style d'interaction en fonction des cinq facteurs mis en évidence par Shneiderman. Notons que Shneiderman dans l'analyse des SI, il prend en considération d'autres facteurs tels que la réversibilité, l'automatisation, la multimodalité, ... Les facteurs décrits ci-dessus, sont les plus importants lors de l'analyse d'une interface. Mais cela ne nous empêchera pas par la suite de tenir en considération aussi d'autres facteurs lors de notre analyse.

Notre étude se concentre donc seulement sur les quatre styles d'interaction manipulables par le stylet :

- Reconnaissance des blocs ou Graffiti® (RB)
- Reconnaissance des lettres ou Jot® (RL)
- Clavier virtuel (CV)
- Transcriber (TR)

Nous avons développé une grille d'analyse. Compte tenu du fait que nous allons analyser les styles d'interaction liés à l'insertion de contenu par le biais d'un stylet nous avons dû procéder à une simplification des critères de Shneiderman en ressemblant plusieurs critères. Nous sommes parvenus à limiter leur nombre à cinq critères :

1. Apprentissage :

- *Apprentissage* : Est-ce que le SI est il facile à apprendre ? Combien de temps faut-il aux utilisateurs pour apprendre à utiliser les commandes pertinentes à la tâche ?
- *Mémorisation* : Est-ce que ce SI est il facile à mémoriser ? Comment les utilisateurs se souviennent-ils de leurs connaissances après un certain laps de temps (par rapport à leur fréquence d'usage) ?
- *Automatisation* : Est-ce qu'il faut une réflexion particulière pendant son usage ou pas ?

2. Erreurs :

- *Erreurs* : Quel est le risque d'erreurs ? Quelle est la fréquence d'erreurs ? Combien et de quel type d'erreurs sont faits lors de l'exécution d'une série de tâches ?
- *Réversibilité* : Est-ce qu'on peut facilement faire « machine arrière » ? Opérations réversibles et incrémentales dont l'impact sur les objets concernés est immédiatement visible.

3. Multimodalité :

- *Multimodalité* : Est-ce qu'on peut facilement changer de SI ?
- *Support* : En cas d'échec, est ce qu'on a un support particulier pour rediriger l'action ?

4. Vitesse des performances :

- Quel est l'impact du SI sur les performances ? Combien de temps ça prend pour effectuer une série de tâches ?

5. Satisfaction subjective :

- Quel est la satisfaction de l'utilisateur vis-à-vis du SI ? Les utilisateurs aiment-ils différents aspects du système ? Ce critère est directement dépendant des 4 autres critères.

Chaque tableau ci-dessous contient l'analyse d'un SI en fonction des critères que nous venons de décrire et de compacter ensemble. Dans la dernière ligne de chaque tableau nous avons donné une note pour chaque SI en fonction du critère utilisé.

Tableau 1: L'apprentissage pour les 4 SI

Reconnaissance de blocs	Reconnaissance des lettres	Clavier Virtuel	Transcriber
<p>Implique un apprentissage plutôt long (Sears et Arora, 2002).</p> <p>Nous pouvons imaginer une mémorisation facile pour les items alphanumériques standard. L'introduction de signes et de symboles complexes pourrait s'avérer plus difficile.</p> <p>L'usage fréquent de ce SI pourrait permettre une automatisation de son usage.</p>	<p>Les débutants dans le domaine des PDA auront un apprentissage facilité impliquant moins d'efforts (Sears et Arora, 2002).</p> <p>Le seul apprentissage réside dans la compréhension de l'interface et dans l'apprentissage des caractères spéciaux.</p>	<p>Aucun apprentissage majeur. Il faut juste localiser les « touches » susceptibles de fournir d'autres caractères (comme la touche « majuscule »)</p> <p>De plus ce SI implique les mêmes contraintes liées à l'usage d'un clavier classique (localisation des touches) et aussi introduits de nouvelles contraintes au plus expérimentés car la localisation de certaines touches diffère d'un clavier classique.</p>	<p>Difficilement automatisable pour des symboles et des caractères spéciaux. Compréhension du fonctionnement du SI complexe.</p> <p>Apprentissage rapide pour des caractères alphanumériques standard.</p>
-	+	++	-

Tableau 2: Les erreurs pour les 4 SI

Reconnaissance de blocs	Reconnaissance des lettres	Clavier Virtuel	Transcriber
<p>Les débutants auront un taux d'erreurs important avec comme conséquence une perte de temps conséquente. Une diminution de ce taux pourrait être corrélé avec une augmentation du niveau d'expertise.</p> <p>On peut toujours corriger les erreurs directement dans la zone d'introduction du texte ou en mettant en évidence le mot directement dans la zone d'affichage.</p> <p>Vitesse de temps réduite si l'on doit effacer un mot entier. Si l'on doit changer une lettre au milieu d'un mot, ça va prendre plus de temps et d'efforts.</p> <p>Au fur et au mesure de l'apprentissage, l'utilisateur développe des stratégies lui permettant de choisir le mode de correction.</p>	<p>Peu d'erreurs pour les débutants.</p> <p>Si on est habitué avec la reconnaissance de blocs ou avec le système Graffiti sur PalmOS, on a tendance à écrire en Graffiti. Ce SI ne reconnaît qu'en partie ce langage. Par conséquent, on peut s'attendre à un taux d'erreurs supérieur chez les plus expérimentés.</p> <p>Il est aussi difficile d'écrire des éléments spéciaux comme + « ° ...</p> <p>En fonction de l'écriture de l'utilisateur, le système comprend ou ne comprend pas correctement la calligraphie. L'utilisateur doit en quelque sorte soigner son écriture.</p> <p>On peut toujours corriger les erreurs directement dans la zone d'introduction du texte ou en mettant en évidence le mot directement dans la zone d'affichage.</p> <p>Vitesse de temps réduite si l'on doit effacer un mot entier. Si l'on doit changer une lettre au milieu d'un mot, ça va prendre plus de temps et d'efforts.</p> <p>Au fur et au mesure de l'apprentissage, l'utilisateur développe des stratégies lui permettant de choisir le mode de correction.</p>	<p>Faibles voir aucun. Le stylet a suffisamment de place pour que l'utilisateur pointe correctement sur la bonne touche.</p> <p>On peut toujours corriger les erreurs directement dans la zone d'introduction du texte ou en mettant en évidence le mot directement dans la zone d'affichage.</p> <p>Vitesse de temps réduite si l'on doit effacer un mot entier. Si l'on doit changer une lettre au milieu d'un mot, ça va prendre plus de temps et d'efforts.</p> <p>Au fur et au mesure de l'apprentissage, l'utilisateur développe des stratégies lui permettant de choisir le mode de correction.</p>	<p>Erreurs au niveau de la reconnaissance car le système ne permet pas de définir la propre calligraphie.</p> <p>Le système prend plus de temps pour mettre en évidence le texte, la partie du texte ou tout autre élément qui doit être modifié. On pourrait imaginer que cette tâche soit une source d'erreurs supplémentaires et de perte de temps importante.</p>
Novices -- / Experts ++	Novices + / Experts +	Novices ++ / Experts ++	Novices -- / Experts --

Tableau 3: La multimodalité pour les 4 SI

Reconnaissance de blocs	Reconnaissance des lettres	Clavier Virtuel	Transcriber
<p>L'utilisateur dispose d'un accès immédiat à un clavier pour dépanner l'utilisateur.</p> <p>La suggestion de mots enregistrés dans une « base de données ». Cette fonctionnalité permet une augmentation des performances.</p> <p>A tout moment de l'interaction le système offre la possibilité de changer de SI.</p>	<p>L'utilisateur a toujours à disposition d'un clavier virtuel simplifié avec les caractères spéciaux.</p> <p>La suggestion de mots enregistrés dans une « base de données ». Cette fonctionnalité permet une augmentation des performances.</p> <p>A tout moment de l'interaction le système offre la possibilité de changer de SI.</p>	<p>La suggestion de mots enregistrés dans une « base de données ». Cette fonctionnalité permet une augmentation des performances.</p> <p>A tout moment de l'interaction le système offre la possibilité de changer de SI.</p>	<p>A tout moment de l'interaction le système offre la possibilité de changer de SI.</p> <p>Accès à des aides instantanées qui sont difficilement compréhensibles pour des débutants ou des intermédiaires.</p>
+	+	+	+

Tableau 4: La vitesse des performances pour les 4 SI

Reconnaissance de blocs	Reconnaissance des lettres	Clavier Virtuel	Transcriber
<p>Haute si le niveau de compétence est élevé.</p> <p>Faible si la fréquence d'usage est basse.</p>	<p>Vélocité si l'utilisateur ne doit pas effectuer d'efforts pour adapter son écriture.</p> <p>Pour faire reconnaître des caractères comme le t ou le i, le système prend plus de temps (par exemple d'abord il faut écrire une t sans le tiret et ensuite le système attend quelques instants pour que l'utilisateur écrive le tiret de telle sorte qu'il puisse encoder correctement la t). Ce laps de temps empêche l'utilisateur de poursuivre l'insertion du texte.</p>	<p>Performances meilleures pour les non expérimentés.</p>	<p>Diminuée par le fait que le système prend beaucoup de temps pour décoder l'écriture (ce laps de temps peut être paramétré).</p> <p>Ce SI offre la possibilité d'insérer ou pas une espace après le dernier mot inséré. Lors des corrections à l'intérieur d'un mot erroné, l'activation cette option engendre un handicap au niveau des performances.</p>
Novices -- / Experts ++	Novices - / Experts +	Novices ++ / Experts ++	Novices -- / Experts -

Tableau 5: La satisfaction subjective pour les 4 SI

Reconnaissance de blocs	Reconnaissance des lettres	Clavier Virtuel	Transcriber
<p>Haute satisfaction si le niveau de compétences est élevé. Basse satisfaction au début.</p> <p>Basse satisfaction lorsqu'il faut à chaque fois utiliser le clavier alphanumérique pour introduire un item qu'on ne connaît pas ou qu'on arrive pas à mémoriser correctement.</p>	<p>Haute pour les utilisateur débutants. Insatisfaction pour les très expérimentés.</p> <p>Meilleure satisfaction lors de l'introduction de caractères spéciaux.</p>	<p>Haute pour les novices. Moyenne pour les plus expérimentés.</p>	<p>Moyenne pour les débutants. Nulle pour les expérimentés.</p>
Novices - / Experts ++	Novices ++ / Experts -	Novices ++ / Experts +- / +-	Novices +- / Experts --

### 3.3 Partie expérimentale

#### 3.3.1 Variables

##### 3.3.1.1 Variables indépendantes

Styles d'interaction :

- Reconnaissances des blocs (RB)
- Reconnaissances des lettres (RL)
- Clavier Virtuel (CV)
- Transcriber (TR)

Niveau d'expertise avec le PocketPC

- Experts ou novices

##### 3.3.1.2 Variables invoquées :

Niveau de maîtrise des styles d'interaction

- Maîtrise ou pas maîtrise de RB, RL, CV ou TR

##### 3.3.1.3 Tâches :

Nous avons conçu principalement deux types de tâches : les tâches *fermées* impliquent quatre séries de tâches où nous forçons les sujets à utiliser à chaque fois un SI différent ; les tâches dites *libres* où les sujets sont libres de choisir le(s) SI de leur préférence.

##### 3.3.1.4 Variables dépendantes

Performances : nombre d'erreurs et temps d'exécution de la tâche

Satisfaction subjective : réponses au questionnaire suivant chaque série des tâches fermées

##### 3.3.1.5 Variables contrôlées

Pendant la phase expérimentale nous contrôlons une série de variables, en effet chaque élément cité ci-dessous, est présenté à l'utilisateur suivant un schéma systématique:

- L'ordre de passation des séries de tâches et des tâches mêmes.
- L'ordre d'usage.

- Nous avons développé 4 items différents pour chaque tâche. Chaque item est présenté une seule fois à l'utilisateur.
- Les deux tâches libres sont composés par deux items inédits.

### 3.3.2 Hypothèses opérationnelles

Avant de présenter les hypothèses opérationnelles que nous avons développées suivant notre analyse des styles d'interaction du PPC, nous allons rappeler la question de recherche générale. Nous postulons que lorsque l'utilisateur doit choisir le style d'interaction le plus approprié, son choix est guidé par plusieurs facteurs distincts mais reliés entre eux comme : le niveau d'expertise avec le PocketPC et la maîtrise des styles d'interaction. Ces deux facteurs sont sensés modifier positivement ou négativement la satisfaction subjective ressentie par l'utilisateur qui en fonction de celle-ci guidera son choix présent et futur.

**H1.** Plus le niveau d'expertise du PPC et de maîtrise des SI de l'utilisateur est haut et plus l'utilisateur utilise des styles d'interaction de haut niveau et inversement (Shneiderman).

Nous pouvons constater cette hypothèse à deux niveaux :

Questionnaire de recrutement : réponses dans la partie 5 « informations relatives au médiums d'insertions de texte » et dans les réponses à la partie 1 « informations concernant votre expérience avec les PDA » (Annexe 1 : Questionnaire d'évaluation des sujets experts).

Tâches libres : choix du SI utilisé pour effectuer les deux tâches libres.

En fonction des tâches, l'utilisateur aura tendance à changer de style d'interaction. Pour des tâches complexes, il utilisera des méthodes simplifiant la tâche.

Nous pensons observer ce phénomène au niveau des tâches libres.

**H2.** Plus l'action de l'utilisateur est entravée par les erreurs, plus il aura tendance à changer de style d'interaction en faveur d'une modalité de plus bas niveau impliquant moins d'efforts cognitifs (Rasmussen – Modèle de Hutchins *et al.*)

Nous pouvons vérifier cette hypothèse en fonction des réponses données dans le questionnaire qu'ils doivent remplir après l'exécution de chaque série de tâches, en particulier nous nous concentrons sur la réponse 1.3 *Par rapport aux tâches que vous avez effectuées, auriez-vous préféré un autre*

*médium d'insertion ?* (Annexe 1 : Questionnaire d'évaluation des sujets experts) et aussi dans l'évaluation du niveau de difficulté de chaque tâche.

**H3.** Pour des tâches fréquentes et de routine, la capacité de reconnaissance de l'écriture du système doit être hautement structurée (Frankish, Hall, Morgan, 1995). Les performances et la satisfaction subjective découlant de l'utilisation de la reconnaissance de blocs ou de la reconnaissance des lettres seront supérieures par rapport au Transcriber. Le CV, en étant hautement structuré et ne nécessitant pas d'une reconnaissance particulière, est le meilleur SI.

$CV < (RB=RL) < (TR)$

Nous pouvons vérifier cette hypothèse à l'aide du questionnaire post-série des tâches ainsi que dans l'analyse des performances.

**H4.** Chaque utilisateur adapte son action en utilisant le SI le mieux adapté en fonction de son niveau d'expertise des SI (Modèle de Hutchins et de Rasmussen);

Nous allons comparer les performances des utilisateurs entre les tâches fermées et les performances avec les tâches libres.

Nous pouvons nous attendre à observer des résultats qui mettent en évidence de meilleures performances des experts et cela principalement pour deux raisons : premièrement parce qu'en ayant une meilleure expertise du PPC ils sont capables de manier d'une façon plus performante tous les SI ; deuxièmement parce que la connaissance des différents SI a été développée depuis plus longtemps que les novices, qui eux sont au stade de la découverte du système et des SI.

**H5.** Plus l'utilisateur est conscient des différences dans les propriétés de chaque style d'interaction (il a des connaissances stables et transférables), plus il aura des capacités à adapter sa conduite (modèle SSOA de Shneiderman, Modèle de Hutchins, et Rasmussen) ;

**H6.** Plus les connaissances acquises sont fréquemment et régulièrement mobilisées et plus le sujet sera capable de manipuler le système efficacement et de s'adapter en conséquence (modèle SSOA de Shneiderman, Modèle de Hutchins *et al.* et de Rasmussen) ;

Nous nous concentrons sur l'analyse des différences entre les experts et les novices.

On se référant à notre analyse des styles d'interaction (suivant les critères de Shneiderman et aux études précédant notre travail) nous pouvons soulever les hypothèses suivantes:

**H7.** Erreurs :

Novices :  $CV < RL < RB < TR$

Experts :  $CV < RB < RL < TR$

**H8.** Vitesse des performances :

Novices:  $CV < RL < RB < TR$

Experts :  $(RB = RL) < CV < TR$

**H9.** Satisfaction subjective :

Novices :  $(RL = CV) > TR > RB$

Experts :  $(RB = RL) > CV > TR$

### **3.4 Partie qualitative**

Comme nous l'avons souligné en précédence, cette partie nous permet de récolter des informations plus liées à l'usage des SI testés. Cela dans le but de mieux comprendre les résultats issus de la partie expérimentale. Cette phase, d'une durée maximale de 30 minutes (mais qui peut varier en fonction des utilisateurs interviewés), s'effectue à la fin le partie des tests, elle suit la phase des tâches libres.

En fonction du groupe d'appartenance du sujet, certaines questions ne seront pas posées. Pour les experts (Annexe 9), nous avons par exemple prévu de demander les raisons qui l'ont amené à utiliser pendant les tâches libres un SI qui diffère de celui qu'il utilise couramment. Cette question ne peut donc pas être posée aux novices (Annexe 8). Pour les deux groupes nous avons demandé aux sujets de commenter les tâches qu'ils ont effectuées, de nous dire quels sont les avantages et les inconvénients de chaque SI, ... Pour plus de détails, dans les annexes sont disponibles les fiches de questionnement.

Rappelons que les entretiens ont été enregistrés et nous avons inséré dans des tableaux les informations principales qui sont ressorties au cours de cette phase (Annexe 10 : Analyse des entretiens).

## 3.5 Méthode

### 3.5.1 Sujets

Nous avons recruté les utilisateurs selon leur niveau d'expertise du PocketPC. Nous avons regroupé nos 16 utilisateurs dans deux groupes : expérimentés (8 hommes, âge moyen 29 ans) et novices (8 hommes, âge moyen 28 ans).

Le recrutement de nos sujets s'est effectué sur Internet par le biais de messages sur des forums, des newsgroups, des mailings listes Suisses et Françaises.

### 3.5.2 Le questionnaire de recrutement

Pour les sujets expérimentés nous avons prévu qu'ils remplissent un questionnaire (Annexe 1) dans lequel nous demandons à nos sujets des informations relatives à leur expérience avec le PocketPC, leur niveau de compétences avec les différents SI, des informations d'usage, ... Une fois le questionnaire rempli nous procédions à les recontacter afin de fixer un rendez-vous pour passer le test et l'entretien. Les sujets novices ont rempli le même questionnaire se limitant à indiquer des informations socio-métriques telles que le genre, l'âge et la profession.

### 3.5.3 Matériel

Un PocketPC Hewlett-Packard Jornada 568 avec les spécifications suivantes :

- Processeur 206 Mhz
- 64 MB d'espace disque
- 32 MB ROM
- Ecran LCD couleur 240x320
- Système opératif : Microsoft PocketPC 2002 version française
- Batterie spéciale en lithium pour augmenter son autonomie (14 heures)
- 1 stylet
- 1 carte de Compact Flash Type I pour communiquer sans fil sur un réseau LAN (D-Link DCF-660W) configurée en mode Ad-Hoc pour communiquer directement avec la carte PCMCIA située dans le PC Portable

Un PC portable:

- Pentium III 1.1 GHz

- 320 MB RAM
- 28 GB d'espace disque
- 1 carte PCMCIA type II pour communiquer sans fil sur un réseau LAN (D-Link DWL-650+) configurée en mode Ad-Hoc pour communiquer directement avec la carte Compact Flash située dans le PocketPC
- Système Opératif : Windows XP Professional
- 1GB d'espace disque libre nécessaire pour stocker l'enregistrement audio et vidéo

Autre :

- Un microphone pour enregistrer l'audio
- Virtual CE 3.0<sup>6</sup>: ce logiciel permet de visualiser l'écran du PocketPC sur un ordinateur portable
- Snag-It ! 6.1.3<sup>7</sup> : Enregistre une vidéo l'interaction entre le sujet et le PocketPC directement depuis l'écran de l'ordinateur portable. Ce logiciel enregistre tout ce que disent l'expérimentateur et le sujet.
- Microsoft MediaPlayer<sup>8</sup> pour regarder les vidéos en format AVI
- Nero 5.5.9.17<sup>9</sup> : pour graver les vidéos des entretiens

### 3.5.4 Procédure

#### Phase 0 : Briefing

En début de la session d'expérimentation, l'observateur explique au sujet le déroulement du test qui se divise en trois parties pour les expérimentés (Tableau 6, Annexe 5) et en quatre parties pour les novices. En effet, pour ces derniers, l'expérimentateur insère une phase de familiarisation de tous les SI avant de commencer le véritable test. La familiarisation dure environ 5 minutes par SI.

Tableau 6: schéma des trois principales phases de l'entretien

Etape 1 : Tâches fermées

Etape 2 : Tâches libres

Etape 3 : Entretien

---

<sup>6</sup> HanaHo BitBank Software : <http://www.hive.speedhost.com/>

<sup>7</sup> TechSmith Corporation: <http://www.techsmith.com/>

<sup>8</sup> Microsoft Corporation: <http://www.microsoft.com/windows/windowsmedia/>

<sup>9</sup> Ahead Software : <http://www.nero.com>

Si nécessaire 2 minutes de re-familiarisation du SI, valable pour tous les sujets.

Passation de quatre séries de tâches.

Chaque série de tâche s'effectue toujours avec le même SI.

A la fin d'une série, le sujet doit remplir un questionnaire.

Une fois le questionnaire rempli, le sujet effectue une nouvelle série de tâches avec un autre SI

Le sujet effectue deux tâches dites libres car il a le choix du SI à utiliser.

L'entretien est utile pour obtenir des informations plus qualitatives quant à l'usage des différents SI par le sujet et pour commenter les tâches effectuées.

Phase de debriefing une fois l'entretien achevé.

### Phase I : Les tâches fermées

On reprend grosso modo les tâches effectuées dans l'étude de Sears et Arora (2002). Lors de l'exécution des tâches nous permettons aux utilisateurs de corriger leurs erreurs et nous les incitons d'essayer d'aller jusqu'au bout dans l'exercice. De plus, les tâches à effectuer ne comportent pas d'accents afin de faciliter leur exécution.

Après chaque tâche nous soumettons à l'utilisateur un questionnaire (Annexe 2) à remplir dont le but est d'évaluer la série de tâches en fonction de sa satisfaction subjective. Les tâches proposées peuvent être des tâches fréquentes ou peu fréquentes, complexes ou peu complexes. Il est clair que chaque utilisateur percevra le niveau de difficulté en fonction de sa propre expérience et de son mode d'interaction avec le PocketPC. Voici en détails les types de tâches présentes dans chaque séries :

#### Tâche de familiarisation :

- Experts PPC : cette tâche facultative d'une durée de deux minutes permet à tous les utilisateurs de se rafraîchir la mémoire avec le SI qu'il utilisera au cours de la série. Nous conseillons à ces utilisateurs d'écrire par exemple l'alphabet dans une note ou bien leur coordonnées, adresse web...
- Novices : l'expérimentateur montre des fiches (Annexe 3) aux sujets avec lesquelles il explique les principales fonctionnalités du SI concerné. Ensuite, pendant environ 5 minutes par SI, il propose au sujet de rentrer un petit texte comme par exemple (Annexe 4):

*A l'école, c'est devenu la norme : chaque année, de la maternelle au primaire, les instituteurs voient arriver leur lot de petits démons qui entendent faire la loi. Voulant tout, tout de suite, refusant toute contrainte. Se roulant par terre ou burlant à la moindre contrariété.*

Avant de démarrer chaque série des tâches, l'expérimentateur incite le sujet à s'exercer à nouveau (facultatif), pendant 2 minutes, en proposant les mêmes tâches que pour les expérimentés.

Tâches fermées (Annexe 6):

1. Insérer un nom et une adresse dans un agenda de contacts : cette tâche implique l'introduction de symboles et de caractères alphanumériques. Par exemple :

*Masiero Fabrice  
Rue de Berne 34  
1247 Anieres*

Nombre de caractères : 37

2. Insérer une adresse Web dans la barre de navigation de Internet Explorer : tâche simple, impliquant une séquence logique de caractères alphanumériques et de symboles. Par exemple :

*http://www.letemps.ch*

Nombre de caractères : 21

3. Insérer un rendez-vous : Tâche a double entrée. Premièrement insérer un mot pour spécifier le sujet du rendez-vous et deuxièmement écrire une brève phrase pour spécifier plus en détail le sujet du rendez-vous. Par exemple :

Objet : *Medecin*  
Note : *Demander les ordonnances*

Nombre de caractères : 29

4. Ecrire un nouveau email : Implique l'utilisation du caractère « @ ». Ecrire ensuite un petit paragraphe. Par exemple :

A : *roland@geneve.ch*

Objet : *Nouveau lien*

Texte :

*Bonjour,*

*J'ai remarqué un lien mort pour accéder à votre page de contact. En effet, la page en question ne s'affiche pas.*

*Salutations.*

Nombre de caractères : 139

Rappelons que pour chaque tâche nous avons prévu quatre items avec un contenu comparable de telle sorte que le sujet n'effectue une tâche sur un item déjà utilisé en précédence.

Après chaque série de tâches le sujet est incité à remplir un questionnaire (Annexe 2) portant sur des questions telles que :

- Le niveau de satisfaction subjective
- Appréciation (agréable ou pas)
- L'envie de changer de style pour effectuer la tâche

#### Phase II : Tâches libres (Annexe 7)

Cette phase est dite libre car l'utilisateur peut choisir le SI qu'il préfère pour effectuer ces deux tâches :

1. Insérer un nom et une adresse dans un agenda de contacts : cette tâche implique l'introduction de symboles et de caractères alphanumériques.
2. Insérer un rendez-vous : Tâche a double entrée. Premièrement insérer un mot pour spécifier le sujet du rendez-vous et deuxièmement écrire une brève phrase pour spécifier plus en détail le sujet du rendez-vous.

Ces tâches sont les mêmes que dans la phase précédente mais les items à insérer diffèrent (pas en difficulté). Sur la base de ces deux tâches la phase suivante entre en jeux.

### Phase III : Entretiens sémi-directifs

Cette phase nous permet de mieux expliciter ce qui s'est passé pendant la première phase expérimentale. Nous profiterons de cette partie pour poser des questions sur des détails qui ont frappé l'expérimentateur au cours de la phase précédente.

### Observation et récolte des données

Nous allons enregistrer toutes les interactions entre le sujet et le PocketPC par le biais de deux logiciels :

- Virtual CE 3.0 : le « client » étant installé sur le PocketPC expérimental. Le « host » installé sur un pc portable. Sur le pc portable on visualise l'interface du PPC et l'interaction du sujet avec son stylet.
- Snag-it !: Permet de capturer en séquences de film (\*.AVI) tout ce qui se passe à l'écran (ou dans une partie de l'écran) du pc portable. Nous pouvons aussi enregistrer par ce moyen tout ce qui est dit pendant l'entretien. Dans cette capture d'écran (ou mieux de la fenêtre) nous ajouterons aussi la date et le temps écoulé au format HH : MM : SS : ms, dans le but de permettre à l'expérimentateur d'ensuite pouvoir évaluer de manière précise les performances du sujet.

## 4. Résultats

Dans cette partie nous allons présenter les résultats des analyses effectuées. Nous avons effectué principalement des analyses de la variance sur des mesures répétées pour tester nos hypothèses et pour comparer les performances de nos sujets. Dans un premier temps nous allons commenter le déroulement de nos tests et dans un deuxième temps nous allons présenter nos analyses par rapport aux hypothèses formulées précédemment. Mais d'abord regardons les réponses au questionnaire soumis aux experts avant le début de l'expérimentation.

### 4.1 Questionnaire experts

#### 4.1.1. Genre et âge :

8 sujets, 8 hommes, âge moyenne de 29 ans (min = 24 ; max 38).

#### 4.1.2. Temps d'utilisation du PPC

En moyenne nos sujets ont une utilisation d'environ 7 mois ( $m = 7.375$  mois, min = 4 mois, max = 12 mois). Trois sujets ont une expérience antérieure avec un autre PPC d'une durée de 12 mois pour un sujet et de 6 mois pour les deux autres. De plus, 6 sujets avaient déjà utilisé un autre PDA (non PPC) pendant une durée moyenne de 34 mois (min = 24 mois, max = 48 mois).

#### 4.1.3. Degré de compétence subjectif PPC

4 sujets estiment être experts dans leur degré de compétence avec le PPC, 3 sujets intermédiaires, 1 sujet novice.

#### 4.1.4. Utilisation du PPC

Tous les sujets utilisent presque tous les jours (6.37 jours par semaine) le PPC au travail et en privé avec une fréquence moyenne par jour de 8.13 (min = 2, max = 30).

Quant aux tâches effectuées (Figure 13), nos sujets utilisent leur PPC principalement pour consulter/éditer le calendrier (agenda), les contacts, des notes ou pour les jeux. La Figure 14 montre que

le calendrier est utilisé avec une fréquence moyenne de 35% sur l'ensemble des tâches, suivi par les contacts (environ 20%) et par d'autres tâches (12%).

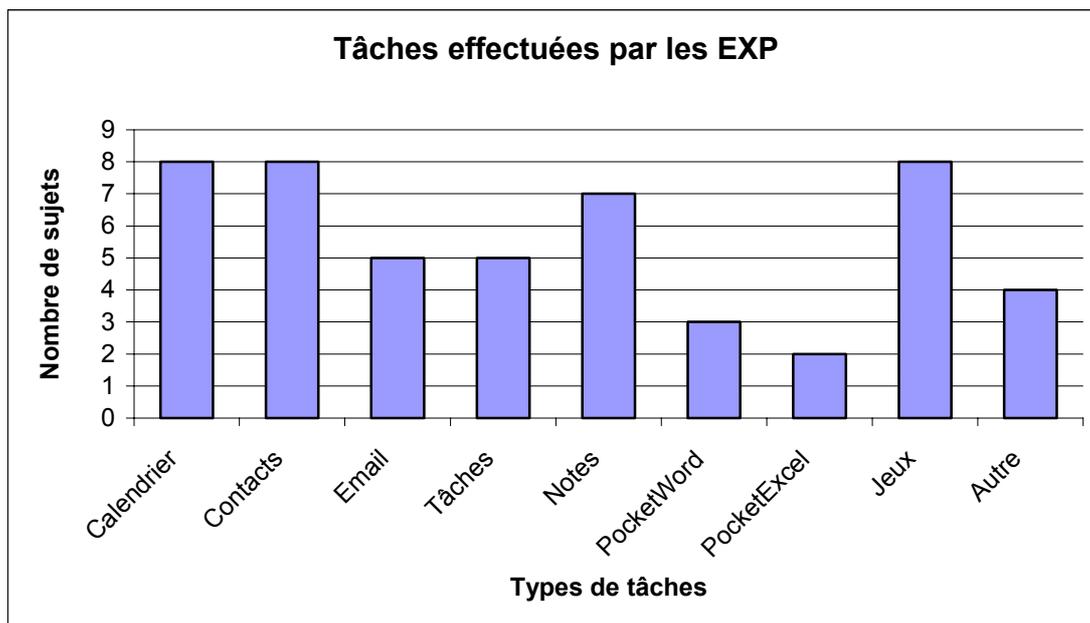


Figure 13: Types de tâches effectuées principalement par les experts.

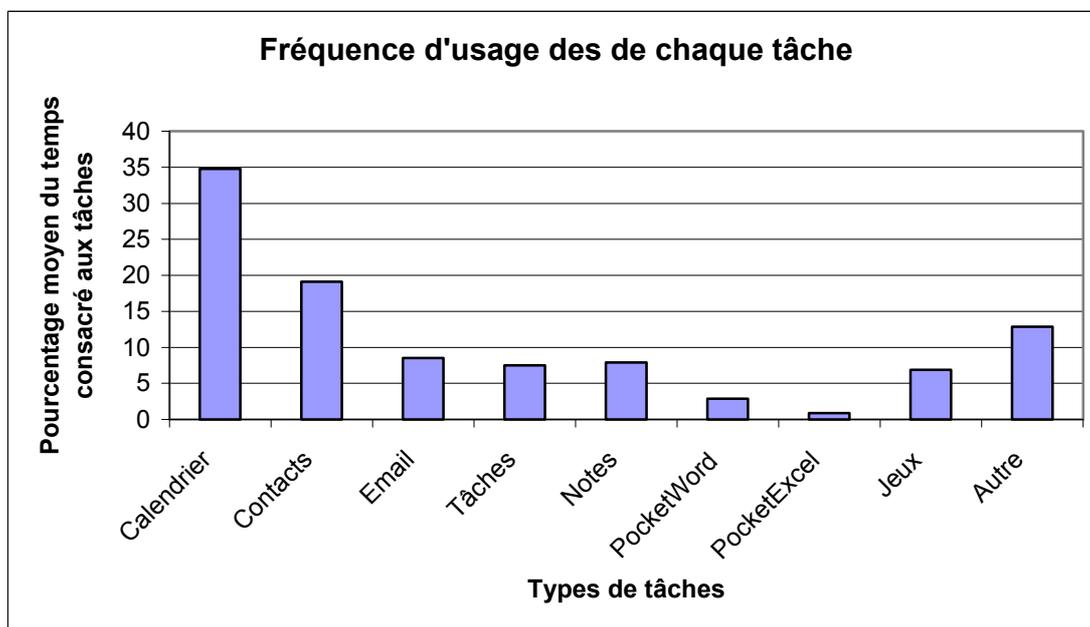


Figure 14: Fréquence d'usage moyen en pourcentage relatif aux principales tâches offertes par le PPC

#### 4.1.5. Les styles d'interaction

La Figure 15 montre l'utilisation réelle des 4 SI de nos 8 sujets experts. La Figure 16 par contre, indique l'évaluation des compétences de nos sujets pour chaque SI.

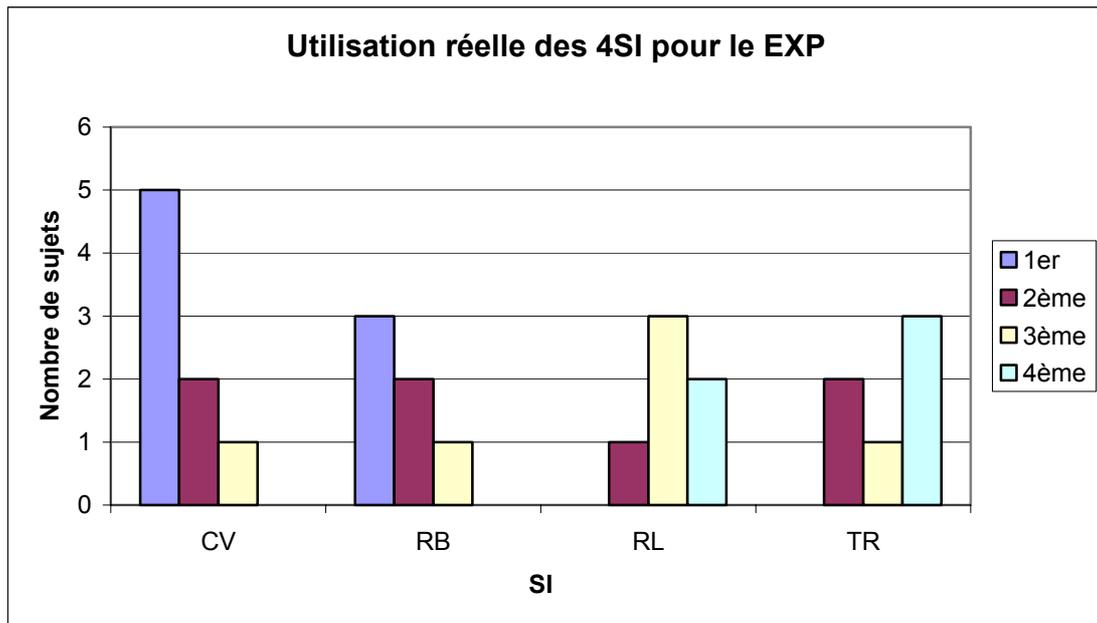


Figure 15: Utilisation réelle des 4 SI. Nous observons que le CV est le SI utilisé en priorité par 5 sujets suivi par la RB (3 sujets). Les SI utilisés comme modalité secondaire varie beaucoup plus.

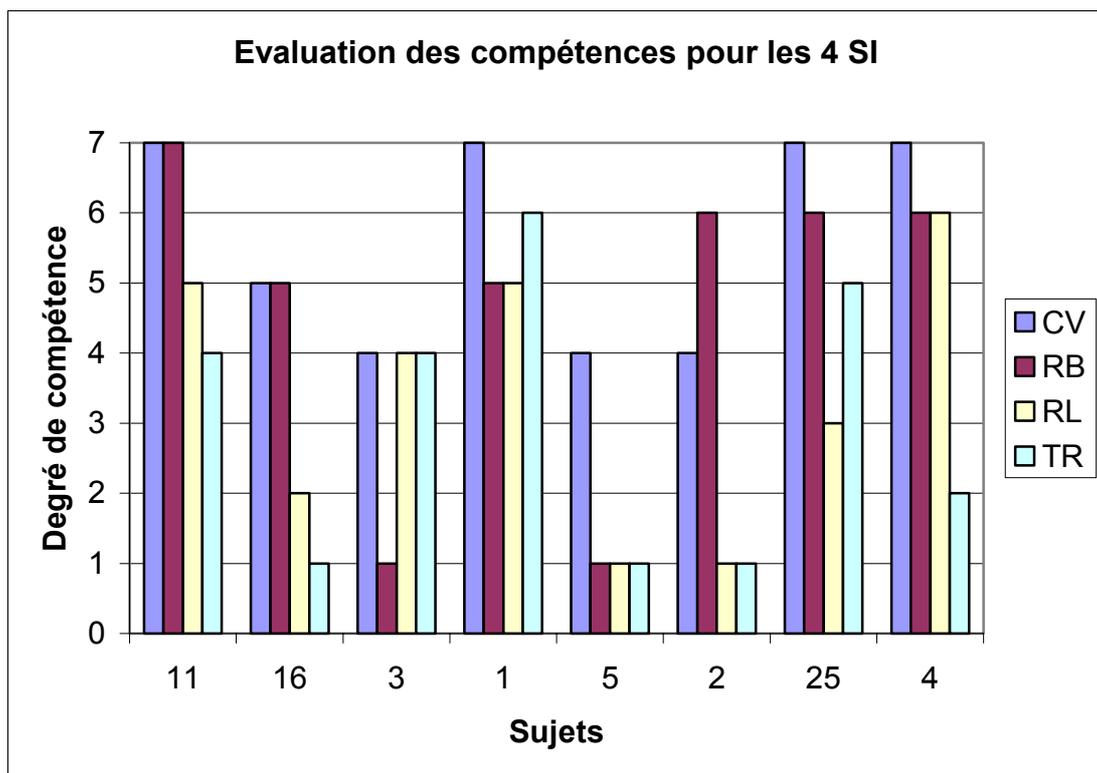


Figure 16: Evaluation des compétence des sujets du groupe expert pour chaque SI (0 = débutant ; 7 = expert)

## 4.2 Résultats partie quantitative

Dans cette partie d'analyse des résultats, nous allons d'abord nous concentrer sur l'analyse de la variance (ANOVA) sur des mesures répétées concernant le temps d'exécution moyen sur les 4 tâches fermées pour les 4 SI testés. Nous présenterons ensuite les résultats du point de vue des erreurs.

### 4.2.1 Temps moyen d'exécution des tâches

Tableau 7: Analyse descriptive du temps moyen d'exécution des 4 tâches fermées avec les 4 SI par les sujets experts

SI:CV Expertise:EXP	SI:RB Expertise:EXP	SI:RL Expertise:EXP	SI:TR Expertise:EXP
Temps	Temps	Temps	Temps
Min: 44.51825	Min: 41.55025	Min: 57.66650	Min: 67.08325
1st Qu.: 46.57500	1st Qu.: 55.00538	1st Qu.: 84.21300	1st Qu.: 100.22100
Mean: 56.28478	Mean: 80.19109	Mean: 90.82375	Mean: 122.97481
Median: 52.75925	Median: 67.49175	Median: 88.87875	Median: 121.50162
3rd Qu.: 66.96044	3rd Qu.: 89.19231	3rd Qu.: 94.57406	3rd Qu.: 140.46556
Max: 72.33325	Max: 162.21675	Max: 131.01675	Max: 186.63325
Total N: 8.00000	Total N: 8.00000	Total N: 8.00000	Total N: 8.00000
NA's : 0.00000	NA's : 0.00000	NA's : 0.00000	NA's : 0.00000
Std Dev.: 11.52840	Std Dev.: 39.28139	Std Dev.: 20.51466	Std Dev.: 38.30630

Tableau 8: Analyse descriptive du temps moyen d'exécution des 4 tâches fermées avec les 4 SI par les sujets novices

SI:CV Expertise:NOV	SI:RB Expertise:NOV	SI:RL Expertise:NOV	SI:TR Expertise:NOV
Temps	Temps	Temps	Temps
Min: 61.53050	Min: 86.70025	Min: 90.29725	Min: 110.36675
1st Qu.: 66.80844	1st Qu.: 118.75369	1st Qu.: 106.56263	1st Qu.: 128.60006
Mean: 74.96384	Mean: 135.31397	Mean: 117.85891	Mean: 159.34916
Median: 72.15100	Median: 139.36937	Median: 120.16675	Median: 146.91688
3rd Qu.: 75.86406	3rd Qu.: 148.28556	3rd Qu.: 131.37138	3rd Qu.: 191.49613
Max: 106.08325	Max: 180.08325	Max: 140.43350	Max: 224.05300
Total N: 8.00000	Total N: 8.00000	Total N: 8.00000	Total N: 8.00000
NA's : 0.00000	NA's : 0.00000	NA's : 0.00000	NA's : 0.00000
Std Dev.: 13.91304	Std Dev.: 28.59217	Std Dev.: 17.61933	Std Dev.: 41.18673

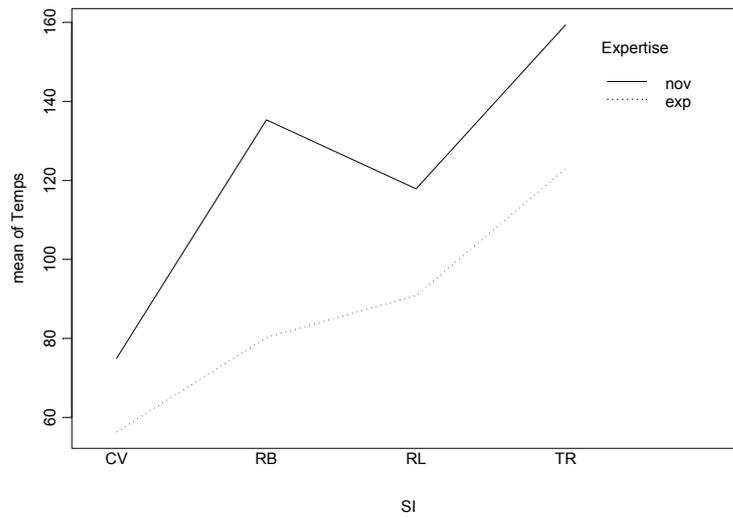


Figure 17: Effet des SI sur le temps moyen d'exécution des tâches

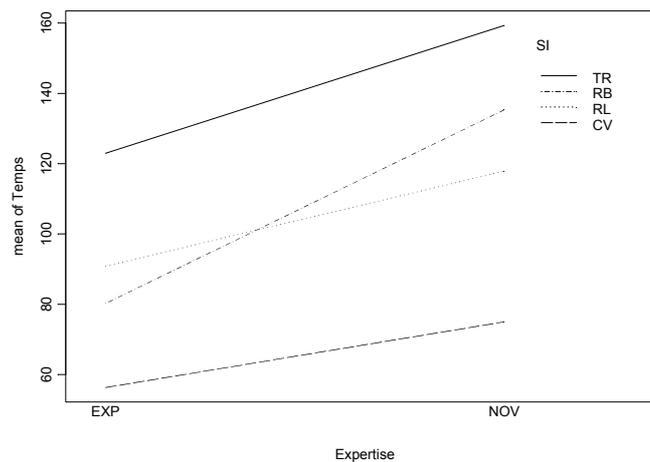


Figure 18: Effet de l'expertise sur les temps moyen d'exécution des tâches

Nous observons que pour chaque SI, le groupe expert a de meilleures performances par rapport aux novices. De plus, pour les deux groupes, CV est la meilleure modalité alors que TR est celle qui implique un temps d'exécution supérieur. Pour les experts, RB est meilleure que RL, par contre, pour les novices, on constate le contraire.

Comme chaque sujet a effectué toutes les tâches fermées avec chaque modalité d'interaction nous allons effectuer une analyse de la variance (ANOVA) sur des mesures répétées, en vérifiant les effets principaux des deux variables de même que leur effet d'interaction.

Nous constatons un effet principal pour l'expertise ( $F = 22.883$  ;  $p < .0000129$ ) et pour le SI ( $F = 18.57717$  ;  $p < .00$ ) sur le temps moyen d'exécution des 4 tâches fermées. Nous n'observons pas un effet d'interaction entre les deux variables ( $F = 1.19061$  ;  $p > .32$ ). Par conséquent, le temps d'exécution diffère de manière significative en fonction du groupe d'appartenance et en fonction du SI utilisé. L'effet du SI utilisé est donc indépendant du niveau d'expertise au PPC. La comparaison multiple sur un seul facteur nous montre des différences significatives ( $p = .05$ ) entre :

- CV-RB
- CV-RL
- CV-TR
- RB-TR
- RL-TR

Nous pouvons conclure que le CV (indépendamment du facteur expertise) est le meilleur SI, suivi par RB et RL qui sont équivalents. TR est le SI qui diffère significativement avec tous les autres SI.

Bien que nous n'ayons pas observé d'effet d'interaction, nous allons quand même effectuer une analyse des effets simples, car, pour vérifier ou infirmer certaines de nos hypothèses, nous avons besoin de connaître ces effets ; d'ailleurs, Howell (1998) suggère que « si l'expérimentateur a une raison a priori d'examiner par exemple l'interaction AB à un niveau particulier de C, il est tout à fait libre de le faire même en l'absence d'une interaction de deuxième ordre significative » (p. 499). L'analyse des effets simples découlant de l'ANOVA sur les temps moyen d'exécution nous montre les différences de performances des SI suivants ( $p = .05$ ) :

Pour les experts :

- CV – TR
- RB - TR

Pour les novices :

- CV – RL
- CV – TR
- CV – RB

Dans les 4 prochaines analyses nous allons nous concentrer sur les données brutes des 4 tâches fermées. Pour chaque tâche nous allons l'analyser du point de vue du temps d'exécution et du nombre

d'erreurs effectués, de même que du point de vue de la satisfaction subjective. Nous nous appuyons encore sur l'ANOVA sur des données répétées.

#### 4.2.2 Tâche fermée 1 : Adresse

##### Temps d'exécution de la tâche :

Tableau 9: Analyse descriptive du temps d'exécution de TF1 avec les 4 SI par les sujets experts

SI:CV Expertise:EXP	SI:RB Expertise:EXP	SI:RL Expertise:EXP	SI:TR Expertise:EXP
TF1	TF1	TF1	TF1
Min: 29.00600	Min: 30.60000	Min: 37.80000	Min: 47.40000
1st Qu.: 32.51400	1st Qu.: 42.07500	1st Qu.: 47.55000	1st Qu.: 67.36650
Mean: 39.94613	Mean: 57.64912	Mean: 59.10850	Mean: 84.23625
Median: 34.12700	Median: 56.01300	Median: 51.60000	Median: 85.44500
3rd Qu.: 48.23350	3rd Qu.: 65.83325	3rd Qu.: 64.69025	3rd Qu.: 98.25000
Max: 60.46700	Max: 94.33400	Max: 94.53300	Max: 120.20000
Total N: 8.00000	Total N: 8.00000	Total N: 8.00000	Total N: 8.00000
NA's : 0.00000	NA's : 0.00000	NA's : 0.00000	NA's : 0.00000
Std Dev.: 11.55462	Std Dev.: 21.78622	Std Dev.: 20.02866	Std Dev.: 25.45212

Tableau 10: Analyse descriptive du temps d'exécution de TF1 avec les 4 SI par les sujets novices

SI:CV Expertise:NOV	SI:RB Expertise:NOV	SI:RL Expertise:NOV	SI:TR Expertise:NOV
TF1	TF1	TF1	TF1
Min: 41.73400	Min: 63.93300	Min: 58.13400	Min: 66.20000
1st Qu.: 48.93375	1st Qu.: 92.40800	1st Qu.: 78.68025	1st Qu.: 116.18650
Mean: 56.80712	Mean: 104.69588	Mean: 81.74850	Mean: 140.65163
Median: 53.33350	Median: 102.96700	Median: 82.03300	Median: 146.50050
3rd Qu.: 59.38025	3rd Qu.: 126.66675	3rd Qu.: 89.88300	3rd Qu.: 172.08325
Max: 88.00000	Max: 132.06700	Max: 95.66700	Max: 191.06700
Total N: 8.00000	Total N: 8.00000	Total N: 8.00000	Total N: 8.00000
NA's : 0.00000	NA's : 0.00000	NA's : 0.00000	NA's : 0.00000
Std Dev.: 14.26546	Std Dev.: 23.44375	Std Dev.: 12.08267	Std Dev.: 43.35874

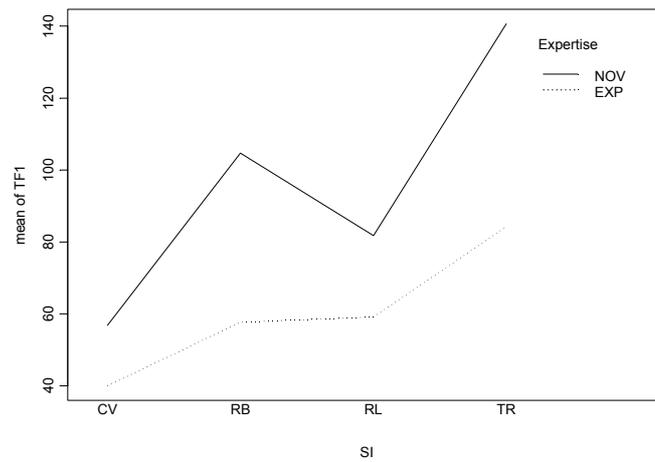


Figure 19: Effet des SI sur le temps d'exécution de TF1

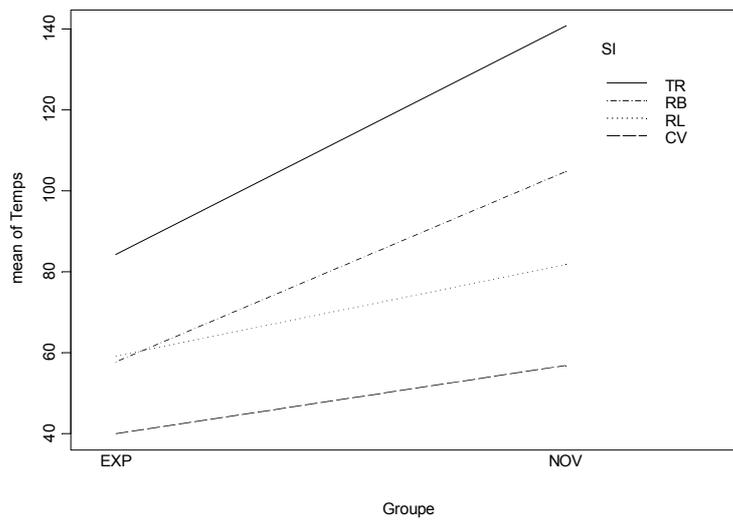


Figure 20: Effet du groupe (expertise) sur le temps d'exécution de TF1

D'après les deux graphiques nous pouvons nous attendre à observer deux effets principaux pour le groupe et pour le SI utilisé sur le temps d'exécution de la tâche. Nous ne pensons pas observer un effet d'interaction entre ces deux variables.

L'ANOVA nous montre comme prédit un effet de l'expertise ( $F = 36.8577$  ;  $p < .00$ ) sur le temps d'exécution de même qu'un effet du SI employé ( $F = 20.498$  ;  $p < .00$ ). Nous observons aussi un faible effet d'interaction entre les deux variables ( $F = 2.6048$  ;  $p < .0607$ ).

Comme nous constatons un effet d'interaction, il est intéressant d'observer quels SI diffèrent significativement entre les experts et les novices. L'analyse des effets simples montrent que RB ( $F =$

15.97 ;  $p < .00019$ ), RL ( $F = 3.7$  ;  $p < .05959$ ) et TR ( $F = 22.96$  ;  $p < .00$ ) diffèrent significativement entre les deux groupes. De plus, les différences dues au groupe d'appartenance apparaissent pour les experts et les novices. Les différences observées ne concernent pas le CV.

SI nous nous concentrons sur le facteur SI indépendamment du facteur expertise nous observons les différences suivantes ( $p = .05$ ) :

- CV – RB
- CV – RL
- CV – TR
- RB – TR
- RL – TR

En intra-groupe, nous constatons, pour le groupe des experts, une différence significative ( $p = .05$ ) entre CV et TR. Pour les sujets novices nous observons quatre différences significatives ( $p = .05$ ) et notamment entre:

- CV – RB
- CV – TR
- RB – TR
- RL – TR

Nombre d'erreurs :

Tableau 11: Analyse descriptive du nombre d'erreurs de TF1 avec les 4 SI par les sujets experts

SI:CV Expertise:EXP	SI:RB Expertise:EXP	SI:RL Expertise:EXP	SI:TR Expertise:EXP
Erreur	Erreur	Erreur	Erreur
Min: 0.000000	Min: 2.000000	Min: 1.000000	Min: 2.000000
1st Qu.: 0.750000	1st Qu.: 2.000000	1st Qu.: 4.250000	1st Qu.: 4.750000
Mean: 1.625000	Mean: 6.375000	Mean: 7.000000	Mean: 6.875000
Median: 1.500000	Median: 4.000000	Median: 6.500000	Median: 6.500000
3rd Qu.: 2.250000	3rd Qu.: 10.250000	3rd Qu.: 11.250000	3rd Qu.: 9.500000
Max: 4.000000	Max: 16.000000	Max: 12.000000	Max: 11.000000
Total N: 8.000000	Total N: 8.000000	Total N: 8.000000	Total N: 8.000000
NA's : 0.000000	NA's : 0.000000	NA's : 0.000000	NA's : 0.000000
Std Dev.: 1.407886	Std Dev.: 5.316752	Std Dev.: 4.342481	Std Dev.: 3.270539

Tableau 12: Analyse descriptive du nombre d'erreurs de TF1 avec les 4 SI par les sujets novices

SI:CV Expertise:NOV	SI:RB Expertise:NOV	SI:RL Expertise:NOV	SI:TR Expertise:NOV
Erreur	Erreur	Erreur	Erreur
Min: 0.000000	Min: 3.000000	Min: 5.000000	Min: 3.000000
1st Qu.: 0.000000	1st Qu.: 4.500000	1st Qu.: 10.500000	1st Qu.: 6.750000
Mean: 1.000000	Mean: 5.375000	Mean: 12.000000	Mean: 10.625000
Median: 0.500000	Median: 5.000000	Median: 12.000000	Median: 10.000000
3rd Qu.: 2.000000	3rd Qu.: 6.250000	3rd Qu.: 14.500000	3rd Qu.: 14.500000
Max: 3.000000	Max: 9.000000	Max: 17.000000	Max: 19.000000
Total N: 8.000000	Total N: 8.000000	Total N: 8.000000	Total N: 8.000000
NA's : 0.000000	NA's : 0.000000	NA's : 0.000000	NA's : 0.000000
Std Dev.: 1.195229	Std Dev.: 1.995531	Std Dev.: 3.891382	Std Dev.: 5.423165

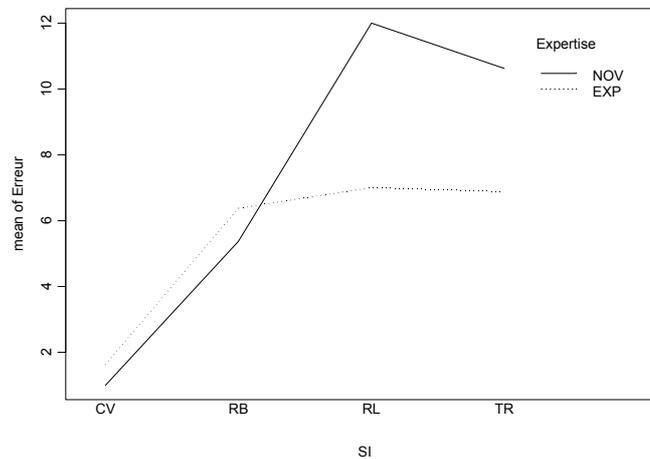


Figure 21: Effet des SI sur le nombre d'erreurs de TF1

Nous observons que les novices et les experts ont effectué pratiquement le même nombre d'erreurs pour CV et RB. Avec RL et TR, les experts commettent moins d'erreurs. Nous pouvons imaginer d'observer un effet du SI sur le nombre d'erreurs pour cette tâche. Nous ne pensons pas observer un effet de l'expertise. Un effet d'interaction pourrait être mis en évidence.

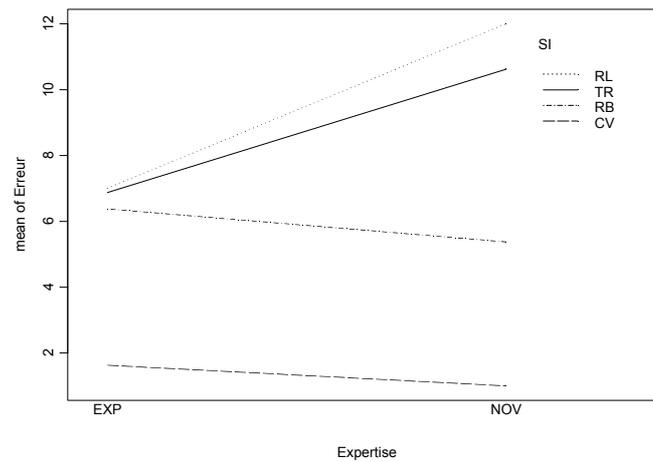


Figure 22: Effet de l'expertise sur le nombre d'erreurs de TF2

L'ANOVA montre un effet du SI ( $F = 16.047$  ;  $p < .00$ ) sur le nombre d'erreurs pour cette tâche. Nous constatons un effet faible de l'expertise ( $F = 3.6998$  ;  $p > .0595$ ) et un effet d'interaction ( $F = 2.69774$  ;  $p > .0544$ ).

Si nous conditionnons la variable expertise nous observons les différences suivantes ( $p = .05$ ) :

- CV – RB
- CV – TR
- CV – RL
- RB – RL

Nous observons aussi des différences significatives suivantes lorsque nous comparons les experts et les novices sur chaque SI (effets simples):

- RL ( $F = 7.29$  ;  $p < .0091$ )
- TR ( $F = 4.1$  ;  $p < .0477$ )

Il est maintenant intéressant de voir quels SI sont différents pour les experts et pour les novices. Nous constatons des différences significatives seulement pour les novices ( $p = .05$ ) :

- CV – RL
- CV – TR
- RB – RL

Satisfaction subjective :

Tableau 13: Analyse descriptive de la satisfaction subjective de TF1 avec les 4 SI par les sujets experts

SI:CV Expertise:EXP	SI:RB Expertise:EXP	SI:RL Expertise:EXP	SI:TR Expertise:EXP
Satisfaction	Satisfaction	Satisfaction	Satisfaction
Min: 8.200000	Min: 5.500000	Min: 3.90000	Min: 1.000000
1st Qu.: 8.575000	1st Qu.: 6.375000	1st Qu.: 6.02500	1st Qu.: 3.100000
Mean: 8.968750	Mean: 7.700000	Mean: 7.22500	Mean: 5.206250
Median: 9.100000	Median: 8.350000	Median: 7.55000	Median: 5.875000
3rd Qu.: 9.337500	3rd Qu.: 8.725000	3rd Qu.: 8.75000	3rd Qu.: 7.125000
Max: 9.500000	Max: 9.400000	Max: 9.50000	Max: 8.800000
Total N: 8.000000	Total N: 8.000000	Total N: 8.00000	Total N: 8.000000
NA's : 0.000000	NA's : 0.000000	NA's : 0.00000	NA's : 0.000000
Std Dev.: 0.4817657	Std Dev.: 1.475514	Std Dev.: 2.05548	Std Dev.: 2.832016

Tableau 14: Analyse descriptive de la satisfaction subjective de TF1 avec les 4 SI par les sujets novices

SI:CV Expertise:NOV	SI:RB Expertise:NOV	SI:RL Expertise:NOV	SI:TR Expertise:NOV
Satisfaction	Satisfaction	Satisfaction	Satisfaction
Min: 3.000000	Min: 6.500000	Min: 6.000000	Min: 2.60000
1st Qu.: 7.600000	1st Qu.: 6.900000	1st Qu.: 7.200000	1st Qu.: 3.87500
Mean: 7.637500	Mean: 7.818750	Mean: 7.700000	Mean: 5.41250
Median: 8.250000	Median: 7.475000	Median: 7.750000	Median: 4.37500
3rd Qu.: 8.675000	3rd Qu.: 9.000000	3rd Qu.: 8.200000	3rd Qu.: 7.58750
Max: 9.500000	Max: 9.500000	Max: 9.500000	Max: 9.00000
Total N: 8.000000	Total N: 8.000000	Total N: 8.000000	Total N: 8.00000
NA's : 0.000000	NA's : 0.000000	NA's : 0.000000	NA's : 0.00000
Std Dev.: 2.094167	Std Dev.: 1.180175	Std Dev.: 1.087592	Std Dev.: 2.37438

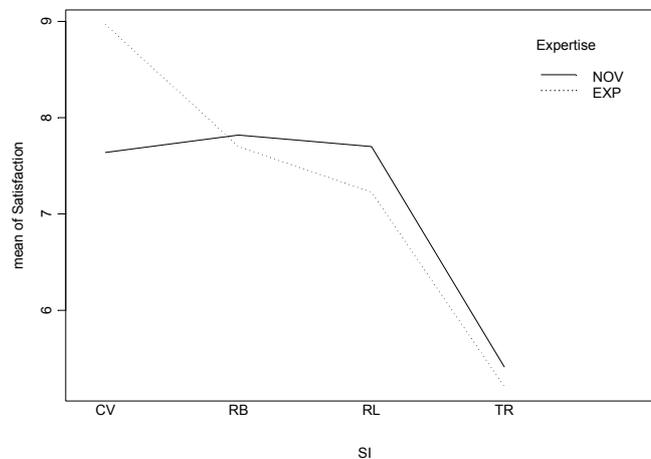


Figure 23: Effet des SI sur la satisfaction subjective de TF1

Nous constatons la même tendance pour les deux groupes à l'exception de la valeur de satisfaction subjective pour CV, qui chez les experts est supérieure à RB ou RL. Nous n'attendons donc pas à observer un effet lié au groupe d'appartenance. Par contre, un effet du SI pourrait être observé.

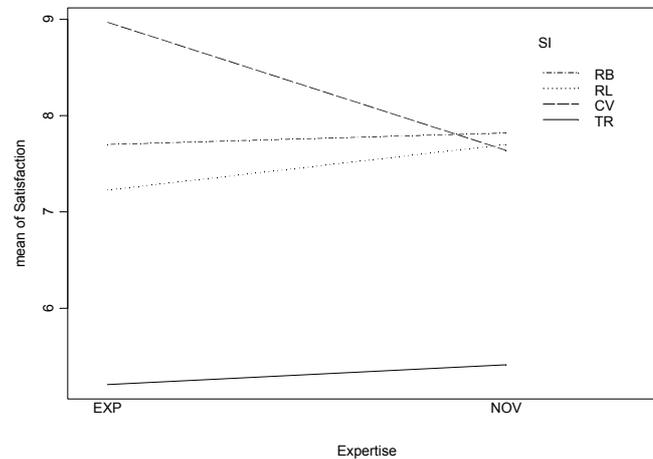


Figure 24: Effet de l'expertise sur la satisfaction subjective de TF1

L'ANOVA montre un effet du SI ( $F = 8.098$  ;  $p < .00$ ). Si nous conditionnons la variable expertise nous obtenons les différences entre SI suivantes ( $p = .05$ ) :

- CV – TR
- RB – TR
- RL - TR

De manière générale nous pouvons dire que cette tâche a mis en évidence, pour le temps d'exécution, pour les erreurs commises et pour la satisfaction subjective, un effet significatif de la variable SI. Nous pouvons donc conclure qu'en fonction du SI utilisé, nos sujets ont obtenu des performances différentes indépendamment du groupe d'appartenance. Pour les erreurs et le temps d'exécution, nos analyses ont démontré aussi des effets liés à l'expertise et à l'interaction entre l'expertise et les SI.

En ce qui concerne l'effet de l'expertise, nous avons constaté qu'il est présent dans les performances en termes de temps d'exécution et du nombre d'erreurs. Nous pouvons donc conclure, d'après les moyennes, que les experts emploient moins de temps pour effectuer cette tâche en commettant aussi moins d'erreurs. Par contre, dans l'évaluation subjective des tâches et du SI utilisé, nous ne rencontrons pas de différences significatives entre les deux groupes.

Si nous ne considérons pas l'expertise comme étant une variable, nos analyses montrent que, du point de vue du temps d'exécution, le CV est le meilleur SI, suivi par RB et RL. Ces dernières n'ayant pas de différences significatives entre eux, et que TR est le SI le moins performant. Du point de vue des erreurs, le CV est encore meilleur SI par rapport aux trois autres, et nous constatons aussi que RB est significativement meilleur que TR. Pour la satisfaction subjective, nous observons des différences significatives d'appréciation entre tous les SI et TR.

Si nous regardons les différences entre les deux groupes, pour le temps d'exécution, nous avons observé que RB, RL et TR diffèrent significativement de CV. Le CV, qui est le meilleur SI, apporte des performances comparables pour les deux groupes. Dans l'analyse du nombre des erreurs commises, CV et RB ne diffèrent pas entre les deux groupes.

Enfin, l'analyse du temps d'exécution nous a permis aussi de regarder à l'intérieur de chaque groupe, quelles étaient les différences significatives au niveau du SI. Pour les experts nous n'avons constaté qu'une différence entre CV et TR. Par contre, chez les novices, tous les SI ont des différences significatives avec TR, de même qu'entre CV et RB. Dans le cas des erreurs, par contre, nous observons des différences significatives seulement à l'intérieur du groupe des novices, notamment entre CV et RL, CV et TR, et entre RB et RL ; par conséquent, au niveau des erreurs, CV est comparable à RB et TR est comparable avec RL.

Il est intéressant souligner que les sujets experts, ont une meilleure évaluation de la difficulté des tâches en évaluant leur satisfaction subjective. En effet, si nous comparons les graphiques des effets des SI pour le temps d'exécution (Figure 19) et pour la satisfaction subjective (Figure 23), nous observons la même tendance. La tendance observée résulte être inversée car pour la satisfaction subjective, plus le score est élevé et plus le SI évalué est satisfaisant. Par conséquent, les deux graphiques, pour les experts, sont comparables. Notons aussi que pour cette tâche, les novices apprécient de manière comparable RB, CV et RL. Cette information résulte contradictoire avec l'analyse du temps d'exécution qui indique que entre ces deux SI il y a une différence significative, où CV est meilleur que RB.

A l'état actuel des analyses, comme d'ailleurs pour l'analyse des autres résultats, nous ne pouvons pas avancer des hypothèses expliquant ce biais d'évaluation de la satisfaction subjective chez les novices. Nous pensons que les données découlant des entretiens nous pourront nous éclairer.

4.2.3 Tâche fermée 2 : Adresse web

Temps d'exécution de la tâche :

Tableau 15: Analyse descriptive du temps d'exécution de TF2 avec les 4 SI par les sujets experts

Expertise:EXP SI:CV	Expertise:EXP SI:RB	Expertise:EXP SI:RL	Expertise:EXP SI:TR
TF2 Min: 12.866000 1st Qu.: 15.717000 Mean: 21.535875 Median: 20.483000 3rd Qu.: 24.140750 Max: 39.000000 Total N: 8.000000 NA's : 0.000000 Std Dev.: 8.331848	TF2 Min: 21.800000 1st Qu.: 25.609750 Mean: 29.498875 Median: 28.533000 3rd Qu.: 34.550500 Max: 37.800000 Total N: 8.000000 NA's : 0.000000 Std Dev.: 5.928953	TF2 Min: 16.733000 1st Qu.: 26.636000 Mean: 33.34575 Median: 30.54450 3rd Qu.: 35.25000 Max: 61.40000 Total N: 8.00000 NA's : 0.00000 Std Dev.: 13.59456	TF2 Min: 29.000000 1st Qu.: 40.38375 Mean: 48.07238 Median: 44.23300 3rd Qu.: 51.70000 Max: 72.11200 Total N: 8.00000 NA's : 0.00000 Std Dev.: 15.12881

Tableau 16: Analyse descriptive du temps d'exécution de TF2 avec les 4 SI par les sujets novices

Expertise:NOV SI:CV	Expertise:NOV SI:RB	Expertise:NOV SI:RL	Expertise:NOV SI:TR
TF2 Min: 23.067000 1st Qu.: 29.980500 Mean: 31.649750 Median: 30.971000 3rd Qu.: 34.500500 Max: 38.667000 Total N: 8.000000 NA's : 0.000000 Std Dev.: 4.868887	TF2 Min: 60.53400 1st Qu.: 65.10075 Mean: 77.39312 Median: 81.84950 3rd Qu.: 86.04975 Max: 92.00000 Total N: 8.00000 NA's : 0.00000 Std Dev.: 12.14345	TF2 Min: 31.66700 1st Qu.: 36.91650 Mean: 53.41313 Median: 44.10000 3rd Qu.: 47.05875 Max: 139.27100 Total N: 8.00000 NA's : 0.00000 Std Dev.: 35.20204	TF2 Min: 78.66600 1st Qu.: 87.33325 Mean: 110.99325 Median: 103.40000 3rd Qu.: 127.80350 Max: 168.06700 Total N: 8.00000 NA's : 0.00000 Std Dev.: 31.64984

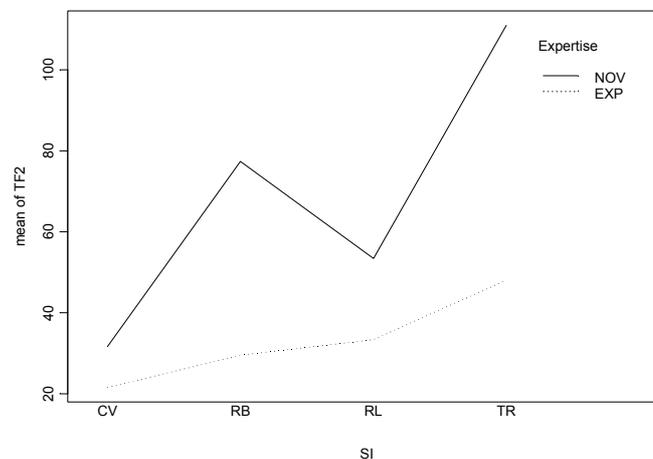


Figure 25: Effet des SI sur le temps d'exécution de TF2

Les écarts des moyennes dans les experts ne semble pas être a priori important. Par contre, des différences sont observées à nouveau chez les novices, où le CV est le SI meilleur suivi par RL, RB et TR.

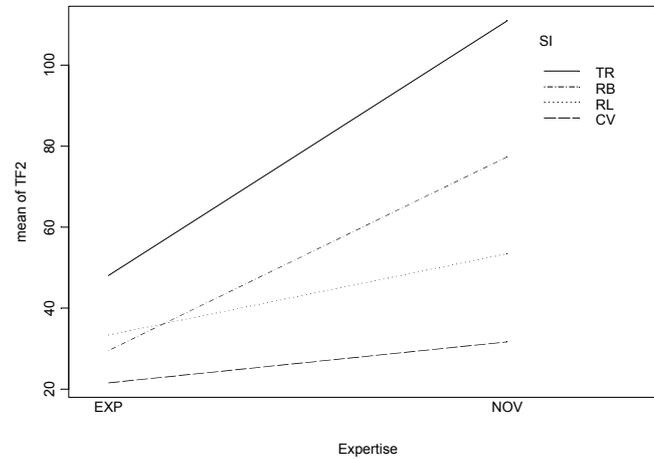


Figure 26: Effet de l'expertise sur le temps d'exécution de TF2

Nous pouvons nous attendre à observer des effets principaux pour les deux variables, de même qu'un effet d'interaction, en raison du fait que l'on constate un croisement des droites dans la Figure 26.

Nous observons que dans chaque modalité, les experts ont de meilleures performances par rapport aux novices. Comme dans le cas précédant, CV est moyennant le meilleur suivi par RL et RB pour les novices et par RB et RL pour les experts. TR est encore une fois le médium le plus onéreux en termes de temps d'exécution.

L'ANOVA nous montre encore une fois un effet de l'expertise ( $F = 54.27368$  ;  $p < 0.00$ ) et un effet du SI ( $F = 21.456$  ;  $p < .00$ ) sur le temps d'exécution. Un effet d'interaction ( $F = 6.508$  ;  $p < .00074$ ) est mis en évidence aussi.

Si nous ne considérons pas le facteur expertise nous observons les comparaisons multiples suivants ( $p = .05$ ):

- CV – RB
- CV – TR
- RB – TR
- RL – TR

Les analyses en intra-groupe ( $p = .05$ ) ne montrent aucune différence dans le groupe des experts alors que chez les novices les fluctuations sont plus importantes :

- CV – RB
- CV – TR
- RB – TR
- RL – TR

Nombre d'erreurs :

Tableau 17: Analyse descriptive du nombre d'erreurs de TF2 avec les 4 SI par les sujets experts

Expertise:EXP SI:CV	Expertise:EXP SI:RB	Expertise:EXP SI:RL	Expertise:EXP SI:TR
Erreur	Erreur	Erreur	Erreur
Min: 0.000000	Min: 0.000000	Min: 0.000000	Min: 0.000000
1st Qu.: 0.000000	1st Qu.: 1.500000	1st Qu.: 2.500000	1st Qu.: 3.750000
Mean: 0.250000	Mean: 2.375000	Mean: 5.000000	Mean: 4.500000
Median: 0.000000	Median: 3.000000	Median: 4.500000	Median: 4.500000
3rd Qu.: 0.000000	3rd Qu.: 3.000000	3rd Qu.: 6.250000	3rd Qu.: 5.500000
Max: 2.000000	Max: 5.000000	Max: 14.000000	Max: 8.000000
Total N: 8.000000	Total N: 8.000000	Total N: 8.000000	Total N: 8.000000
NA's : 0.000000	NA's : 0.000000	NA's : 0.000000	NA's : 0.000000
Std Dev.: 0.7071068	Std Dev.: 1.685018	Std Dev.: 4.342481	Std Dev.: 2.44949

Tableau 18: Analyse descriptive du nombre d'erreurs de TF2 avec les 4 SI par les sujets novices

Expertise:NOV SI:CV	Expertise:NOV SI:RB	Expertise:NOV SI:RL	Expertise:NOV SI:TR
Erreur	Erreur	Erreur	Erreur
Min: 0.000000	Min: 2.000000	Min: 0.000000	Min: 4.000000
1st Qu.: 0.000000	1st Qu.: 4.000000	1st Qu.: 0.000000	1st Qu.: 8.000000
Mean: 0.250000	Mean: 5.000000	Mean: 3.000000	Mean: 9.375000
Median: 0.000000	Median: 4.500000	Median: 1.000000	Median: 8.500000
3rd Qu.: 0.250000	3rd Qu.: 6.250000	3rd Qu.: 2.250000	3rd Qu.: 11.500000
Max: 1.000000	Max: 8.000000	Max: 17.000000	Max: 14.000000
Total N: 8.000000	Total N: 8.000000	Total N: 8.000000	Total N: 8.000000
NA's : 0.000000	NA's : 0.000000	NA's : 0.000000	NA's : 0.000000
Std Dev.: 0.46291	Std Dev.: 1.927248	Std Dev.: 5.781745	Std Dev.: 3.20435

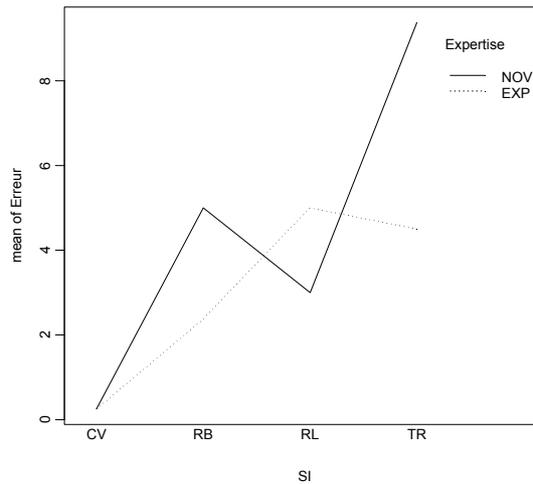


Figure 27: Effet des SI sur le nombre d'erreurs de TF2

Nous constatons que le nombre d'erreurs avec le CV est comparable pour les deux groupes de sujets. Par contre, les experts ont de meilleures performances avec RB et TR par rapport aux novices, qui eux, sont meilleurs avec RL. Nous nous attendons à observer un effet d'interaction, ainsi qu'un effet du SI et un effet faible de l'expertise.

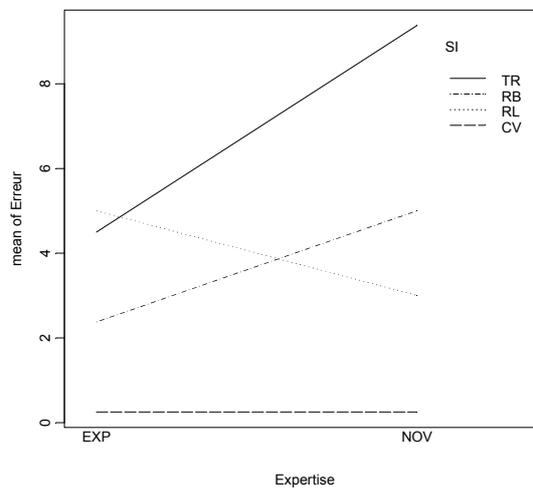


Figure 28: Effet de l'expertise sur le nombre d'erreurs de TF2

L'ANOVA montre un effet du SI ( $F = 12.646$  ;  $p < .00$ ) et un effet d'interaction ( $F = 3.812$  ;  $p < .015$ ). Aucun effet de l'expertise ( $F = 3.192$  ;  $p < .079$ ). Regardons les effets simples liés à l'interaction observée : nous constatons seulement une différence liée au groupe pour TR ( $F = 10.03$  ;  $p < .0025$ ). Les comparaisons multiples à l'intérieur de chaque groupe nous montrent que pour les experts la seule

différence significative se situe entre CV et RL ( $p = .05$ ) ; pour les novices nous en observons trois et plus précisément ( $p = .05$ ):

- CV - RB
- CV - TR
- RL - TR

Si maintenant nous conditionnons la variable expertise, nous observons les différences suivantes ( $p = .05$ ):

- CV – RB
- CV – RL
- CV – TR
- RB – TR
- RL – TR

Satisfaction subjective :

Tableau 19: Analyse descriptive de la satisfaction subjective de TF2 avec les 4 SI par les sujets experts

SI:CV Expertise:EXP	SI:RB Expertise:EXP	SI:RL Expertise:EXP	SI:TR Expertise:EXP
Satisfaction	Satisfaction	Satisfaction	Satisfaction
Min: 8.0500000	Min: 5.5000000	Min: 3.0000000	Min: 1.1000000
1st Qu.: 8.8500000	1st Qu.: 6.5125000	1st Qu.: 6.0500000	1st Qu.: 3.0500000
Mean: 8.9812500	Mean: 7.3812500	Mean: 7.0875000	Mean: 4.8812500
Median: 9.0000000	Median: 7.6500000	Median: 7.5000000	Median: 5.2500000
3rd Qu.: 9.3000000	3rd Qu.: 8.1750000	3rd Qu.: 8.5000000	3rd Qu.: 6.6500000
Max: 9.6000000	Max: 8.8000000	Max: 9.7000000	Max: 8.4500000
Total N: 8.0000000	Total N: 8.0000000	Total N: 8.0000000	Total N: 8.0000000
NA's : 0.0000000	NA's : 0.0000000	NA's : 0.0000000	NA's : 0.0000000
Std Dev.: 0.5371469	Std Dev.: 1.209468	Std Dev.: 2.268062	Std Dev.: 2.576534

Tableau 20: Analyse descriptive de la satisfaction subjective de TF2 avec les 4 SI par les sujets novices

SI:CV Expertise:NOV	SI:RB Expertise:NOV	SI:RL Expertise:NOV	SI:TR Expertise:NOV
Satisfaction	Satisfaction	Satisfaction	Satisfaction
Min: 8.000000	Min: 5.500000	Min: 2.000000	Min: 0.500000
1st Qu.: 8.275000	1st Qu.: 6.675000	1st Qu.: 4.650000	1st Qu.: 1.575000
Mean: 8.5937500	Mean: 7.312500	Mean: 5.950000	Mean: 2.437500
Median: 8.4750000	Median: 7.225000	Median: 6.700000	Median: 2.700000
3rd Qu.: 8.7250000	3rd Qu.: 8.050000	3rd Qu.: 7.300000	3rd Qu.: 3.525000
Max: 9.6000000	Max: 8.850000	Max: 9.000000	Max: 4.000000
Total N: 8.0000000	Total N: 8.000000	Total N: 8.000000	Total N: 8.000000
NA's : 0.0000000	NA's : 0.000000	NA's : 0.000000	NA's : 0.000000
Std Dev.: 0.5240212	Std Dev.: 1.091117	Std Dev.: 2.573491	Std Dev.: 1.344765

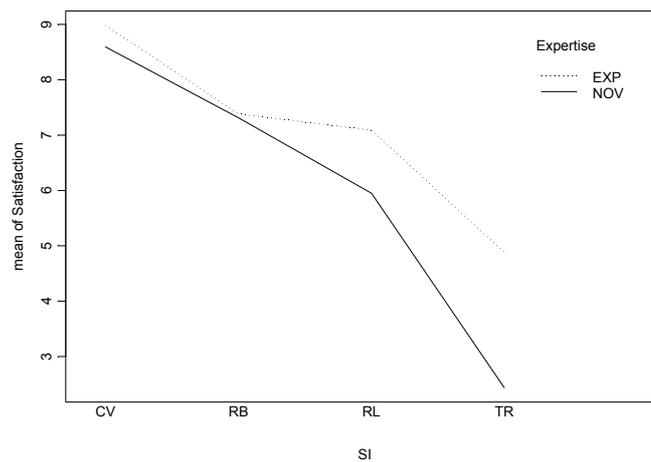


Figure 29: Effet des SI sur la satisfaction subjective de TF2

Aussi dans cette tâche les deux groupes ont la même tendance. Comme pour TF1 nous n'attendons à observer qu'un effet de SI sur la satisfaction subjective.

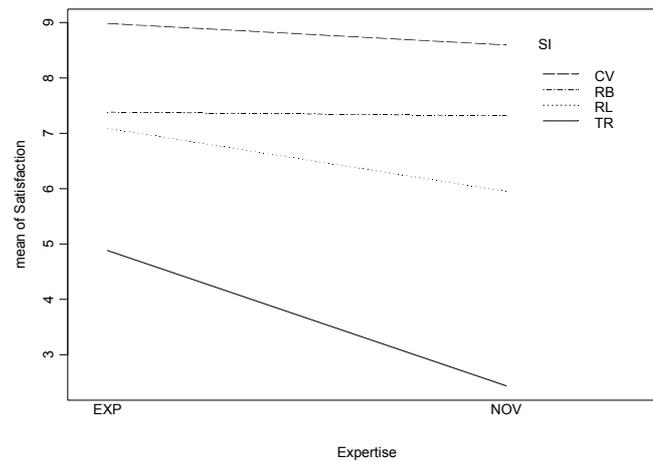


Figure 30: Effet de l'expertise sur la satisfaction subjective de TF2

L'ANOVA montre un effet de SI ( $F = 25.485$ ;  $p < .00$ ) comme prédit. Mais nous observons aussi un effet de l'expertise: ( $F = 5.566$ ;  $p < .022$ ). Si nous conditionnons la variable expertise nous obtenons les différences suivantes lors de nos comparaisons ( $p = .05$ ):

- CV – RL
- CV – TR
- RB – TR
- RL - TR

Pour résumer les analyses effectuées pour la tâche fermée 2, nous avons constaté dans les trois conditions évaluées, temps d'exécution, erreurs et satisfaction subjective, un effet lié au SI utilisé. Pour le temps d'exécution et la satisfaction subjective, nous avons mis en évidence aussi un effet de l'expertise. Un effet d'interaction a été aussi observé pour le temps d'exécution et le nombre d'erreurs.

La comparaison plus approfondie des trois analyses, nous montre que si nous ne considérons pas la variable expertise, nous observons que dans toutes les analyses, CV, RB et RL diffèrent toujours significativement de TR. Dans l'analyse des erreurs, le CV diffère aussi de RB, comme dans l'analyse du temps d'exécution, et de RL, comme dans l'analyse de l'évaluation subjective. Dans les trois analyses, RB et RL ne diffèrent jamais significativement.

Les résultats intra-groupe, montrent encore une fois une plus grande fluctuation des résultats chez les novices, en particulier, pour le temps d'exécution tous les SI diffèrent de TR et en plus CV de RB ; dans l'analyse des erreurs, seulement CV et RL diffèrent significativement de TR, de plus CV diffère aussi de

RB. Chez les experts par contre, nous n'observons qu'une différence entre CV et RL dans l'analyse des erreurs.

Les différences inter-groupe observées, se situent au niveau de TR pour le nombre d'erreurs.

Encore une fois, nous pouvons conclure que CV est le meilleur SI et que TR est le moins performant. Par contre, RB et RL sont toujours assez comparables se situant au milieu des deux autres. Il est intéressant aussi d'observer que les différences de SI à l'intérieur du groupe des novices sont pratiquement les mêmes dans la condition des erreurs et du temps d'exécution, à l'exception de RB-TR présente seulement dans cette dernière condition. Comme nous n'avons pas observé d'effet d'interaction dans la condition de la satisfaction subjective, nous ne pouvons pas observer les attitudes des deux groupes séparément. Néanmoins, nous pouvons souligner que si nous ne considérons pas la variable expertise, dans les trois conditions tous les SI sont significativement différents de TR.

#### 4.2.4 Tâche fermée 3 : Rendez-vous

##### Temps d'exécution de la tâche :

Tableau 21: Analyse descriptive du temps d'exécution de TF3 avec les 4 SI par les sujets experts

Expertise:EXP SI:CV	Expertise:EXP SI:RB	Expertise:EXP SI:RL	Expertise:EXP SI:TR
Temps	Temps	Temps	Temps
Min: 14.53300	Min: 17.86700	Min: 26.40000	Min: 29.73300
1st Qu.: 21.96725	1st Qu.: 25.27325	1st Qu.: 36.11600	1st Qu.: 36.86675
Mean: 25.17638	Mean: 47.45950	Mean: 46.69987	Mean: 45.13212
Median: 25.13300	Median: 33.33350	Median: 42.13350	Median: 41.39450
3rd Qu.: 27.63300	3rd Qu.: 67.66675	3rd Qu.: 55.56650	3rd Qu.: 43.35050
Max: 34.12300	Max: 109.26600	Max: 79.40000	Max: 91.13400
Total N: 8.00000	Total N: 8.00000	Total N: 8.00000	Total N: 8.00000
NA's : 0.00000	NA's : 0.00000	NA's : 0.00000	NA's : 0.00000
Std Dev.: 6.37833	Std Dev.: 31.92405	Std Dev.: 16.71809	Std Dev.: 19.26677

Tableau 22: Analyse descriptive du temps d'exécution de TF3 avec les 4 SI par les sujets novices

Expertise:NOV SI:CV	Expertise:NOV SI:RB	Expertise:NOV SI:RL	Expertise:NOV SI:TR
Temps	Temps	Temps	Temps
Min: 18.400000	Min: 32.46700	Min: 42.12200	Min: 19.00000
1st Qu.: 24.783000	1st Qu.: 61.68325	1st Qu.: 52.48075	1st Qu.: 26.18350
Mean: 27.790250	Mean: 66.61675	Mean: 57.73313	Mean: 47.96412
Median: 25.966500	Median: 68.10000	Median: 56.53300	Median: 45.22800
3rd Qu.: 33.031000	3rd Qu.: 79.26675	3rd Qu.: 59.24975	3rd Qu.: 63.97550
Max: 36.333000	Max: 83.26700	Max: 83.93300	Max: 90.53300
Total N: 8.000000	Total N: 8.00000	Total N: 8.00000	Total N: 8.00000
NA's : 0.000000	NA's : 0.00000	NA's : 0.00000	NA's : 0.00000
Std Dev.: 6.052051	Std Dev.: 16.70437	Std Dev.: 12.14633	Std Dev.: 26.20552

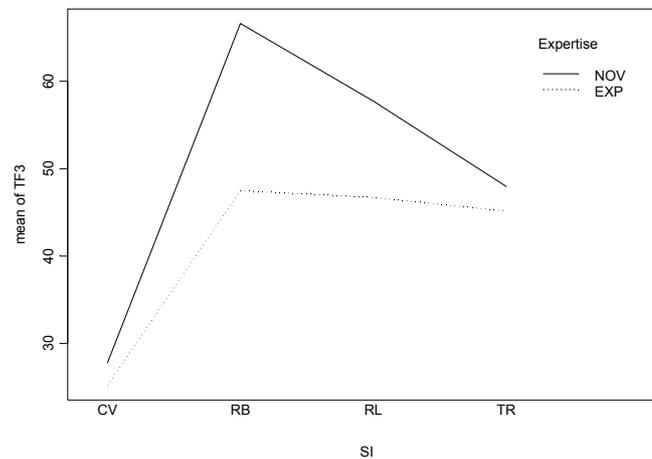


Figure 31: Effet des SI sur le temps d'exécution de TF3

La Figure 31 montre deux courbes inédites. Les novices ont comme pour les tâches fermées précédentes les meilleurs performances avec CV, mais ils sont suivi par TR, RL et RB. Les experts ont eux aussi effectués les meilleures performances avec CV ; les trois autres SI sont comparables. D'après cette figure nous pensons observer un effet principal du facteur SI.

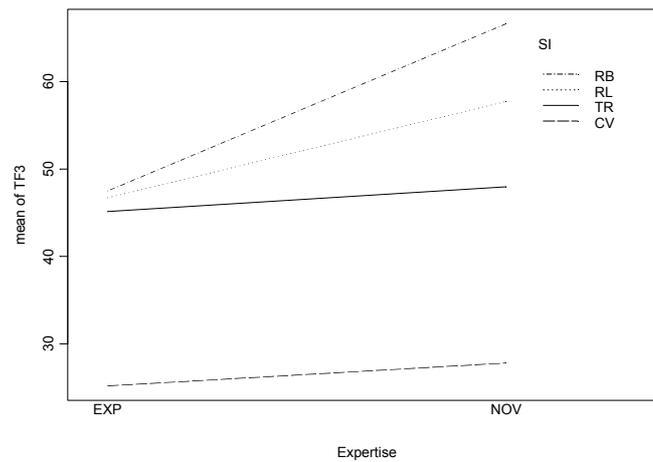


Figure 32: Effet de l'expertise sur le temps d'exécution de TF3

La Figure 32 nous laisse penser que nous n'observerons pas d'effets liés au groupe, car les différences dans les moyennes des deux groupes sont moindres, à l'exception de RB. Il est intéressant d'observer que pour cette tâche les performances entre les deux groupes avec CV et TR sont pratiquement identiques. Nous ne pensons donc pas observer un effet principal du facteur expertise, de même qu'un effet d'interaction.

L'ANOVA confirme nos prédictions : un constate seulement un effet du SI sur le temps d'exécution de la tâche fermée 3 ( $F = 8.0688$  ;  $p < .0001477$ ). Les comparaisons multiples significatives à l'intérieur du facteur SI sont les suivantes ( $p = .05$ ) :

- CV – RB
- CV – RL
- CV - TR

Le CV est donc significativement le meilleur SI pour cette tâche.

Nombre d'erreurs :

Tableau 23: Analyse descriptive des erreurs de TF3 avec les 4 SI par les sujets experts

SI:CV Expertise:EXP	SI:RB Expertise:EXP	SI:RL Expertise:EXP	SI:TR Expertise:EXP
Erreur Min: 0.000000	Erreur Min: 2.000000	Erreur Min: 1.000000	Erreur Min: 2.000000
1st Qu.: 0.750000	1st Qu.: 2.000000	1st Qu.: 4.250000	1st Qu.: 4.750000
Mean: 1.625000	Mean: 6.375000	Mean: 7.000000	Mean: 6.875000
Median: 1.500000	Median: 4.000000	Median: 6.500000	Median: 6.500000
3rd Qu.: 2.250000	3rd Qu.: 10.250000	3rd Qu.: 11.250000	3rd Qu.: 9.500000
Max: 4.000000	Max: 16.000000	Max: 12.000000	Max: 11.000000
Total N: 8.000000	Total N: 8.000000	Total N: 8.000000	Total N: 8.000000
NA's : 0.000000	NA's : 0.000000	NA's : 0.000000	NA's : 0.000000
Std Dev.: 1.407886	Std Dev.: 5.316752	Std Dev.: 4.342481	Std Dev.: 3.270539

Tableau 24: Analyse descriptive des erreurs de TF3 avec les 4 SI par les sujets novices

SI:CV Expertise:NOV	SI:RB Expertise:NOV	SI:RL Expertise:NOV	SI:TR Expertise:NOV
Erreur Min: 0.000000	Erreur Min: 3.000000	Erreur Min: 5.000000	Erreur Min: 3.000000
1st Qu.: 0.000000	1st Qu.: 4.500000	1st Qu.: 10.500000	1st Qu.: 6.750000
Mean: 1.000000	Mean: 5.375000	Mean: 12.000000	Mean: 10.625000
Median: 0.500000	Median: 5.000000	Median: 12.000000	Median: 10.000000
3rd Qu.: 2.000000	3rd Qu.: 6.250000	3rd Qu.: 14.500000	3rd Qu.: 14.500000
Max: 3.000000	Max: 9.000000	Max: 17.000000	Max: 19.000000
Total N: 8.000000	Total N: 8.000000	Total N: 8.000000	Total N: 8.000000
NA's : 0.000000	NA's : 0.000000	NA's : 0.000000	NA's : 0.000000
Std Dev.: 1.195229	Std Dev.: 1.995531	Std Dev.: 3.891382	Std Dev.: 5.423165

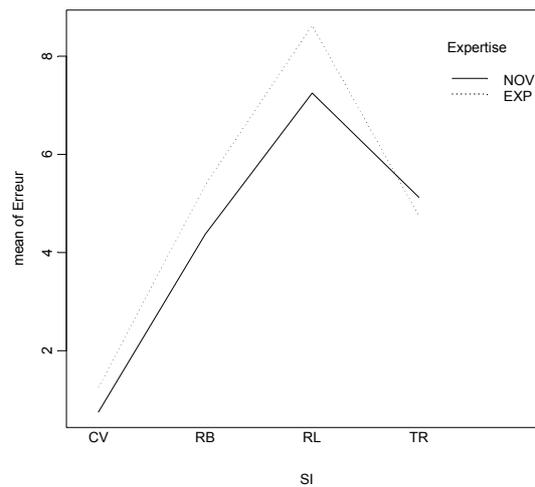


Figure 33: Effet des SI sur les erreurs de TF3

D'après les deux graphiques, nous pouvons observer que les deux groupes ont le même type de trend avec des performances comparables. Le seul effet que nous nous attendons d'observer est celui du SI.

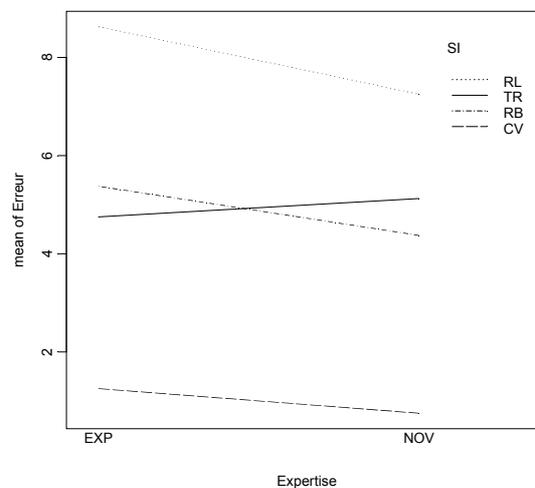


Figure 34: Effet de l'expertise sur les erreurs de TF3

L'ANOVA sur montre que seul l'effet du SI est présent ( $F = 6.632 ; p < .0007$ ). Si nous conditionnons la variable expertise, nous observons que la seule différence significative se situe entre CV et RL ( $p = .05$ ).

Satisfaction subjective :

Tableau 25: Analyse descriptive de la satisfaction subjective de TF3 avec les 4 SI par les sujets experts

Expertise:EXP SI:CV	Expertise:EXP SI:RB	Expertise:EXP SI:RL	Expertise:EXP SI:TR
Satisfaction	Satisfaction	Satisfaction	Satisfaction
Min: 7.900000	Min: 5.500000	Min: 3.900000	Min: 0.900000
1st Qu.: 8.775000	1st Qu.: 6.787500	1st Qu.: 6.050000	1st Qu.: 3.125000
Mean: 8.987500	Mean: 7.900000	Mean: 7.275000	Mean: 5.650000
Median: 9.100000	Median: 8.375000	Median: 8.200000	Median: 7.050000
3rd Qu.: 9.375000	3rd Qu.: 8.950000	3rd Qu.: 8.525000	3rd Qu.: 7.750000
Max: 9.600000	Max: 9.600000	Max: 9.400000	Max: 8.700000
Total N: 8.000000	Total N: 8.000000	Total N: 8.000000	Total N: 8.000000
NA's : 0.000000	NA's : 0.000000	NA's : 0.000000	NA's : 0.000000
Std Dev.: 0.5890367	Std Dev.: 1.449384	Std Dev.: 1.963779	Std Dev.: 3.053803

Tableau 26: Analyse descriptive de la satisfaction subjective de TF3 avec les 4 SI par les sujets novices

Expertise:NOV SI:CV	Expertise:NOV SI:RB	Expertise:NOV SI:RL	Expertise:NOV SI:TR
Satisfaction	Satisfaction	Satisfaction	Satisfaction
Min: 7.000000	Min: 5.450000	Min: 5.500000	Min: 2.500000
1st Qu.: 8.262500	1st Qu.: 6.625000	1st Qu.: 6.500000	1st Qu.: 3.725000
Mean: 8.500000	Mean: 7.506250	Mean: 7.662500	Mean: 5.500000
Median: 8.475000	Median: 7.750000	Median: 8.050000	Median: 4.900000
3rd Qu.: 8.825000	3rd Qu.: 8.500000	3rd Qu.: 8.550000	3rd Qu.: 7.750000
Max: 9.600000	Max: 9.600000	Max: 9.500000	Max: 9.000000
Total N: 8.000000	Total N: 8.000000	Total N: 8.000000	Total N: 8.000000
NA's : 0.000000	NA's : 0.000000	NA's : 0.000000	NA's : 0.000000
Std Dev.: 0.8241879	Std Dev.: 1.469314	Std Dev.: 1.354292	Std Dev.: 2.580698

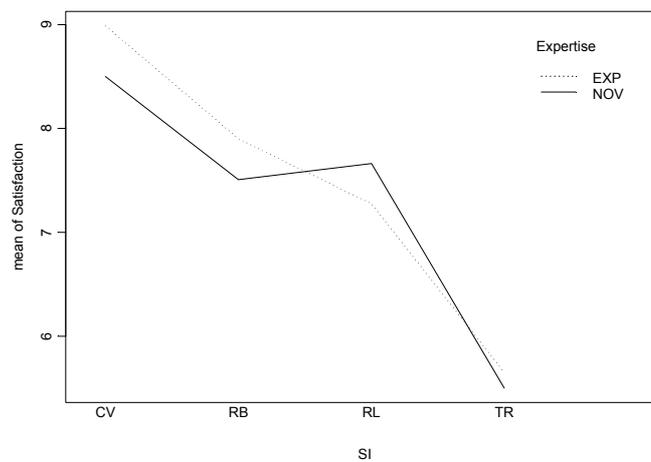


Figure 35: Effet du SI sur la satisfaction subjective de TF3

D'après les deux graphiques affichant les effets de SI et de l'expertise, nous nous attendons à observer que l'effet de SI.

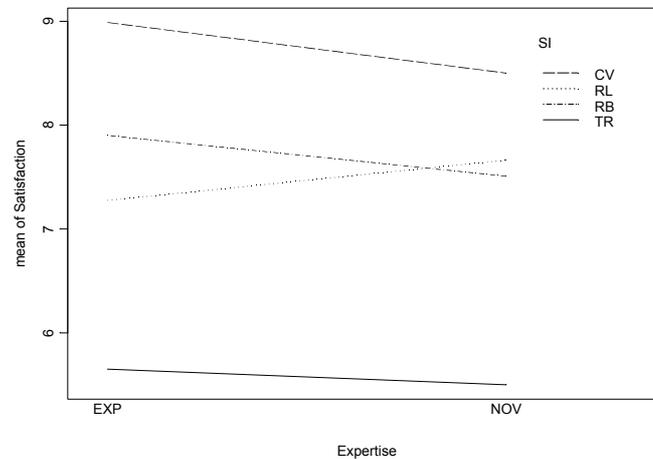


Figure 36: effet de l'expertise sur la satisfaction subjective

L'ANOVA indique que il y a seulement l'effet de SI comme nous l'avons prédit ( $F = 8.276$  ;  $p < .00$ ). Si nous ne considérons pas la variable expertise nous obtenons les différences suivantes ( $p = .05$ ) :

- CV – TR
- RB – TR
- RL - TR

Cette tâche est difficilement analysable car dans les trois conditions, nous n'observons qu'un effet significatif des SI. Si nous considérons seulement cette dernière variable, nous n'obtenons pas d'analyses comparables pour les trois condition. Néanmoins, notons que pour le temps d'exécution CV diffère significativement avec les trois autres SI ; dans l'analyse des erreurs, seulement la différence entre CV et RL a été mise en évidence ; dans le cas de la satisfaction subjective, par contre tous les SI diffèrent significativement de TR.

Malgré le faible nombre d'analyses pour cette tâche il est intéressant d'observer que de manière générale, TR n'est pas apprécié par nos 16 sujets, alors que dans le temps d'exécution ou dans le nombre d'erreurs cette tendance n'a pas été mise en évidence. De plus, ce type de tâche ne montre pas de différences significatives liées au temps d'exécution entre nos deux groupes expérimentaux.

4.2.5 Tâche fermée 4 : Email

Temps d'exécution de la tâche :

Tableau 27: Analyse descriptive du temps d'exécution de TF4 avec les 4 SI par les sujets experts

Expertise:EXP SI:CV	Expertise:EXP SI:RB	Expertise:EXP SI:RL	Expertise:EXP SI:TR
Temps	Temps	Temps	Temps
Min: 108.66700	Min: 94.8670	Min: 149.73300	Min: 145.6000
1st Qu.: 120.55025	1st Qu.: 130.3800	1st Qu.: 215.13350	1st Qu.: 202.6168
Mean: 138.48075	Mean: 186.1569	Mean: 224.14088	Mean: 314.4585
Median: 131.79500	Median: 147.8000	Median: 218.36350	Median: 314.1670
3rd Qu.: 148.23075	3rd Qu.: 199.4497	3rd Qu.: 221.38300	3rd Qu.: 370.6835
Max: 181.93300	Max: 407.4670	Max: 335.26700	Max: 568.9330
Total N: 8.00000	Total N: 8.0000	Total N: 8.00000	Total N: 8.0000
NA's : 0.00000	NA's : 0.0000	NA's : 0.00000	NA's : 0.0000
Std Dev.: 26.63227	Std Dev.: 101.3689	Std Dev.: 51.05961	Std Dev.: 136.4244

Tableau 28: Analyse descriptive du temps d'exécution de TF4 avec les 4 SI par les sujets novices

Expertise:NOV SI:CV	Expertise:NOV SI:RB	Expertise:NOV SI:RL	Expertise:NOV SI:TR
Temps	Temps	Temps	Temps
Min: 145.06700	Min: 189.86700	Min: 223.40000	Min: 254.40000
1st Qu.: 157.39950	1st Qu.: 255.06700	1st Qu.: 242.98175	1st Qu.: 267.16700
Mean: 183.60825	Mean: 292.55012	Mean: 278.54088	Mean: 337.78762
Median: 168.86650	Median: 289.86650	Median: 257.23350	Median: 326.96700
3rd Qu.: 177.51675	3rd Qu.: 310.11700	3rd Qu.: 310.93325	3rd Qu.: 384.67500
Max: 313.73300	Max: 428.33300	Max: 368.00000	Max: 478.73300
Total N: 8.00000	Total N: 8.00000	Total N: 8.00000	Total N: 8.00000
NA's : 0.00000	NA's : 0.00000	NA's : 0.00000	NA's : 0.00000
Std Dev.: 54.52867	Std Dev.: 70.74161	Std Dev.: 52.63238	Std Dev.: 81.05466

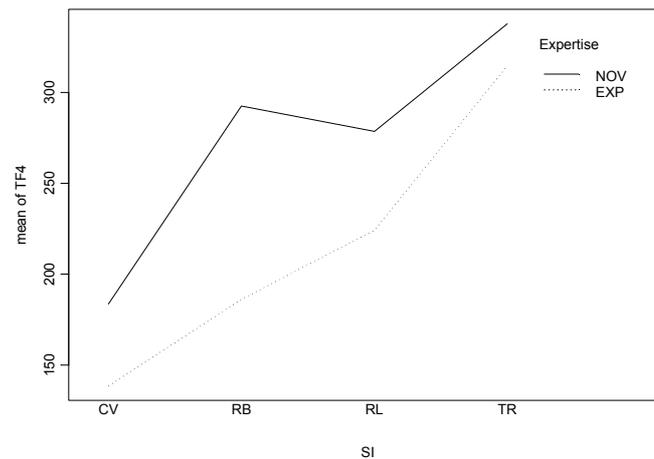


Figure 37: Effet des SI sur le temps d'exécution de TF4

Nous retrouvons dans cette tâche la même tendance que les deux premières tâches pour les deux groupes (Figure 37). Constatons que les deux groupes ont des performances comparables avec TR qui implique moyennement le plus de temps d'exécution de la tâche. Pour les experts le meilleur SI est CV suivi par RB et RL ; pour les novices, CV est le meilleur suivi par RL et RB. Nous pouvons nous attendre à observer un effet principal lié au SI de même qu'un effet d'interaction.

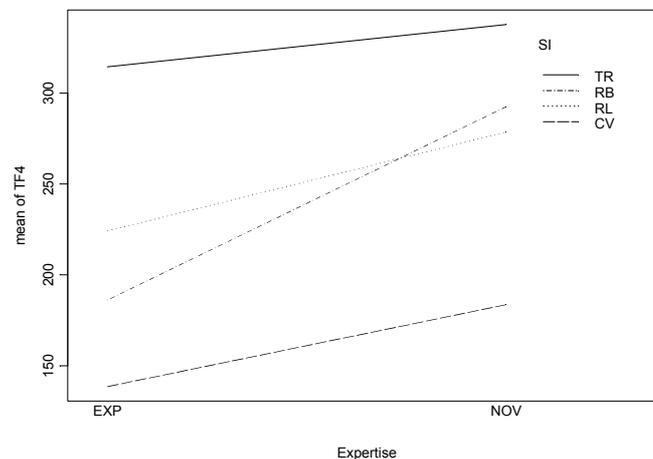


Figure 38: Effet de l'expertise sur le temps d'exécution de TF4

De manière générale, nous observons (Figure 38) que les sujets experts ont à nouveau de meilleures performances sur l'ensemble des SI utilisés.

L'ANOVA montre un effet de l'expertise ( $F = 8.49015$  ;  $p < .0051$ ) et un effet du SI ( $F = 11.804$  ;  $p < .00$ ) sur le temps d'exécution pour la tâche fermée 4. L'analyse des contrastes montre, indépendamment du facteur expertise, les différences suivantes :

- CV – RB
- CV – RL
- CV – TR
- RB – TR
- RL – TR

Nombre d'erreurs :

Tableau 29: Analyse descriptive du nombre d'erreurs de TF4 avec les 4 SI par les sujets experts

SI:CV Expertise:EXP	SI:RB Expertise:EXP	SI:RL Expertise:EXP	SI:TR Expertise:EXP
Erreur	Erreur	Erreur	Erreur
Min: 2.000000	Min: 11.000000	Min: 16.000000	Min: 9.000000
1st Qu.: 2.750000	1st Qu.: 14.750000	1st Qu.: 23.250000	1st Qu.: 22.500000
Mean: 5.625000	Mean: 21.750000	Mean: 33.250000	Mean: 36.625000
Median: 4.500000	Median: 19.000000	Median: 26.000000	Median: 35.500000
3rd Qu.: 6.500000	3rd Qu.: 26.750000	3rd Qu.: 38.500000	3rd Qu.: 44.000000
Max: 15.000000	Max: 41.000000	Max: 71.000000	Max: 83.000000
Total N: 8.000000	Total N: 8.000000	Total N: 8.000000	Total N: 8.000000
NA's : 0.000000	NA's : 0.000000	NA's : 0.000000	NA's : 0.000000
Std Dev.: 4.307386	Std Dev.: 9.852483	Std Dev.: 17.89453	Std Dev.: 22.96542

Tableau 30: Analyse descriptive du nombre d'erreurs de TF4 avec les 4 SI par les sujets novices

SI:CV Expertise:NOV	SI:RB Expertise:NOV	SI:RL Expertise:NOV	SI:TR Expertise:NOV
Erreur	Erreur	Erreur	Erreur
Min: 1.000000	Min: 8.000000	Min: 15.000000	Min: 8.000000
1st Qu.: 2.500000	1st Qu.: 8.750000	1st Qu.: 29.500000	1st Qu.: 17.250000
Mean: 4.750000	Mean: 14.625000	Mean: 36.750000	Mean: 27.125000
Median: 3.000000	Median: 14.000000	Median: 38.000000	Median: 27.000000
3rd Qu.: 4.750000	3rd Qu.: 17.500000	3rd Qu.: 46.750000	3rd Qu.: 35.500000
Max: 16.000000	Max: 26.000000	Max: 54.000000	Max: 46.000000
Total N: 8.000000	Total N: 8.000000	Total N: 8.000000	Total N: 8.000000
NA's : 0.000000	NA's : 0.000000	NA's : 0.000000	NA's : 0.000000
Std Dev.: 4.920801	Std Dev.: 6.631903	Std Dev.: 13.45628	Std Dev.: 13.3142

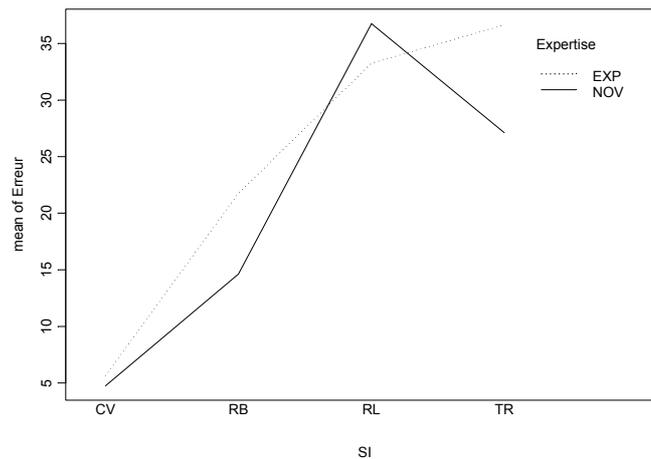


Figure 39: Effet des SI sur le nombre d'erreurs de TF4

Nous pouvons observer que les deux groupes suivent la même tendance du point de vue du nombre d'erreurs commis lors de l'exécution de la tâche fermée 4. Notons aussi que contrairement aux novices, les experts effectuent plus d'erreurs avec TR qu'avec RL. Pour les deux groupes CV est le meilleur SI suivi par RB.

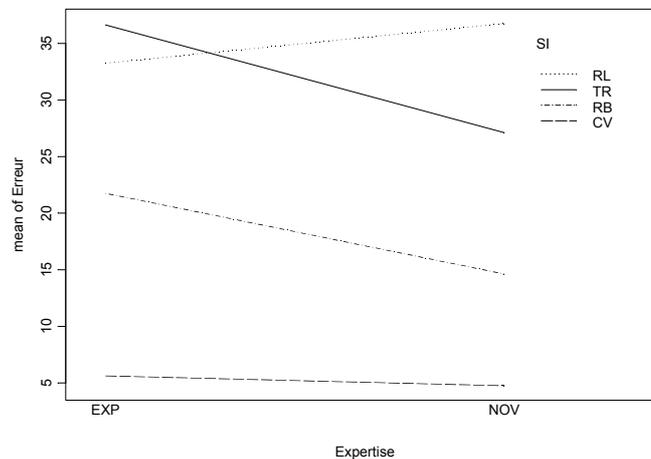


Figure 40: Effet de l'expertise sur le nombre d'erreurs de TF4

Nous pouvons nous attendre à observer un effet du SI mais pas d'effet d'interaction ni de l'expertise.

L'ANOVA montre en effet que seul l'effet dû au SI n'est présent ( $F = 17.267$  ;  $p < .00$ ). En conditionnant l'expertise nous observons les différences significatives suivantes ( $p = .05$ ) :

- CV – RB
- CV – RL
- CV – TR
- RB – RL
- RB – TR
- RL - TR

Satisfaction subjective :

Tableau 31: Analyse descriptive de la satisfaction subjective de TF4 avec les 4 SI par les sujets experts

SI:CV Expertise:EXP	SI:RB Expertise:EXP	SI:RL Expertise:EXP	SI:TR Expertise:EXP
Satisfaction	Satisfaction	Satisfaction	Satisfaction
Min: 8.000000	Min: 2.900000	Min: 2.900000	Min: 0.900000
1st Qu.: 8.450000	1st Qu.: 6.212500	1st Qu.: 6.175000	1st Qu.: 2.850000
Mean: 8.737500	Mean: 7.006250	Mean: 6.850000	Mean: 5.012500
Median: 8.700000	Median: 7.950000	Median: 7.300000	Median: 5.500000
3rd Qu.: 8.975000	3rd Qu.: 8.250000	3rd Qu.: 7.950000	3rd Qu.: 7.525000
Max: 9.600000	Max: 8.700000	Max: 9.400000	Max: 8.500000
Total N: 8.000000	Total N: 8.000000	Total N: 8.000000	Total N: 8.000000
NA's : 0.000000	NA's : 0.000000	NA's : 0.000000	NA's : 0.000000
Std Dev.: 0.5153016	Std Dev.: 1.982141	Std Dev.: 2.026256	Std Dev.: 2.965727

Tableau 32: Analyse descriptive de la satisfaction subjective de TF4 avec les 4 SI par les sujets novices

SI :CV Expertise :NOV	SI:RB Expertise:NOV	SI:RL Expertise:NOV	SI:TR Expertise:NOV
Satisfaction	Satisfaction	Satisfaction	Satisfaction
Min : 2.000000	Min: 3.100000	Min: 1.000000	Min: 1.4000
1 <sup>st</sup> Qu.: 7.375000	1st Qu.: 5.950000	1st Qu.: 5.537500	1st Qu.: 3.2750
Mean: 7.362500	Mean: 6.737500	Mean: 6.362500	Mean: 4.3250
Median: 8.000000	Median: 7.250000	Median: 6.925000	Median: 4.0000
3 <sup>rd</sup> Qu.: 8.425000	3rd Qu.: 8.050000	3rd Qu.: 7.625000	3rd Qu.: 6.1750
Max: 9.500000	Max: 8.500000	Max: 9.500000	Max: 6.6000
Total N: 8.000000	Total N: 8.000000	Total N: 8.000000	Total N: 8.0000
NA's : 0.000000	NA's : 0.000000	NA's : 0.000000	NA's : 0.0000
Std Dev.: 2.289066	Std Dev.: 1.799157	Std Dev.: 2.525407	Std Dev.: 1.8919

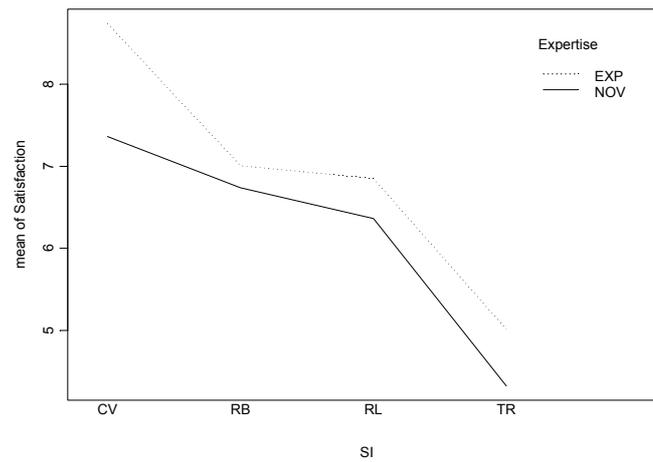


Figure 41: Effet de SI sur TF4

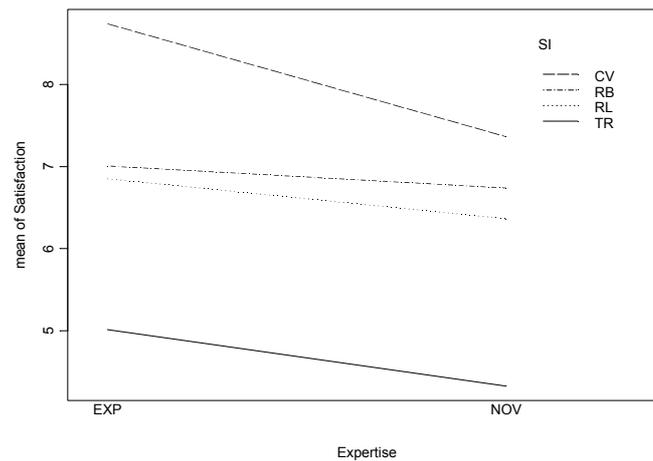


Figure 42: Effet de l'expertise sur TF4

Nous pouvons prévoir d'observer, d'après les deux figures ci-dessus, seulement un effet de SI sur la satisfaction subjective. En effet, l'ANOVA nous le confirme ( $F = 7.08$  ;  $p < .00$ ). Si nous ne considérons pas la variable expertise nous obtenons les différences suivantes ( $p = .05$ ) :

- CV – TR
- RB - TR

Pour cette tâche, nous avons constaté pour les trios conditions seulement un effet du SI, mis à par pour le temps d'exécution qui a aussi un effet lié à l'expertise. Si nous ne considérons pas la variable expertise, nous observons que, dans les conditions temps d'exécution et erreurs, les deux montrent les mêmes

différences, à savoir que CV diffère des trois autres SI, et que RB et RL diffèrent de TR ; de plus, dans la condition des erreurs, RB a aussi significativement moins d'erreurs que RL. La condition de la satisfaction subjective, montre seulement les différences entre CV et TR et RB et TR.

#### 4.2.6 Nombre moyen d'erreurs

Nous avons calculé le nombre moyen d'erreur sur les 4 tâches fermées. Nous voulons maintenant analyser les erreurs, afin de voir si nous obtenons les mêmes résultats que pour le temps d'exécution.

Tableau 33: Statistique descriptive du nombre moyen d'erreurs du groupe expert pour les 4 tâches avec chaque SI

Expertise:EXP SI:CV	Expertise:EXP SI:RB	Expertise:EXP SI:RL	Expertise:EXP SI:TR
Erreurs	Erreurs	Erreurs	Erreurs
Min: 0.750000	Min: 4.250000	Min: 6.000000	Min: 4.250000
1st Qu.: 1.687500	1st Qu.: 5.937500	1st Qu.: 8.875000	1st Qu.: 10.312500
Mean: 2.187500	Mean: 8.968750	Mean: 13.468750	Mean: 13.187500
Median: 2.000000	Median: 6.500000	Median: 10.625000	Median: 12.375000
3rd Qu.: 2.125000	3rd Qu.: 12.500000	3rd Qu.: 15.312500	3rd Qu.: 15.187500
Max: 5.000000	Max: 17.750000	Max: 29.750000	Max: 26.500000
Total N: 8.000000	Total N: 8.000000	Total N: 8.000000	Total N: 8.000000
NA's : 0.000000	NA's : 0.000000	NA's : 0.000000	NA's : 0.000000
Std Dev.: 1.244631	Std Dev.: 4.839269	Std Dev.: 7.752232	Std Dev.: 6.659781

Tableau 34: Statistique descriptive du nombre moyen d'erreurs du groupe novice pour les 4 tâches avec chaque SI

Expertise:NOV SI:CV	Expertise:NOV SI:RB	Expertise:NOV SI:RL	Expertise:NOV SI:TR
Erreurs	Erreurs	Erreurs	Erreurs
Min: 0.250000	Min: 3.750000	Min: 6.500000	Min: 7.500000
1st Qu.: 0.875000	1st Qu.: 5.000000	1st Qu.: 11.500000	1st Qu.: 9.187500
Mean: 1.687500	Mean: 7.343750	Mean: 14.750000	Mean: 13.062500
Median: 1.375000	Median: 7.000000	Median: 16.500000	Median: 11.625000
3rd Qu.: 2.062500	3rd Qu.: 8.937500	3rd Qu.: 17.875000	3rd Qu.: 17.375000
Max: 4.250000	Max: 11.750000	Max: 20.000000	Max: 19.500000
Total N: 8.000000	Total N: 8.000000	Total N: 8.000000	Total N: 8.000000
NA's : 0.000000	NA's : 0.000000	NA's : 0.000000	NA's : 0.000000
Std Dev.: 1.334635	Std Dev.: 3.082026	Std Dev.: 4.773438	Std Dev.: 4.810238

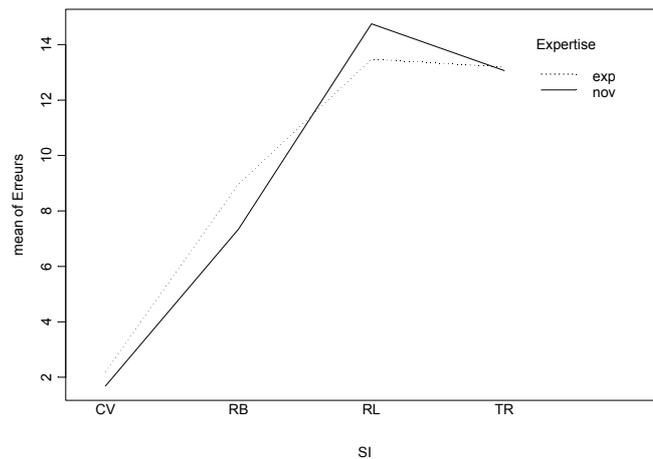


Figure 43: Graphique d'interaction entre le nombre moyen d'erreurs des novices et des experts avec les 4 SI

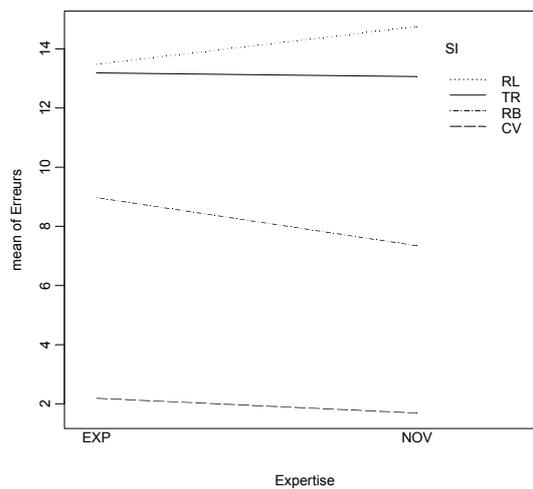


Figure 44: Effet de l'expertise sur le nombre moyen d'erreurs

D'après la Figure 44 nous constatons que les utilisateurs ont moyennement sur les 4 tâches le même nombre d'erreurs. Nous n'attendons donc pas d'observer d'effet principal de l'expertise. Par contre, dans la Figure 43, nous pouvons nous attendre à observer à un effet lié au SI utilisé.

Nous pouvons aussi observer (Tableau 33, Tableau 34) que de manière générale CV est le SI qui provoque le moins d'erreurs chez les sujets des deux groupes expérimentaux suivi par RB. Nous observons chez les experts que RL et TR ont une moyenne comparable alors que chez les novices le TR est légèrement meilleur par rapport à RL.

L'ANOVA nous montre le seul effet principal lié au SI sur le nombre d'erreurs moyen ( $F = 21.324$  ;  $p < .00$ ). Les comparaisons multiples sur le facteur SI nous affichent les différences suivantes ( $p = .05$ ):

- CV – RB
- CV – RL
- CV – TR
- RB – RL
- RB – TR

Bien que nous n'ayons pas observé d'effet d'interaction, nous allons quand même effectuer une analyse des effets simples, car, pour vérifier ou infirmer certaines de nos hypothèses, nous avons besoin de connaître ces effets ; d'ailleurs, Howell (1998) suggère que « si l'expérimentateur a une raison a priori d'examiner par exemple l'interaction AB à un niveau particulier de C, il est tout à fait libre de le faire même en l'absence d'une interaction de deuxième ordre significative » (p. 499). L'analyse des effets simples découlant de l'ANOVA sur les erreurs moyens nous montre les différences de performances des SI suivantes ( $p = .05$ ) :

Pour les experts :

- CV – RL
- CV – TR

Pour les novices :

- CV – RL
- CV – TR
- RB – RL

#### 4.2.7 Tâches libres

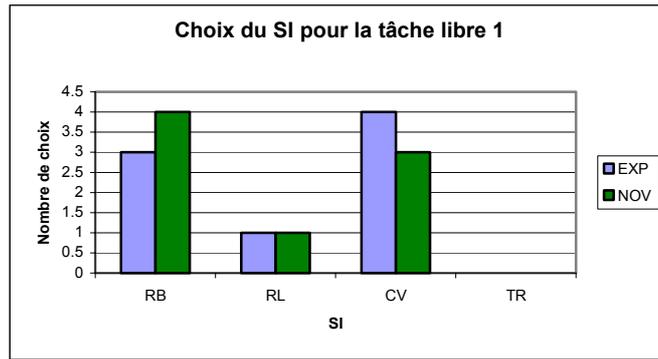
Dans les tâches libres, nous demandons aux sujets des deux groupes d'effectuer des tâches (insérer un contact et un rendez-vous<sup>10</sup>) à l'aide du SI de leur préférence. Comme dans le cas de la passation des tâches fermées nous retenons le temps et le nombre d'erreurs effectués de même que le SI utilisé.

Nous allons d'abord montrer le choix de nos sujets du SI utilisé pour l'exécution des deux tâches. Ensuite nous allons analyser et comparer le temps d'exécution et le nombre moyen d'erreurs avec les données issues des tâches fermées.

Tableau 35: Choix du SI pour la tâche libre 1 (Adresse)

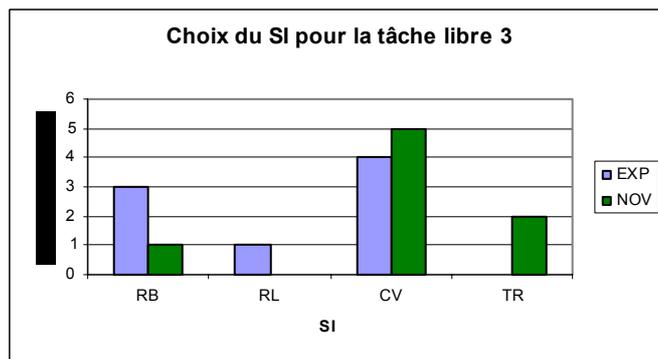
---

<sup>10</sup> Ces deux tâches correspondent respectivement aux tâches fermées 1 et 3



Nous constatons que 4 expérimentés choisissent CV comme SI pour effectuer la première tâche libre (rendez-vous), 3 sujets choisissent la RB et 1 seul la RL. Chez les novices, 4 d'entre eux choisissent RB, 3 CV et 1 RL.

Tableau 36: Choix du SI pour la tâche libre 2 (rendez-vous)



Encore plus fort ici le choix du CV de la part des sujets des deux groupes.

Il est maintenant intéressant de comparer les performances, au niveau du temps et des erreurs, des deux tâches libres avec les mêmes tâches en situation « fermée ».

#### 4.2.8 Tâche libre 1 : Adresse

##### Temps d'exécution :

Nous allons d'abord nous concentrer sur la première tâche libre (TL) qui implique l'introduction d'une adresse civique. Comme nous avons déjà reporté la statistique descriptive pour cette même tâche dans les 4 séries de tâches fermées (page 50) pour les deux groupes, nous nous limitons à présenter ici les données de la TL.

Tableau 37: Analyse descriptive du temps d'exécution de TL1 avec les 4 SI par les sujets experts et novices

SI:TL Expertise:EXP Temps Min: 29.534000 1st Qu.: 30.750250 Mean: 36.093375 Median: 34.100000 3rd Qu.: 38.253500 Max: 49.666000 Total N: 8.000000 NA's : 0.000000 Std Dev.: 7.307103	SI:TL Expertise:NOV Temps Min: 47.200000 1st Qu.: 54.27475 Mean: 70.10425 Median: 61.86700 3rd Qu.: 84.16675 Max: 106.20000 Total N: 8.000000 NA's : 0.000000 Std Dev.: 21.43423
---	---

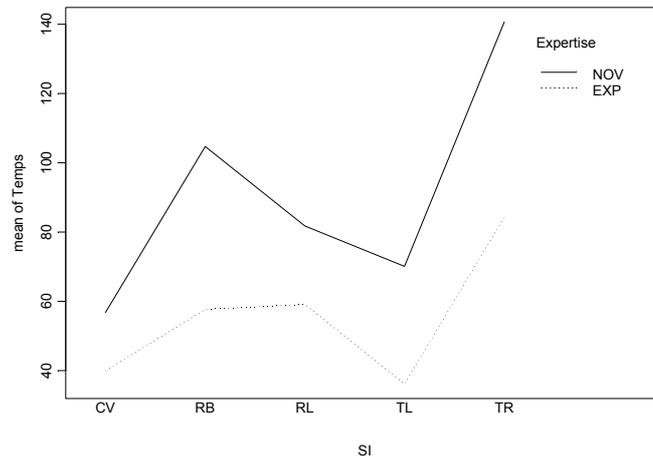


Figure 45: Effet des SI et de TL sur le temps d'exécution de TF1 et TL1

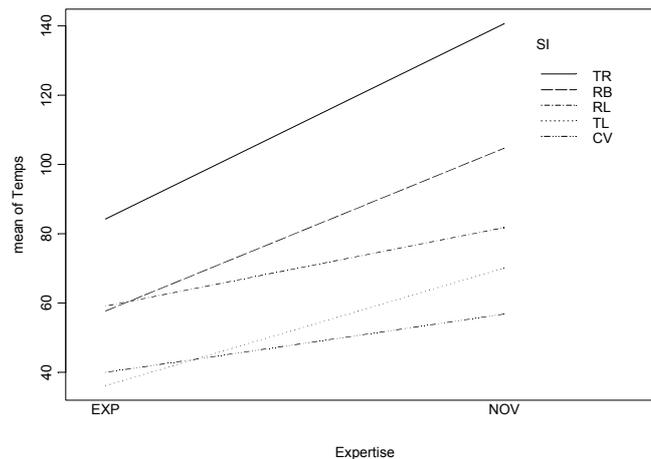


Figure 46: Effet du groupe (expertise) sur le temps d'exécution de TF1 et de TL1

Par rapport à la première tâche libre, nous pouvons constater que les experts ont un temps comparable à celui effectué dans la condition *fermée* avec le CV. Les novices par contre, ont une performance qui se situe entre celle obtenue avec le CV et avec le RL en condition *fermée*.

Comme dans l'analyse de la tâche fermée 1, nous pouvons nous attendre à observer un effet principal lié au groupe et un effet principal lié au SI.

L'ANOVA ne montre pas d'effet d'interaction. Par contre, un effet de l'expertise ( $F = 50.628 ; p < .00$ ) et un effet du SI ( $F = 21.239 ; p < .00$ ) sont mis en évidence. Si nous conditionnons l'expertise nous pouvons observer les différences entre SI suivantes :

- CV – RB
- CV – RL
- CV – TR
- RB – RL
- RB – TR
- RL – TR
- TL - TR

Par conséquent, les performances des deux groupes dans cette tâche libre sont comparables avec CV ou RL. Il est intéressant de constater que dans cette tâche libre, les experts ont encore une fois de meilleures performances que les novices. Néanmoins, les deux groupes ont su choisir de manière générale, des SI qui soient efficaces impliquant le moins de temps d'exécution.

En intra-groupe nous observons les différences suivantes ( $p = .05$ ):

Experts :

- CV – TR
- TL - TR

Novices :

- CV – RB
- CV – TR
- RB – TR
- RL – TR
- TL – TR

Erreurs :

Tableau 38: Analyse descriptive des erreurs de TL1 pour les sujets experts et novices

SI:TL Expertise:EXP	SI:TL Expertise:NOV
Erreur	Erreur
Min: 0.000000	Min: 1.000000
1st Qu.: 0.750000	1st Qu.: 3.750000
Mean: 1.750000	Mean: 6.000000
Median: 1.000000	Median: 4.500000
3rd Qu.: 2.500000	3rd Qu.: 8.750000
Max: 5.000000	Max: 12.000000
Total N: 8.000000	Total N: 8.000000
NA's : 0.000000	NA's : 0.000000
Std Dev.: 1.832251	Std Dev.: 3.927922

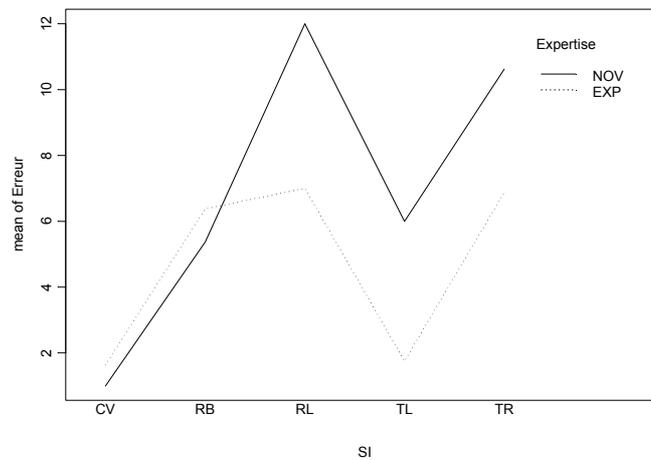


Figure 47: Effet des SI et de TL sur les erreurs de TF1 et TL1

Pour les sujets experts, le nombre d'erreurs se rapproche des performances observées avec le CV ; pour les novices, par contre, il est comparable avec RB en situation des tâches fermées. Nous pensons observer un effet des SI et de l'expertise de même qu'un effet d'interaction.

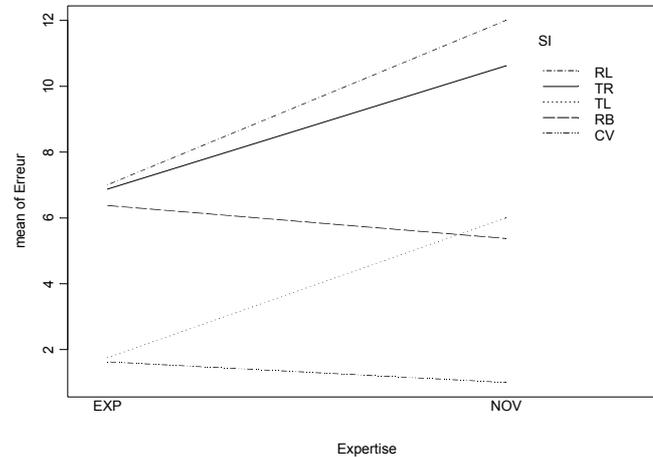


Figure 48: Effet du groupe (expertise) sur les erreurs de TF1 et de TL1

Nos attentes sont confirmées par l'ANOVA. Nous observons un effet du SI ( $F = 14.382$  ;  $p < .00$ ), un effet de l'expertise ( $F = 8.052$  ;  $p < .0059$ ) et un effet d'interaction ( $F = 2.539$  ;  $p < .047$ ). Comme nous avons un effet d'interaction, nous pouvons analyser de plus près les effets simples par rapport aux différences des SI à l'intérieur de chaque groupe. RL ( $F = 7.78$  ;  $p < .0068$ ), TR ( $F = 4.38$  ;  $p < .04$ ) et les données issus de la tâche libre ( $F = 5.62$  ;  $p < .021$ ) sont significativement différents entre les deux groupes.

Si nous conditionnons la variable expertise, nous obtenons des différences entre SI ( $p = .05$ ) :

- CV – RB
- CV – RL
- CV – TR
- RB – RL
- RL – TL
- TL – TR

Comme il y a un effet d'interaction, nous observons seulement les différences intra-groupe pour les novices suivantes ( $p = .05$ ) :

- CV – RL

- CV – TR
- RB – RL
- RL - TL

#### 4.2.9 Tâche libre 2 : Rendez-vous

##### Temps d'exécution :

Figure 49: Analyse descriptive du temps d'exécution de TF1 avec les 4 SI par les sujets novices et experts

SI:RL Expertise:EXP Temps Min: 37.80000 1st Qu.: 47.55000 Mean: 59.10850 Median: 51.60000 3rd Qu.: 64.69025 Max: 94.53300 Total N: 8.00000 NA's : 0.00000 Std Dev.: 20.02866	SI:TL Expertise:NOV Temps Min: 47.20000 1st Qu.: 54.27475 Mean: 70.10425 Median: 61.86700 3rd Qu.: 84.16675 Max: 106.20000 Total N: 8.00000 NA's : 0.00000 Std Dev.: 21.43423
---	--

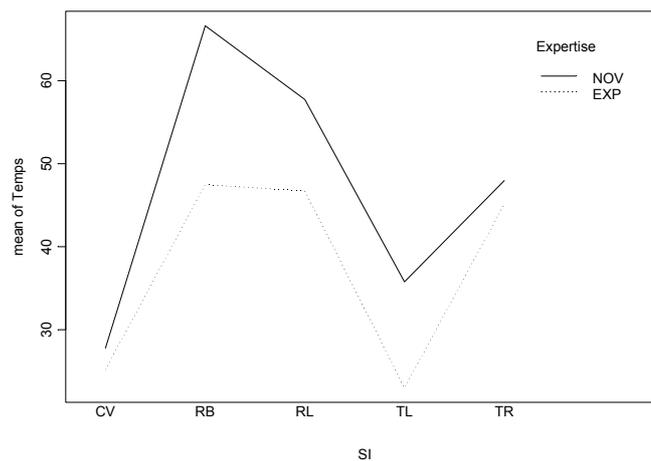


Figure 50: Effet des SI et de TL sur le temps d'exécution de la tâche rendez-vous

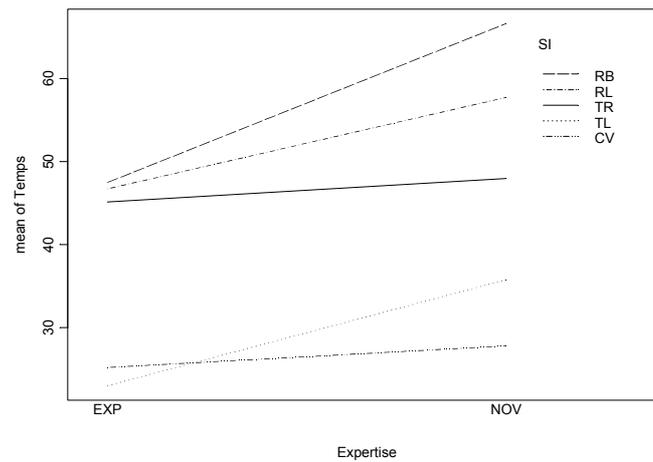


Figure 51: Effet du groupe (expertise) sur le temps d'exécution de la tâche rendez-vous

L'ANOVA nous indique un effet principal du SI ( $F = 10.192$  ;  $p < .00$ ) et un effet de l'expertise ( $F = 3.365$  ;  $p < .0139$ ). Si nous comparons les SI indépendamment de l'expertise nous obtenons les différences suivantes ( $p = .05$ ):

- CV – RB
- CV – RL
- CV – TR
- RB – TL
- RL – TL
- TL – TR

Dans cette deuxième tâche libre aussi les experts ont de meilleures performances par rapport aux novices. De plus, les temps d'exécution de la tâche libre, indépendamment du groupe d'appartenance, est significativement meilleure par rapport à la même tâche fermée effectuée avec RB, RL et TR. Par conséquent, nos sujets ont à nouveau choisi le meilleur SI, dont les performances sont comparables à celles observées en situation fermée avec CV.

Si nous regardons les différences intra-groupe nous n'observons des différences que pour les novices ( $p = .05$ ) :

- CV – RB
- CV – RL
- RB – TL

Erreurs :

Tableau 39: Analyse descriptive des erreurs de TL2 pour les sujets experts et novices

<pre> SI:TL Expertise:EXP       Erreur       Min: 0.00000 1st Qu.: 0.00000       Mean: 1.37500       Median: 0.50000 3rd Qu.: 1.50000       Max: 6.00000 Total N: 8.00000   NA's : 0.00000 Std Dev.: 2.13391         </pre>	<pre> SI:TL Expertise:NOV       Erreur       Min: 0.0000000 1st Qu.: 0.7500000       Mean: 1.2500000       Median: 1.5000000 3rd Qu.: 2.0000000       Max: 2.0000000 Total N: 8.0000000   NA's : 0.0000000 Std Dev.: 0.8864053         </pre>
---	---

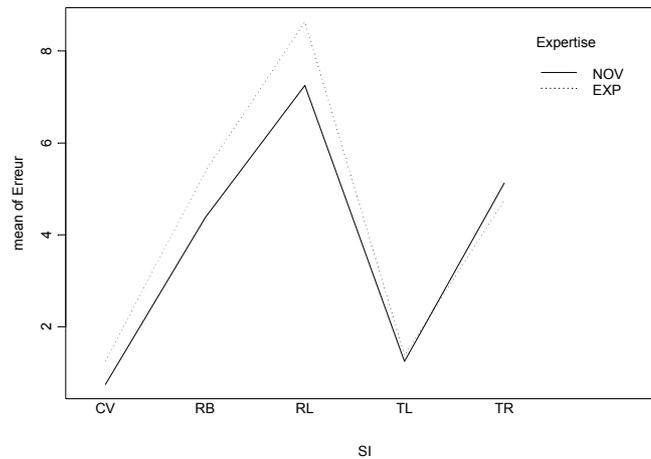


Figure 52: Effet des SI sur les erreurs de la tâche *rendez-vous*

Nous observons que lors de cette tâche, les deux groupes ont pratiquement la même moyenne. Le trend représentant aussi les 4 SI, sont semblables. Nous nous attendons à observer seulement un effet du SI sur le nombre d'erreurs effectués lors de l'exécution de la tâches *rendez-vous*.

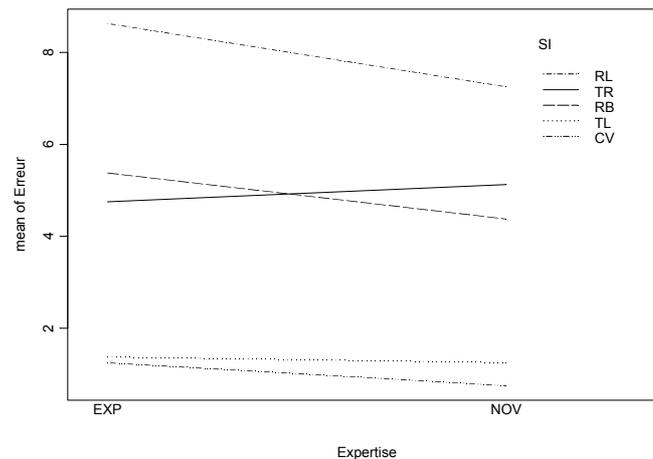


Figure 53: Effet de l'expertise sur les erreurs de la tâche *rendez-vous*

L'ANOVA montre comme prédit le seul effet principal ( $F = 8.27$  ;  $p < .00$ ) des SI sur le nombre d'erreurs commis lors de l'exécution de la tâche *rendez-vous*. Si nous conditionnons la variable expertise nous obtenons les différences suivantes ( $p = .05$ ) :

- CV – RL
- RL – TL

Nous avons observé les différences intra-groupe suivantes ( $p = .05$ ):

Experts :

- CV – RL
- RL - TL

Novices :

- CV – RL

### 4.3 Réponses au questionnaire post-série

Le questionnaire que nos sujets ont dû remplir à la suite de chaque série de tâches comportait une série de réponses concernant leur évaluation du médium utilisé, des tâches effectuées et de leur préférence pour un SI différent à la place de celui qu'ils venaient d'utiliser. Lors de l'analyse des 4 tâches fermées, nous avons déjà pris en compte certaines données issues de ces questionnaires ; en effet, l'analyse de l'évaluation subjective était demandé à ce moment de l'expérimentation.

4.3.1 Satisfaction du SI utilisé

Tableau 40: Satisfaction subjective des tâches fermées pour les experts

Expertise:EXP SI:CV	Expertise:EXP SI:RB	Expertise:EXP SI:RL	Expertise:EXP SI:TR
Satisfaction	Satisfaction	Satisfaction	Satisfaction
Min: 8.4000000	Min: 4.5000	Min: 4.000000	Min: 0.450000
1st Qu.: 8.9750000	1st Qu.: 6.6750	1st Qu.: 6.750000	1st Qu.: 3.175000
Mean: 9.2687500	Mean: 7.5500	Mean: 7.400000	Mean: 4.706250
Median: 9.3000000	Median: 8.0000	Median: 7.450000	Median: 5.000000
3rd Qu.: 9.7125000	3rd Qu.: 8.8250	3rd Qu.: 8.625000	3rd Qu.: 6.750000
Max: 9.8000000	Max: 9.3000	Max: 9.400000	Max: 7.700000
Total N: 8.0000000	Total N: 8.0000	Total N: 8.000000	Total N: 8.000000
NA's : 0.0000000	NA's : 0.0000	NA's : 0.000000	NA's : 0.000000
Std Dev.: 0.4992405	Std Dev.: 1.7337	Std Dev.: 1.704616	Std Dev.: 2.615126

Tableau 41: Satisfaction subjective des tâches fermées pour les novices

Expertise:NOV SI:CV	Expertise:NOV SI:RB	Expertise:NOV SI:RL	Expertise:NOV SI:TR
Satisfaction	Satisfaction	Satisfaction	Satisfaction
Min: 2.000000	Min: 3.10000	Min: 1.000000	Min: 0.500000
1st Qu.: 8.412500	1st Qu.: 5.35000	1st Qu.: 6.475000	1st Qu.: 2.875000
Mean: 7.950000	Mean: 6.21250	Mean: 6.550000	Mean: 3.406250
Median: 8.625000	Median: 6.90000	Median: 7.000000	Median: 3.350000
3rd Qu.: 8.975000	3rd Qu.: 7.20000	3rd Qu.: 8.000000	3rd Qu.: 3.750000
Max: 9.500000	Max: 8.00000	Max: 8.500000	Max: 6.550000
Total N: 8.000000	Total N: 8.00000	Total N: 8.000000	Total N: 8.000000
NA's : 0.000000	NA's : 0.00000	NA's : 0.000000	NA's : 0.000000
Std Dev.: 2.455751	Std Dev.: 1.65912	Std Dev.: 2.382676	Std Dev.: 1.714317

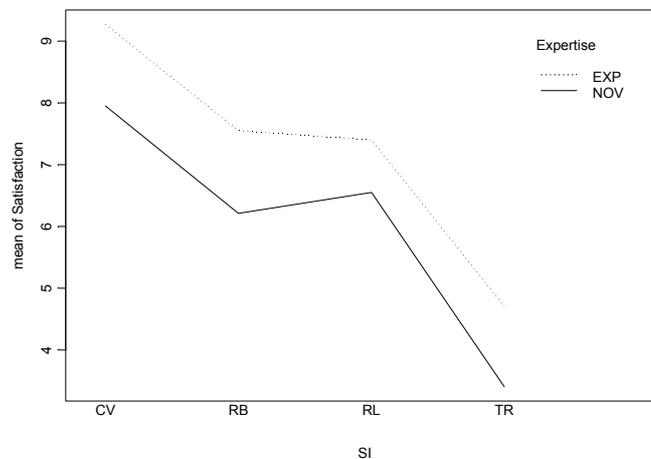


Figure 54: Effet des SI sur la satisfaction subjective

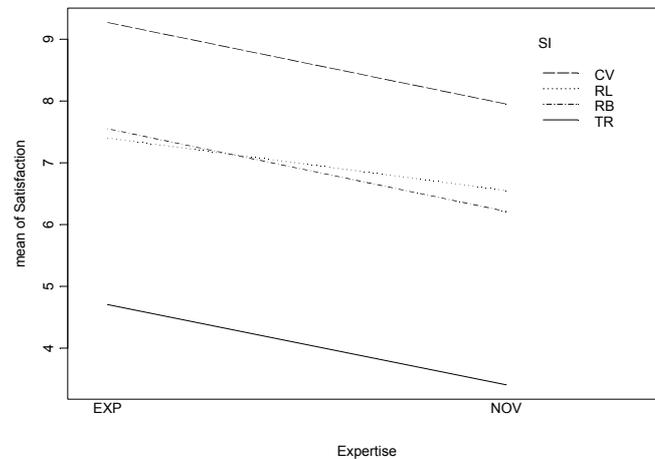


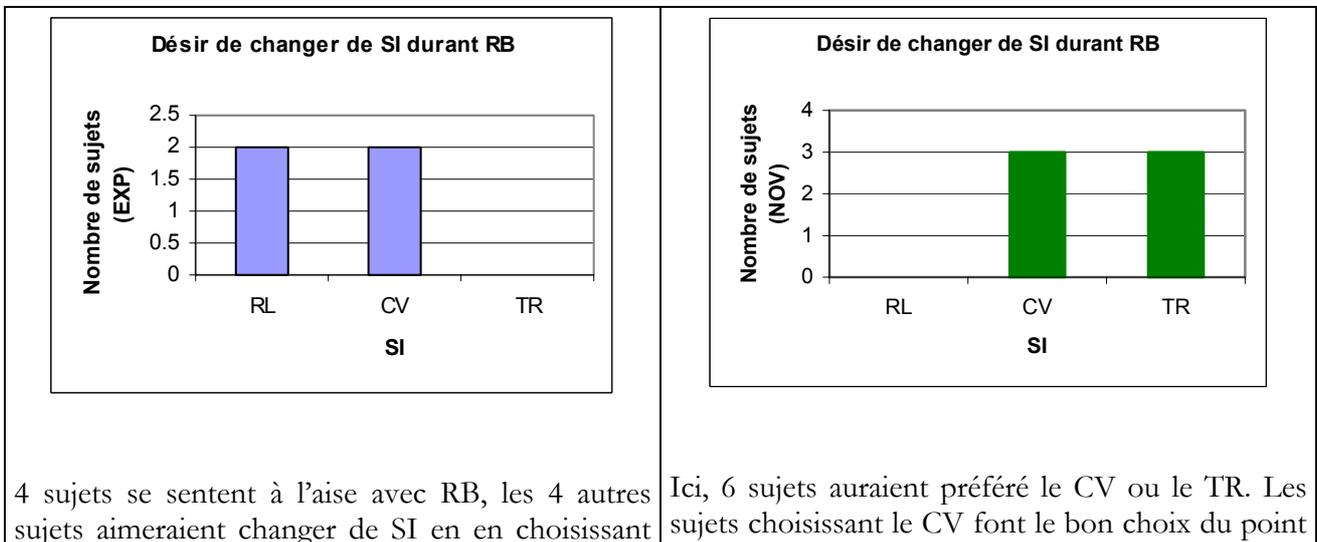
Figure 55: Effet de l'expertise sur la satisfaction subjective

L'ANOVA montre un effet de l'expertise ( $F = 6.079 ; p < .017$ ), un effet de SI ( $F = 15.052 ; p = .00$ ). La comparaison des SI en conditionnant la variable expertise montre des différences significatives ( $p = .05$ ) entre CV et TR, RB et TR, et RL et TR.

#### 4.3.2 Préférence d'un autre SI

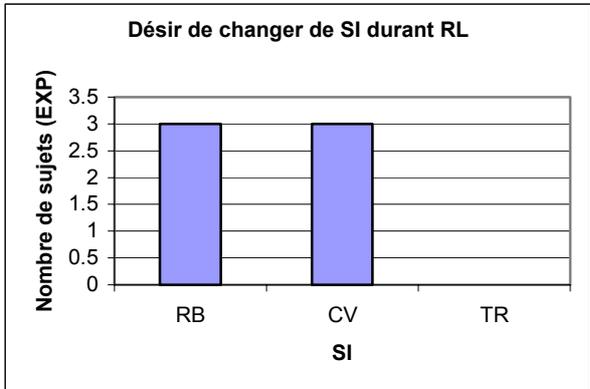
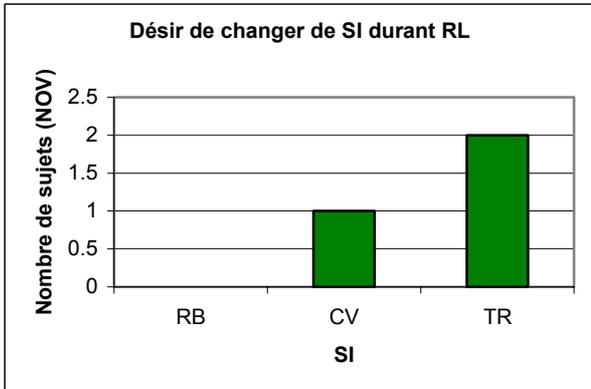
Nos sujets ont dû répondre à la question « Par rapport aux tâches que vous avez effectuées, auriez vous préféré l'usage d'un autre médium ? Si oui lequel ». Dans les tableaux ci-dessous nous avons représenté les réponses pour chaque SI.

Dans la RB :

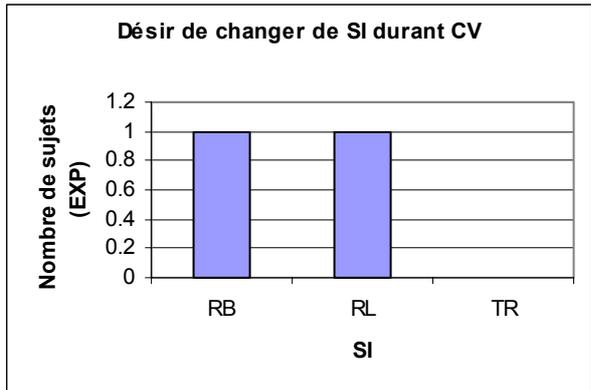
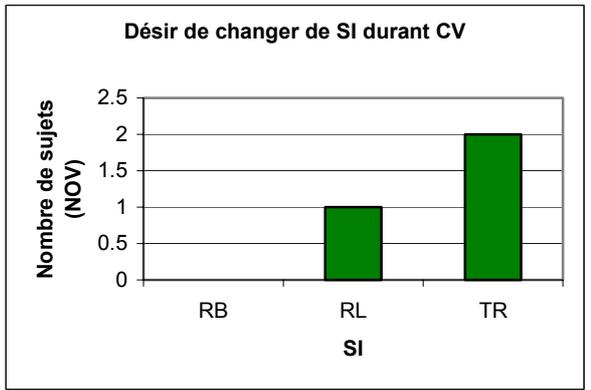


<p>un de plus bas niveau.</p> <p>Deux sujets préfèrent passer au CV ce qui a statistiquement du sens (moins d'erreurs mais même performances au niveau du temps d'exécution). Les deux autres font le mauvais choix (RL), toujours du point de vue statistique.</p> <p>Les 4 sujets qui ne veulent pas changer de SI ne sont pas donc influencés par les erreurs.</p>	<p>de vue du nombre moyens d'erreurs et du temps d'exécution qui sont inférieurs par rapport à RB.</p> <p>Par contre, les 3 sujets qui ont choisi le TR comme substitut, se trompent en choisissant celui qui implique le plus d'erreurs et de temps d'exécution des tâches.</p>
---	--

Dans RL :

 <p>3 sujets choisissent un SI de plus bas niveau (CV), les 3 autres choisissent la RB. On peut induire que les premiers choisissent le médium de plus bas niveau impliquant le moins d'erreurs et de temps d'exécution, les deuxièmes se concentrant sur le temps d'exécution.</p>	 <p>Ici, seulement trois sujets ne sont pas à l'aise avec la RL. Deux sujets font confiance au TR qui demeure être le moins fiable des 4 SI au niveau du temps d'exécution des tâches et du nombre d'erreurs.</p> <p>Personne ne choisit un SI de plus haut niveau.</p>
---	--

Dans CV :

	
---	--

<p>6 sujets n'ont pas l'intention de changer de SI. Les deux autres choisissent un SI de plus haut niveau.</p>	<p>Les NOV aussi sont à l'aise avec le CV. Nous constatons que 2 sujets aimeraient tout de même effectuer ces tâches avec le TR (bas niveau) et une personne avec la RL.</p>
--	--

Dans TR :

<div data-bbox="167 560 758 952" data-label="Figure"> <p>Désir de changer de SI durant TR</p> <table border="1"> <thead> <tr> <th>SI</th> <th>Nombre de sujets (EXP)</th> </tr> </thead> <tbody> <tr> <td>RB</td> <td>3</td> </tr> <tr> <td>RL</td> <td>2</td> </tr> <tr> <td>CV</td> <td>3</td> </tr> </tbody> </table> </div> <p>Tous les sujets décident de changer de SI (s'ils auraient la possibilité). Il est intéressant d'observer que 6 sujets se partagent dans RB et CV. Deux choisissent la RL.</p> <p>Ici il est impossible de choisir une SI de plus bas niveau.</p>	SI	Nombre de sujets (EXP)	RB	3	RL	2	CV	3	<div data-bbox="821 560 1412 952" data-label="Figure"> <p>Désir de changer de SI durant TR</p> <table border="1"> <thead> <tr> <th>SI</th> <th>Nombre de sujets (NOV)</th> </tr> </thead> <tbody> <tr> <td>RB</td> <td>2</td> </tr> <tr> <td>RL</td> <td>2</td> </tr> <tr> <td>CV</td> <td>4</td> </tr> </tbody> </table> </div> <p>Ici aussi, tous les sujets n'ont pas apprécié le TR. La moitié choisit comme substitut le CV et les autres se partagent dans RB et RL.</p>	SI	Nombre de sujets (NOV)	RB	2	RL	2	CV	4
SI	Nombre de sujets (EXP)																
RB	3																
RL	2																
CV	3																
SI	Nombre de sujets (NOV)																
RB	2																
RL	2																
CV	4																

## 4.4 Résultats partie qualitative

### 4.4.1 Evaluation des performances

Nous estimons intéressant de comparer ce que nos sujets ont dit par rapport à leur performances avec leurs performances effectives du point de vue du temps d'exécution moyen des 4 tâches. Cet indice nous permet de voir si les sujets sont capables d'évaluer leurs performances avec les 4 SI.

Dans un deuxième temps nous allons restituer l'avis de nos sujets sur les 4 SI, notamment du point de vue des avantages et désavantages de chaque modalité.

### Experts :

- Par rapport au temps d'exécution des tâches :

Evaluation parfaite de leurs performances : **Sujets 1, 3, 11,4**

Evaluation différente :

Le **sujet 5** estime que RB et RL sont comparables alors que moyennement RL est plus rapide de 20 secondes ( $m_{RL} = 92.450$  sec;  $m_{RB} = 111.10$  sec). Le **sujet 16** a bien interprété ses deux meilleurs SI (RB et CV) ; lors de l'entretien il indique que TR ( $m_{TR} = 186.633$  sec) est le meilleur par rapport à RL ( $m_{RL} = 131.017$  sec) alors que la réalité voit cette estimation inversée. Le **sujet 2** pense que RL et TR sont comparables et il les classe comme étant le 3<sup>ème</sup> SI. En réalité TR ( $m_{TR} = 160.76$  sec) est moyennement plus lent d'environ d'une minute par rapport à TL ( $m_{RL} = 100.95$  sec). Le **sujet 25** invertit lui aussi TR ( $m_{TR} = 88.48$  sec) avec RL ( $m_{RL} = 81.70$  sec) bien que le premier diffère du deuxième de seulement 6.7 secondes environ.

### Novices :

- Par rapport au temps d'exécution des tâches :

Evaluation parfaite de leurs performances : **sujets 13, 22,**

Evaluation différente : le **sujet 12** n'a pas su évaluer correctement ses performances, en effet, d'après son estimation le meilleur SI est RB suivi par CV, RL et TR. Ses performances réelles montrent par contre que CV est le meilleur ( $m_{CV} = 74.167$  sec) suivi par RL ( $m_{RL} = 124.134$  sec), RB ( $m_{RB} = 139.767$  sec) et TR ( $m_{TR} = 224.053$  sec). Le **sujet 14** a aussi mal perçu ses performances dans l'ordre RL, RB, CV et TR contre CV ( $m_{CV} = 74.231$  sec), RL ( $m_{RL} = 90.297$  sec), RB ( $m_{RB} = 120.600$  sec) et TR ( $m_{TR} = 129.100$  sec). Même chose pour le **sujet 21** qui a défini dans l'ordre RB, CV, RL et TR à la place de CV ( $m_{CV} = 67.217$  sec), RL ( $m_{RL} = 116.200$  sec), TR ( $m_{TR} = 127.1$  sec) et RB ( $m_{RB} = 144.984$  sec). Le **sujet 15** pense que TR ( $m_{TR} = 133.68$  sec) est le 2<sup>ème</sup> meilleur SI alors qu'il précède en 3<sup>ème</sup> place RL ( $m_{RL} = 98.36$  sec). Le **sujet 17** a pratiquement bien jugé ses performances bien qu'il a inversé RB ( $m_{RB} = 138.97$  sec) et RL ( $m_{RL} = 133.48$  sec), qui sont comparables et ne diffèrent que d'environ 5 secondes. Même chose pour le **sujet 18**, dont la différence entre RB ( $m_{RB} = 113.21$  sec) et RL ( $m_{RL} = 109.3$  sec) est de seulement 4 secondes.

De manière générale, nous constatons que les sujets ayant une expertise préalable du PPC ont tendanciellement des estimations liées à leur performances meilleures par rapport au sujets novices dans ce domaine.

#### 4.4.2 Analyse des SI

##### Reconnaissance des blocs (RB) :

D'après les commentaires de nos sujets novices (sujets 12, 13), c'est peut être le SI le plus attrayant et fascinant (pas son alphabet spécial et simplifié), car c'est une manière simple pour écrire, il est rapide, il y a beaucoup d'espace à disposition, la ponctuation n'est pas si difficile à mémoriser. Le seul problème handicapant les performances de ces utilisateurs étant la mémorisation initiale qui empêche ces sujets de profiter un maximum de toutes les caractéristiques offertes par ce SI (sujets 13, 12, 14, 17, 18). De plus, nos sujets estiment qu'il est impératif d'avoir un usage constant de ce SI afin de ne pas oublier des symboles et des caractères (sujet 14). « Sur le plan cognitif c'est dur d'apprendre un nouveau langage, il faut continuellement se référer au nouveau alphabet », de plus les caractères spéciaux sont difficiles à les mémoriser (sujet 22). Un usage constant et répété permettra selon eux d'aller très vite avec le moindre d'erreurs et à long terme il pourrait même être le meilleur SI (sujets 15, 17).

Cette modalité fascine toujours aussi les sujets plus experts (sujet 1, 16), car même s'ils ne le pratiquent pas couramment, « ça revient assez vite » (sujet 1) par sa facilité et simplicité d'écriture et rétention.

##### Clavier Virtuel (CV) :

Les experts aiment cette modalité et à différence des novices ils ne sous-estiment pas son utilité et efficacité. En effet, les experts tendent à préférer et à l'utiliser dans le quotidien car ils observent que du point de vue de la rapidité et de la précision le CV n'a pas de concurrents (sujets 1, 3, 5). Mais tout de même dans le groupe des experts il y a des sujets qui rejoignent le point de vue des novices en soulevant le fait que « je n'aime pas viser chaque lettre » (sujet 16 ou chez les novices sujet 14).

D'autres points ont été mis en évidence par les novices, notamment le fait qu'il est précis pour un apprentissage à court terme mais pénalisant pour de longs textes (sujet 13) ; d'autres sont gênés par le fait que pour accéder à des symboles ou caractères spéciaux il faut basculer sur un autre clavier (sujet 22). Pour ce groupe, le fait qu'il soit rapide et précis n'est pas suffisant car pointer sur un clavier sur un petit écran est ridicule et « la jonglette » entre les 2 claviers est pénalisant.

Notons que les deux groupes sont un peu dérangés par le fait que certains éléments de ponctuation tels que le point surtout, n'est pas présent sur l'écran par défaut du CV ou bien que le clavier AZERTY est ridicule et implique une perte de temps dues à une recherche supplémentaire à faire, recherche moindre pour les experts qui ne sont pas trop gênés. Un sujet (15) aimerait que sur le clavier principal soit présente une ligne avec les chiffres.

#### Reconnaissance des lettres (RL) :

De manière générale, les experts ne prennent pas en considération cette modalité qui même s'il a une bonne reconnaissance, il faut toujours adapter la propre écriture parce que certains caractères sont mal interprétés par le système (sujet 1). D'autres ont de la peine avec notamment les « i, t, f » qui nécessitent deux traits pour indiquer au système ce que l'on veut écrire (pour le « i » il faut d'abord écrire un trait vertical et ensuite rajouter un petit point – comme dans l'écriture courante). Or pour certains, qui sont habitués ou ont été habitués avec RB ou Graffiti sur Palm, ce n'est pas naturel (sujet 5). De plus, certains indiquent néanmoins que s'il n'y aurait pas la RB probablement ils l'utiliseraient bien qu'il soit plus contraignant et laborieux que RB (sujet 16). Le sujet 4 a trouvé qu'il est plus « mécanique » par rapport à RB et il trouve que les deux sont des doublons et qu'il vaut mieux connaître un ou deux SI plutôt que les quatre.

Pour les novices il est assez comparable avec RB, bien qu'il soit moins inventif et plus banal que ce dernier (sujets 12, 17, 18). Le sujet 22 nous dit que « il est plus facile que RB. Il y a une simplification de la zone sensible et le langage libre est pas mal mais à nouveau il faut apprendre des trucs ... mais à terme c'est peut être plus rapide que RB » de plus « il est plus rapide, structuré et paramétré. Il se rapproche plus de l'écriture qu'on a ... moins besoin de s'adapter nous que la RB où c'est nous qu'il faut s'adapter à la machine ... pratique pour une utilisation à court terme » (sujet 14, le sujet 15 a des propos comparables).

#### Transcriber (TR):

Les experts n'aiment pas du tout cette modalité qui n'est pas vraiment utilisable et adaptée pour un petit écran (sujet 1, 5), pour l'épaisseur du PPC, la direction de l'écran en vertical (sujet 5). Il est difficile pour corriger les fautes (sujet 2). Le sujet 4 l'a trouvé intéressant par le fait que son écriture était facilement interprétable par le système. Ou encore le sujet 25 a trouvé que la sensibilité du PPC utilisé lors de l'expérimentation n'avait pas la même sensibilité de celui qu'il utilise couramment.

Les novices par contre sont plus ambigus dans leur jugement : il est plus compliqué, ponctuation et retour au chariot difficile, incompréhensions entre les majuscules et les minuscules, peu intuitif, bonne

reconnaissance mais il y a des problèmes pour modifier ou pour la ponctuation (sujet 13) ; trop de fautes, il faudrait des fonctionnalités en plus, notamment de ponctuation, présentes à l'écran mais l'avantage d'écrire un mot en entier ou une phrase est intéressant (sujet 22) ; catastrophique pour les erreurs (sujets 21, 17, 15). « Sympathique au début, parfois pénible pour les erreurs ... il faudrait plus de fonctionnalités ou plus d'options comme le point par exemple qui soit visible tout le temps » (sujet 15).

Il est difficile de tirer des conclusions sur les commentaires provenant de 16 sujets partagés en deux groupes séparés. Néanmoins, ces quelques indications pourraient être utiles pour des recherches ultérieures dans ce même domaine ou pourraient être des renseignements utiles pour les concepteurs de SI pour PPC ou pour d'autres systèmes dotés d'un stylet comme médium d'interaction privilégié.

## 5. Discussion des résultats

Dans cette partie, nous allons vérifier si nos hypothèses ont été vérifiées ou infirmées par nos résultats. Nous allons reprendre chaque hypothèse dans l'ordre selon lequel nous les avons exposées dans la partie de la méthode expérimentale (3.3.2 Hypothèses opérationnelles, page 35), nous allons rappeler le contenu de chaque hypothèse et la mettre en relation avec nos résultats.

### 5.1 H1 : rôle de l'expertise dans le choix des SI

Dans cette hypothèse nous postulons que plus le niveau d'expertise est haut et plus nous nous attendons à observer à ce que les utilisateurs choisissent des SI de haut niveau et inversement. Nous considérons que RB comme étant le SI de plus haut niveau, suivi par RL, CV et TR.

Cette hypothèse peut être vérifiée au niveau qualitatif à plusieurs niveaux de notre expérimentation :

1. Questionnaire de recrutement des sujets experts : en particulier dans la question 5.1 où nous demandons de nous indiquer les SI qu'ils utilisent couramment.
2. Un autre indicateur nous permettant de vérifier notre hypothèse se situe au niveau des SI utilisés lors de l'exécution des deux tâches libres pour les deux groupes.

Nous constatons que cette hypothèse n'est pas vérifiée. En effet, nous constatons que d'après le questionnaire de recrutement, seulement trois experts utilisent RB de manière prioritaire, alors que les 5 autres utilisent CV (cf. 4.1.5. Les styles d'interaction, page 45). Cette tendance est aussi observée lors de l'exécution des tâches libres : dans les deux tâches libres, les experts utilisent toujours le même SI et nous observons que 4 sujets choisissent CV, 3 RB et 1 seul RL. Chez les novices nous observons que pour la première tâche, 4 sujets choisissent RB, 3 sujets préfèrent CV et un seul RL ; dans la tâche libre 2, par contre, nous constatons qu'un seul sujet choisit RB, 5 sujets choisissent CV et 2 TR (cf. Tableau 35 et Tableau 36).

Malgré que cette hypothèse ne soit pas vérifiée, nous pouvons constater que finalement les experts optent par un SI qui soit à la fois facile d'usage, précis et efficace. Nos experts, sont donc partagés principalement dans le choix du CV et de RB. Le premier a l'avantage d'offrir à l'écran toute l'information et de ne devoir donc pas utiliser trop de processus mentaux pour l'utiliser efficacement (les novices aussi ont de meilleures performances avec ce SI) ; le deuxième résulte aussi très efficace à condition qu'il soit utilisé régulièrement. Le fait que 6 sujets avaient déjà un autre PDA (non PPC, cf.

4.1.2. Temps d'utilisation du PPC, page 44) pendant une durée moyenne de 32 mois, fait qu'ils étaient déjà expérimentés avec RB qui est présent chez les PDA de la marque Palm (nous ne pouvons qu'imaginer que la majorité de ces 6 sujets avaient un Palm comme PDA). Si, RB n'est pas utilisé régulièrement, trop d'efforts cognitifs sont engendrés et par conséquent, le recours à un SI de plus bas niveau serait nécessaire.

## **5.2 H2 : rôle des erreurs**

Cette hypothèse postule que plus les sujets sont entravés par les erreurs et plus ils auront tendance à changer de SI, en faveur d'une modalité de plus bas niveau. Le questionnaire rempli par nos sujets à la fin de chaque série de tâches, nous a permis de quantifier le nombre de personnes ayant voulu remplacer le SI testé avec un autre pour l'exécution des tâches soumises (4.3.2 Préférence d'un autre SI, page 91). L'ANOVA sur les erreurs moyens (page 78) nous montrent que pour les experts CV diffère significativement avec RL et TR ; pour les novices nous observons la même tendance que pour les experts en plus de la différence entre RB et RL. Si nous ne considérons pas la variable expertise, nous obtenons des différences entre CV et les trois autres SI, RB et RL, RB et TR.

Dans le cas de RB, le SI de plus haut niveau, seulement 4 sujets auraient voulu changer de SI (2 pour RL, 2 pour CV) ; chez les novices, trois sujets préfèrent le CV et trois autres TR. Dans les deux groupes, seulement la moitié d'entre eux auraient préféré un SI, en l'occurrence le CV, plus performant du point de vue des erreurs. Les autres sujets auraient choisi un SI qui aurait impliqué plus d'erreurs que RB.

Dans RL, nous avons observé chez les experts un plus grand nombre de personnes qui auraient la tendance de changer de SI (3 pour RB et 3 pour CV). Dans les deux cas, ils auraient choisi un SI plus performant, mais seulement le CV comme étant de plus bas niveau par rapport à RL. Les novices, par contre se sentent plus à l'aise avec ce SI. Une personne préfère le CV et deux autres TR. Les novices choisissent un SI de plus bas niveau, alors que chez les experts seulement la moitié aurait préféré le CV.

Avec le CV, les experts se sentent à l'aise, en effet seulement deux sujets auraient choisi un SI de plus haut niveau (RB ou RL) ; chez les novices, par contre, deux sujets auraient opté pour TR ou pour RL.

Enfin, TR pose le plus de problèmes à tous nos sujets. Notons que chez les experts, seulement trois sujets auraient choisi le CV (3 RB et 2 RL), chez les novices la moitié (2 RB et 2 RL).

Nous pouvons avancer que cette hypothèse est difficilement vérifiable, néanmoins, nous constatons que chez les novices ont plus la tendance à choisir un SI de plus bas niveau en cas d'erreurs.

Suite à ces résultats nous pouvons nous poser deux questions :

1. Est-ce que pour les experts, RB doit être considéré de plus bas niveau, compte tenu qu'il est automatisé chez certains de ces sujets, alors que RL ou TR posent des problèmes plus grands ? Si nous considérons cette remarque comme vraie, notre hypothèse serait vérifiée aussi pour les experts.
2. Est-ce que les erreurs sont un facteur prépondérant du choix pour un autre SI ou le temps d'exécution est plus important ? Est-ce que d'autres facteurs entrent en jeu ?

### **5.3 H3 : reconnaissance du système**

Nous postulons ici que pour des tâches fréquentes et de routine, telles que celles présentées dans les tâches fermées, la capacité de reconnaissance de l'écriture de la part du système est capitale. Plus le SI utilisé est efficace et performant, et plus le niveau de satisfaction subjective sera grand. De manière générale, indépendamment des deux groupes, nous avons prédit que pour la satisfaction subjective :  $CV > (RB=RL) > (TR)$ .

Du point de vue des analyses de performances en termes de temps (page 47) et en terme d'erreurs (page 78) commis nous avons obtenu (en conditionnant la variable expertise) :

- Temps :  $CV < (RB = RL) < TR$
- Erreurs :  $CV < RB < RL \leq TR$

Ces deux informations nous indiquent de manière générale quels sont les SI les plus performants. Les résultats de l'ANOVA sur la satisfaction générale du médium utilisé (page 90) nous indique que tous les SI sont préférés à TR :

- Satisfaction subjective  $(CV = RB = RL) > TR$

Cette hypothèse est partiellement vérifiée. Si d'un côté le CV est indiscutablement le meilleur, il ne l'est pas statistiquement du point de vue de la satisfaction subjective.

#### **5.4 H4, H5 & H6 : Tâches fermées vs. Tâches libres**

Les trois hypothèses considérées ici, sont toutes reliées par le terme *adaptation*. Chaque hypothèse se caractérise par le fait qu'elle apporte de nouveaux facteurs : H4 se limite à postuler que l'utilisateur adapte son action en utilisant le SI le mieux adapté en fonction de son niveau d'expertise ; H5, indique que plus l'utilisateur est conscient des différences dans les propriétés de chaque SI, plus il aura des capacités d'adapter sa conduite. Pour terminer, H6 postule que plus les connaissances acquises sont fréquemment et régulièrement mobilisées et plus le sujet sera capable de manipuler le système efficacement et de s'adapter en conséquence.

Indiscutablement, nos analyses montrent qu'à tous les niveaux (moyennes des temps et d'erreurs dans les tâches fermées), un effet de l'expertise a été mis en évidence. Pratiquement sur la totalité de nos analyses, les experts ont obtenu de meilleures performances que les novices. H6 est donc confirmée, il est maintenant intéressant d'observer si la même tendance se maintient aussi dans les tâches libres.

Mais nous devons maintenant nous concentrer sur les tâches libres, où nos sujets ont pu choisir le SI à utiliser. Si d'un côté les experts connaissent déjà tous les SI et leurs propriétés, les novices ont pu, au cours des tâches fermées, se familiariser avec chaque SI et développer leur opinion à leur sujet.

La comparaison des tâches libres avec les mêmes tâches fermées, nous permet de vérifier si les novices comme les experts ont su choisir le meilleur SI.

Pour la première tâche libre (adresse, page 81), nous constatons que les experts ont un temps d'exécution comparable à celui du CV ; les novices, par contre, ont un temps qui se situe entre le CV et RL. Dans la deuxième tâche libre (rendez-vous, page 86), les novices ont un temps d'exécution qui est légèrement supérieur à CV ; les experts ont un temps inférieur ou égal à celui du CV.

Nous pouvons donc affirmer que les trois hypothèses sont vérifiées :

- H4 : L'hypothèse selon laquelle les utilisateurs adaptent leur action en utilisant le meilleur SI en fonction de leur expertise est vérifiée, comme les statistiques le montrent. Au cours des deux

tâches libres, les deux groupes utilisent un SI qui produit des performances comparables à celle obtenues en situation des tâches fermées avec le CV, qui résulte être le meilleur SI. Quant au choix des SI soulignons que les experts ont la tendance à utiliser CV et RB comme SI préférentiel aussi lors des tâches libres ; ce choix est principalement guidé par leur expérience antérieure. Chez les novices nous observons que certains sujets changent de SI lors de l'exécution de la tâche libre 2. Ce changement (cf. page 80 pour voir les choix des SI dans les tâches libres) montre que les performances s'améliorent lors de la deuxième tâche libre.

- H5 : L'utilisateur est conscient des propriétés de chaque SI et adapte ainsi son action. Les experts ont déjà construit leur modèle d'interaction avec le PPC. Ils interagissent principalement avec un seul SI ; chez certains novices il est intéressant d'observer que pour la deuxième tâche libre ils changent de SI ; chez ces sujets il y a des raisons pertinentes, comme le fait d'utiliser TR dans cette situation, car « elle n'implique pas trop de texte et que pendant le test ça c'est bien passé ». Après quelques dizaines de minutes d'interaction, ces sujets commencent à entrevoir les avantages et désavantages de chaque SI pour enfin choisir celui qui estiment être le plus performant.
- H6 : plus on utilise fréquemment le PPC et plus on arrive à le manipuler efficacement. Au-delà du choix du SI, les experts ont toujours de meilleures performances que les novices.

Les trois prochaines hypothèses sont issues de notre analyse des SI (page 29).

### **5.5 H7 : les erreurs**

Nous considérons que les experts et les novices rencontrent des difficultés différentes avec chaque SI, par conséquent nous avons prédit des hypothèses différentes :

- Novices : CV < RL < RB < TR
- Experts : CV < RB < RL < TR

Dans les deux cas, le CV est le SI qui implique le moins d'erreurs pendant l'exécution de la tâche, le TR le pire SI.

D'après l'ANOVA sur la moyenne des erreurs sur les quatre tâches fermées (page 78), nous ne pouvons ou ne devrions pas séparer les deux groupes, car nous n'observons pas d'effet principal de l'expertise. Si nous ne considérons pas l'expertise, nous obtenons le même résultat postulé dans l'hypothèse exprimée pour les experts :

**CV < RB < RL < TR**

Nous avons donc été étonnés d'observer que nos deux groupes ont le même trend quant aux erreurs commises. Au cours des tests effectués, nous avons observé que la plupart des experts avait la tendance de commettre des erreurs dus principalement à la vitesse, alors que les experts commettaient des erreurs de *jeunesse*. Peut être le fait de tester nos experts comme étant des experts les mettaient dans une position d'évaluation de leur performances, pression qui les incitaient à aller vite et à commettre des fautes.

L'analyse de chaque tâche fermée montre que les experts ont une conduite d'erreurs plus cohérente que les novices. Si d'un côté leur interaction confirme que **CV < RB < RL < TR**, de l'autre côté les novices ont la tendance à commettre plus d'erreurs avec RL pour les tâches telles que la rédaction d'un email ou l'introduction d'un rendez-vous. Soulignons aussi que seulement lors de la tâche fermée 1 nous avons observé un effet dû à l'expertise.

Cette hypothèse est donc vérifiée qu'en partie. De plus, nous obtenons une différence significative entre RB et RL alors que dans l'étude de Sears et Arora (2002), ils ne constatent pas de différences entre les deux.

## **5.6 H8 : vitesse des performances**

Comme pour l'hypothèse précédente et la suivante aussi, nous avons séparé pour les deux groupes le classement des SI les plus performants en terme de temps d'exécution.

- Novices: **CV < RL < RB < TR**
- Experts : **(RB = RL) < CV < TR**

Dans cette hypothèse nous avons été complètement étonnées dans le comportement de nos experts (cf. page 47): d'un côté par le fait que la moitié d'entre eux utilise dans sa vie quotidienne le CV ; de l'autre côté qu'ils obtiennent les meilleures performances avec un SI de bas niveau.

Pour les experts, nos résultats nous montrent le classement suivant :

- $CV \leq RB \leq RL$
- $(CV, RB) < TR$

Pour les novices par contre nous avons obtenu des résultats semblables à ceux attendus :

- $RL \leq RB \leq TR$
- $CV < (RB, RL, TR)$

Notre hypothèse n'est vérifiée qu'en partie. De plus, nous n'avons pas obtenu des différences significatives entre RB et RL, néanmoins RL semble être plus facile à utiliser pour les novices comme le soulignent aussi Sears et Arora (2002) dans leur étude.

### **5.7 H9 : satisfaction subjective**

Encore une fois nous avons séparé les deux groupes en postulant que :

- Novices :  $(RL = CV) > TR > RB$
- Experts :  $(RB = RL) > CV > TR$

Comme pour l'hypothèse précédente, nous avons été étonnés par les résultats obtenus. De manière générale (cf. page 90), les deux groupes ont le même trend à savoir :  $CV \geq RB \geq RL > TR$ .

L'hypothèse est donc rejetée. De plus nous obtenons des résultats différents que Sears et Arora (2002) qui indiquent que RL (Jot) est plus satisfaisant que Graffiti.

Les résultats obtenus pour tester nos trois dernières hypothèses sont intéressants. Notons que pratiquement dans chaque situation, CV et TR sont respectivement le meilleur et le pire SI. Ce qui change est la position de RB ou RL.

## 6. Conclusion

Malgré le fait que beaucoup d'hypothèses ont été vérifiées seulement en partie, nous avons pu extraire des tendances très intéressantes notamment au niveau du comportement réel des experts qui utilisent de préférence CV et RB. Nous avons été surpris par le taux élevé d'experts utilisant un SI de bas niveau tel que le clavier virtuel. Pourtant, leur choix est tout à fait compréhensible car les résultats de la partie quantitative montrent que CV est indiscutablement le meilleur SI. Dans ce contexte, nous pouvons avancer que l'expérience de nos experts, leur ont permis de développer des connaissances au niveau syntaxique concernant chaque SI, de choisir celui le plus performant selon leur caractéristiques visant à rendre leur interaction la plus directe possible et en conséquence réduire la distance articulatoire et sémantique lors de la planification et la mise en place d'une action. De plus, d'après les propos de nos sujets, nous avons pu observer que la plupart de ces sujets n'utilisent qu'un seul SI et qu'il ne le changent que très rarement. La partie qualitative a aussi montré des avantages et désavantages pour chaque SI. Malgré que CV soit le meilleur SI, nos sujets ont remarqué des problèmes ergonomiques qui pourraient être améliorés dans le but de rendre cette modalité encore plus performante. De plus, notre étude a mis en évidence que, de manière générale, les experts ont de meilleures évaluations et de meilleures performances (temps et erreurs) que les novices.

Nous aimerions aussi souligner que notre étude montre que RB est généralement meilleur que RL. Or, il y a quelques mois, en raison d'une plainte pénale, PALM s'est vu contraint à abandonner Graffiti (RB) pour ensuite insérer dans leur système une autre modalité, Jot, qui correspond à la reconnaissance des lettres. Ce changement au sein des systèmes hébergés par PALM ne nous laisse pas indifférents et nous nous demandons s'ils n'auraient pas dû effectuer un autre choix, vers par exemple un clavier virtuel. Mais d'un côté, nous comprenons ce choix, vu la méfiance que les novices ont montré à l'égard de CV. En effet, malgré que CV soit le meilleur SI, ce groupe de sujets n'apprécie pas le fait de devoir pointer sur des touches plutôt que d'utiliser l'écriture humaine qui leur laisse un plus grand marge de manœuvre associé à un aspect ludique de l'instrument.

Nous devons mettre en évidence quelques critiques à l'égard de notre étude. Tout d'abord, le faible nombre de sujets présents dans les deux groupes doit nous faire prendre les résultats avec une certaine précaution. Un plus grand échantillon aurait pu montrer d'autres facteurs et effets. De plus, le fait d'avoir testé que des sujets de sexe masculin, peut biaiser les interprétations. Un effet lié au genre pourrait être observable. Notons aussi qu'une partie des sujets experts ne connaissaient pas forcément

tous les SI, en particulier TR ; nous ne les avons pas soumis à une phase de familiarisation comme pour les novices.

Une extension de ce mémoire pourrait être envisageable. Surtout en ce moment que de nouveaux PocketPC, SmartPhones, TabletPC voient le jour avec toujours de nouvelles fonctionnalités comme un téléphone portable de MyOrigo qui peut être contrôlé avec les mouvements effectués sur celui-ci (voir *Embodied User Interfaces*). Ou encore, il serait intéressant de se concentrer ultérieurement sur l'un des deux groupes étudiés : les experts pour vérifier avec un échantillon plus représentatif si les tendances que nous avons soulignées se maintiennent ou encore mieux analyser les différences entre les différents SI en testant d'autres modalités payantes qui ont vu le jour pendant la rédaction de ce mémoire (comme par exemple Calligrapher, MyScript, PenReader,, Fitaly, Spb Full Screen Keyboard,...). Etudier les novices serait aussi intéressant pour observer leur apprentissage, leur développement dans la compréhension des propriétés des SI testés et éventuellement leur changement dans la prise en compte du CV comme étant la meilleure solution à utiliser.

## 7. Bibliographie

- Beaudoin-Lafon, M.. *Instrumental Interaction: An Interaction Model for Designing Post-WIMP User Interfaces*. In Proceedings of CHI 2000, pages 446-453, The Hague, The Netherlands, April 2000. Association for Computing Machinery.
- Calvet, G., Kahn, J., Zouinar, M., Salembier, P., Briois, J.-C., L. Nigay, Rey, G., Pasqualetti, L. : *Etude empirique de l'usage de la multimodalité avec un ordinateur de poche*. Conférence IHM-HCI2001, Lille 10-14 sept., CEPADUES Publ., 2001, pp. 5-8.
- Chang, L., & MacKenzie, I. S. (1994). A comparison of two handwriting recognizers for pen-based computers. Proceedings of CASCON '94, pp. 364-371. Toronto: IBM Canada.
- Frankish, C., Hull, R. & Morgan, P. (1995), Recognition Accuracy and User Acceptance of Pen Interfaces, in *Proceedings of ACM CHI'95*, pp.503-510.
- Frohlich, D.M. (1997). Direct Manipulation and Other Lessons. Chapter 8. In: M. Helander, T. Landauer, P. Prabhu (Eds.). *Handbook of Human-Computer Interaction*. North Holland: Elsevier Science Publishers C.V.
- Harrison, B. L., Fishkin, K. P., Gujar, A., Mochon, C., and Want, R. Squeeze Me, Hold Me, Tilt Me! An Exploration of Manipulative User Interfaces. *Proceedings of CHI'98*, Los Angeles, CA, April 18-23, pp. 17-24.
- Howell, C. H. (1998). *Méthodes statistiques en sciences humaines*. Paris : De Boeck Université s.a.
- Hynckley, K., Pierce, J., Sinclair, M., Horvitz, E. (2000). Sensing Techniques for Mobile Interaction. *ACM UIST 2000. Symposium on User Interface Software and Technology*, CHI-Letter 2 (2), pp. 91-100.
- Hutchins, E.L., Holland, J.D., Norman, D.A (1986). Direct Manipulation Interfaces. In: Norman, D.A., Draper, S.W. (Eds.). *User Centered System Design : New Perspectives on Human-Computer Interaction*. Lawrence Erlbaum Associates, Publishers: London.
- Kenneth P. Fishkin, Thomas P. Moran, and Beverly L. Harrison. "Embodied User Interfaces: Towards Invisible User Interfaces". Proceedings of EHCI '98 (Heraklion, Crete, September 13-18). In Press.
- Jacob, R.J.K. *User Interfaces*. in *Encyclopedia of Computer Science*, Fourth Edition, ed. By A. Ralston, E.D. Reilly, and D. Hemmendinger, Grove Dictionaries Inc., 2000.
- MacKenzie, I. S., & Zhang, S. (1997). The immediate usability of Graffiti. Proceedings of Graphics Interface '97, pp. 129-137. Toronto: Canadian Information Processing Society.
- Mitrani, Marc. *13 assistants numériques passés à la loupe*. In PC Magazine, 169, 26-39, 2002.

- Norman, D.A (1986). Cognitive Engineering. In: Norman, D.A., Draper, S.W. (Eds.). *User Centered System Design Design : New Perspectives on Human-Computer Interaction*. Lawrence Erlbaum Associates, Publishers: London.
- Piaget, J. (1964). *Six études de psychologie*. France : Editions Denoël.
- Sears, A, Arora R. Data entry for mobile devices: an empirical comparison of novice performance with Jot and Graffiti. *Interacting with Computers*. 14, 413-433, 2002.
- Shneiderman, B. (1992). *Designing the User Interface: Strategies for Effective Human-Computer Interaction*. Addison-Wesley Publishing Company: USA.
- Smith, D.C, Irby, C., Kimball, R., Verplank, W., Harslem, E.. *Designing the Star User Interface*", Byte 7(4):242-282. Copyright 1982, McGraw-Hill, Inc.
- St. Amant, R. User interface affordances in a planning representation. *Human-Computer Interaction*, 14(3): 317-354. 1999.
- Van Oostendorp, H., Walbeehn, B.J. Towards modeling exploratory learning in the context of direct manipulation interfaces. *Interacting with computers*, 7(1), 3-24, 1995.
- Vigouroux, N., Poirier, F., Nigay, L. (1996). GT-M3 (GT5) : Modélisation de la Multimodalité. Extrait du Web le 23 septembre 2002 : [http://www.irit.fr/M3/Textes/Rapports/GT\\_M3\\_Rapp95\\_96.html](http://www.irit.fr/M3/Textes/Rapports/GT_M3_Rapp95_96.html).

## **Annexes**

## Annexe 1 : Questionnaire d'évaluation des sujets experts

Les résultats de ce questionnaire sont anonymes. Nous ne ferons jamais référence à votre nom et prénom lors de la présentation et la discussion de nos résultats. Pour des raisons d'organisation, à chaque utilisateur nous avons associé un numéro qui nous permettra de ne pas perdre les traces de vos résultats. Ce questionnaire fait partie d'une étude universitaire et il nous permettra de mieux connaître les utilisateurs ayant accepté de participer au test/entretien sur le sujet des PocketPC.

Pour plus d'informations vous pouvez contacter Roberto ORTELLI [rortelli@ticino.com](mailto:rortelli@ticino.com)  
Cette étude est valable pour le diplôme DES STAF de l'unité TECFA (<http://tecfa.unige.ch>) de l'Université de Genève (<http://www.unige.ch>).

Notes: Lisez attentivement les questions. Comme ce questionnaire ne s'effectue pas en présence d'un observateur, en cas de doutes veuillez contacter Monsieur Roberto Ortelli.  
Le questionnaire peut être rempli directement sur Microsoft Word et renvoyé par mail à l'adresse [rortelli@ticino.com](mailto:rortelli@ticino.com) ou bien vous pouvez l'imprimer et le remplir pour ensuite l'envoyer à cette adresse:

Roberto Ortelli  
27, chemin Sous-Le-Têt  
CH-1233 Bernex

Les frais d'envoi par courrier seront remboursés.

### 1. Informations générales:

Age:	.....	Profession:	.....
Genre:	<input type="checkbox"/> Homme <input type="checkbox"/> Femme <i>Remplacer la case par un X pour la réponse choisie (dans MS Word)</i>		

## 2. Informations concernant votre expérience avec les PDA

2.1 Quel modèle de PocketPC utilisez-vous actuellement?

.....

2.2 Depuis combien de temps possédez-vous ce PocketPC (2002)?

.....

2.3 Aviez-vous d'autres PocketPC (2002 ou autre) avant le modèle actuel?

*Remplacer la case par un X pour la réponse choisie (dans MS Word)*

Oui  Non

2.4 Si oui (question 2.3), combien de temps avez-vous utilisés ce(s) PocketPC?

.....

2.5 Aviez-vous d'autres PDA (PocketPC exclus) avant le PocketPC actuel?

*Remplacer la case par un X pour la réponse choisie (dans MS Word)*

Oui  Non

2.6 Si oui (question 2.5), combien de temps avez-vous utilisés ce(s) PDA?

.....

2.7 Comment estimez-vous votre degré de compétences global avec le PocketPC (2002)?

*Remplacer la case par un X pour la réponse choisie (dans MS Word)*

<input type="checkbox"/> Débutant	<input type="checkbox"/> Intermédiaire	<input type="checkbox"/> Expert
-----------------------------------	--	---------------------------------

### 3. Informations contextuelles

<p>3.1 Dans quel contexte utilisez-vous le PocketPC (2002)?  <i>Remplacer la case par un X pour la réponse choisie (dans MS Word)</i></p>		
<input type="checkbox"/> Au travail (professionnel)	<input type="checkbox"/> A la maison (privé)	<input type="checkbox"/> Au travail et à la maison
<input type="checkbox"/> Autre:		
<p>3.2 Quelle est votre fréquence d'usage de votre PocketPC (2002)?  <i>Remplacer la case par un X pour la réponse choisie (dans MS Word) ou ajouter le nombre de jours à la places des ...</i></p>		
<input type="checkbox"/> Tous les jours	... jours par semaine (indiquer le nombre de jours par semaine)	
<p>3.3 Quelle est la fréquence journalière d'usage du PocketPC (2002)? Faites une estimation.  <i>Remplacer les ... par un chiffre (dans MS Word)</i></p>		
J'utilise le PocketPC (2002) ... fois par jour		

### 4. Informations d'usage

<p>4.1 Quels sont les tâches que vous effectuez avec le PocketPC (2002)? <b>Plusieurs</b> réponses sont possibles.  <i>Remplacer la case par un X pour la(es) réponse(s) choisie(s) (dans MS Word)</i></p>		
<input type="checkbox"/> Rendez-vous/Calendrier	<input type="checkbox"/> Consulter/Editer les contacts	<input type="checkbox"/> Email
<input type="checkbox"/> Tâches	<input type="checkbox"/> Notes	<input type="checkbox"/> PocketWord
<input type="checkbox"/> PocketExcel	<input type="checkbox"/> Jeux	
<input type="checkbox"/> Autres (les énumérer):		
<p>4.2 Pouvez-vous quantifier en pourcentage le taux d'usage de chaque tâche par rapport à votre usage global du PocketPC (2002)?  <i>Remplacer les ... par un chiffre (dans MS Word)</i></p>		
...% Rendez-vous/Calendrier	...% Consulter/Editer les contacts	...% Email
...% Tâches	...% Notes	...% PocketWord
...% PocketExcel	...% Jeux	
...% Autres (les énumérer et attribuer un pourcentage pour chaque tâche):		

--

5. Informations relatives au médiums d'insertion du texte

5.1 Avec quel(s) médium(s) insérez-vous du texte? **Plusieurs** réponses sont possibles.

*Remplacer la case par un X pour la(es) réponse(s) choisie(s) (dans MS Word)*

<input type="checkbox"/> Clavier Virtuel (dans PocketPC)	<input type="checkbox"/> Reconnaissance des blocs
<input type="checkbox"/> Reconnaissance des lettres	<input type="checkbox"/> Transcrire
<input type="checkbox"/> Autres:	

5.2 Insérez à côté de chaque médium d'insertion de texte un chiffre indiquant votre usage réel et préférentiel des médiums que vous utilisez. (1 = médium le plus utilisé, 2 = le deuxième médium le plus utilisé, 3 = ..., et 0 = jamais utilisé ce médium)

*Remplacer les ... par des chiffres pour les réponses (dans MS Word)*

... Clavier (dans PocketPC)	... Reconnaissance des blocs
... Reconnaissance des lettres	... Transcrire
... Autres:	

5.3 Comment estimez-vous vos compétences par rapport à chaque médium d'insertion de texte?

Cochez une case allant de débutant à expert pour chaque médium.

*Remplacer les cases par un X pour les réponses choisies (dans MS Word)*

Clavier (dans PocketPC):

Débutant	Expert
<input type="checkbox"/>	<input type="checkbox"/>

Reconnaissance des blocs :

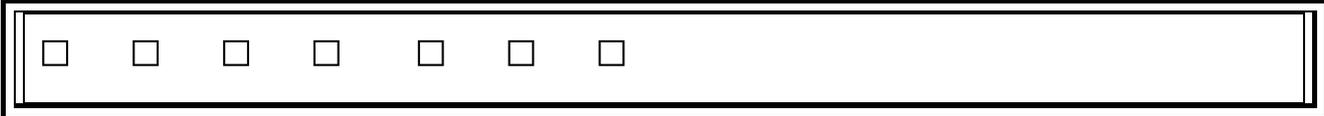
Débutant	Expert
<input type="checkbox"/>	<input type="checkbox"/>

Reconnaissance des lettres :

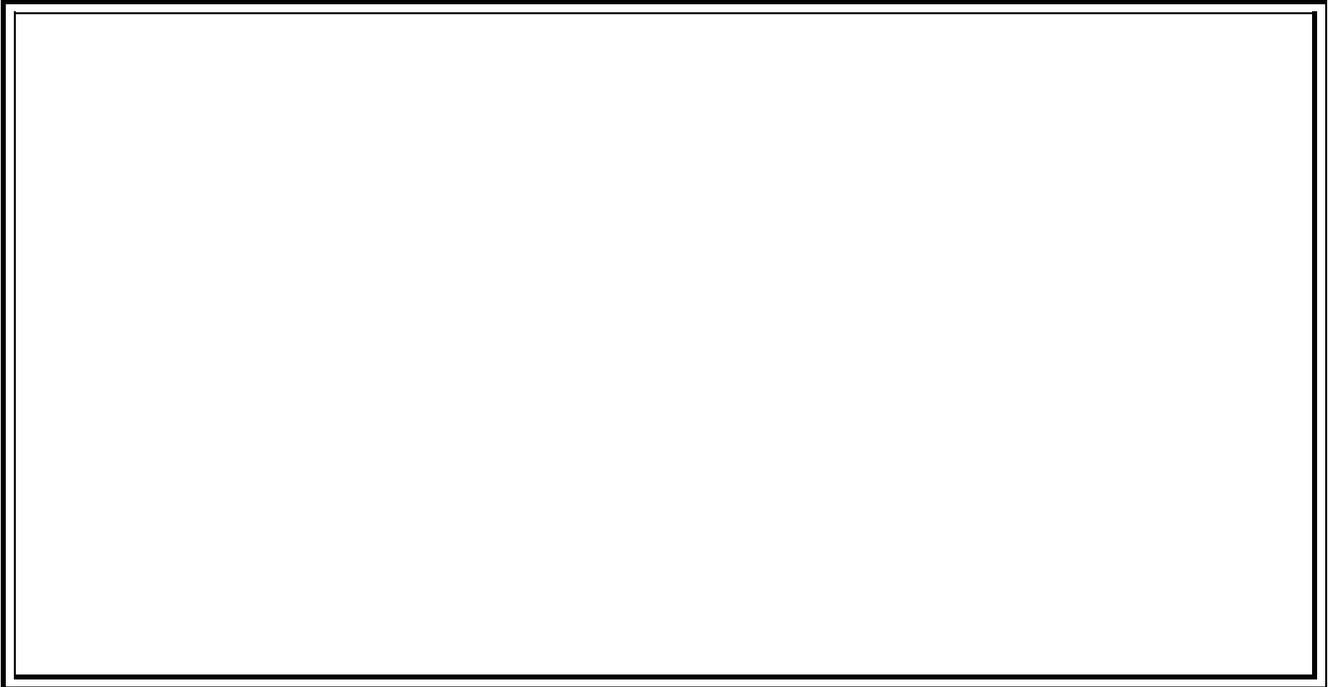
Débutant	Expert
<input type="checkbox"/>	<input type="checkbox"/>

Transcrire :

Débutant	Expert
<input type="checkbox"/>	<input type="checkbox"/>



6. Commentaires



Merci de votre collaboration !

## Annexe 2 : Questionnaire post-tâches fermées

Questionnaire soumis à tous les sujets après la passation de chaque série de tâches

### Partie I

1.1 Estimez le degré de difficulté des tâches que vous venez d'effectuer :

*Insérez une croix sur la ligne*

difficile \_\_\_\_\_ pas difficile

1.2 Estimez le degré de difficulté du médium utilisé :

*Insérez une croix sur la ligne*

difficile \_\_\_\_\_ pas difficile

1.3 Par rapport aux tâches que vous avez effectué, auriez vous préféré l'usage d'un autre médium ?

Oui  Non

1.3.1 Si oui, lequel ?

*Une seule réponse est possible*

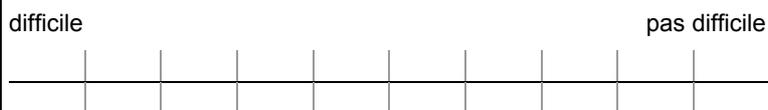
<input type="checkbox"/> Reconnaissance des blocs	<input type="checkbox"/> Reconnaissance des lettres
<input type="checkbox"/> Transcrire	

## Partie II

2.1 Pour chaque tâche effectuée évaluez leur degré de difficulté par rapport au médium utilisé :

2.1.1 Insérer un contact :

*Insérez une croix sur la ligne*



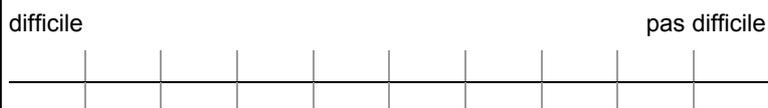
2.1.2 Insérer une adresse Web :

*Insérez une croix sur la ligne*



2.1.3 Insérer un rendez-vous :

*Insérez une croix sur la ligne*



2.1.4 Ecrire un e-mail :

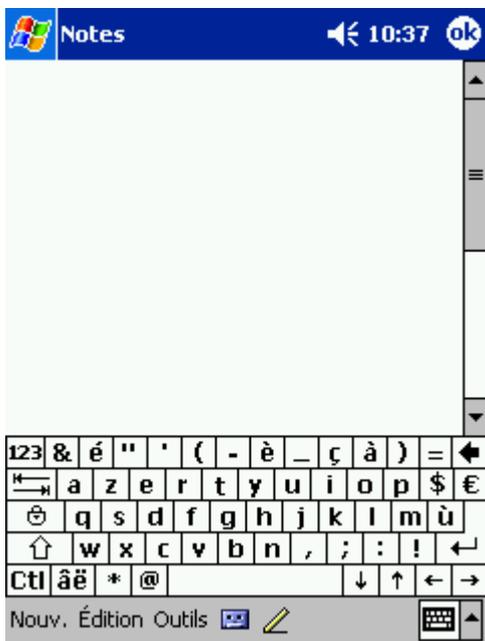
*Insérez une croix sur la ligne*



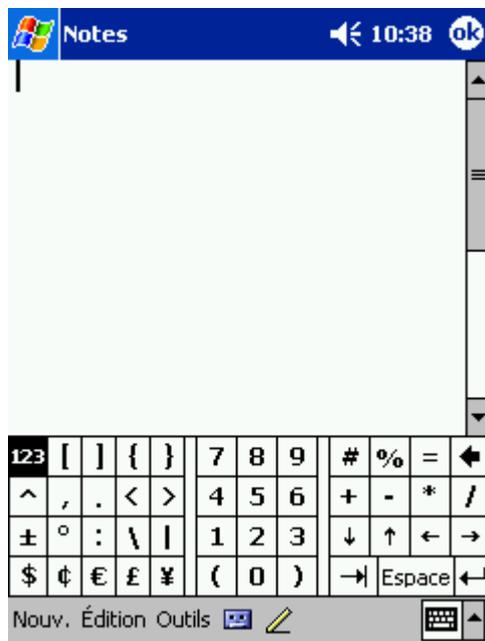
## Annexe 3 : Fiches SI

Fiche d'explication pour les novices du fonctionnement de chaque SI

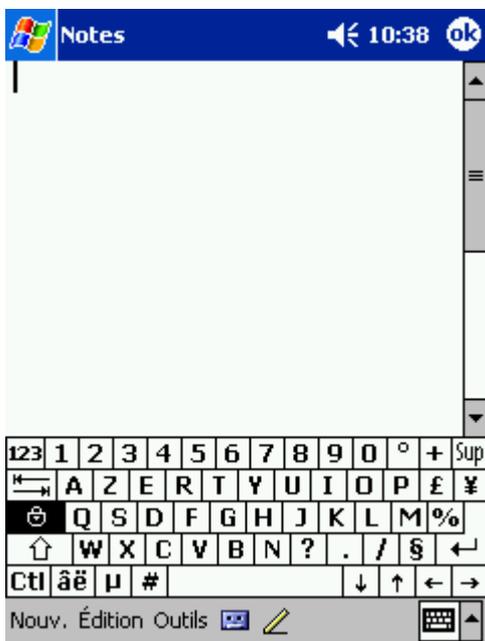
Clavier



Etat normal

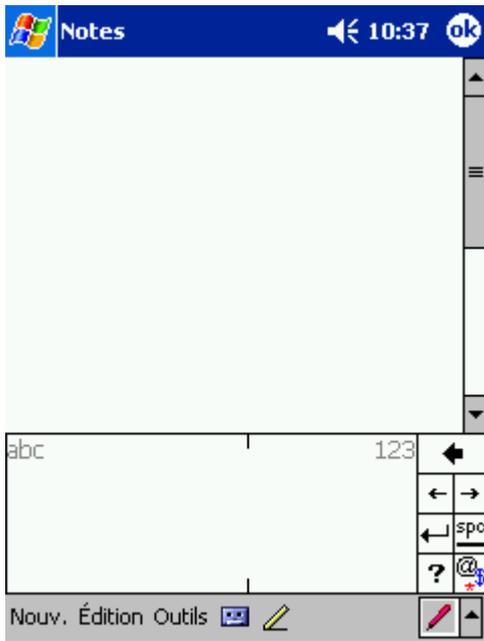


Appuyer sur la touche « 123 » pour accéder aux touches numériques

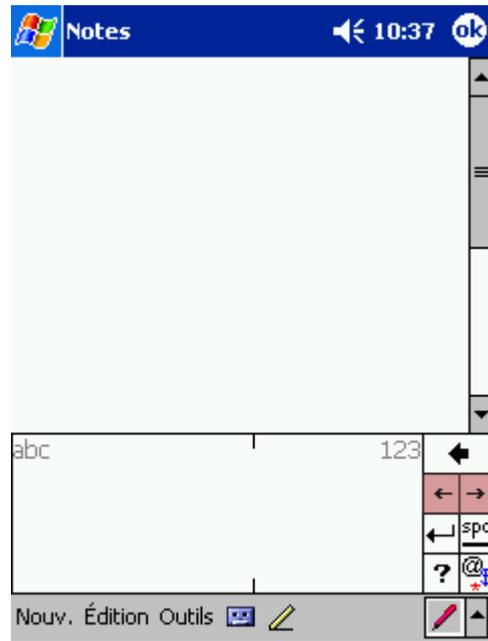


Appuyer sur la touche majuscule ou sur le cadenas (majuscule permanent)

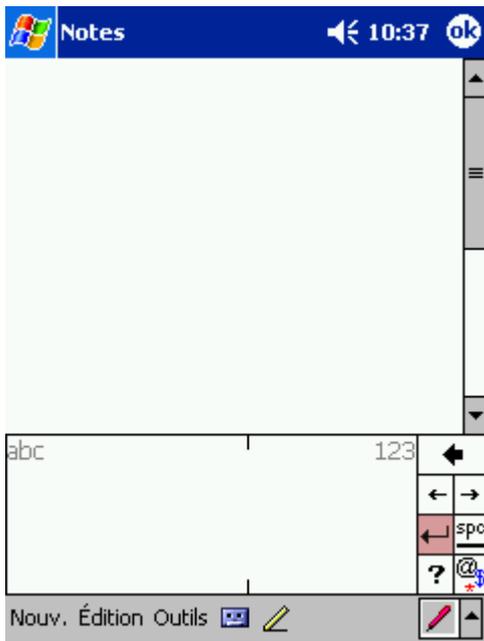
## Reconnaissance des blocs



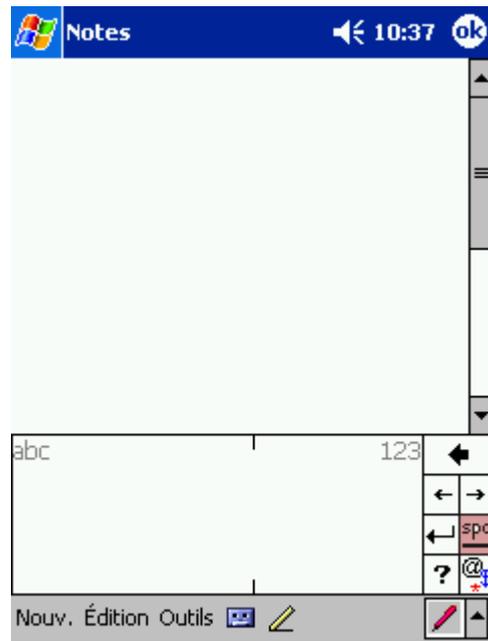
Etat normal



Déplacer le curseur à gauche ou a droite

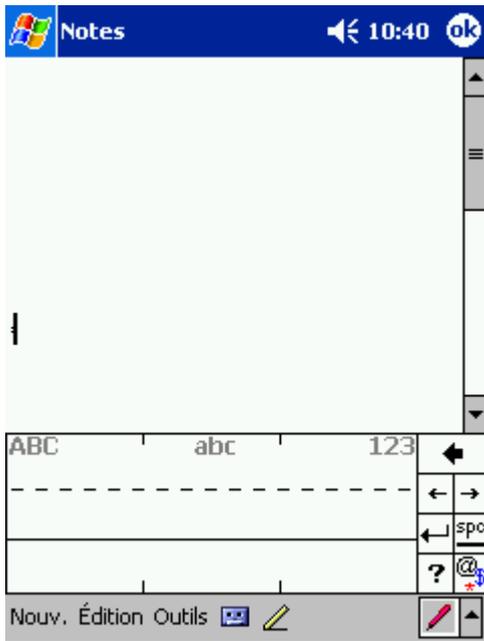


Aller à la ligne

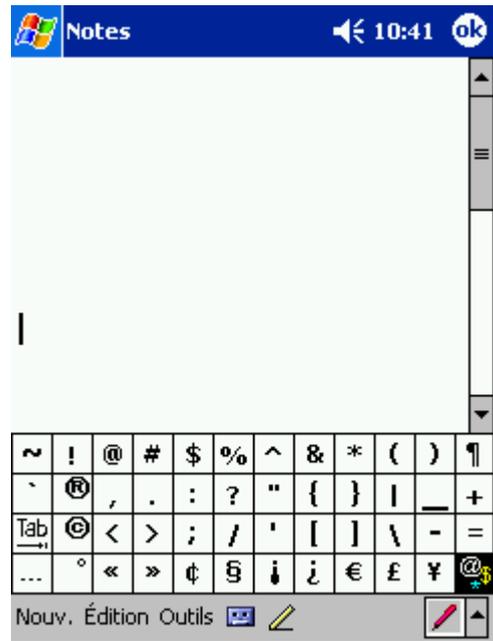


Créer un espace

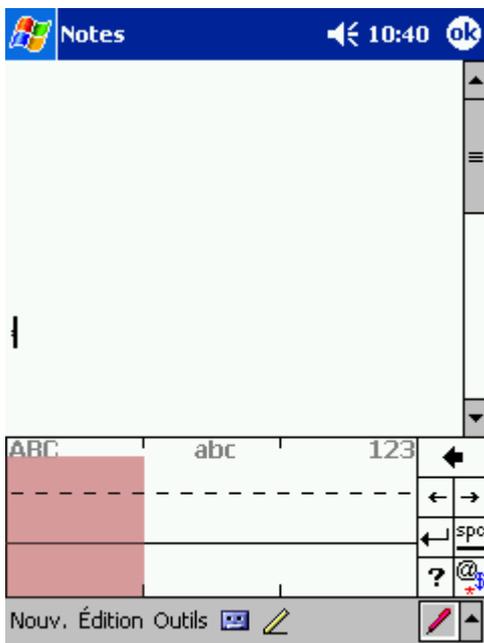
## Reconnaissance des lettres



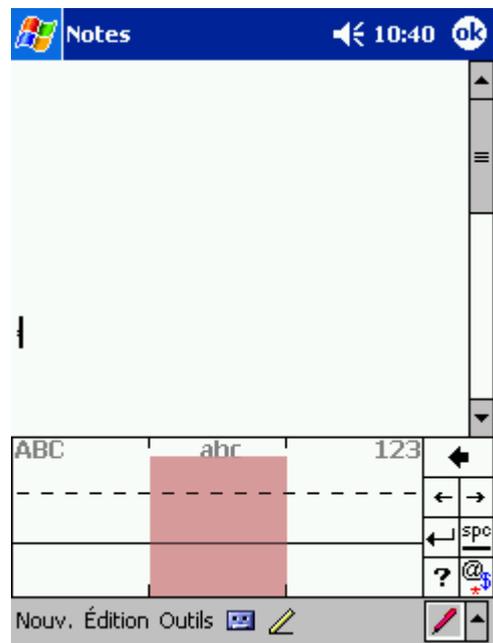
Etat normal



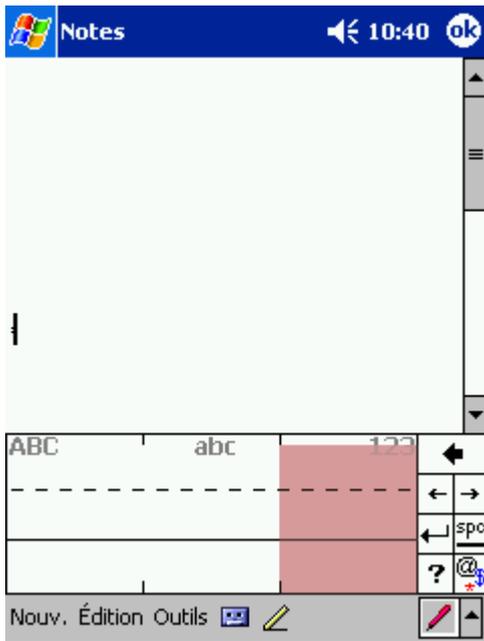
Appuyer sur la touche « 123 » pour accéder aux touches numériques



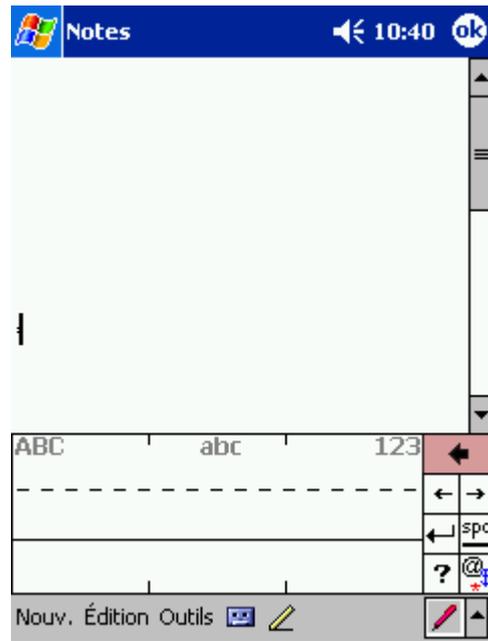
Ecrire les lettres en majuscule



Ecrire les lettre en minuscule

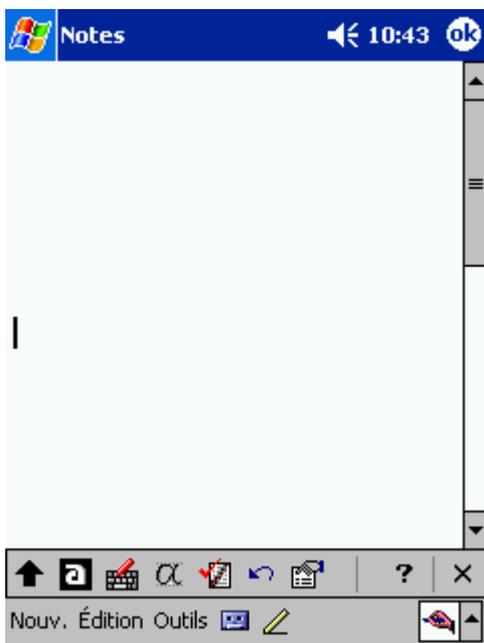


Ecrire des chiffres et des éléments de ponctuation

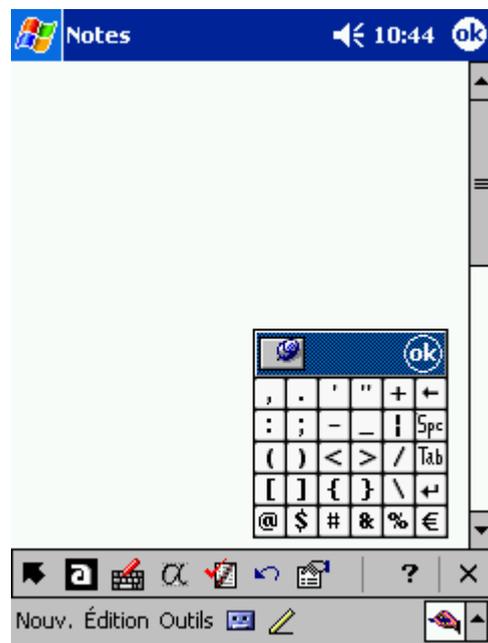


Effacer

### Transcrire



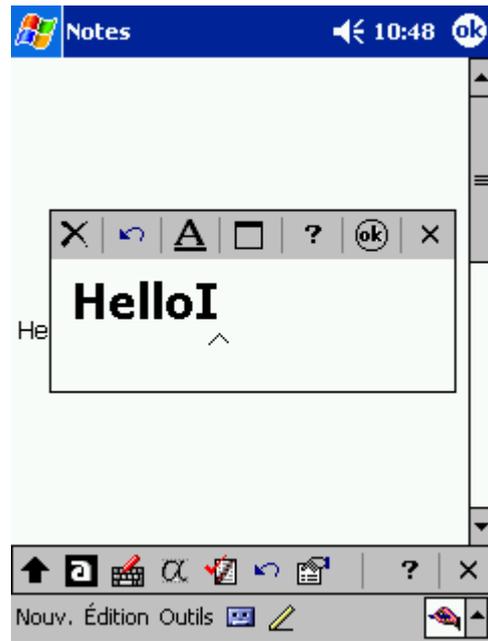
Etat normal



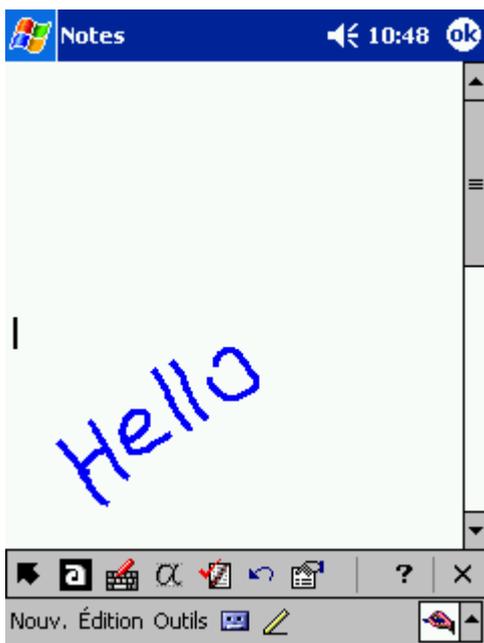
Symboles



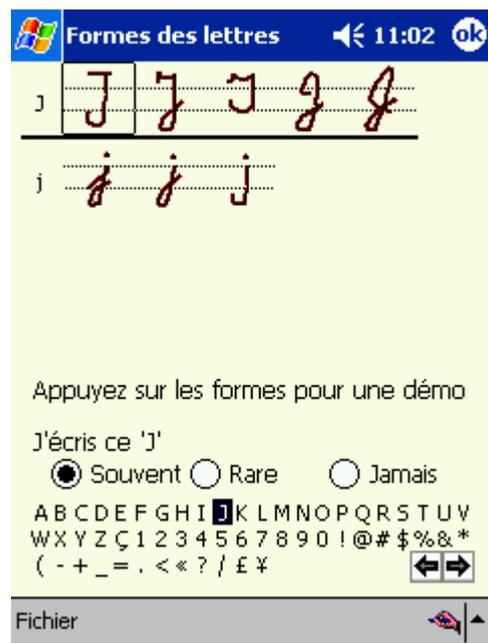
Ecrire en horizontal



Corriger ce qu'on vient d'écrire



Ecrire en diagonale



Aide

#### **Annexe 4 : Textes de familiarisation**

Exemple de texte à insérer pour se familiariser avec chaque SI

A l'école, c'est devenu la norme : chaque année, de la maternelle au primaire, les instituteurs voient arriver leur lot de petits démons qui entendent faire la loi. Voulant tout, tout de suite, refusant toute contrainte. Se roulant par terre ou hurlant à la moindre contrariété.

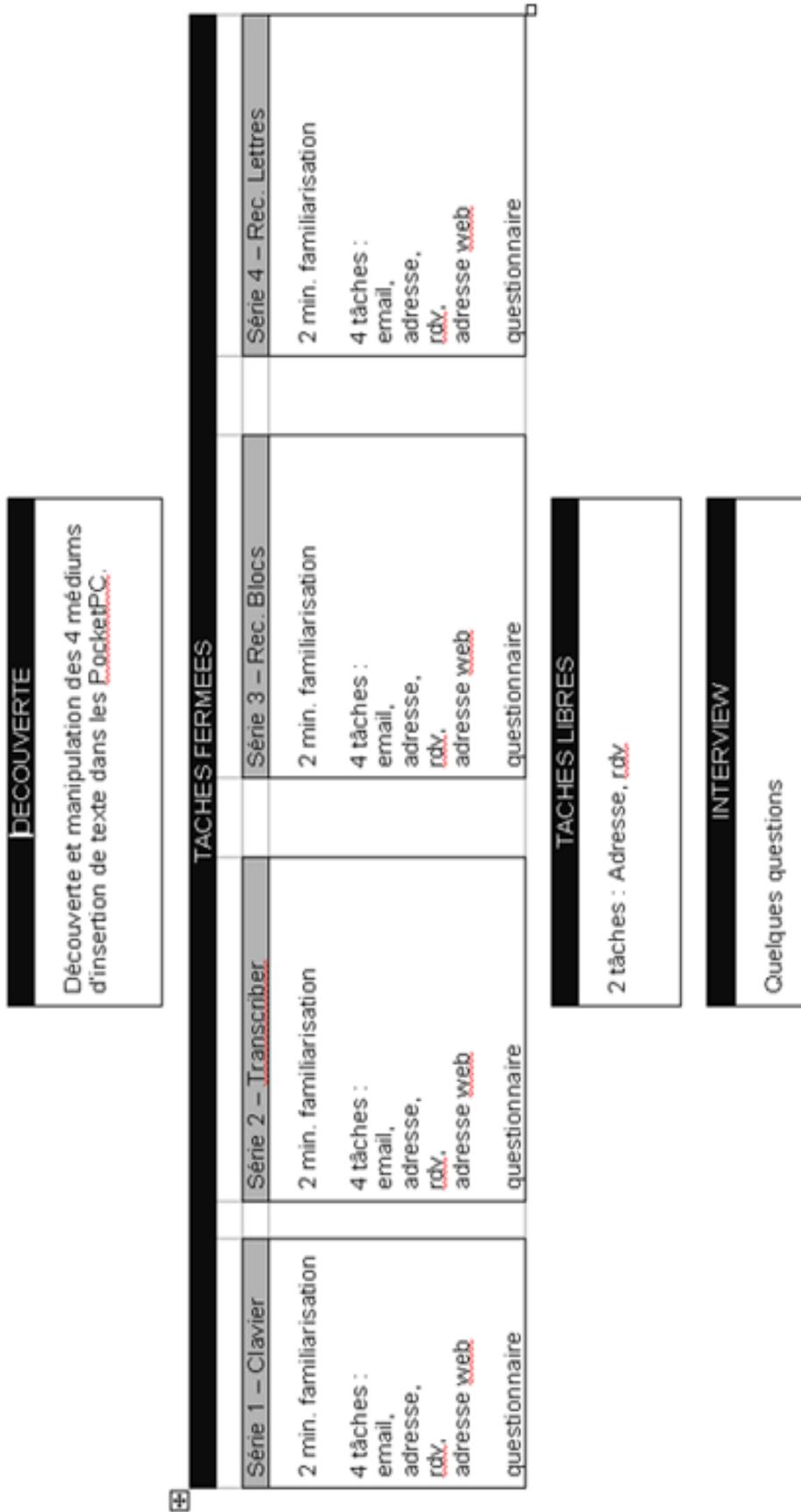
Dans la majorité des cas, l'école n'est en effet que le révélateur de ce qui se passe dans l'intimité du foyer. Un univers où l'enfant roi, désiré et choyé, placé au centre de tous les regards et adulé par notre société de consommation, fait de plus en plus souvent régner la terreur. Allant parfois jusqu'à insulter et frapper à répétition ses parents.

Mais le poison est là, en germe, qui ne demande qu'à grandir. Un danger dont les pédopsychiatres de l'hôpital Robert-Debré furent parmi les premiers à s'alarmer, en dressant dans *La Revue du praticien - Médecine générale* (tome XV, no 532), en mars 2001, un tableau inquiétant de cette forme particulière de "tyrannie intra-familiale".

Même réunis, ces "facteurs de risque" ne génèrent pas toujours des petits monstres. Tant s'en faut. Mais ils constitueraient un terreau favorable à l'omnipotence infantile, voire à la maltraitance "psychologique".

## Annexe 5 : Fiche de briefing

Fiche expliquant aux novices le déroulement du test. Le même schéma sans l'encadré découverte est présenté aux experts.



**Annexe 6 : Items tâches fermées**

1. Nom & Adresse

Masiero Fabrice  
Rue de Berne 34  
1247 Anieres

2. Adresse Web

<http://www.letemps.ch>

3. Rendez-vous

Objet : Medecin

Note : Demander les  
ordonnances

Date : celle proposée par  
défaut

4. Email

A :           roland@geneve.ch

Objet :      Nouveau lien

Texte :

Bonjour,

J'ai remarque un lien mort pour  
accéder a votre page de contact.

En effet, la page en question ne  
s'affiche pas.

Salutations.

1. Nom & Adresse

Nova Elia  
Route de Florissant 12  
1257 Morges

2. Adresse Web

<http://www.lematin.ch>

3. Rendez-vous

Objet : Cabaret

Note : Spectacle de Jean  
Pierre

Date : celle proposée par  
défaut

4. Email

A : nicole@valais.ch

Objet : Louer maison

Texte :

Bonjour,

J'ai vu que vous louez une maison. J'aimerais pouvoir me rendre sur place pour voir de mes propres yeux la maison en question.

1. Nom & Adresse

Oberson Daniel  
Rue de Geneve 56  
1324 Vernier

2. Adresse Web

<http://www.lameteo.ch>

3. Rendez-vous

**Objet : Courses**

**Note : Chercher la creme  
fraiche**

**Date : celle proposée par  
défaut**

4. Email

A : mireille@vaud.ch

Objet : Recherche chat

Texte :

Bonjour,

Sur votre site vous dites que vous avez treize chats a vendre.

Nous sommes interesses a pouvoir venir chez vous pour les voir.

1. Nom & Adresse

Lauper Claude  
Rue de Vermont 78  
1227 Carouge

2. Adresse Web

<http://www.liberte.ch>

3. Rendez-vous

Objet : Fitness

Note : Préparer les  
serviettes

Date : celle proposée par  
défaut

4. Email

A : elia@fribourg.ch

Objet : Recherche jobs

Texte :

Bonjour,

Sur le journal j'ai vu que vous cherchez un collaborateur. Je suis intéressé au poste. Je reste à votre entière disposition.

**Annexe 7 : Items des tâches libres**

2. Rendez-vous

Objet : Annonce

Note : A rendre les  
descriptions

Date : celle proposée par  
défaut

1. Adresse

Curchod Robert  
Rue du Lievre 14  
1326 Meinier

## **Annexe 8 : Entretien novices**

Questions pour les novices lors de la partie des entretiens

Questions liées aux tâches :

1. Pourquoi dans les tâches libres avez-vous choisi ce médium ?
2. Dans l'ensemble des tâches que vous venez d'effectuer, avec quels médium estimez-vous avoir été le plus performant ?
3. Quel est votre évaluation par rapport à tous les médiums utilisés ? Quels sont les avantages et les inconvénients de chaque médium ?
4. Quel médium préférez-vous. Essayez de classer les quatre médiums utilisés.

Questions liées à l'utilisation :

5. Quels facteurs influencent votre choix du médium à utiliser ? Est-ce qu'il y a des facteurs susceptible modifier votre comportement ?

C'est difficile à estimer vu que vous n'avez pas une expérience, mais essayer d'imaginer des facteurs qui pourraient influencer votre comportement.

## **Annexe 9 : Entretien experts**

Questions pour les experts lors de la partie des entretiens

Questions liées aux tâches :

6. Voir si le sujet a choisi le médium de sa préférence

Sujet X : ...

Si il a utilisé un autre ou s'il a utilisé différents médiums selon la tâche, lui demander le pourquoi.

7. Évaluez les dernières tâches que vous venez d'effectuer
8. Globalement, avec quel médium pensez vous avoir été le plus performant ? Pourquoi ?
9. Quel est votre évaluation par rapport à tous les médiums utilisés ? Quels sont les avantages et les inconvénients de chaque médium ?

Questions liées à l'utilisation :

10. Quels facteurs influencent votre choix du médium à utiliser ? Est-ce qu'il y a des facteurs susceptibles de modifier votre comportement ?
11. Est-ce qu'il y a des situations dans lesquelles vous choisissez de changer de médium ?
12. Comment vous vous comportez lorsque vous effectuez des erreurs ? Est-ce que vous changez de médium par exemple ?

Autres questions :

13. Est-ce que après ce test vous envisagez d'apprendre ou d'utiliser en futur un autre médium ?  
Si oui, pourquoi ?
14. Une fois acheté le PocketPC, comment avez vous découvert les médiums et comment vous les avez appris ? Avec quels critères vous avez décidé d'utiliser de préférence le ... clavier ... et la reconnaissance des blocs en écartant les autres ?

## Annexe 10 : Analyse des entretiens

Nous avons séparé les réponses des experts et des novices pour mieux les analyser. Nous n'avons pas retranscrit toutes les discussions mais nous nous sommes contentés de retenir les éléments pertinents.

### 1a. NOV: Pourquoi avez-vous choisi cette (ces) modalité(s) dans les tâches libres ?

Sujet 12

RB : quand il écrit bcp il préfère car il arrive à lire et à écrire en même temps. Bonne mémorisation. Plus libre que CV.

CV : peut pas lire et écrire en même temps. Pour les chiffres il faut toujours changer de clavier, faut des aller et retour continue. Il va vite mais il faut plus d'attention.

Sujet 14

RL : le plus simple à utiliser

Sujet 21 :

RB : parce que le dernier utilisé et il était à l'aise

TR : c'était du texte et je pensais que ça allait bien

Sujet 15

TR : j'ai vu que pour cette même tâche ça allait relativement vite

CV : si il y a des chiffres c'est mieux le CV que le TR

Sujet 17

CV : j'ai pas eu de gros problèmes

Sujet 18

CV : facile, simple à utiliser

### 1a. EXP: Pourquoi avez-vous choisi cette (ces) modalité(s) dans les tâches libres ? Est-ce que le sujet a t'il utilisé une modalité différente par rapport à ce qu'il a indiqué dans le questionnaire ?

Sujet 1

A choisi le CV car :

- c'est le moyen le plus simple, efficace, rapide
- avec Messenger c'est le meilleur moyen, j'ai essayé un peu tout, le clavier c'est ce qui va mieux

<p>Sujet 3 :</p> <p>C'est celui que il utilise le plus souvent</p>
<p>Sujet 5</p> <p>CV : plus facile, utilisation courante. Il est fort probable que si je m'habituerai à un autre médium, je l'utiliserais aussi bien que le CV</p>
<p>Sujet 11</p> <p>RB : pas réfléchi, l'habitude, je n'utilise que rarement les autres</p>
<p>Sujet 2</p> <p>RB : plus performant, connaît le mieux</p>
<p>Sujet 25</p> <p>CV : il pense que comme c'est celui qu'il utilise tout le temps, c'est forcément le meilleur, c'est aussi l'impression qu'il a eu au cours du test</p>
<p>Sujet 4</p> <p>CV : indécis avec RB, car il les utilise couramment les deux, pour des petits trucs il utilise RB, mais s'il doit écrire des emails il préfère utiliser CV, car il doit être sûr d'écrire tout correctement.</p>

### 3. EXP: Globalement avec quelle modalité estimez-vous avoir été le plus performant ?

<p>Sujet 1</p> <ol style="list-style-type: none"><li>1. CV : meilleur, en plus, normalement il utilise la liste prédictive, donc c'est assez rapide</li><li>2. RB : manière d'écrire assez intuitive</li><li>3. RL : s'adapte à une écriture assez standard. La séparation majuscules, minuscules, chiffres assez pratique</li><li>4. TR : il n'est pas vraiment utilisable. Peut être utilisable sur une TabletPC, avec un écran plus grand on aurait vraiment de la place pour écrire</li></ol>
<p>Sujet 3</p> <p>RL et Clavier : sont les deux qu'il utilise couramment</p>

Sujet 5

4. TR : il serait bien si moins épais (le ppc) et l'écran disposé différemment. Il aime le principe mais à son avis c'est un peu long. Car il faut attendre. Si on écrit une phrase il faut penser en termes d'espace qui va prendre à l'écran

1. CV : facile

2. RB : bien, rapide à condition qu'on l'utilise souvent et beaucoup

2. RL : assez bonne, mais il l'utilise pas

Sujet 16

RB (1) le plus performant car il a aussi une expérience du Palm. Il convaincu que même si l'on commence de 0, entre RB et RL (4), RB l'emporte parce qu'il est plus tolérant et en plus c'est facile à mémoriser les caractères de RB. TR (3) va très bien mais seulement beaucoup d'exercice, à long terme il saura le meilleur. CV (2) pour certaines chose il va bien mais je l'utilise jamais.

Sujet 11

1. RB : meilleur, toujours utilisé, libre de faire ce qu'il veut, pas de contrainte des touches

2. CV : précis, ça va vite, attention à ne pas appuyer sur d'autres touches

3. RL : pas mal, implique plus d'efforts que RB, bon substitut de RB

4. TR : pas mal d'habitude, parfois adapter son écriture, prends du temps pour reconnaître, pénible pour les erreurs

Sujet 2

1. RB

2. CV

3. RL

3. TR

Sujet 25

1. CV : faut rien apprendre, tout est là, utilisation de raccourcis clavier. L'absence de point dans le clavier principal ça le dérange.

2. RB : vite appris lorsqu'il a acheté le premier PPC. Connaît pas des symboles spéciaux ou comment créer les raccourcis tels que copier-coller. Mais il ne l'utilise pas souvent. Comme CV est toujours le SI par défaut il l'utilise sans se poser la question d'en utiliser un autre.

3. TR : il est habitué à l'utiliser, mais il lui semblait que la sensibilité de l'écran était différente, que celui expérimentale était moins sensible. Parfois trop d'erreurs, ça l'énerve. Faut rien apprendre, il y a un aide pour corriger les erreurs, peut proposer une liste de mots pour corriger les fautes, le clavier est pas mal même si n'est pas standard par rapport à ceux qui se trouvent dans les 3 autres SI

4. RL : connaît pas beaucoup, il semble qu'il a fait pas mal d'erreurs, il avait de la peine avec les « i » et le « e ». Il n'aime pas trop, il a préféré focalisé son apprentissage sur RB car selon lui il est inutile d'apprendre trop de SI.

Sujet 4

1. CV ou RB
3. RL
4. TR

3. NOV: Globalement avec quelle modalité estimez-vous avoir été le plus performant ?

Sujet 12

1. RB:
2. CV :
3. RL :
4. TR :

Sujet 13

TR (4): plus compliqué. Faut se souvenir à chaque fois la ponctuation, aller à la ligne, majuscules difficile, peu intuitif, pour modifier, le point pas efficace. Bonne reconnaissance mais problèmes pour modifier ou la ponctuation.

RL (2): pas mal, easy

CV (1) : plus à l'aise, intuitif, azerty un peu un problème. Mêmes touches que clavier classique

RB (3) :

Sujet 22

2 RB : même si des fautes

4 RL : pas aimé. Intéressant car tout à disposition, même facilités que RB mais pour certaines lettres très difficile

1 CV : pas de fautes

3 TR : beaucoup trop de fautes. Il faudrait des fonctionnalités comme dans le traitement de texte (points ou virgule...). Il faudrait un mélange avec RB ou des fonctionnalités en plus. Faudrait améliorer l'interface. Le clavier est différent que RB, apprendre de nouvelles notations. Ecriture naturelle pas une lettre à la fois.

Sujet 14

- 1 RL
- 2 RB
- 3 CV
- 4 TR

Sujet 21

2 CV : moins d'erreurs, peu de risque mais pas super rapide

1 RB : faut beaucoup apprendre mais après peu de chances de se tromper

3 RL : difficile pour ponctuation  
4 TR : catastrophique pour les erreurs

Apprentissage

RB apprentissage assez sympa, lettre simples, une fois que tu les connais ça marche bien mais il faut apprendre caracteres speciaux. Si t'as rien sous la main, t'es un peu emmerdé  
TR si tu fais des erreurs et des caractères spéciaux tu t'es sors plus. Tu écris vraiment.  
CV pas d'apprentissage intéressant ou particulier  
RL finalement il marche pas mal mais je ne sais pas pourquoi, pas convaincu. C un compromis entre RB et TR. Je vois pas finalement l'utilité.

Sujet 15

1. CV
2. TR
3. RL
4. RB

Sujet 17

1. CV
2. RB
3. RL
4. TR

Sujet 18

1. CV
2. RB
3. RL
4. TR

**4. EXP: Évaluez toutes les modalités utilisées au cours du test. Identifiez les avantages et inconvénients de chaque modalité.**

**Sujet 1**

CV

- si tu vas vite, le CV fais des espaces pas souhaités (avec PPC personnel et aussi avec celui expérimental)
- rapide, pas d'apprentissage
- les lettres sont définies
- une fois qu'on a l'habitude à utiliser un ou deux raccourcis c pratique

TR

- ce qui me dérange c'est encore une fois que comme j'écris assez mal, j'écris assez grand, je suis limité par la taille de l'écran
- à chaque fois qu'on écrit un mot il faut toujours attendre qu'il reconnaisse, ça ralentit passablement la saisie quand même
- il suffit qu'il reconnaisse mal une lettre et c'est la catastrophe, revenir, effacer la lettre, changer la lettre, c'est vraiment peu pratique. Avec les autres, on peut corriger toute suite la faute
- Il faudrait un dictionnaire plus précis pour reconnaître les erreurs : il rajoute les accents mais si on met un o à la place d'un a, le système est perdu

RB

- c'est peut être la reconnaissance que je préfère
- j'ai perdu un peu l'habitude mais ça revient assez vite
- manière assez simple à écrire

RL

- ça se rapproche de mon écriture normale, mais comme j'écris assez mal, j'ai de la peine avec certains caractères

Sujet 3

CV : rapide

RL : si on écrit trop vite la reconnaissance ne suit pas,

TR : très sensible, écran trop petit, il ne l'utilise jamais

RB : ça va

Sujet 5

RL : il préfère le moins cette modalité. Il se fait toujours avoir par le « i », mais il faut l'apprendre. C pas naturel par rapport à RB, il aimerait bien retrouver le même « i » de RB dans RL.

TR : épaisseur de l'appareil empêche une bonne manipulation de cette modalité. Il faudrait pouvoir l'utiliser en horizontal pour mieux en profiter. Pas mal le clavier.

RB : tous les caractères spéciaux en un seul emplacement

CV : le fait que soit en azerty, des fois il doit chercher les lettres. Lui de toute façon il n'écrit jamais de longs textes. L'habitude fait que c'est bien. Certains caractères mal placés comme le point et la virgule devraient être côte à côte ... il faudrait pas changer de « clavier » pour accéder au point.

Faudrait une ligne supplémentaire avec les chiffres

Sujet 16

RB : on écrit des symboles qui rappellent des lettres, écrire majuscule/minuscule très simple. Faut l'apprendre mais facile à retenir

CV : va bien mais pour écrire vite c'est pas bon, il n'aime pas à chaque viser chaque lettre. Alors il préfère acheter un clavier classique et le brancher au ppc.

TR : pratique mais le fait d'écrire entre le texte ... attente gênante. Il faut une bonne écriture.

RL : s'il n'y aurait pas RB je l'utiliserais. C'est laborieux, t'es pas libre, c'est un peu plus contraignant par rapport au RB

Sujet 11

RB : vitesse, simplification de l'écriture, amusant, ludique, précis, pas d'efforts, un peu écrire en grand ou en petit, apprentissage se fait vite, mémorisation de l'écriture facile. Pas de possibilité d'annuler lorsqu'on sélectionne la modalité « symboles » (le point)

CV : concentré sur l'écran continuellement (-), précision, modifier taille des touches, AZERTY (-), existence de raccourcis clavier

RL : un peu d'efforts pour des lettres comme le i, t, f ... RB plus simple. Moins d'espace pour écrire que RB. Pas d'apprentissage

TR : difficile au début, on peut paramétrer l'écriture, profiter de tout l'écran

Sujet 2

RB : connaît depuis beaucoup de temps, simple, efficace, peu d'erreurs, intuitif

CV : c un clavier normal, rien à apprendre

RL : connaissait pas, plus d'actions à effectuer, ridicule la séparation en trois de la zone de saisie, pas compris toute l'interface

TR : difficile pour corriger, trop de fautes, faudrait de la ponctuation sur la barre en bas

Sujet 4

TR : il ne le connaissez pas avant, quand il a acheté le nouveau iPaq il l'a remarqué mais il ne l'a pas testé. Pourtant il a trouvé intéressant à utiliser. Il pense que le fait qu'il a une bonne écriture il avantage la reconnaissance de son écriture. Mais il ne sait s'il pourrait l'utiliser tout le temps. Il a apprécié qu'on peut écrire une phrase, il aime aussi exploiter tout l'écran du PPC.

CV : rien à dire

RB : il est intéressant à utiliser, il dit qu'il n'a pas eu de la peine pour l'apprendre.

RL : il trouve un peu « mécanique », moins intéressant que RB, il trouve que c'est un doublon, mieux vaut en connaître un SI que tous les SI et se confondre par la suite. Il pense qu'au début il permet avec le CV d'interagir facilement avec le PPC.

4. NOV: Evaluez toutes les modalités utilisées au cours du test. Identifiez les avantages et inconvénients de chaque modalité.

Sujet 12

CV : connaît bien la disposition des touches, organisé, petit espace, on trouve vite à repérer les touches mais parfois ça peut mener a la confusion. Ponctuation, le point ne figure pas, faut aller le chercher. Il suffit de cliquer pour faire apparaître la lettre, pas mal.

RB : compris le fonctionnement des lettres. Mais confusion. Faut comprendre comment écrire et le mémoriser. Ponctuation difficile. Faut être précis. Rapidité, biens séparées les deux zones, beaucoup

d'espace. Original en plus que RL. Aller à la ligne pas mal.

RL : pas mal, RB et RL mêmes performances d'usage, trop banal celui-ci. RB plus fasciné.

Pragmatique car trois zones, RB plus inventif. Trois zones, moins d'espace.

TR : obscène, limitation de l'espace à disposition, réfléchir comment écrire par rapport à l'espace à disposition. J'étais sous pression, car il fallait écrire tout le mot, car sinon il rajoute l'espace après je dois effacer, c'est difficile, ...

#### Sujet 13

CV : intuitif, commandes prédéfinies, on ne peut pas se tromper. A court terme plus efficace. Pour la précision c'est le plus efficace. Pénalisant pour longs textes.

RB : - rappeler les symboles

RL : moins d'erreurs que TR, donc moins modifications

TR : moins intuitif pour la ponctuation, même si il y a une bonne reconnaissance mais pas pour tout. A moyen terme plus efficace car on écrit directement tout le mot ou toute une phrase, pour des longs textes ça pourrait être plus efficace.

#### Sujet 22

RB : ça vaut la peine d'apprendre un nouveau langage car tu vas avoir moins d'erreurs. Mécanique. Plus proche de notre écriture naturelle. Sur le plan cognitif c'est dur d'apprendre un nouveau langage, il faut continuellement se référer au nouveau alphabet. Caractères spéciaux difficile sans les aides. Ça va vite si t'as binet intègre mais ce qui handicape c'est le fait d'écrire lettre après lettre et pas un mot en entier.

CV : chaque fois une recherche, beaucoup plus travailler, fatigant à la fin. A la fin je sentais que ça allé vite, mais le fait de à chaque fois changer de clavier (symboles vs classique). Le point mal placé. Le fait de devoir aller rechercher l'endroit, faire le chemin sur le clavier, mais ça à l'air d'être plus rapide, t'as pas à faire tout le truc..., peu d'erreurs j'imagine, mais la jonglette entre ces deux interfaces la c'est négatif, c'est pas évident. Azerty ridicule.

TR : c plus rapide si t'as une bonne écriture, tu dois te forcer de bien écrire, tu peux écrire plus vite et plus facilement, c'est plus spontané mais t'as pas de contrôle des fautes d'interprétation donc autant mieux choisir RB. TR à terme c'est le summum, mais c'est pas encore suffisamment performant. C'est insupportable de corriger les erreurs ce qui prends énormément de temps.

RL : plus facile que RB. Simplification de la zone sensible et le langage libre pas mal mais à nouveau il faut apprendre des trucs. Mais à terme c'est peut être plus rapide que RB.

#### Sujet 14

RL : plus rapide, structure, paramétré. Il se rapproche le plus de l'écriture qu'on a. Moins de besoin de s'adapter nous que RB ou s'est nous qu'il faut s'adapter à la machine. Mais quand même mes T ou mes i il fallait les répéter plusieurs fois. Utilisation à court terme.

TR : trop facile de se tromper si on ne s'habitue pas. C'est un peu plus brouillon. Plus flexible (direction, en fonction de comment on le tient). Toucher les boutons par erreur. Je pense que pour quelqu'un qu'il utilise beaucoup c'est celui là qui va choisir, je pense. Une fois qu'on sait habitué peut être que c'est celui plus rapide et confortable par rapport à la position des mains. Le problème c'est lorsqu'on faisait des erreurs, c'était difficile de les corriger, j'étais presque incapable. J'hésite si je l'utiliserais.

RB : faut bien mémorisé, une fois utilisé plus je pense que c'est un peu pareil que RL. Une fois

qu'on connaît par cœur l'écriture ça peut aller très vite. En plus les lettres sont facilitées, en perd moins de temps, par exemple pour un A qu'un A normal. Je pense qu'il pourrait être même plus rapide que TR. Si on ne l'utilise pas souvent on oublie par conséquent on fait plus d'erreurs. Utilisation à court-moyen terme car le plus performant.

CV : préfère taper sur un clavier classique qu'à l'écran, mais c'est facile, moins de chances de faire des erreurs, mais c'est long. Azerty, recherche des lettres. On a accès à tout. Lenteur. Pointer sur chaque lettre... Je ne l'utiliserais pas.

#### Sujet 15

CV : azerty dérange un peu, ponctuation, surtout le point, faudrait une ligne avec les chiffres, le reste facile à apprendre, comme un clavier normal

TR : sympathique au début, mais parfois pénible pour les erreurs, faudrait plus de fonctionnalités ou plus d'options comme le point par exemple qui soit visible tout le temps

RB : pense qu'il soit mieux que RL à long terme, car moins sensible, faut s'adapter au système

RL : machine s'adapte au système, facile à utiliser, presque rien à apprendre

#### Sujet 17

CV : rien à dire si ce n'est pour le point

RB : pas compliqué une fois appris les différents symboles, faudrait plus d'apprentissage et continuité pour vraiment l'exploiter, probablement plus efficace que CV

RL : pas mal, faut faire vite pour certaines lettres où il faut deux traits pour indiquer au système qu'il faut qu'il reconnaisse le t par exemple, mais sinon c'est proche de RB

TR : pas vraiment aimé, surtout pour l'adresse web. Je me suis plus sorti avec ces erreurs, à la fin je ne les corrigeais plus

#### Sujet 18

CV : je pense que c'est le médium que j'utiliserais au début une fois acheté car il ne faut rien apprendre

RB : je préfère, je l'avais déjà vu sur palm mais il faut l'apprendre, je ne sais pas si je pourrais, j'aimerais utiliser toute suite l'engin.

TR : j'ai une écriture difficile à encoder pour le PPC, il avait du mal, je devais parfois changer d'écriture, faudrait un système ou on dit au système comment l'on écrit pour qu'il soit plus précis

RL : c'est presque comme RB, il est plus flexible, mais parfois j'avais de la peine pour le retour au chariot ou d'autres éléments de ponctuation, alors que RB c'était plus intuitif sous cet aspect.

Questions liées à l'utilisation :

5. EXP: Quels facteurs influencent votre choix du médium à utiliser ? Est-ce qu'il y a des facteurs susceptible modifier votre comportement ?

Sujet 1 :

CV : 90% d'usage  
RB : pour essayer, pour voir si ça va mieux, mais il revient toujours au CV  
La rapidité influence son choix, beaucoup plus rapide avec le clavier

Sujet 3

Il utilise le dernier SI utilisé en précédence. Normalement c'est le CV. Il passe de RL à CV en cas d'erreurs ou s'il n'arrive pas à effectuer un caractère

Sujet 11

Presque toujours RB, parfois TR pour des notes, CV lorsque je veux être précis avec les accents

Sujet 2

Toujours RB, dans le train ou en voiture CV, à cause des vibrations

Sujet 25

Je me pose pas la question du médium à utiliser car j'utilise le premier qui se présente à l'écran, en l'occurrence il s'agit presque toujours le CV (le système enregistre le dernier SI utilisé)

Sujet 4

CV pour de longs textes ou si il veut être précis, RB pour le reste

**5. NOV: Quels facteurs influencent votre choix du médium à utiliser ? Est-ce qu'il y a des facteurs susceptible modifier votre comportement ?**

Sujet 13

CV : adresse web  
TR : adresse web très, trop difficile. Pour long texte plus efficace

Sujet 21

Ponctuation, dès qu'il y a trop de points c'est difficile de faire des phrases, je ne sais pas. Moins de prise de tête avec TR car t'as des raccourcis. Une fois que t'es dedans ça va vite. Pour être précis, je ne sais pas. Pour corriger dans TR c'est difficile, y a l'utilitaire qui est pas mal.

Sujet 15

Sujet 17

Sujet 18

**6. EXP: Est-ce qu'il y a des situations dans lesquelles vous choisissez de changer de médium ?**

Sujet 1 :

Rendez vous : RB pourrait aller, mais je serais de toute façon plus lent qu'avec le clavier.  
Pour de longs texte CV beaucoup plus performant

Sujet 3

CV : adresses car ça permet de switcher vite aux symboles et aux caractères spéciaux. En plus normalement c'est le clavier a apparaître comme modalité par « défaut ». J'utilise le machin qui est déjà ouvert

RL : pour des notes, personnelle, faut pas faire trop attention aux fautes. Il switche au clavier s'il n'arrive pas à faire un caractère mais normalement il ne passe jamais d'une modalité à une autre

Sujet 16

Il utilise toujours RB mais sinon, s'il a des oublis il utilise l'aide sous forme de clavier, mais il ne change pas de SI

**7. EXP: Comment vous vous comportez lorsque vous effectuez des erreurs ? Est-ce que vous changez de médium par exemple ?**

Sujet 1

En règle générale il corrige toute suite. Avec le clavier c'est très simple à corriger alors qu'avec les autres, je trouve que c'est moins pratique

Sujet 11

Si j'arrive pas avec Rb je passe au CV

Sujet 2

Change jamais de SI

Sujet 25

Normalement je sélectionne avec le stylet l'erreur et je corrige la faute, ou bien j'efface de droite vers la gauche le mot entier pour reprendre du début du mot.

Autres questions :

8. EXP: Est-ce que après ce test vous envisagez d'apprendre ou d'utiliser en futur un autre médium ?  
Si oui, pourquoi ?

Sujet 4

TR il est intéressant, de plus TR est pas mal pour de longs textes, à condition qu'il ne se trompe pas à interpréter mon écriture

9. EXP: Une fois acheté le PocketPC, comment avez vous découvert les médiums et comment vous les avez appris ? Avec quels critères vous avez décidé d'utiliser de préférence le ... clavier ... et la reconnaissance des blocs en écartant les autres ?

Sujet 1

Avec Palm j'étais d'abord très Graffiti, pour finir je suis passé au clavier (Fitaly)  
Sur PPC il a toute suite testé le TR pour voir la performance, vite déçu  
RB : moins précis que Graffiti par rapport à Graffiti sur Palm. Il y a une petite différence. Palm arrive à mieux reconnaître les petites erreurs  
RL : j'ai jamais été beaucoup attiré par ça

Sujet 3

CV : parce qu'il n'y pas d'apprentissage à faire  
RL : simple, pas long apprentissage

But ; utiliser l'outil toute suite

Sujet 5

Il sait pas pourquoi il a laissé tomber la RB en faveur du CV.

Sujet 11

RL et TR : connaissait déjà RB et CV, par conséquent il a voulu voir ce qu'ils offraient. Mais RB plus performant que les deux nouveaux.

Sujet 2

Au début CV, car pas le temps et envie d'apprendre de nouvelles écritures. Après il a essayé le RB sur conseil d'un ami et il a trouvé plus sympa. Les 2 autres il ne les aimait pas. Mais peut être qu'il va voir un peu de plus près RL. TR n'aime pa du tout.

Critiques

Sujet 3

CV : mettre tous les signes de ponctuation au début, sur l'écran principal

RL ou Transcriber: meilleure reconnaissance de l'écriture

RB :

TR : intéressant d'écrire sur tout l'écran mais il faut soigner l'écriture et faire un effort pour que le système la reconnaisse. On n'est pas dans la possibilité de tenir le machin pour écrire correctement etc. Après c'est de la chance on espère que ça tienne. Il faudrait un outil qui permette de définir mon écriture. Mais ça c'est un problème car ça va prendre du temps et moi je suis quelqu'un qui face à une machine, je veux utiliser toute suite les choses sans prise de tête