UNIVERSITE DE TECHNOLOGIE DE COMPIEGNE

ECOLE DOCTORALE « SCIENCES POUR L'INGENIEUR »

Thèse de Psychologie

Présentée pour obtenir le grade de Docteur de l'UTC

en spécialité

« Technologies Cognitives, Management de l'Innovation et Systèmes Complexes »

Loïc DESCHAMPS

- Année 2013 -

SUPPLEANCE PERCEPTIVE ET COGNITION SOCIALE

Etude des interactions tactiles minimalistes

Sous la co-direction de Charles Lenay et Katia Rovira

Soutenue publiquement le 07 Mai 2013

Composition du Jury :

René PRY - *Rapporteur* - Professeur des Universités - Université Paul Valéry (Montpellier 3)

Hélène TREMBLAY - *Rapporteur* - Professeur des Universités - Université de Rouen

Nadine VIGOUROUX - *Examinateur* - Chargée de Recherche au CNRS-IRIT - Toulouse

Olivier GAPENNE - *Examinateur*- Professeur des Universités – Université de Technologie de Compiègne

Charles LENAY - *Directeur* - Professeur des Universités - Université de Technologie de Compiègne

Katia ROVIRA - *Directeur* - Maître de Conférences / HDR - Université de Rouen

REMERCIEMENTS

Cette thèse est le fruit d'un long travail qui n'aurait pu voir le jour sans l'intervention, le soutien et la bienveillance de nombreuses personnes, qui m'ont accompagné tout au long de mon doctorat, et qui ont contribué à la réalisation de ce mémoire. C'est avec beaucoup de reconnaissance que je me tourne vers elles, que je leur adresse ces quelques mots, et que je les remercie infiniment.

Tout d'abord, je remercie chaleureusement René Pry et Hélène Tremblay, pour avoir accepté d'être les rapporteurs de mon travail et d'en évaluer la première version, ainsi que Nadine Vigouroux et Olivier Gapenne, qui ont accepté d'en être les examinateurs. J'espère que la soutenance de mon travail fut aussi agréable pour vous qu'elle ne l'a été pour moi.

Mes plus vifs remerciements reviennent à mes deux directeurs de thèse, Charles Lenay et Katia Rovira, qui ont encadré et soutenu mon travail, m'ont transmis leur passion de la recherche et m'ont ouvert de nombreuses portes pendant ces quelques années. C'est quand je regarde en arrière que je constate à quel point leur rencontre dépasse de loin le cadre de ce doctorat. Vous m'avez permis un épanouissement personnel, de par la confiance et la liberté que vous m'avez accordées. Vous m'avez toujours poussé vers l'avant, malgré mes innombrables remises en cause. Ce qui est là, entre vos mains, ce n'est qu'une partie de ce que je vous dois, et j'espère pouvoir un jour justifier tout ce que vous m'avez donné.

Je remercie également le laboratoire COSTECH de l'Université de Technologie de Compiègne, et en particulier l'équipe CRED, qui a su m'accueillir et me donner l'opportunité de réaliser cette recherche dans les meilleures conditions. La pluridisciplinarité, qui fait la richesse de cette équipe en particulier et de ce laboratoire en général, m'a permis des réflexions profondes, tant sur le plan épistémologique que sur le plan méthodologique, m'offrant ainsi l'inspiration et la motivation qui m'ont été nécessaires pour aller jusqu'au bout. Mais au-delà des compétences et des connaissances que j'ai pu y trouver, ce laboratoire a été avant tout le lieu de nombreuses rencontres, qui ont contribué étroitement à faire de cette aventure un plaisir. Je remercie chaque membre pour avoir fait émerger cette atmosphère amicale et chaleureuse.

Si je ne peux pas citer tout le monde, je tiens à remercier tout particulièrement Dominique Aubert pour ses nombreux développements techniques, dépannages et explications, mais aussi pour la disponibilité, l'écoute et l'amitié qu'il m'a témoignées et qui ont été déterminantes dans la réalisation de mon travail. J'espère que tu trouveras un nouveau révélateur de bug, sans quoi tout risque maintenant de fonctionner comme sur des roulettes.

De la même façon, je remercie particulièrement Alexandre Lang, pour ses conseils et ses encouragements, pour ses leçons de billard et pour tous les moments sympathiques que nous avons partagé, ainsi que Jocelyne Morel et Christine Devaux pour leur présence amicale et pour avoir notamment facilité les diverses démarches administratives qui incombent au doctorant.

Je remercie également les membres de l'équipe FIACRE du laboratoire PSY.NCA de l'Université de Rouen pour leur accueil et pour m'avoir accompagné dans la dernière ligne droite de mon travail. Je tiens à remercier particulièrement Emilie Chanoni, pour ses conseils précieux concernant la préparation de ma soutenance.

Pour leur aide inestimable dans l'élaboration et la réalisation des expérimentations, je remercie John Stewart, Gabrielle Le Bihan, Camille Barot, Cécile Laborde, Pierre-Alexandre Kofron et Caroline Laroche, ainsi que tous les sujets qui ont accordé de leur temps pour des passations parfois longues. Je remercie vivement Jérémy Terrien, pour avoir « mis les mains dans le programme » et participé aux analyses de trajectoires, ainsi que Régis Thouvarecq et Vincent Roy pour leurs conseils précieux en termes d'analyse statistique.

De plus, je remercie chaleureusement les personnes sans qui ce travail n'aurait pas de sens. Vous m'avez témoigné un enthousiasme, un engagement et une motivation qui ont stimulé ma recherche, et qui a fait avancer la conception de nos dispositifs. Parmi eux, je remercie particulièrement Christopher, Yamine, Corentin, Yves, Sabine, Sana, Linda, Laurence et Brigitte, ainsi que tous les collégiens et personnels du centre Normandie-Lorraine et les membres de l'association APICADEV. Dans le même temps, je remercie les membres des projets PORTINTERTACT et ITOIP, qui m'ont offert un cadre de travail passionnant et enrichissant.

J'adresse également un énorme merci à mes parents, mon frère, ma sœur, sans oublier vos moitiés et vos quarts, ainsi qu'à mes amis, et en particulier les malous, dont la liste pourrait dépasser l'entendement. Vous vous êtes toujours montrés présents et votre amour, l'attention constante et l'appui que vous m'avez offerts m'ont donné la confiance et les ressources pour en arriver là. Je ne serai pas grand-chose sans vous et sans votre compréhension, malgré la distance qui s'est parfois immiscée. J'ai hâte de vous présenter une sérénité retrouvée.

Enfin, je ne pourrai jamais assez remercier celle qui partage ma vie et qui fait le bonheur de mon quotidien. Je m'excuse du fond du cœur pour les moments les plus pénibles, pour t'avoir fait absorber mes angoisses et mes souffrances dans les derniers moments de cette thèse. Je ne peux pas dire à quel point tu as transformé ma vie, car il faut le vivre, et je préfère qu'on le partage jalousement tout au long de notre avenir.

RESUME

Cette recherche s'inscrit dans le cadre de l'élaboration d'espaces numériques d'interaction tactile, en tant qu'ils sont rendus possibles par la connexion en réseau d'un dispositif de suppléance perceptive. Dans ce contexte, nous avons articulé une recherche appliquée, centrée sur les analyses d'usage du dispositif dans des contextes écologiques, à une recherche fondamentale, dirigée par les questions théoriques soulevées par les usages eux-mêmes. Pour cela, la méthodologie minimaliste nous offre une occasion originale d'étudier la constitution même des rencontres interpersonnelles, à travers un espace de couplage sensori-moteur inédit.

Nos résultats généraux suggèrent que le processus de l'interaction est une dynamique relationnelle autonome qui émerge de l'engagement mutuel des participants. Dans une approche énactive et interactionniste de la cognition sociale, il s'agit alors de considérer que le croisement perceptif entre deux sujets, même réduit à son expression la plus simple, consiste en la rencontre de deux activités perceptives construisant du sens. Dans le contexte de rencontres strictement dyadiques, nous tentons de caractériser cette dynamique, de façon à en isoler les éléments constitutifs. Nous montrons alors que la coordination interpersonnelle résulte d'un processus actif de co-ajustements dynamiques qui se joue autant à un niveau microscopique (qualité de l'accroche perceptive) qu'à un niveau macroscopique (organisation de séquences d'interaction). Dans le contexte de l'exploration mutuelle d'un contenu numérique, nous montrons que cette dynamique d'interaction permet aux participants de faire sens de leurs engagements respectifs en fonction des objets présents. La coordination des activités perceptives se présente alors comme le support de la co-constitution d'un monde partagé de significations, à travers une compréhension interpersonnelle ancrée dans un contexte pragmatique.

En outre, l'articulation des expérimentations et des analyses d'usage écologiques nous a conduits à proposer des spécifications techniques et fonctionnelles pour les espaces numériques partagés, de façon à proposer un dispositif pertinent pour les utilisateurs déficients visuels.

Mots-clés : Suppléance perceptive ; Interactions tactiles ; Minimalisme ; Analyse d'usage ; Cognition sociale

ABSTRACT

Our research is part of the design of tactile interaction digital spaces, as they are made possible by a network connection between perceptual supplementation devices. Within this framework, we articulated an applied research, focused on the analysis of use of the device in ecological contexts, with a fundamental one, led by theoretical issues raised by the usage itself. In this respect, minimalist methodology gives us an opportunity to study the very constitution of interpersonal encounters, through a new space of sensorimotor coupling for users.

Our overall results suggest that interaction process is an autonomous relational dynamics that emerges from the mutual engagement of participants. Within an enactive and interactionist approach of social cognition, we consider that this perceptual crossing, even when reduced to its simplest expression, involves the meeting of two perceptual activities which inherently produce sense. In the context of strictly dyadic encounters, we try to characterize this dynamic, so as to isolate its fundamental features. We show that interpersonal coordination results from an active process of dynamic co-adjustments that unfolds both on a microscopic level (quality of the perceptual coupling) and on a macroscopic level (organization of interaction sequences). In the case of mutual exploration of digital content, we show that this dynamic interaction allows participants to make sense of their respective engagement relatively to the present objects. The coordination of perceptual activities is then presented as a support for the co-constitution of a shared world of meanings, from an interpersonal understanding rooted in a pragmatic context.

In addition, these experiments, coupled to ecological usage analysis, have led us to propose technical and functional specifications for shared digital spaces, so as to provide a relevant device for visually impaired users.

Keywords: Perceptual supplementation; Tactile interactions; Minimalism; Usage analysis; Social cognition

SOMMAIRE

INTRODUCTION 15

PARTIE 1 : SUPPLEANCE PERCEPTIVE ET INTERACTION 17

1. La suppléance perceptive 17

1.1. Présentation générale 17

1.1.1. De la substitution sensorielle… 17

1.1.2. …à la suppléance perceptive 20

1.2. Le dispositif Tactos 21

1.2.1. Présentation générale et objectifs 21

1.2.2. Principe d'utilisation 24

1.2.3. Parallélisme des champs récepteurs 25

1.3. Analyses d'usage du dispositif Tactos 27

1.3.1. Objectifs 27

1.3.2. Méthodologie 29

1.3.3. Observations et analyses d'usage 31

*1.3.3.1. Bilan du suivi de **C.**(utilisateur expert) 31*

*1.3.3.2. Bilan du suivi de **S.**(utilisateur novice) 33*

*1.3.3.3. Bilan du suivi de **Y.** (utilisateur novice) 36*

1.4. Le problème de l'usage individuel 37

2. Vers des espaces d'interaction tactile 41

2.1. Etude de faisabilité via des analyses d'usage préliminaires 41

2.1.1. Effecteurs 43

2.1.2. Supports pédagogiques numériques 48

2.1.3. Emoticônes tactiles 53

2.1.4. Conclusion et perspectives de l'étude de faisabilité 55

2.2. Le dispositif Intertact ou Tactos-Réseau 56

2.2.1. Présentation générale et objectifs 56

2.2.2. Prise en main du module d'interaction tactile 61

2.2.3. Analyses d'usage du module d'interaction tactile 65
2.2.4. Conclusion de l'analyse d'usage du module d'interaction tactile 67
3. Etudes préliminaires concernant les interactions tactiles 68
3.1. Observations préliminaires 68
3.1.1. Configurations symétriques 70
3.1.2. Configurations asymétriques 76
3.1.3. Synthèse des observations préliminaires 76
3.2. Les rencontres interpersonnelles : une problématique d'usage 77
3.3. Discrimination interpersonnelle (1) 81
3.3.1. Population 82
3.3.2. Matériel 82
3.3.3. Procédure 84
3.3.4. Résultats 87
3.3.5. Discussion 89
3.4. Discrimination interpersonnelle (2) 90
3.4.1. Population 91
3.4.2. Matériel 92
3.4.3. Procédure 93
3.4.4. Résultats 96
3.4.5. Analyse des trajectoires 101
3.4.6. Discussion 106
4. Vers des questionnements théoriques 109
4.1. L'interaction sociale comme objet d'étude 109
4.2. Le minimalisme comme cadre d'analyse 111
PARTIE 2 : COGNITION SOCIALE ET INTERACTION 115
1. Des positions classiques 115

1.1. La théorie de la théorie (TT) 115
1.2. La théorie de la simulation (ST) 117
1.3. Une détection de contingences 118
1.4. Vers une alternative : la théorie de l'interaction 121
2. La compréhension sociale en l'absence de vision 123
2.1. Les conséquences de la cécité sur l'interaction sociale 124
2.2. Des modes d'interaction alternatifs 126
2.3. L'interaction comme support de la compréhension sociale 128
3. L'interaction comme engagement mutuel 129
3.1. Une critique fondamentale des positions classiques 129
3.2. L'ancrage de la compréhension sociale dans l'action 131
3.3. L'interaction comme "*participatory sense-making*" 132
4. Croisement perceptif et dynamique collective 134
5. Etude expérimentale : test de Turing minimaliste non-verbal 139
5.1. Population 140
5.2. Matériel 141
5.3. Procédure 143
5.4. Hypothèses opérationnelles 145
5.5. Résultats 145
5.5.1. Type d'interactant 145
5.5.2. Apprentissage 147
5.5.3. Type de robots 148
5.5.4. Reconnaissance mutuelle 149
5.6. Analyses de trajectoires 152
5.6.1. Proportion d'accroche mutuelle 152
5.6.2. Variabilité de l'accroche mutuelle 155
5.6.3. Désengagements de l'accroche mutuelle 158
5.7. Discussion 161
6. Retour sur les usages : le cas des interactions dyadiques 168

6.1. Scénario 1 : rencontre et suivi dans un espace vide 169
6.1.1. Configuration « Mono » 170
6.1.2. Parallélisme de type I 172
6.1.3. Parallélisme de type II 174
6.2. Scénario 2 : rencontre et suivi finalisé 176
6.2.1. Configuration « Mono » 176
6.2.2. Parallélisme de type I 177
6.2.3. Parallélisme de type II 179
6.3. Synthèse des observations contrôlées 180

PARTIE 3 : LA PERCEPTION MUTUELLE 183

1. Introduction 183
2. Position du problème 184
3. L'attention conjointe comme coordination perceptive 187
4. Le rôle de la dynamique du mouvement 191
5. L'attention conjointe sans la vision ? 194
6. L'attention conjointe comme pratique sociale incarnée 199
7. Etude expérimentale : objets communs et objets privés 203
7.1. Population 203
7.2. Matériel 203
7.3. Procédure 204
7.4. Hypothèses opérationnelles 207
7.5. Résultats 208
7.5.1. Analyse des réponses individuelles 208
7.5.2. Analyse des réponses collectives 212
7.6. Analyse des trajectoires 217
7.6.1. Interaction des sujets avec les deux types d'objets 217
7.6.2. Interaction entre les sujets autour des deux types d'objets 222
7.6.3. Interaction entre les sujets en fonction de la qualité des clics 228

7.7. Discussion 232

8. Retour sur les usages : exploration mutuelle d'un contenu partagé 240

8.1. Scénario 1 : exploration mutuelle simple 241

8.1.1. Parallélisme de type I 241

8.1.2. Parallélisme de type II 242

8.2. Scénario 2 : exploration mutuelle complexe 243

8.2.1. Parallélisme de type I 245

8.2.2. Parallélisme de type II 246

8.3. Synthèse des observations contrôlées 248

SYNTHESE GENERALE 251

CONCLUSION ET PERSPECTIVES 259

BIBLIOGRAPHIE 263

ANNEXES 285

INTRODUCTION

Notre travail se place dans le cadre de l'élaboration d'un dispositif de suppléance perceptive (Tactos). Ce dispositif, destiné en premier lieu à des utilisateurs déficients visuels, permet d'accéder aux informations graphiques d'un écran d'ordinateur par l'intermédiaire de la modalité tactile. Il a donc vocation à permettre une accessibilité au monde numérique, particulièrement dans ses dimensions spatiales, de façon à favoriser l'intégration et l'autonomie des utilisateurs ciblés. Dans la première partie de ce travail, nous commencerons par décrire en détails ce dispositif et son principe de fonctionnement. Nous présenterons également la méthodologie de conception, centrée sur la mobilisation des utilisateurs dans des suivis longitudinaux et des analyses d'usage en contexte écologique. Il s'agit alors d'étudier l'appropriation d'un nouveau mode de couplage sensori-moteur par l'utilisateur, tout en encourageant l'amélioration du dispositif par l'intermédiaire de spécifications matérielles et fonctionnelles.

Mais au-delà de son usage individuel, ce dispositif original peut également être connecté en réseau (Intertact), et ainsi permettre l'interaction tactile d'utilisateurs distants. Nous pensons qu'il s'agit là d'une opportunité intéressante pour inscrire ce type de dispositif dans une pratique partagée et constituer une communauté d'usage. Nous présenterons alors l'étude de faisabilité d'un tel dispositif, qui a permis d'aboutir à l'élaboration de premiers prototypes fonctionnels et à la définition de spécifications pour la conception d'espaces numériques partagés. Ceux-ci ont alors donné lieu aux premières observations d'interactions tactiles distales, de façon à poursuivre un processus d'évaluation itératif et correctif basé sur des analyses d'usage écologiques. Parallèlement, nous avons entrepris une première étude expérimentale visant à explorer les conditions de possibilité de discrimination interpersonnelle dans des espaces numériques partagés. Si cette expérience est d'abord basée sur une problématique d'usage, les résultats soulèvent des questions théoriques fondamentales. Les utilisateurs semblent en effet capables de s'identifier les uns des autres en vertu de l'émergence d'une dynamique d'interaction spécifique. Notre dispositif offre alors une occasion originale d'étudier la constitution même de cette dynamique, à travers l'engagement des utilisateurs dans un mode de couplage interpersonnel inédit.

Dans la deuxième partie de ce travail, nous présenterons une approche énactive, incarnée et située de la cognition sociale, qui permet de placer le processus même de l'interaction au cœur de la compréhension interpersonnelle. Nous questionnerons ainsi le

statut même de l'interaction, classiquement entendue en tant que facteur contextuel traité individuellement par des mécanismes internes, en y opposant une approche pleinement interactionniste, considérant la dynamique d'interaction elle-même comme constitutive de la cognition sociale. Notre objectif consistera alors à tenter de caractériser cette dynamique dans un contexte strictement dyadique. Nous proposerons pour cela un test de Turing non verbal, dans lesquels des sujets devront interagir tactilement avec d'autres sujets et avec des robots censés imiter le comportement humain. Parallèlement, il sera question de mettre à l'épreuve la dynamique d'interaction dans des contextes d'usage plus écologiques. Nous présenterons alors des analyses d'usage mettant en jeu deux utilisateurs dans des situations de guidage interpersonnel.

Dans la troisième partie de ce travail, nous nous intéresserons à la perception mutuelle d'objets communs. Cette question est essentielle, dans la mesure où les espaces numériques partagés impliquent non seulement des rencontres interpersonnelles, mais aussi une coordination des activités perceptives pour l'exploration de contenus numériques partagés. Il sera alors question de comprendre comment s'organisent les activités perceptives pour que les utilisateurs puissent savoir qu'ils partagent un monde commun. Nous nous référerons alors essentiellement au cadre théorique de l'attention conjointe, en proposant une approche sensori-motrice destinée à replacer cette capacité dans le cadre d'une pratique sociale incarnée. Nous présenterons alors une étude expérimentale visant à montrer que cette connaissance émerge de la dynamique des activités perceptives. Enfin, dans l'optique de confronter cette hypothèse à l'épreuve du concret, nous exposerons des analyses d'usage dans le cadre de l'exploration mutuelle de contenus partagés.

PARTIE 1 : SUPPLEANCE PERCEPTIVE ET INTERACTION

1. La suppléance perceptive

1.1. Présentation générale

1.1.1. De la substitution sensorielle…

Les technologies d'assistance offrent de plus en plus de solutions pour les personnes en situation de handicap, en favorisant notamment l'accessibilité à l'information, l'autonomie des utilisateurs et leur insertion pédagogique ou professionnelle. Dans le cas de la déficience sensorielle, et en particulier dans celui des déficiences visuelles, des dispositifs techniques spécifiques ont vu le jour dès la fin des années 60. Il s'agit des dispositifs dits de substitution sensorielle (pour des exemples, voir Auvray, 2004 ou Ali Ammar, 2005). Pour des raisons que nous détaillerons plus tard, nous préférons appeler ces systèmes techniques des dispositifs de suppléance perceptive (Lenay, Gapenne, Hanneton, Marque & Genouëlle, 2000). Parmi eux, le Tactile-Vision Substitution System (TVSS) est considéré comme le pionnier en la matière (Bach-Y-Rita, Collins, Saunders, White & Scadden, 1969 ; Bach-Y-Rita, 1972).

L'objectif général de ces dispositifs dits de "substitution sensorielle", est de rendre accessible des informations provenant d'une modalité sensorielle donnée par l'intermédiaire d'une autre modalité sensorielle. Par exemple, des informations visuelles sont converties en informations tactiles. Pour cela, le principe est de mobiliser conjointement trois éléments principaux : des capteurs assurant l'entrée de l'information, des stimulateurs en sortie, et un système de couplage qui assure la liaison technique des deux premiers éléments cités (Lenay et al., 2000).

Ainsi, dans sa première version, le TVSS est composé d'une caméra fixée sur un trépied, d'un convertisseur analogique pour constituer une « image tactile » à partir des informations visuelles captées par la caméra, et d'une matrice de 20 x 20 stimulateurs électro-tactiles implémentée dans le dossier d'un fauteuil (Bach-Y-Rita & al., 1969) (voir Figure 1).

Avec ce dispositif, des utilisateurs non-voyants ou voyants aveuglés sont capables très rapidement de détecter des cibles simples et de s'orienter vers elles (Bach-Y-Rita, 1972). De même, ils peuvent discriminer des lignes présentant différentes orientations dans l'espace, ainsi que détecter la direction du mouvement de cibles mobiles (ibid.). Seulement, très vite, des limites se posent. Avec la caméra immobile, les utilisateurs font preuve de grandes

difficultés dans la discrimination de formes géométriques simples ou d'objets usuels présentés selon différents angles à la caméra. Ainsi, une cinquantaine d'essais est nécessaire pour obtenir 100% de reconnaissance à la présentation de formes géométriques simples. Dans le cas de la reconnaissance d'objets usuels, l'apprentissage est encore plus conséquent, et nécessite environ une dizaine d'heures de pratique. Pour les utilisateurs, l'attention reste alors fixée sur les stimuli perçus à la surface de la peau (ibid.).

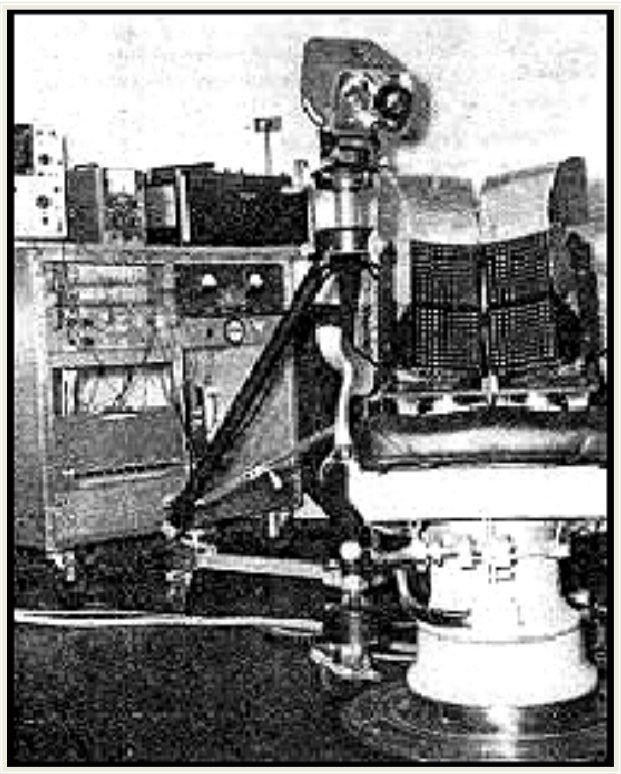


Figure 1 : Le Tactile-Vision Substitution System (TVSS) développé par Bach-Y-Rita à la fin des années 60

Dans les versions suivantes, le dispositif a subi des améliorations techniques notables. La matrice de stimulateurs a été notamment miniaturisée, et a quitté le dossier du fauteuil pour être appliquée sur le thorax, l'abdomen ou sur le front de l'utilisateur. De même, la caméra est maintenant manipulable plus facilement, et est soit tenue en main, soit implémentée dans des lunettes que porte l'utilisateur (Collins & Bach-Y-Rita, 1973) (voir Figure 2).

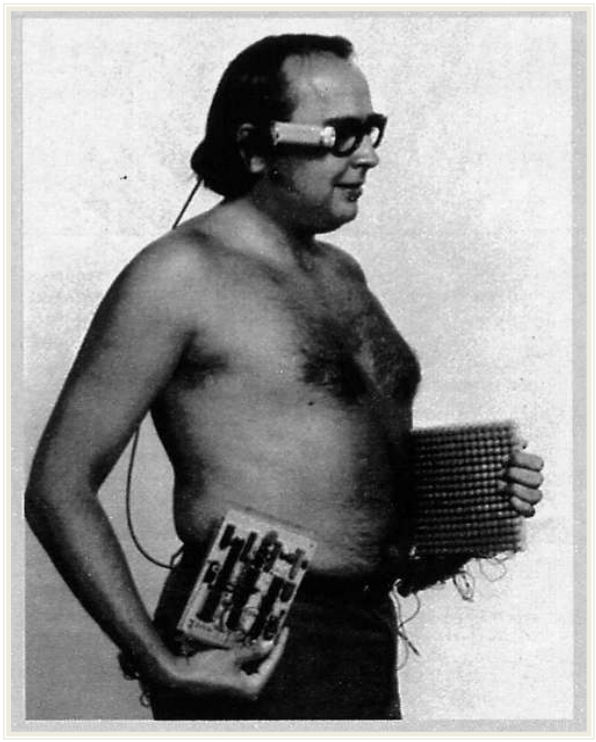


Figure 2 : Exemple d'un TVSS miniaturisé. La matrice de stimulateurs est appliquée sur l'abdomen, tandis que la caméra est fixée sur les lunettes de l'utilisateur

Lorsque l'utilisateur peut manipuler activement la caméra, la reconnaissance de formes est spectaculaire. Les stimulations tactiles ne sont plus perçues sur la peau, à l'endroit où est appliquée la matrice de stimulateurs, mais sont maintenant attribuées directement à l'objet perçu, situé dans l'espace perceptif de l'utilisateur (Bach-Y-Rita, 1982).

Avec le système TheVoice, un dispositif similaire basé sur une transformation des informations visuelles en informations auditives (Meijer, 1992), il a été argumenté que cette capacité d'appropriation active du dispositif de substitution sensorielle passe par cinq étapes distinctes (Auvray, 2004) :

i) l'utilisateur apprend à détecter les régularités entre ses actions et les retours sensoriels qui en résultent. De cette façon, l'utilisateur est capable de mettre en place un contact perceptif avec l'objet, de le stabiliser et de le maintenir dans le temps.

ii) à travers l'exercice de ce contact perceptif, l'utilisateur comprend que les stimulations sont dues à la rencontre d'un objet situé dans l'espace distal. De cette façon, il y a extériorisation de la source de stimulation.

iii) l'utilisateur apprend à faire varier ses « points de vue » en fonction de la variation des stimulations qu'il reçoit. En d'autres termes, le sujet maîtrise progressivement l'espace distal, même si la localisation des objets reste égocentrée.

iv) l'expérience de l'espace est fiabilisée via une automatisation du couplage entre les actions du sujet et les stimulations tactiles qui en résultent. La localisation de l'objet n'est plus dépendante du corps lui-même, mais s'inscrit dans un espace perceptif distal.
v) cette maîtrise du couplage donne lieu à la constitution d'une expérience distale, à travers laquelle une immersion dans l'espace perceptif est rendue possible (ibid.).

Cette immersion est effective avec l'utilisation du TVSS, comme en témoigne une observation de Bach-Y-Rita qui, en manipulant le zoom de la caméra à l'insu de l'utilisateur, provoque un comportement d'évitement caractéristique de la part du sujet (Bach-Y-Rita, 1972). Tout se passe comme si le sujet, immergé dans l'espace perceptif, percevait l'expansion de l'image tactile comme un mouvement de l'objet lui-même vers lui.

A travers les étapes d'appropriation décrites ci-dessus, nous voyons bien que la mise en extériorité des percepts et la reconnaissance des formes qui en découle n'est rendue possible qu'au travers d'une exploration active de l'espace distal par l'utilisateur. En d'autres termes, il n'y a pas de perception sans action. La possibilité de faire varier les entrées sensorielles en fonction des activités exploratoires permet d'établir un couplage sensori-moteur minimal (Lenay, Canu & Villon, 1997), via la stabilisation d'une loi de contingence sensori-motrice (O'Regan & Noé, 2001). L'action est alors constitutive de la perception, et l'entrée sensorielle seule ne suffit pas.

1.1.2. …à la suppléance perceptive

La nécessité de considérer l'activité perceptive de l'utilisateur vient alors ébranler le terme même de « substitution sensorielle ». En effet, quand la vision fait défaut, il ne suffit pas de modifier la nature de l'entrée sensorielle pour permettre une immersion dans une expérience perceptive visuelle. Ce qui est perçu par l'utilisateur « n'est pas issu des invariants de la sensation », qu'il s'agirait alors de présenter via un autre canal, mais est issu « des invariants de cercles sensori-moteurs inséparables de l'activité du sujet » (Lenay et al., 2000, p. 296). Ainsi, s'il y a substitution, celle-ci est sensori-motrice, la perception ne se réalisant qu'au travers du couplage entre les actions et les sensations (ibid.).

Mais il ne s'agit pas non plus d'une « substitution », car l'expérience perceptive qui résulte de l'usage du dispositif n'est en rien une expérience visuelle à proprement parler. Les utilisateurs non-voyants eux-mêmes sont particulièrement déçus de ce couplage artificiel, qui leur promettait pourtant l'accès au monde des voyants. Ici, malgré l'efficacité relative du dispositif, il n'y a pas de couleurs, les percepts sont grossiers, les possibilités d'action sont

limitées et la reconnaissance de formes est lente et fastidieuse. Nous reviendrons plus tard sur la différence essentielle entre ces deux types d'expériences perceptives. Ainsi, les utilisateurs ne remplacent pas la modalité sensorielle manquante par une autre modalité existante. Au contraire, comme nous l'avons vu, ils sont immergés dans un espace perceptif qui est ouvert par le dispositif lui-même (Auvray, 2004 ; Khatchatourov & Auvray, 2005). Plutôt que d'une substitution, il s'agit alors d'établir et de consolider une nouvelle forme de couplage sensori-moteur avec le monde (Lenay et al., 2000). Nous préférons alors parler de « suppléance perceptive », de façon à souligner le fait que ces dispositifs ne viennent pas simplement combler un manque, mais qu'ils ouvrent au contraire une modalité perceptive originale, reposant sur des lois de couplage sensori-moteur inédites (ibid.).

Si de tels dispositifs sont conçus pour l'accessibilité de l'espace physique, d'autres dispositifs, comme celui que nous développons à l'Université de Technologie de Compiègne depuis plus de quinze ans, sont spécialement destinés à l'accessibilité des espaces numériques pour les personnes en situation de handicap visuel. L'accessibilité numérique représente en effet aujourd'hui un enjeu crucial pour l'insertion et l'autonomie pour tous. Mais au-delà de l'accès à l'information, déjà assurée en partie par les technologies existantes, l'accès à la spatialisation de cette information fait encore défaut.

1.2. Le dispositif Tactos

1.2.1. Présentation générale et objectifs

Classiquement, les logiciels de synthèse vocale ou les barrettes de cellules brailles électroniques permettent de donner aux personnes aveugles un accès aux informations numériques. Dans les deux cas, les informations textuelles présentes à l'écran sont délivrées à l'utilisateur, via la modalité auditive ou via la modalité tactile, ce qui lui permet d'utiliser les dispositifs informatiques. De plus, si les métadonnées le permettent (les info-bulles dans le cas de la navigation dans les interfaces graphiques des systèmes d'exploitation informatiques ou les rubriques html dans le cas de la navigation sur les pages web), il est possible d'accéder à une description textuelle des images. Cependant, ces systèmes ne peuvent délivrer qu'une information strictement temporelle, linéaire, suivant la succession des mots du texte ou de la description de l'écran, dans le sens de la lecture. Si l'information délivrée peut être plus ou moins exhaustive, elle ne permet pas de comprendre l'organisation spatiale de l'information : les positions des icônes et des fenêtres, les graphes et les tableaux de données, les schémas et les cartes. Pourtant, toutes ces informations bidimensionnelles font la richesse des interfaces

graphiques en spatialisant l'information. Par ailleurs, outre le fait que les interfaces numériques sont de plus en plus imagées, les images sont utiles pour les aveugles (Heller, 2000 ; 2002), ces dernières pouvant par exemple fournir des informations spatiales sous formes de cartes ou encore servir de support aux formalisations mathématiques.

Le système Tactos (correspondant à un logiciel dédié couplé à un dispositif tactile) a été développé pour rendre accessible aux aveugles l'organisation spatiale des informations présentes sur l'écran de l'ordinateur (Hanneton, Gapenne, Genouëlle, Lenay & Marque, 1999 ; Sribunruangrit, Marque, Lenay, Gapenne & Vanhoutte, 2002 ; 2004). Ce système consiste essentiellement en un dispositif de commande de stimulateurs tactiles (cellules Braille piézoélectriques générant l'activation de picots mécaniques) en fonction des déplacements du curseur sur l'écran d'un ordinateur. Un effecteur manipulé par l'utilisateur commande les déplacements d'un curseur qui correspond à un champ récepteur élémentaire (voir Figure 3). Quand ce champ récepteur croise au moins un pixel noir à l'écran, cela déclenche l'activation en tout ou rien d'un stimulateur tactile (en l'occurrence, un picot d'une cellule Braille piézoélectrique). Le sujet (non-voyant, malvoyant ou voyant avec les yeux bandés), tient l'effecteur avec sa main dominante, le stimulateur tactile étant installé sous sa main libre. L'effecteur constitue le moyen d'action de l'utilisateur dans l'espace bidimensionnel de l'écran. Dans la version classique du dispositif Tactos, nous utilisons un stylet informatique appliqué sur une tablette graphique[1]. D'autres effecteurs peuvent être utilisés, à condition que la surface d'action entretienne un rapport bijectif avec l'espace de l'écran : l'utilisateur doit pouvoir commander une position *absolue* de son curseur à l'écran, la position de son effecteur sur la surface d'action devant correspondre à la position du curseur à l'écran. Ainsi, l'utilisateur accède à une connaissance intuitive de l'espace en mobilisant des informations proprioceptives directement liées à son geste d'exploration (Auvray, Lenay, O'Regan & Lefèvre, 2005). De cette façon, les effecteurs *relatifs*, qui commandent un déplacement plus qu'une position (comme une souris informatique classique) ne sont pas préconisés, de par l'absence de corrélation entre les informations proprioceptives du geste d'exploration et la position exacte du curseur à l'écran (ibid.). Notons toutefois que dans le cas de certaines tâches, qui se focalisent sur la nature des mouvements, l'utilisation d'effecteurs relatifs peut s'avérer pertinente.

[1] Nous utilisons une tablette Wacom Intuos 2© au format A4

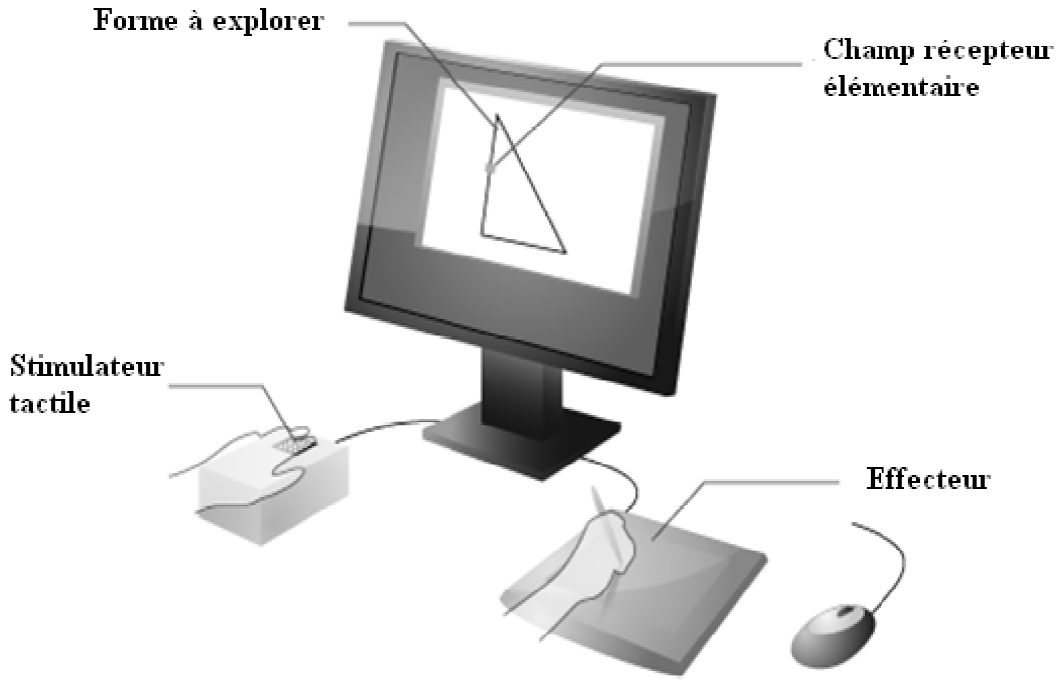


Figure 3 : le dispositif « Tactos ». Les formes inscrites dans l'espace numérique de l'écran sont perçues tactilement : lorsque le champ récepteur élémentaire croise au moins un pixel noir, un stimulateur tactile s'active en tout ou rien

Le dispositif de suppléance perceptive Tactos permet donc l'exploration d'une image tactile virtuelle. De nombreuses observations ont permis de souligner la capacité des sujets à localiser, identifier, reconnaître et catégoriser des formes plus ou moins complexes (Gapenne, Rovira, Ali Ammar & Lenay, 2003 ; Rovira & Gapenne, 2009). Ces objets ne sont pas donnés d'un coup au système sensoriel comme une forme qui serait appliquée sur la peau. Il n'y a qu'un seul champ récepteur élémentaire, donc une donnée sensorielle très pauvre correspondant à un seul bit d'information à chaque instant. Conformément aux résultats classiques du TVSS, si les sujets réussissent à reconnaître des formes dans un espace, ce ne peut être que par leur exploration active, en intégrant dans le temps leurs mouvements et les sensations tactiles résultantes (Bach-Y-Rita, 1972). Par cette activité perceptive, les sujets réalisent une forme de perception tactile des objets numériques. Ce type de perception est de même nature que la perception haptique "naturelle" puisque dans le cadre de travaux concernant la perception de l'espace, il a été montré que la compréhension des configurations spatiales s'appuie en priorité sur les informations kinesthésiques des trajets manuels (Millar, 1994). Le système Tactos donne ainsi accès à l'information spatialisée : lecture et production de dessins, de graphiques, de figures géométriques, lecture de l'interface graphique, différenciation des fenêtres actives ou non, etc. Nous allons maintenant détailler le principe

d'utilisation du dispositif, de façon à mieux comprendre la manière dont les utilisateurs parviennent à organiser leurs gestes d'exploration pour accéder à des percepts signifiants.

1.2.2. Principe d'utilisation

Tout comme le TVSS avant lui, le dispositif Tactos met en évidence le caractère constitutif de l'action dans la perception. En effet, par les possibilités d'action qui lui sont offertes, l'utilisateur peut déployer une exploration active de l'environnement numérique, et ainsi reconnaître les formes bidimensionnelles qui y sont inscrites (Hanneton et al., 1999). En d'autres termes, ce sont les possibilités d'action dans l'espace numérique, couplées aux retours tactiles, qui lui permettent de constituer de nouveaux percepts. La question est maintenant de savoir comment se constituent ces percepts, ou plutôt, de quelle manière doit se déployer l'exploration pour permettre la constitution de percepts qui ont du sens.

Grâce à l'enregistrement des trajectoires perceptives rendu possible par le dispositif numérique, nous pouvons observer de façon fine la manière dont les sujets s'y prennent pour réussir la tâche. Il apparaît que la constitution des percepts via la médiation technique fait appel à un répertoire sensori-moteur caractéristique (ibid.). En effet, l'analyse des stratégies d'exploration mises en œuvre par les sujets révèle que les actions de l'utilisateur s'organisent de façon dynamique en fonction des retours sensoriels. Trois dynamiques exploratoires ont pu être ainsi isolées.

Tout d'abord, le ***mouvement de balayage*** est une stratégie qui consiste à effectuer des oscillations dans l'espace perceptif. D'abord de grande amplitude (*macro-balayage*), le geste tend à décroître et à se stabiliser autour d'une source de stimulation localisée (*micro-balayage*), en l'occurrence lorsque le sujet « accroche » une figure présente dans l'espace. Avant toute stimulation, le sujet « erre » dans un espace qui n'existe pas : il recherche alors un point de stimulation qui va pouvoir fournir un premier repère spatial. Lorsqu'il rencontre une telle stimulation tactile, le sujet tente alors de la reproduire en effectuant le mouvement inverse. Petit à petit, le sujet à parvient à anticiper l'activation du retour sensoriel en fonction du mouvement accompli via l'effecteur du système (Lenay, 1997 ; Lenay et al., 1997). En d'autres termes, le point de stimulation local constitue un attracteur qui va guider l'activité et permettre de stabiliser le mouvement d'exploration via une anticipation dynamique du retour sensoriel. Notons que ce mouvement d'exploration se stabilise le plus souvent de façon à croiser périodiquement et de façon orthogonale le contour de la forme explorée. En fait, l'espace ne se constitue qu'à partir du moment où le sujet peut aller et venir autour d'une singularité, c'est-à-dire passer de l'absence de stimulation à la stimulation réversiblement et

de façon intentionnelle. De cette façon, nous pouvons décrire cette activité d'exploration comme une activité de constitution de l'espace (Lenay, 2005).

La deuxième stratégie consiste en un ***suivi de contour*** de la forme explorée. Ici, la stabilisation du micro-balayage orthogonal au contour permet d'« accrocher » la source de stimulation. Il s'agit maintenant pour le sujet de ne plus simplement anticiper la localisation de la source de stimulation dans l'espace, mais d'explorer cette source par un mouvement « longitudinal, tangentiel, suivant des directions élémentaires » (Hanneton et al., 1999, p. 4). Cette stratégie est caractérisée par la reconnaissance de traits constitutifs de la forme.

La troisième stratégie est caractérisée par une amélioration qualitative de la précision et des anticipations de la séquence gestuelle dynamique conduisant à la reconnaissance de la forme. En d'autres termes, l'exploration de la forme est comparable à une ***anticipation dynamique*** semblable au geste qui aurait produit cette même forme (ibid.). La dynamique gestuelle mise en œuvre par le sujet semble être déployée dans l'optique d'une vérification des stimulations préalablement anticipées, rendant comparable le geste d'exploration au geste d'écriture ou de production de la forme (Ploux, 1994, cité par Hanneton et al., 1999).

Cette analyse fine des trajectoires perceptives conduisant à une reconnaissance effective des formes explorées permet de caractériser l'activité même de l'exploration en termes de dynamiques sensori-motrices. Ces dynamiques sont essentielles à la constitution de l'espace et à la reconnaissance de formes, dans la mesure où elles permettent d'établir un couplage sensori-moteur nécessaire à la constitution de l'espace et à la construction de sens (Lenay, 2005 ; Lenay et al., 1997).

1.2.3. Parallélisme des champs récepteurs

Même si des capacités effectives de reconnaissance de formes sont observées avec la configuration la plus minimale du dispositif Tactos (soit un unique champ récepteur qui ne permet d'accéder qu'à un seul bit d'information à chaque instant), il a été question d'enrichir progressivement les possibilités d'action et/ou la richesse des stimulations tactiles. En effet, le dispositif Tactos étant destiné à l'exploration d'interfaces riches contenant des formes nombreuses et complexes, il est nécessaire d'enrichir l'accès à l'information pour garantir une utilisabilité optimale.

Par exemple, des études antérieures ont montré que l'augmentation du nombre de champs récepteurs utilisés conjointement permettait d'améliorer la perception de l'environnement (Mauris, Benoit & Foulloy, 1998). L'utilisation de ce parallélisme dans le cadre de dispositifs de perception médiatisée pourrait en effet permettre d'améliorer la qualité

et la richesse de la perception dans un espace numérique (Sribunruangrit, 2004). Pour vérifier cette hypothèse, une étude a porté sur l'effet de la variation des propriétés du dispositif de couplage (Sribunruangrit, 2001). Il a été question de faire varier la taille et le nombre de champs récepteurs, ainsi que le nombre de stimulateurs associés. Les meilleures performances ont été obtenues lorsque l'utilisateur déplace une matrice carrée de 16 champs récepteurs élémentaires (4x4 champs récepteurs de 2x2 pixels chacun), chacun de ces champs étant relié au picot correspondant d'une matrice de 16 stimulateurs tactiles (4x4 picots, soit deux cellules Braille piézoélectriques juxtaposées) (ibid.). De cette façon, le parallélisme entre les champs récepteurs et les stimulateurs tactiles permet de rendre accessible un « mouvement déjà fait », présent dès l'entrée sensorielle (voir Figure 4). Dans le cas de la vision, déjà, le parallélisme induit par la multiplicité des photorécepteurs de la rétine donne l'illusion d'une perception absolue et directe ; pourtant, « l'utilisation de la convergence binoculaire pour évaluer la distance d'un objet est formellement équivalente au changement d'orientation d'un unique œil à la suite d'un léger basculement de la tête » (Lenay, 2006a, p. 35-36). De ce fait, le parallélisme des entrées sensorielles produit une économie de mouvement et de mémoire qui permet d'augmenter la vitesse perceptive, améliorant sensiblement dans le cas de Tactos la qualité de l'exploration et la reconnaissance des formes inscrites à l'écran (Gapenne, Lenay, Stewart, Bériot & Meidine, 2001 ; Sribunruangrit, 2004).

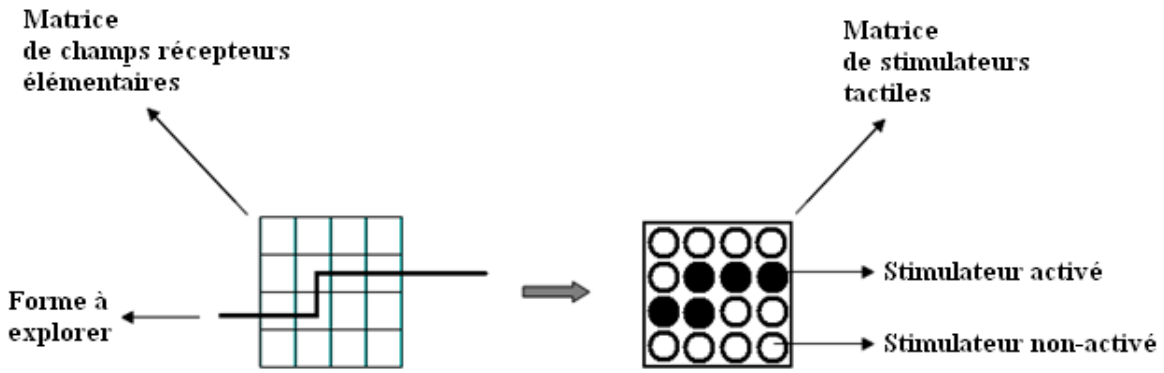


Figure 4 : Les 16 champs récepteurs élémentaires forment une matrice. Chaque champ récepteur active le picot correspondant sur la matrice de 16 stimulateurs tactiles. Il y a donc parallélisme entre les champs récepteurs et les stimulateurs tactiles.

Ainsi, les diverses situations expérimentales et les observations *in situ* ont montré que l'utilisateur de Tactos est capable de discriminer et de reconnaître les formes rencontrées dans l'espace numérique (Ali Ammar, 2005 ; Deschamps, Rovira, Lenay & Gapenne, 2009). De plus, le dispositif rend accessible la configuration spatiale des informations présentes à

l'écran, ce qui permet la localisation des objets numériques (icônes, images, fenêtres etc.) - et donc la navigation dans les interfaces graphiques -, ainsi que la lecture d'objets spatialisés tels que des tableaux ou des graphiques. De cette façon, couplé aux « screen-readers », le dispositif Tactos permet d'en dépasser la linéarité (Gapenne et al., 2003), et laisse envisager une accessibilité effective des environnements numériques pour les personnes déficientes visuelles.

1.3. Analyses d'usage du dispositif Tactos

1.3.1. Objectifs

Afin de mettre à l'épreuve le dispositif par les utilisateurs eux-mêmes, parallèlement aux expérimentations de laboratoire, des suivis longitudinaux en contexte écologique ont été mis en place dès les premières étapes de la conception de Tactos (voir notamment Ali Ammar, 2005 ; Rovira, Gapenne & Ali Ammar, 2010). En effet, développer un tel dispositif nécessite la participation active des utilisateurs ciblés, de façon i) à prendre en compte l'utilisateur final dès les premières étapes de conception, et ii) à valider l'ensemble des étapes et vérifier leur pertinence dans des contextes d'usage réel.

Cette démarche de conception participative permet d'intégrer les facteurs humains dans le processus même de l'élaboration du système technique (Buisine & Roussel, 2008). En cela, elle est une alternative pertinente aux modèles de conception classiques, qui partent de l'analyse d'un besoin pour finir par un processus d'évaluation (Pahl & Beitz, 1996). En effet, dans ces processus linéaires, les utilisateurs ne sont impliqués qu'après la réalisation technique, ce qui conduit à un paradoxe de l'ergonomie de conception (Theureau & Pinsky, 1984). Plutôt que de guider la conception par des considérations purement techniques, il semble en effet nécessaire de prendre en compte l'utilisateur dès le départ. Cette démarche de conception centrée utilisateur (norme ISO 13407, 1999) permet alors de penser ensemble l'objet technique et son usage (Jouët, 1993).

Analyser l'usage d'un dispositif technique consiste alors à observer et interpréter les actions de l'utilisateur, dans des contextes d'emploi réels, avec pour objectif de comprendre son fonctionnement et son intégration par l'utilisateur lui-même. Il s'agit donc, conjointement au déploiement de tests-utilisateur en situation contrôlée (simulation d'un usage donné via un scénario prédéterminé), de mener une investigation portant sur l'usage lui-même (observation de situations d'usage libre). Procéder à une telle méthode à l'intérieur même du processus de conception technique, c'est alimenter les hypothèses de conception, proposer des pistes de

solution pertinentes et se donner les moyens de préciser conjointement l'utilité et l'utilisabilité à travers la mise à l'épreuve du concret (Bationo-Tillon, 2007 ; Maillet, 2010). De cette façon, l'analyse d'usage rend possible l'étude de l'appropriation du dispositif, sa prise en main et son intégration dans l'activité de l'utilisateur, ainsi que la stabilisation même de cette pratique. En outre, il permet de spécifier sa pertinence écologique, et de préciser ses situations d'usage, d'observer l'évolution de ces situations, voire l'apparition d'usages non anticipés (Von Hippel, 2005).

Dans ce contexte, il a été question de poursuivre les suivis longitudinaux, de façon à analyser l'usage du dispositif Tactos, en recueillant un ensemble de données quantitatives et qualitatives concernant l'appropriation et l'utilisabilité du système dans des contextes écologiques. Pour cela, nous avons recontacté un des premiers utilisateurs (**C.**) mobilisé dès le début du projet Tactos. **C.** est né avec la rétinite de Coats, qui s'est traduite par une déficience visuelle profonde dès la naissance et qui a conduit à la cécité complète au cours du 19e mois. Notons que **C.** est très intéressé par les nouvelles technologies, et maîtrise l'environnement informatique grâce à l'utilisation d'un logiciel de synthèse vocale. **C.** a été suivi dans le cadre de l'étude de l'appropriation du système Tactos pendant deux ans (classe de 5e et de 4e), pendant lesquels il a pu bénéficier d'une pratique régulière du dispositif, et développer ainsi une certaine expertise dans l'utilisation de celui-ci. En classe de 3e, il a eu l'occasion d'utiliser le dispositif de façon plus ponctuelle. Lorsque nous reprenons le suivi, **C.** est en classe de Terminale, et deux ans d'interruption totale de pratique sont donc à déplorer.

Parallèlement à cette reprise du suivi pour **C.**, deux nouveaux utilisateurs ont été contactés par l'intermédiaire du centre Normandie Lorraine[2], institution partenaire de l'équipe du projet depuis le début de la conception de Tactos. Naturellement intéressés par le développement de nouvelles technologies, notamment en ce qui concerne l'accessibilité des supports numériques, ce centre propose aux jeunes déficients visuels, sur la base du volontariat, de participer au projet Tactos afin de contribuer à la conception de cet outil. Deux élèves se sont donc portés volontaires, en accord avec leurs familles, pour participer au projet. Il s'agit d'une élève de 4e (**S.**), aveugle complète de naissance (glaucome congénital), et d'un élève de 5e (**Y.**), aveugle complet depuis l'âge de 7 ans suite à un accident de la circulation.

Les évaluations du dispositif permettent d'orienter les choix de conception, en adaptant au mieux les différents outils aux utilisateurs. Pour chaque étape de conception, il a été

[2] Le centre Normandie Lorraine est un centre situé dans la commune de Mesnil-Esnard (76). Il assure la prise en charge médicale, psychologique et pédagogique de jeunes déficients visuels âgés entre 4 et 20 ans, que ce soit dans les sections internes spécialisées ou en intégration dans les établissements scolaires des départements de Seine-Maritime (76) ou de l'Eure (27).

question de recueillir un ensemble de données qualitatives et quantitatives concernant l'usage du dispositif, puis à les analyser afin de proposer des spécifications matérielles et fonctionnelles, en collaboration avec l'équipe du projet. Le but de ces spécifications concerne i) l'amélioration qualitative de Tactos, afin de pallier les difficultés particulières rencontrées par les utilisateurs et ainsi favoriser l'utilisabilité générale du dispositif, et ii) la perspective de nouvelles extensions du dispositif initial, notamment dans le cadre d'un élargissement des contextes de son utilisation. Dans cette dernière optique, il s'agit principalement d'anticiper de nouvelles pratiques, et de proposer de nouveaux outils et de nouvelles fonctionnalités en adéquation avec les attentes des utilisateurs.

1.3.2. Méthodologie

Les suivis longitudinaux ont été déployés à raison d'une séance hebdomadaire à domicile pendant les périodes scolaires. Chaque séance dure entre 1h30 et 2h. De façon générale, les suivis longitudinaux permettent d'analyser l'usage du dispositif en contexte écologique. Pour cela, les séances de suivi consistent en la réalisation d'exercices spécifiques (dans le cadre de la prise en main du dispositif), mais aussi de tests-utilisateur en ce qui concerne les évolutions techniques, matérielles et logicielles proposées par l'équipe du projet ou suggérées par les utilisateurs au fur et à mesure du suivi. Ainsi, toutes les fonctionnalités développées sont mises à l'épreuve de l'utilisateur, ce qui permet une évaluation itérative et corrective du dispositif. De cette façon, chacune des séances vise le recueil de données d'usage objectives, tout comme la discussion du système par les utilisateurs, son utilité, son utilisabilité et la satisfaction liée à son usage.

Pour les nouveaux utilisateurs, la première séance a consisté en une présentation générale du projet, de ses objectifs et de son principe d'utilisation. Nous en avons profité notamment pour discuter de la pertinence du dispositif, que ce soit dans un contexte pédagogique, pratique ou ludique. Les deux utilisateurs ont alors manifesté de l'enthousiasme envers le projet, et ont exprimé leur volonté de commencer le suivi. Un plan pédagogique initial a été alors élaboré, en accord avec les premières observations de terrain effectuées lors des premiers suivis longitudinaux (Ali Ammar, 2005 ; Rovira et al., 2010).

Le plan pédagogique vise en premier lieu la prise en main initiale du dispositif, notamment en ce qui concerne la manipulation de l'effecteur, mais aussi les premiers exercices, qui concernent l'exploration de lignes simples orientées orthogonalement. Notons que pour les premières étapes d'exploration, la main de l'utilisateur novice est guidée jusqu'à la figure, de façon à ce que le curseur soit en contact avec la figure à explorer. Petit à petit, en

fonction du niveau d'aisance du sujet, nous limitons ces interventions directes, en les remplaçant par des instructions verbales, puis en ne délivrant plus aucune aide à la localisation de la figure dans l'espace numérique.

De façon générale, chaque séance est organisée en fonction d'un objectif défini à l'avance. Cet objectif s'inscrit dans le plan pédagogique préalablement constitué, mais est redéfini de façon continue pour chaque utilisateur en fonction des observations antérieures. Pour atteindre cet objectif, un ensemble d'exercices a été préalablement préparé. Par exemple, si l'objectif d'une séance est de travailler sur la discrimination entre des segments et des courbes, un ensemble d'images présentant un ou plusieurs segments et une ou plusieurs courbes est élaboré. Parmi ces images, les segments et les courbes peuvent être présentés successivement ou simultanément, dans un ordre de difficulté croissante ou non. L'utilisateur explore d'abord seul certaines de ces images, et ses activités sont observées. Ensuite, il rapporte son expérience, dans un entretien semi-directif. L'objectif de cet entretien consiste à lui faire verbaliser ses difficultés éventuelles, son opinion sur la faisabilité d'un tel exercice, mais aussi sur la satisfaction liée à l'usage ou sur la pertinence de l'état actuel du système pour réaliser la tâche. Nous observons alors les utilisateurs en situation d'appropriation du dispositif, et nous évaluons de façon continue l'efficience de son utilisation lors des exercices proposés. L'analyse est avant tout qualitative, basée sur l'observation, même si parfois, des données quantitatives sont recueillies, que ce soit en termes de performances ou en termes de stratégies d'exploration. Après chaque séance, nous réunissons l'ensemble de ces données, de façon à mener une analyse globale de la séance (observations, données recueillies, verbalisations). Cette analyse conduit à une redéfinition du plan pédagogique de façon à proposer des exercices personnalisés lors de la séance suivante. De cette façon, chaque séance est basée sur les difficultés identifiées ou exprimées par l'utilisateur lors des séances antérieures, dans une optique d'évaluation formative.

L'ensemble du processus d'appropriation et d'apprentissage du dispositif est centré sur l'utilisateur. L'objectif est en effet de concevoir des exercices adaptés à ces difficultés, de façon à entraîner l'utilisateur à améliorer voire à modifier certaines stratégies d'exploration. Pour cela, deux cas de figure peuvent se présenter. Premièrement, nous proposons à l'utilisateur une série d'exercices adaptés et de difficulté croissante, de façon à ce que la stratégie d'exploration utilisée préférentiellement soit amenée à être affinée par l'utilisateur lui-même, en le confrontant à une répétition des routines d'exploration ainsi qu'à des difficultés spécifiques. Deuxièmement, il est parfois nécessaire de suggérer des stratégies à

l'utilisateur, particulièrement dans le cas où le sujet a automatisé des routines d'exploration non-adaptées à la tâche.

Ponctuellement, en fonction des étapes de conception du dispositif (qu'elles aient été suggérées par les utilisateurs ou par l'équipe du projet), des tests-utilisateur sont proposés. Dans ceux-ci, il s'agit d'évaluer certaines fonctionnalités ou certains outils dans des situations contrôlées. Pour cela, un scénario d'usage est élaboré. L'objectif consiste à recueillir un ensemble de données qualitatives et quantitatives, que ce soit au niveau des performances réalisées dans les scénarii proposés qu'au niveau de la satisfaction exprimée par les utilisateurs concernant les fonctionnalités ou outils testés. Ces données permettent un retour objectif à l'équipe de conception, de façon à valider, ou, le cas échéant, à invalider le dispositif en vue d'un nouveau test. Dans ce dernier cas de figure, des spécifications techniques et/ou fonctionnelles sont élaborées, en collaboration avec les utilisateurs, de façon à proposer une nouvelle version prenant en compte les observations et données antérieures.

1.3.3. Observations et analyses d'usage

*1.3.3.1. Bilan du suivi de **C.**(utilisateur expert)*

Concernant le suivi de l'utilisateur expert, le premier constat que nous pouvons mentionner ici est que la réadaptation au dispositif a été immédiate. Lors des premières séances, cet utilisateur a pu explorer et reconnaître des figures géométriques de façon très efficace, soulignant ainsi que l'utilisation de ce dispositif fait appel à un savoir-faire pratique, qui ne souffre pas des interruptions dans le temps. Dès les premières explorations, **C.** manifeste une grande expertise à travers des gestes assurés et précis. L'exploration des formes est très rapide, et la stratégie utilisée spontanément est un suivi de ligne continu. En effet, il localise très facilement les formes dans l'espace, en procédant à de rapides mouvements de grande amplitude. Dès le premier contact, le curseur est stabilisé sur la source de stimulation, et les premières informations concernant celle-ci sont interprétées très rapidement à travers les picots activés de la matrice de stimulateurs. L'orientation des lignes touchées est notamment immédiatement détectée, ce qui permet à l'utilisateur d'anticiper la direction à suivre pour garantir une exploration totale de la figure. Lors des incidents géométriques (les angles, les courbes ou les extrémités d'une ligne explorée), une stratégie déjà identifiée est retrouvée instantanément : celle-ci consiste à procéder à un mouvement circulaire autour de la dernière stimulation reçue de façon à rencontrer le cas échéant la suite de la figure (voir Figure 5).

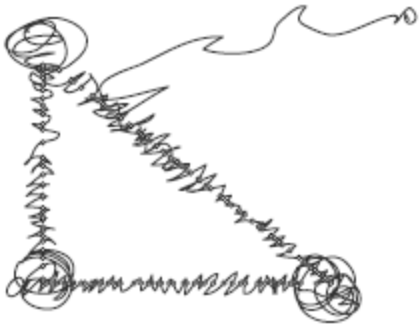


Figure 5 : Trajectoire d'exploration d'un triangle, qui met en évidence la stratégie des mouvements circulaires en cas d'incident géométrique (tiré de Maillet, 2010)

L'essentiel du travail effectué avec cet utilisateur expert a alors consisté à lui faire expliciter ses stratégies d'exploration dans l'optique de mieux comprendre ses performances, et de conseiller efficacement les utilisateurs novices sur la façon de surmonter certaines difficultés, et plus généralement sur la façon d'utiliser le système. Ainsi, les difficultés identifiées par les utilisateurs novices, qu'elles soient générales ou spécifiques à des figures ou des exercices particuliers, ont été soumises à l'expert, de façon à dégager un ensemble de recommandations et de solutions possibles pour les dépasser. De ce point de vue, **C.** s'est vu donner un rôle de consultant, son expertise s'étant révélée être un atout notable pour améliorer les suivis longitudinaux.

Parallèlement, son expertise a été un avantage pour explorer des pistes d'amélioration du dispositif. En effet, son niveau d'expertise nous a semblé être une bonne opportunité pour envisager des outils et fonctionnalités supplémentaires qui pourraient étendre l'usage du dispositif au-delà de la seule exploration de formes géométriques simples. Pour cela, sa maîtrise de l'environnement informatique a été particulièrement utile, dans le sens où des attentes spécifiques concernant l'accessibilité des contenus numériques ont pu être formulées. Par exemple, des observations de son usage de l'informatique pour naviguer sur Internet ont été menées, de façon à dégager un ensemble de spécifications pour envisager l'utilisation de Tactos à cette fin. Ces spécifications ont conduit l'équipe de conception à définir une méthode de navigation spécifique, qui tire parti de l'accès à la spatialisation de l'information (voir Partie 1, section 2.2.1, p. 56).

Enfin, nous avons pu travailler sur la complexification progressive des contenus numériques. L'objectif a consisté à passer de l'exploration de figures géométriques à l'exploration de contenus riches habituellement explorés via les supports traditionnels (impressions en relief thermoformé). Dans cette optique, nous avons observé ses stratégies

dans le contexte de l'exploration de ces supports, de façon à discuter de la pertinence de Tactos et des développements à prévoir pour garantir un accès à ces contenus riches. De cette façon, des contenus numériques spécifiques ont été élaborés, comme des cartes de géographie ou des plans de transports, et ont été systématiquement testés dans une perspective d'évaluation corrective (voir Partie 1, section 2.1.2, p. 48).

*1.3.3.2. Bilan du suivi de **S.**(utilisateur novice)*

Nous pouvons tout de suite préciser que le suivi de **S.** a pris fin au bout de 11 séances. Malgré l'enthousiasme initial manifesté suite à la présentation du projet, l'usage effectif du dispositif n'a pas été concluant. Lorsqu'elle exprime son désir d'arrêter le suivi, **S.** déclare qu'elle ne trouve pas l'intérêt d'explorer des images, ni d'utiliser Tactos pour naviguer sur Internet. Selon elle, les supports pédagogiques traditionnels (essentiellement des impressions en relief thermoformées) sont suffisants, et l'accès à l'informatique est déjà amplement permis par l'usage des logiciels de synthèse vocale ou des barrettes Braille.

En réalité, de grandes difficultés de prise en main sont survenues dès les premières séances. En effet, **S.** a d'emblée manifesté des difficultés concernant la saisie et la manipulation du stylet : la tenue n'est pas assez ferme, ce qui entrave la précision des gestes d'exploration. En fait, le stylet semble être tenu comme un pinceau. Cette première observation n'est pas étonnante, dans la mesure où la déficience visuelle induit souvent une immaturité motrice, se manifestant en particulier par une tendance à l'inactivité manuelle (Fraiberg, 1977). Celle-ci est expliquée par les caractéristiques inhérentes au sens du toucher. Le champ perceptif est restreint (fovéa tactile), et en tant que modalité sensorielle proximale, il n'y a pas d'appel sur l'environnement distal incitant à une exploration systématique des objets (Hatwell, 2003). De ce fait, l'expérience de la motricité fine liée à la manipulation s'en trouve réduite, et certaines coordinations intermodales ne peuvent se mettre en place (ibid.). En outre, l'absence de contrôle visuel ne permet pas l'exercice de la motricité fine au travers d'activités comme la production de dessins ou l'écriture[3], activités fortement incitées et de façon très précoce chez les enfants voyants.

Dans le cadre de Tactos, la difficulté de saisie du stylet rend difficile l'exploration des lignes, surtout dans le cas des verticales. De ce fait, **S.** a tendance à perdre systématiquement

[3] A ce propos, nous pouvons tout de même mentionner un outil proposé dès les années 60 par la Swedish Foundation for the Blind et qui est très utilisé en contexte pédagogique. Il s'agit d'une « planche molle » en caoutchouc sur laquelle on dispose un type de feuille particulier : le passage de la pointe d'un stylo ou d'un crayon permet de « tracer » en relief, ce qui donne la possibilité à l'utilisateur de contrôler tactilement sa production.

le contact avec les formes explorées, et la difficulté de contrôle du geste d'exploration ne lui permet pas de les retrouver facilement en produisant le mouvement inverse. Pour tenter de résoudre ce problème, les quatre premières séances ont consisté à tenter de corriger la position de la main, tout en proposant des exercices de discrimination simple entre des lignes horizontales, verticales et obliques. Sans amélioration notable, nous avons décidé lors de la 5^{e} séance de tester deux nouveaux effecteurs : une souris informatique classique, et un stylet aménagé présentant un bord plus large (voir Figure 6).

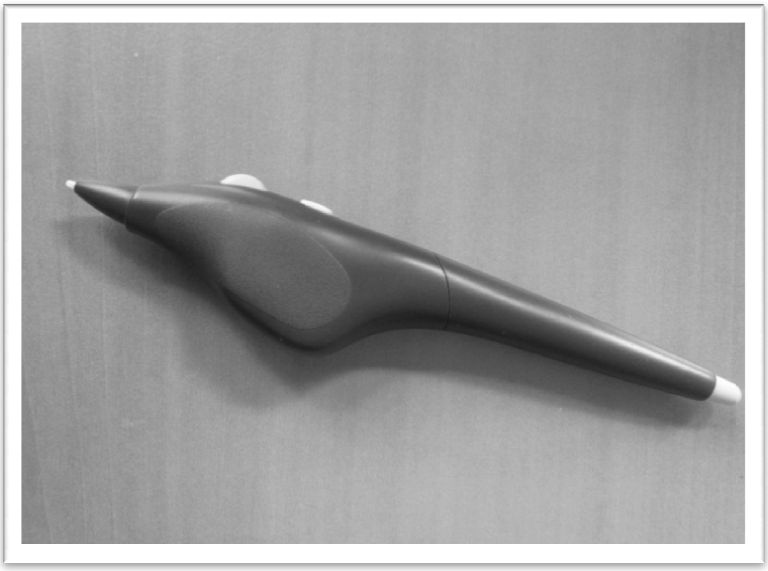


Figure 6 : Stylet aménagé, commercialisé par la société Wacom©. La manipulation de ce modèle de stylet est facilitée par un bord plus large, qui favorise une saisie manuelle plus ferme

Le test de la souris pour le suivi d'une ligne horizontale n'est pas concluant. **S.** perd souvent la figure. Elle nous indique que la souris est "*trop grosse, pas facile à tenir et à bouger*". De plus, l'utilisation d'un effecteur relatif, c'est-à-dire qui commande un déplacement plutôt qu'une position, soulève de nouveaux problèmes (voir Partie 1, section 2.1.1, p. 43), qui ne favorisent pas une utilisabilité effective du dispositif. En revanche, le stylet aménagé paraît plus maniable. **S.** parvient à le saisir plus franchement : les doigts englobent le stylet, à la manière d'un manche, et ce mode de saisie lui permet de maîtriser davantage ses gestes en étant plus tonique. De cette façon, le déplacement sur la tablette est plus souple. Notons que pour les lignes horizontales, la stratégie utilisée est le suivi continu : la ligne est explorée de bout en bout relativement lentement, mais en ne déplorant que très peu de pertes de contact. Lors de l'exploration de la verticale, nous observons que **S.** déploie une oscillation horizontale orthogonale au contour, avec une grande amplitude. Les mouvements sont rapides, et lorsque nous lui demandons de s'arrêter sur la ligne, elle n'y parvient pas. Il

semble que le gain de tonus induit par le mode de saisie de ce stylet soit obtenu au prix de la précision du geste d'exploration. De plus, quel que soit le type de lignes explorées, les gestes d'exploration sont toujours horizontaux : **S.** passe donc du suivi continu pour les lignes horizontales, au balayage latéral pour les autres lignes. Au final, **S.** préfère reprendre le travail avec le stylet initial, avec lequel elle a pris des repères depuis le début de l'apprentissage malgré les difficultés rencontrées.

D'une façon générale, **S.** a tout de suite compris que "*plus elle sent de picots, plus elle est sur la figure*". Mais son attention se focalise davantage sur la forme donnée par les picots que sur les mouvements de sa main, ce qui la conduit à explorer les stimulateurs tactiles par des micro-mouvements avec sa main non-dominante plutôt que d'organiser le geste d'exploration de sa main dominante par rapport à la forme qui les produit. De cette façon, il semble que **S.** ne s'immerge pas dans l'espace perceptif ouvert par l'outil, ce qui est une étape clé de l'appropriation du dispositif (Auvray, 2004 ; Khatchatourov & Auvray, 2005). En outre, cette stratégie présente ses limites dès lors qu'il s'agit d'explorer une oblique ou une courbe. En effet, la matrice de stimulateurs étant orthogonale, l'activation des picots résultant de la rencontre avec une oblique ou une courbe ne produit pas d'information directement « lisible », contrairement aux lignes horizontales et verticales (voir Figure 7). Encore une fois, la reconnaissance de formes via le dispositif Tactos résulte plus de la variation des stimulations en fonction du geste d'exploration que de la simple « lecture » des formes via la matrice de stimulateurs, même si celles-ci peuvent effectivement fournir des indices. En d'autres termes, les stimulations guident le geste, mais ne sont pas en soi des représentations de la forme qu'il s'agirait d'interpréter.

Jusqu'à la 11^{e} séance, qui marquera la fin prématurée du suivi, les exercices ont consisté à tenter d'affiner le geste d'exploration, de façon à réduire l'amplitude des oscillations latérales dans le cas de l'exploration des lignes verticales et des obliques. Pour cela, nous alternons entre des exercices « guidés » (nous tenons la main de **S.** qui tient le stylet) avec des exercices d'exploration libre. Petit à petit, au vu des progrès de **S.**, nous avons introduit les angles et les intersections de lignes. A la fin du suivi, elle est parvenue à reconnaître quelques figures simples (triangle, carré etc.), et même à les dessiner (mais avec une taille beaucoup plus grande).

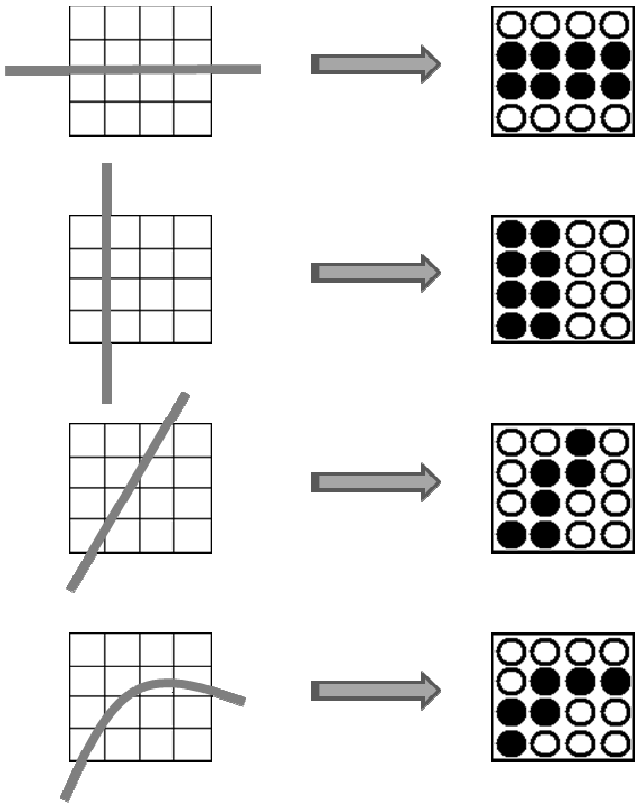


Figure 7 : Stimulations tactiles (à droite) résultant de la rencontre de la matrice de champs récepteurs (à gauche) avec différents types de lignes (en gris).Si les lignes orthogonales permettent une lecture directe de la forme des picots, les lignes obliques et les courbes ne le permettent pas.

*1.3.3.3. Bilan du suivi de **Y.** (utilisateur novice)*

Contrairement à **S.**, **Y.** ne montre pas de difficultés initiales en ce qui concerne la tenue du stylet : celui-ci est saisi de façon ferme et assurée, à la manière d'un stylo classique. En effet, **Y.** ayant perdu la vue tardivement, il a pu bénéficier d'une expérience visuelle précoce, qui lui a permis notamment un développement typique des habiletés motrices (comme par exemple la motricité fine liée à la manipulation d'un crayon) et des coordinations intermodales. Nous pouvons préciser d'emblée que cette expérience visuelle, couplée à des compétences motrices spécifiques telles que la tenue d'un stylo, ont nettement facilité la prise en main du dispositif.

En effet, dès les premières séances, **Y.** a montré beaucoup d'aisance dans l'exploration des lignes et des courbes. Les gestes sont précis, et la stratégie préférentielle est le balayage orthogonal au contour, quelle que soit l'orientation de la ligne explorée. Dès la deuxième séance, il trouve de lui-même la stratégie qui consiste à faire un mouvement circulaire en bout de ligne pour savoir s'il y a autre chose à explorer (par exemple, un autre segment dans le cas d'un angle) (voir Figure 5, p. 32). Toutefois, la tendance à utiliser spontanément une stratégie de balayage lui pose des difficultés dans certaines situations (figures de petite taille,

discrimination fine entre les angles aigus et obtus, et entre les angles et les courbes). Ainsi, l'essentiel de l'apprentissage a consisté à amener **Y.** à adapter sa stratégie lors de l'exploration des incidents géométriques (courbes / angles), notamment en tentant d'appliquer un suivi de ligne plus continu, ou du moins à réduire l'amplitude des oscillations. Pour cela, des exercices présentant des figures géométriques de plus en plus complexes ont été proposés. Malgré une certaine résistance au changement de stratégie, qui semble difficile à modifier après stabilisation, les dernières observations semblent démontrer chez **Y.** la mise en place d'une vicariance des stratégies, qui sont tour à tour utilisées selon un "ordre hiérarchique" en fonction des hésitations ressenties et des tâches proposées.

De même, contrairement à **S.**, avec qui il était nécessaire de placer le curseur sur la figure pour permettre l'exploration, **Y.** a été amené rapidement à trouver de lui-même les figures à explorer, et donc à localiser les objets numériques dans l'espace de l'écran.

Le suivi de **Y.** a duré deux ans. Ces deux ans de pratique lui ont permis de développer une maîtrise certaine du dispositif, et à dépasser les difficultés rencontrées initialement. A la fin du suivi, **Y.** a pu notamment faire preuve de plus de flexibilité, en adaptant la stratégie préférentielle initiale au profit d'un répertoire d'activités d'exploration mobilisées en fonction de la tâche. Au vu de la progression de ses performances, nous avons alors proposé à **Y.** de participer à différents tests-utilisateur, que ce soit en ce qui concerne les évolutions matérielles du dispositif (voir Partie 1, section 2.1.1, p. 43) ou en ce qui concerne l'élaboration de contenus riches à visée pédagogique (voir Partie 1, section 2.1.2, p. 48).

1.4. Le problème de l'usage individuel

Comme nous l'avons vu, et conformément aux observations antérieures, la maîtrise d'un dispositif de suppléance perceptive tel que Tactos nécessite une prise en main et un apprentissage plus ou moins long. L'appropriation d'un tel dispositif passe par plusieurs étapes, qui nécessitent un accompagnement personnalisé de façon à guider au mieux le processus d'apprentissage, en tenant compte des difficultés spécifiques des utilisateurs. Ce processus conduit à l'ouverture d'une modalité perceptive originale, qui permet, dans le cas de Tactos, d'accéder aux informations spatiales des interfaces graphiques numériques. Cet accès est crucial pour les personnes déficientes visuelles dans la mesure où il favorise l'insertion pédagogique et professionnelle. Toutefois, malgré l'enthousiasme des utilisateurs lors des étapes d'apprentissage, il apparaît que le dispositif Tactos n'est pas utilisé en dehors des séances de suivi. De cette façon, les logiciels de synthèse vocale et les barrettes Braille restent les moyens privilégiés des utilisateurs déficients visuels pour accéder aux informations

numériques, en navigant dans les interfaces de façon séquentielle via le clavier de l'ordinateur.

Dans notre cas particulier, il est important de noter que le système Tactos reste un dispositif expérimental, qui ne favorise pas encore un usage autonome. En effet, l'utilisation de Tactos nécessite un ensemble de manipulations préalables de la part de l'utilisateur (installation de drivers, branchement des périphériques, lancement de modules dédiés et paramétrage via une navigation dans les menus etc.), qui implique une maîtrise initiale de l'environnement informatique. Pourtant, si cette maîtrise n'était pas atteinte par les deux utilisateurs novices qui ont été suivis, l'utilisateur le plus expérimenté s'est révélé capable de procéder à l'ensemble des opérations informatiques nécessaires à l'utilisation de Tactos en réelle autonomie. En effet, cet utilisateur a effectivement exploré toutes les possibilités des menus (des questions ont été posées à plusieurs reprises sur des aspects découverts lors de l'exploration des menus), et a rapporté quelques séances d'entraînement autonome, notamment lors des périodes de vacances scolaires. Malgré tout, ces séances d'usage autonome restent ponctuelles, et ne conduisent pas à une adoption définitive du dispositif dans le cadre d'un usage pourtant quotidien de l'informatique. Il semble donc que la difficulté d'accès autonome au dispositif n'explique pas à elle seule la non-adoption par les utilisateurs.

De façon générale, il avait déjà été noté que les dispositifs de suppléance perceptive, malgré leur efficacité, sont délaissés par les utilisateurs. En effet, l'ensemble des entreprises visant l'élaboration et la diffusion de ces dispositifs a conduit à un échec économique et social (Lenay, 2006b), en dépit de données empiriques pourtant très encourageantes. Si les utilisateurs manifestent effectivement de la satisfaction lors de la prise en main de ces dispositifs, ils éprouvent au final une grande déception, et finissent par les abandonner. Malgré des capacités de discrimination et de catégorisation, ils n'accèdent pas à une expérience perceptive à proprement parler. Ce qu'il manque cruellement aux utilisateurs, c'est l'accès aux *qualia*, c'est-à-dire à la valeur émotionnelle des percepts (Bach-Y-Rita, 2002 ; Lenay, 2006b). Un exemple célèbre fait état de la déception d'un utilisateur, qui, après un long apprentissage du TVSS, n'éprouve aucune émotion lors de l'exploration du visage de sa femme (Bach-Y-Rita & Kercel, 2003). De la même façon, la présentation d'images de pin-ups dénudées ne provoque pas de réaction émotionnelle chez des étudiants (ibid.). Les mêmes constats sont observés chez les personnes aveugles qui retrouvent subitement la vue suite à un abattement de la cataracte. Les entretiens menés avec ces nouveaux voyants révèlent une déception similaire : les couleurs ne présentent pas de qualité affective, et les visages ne transmettent pas de message émotionnel (Gregory, 1990). Ces observations ont des

implications cruciales pour comprendre la constitution d'une expérience perceptive. En effet, les *qualia* ne sont pas des informations contenues dans le monde, qu'il s'agirait simplement de capter via un canal sensoriel, qu'il soit naturel ou artificiel (Lenay et al., 2000). Encore une fois, les dispositifs de suppléance perceptive ne remplacent pas simplement une modalité perceptive qui fait défaut, mais permettent d'ouvrir un mode de couplage sensori-moteur inédit. Ainsi, le sens, et notamment la signification émotionnelle des percepts, doit se construire au travers de l'exercice du couplage sensori-moteur, de façon à aboutir à la constitution d'une expérience perceptive par les utilisateurs.

Ainsi, la déception liée à l'utilisation des dispositifs de suppléance perceptive pourrait être liée au manque fondamental d'un apprentissage précoce et/ou durable, qui permettrait la constitution d'une expérience perceptive porteuse de significations et de valeurs émotionnelles. Il en est de même pour l'apprentissage d'une seconde langue : la maîtrise des aspects émotionnels d'une langue nécessite une longue expérience en contexte qui permet une immersion culturelle propice à l'émergence de valeurs perçues (Bach-Y-Rita, 2002). Concernant les dispositifs de suppléance perceptive, des résultats expérimentaux à propos de leur usage précoce par des nourrissons aveugles semblent aller dans ce sens (Mellier, 2005). En effet, des enfants aveugles qui ont l'expérience précoce d'un dispositif de suppléance visuo-auditif (le GUS : Guide UltraSonique ; voir Kay, 1974) manifestent des réponses comportementales immédiates au flux sonore telles que la diminution de l'agitation motrice, ou encore des changements de rythme respiratoire (Bullinger & Mellier, 1988). A la suite d'un apprentissage nécessaire (Sampaio, 1989), des orientations posturales, faciales et des comportements d'atteinte et de préhension à la présentation d'objets dans le champ perceptif, mais aussi des réactions émotionnelles à certains stimuli, comme des sourires en présence de certains objets ou de la mère sont observés (Sampaio & Dufier, 1988). Ces manifestations comportementales témoignent d'un couplage sensori-moteur minimal (Bullinger & Mellier, 1988), accompagné de valeurs émotionnelles (Sampaio & Dufier, 1988).

Seulement, il est également constaté que si ces dispositifs sont utiles aux nourrissons pour construire des actions, ils sont rejetés lorsque la conduite est acquise (Bullinger et Mellier, 1988). Il semble alors que l'expérience précoce de l'utilisation du GUS permet avant tout de favoriser les coordinations intermodales précoces (Streri, 2000). L'ouverture de cette nouvelle modalité perceptive permet alors de fournir un support à certaines étapes du développement, en permettant un accès à des sources d'information jusque-là inaccessibles à cause de la déficience sensorielle. Par ailleurs, l'utilisation régulière du GUS, soit environ 20 minutes par jour, a des conséquences sur le comportement moteur du jeune aveugle lorsqu'il

en est équipé : diminution des blindismes[4], amélioration de l'orientation dans l'espace et de la manipulation des objets (ibid.). En revanche, ces effets ne sont pas durables, et disparaissent la plupart du temps dès que le dispositif n'est plus utilisé (Ferrell, 1980). Les résultats sont donc mitigés, et il est difficile de conclure à un réel impact positif de l'exercice précoce d'un dispositif de suppléance perceptive chez le nourrisson aveugle, si ce n'est en ce qui concerne l'acquisition de certaines conduites spécifiques mais isolées (Hatwell, 2003). Pire encore, il est parfois argumenté que l'utilisation précoce de ces dispositifs peut avoir des effets néfastes sur le développement. Par exemple, l'utilisation détournée de la modalité auditive peut perturber le développement normal du nourrisson, notamment en retardant les coordinations intermodales impliquant l'ouïe (Ferrell, 1980).

Ainsi, l'hypothèse d'un manque d'apprentissage précoce ou durable ne semble pas expliquer à elle seule l'absence de qualia dans les percepts qui conduit à la déception des utilisateurs et à l'échec des dispositifs de suppléance perceptive. Une seconde hypothèse, que nous allons défendre tout au long de cette thèse, porte sur le fait que l'absence de qualia résulte d'un emploi exclusivement individuel de ces dispositifs, qui conduit à un isolement de l'utilisateur dans la modalité perceptive ainsi ouverte (Lenay, 2006b). Or, les valeurs perceptives ne semblent pouvoir émerger que dans le cadre d'une histoire et d'une mémoire collective, qui impliquent un jeu complexe d'interactions entre des individus (ibid.). En effet, « un sujet isolé ne peut attribuer, dans le présent d'un mode de perception nouveau, une signification existentielle aux objets et évènements qu'il perçoit » (Lenay et al., 2000, p.298). Ainsi, l'attribution de qualia aux percepts ne semble possible qu'au travers d'une expérience perceptive partagée entre plusieurs individus. Si, l'acceptation de ces « prothèses perceptives » par les utilisateurs est rendue difficile par l'appareillage nécessaire qui s'impose aux yeux des autres (Lenay, 2006b), l'isolement de l'utilisateur dans un mode de perception propre ne favorise pas l'établissement de relations intersubjectives propices à la constitution de valeurs partagées (Lenay et al., 1997). Pour que le dispositif puisse s'oublier au profit de l'expérience perceptive qu'il permet, il semble nécessaire qu'il s'offre à une communauté d'utilisateurs, de façon à constituer une histoire affective collective et partagée qui rende possible la constitution des qualia (ibid.). Le partage d'un monde commun serait alors la

[4] Les blindismes désignent des comportements moteurs stéréotypés caractéristiques des personnes atteintes de cécité. Il s'agit de balancements corporels (tête, membres, tronc), d'écholalies (répétitions verbales) ou encore de comportements moteurs spécifiques (poings enfoncés dans les orbites) (Hatwell, 2003). Ces comportements sont interprétés comme une forme d'autostimulation qui vient combler l'absence ou le manque de stimulations perceptives ou sociales extérieures (Burlingham, 1967, cité par Hatwell, 2003).

condition nécessaire à l'acceptation individuelle des dispositifs de suppléance perceptive, tout comme à leur adoption sociale par une communauté d'utilisateurs.

2. Vers des espaces d'interaction tactile

C'est cette deuxième hypothèse qui a motivé l'ensemble du travail entrepris dans le cadre de ce travail, ainsi que l'évolution du projet Tactos. L'objectif consiste alors à passer d'un usage strictement individuel (navigation dans les interfaces graphiques numériques, exploration de formes inscrites à l'écran etc.), à un usage collectif, où sont rendus possibles l'interaction entre des utilisateurs distants, l'accès à des contenus partagés et la communication interpersonnelle à travers la modalité perceptive inédite ouverte par le dispositif. Pour cela, il s'agit de constituer des espaces numériques partagés, hébergés sur un serveur dédié, et accessibles localement via une connexion Internet. L'équipe de conception propose alors l'élaboration d'un portail Internet d'interaction tactile, ce qui a donné lieu au projet « Portintertact »[5]. Dans cette optique, nous visons concrètement l'établissement d'une communauté d'usage à travers laquelle les utilisateurs vont pouvoir constituer une histoire collective empreinte de significations et de valeurs partagées.

Mais cette ouverture de l'usage pose un ensemble de problèmes, qu'il s'agit maintenant d'évaluer à travers des études de faisabilité. Si d'un point de vue technique, l'élaboration d'espaces numériques partagés pose la question du mode de connexion des dispositifs Tactos et de l'architecture même de ce réseau, des questions d'usage sont également posées. Ainsi, des analyses d'usage préliminaires ont été menées, de façon à orienter les choix de conception et d'adapter au mieux le dispositif Tactos pour permettre des interactions tactiles distales.

2.1. Etude de faisabilité via des analyses d'usage préliminaires[6]

Comme nous l'avons vu, nous avons entrepris d'étudier l'appropriation du dispositif Tactos et son utilisabilité dans des situations écologiques. Afin d'envisager au mieux l'interaction tactile sur le réseau, des analyses d'usage des « extensions » de Tactos (matériel, fonctionnalités, applications etc.) ont été également menées. L'objectif de ces analyses

[5] Le projet « Portintertact » a été financé par la région Picardie de Décembre 2007 à Décembre 2009.

[6] Cette étude a donné lieu à une publication et à une communication orale :

Deschamps, L., Rovira, K. & Lenay, C., (2008), The appropriation of prosthetic devices for reading computer screens and network interactions via tactile perception, *Sciences et Technologies pour le Handicap, n°spécial : Technologie Tactile et ses Applications pour les Handicaps*, 2(2), p. 159-173

Deschamps, L., (2008), From traditional tools to computer devices, a tactile perception study to adapt digital teaching aids, Tactile Graphics 2008, *The Fourth International Conference on Tactile Diagrams, Maps and Pictures*, 3-5 décembre 2008, Birmingham, UK

d'usage consiste à évaluer les conditions techniques et fonctionnelles devant permettre le passage d'une utilisation locale du dispositif à une utilisation en réseau, que nous pensons propice à une adoption collective du dispositif. Nous tentons donc de rendre compte de l'utilisation de Tactos et de ses extensions dans une démarche évaluative, puis d'orienter la conception d'un nouveau dispositif en fonction des résultats obtenus et des questions soulevées.

Dans cette optique, nous avons mis en place une série de tests-utilisateurs, dans le cadre des suivis longitudinaux que nous avons présentés plus haut (voir Partie 1, section 1.3, p. 27). Parmi les sujets suivis, seuls **C.** (l'utilisateur expert) et **Y.** (utilisateur novice) ont présenté les compétences initiales requises pour prendre part à ces tests. Notons qu'au moment de ces tests, **C.** a 18 ans et **Y.** a 15 ans.

Trois questions initiales ont guidé ces premières observations. D'une part, il a été question d'envisager un nouvel effecteur, qui serait spécialement destiné aux interactions tactiles sur le réseau. Parmi les pistes envisagées par l'équipe de conception, un contact direct de l'utilisateur avec la surface d'action est privilégié, dans la mesure où cela pourrait favoriser l'immersion dans les espaces tactiles partagés, tout comme le sentiment de présence lors de la rencontre avec d'autres utilisateurs. En effet, un contact direct nous semble plus proche d'une perception tactile naturelle, modalité privilégiée dans le cadre émotionnel, comme en témoigne une analyse sémantique du « touchant » (Lenay, 2010). De cette façon, nous nous orientons vers un Touchpad, qui permettrait en outre de s'affranchir de la tenue d'un stylet. Comme nous l'avons vu, le stylet peut poser des problèmes pour certains utilisateurs, notamment ceux qui n'ont pas eu d'expérience visuelle ni d'expérience de la tenue d'un crayon. Parallèlement, le choix d'un Touchpad nous semble tout à fait pertinent, dans la mesure où il permettrait de miniaturiser l'effecteur, rendant possible la conception d'un module portable.

D'autre part, parallèlement à la possibilité d'interagir de façon tactile sur le réseau, l'ouverture de Tactos vers une utilisation en réseau nous permet d'envisager un portail internet dédié, qui pourrait héberger un ensemble de contenus numériques à visée pratique (cartes, plans, etc.), pédagogique (géométrie, géographie, arts, etc.) et ludique (jeux tactiles par exemple). Dans cette optique, une réflexion sur l'accessibilité de contenus riches s'est imposée, de façon à élaborer un ensemble de spécifications destinées à concevoir au mieux ces différents contenus. Lors de cette analyse préliminaire, quelques contenus ont été élaborés en guise d'exemple, et ont été testés par les utilisateurs dans une visée évaluative.

Enfin, outre la rencontre tactile d'utilisateurs distants, les espaces numériques partagés posent la question de la communication tactile au sens large. Plus précisément, nous nous demandons si les stimulations tactiles peuvent véhiculer un contenu émotionnel en dehors du simple contact entre les utilisateurs. Dans le contexte de l'élaboration d'une application de communication asynchrone de type Messenger, nous avons testé l'usage de patterns de stimulations tactiles destinés à se substituer aux émoticônes visuelles. Le but est de pouvoir enrichir le canal de communication en utilisant les retours tactiles.

Pour tenter de répondre à ces différentes questions, nous privilégions une approche qualitative (mais pas uniquement) en nous centrant sur l'appropriation, l'utilisabilité et la satisfaction liée à l'usage. Parallèlement, quelques données quantitatives, basées sur l'étude de la performance à proprement parler ont été analysées.

2.1.1. Effecteurs

Quatre effecteurs ont été testés durant ces analyses d'usage préliminaires. Dans une visée comparative, des tests-utilisateurs ont été élaborés, de façon à évaluer la pertinence de différents types d'effecteurs. Ces tests ont concerné d'une part, deux effecteurs *absolus* (la surface d'action proposée à l'utilisateur correspond à la surface de perception sur l'écran dans un rapport bijectif ; le sujet commande donc une position du curseur à l'écran) : il s'agit du couple Tablette / Stylet, et d'un Touchpad configuré en absolu. Et d'autre part, deux effecteurs *relatifs* (la surface d'action correspond à une partie de l'espace de perception sur l'écran ; le sujet commande donc un déplacement du curseur à l'écran) : il s'agit d'une Souris, et d'un Touchpad configuré en relatif. Pour les effecteurs *relatifs* (Souris et Touchpad *relatif*), les bords de l'écran sont matérialisés par des bords de couleur produisant un son lorsque le curseur les rencontre.

Deux tâches ont été proposées pour chaque effecteur : une tâche d'identification et une tâche de localisation. Tout d'abord, deux figures par effecteur sont proposées : un angle droit et une figure curviligne. Le sujet est placé sur la figure et n'a aucune contrainte de temps pour son exploration. S'il le juge nécessaire, il peut être replacé sur la figure à sa demande. Le sujet doit identifier la figure. Nous relevons la qualité de la réponse, le temps d'exploration, la qualité de l'exploration (exploration intégrale vs exploration partielle), la stratégie utilisée, le nombre de replacements sur la figure et les verbalisations produites. Dans un deuxième temps, nous disposons deux carrés pleins dans l'espace de l'écran. Sans aucun guidage, le sujet doit les retrouver, puis les localiser (en indiquant leur position par rapport à l'espace de l'écran).

Nous relevons le nombre de stimuli trouvés, le nombre de stimuli correctement localisés, le temps d'exploration, la stratégie utilisée et les verbalisations produites.

Dans un premier temps, nous nous intéressons aux scores d'identification des figures géométriques proposées (un angle et une courbe). Nous observons une différence entre les deux sujets. En effet, quand le sujet expert ne semble pas perturbé par le changement d'effecteur (excepté pour l'effecteur Souris qui a donné lieu à une erreur), le sujet novice présente des performances d'identification différentes en fonction de la nature absolue ou relative du Touchpad (voir Tableau 1). Par ailleurs, nous observons que quel que soit le niveau d'expertise, le Touchpad configuré en absolu offre les meilleurs résultats.

Effecteur	Expert	Novice
Tablette/stylet	2/2	1/2
Touchpad *absolu*	2/2	2/2
Souris	1/2	1/2
Touchpad *relatif*	2/2	0/2

Tableau 1 : Scores d'identification obtenus par les deux sujets avec les quatre effecteurs testés

Au début de la tâche d'identification, la main du sujet tenant le stylet est placée sur la tablette, de façon à ce que le curseur soit en contact avec la configuration géométrique. La tâche débute lorsque nous lâchons la main du sujet et que ce dernier commence l'exploration de la forme. Si le sujet perd le contact avec la figure au cours de son exploration et qu'il a des difficultés à la retrouver seul, il a la possibilité d'être replacé dessus par l'expérimentateur. Nous appelons donc « score de désorientation » le nombre de replacements sur la figure demandés par l'utilisateur, ce qui témoigne de la difficulté à s'orienter dans l'espace de l'écran et à retrouver la figure perdue.

De façon générale, le sujet expert perd rarement le contact avec la forme lors de son exploration sauf avec le Touchpad *relatif.* Mais dans ce cas, il parvient toujours à reprendre contact avec la figure (voir Tableau 2). Seul l'effecteur Souris entraîne 3 demandes d'aide pour le replacement. Le sujet novice perd plus souvent le contact mais il parvient à retrouver la forme avec les effecteurs *absolus*. En revanche, il ne parvient pas à retrouver les figures avec les effecteurs *relatifs* comme en témoigne le nombre élevé de demande de replacements lors du test des effecteurs relatifs. Nous avons même jusqu'à 12 demandes de replacements avec le Touchpad *relatif.*

Effecteur	Expert	Novice
Tablette/stylet	0	0
Touchpad *absolu*	0	1
Souris	3	9
Touchpad *relatif*	0	12

Tableau 2 : Scores de désorientation obtenus par les deux sujets avec les quatre effecteurs

Nous nous intéressons maintenant aux temps moyens d'exploration par objet correctement identifié. Les résultats indiquent que l'identification correcte des objets explorés par les deux utilisateurs prend moins de temps avec les effecteurs *absolus* qu'avec les effecteurs *relatifs* (voir Tableau 3).

Effecteur	Expert	Novice
Tablette/stylet	25,5	45
Touchpad *absolu*	46	87,5
Souris	40	195
Touchpad *relatif*	145	X

Tableau 3 : Temps moyens en secondes d'exploration avec les quatre effecteurs

Par ailleurs, les temps moyens d'exploration reflètent également un net écart entre les deux niveaux d'expertise : l'identification correcte des objets par le sujet expert prend moins de temps que l'identification correcte des objets par le sujet novice, et ce quel que soit l'effecteur. De plus, le sujet novice n'a pas réussi à identifier les figures avec le Touchpad *relatif* (d'où l'absence de donnée temporelle dans le Tableau 3).

Ceci souligne deux points qui nous paraissent intéressants. D'une part, l'appropriation de Tactos nécessite le développement d'un savoir-faire particulier, qui permet une utilisation efficiente de l'outil (Ali Ammar, 2005; Stewart et Gapenne, 2004). D'autre part, l'expertise de l'outil développée à partir d'un effecteur particulier (en l'occurrence l'effecteur Tablette) peut être transférée à d'autres effecteurs, à condition que le niveau de compétence initial soit suffisant. Ceci ne fait que confirmer des résultats précédemment obtenus auprès d'adultes voyants aveuglés devant tester différents effecteurs tels que la souris, la tablette ou le joystick (Sribunruangrit, 2004). Dans cette étude, l'ensemble des sujets étaient des "novices" face à Tactos et une très nette hiérarchie en faveur de la tablette avait été observée.

Lors de la deuxième tâche, le sujet doit retrouver deux objets (deux carrés pleins) placés aléatoirement sur l'écran et les localiser dans l'espace de celui-ci. Le Tableau 4 présente ainsi le score de découverte de la cible ainsi que celui de sa localisation (scores entre parenthèses).

Effecteur	Expert	Novice
Tablette/stylet	2 (2)	2 (1)
Touchpad *absolu*	2 (2)	2 (2)
Souris	2 (2)	1 (0)
Touchpad *relatif*	1 (1)	1 (0)

Tableau 4 : Scores de découverte et (localisation) obtenus avec les 4 effecteurs

Nous pouvons remarquer que les sujets ne se différencient pas en ce qui concerne la découverte des cibles sur l'écran (voir Tableau 4). Par contre, le niveau d'expertise a une influence sur la capacité à localiser l'objet relativement au cadre de l'écran. Le sujet expert trouve et localise correctement les objets, sauf dans le cas du Touchpad *relatif*, qui ne paraît pas approprié pour ce type de tâche malgré les bords d'écran sonores. En revanche, le sujet novice éprouve de nettes difficultés avec la Souris et le Touchpad *relatif* comparativement aux effecteurs absolus, la meilleure performance étant obtenue avec le Touchpad *absolu*. Dans cette situation, les objets doivent être localisés en fonction d'un cadre de référence spatial qui est ici délimité par les bords de l'écran. Le référentiel spatial est soit de type égocentré, soit de type allocentré (Paillard, 1971). Le deuxième type de référentiel est indépendant du point de vue du sujet et permet ainsi d'appréhender les relations entre objets. Nous pouvons faire l'hypothèse que le sujet novice a des difficultés à construire ce type de cadre de référence au cours de ses explorations lorsque les effecteurs sont relatifs. Toutefois, il est difficile d'attribuer au seul niveau d'expertise les différences entre les deux sujets. Le sujet expert étant plus âgé que le sujet novice, sa capacité à utiliser un cadre de référence allocentré pourrait également être expliqué par un niveau de développement plus avancé.

Par ailleurs, en ce qui concerne les stratégies d'exploration, nous pouvons dire que celles-ci sont largement dépendantes de la configuration de l'effecteur (*relatif* vs *absolu*). Avec les effecteurs *relatifs*, la stratégie a principalement consisté à repérer et suivre les bords d'écran sonores, ce qui a empêché les sujets de s'aventurer trop loin de ceux-ci, et donc d'explorer la totalité de l'écran. Nous retrouvons ainsi une stratégie d'exploration mise en place par les aveugles lors de la découverte d'une pièce inconnue : la "stratégie du périmètre"

qui consiste à recueillir des informations sur la forme et la taille de la pièce (Hill & Ponder, 1976) avant d'aller explorer les objets qui s'y trouvent.

Les résultats obtenus lors du test des effecteurs nous paraissent important pour la bonne conduite du projet d'élaboration d'un portail d'interactions tactiles. En effet, avec l'effecteur original (i.e. Tablette / Stylet), une source de difficulté a été soulevée grâce aux analyses d'usage. En effet, un utilisateur non-voyant de naissance n'a que très rarement une expérience de la tenue d'un stylo, ce qui peut empêcher l'utilisation efficiente du dispositif par une contrainte extrinsèque. De cette façon, le Touchpad *absolu* nous paraît être une alternative intéressante. D'une part, cet effecteur permet des performances comparables à celles obtenues avec l'effecteur original. D'autre part, son utilisation semble plus intuitive que la manipulation du stylet sur la tablette, en se rapprochant davantage de la perception tactile naturelle.

Quoi qu'il en soit, ces études préliminaires en situation d'usage nous ont permis d'élaborer des spécifications quant à la conception d'un module d'interaction tactile spécialement dédié à l'usage de Tactos dans le cadre du portail Intertact. Ces spécifications nous ont amené à ce moment du projet à lancer le prototypage d'un module portable comprenant un Touchpad *absolu* (voir Figure 8).

Figure 8 : Le module d'interaction tactile (face avant)

En outre, ce premier prototype est composé de deux matrices de stimulateurs tactiles (chacune composée de 2 cellules Braille piézoélectriques) disposées de chaque côté du Touchpad. Notons que les cellules Braille implémentées disposent de deux boutons chacune, ces derniers pouvant servir d'actions-utilisateur dédiées en fonction de la tâche. Par ailleurs,

suite à un travail préalable de l'équipe, une nappe de capteurs de pression (FSR) a été implémentée sur la deuxième face du module (voir Figure 9). Cette nappe serait spécialement destinée à un mode d'interaction direct, comparable à une « caresse tactile distale » (Maillet, 2010).

Figure 9 : La deuxième face du module d'interaction tactile

Ce module a donné lieu à de nouvelles analyses d'usage, que nous relaterons dans les sections 2.2.2 et 2.2.3 ci-après.

2.1.2. Supports pédagogiques numériques

Tactos offrant la possibilité d'explorer des images numériques par la modalité tactile, un portail Internet tactile nous paraît approprié pour la diffusion de contenus numériques divers. Dans le cadre de l'appropriation du dispositif, nous proposons aux jeunes utilisateurs d'explorer des formes géométriques sélectionnées lors de précédentes études longitudinales (Ali Ammar, Gapenne, Blomme et Rovira, 2005 ; Ali Ammar, 2005). Ces figures sont explorées, discriminées et reconnues avec plus ou moins de difficulté, mais avec une certaine efficience. Ces résultats d'observations nous permettent donc d'envisager la lecture via Tactos de supports variés d'information, tels que des cartes de géographie, des diagrammes, ou encore des graphiques, qui pourraient être disponibles en ligne via le portail d'interactions tactiles. Dans cette optique, une attention particulière a été portée sur les stratégies d'exploration tactile des supports traditionnels (impressions en relief thermoformé). Pour cela, nous avons sélectionné certains de ces supports, et nous les avons présentés à nos deux

utilisateurs, dans l'optique de guider l'élaboration de contenus numériques accessibles via notre dispositif. Si nous avons observé leurs stratégies haptiques principalement dans le cadre de l'exploration de contenus pédagogiques, notons que nous avons également proposé des tâches plus complexes, comme des tâches de rotation mentale. Nous avons alors rejoint et complété des expérimentations antérieures sur ce sujet[7].

Pour évaluer l'accessibilité de contenus riches via le dispositif Tactos, une première carte de géographie a été élaborée par nos soins et testée lors des suivis longitudinaux. Il s'agit d'une carte de France relativement épurée, minimaliste, présentant uniquement les contours du territoire et sept villes. Celles-ci sont représentées par des points dans l'espace (voir Figure 10). Lorsque le curseur rencontre un de ces points, le nom de la ville est restitué via un dispositif de synthèse vocale (« *Jaws*© » de la société *Freedom Scientific*[8] pour Windows©). Notons que les noms des villes sur la carte sont d'une couleur qui n'est pas détectable par la matrice de champs récepteurs. Nous avons demandé aux sujets d'explorer librement cette carte, et d'exprimer leur satisfaction ou leurs difficultés.

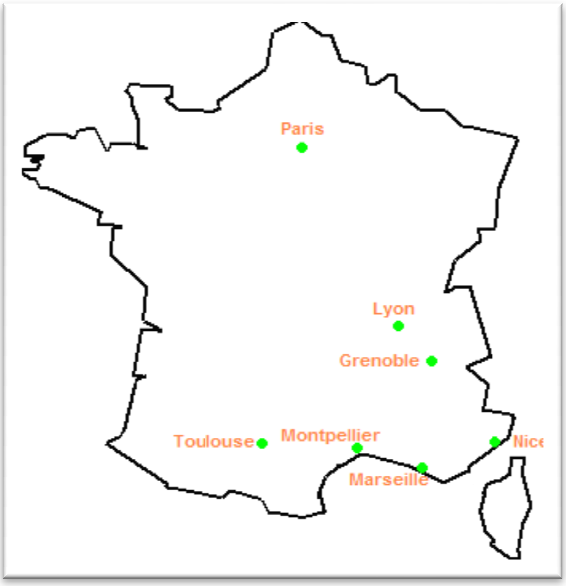


Figure 10 : Carte de France utilisée pour l'analyse de l'exploration de contenus riches

Les résultats de ces tests sont de deux ordres. Premièrement, les observations sont encourageantes, dans la mesure où ce type de support minimaliste permet un accès à

[7] Dans ce cadre, nous pouvons mentionner une publication et une communication affichée :
Rovira, K., Deschamps, L. & Baena-Gomez, D., (2011), Mental rotation in blind and sighted adolescents: The effects of haptic strategies, *European Review of Applied Psychology*, 61, p. 153-160
Rovira, K. & Deschamps, L., (2008), Connaissance de l'espace bidimensionnel chez l'adolescent déficient visuel, Exemple de la rotation mentale, *ARCo-08, Connaissances : Genèse, Nature et Fonction*, 3-5 décembre 2008, Lyon, France (Poster)

[8] www.freedomscientific.com

l’information. Les sujets ont en effet pu localiser quelques villes et explorer le contour de la carte. En revanche, l’analyse d’usage et les verbalisations ont soulevé un problème de guidage au sein de l’espace. En effet, il s’avère difficile de discriminer l’intérieur de l’extérieur de la carte. De plus, il apparaît que l’exploration se fait « au hasard », jusqu’à « rencontrer » une ville de façon fortuite. De cette façon, l’exploration ne permet pas la recherche intentionnelle d’une information, à moins d’avoir une connaissance préalable du contenu. De son côté, l’utilisateur novice n’a pu localiser que les villes situées au contact proche du contour. Au contraire, l’utilisateur expert n’a pas hésité à s’aventurer dans l’espace, ce qui lui a permis de rencontrer davantage de villes. Lorsque l’une d’entre elles est trouvée, sa stratégie a consisté à retrouver le contour pour tenter de localiser la ville relativement à celui-ci. Cela montre encore une fois que les sujets transposent à la situation Tactos les stratégies mises en place lors de l'exploration d'un espace plus large. En effet, dans le cadre de la localisation d'objets dans une pièce, certaines personnes non-voyantes choisissent un objet comme repère pour aborder les relations spatiales entre cet objet et les autres (stratégie du "point de référence", Tellevik, 1992).

Ces premières observations nous ont conduits à élaborer une nouvelle carte, enrichie par un système d’aide à la localisation des villes. Pour cela, l’utilisateur peut saisir à l’aide de son clavier le nom d’une ville (ou plutôt dans notre cas, sélectionner une des villes présentes par l’intermédiaire d’un menu déroulant), de façon à faire converger des « lignes-guides » depuis le contour de la carte jusqu’à la ville recherchée (voir Figure 11).

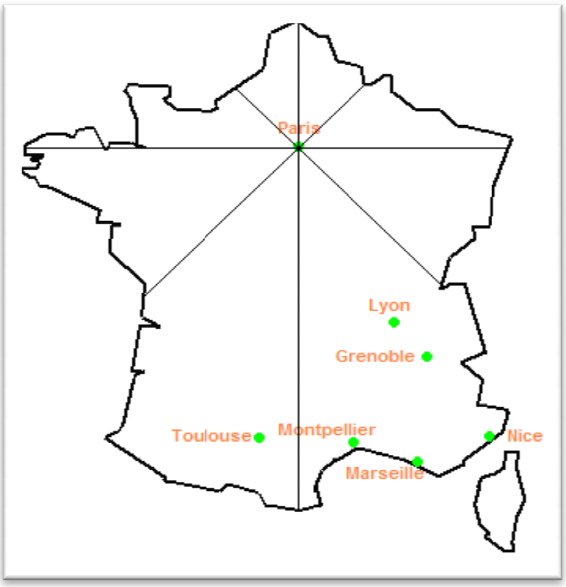


Figure 11 : Carte de France enrichie de "lignes-guides" convergeant vers la ville de Paris

Ce système de guidage perceptif permet en effet aux utilisateurs d'explorer l'intérieur de la carte de façon systématique. Toutefois, les utilisateurs attirent notre attention sur le fait que les stimulations résultant de la rencontre des « lignes-guides » peuvent interférer avec l'exploration du contenu en tant que tel. Pourtant, il est évident que le système d'aide ne doit pas venir gêner l'exploration du contenu, sous peine de devenir inefficace voire néfaste. De plus, le recours aux « lignes-guides » ne favorise pas la découverte des informations, en obligeant l'utilisateur à une exploration séquentielle du contenu, ville par ville.

Le recours à la bimodalité (tactile + auditif), d'ailleurs déjà repéré dans le contexte de l'exploration de formes mathématiques (Ali Ammar, Gapenne, Lenay et Stewart, 2002), semble alors fournir des pistes de conception intéressantes. Au-delà de la lecture d'un point d'information localisé dans l'espace (comme une ville), des informations auditives pourraient désigner des surfaces sur lesquelles évolue le curseur. Dans cette optique, nous « remplissons » la carte de couleurs, qui, lorsque le curseur rentre en contact avec elles, émettent une tonalité sonore. De cette façon, quelle que soit la position du curseur dans le contenu, une information est toujours disponible. Plus encore, dans le cadre de l'exploration de cartes géographiques, ces « couleurs sonores » peuvent également véhiculer des informations situées telles que le nom des régions ou des départements, des fleuves ou encore des mers. Pour tester cette possibilité, nous élaborons une nouvelle carte de France (voir Figure 12).

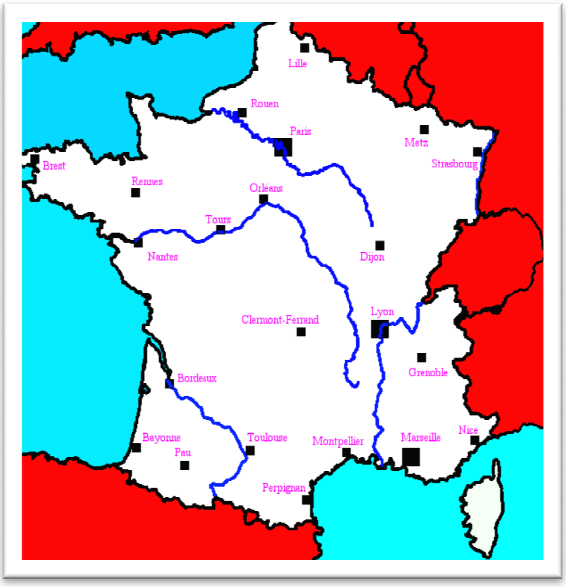


Figure 12 : Carte de France enrichie de "couleurs sonores"

Dans celle-ci nous figurons les océans, les mers et les fleuves par des tons de bleu différents : la rencontre du curseur avec un de ces bleu déclenche la lecture du nom correspondant (ex : « La Seine »). Les pays avoisinants sont figurés par du rouge, et renvoient une tonalité sonore continue. La couleur blanche attribuée à l'intérieur du territoire français renvoie l'information « La France », à l'exception de la Corse, qui renvoie l'information « La Corse ». La lecture de chaque son est déclenchée lors de l'entrée du curseur dans la zone de couleur. Seule la couleur noire déclenche des stimulations tactiles. Parallèlement aux couleurs sonores, nous figurons les villes par des carrés, qui renvoient non seulement une stimulation tactile, mais aussi le nom de la ville pré-enregistré.

L'exploration de cette carte s'est avérée être un succès auprès des utilisateurs. L'enrichissement du contenu par l'ajout d'informations sonores semble conforter le repérage dans l'espace. En revanche, l'utilisation de tonalités sonores continues ne s'est pas montrée satisfaisante : le fait d'entendre un son sans interruption lors de l'exploration a été jugé fatigant, voire « oppressant ». Si nous voulions au départ « forcer » l'utilisateur à explorer le territoire français et pas les autres territoires, nous n'avions pas prévu que l'exploration des frontières de la France avec les pays limitrophes déclencherait également la lecture du son, qui plus est chez l'utilisateur novice, qui déploie préférentiellement une stratégie de balayage orthogonal au contour. De cette façon, l'exploration du contour peut être parasitée par le déclenchement des sons associés aux zones juxtaposées. De plus, le déclenchement de la lecture des informations situées par l'entrée du curseur en contact avec la couleur peut également dans certains cas s'avérer gênant. Selon les utilisateurs, il serait judicieux de spécifier ces zones par une autre information, comme un retour tactile spécifique ou un bip sonore ponctuel, qui inviterait le sujet à déclencher de lui-même la lecture de l'information par l'intermédiaire d'une action-utilisateur, comme la pression d'un bouton dédié (situé idéalement sur l'effecteur). De cette façon, l'information ne serait lue qu'à la demande de l'utilisateur.

Parallèlement à l'utilisation de couleurs sonores, une autre piste, inspirée des supports pédagogiques classiques (impressions en relief thermoformées), a été étudiée. Il s'agit de l'utilisation de textures tactiles, pouvant servir à spécifier l'intérieur d'une carte et les différences entre des zones juxtaposées. Un travail préalable a été mené dans ce sens, de façon à vérifier que des textures du même type, mais présentant des densités différentes (par exemple, une texture composée de hachures verticales espacées de 2 pixels et une texture composée de hachures verticales espacées de 6 pixels) pouvaient être discriminées via le dispositif Tactos (Tao, 2009). 24 sujets voyants aveuglés, âgés de 23 à 55 ans ont été

confrontés à 18 images de 640x480 pixels comportant deux textures juxtaposées (par exemple, voir Figure 13).

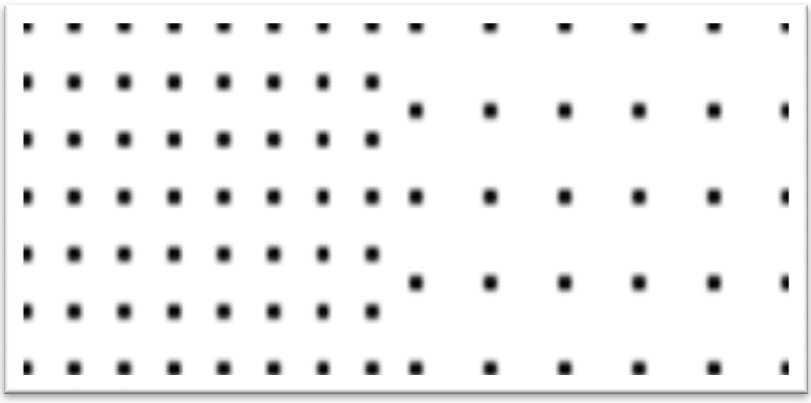


Figure 13 : Exemple d'image utilisée dans l'étude de Tao (2009). Dana la texture de gauche, la densité des points est de 6 pixels, alors que dans la texture de droite, la densité des points est de 10 pixels.

Les résultats de cette étude montrent que quelle que soit la configuration du curseur (un seul champ récepteur de 2x2 pixels ou une matrice de 16 champs récepteurs de 2x2 pixels chacun), deux textures juxtaposées sont correctement discriminées si leur densité présente une différence d'au moins 2 pixels (ibid.). Par ailleurs, lors de tests effectués avec les utilisateurs non-voyants lors de ces analyses d'usage préliminaires, nous avons observé que des textures de type différent avec une même densité (par exemple, une texture ponctiforme et une texture composée de hachures horizontales) sont facilement discriminées, et ce, même par l'utilisateur novice. Par conséquent, l'utilisation conjointe de textures différentes et de densités différentes, parallèlement à l'utilisation d'informations situées sonores et/ou verbales, est alors possible dans le cadre de l'élaboration de contenus numériques riches.

2.1.3. Emoticônes tactiles

L'interaction tactile sur le réseau, médiatisée par le dispositif Tactos, soulève de nombreuses questions. Parmi elles, nous nous demandons quelle valeur émotionnelle peut être attribuée à la stimulation tactile, de façon à permettre une interaction porteuse de sens entre deux ou plusieurs utilisateurs. Plus précisément, dans le cadre d'une communication asynchrone de type Messenger, la question est de savoir ce que pourrait apporter la modalité tactile à la dynamique interactionnelle.

Nous avons donc sélectionné quelques émoticônes tactiles, préalablement conçues dans le cadre d'un travail exploratoire sur les interactions haptiques (Maillet, 2010). L'objectif de ce travail consiste à trouver une alternative aux schémas visuels utilisés dans la

communication asynchrone sur Internet (les « Smileys » qui traduisent l'état émotionnel de l'émetteur). Le but est de pouvoir figurer les indices non-verbaux et paraverbaux de la communication via la modalité tactile, de façon à enrichir les interactions verbales (par exemple, voir Figure 14).

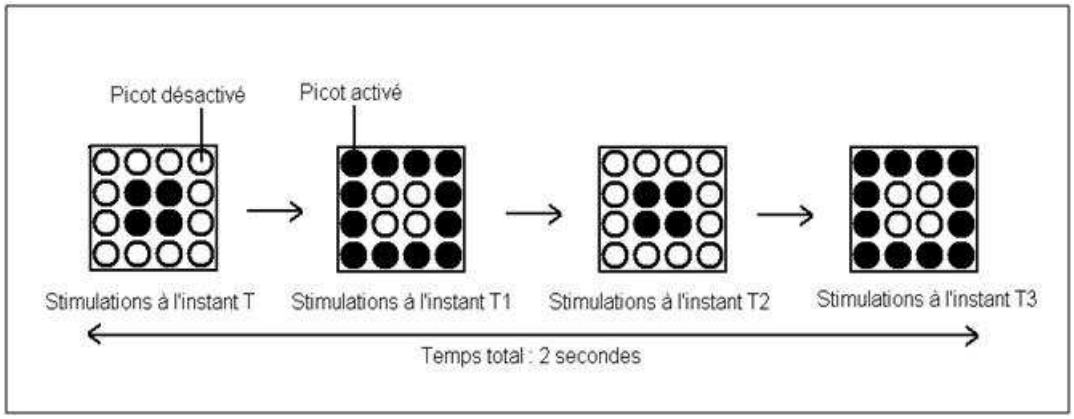


Figure 14 : Exemple d'émoticône tactile, ou « tact-icône ». Ici, le pattern séquentiel de stimulations représente la joie

La Figure 14, ci-dessus, fournit l'exemple de l'émoticône tactile de la joie. L'état émotionnel de "joie" est figuré par une alternance d'expansion et de contraction du champ de stimulation. Après un premier test auprès des deux adolescents, il apparaît que ces séquences de stimulations n'ont pas de sens en dehors de l'utilisateur qui les a produites : l'émergence d'un sens est possible dans le cadre d'une convention entre *n* utilisateurs, mais pas dans celui d'une « norme » de communication que l'on imposerait aux sujets. D'où plusieurs questions soulevées, comme par exemple celle qui concerne les conditions d'émergence d'une signification à partir d'une stimulation tactile arbitraire. Tout au long de ce travail, nous tenterons de préciser ces conditions dans différents contextes d'interaction. Quoi qu'il en soit, il ressort de ces observations préliminaires que les utilisateurs ont exprimé leur souhait d'avoir accès à ces émoticônes, et éventuellement de pouvoir en créer eux-mêmes. Une telle fonction laisserait ainsi la possibilité aux utilisateurs de convenir de leur signification dans le cadre d'une interaction. Par ailleurs, plusieurs remarques nous ont conduits à imaginer une extension de l'utilisation de ces signes au-delà de leur caractère émotionnel. En effet, de tels signes pourraient se révéler utiles pour la navigation dans les interfaces informatiques, de façon à figurer tactilement différents états du système (comme une jauge d'attente lors d'un processus en cours par exemple), ainsi que différentes fonctions ou éléments affichés à l'écran (liens activables ou boutons par exemple). Ainsi, d'une exploration de la communication émotionnelle, nous en sommes arrivés à la conception de « tact-icônes » plus générales.

En outre, les différentes remarques présentées par les utilisateurs ouvrent également sur la question de la définition des « corps-image » (c'est-à-dire des avatars tactiles, perceptibles par les champs récepteurs d'autres utilisateurs). En effet, le fait de pouvoir élaborer sa propre configuration de stimulateurs tactiles permettrait une définition personnalisée des avatars tactiles, de façon à individualiser ce que l'utilisateur laisse à percevoir pour les autres utilisateurs. De cette façon, la reconnaissance d'autrui au sein d'un espace partagé pourrait être facilitée, et l'interaction pourrait s'en trouver enrichie.

2.1.4. Conclusion et perspectives de l'étude de faisabilité

La première étape du projet a consisté à mener un ensemble d'études préliminaires. En effet, afin de permettre le passage d'une utilisation locale à une utilisation en réseau, les « extensions » de Tactos (fonctionnalités, applications et matériel) ont été testées au fur et à mesure par deux utilisateurs (un sujet expert et un sujet novice) (Deschamps, Rovira & Lenay, 2008 ; Deschamps et al., 2009). Le but de ces analyses d'usage préliminaires était de rendre compte de l'utilisation de Tactos et de ses extensions dans une démarche évaluative, puis d'orienter la conception du dispositif Intertact en fonction des résultats obtenus et des questions soulevées. Les résultats obtenus ont permis de proposer et d'implémenter des spécifications techniques et logicielles, tout en respectant trois axes majeurs : la pertinence du dispositif pour les tâches visées (principalement l'interaction), l'efficacité de son utilisation et la satisfaction des utilisateurs (ibid.).

La convocation des utilisateurs dans le processus même de conception nous a donc permis de développer les premiers prototypes du dispositif Intertact, et d'envisager la poursuite de l'analyse systématique des nouveaux usages via une approche qualitative et quantitative. Le projet d'étendre l'usage de Tactos à une situation d'interaction en réseau nécessite en effet tout un ensemble de tests des possibilités matérielles et fonctionnelles. Ainsi, des analyses d'usage en contexte ont permis de différencier les difficultés qui relèvent des limites du matériel, de celles liées aux difficultés propres des utilisateurs. Le fait que notre travail s'insère dans une étude longitudinale permet justement le développement de compétences chez l'utilisateur. Il semble important que ce travail se fasse, dans un premier temps, sous la forme d'un suivi individuel afin de repérer les intérêts propres au sujet. Ceci, ajouté à la collaboration avec l'équipe enseignante, permet d'utiliser un matériel en adéquation avec le niveau scolaire et les intérêts du jeune.

Nous avons pu noter qu'un utilisateur peut changer d'effecteur (passer de la tablette au Touchpad, par exemple) à condition qu'il ait atteint un certain niveau de compétences en

termes de stratégies d'exploration. De cette façon, si l'utilisateur expert semble pouvoir changer d'effecteur sans que cela n'altère sa performance, l'utilisateur novice a permis de mettre en évidence des différences plus saillantes entre les effecteurs testés. Ainsi, les résultats obtenus soulignent la nécessité d'utiliser des effecteurs absolus. Parallèlement, les résultats obtenus avec le Touchpad absolu nous ont permis de conclure à une utilisabilité satisfaisante, ce qui a rendu possible le prototypage d'un module d'interaction tactile portable.

En ce qui concerne les évolutions à apporter au dispositif, le recours à la bimodalité (visuelle et tactile) semble prometteur dans le contexte du repérage spatial. De cette façon, il s'agit d'une piste sérieuse en ce qui concerne l'élaboration de contenus numériques graphiques accessibles. Parallèlement, l'utilisation de textures tactiles nous semble pertinente, au vu des résultats obtenus, et nous encourage à poursuivre l'étude de l'accessibilité de contenus enrichis. A terme, les différents contenus seront hébergés sur le portail et pourront être explorés seuls ou à plusieurs. De même, il s'agit d'intégrer des techniques déjà validées auprès de voyants aveuglés afin d'améliorer l'usage de Tactos. Par exemple, la technique du zoom haptique permettrait des changements d'échelle intéressants dans le cadre de l'utilisation des cartes (Ziat, Gapenne, Lenay, Lecolinet, Mouret & Stewart, 2006; Ziat, Gapenne, Stewart & Lenay, 2007).

Les observations présentées ici ne sont qu'une des premières étapes dans un processus itératif de développements techniques et d'analyse des usages sur lesquels nous allons revenir tout au long de ce travail.

2.2. Le dispositif Intertact ou Tactos-Réseau

2.2.1. Présentation générale et objectifs

Comme nous l'avons vu, le système Tactos est un dispositif de suppléance perceptive qui permet l'exploration et la reconnaissance de formes numériques bidimensionnelles. Mais un tel système permet aussi de « toucher » d'autres utilisateurs dès lors que l'on crée un espace numérique partagé. Pour ce faire, il a été question de connecter deux systèmes Tactos (voir Figure 15). Ainsi, l'espace numérique est partagé via le réseau, et rend possible des rencontres tactiles entre des utilisateurs distants.

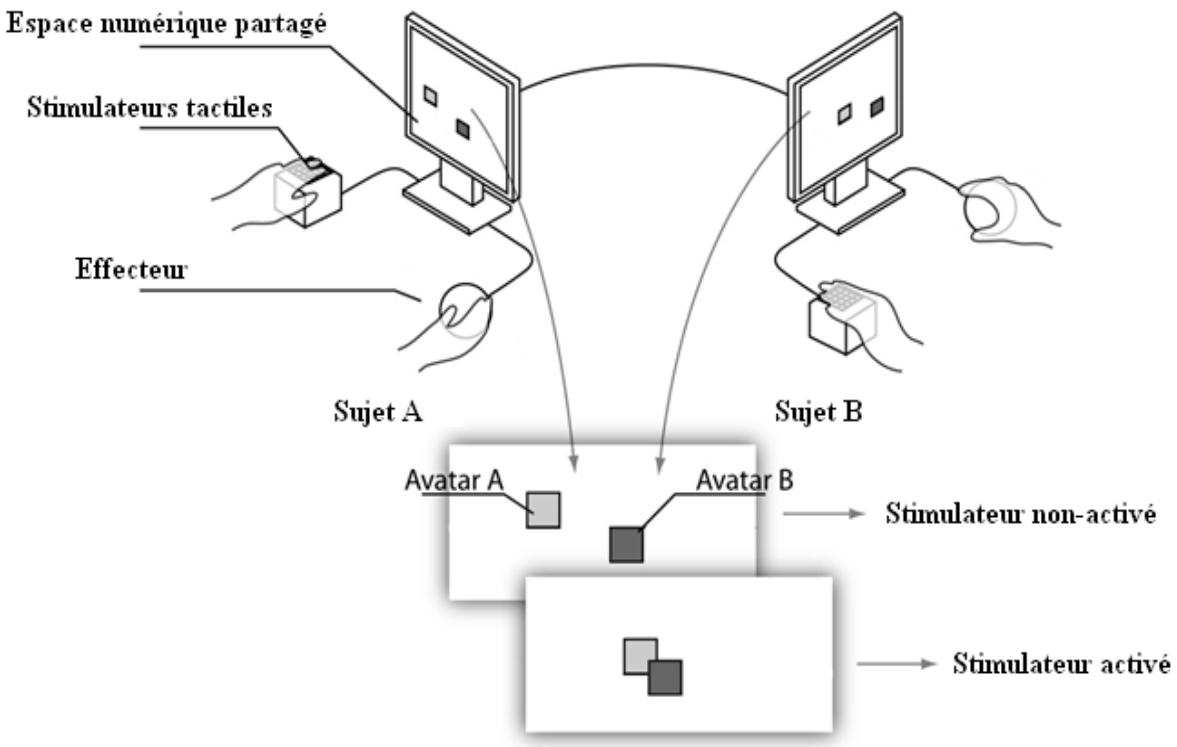


Figure 15 : Le dispositif Intertact. Deux systèmes Tactos connectés en réseau permettent l'exploration d'un espace numérique partagé et la rencontre d'utilisateurs distants

Les utilisateurs peuvent interagir de façon tactile : il s'agit du dispositif Intertact. En d'autres termes, nous rendons possible le passage d'un système à usage strictement local (navigation dans les interfaces graphiques et exploration des objets numériques de façon individuelle) à un système où l'usage s'ouvre sur le partage d'espaces tactiles numériques dédiés à des applications pratiques, pédagogiques et ludiques (Deschamps et al., 2008 ; Deschamps et al., 2009)[9].

Pour rendre possible l'interaction tactile entre des utilisateurs distants, un « corps-image a été ajouté au curseur déplacé par chaque participant. Ainsi, par le biais de son effecteur, chaque utilisateur commande à la fois : i) les mouvements d'un « corps-percevant », soit une matrice de champs récepteurs couplée à autant de stimulateurs tactiles ; et ii) les déplacements d'un « corps-image », c'est-à-dire d'un corps virtuel (avatar) que les autres utilisateurs peuvent percevoir via leurs propres champs récepteurs. De cette façon, chaque utilisateur peut entrer en contact, soit avec des objets présents dans l'environnement partagé, soit avec le corps-image (avatar) d'un autre utilisateur. On permet ainsi la constitution, pour chaque

[9] Notons que ce dispositif a également été présenté au moyen d'une communication affichée :
- Deschamps, L., Lenay, C. & Rovira, K., (2009), Intertact : un dispositif de suppléance perceptive sur le réseau destiné aux interactions tactiles distales, *ARCo-09, Interprétation et problématiques du sens*, 9-11 décembre 2009, Rouen, France (Poster)

utilisateur, d'un espace commun où peuvent se coordonner les activités perceptives. Dans la mesure où, quand un participant « touche » l'autre, il est en même temps « touché » par lui, on peut parler d'une forme de croisement perceptif tactile à travers le réseau (Deschamps et al., 2008 ; Lenay, 2010).

Le projet Intertact a pour objectif de mettre à disposition un serveur dédié rendant possible la connexion simultanée d'utilisateurs distants dans des espaces numériques partagés. S'il demeure pour l'instant à un stade de prototype, non encore accessible sur le World Wide Web, il donnera naissance à terme à un portail Internet tactile permettant l'hébergement de contenus numériques accessibles, l'exploration et le partage de ces informations, et plus généralement, l'interaction tactile entre des utilisateurs distants. Ce portail Internet correspondra à un site web tactile répondant aux critères W3C/WAI[10], de façon à ce que les utilisateurs non-voyants non équipés du dispositif Tactos puissent accéder à une description du contenu, et notamment à la description du dispositif lui-même. De plus, de façon à tirer parti de l'accès à la spatialisation de l'information pour les utilisateurs équipés de Tactos, la navigation à l'intérieur du portail sera favorisée via une organisation spatiale de l'information inspirée de la méthode du « Treemapping »[11] développé via la plateforme Java©. Ce choix de conception découle directement des analyses d'usage du dispositif Tactos, notamment celles qui ont été conduites avec l'utilisateur expert (voir Partie 1, section 1.3.3.1, p. 31). Rappelons que si ce dernier est à ce jour l'utilisateur le plus performant avec le dispositif Tactos, il est aussi un utilisateur avisé de l'informatique en général, ce qui lui a permis de formuler des attentes précises en fonction des limitations auxquelles il est confronté quotidiennement. De cette façon, avec sa collaboration, nous avons défini que pour garantir une navigation efficace dans un tel portail via le dispositif Tactos, l'ensemble des informations devait être organisé spatialement et hiérarchiquement sur l'espace de l'écran, via des « salles de navigation » dédiées (voir Figure 16).

[10] Le W3C (World Wide Web Consortium) est un organisme international visant la normalisation du Web. Un de ses premiers objectifs est de rendre l'Internet accessible à tous, et ce quels que soient leur dispositif informatique, leurs logiciels, l'infrastructure de leur réseau, mais aussi leur langage, leur culture, leur localisation géographique ou leur habileté physique ou mentale. Le WAI (Web Accessibility Initiative) est un projet mené par le W3C qui promeut notamment un haut degré d'accessibilité du Web pour les personnes handicapées, à travers le développement technique de l'informatique et l'élaboration de recommandations ergonomiques pour les développeurs et les éditeurs de contenus.

[11] Le treemapping est une technique de représentation de données créée par Ben Shneiderman en 1990. Il s'agit d'organiser spatialement et hiérarchiquement les données numériques via des zones délimitées et activables. (source : http://fr.wikipedia.org/wiki/Treemapping, consulté en septembre 2012)

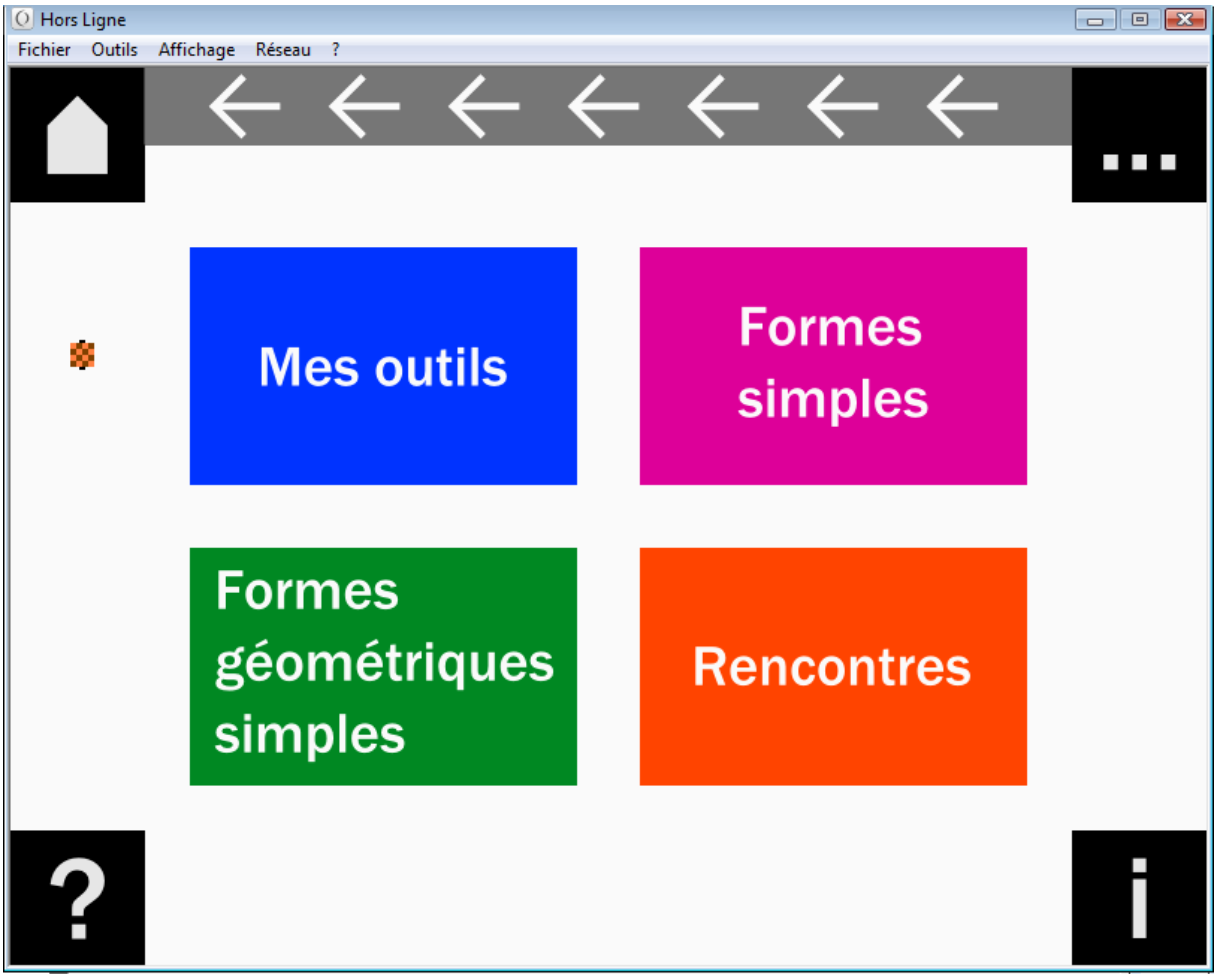


Figure 16 : Exemple de salle de navigation du portail Intertact, dont l'organisation spatiale est inspirée de la méthode du Treemapping.

Dans les salles de navigation, l'information est présentée sous formes de cadres hiérarchiques représentant une certaine catégorie de contenus. Chaque cadre est en réalité une « porte », soit un lien activable qui permet d'accéder à une salle de navigation d'un niveau de profondeur supérieur ou à une salle de contenu, qui correspond toujours au niveau hiérarchique le plus profond pour chaque catégorie définie par les portes correspondantes (par exemple, l'image d'un carré sera accessible via la porte « Encyclopédie », puis « Formes simples », puis « Carré »). De cette façon, chaque ressource accessible sur le portail correspondra à une « salle de contenu » dédiée, spatialement déterminée dans l'espace du portail. Chaque salle de contenu est alors définie par un contenu numérique explorable individuellement via le système Tactos, ou collectivement via le réseau (voir Figure 17).

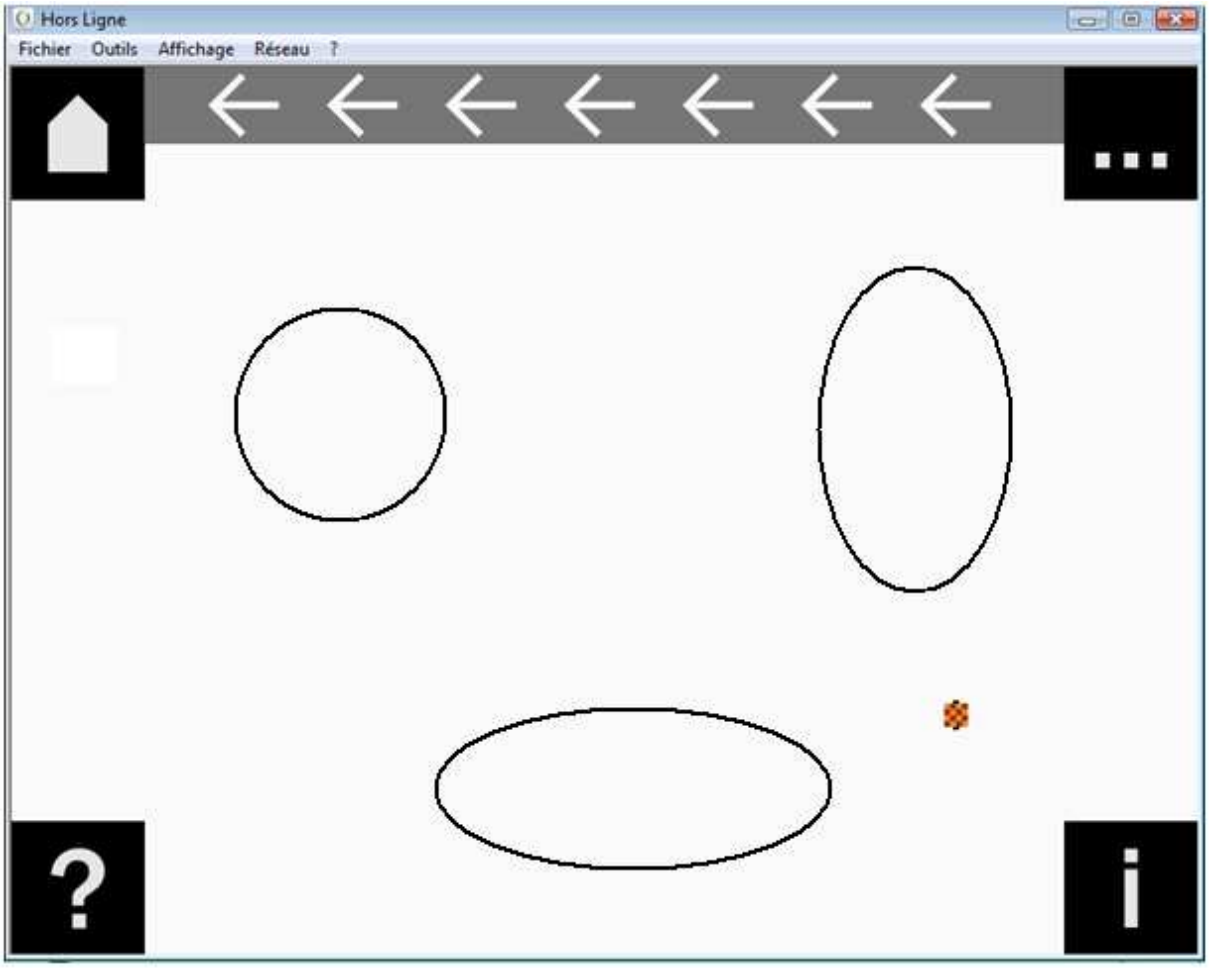


Figure 17 : Exemple de salle de contenu du portail Intertact, ici « Cercles et Ellipses »

D'un point de vue technique, les salles de contenu sont des ressources éditées en xml[12], qui permettent d'associer aux images explorables différents types d'informations situées, qu'elles soient tactiles (pixels de couleur détectables via le corps-percevant ou tact-icônes) et/ou sonores (informations auditives ou verbales). Par ailleurs, ces contenus peuvent être enrichis via des scripts LUA[13], de façon à délivrer dynamiquement ces informations en fonction de différentes actions-utilisateur (par exemple, entrer dans une zone de l'image avec le curseur ou presser un des boutons de l'effecteur au contact d'une source de stimulation).

Notons que pour faciliter la navigation et l'exploration des contenus, une structure commune est surimposée à chacune des salles du portail, qu'elles soient de navigation ou de contenu. Il s'agit des quatre carrés disposés aux quatre coins de l'écran, ainsi que la bande supérieure de l'écran (voir Figure 16 et Figure 17) :

[12] Le xml (Extensible Markup Language) est un langage informatique générique
[13] Le LUA est un langage de script libre et compact, écrit en C ANSI strict, et compilable sur une grande variété de systèmes informatiques

- Le carré supérieur gauche est une porte qui permet de revenir à la salle d'accueil du portail.
- Le carré supérieur droit fournit le titre de la salle actuellement explorée
- Le carré inférieur droit fournit une description du contenu de la salle explorée
- Le carré inférieur gauche fournit une aide contextuelle à la salle explorée (Par exemple, le nombre de figures présentes ainsi que leur localisation exacte).
- La bande supérieure permet de revenir à la salle de navigation précédente

De cette façon, il est possible d'élaborer un ensemble de contenus numériques hébergés sur le portail et mettant en œuvre un ensemble de fonctionnalités nouvelles, rendues possibles par l'accès à la spatialisation de l'information couplée aux possibilités d'interactions tactiles distales : exploration individuelle ou collective de données tactilo-sonores ; espaces de rencontre et de chat tactile ; jeux d'interaction etc. (voir Figure 18).

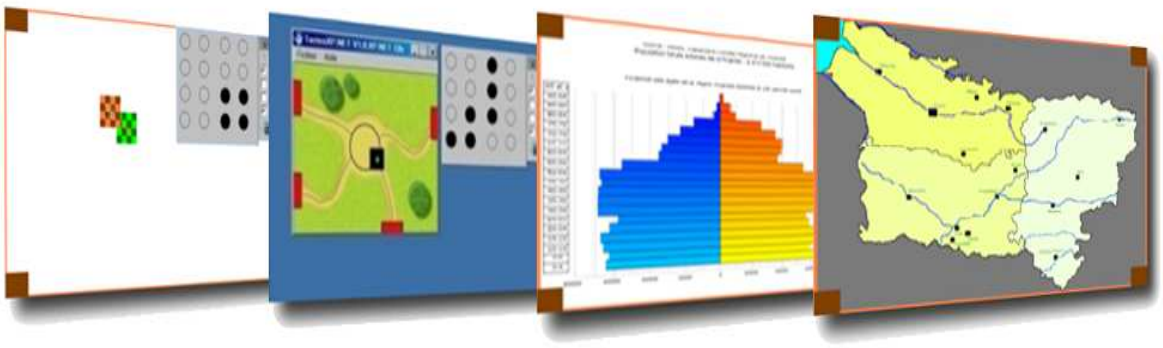


Figure 18 : Exemples d'espaces numériques partagés, hébergés sur le portail Intertact.

Notons que si ce dispositif est destiné en premier lieu aux personnes ayant un déficit visuel pour l'accessibilité partagée des contenus informatiques, il sera accessible de la même façon aux personnes voyantes. Ainsi, la question de l'accessibilité des contenus et des applications est centrale, mais n'est pas destinée à « isoler » les utilisateurs déficients visuels dans une partie dédiée du web. Il est plutôt question de constituer une communauté d'usage qui dépasse le cadre du handicap. Nous inscrivons donc ce projet dans le principe d'un « design for all ».

2.2.2. Prise en main du module d'interaction tactile

Dans le cadre des suivis longitudinaux des utilisateurs déficients visuels, nous avons procédé à des observations concernant la prise en main et l'utilisabilité du premier prototype

de module d'interaction tactile (MIT). Comme lors des études de faisabilité, ces observations ont été conduites avec **C.** (l'utilisateur expert), et **Y.**, (l'utilisateur novice). Lors d'une séance de suivi, un MIT est déposé sur une table devant l'utilisateur, et il lui est demandé de le saisir, de l'explorer, puis de le tenir de la façon qui lui semble la plus adéquate dans le cadre d'un usage de Tactos. Lors de la séance, nous demandons aux sujets de commenter cette saisie, de façon décrire leurs impressions générales quant au module.

Après une première manipulation, qui a consisté pour les deux utilisateurs à explorer le volume global et les différents composants du MIT, il apparaît que la tenue préférentielle spontanée est identique pour les deux sujets : la tenue du module est bi-manuelle, avec un soutien global effectué par la main non-dominante. Il semble que la présence de la connectique (deux câbles USB destinés à l'alimentation en énergie et au contrôle logiciel des composants) influence la saisie pour une orientation verticale du module, les câbles étant dirigés vers le bas. L'index de la main dominante est posé sur le Touchpad, tandis que l'index de la main non-dominante est posé sur la matrice de stimulateurs tactiles situé à gauche du module (voir Figure 19).

Figure 19 : Prise en main préférentielle spontanée du module d'interaction tactile : la saisie est bi-manuelle, avec l'index de la main dominante sur le Touchpad, et l'index de la main non-dominante sur une des matrices de stimulateurs tactiles.

Suite à cette première saisie spontanée, nous demandons aux utilisateurs de tenter de faire varier la position de leurs mains, de façon à explorer et à commenter les différentes prises en main possibles. Nous accompagnons alors cette manipulation, en discutant avec les sujets des différents composants qu'il s'agirait d'atteindre dans le cadre d'un usage effectif du

module. Comme nous l'attendions, il apparaît que le MIT peut effectivement être tenu de multiples façons, sans nécessairement compromettre le confort de la prise en main et l'utilisabilité (voir Figure 20 pour des exemples) :

- Globalement, la saisie peut être bi-manuelle ou uni-manuelle.
- Le pouce, l'index et/ou le majeur peuvent être mobilisés selon les cas sur l'effecteur ou sur les matrices de stimulateurs tactiles.
- Les deux plages de cellules Braille peuvent être atteintes simultanément ou non[14].
- Le module peut être tenu verticalement ou horizontalement.
- Le module peut être tenu dans toutes les orientations possibles, la connectique étant orientée selon les cas à gauche, à droite, en haut ou en bas.

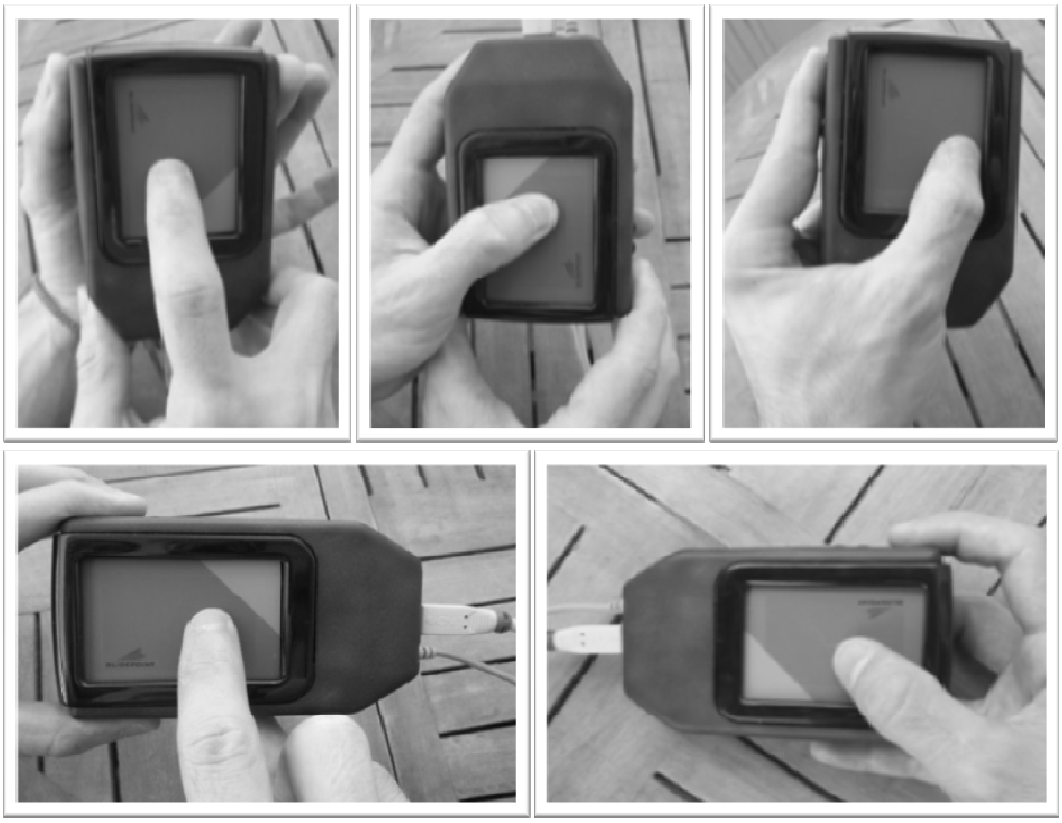


Figure 20 : Exemples de prises en main alternatives du MIT

[14] L'accès à la deuxième plage de cellule Braille pourrait en effet s'avérer intéressante pour enrichir le signal tactile résultant de l'exploration des formes numériques, ou encore pour délivrer des informations complémentaires aux informations tactiles résultant de l'exploration de ces formes (tact-icônes, lettres en Braille etc.). Toutefois, nous pouvons noter que l'accès simultané aux deux plages de cellules Braille pose problème pour le plus jeune des utilisateurs. En effet, le volume du MIT est une contrainte pour l'envergure de la main de l'adolescent, qui peut atteindre la deuxième plage de cellules Braille mais au prix d'une tenue moins confortable.

Spontanément, les sujets ont tenté différentes positions de façon à atteindre préférentiellement le Touchpad, qu'ils avaient déjà testé lors d'analyses d'usage antérieures. Nous attirons leur attention sur la deuxième face du module, qui comprend la nappe de capteurs de pression (FSR). Après leur avoir expliqué l'intérêt de cet effecteur (déformation dynamique du corps-image, mode « caresse tactile »), nous leur demandons à nouveau de prendre en main le module de façon à pouvoir l'utiliser. Il apparaît alors que les tenues préférentielles impliquent une alternance entre un mode « déplacement » (doigt sur le Touchpad) et un mode « caresse » (doigt sur le FSR). Pour cela, le module est pivoté dans l'espace, et nous retrouvons les mêmes principes d'utilisation que précédemment (Bi-manualité ou uni-manualité ; atteinte d'une matrice de stimulateurs ou des deux simultanément ; pouce ou index sur l'effecteur selon le nombre de mains impliquées ; tenue verticale ou horizontale). Nous leur suggérons alors qu'il serait intéressant de pouvoir utiliser conjointement les deux faces du module (se déplacer dans l'espace tout en déformant le corps-image de façon dynamique). Là encore, le volume du module pose problème pour le plus jeune des utilisateurs. Pour l'utilisateur expert, l'accès simultané semble possible, même si ce dernier attire notre attention sur la difficulté présumée d'utiliser conjointement les deux effecteurs. La tenue est alors bi-manuelle, la main non-dominante servant de support global avec l'index sur la matrice de gauche, tandis que la main dominante englobe le module, le pouce étant disposé sur le Touchpad tandis que l'index et le majeur atteignent le FSR (voir Figure 21).

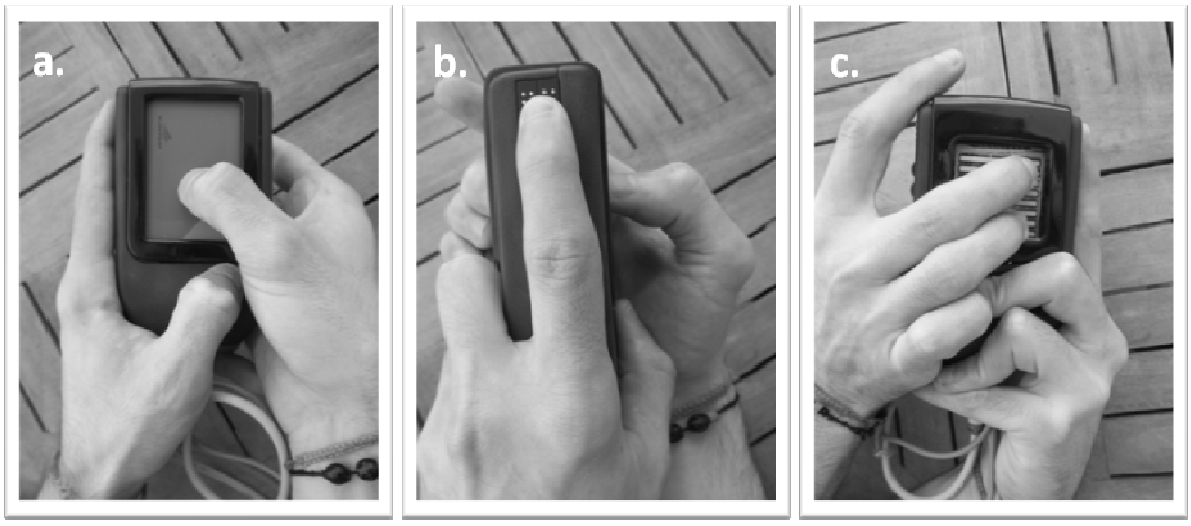


Figure 21 : Prise en main bi-manuelle du module pour l'accès simultané aux deux effecteurs. La Figure 22a présente la vue de face du module ; la Figure 22b présente la vue de côté ; La figure 22c présente la vue de la face arrière du module

2.2.3. Analyses d'usage du module d'interaction tactile

Malgré une prise en main préférentielle identifiée lors de l'étape précédente, nous ne posons pas de conclusion définitive sur la façon optimale de saisir et de manipuler le MIT. En effet, cette analyse est strictement qualitative, et la prise en main spontanée pourrait dépendre de préférences individuelles que nous ne retrouverons pas nécessairement chez d'autres utilisateurs. Dans cette optique, nous avons pu constater des modes de saisie alternatifs, ce qui souligne une certaine flexibilité du module rendant possible une prise en main adaptable selon les utilisateurs et/ou selon la tâche. Toutefois, les différents modes de saisie soulèvent des questions d'usage, qu'il s'agit d'explorer ici.

Dans un premier temps, il est important de noter que le mode de saisie du MIT par l'utilisateur peut induire une configuration logicielle préalable. Par défaut, le Touchpad est un effecteur orienté horizontalement. De cette façon, il est nécessaire de configurer de façon logicielle l'orientation de la surface d'action en fonction de l'orientation physique du Touchpad dans la main. De la même façon, les matrices de stimulateurs tactiles ont une orientation par défaut, qu'il s'agit de configurer si la prise en main l'exige. Ainsi, la rotation du Touchpad et la rotation des cellules doivent être prises en compte avant l'usage. Dans les tests qui vont suivre, ces manipulations logicielles, accessibles via les menus de l'interface Tactos, ont été effectuées par nos soins.

Tout d'abord, nous proposons aux sujets une tâche de reconnaissance de figures géométriques simples avec la tenue préférentielle préalablement identifiée. Si après usage, les deux utilisateurs se déclarent satisfaits du MIT, des limitations sont tout de même observées quant à leur performance. En effet, la tenue verticale du Touchpad a des implications sur le format des formes explorées, l'espace d'exploration étant par défaut orienté horizontalement. En conséquence, les déplacements de l'index sur le Touchpad tenu verticalement ne produisent pas des déplacements strictement analogiques du curseur dans l'espace absolu de l'écran. Dans cette configuration, les déplacements horizontaux sont en effet accélérés dans l'espace de Tactos, alors que les déplacements verticaux sont ralentis. En conséquence, les figures explorées sont déformées, les segments horizontaux paraissant plus courts et les segments verticaux plus longs qu'ils ne le sont « en réalité », c'est-à-dire ici à l'écran. Par exemple, un carré est perçu comme un rectangle, un cercle comme une ellipse.

De façon à éviter ces déformations, plusieurs solutions sont envisagées. Premièrement, il est possible de délimiter une surface active sur le Touchpad par l'intermédiaire d'un module de calibration dédié. La calibration de cette surface peut être effectuée par l'intermédiaire

d'un mouvement du doigt sur le Touchpad, ou en choisissant parmi des formats pré-enregistrés. De cette façon, l'utilisateur peut définir une surface active rectangulaire orientée horizontalement, de façon à obtenir les mêmes dimensions relatives que la surface de l'écran. Lors des essais, cette solution ne s'est pas avérée satisfaisante, et ce pour plusieurs raisons. D'une part, le processus de calibration n'est pas facilement accessible par les utilisateurs eux-mêmes (une navigation dans les menus est nécessaire, ce qui ne favorise pas un usage autonome pour les personnes peu à l'aise avec l'informatique). D'autre part, la surface active résultante est alors trop petite, ce qui oblige l'utilisateur à procéder à des micro-mouvements impliquant un contrôle plus important des gestes d'exploration. De plus, dans la mesure où les zones inactives du Touchpad ne sont pas signalées lors de l'exploration, l'utilisateur peut déplacer son doigt hors de la zone active sans s'en rendre compte. Dans ce cas, le caractère absolu de l'effecteur n'est pas garanti.

Une solution plus simple consiste alors à modifier la configuration de l'espace Tactos par défaut, de façon à passer d'un format rectangulaire horizontal à un format rectangulaire vertical. Dans ce cas, les performances des utilisateurs sont en effet satisfaisantes de par l'absence de déformation et la taille acceptable de la surface d'action du Touchpad.

Enfin, il est important de noter que la position des matrices de stimulateurs tactiles par rapport au Touchpad ne semble pas poser de problème pour les utilisateurs. En effet, avec l'effecteur classique de Tactos (tablette/stylet), la matrice de stimulateurs tactiles est posée sur une table sur le même plan géométrique que la surface d'action proposée à l'utilisateur (soit un plan transversal). Sur le module d'interaction tactile, les matrices sont situées perpendiculairement au Touchpad. Ce changement de plan ne semble pas avoir eu d'incidence sur la compréhension des stimulations tactiles reçues par les utilisateurs, les figures ayant pu être globalement identifiées quel que soit le niveau d'expertise des sujets.

Concernant les autres modes de saisie, les performances des utilisateurs sont plus mitigées. Tout d'abord, avec une prise en main uni-manuelle du MIT en position verticale, les performances de reconnaissance de l'utilisateur expert ne semblent pas affectées, même si l'exploration du Touchpad avec le pouce semble moins fluide et assurée. En revanche, l'utilisateur novice rencontre des difficultés, qui semblent essentiellement dues à la taille du module : l'envergure limitée de la main de l'adolescent empêche une atteinte aisée de l'ensemble de la surface d'action via le pouce. Enfin, pour les prises en main horizontales, tout comme pour la prise en main verticale avec la connectique orientée vers le haut, la question de la rotation des cellules s'avère être centrale. En effet, si le fait que les cellules ne soient pas sur le même plan que l'effecteur n'a pas gêné les deux utilisateurs lors de la saisie

verticale (connectique vers le bas), la rotation du MIT dans l'espace nécessite de paramétrer en conséquence l'orientation des cellules. Par exemple, si le MIT est tenu horizontalement avec la connectique à droite, croiser une ligne horizontale en déplaçant son curseur du bas vers le haut a pour effet d'activer les cellules de la gauche vers la droite. En conséquence, il paraît nécessaire d'intervenir de façon logicielle pour indiquer au système dans quel sens est tenu le module, de façon à ce que les cellules s'activent de façon analogique au sens de déplacement. Mais malgré ce paramétrage, les deux utilisateurs présentent des performances différentes. Si l'utilisateur expert est capable de s'adapter très rapidement et de reconnaître facilement les formes explorées, l'utilisateur novice présente de grandes difficultés. Il semble que l'orientation du doigt sur les cellules est la source de ces difficultés. Lorsque le module est saisi horizontalement, par exemple, le doigt se trouve orienté perpendiculairement aux stimulateurs, ce qui n'est pas le cas des saisies verticales, où le doigt est aligné avec les cellules. Là encore, il est difficile d'attribuer cette différence au seul niveau d'expertise : la différence d'âge pourrait entrer en jeu, la rotation du MIT dans l'espace bouleversant le cadre de référence spatial auquel le sujet se réfère.

2.2.4. Conclusion de l'analyse d'usage du module d'interaction tactile

Nous retenons de cette première analyse d'usage que ce prototype de module d'interaction tactile permet de multiples modes de saisie, ce qui en fait un module particulièrement flexible pour les utilisateurs, adaptable selon les préférences individuelles. Toutefois, cette flexibilité va de paire avec des limitations spécifiques que nous venons d'identifier. En effet, si la prise en main préférentielle a rendu possible des performances d'identification satisfaisantes chez nos deux utilisateurs, la modification de la saisie a entraîné des problèmes d'usage, et ce particulièrement chez l'utilisateur novice.

La première limitation concerne la taille du module, qui ne semble pas adaptée à toutes les mains, notamment celles des plus jeunes. Cette constatation est plutôt gênante, dans la mesure où la population des adolescents déficients visuels fait partie du public visé en priorité, le dispositif Tactos ayant une visée pédagogique pour l'accès à l'information graphique numérique. De plus, les interactions tactiles distales pourraient améliorer les relations numériques quand les plateformes existantes ne sont pas forcément accessibles[15]. Les prochains prototypes devront alors tenter de miniaturiser le volume global, de façon à

[15] Dans cette optique, nous pouvons noter des problèmes de compatibilité entre les applications de type Messenger ou les réseaux sociaux numériques et les logiciels de synthèse vocale. Si ces problèmes finissent par être résolus par les éditeurs ou la communauté des utilisateurs, la compatibilité est constamment remise en cause par les mises à jour de ces applications.

favoriser la prise en main du module par les plus jeunes. Toutefois, cette miniaturisation ne doit pas concerner la surface allouée au Touchpad : si celui-ci est trop petit, cela semble avoir une incidence négative sur les performances d'exploration et de reconnaissance.

Une seconde limitation provient de la nécessité d'intervenir sur le paramétrage du module de façon logicielle, de façon à adapter son fonctionnement relativement à la prise en main de l'utilisateur. En effet, si son design permet effectivement de multiples modes de saisie, chacune des prises en main nécessite une adaptation logicielle pour garantir un usage efficient. Or, le gain de flexibilité ne doit pas conduire à une augmentation de la complexité, induite par la nécessité de naviguer dans les menus pour que la configuration du module soit conforme à la prise en main. Au contraire, de façon à ce que cette flexibilité soit transparente pour l'utilisateur, qui pourrait modifier sa saisie en fonction de la tâche ou des difficultés rencontrées, il semble nécessaire de réfléchir à une automatisation du paramétrage en fonction de l'orientation du module dans l'espace. Dans cette optique, pour les prochaines versions du module, l'implémentation d'un gyroscope ou d'un accéléromètre est à l'étude. Ceux-ci permettraient en effet une détection automatique de l'orientation du module, de façon à adapter de façon dynamique les paramètres de configuration, que ce soit l'orientation du Touchpad ou celle des matrices de stimulateurs.

Enfin, si des tests à plus long terme pourront préciser l'intérêt de ce premier prototype de module d'interaction tactile, et rendre compte d'une éventuelle appropriation par les utilisateurs, il nous semble maintenant nécessaire de procéder à une évaluation de son utilisabilité dans le cadre des interactions tactiles distales.

3. Etudes préliminaires concernant les interactions tactiles

3.1. Observations préliminaires

Dans un premier temps, des observations préliminaires de l'utilisabilité du module d'interaction tactile ont été conduites. Il s'agit là la suite logique du travail effectué lors de la conception et de l'évaluation corrective de Tactos, qui a permis de définir avec précision les modalités techniques et logicielles pour garantir l'ergonomie du système. Ici, nous avons mobilisé notre utilisateur expert, **C.** (voir Partie 1, section 1.3.3.1, p. 31), dans des situations d'interaction avec un sujet novice – en l'occurrence l'expérimentateur avec les yeux bandés – via une connexion locale entre deux systèmes Tactos. Cette situation nous paraît intéressante pour plusieurs raisons :

- Premièrement, le fait de prendre part activement à ces situations permet de mener une analyse en première personne de l'utilisabilité de notre système. De plus, l'engagement dans les tâches a rendu possible une collaboration active avec l'utilisateur expert, en discutant du système et en commentant nos activités tout en effectuant les différentes tâches. Toutefois, lors de celles-ci, la consigne était de privilégier les interactions tactiles, et d'éviter de se donner des instructions verbales directement liées à la tâche.
- Deuxièmement, le fait de faire intervenir un utilisateur expert nous permet de recueillir un ensemble de données qualitatives particulièrement intéressantes sur la situation d'interaction et sur les possibilités offertes par le dispositif. En effet, l'utilisateur expert nous semble être le mieux placé pour juger de la pertinence de certaines fonctionnalités, ainsi que pour exprimer ses attentes, ses points de satisfaction comme ses sources de difficultés.
- Enfin, il s'agit d'une occasion intéressante de mettre à l'épreuve le principe des interactions tactiles distales, en étant notamment sensible à la qualité de l'interaction (par exemple, dans une situation de guidage dans l'espace) et à un éventuel transfert de compétence entre l'utilisateur expert et l'utilisateur novice. En outre, ces situations offrent un terrain propice à l'observation qualitative des stratégies mises en œuvre par les utilisateurs, parallèlement à l'exploration des conditions techniques nécessaires et suffisantes pour permettre une interaction signifiante.

Deux scénarii d'usage ont été élaborés :

- Le ***scénario d'exploration*** consiste pour l'utilisateur expert à m'assister lors de l'exploration de formes géométriques simples (par exemple, un rectangle ou un triangle). Ces figures ont été constituées lors des analyses d'usage du dispositif Tactos (voir Partie 1, section 1.3, p. 27). De façon à ce que nous ne puissions pas savoir à l'avance quelle figure est présente dans l'espace, un tiers est chargé de les sélectionner et de les afficher à ma demande.
- Le ***scénario de guidage*** consiste pour l'utilisateur expert à me guider dans un ensemble de lignes géométriques enchevêtrées (voir Figure 22). Pour cela, l'utilisateur expert a pour consigne de me guider dans l'espace tout en parcourant la forme, et de choisir la direction à suivre lorsque nous arrivons aux différents embranchements.

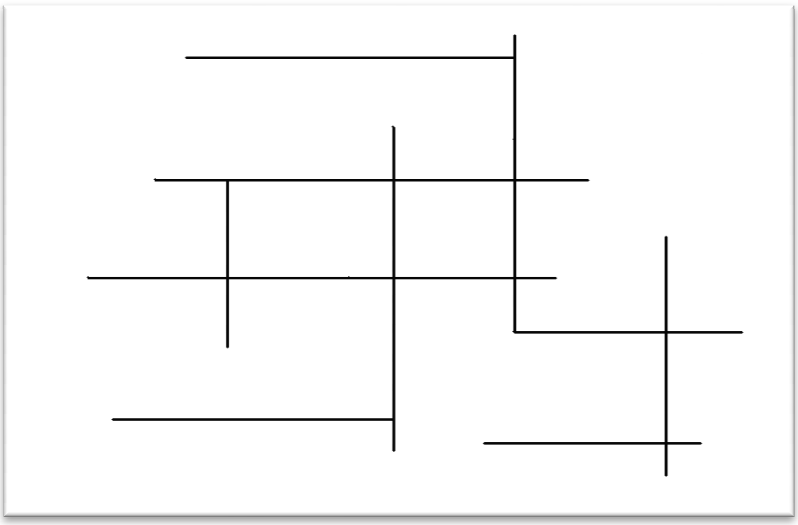


Figure 22 : Exemple de parcours utilisé lors du scénario de guidage.

Dans ces deux scénarii, la taille des lignes utilisées pour chaque image est de 2 pixels, et l'espace total d'exploration est de 800 x 600 pixels.

L'objectif de ces analyses d'usage en situation écologique vise à analyser les interactions tactiles entre C. et moi-même, tout en faisant varier les propriétés du dispositif (en l'occurrence les propriétés des « corps-percevant » et des « corps-image » respectifs) de façon à commenter et à discuter des effets que ces variations produisent sur nos interactions. Pour faciliter l'analyse, l'ensemble des situations ont été filmées.

3.1.1. Configurations symétriques

Premièrement, nous avons opté pour une configuration symétrique entre les deux utilisateurs : les « corps-percevant » et les « corps-image » sont identiques, à savoir que chacun déplace une matrice de 4x4 champs récepteurs élémentaires de 2x2 pixels chacun, soit une surface sur l'écran de 8x8 pixels, auquel correspond topologiquement une surface carrée de 8x8 pixels perceptibles par les champs récepteurs du partenaire (voir Figure 23).

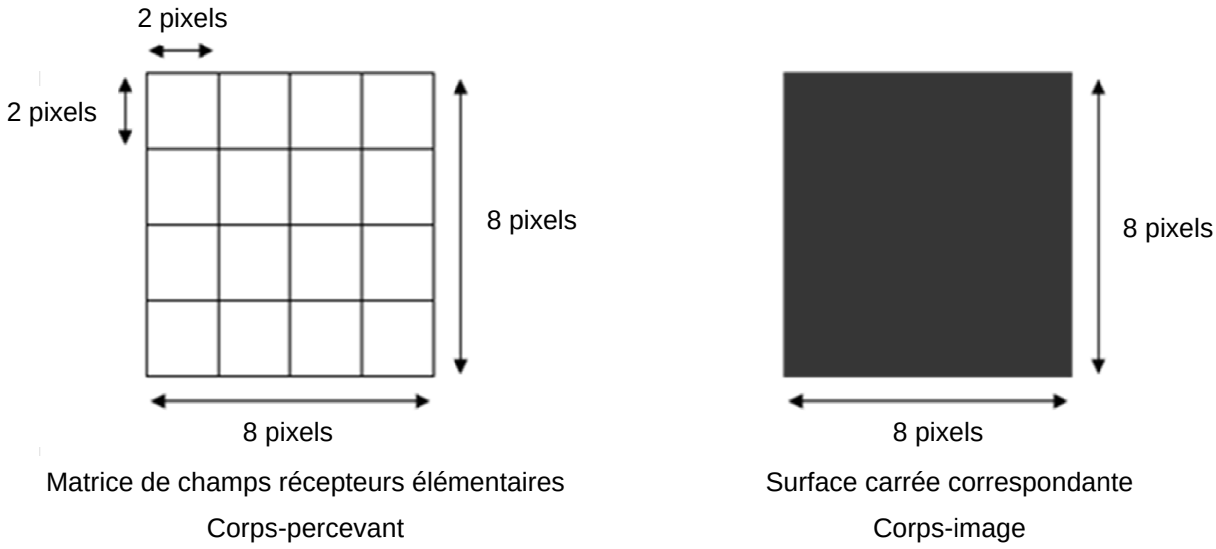


Figure 23 : Propriétés des "corps-percevant" et des "corps-image" des utilisateurs. La matrice de champs récepteurs et la surface carrée correspondante sont "superposées" pour chaque utilisateur

Avec cette configuration, la probabilité de se rencontrer dans l'espace est relativement forte. En effet, puisque chacun déplace un corps-image de 8x8 pixels (soit 64 pixels de surface) sur lequel est calqué le corps-percevant, un seul pixel en commun parmi une surface de 63x63 pixels au total (soit la surface cumulée du corps-image d'un utilisateur et du corps-percevant de l'autre utilisateur moins un pixel en commun) permet un toucher réciproque. De cette façon, dans un espace de 800x600 pixels, nous avons pu nous rencontrer facilement, même si l'utilisation de points de rendez-vous s'est avérée parfois nécessaire[16]. Une fois le contact établi, il a été également possible de se suivre dans l'espace, l'objectif pour **C.** étant de me guider vers la forme à explorer lors du premier scénario d'usage. Dans cette étape de suivi, l'expertise de **C.** dans l'utilisation de Tactos s'est révélée être un atout indispensable. En effet, la rapidité et la précision de ses mouvements lui ont permis de localiser la forme très facilement, et d'aller et venir entre celle-ci et mon corps-image sans être désorienté dans l'espace et sans perdre le contact suffisamment longtemps pour que je me sente moi-même perdu. Ainsi, **C.** a pu prendre connaissance de la direction et de la distance de mon corps-image par rapport à la forme à explorer, en essayant de m'y entrainer progressivement lors des phases de toucher réciproque.

En revanche, lorsque la forme a été atteinte, des difficultés sont survenues. Si **C.** a pu explorer très rapidement la forme pour l'identifier, j'ai pour ma part éprouvé de grandes difficultés pour faire la différence entre les stimuli résultant de contacts avec la forme des stimuli résultant de contacts avec le corps-image de **C.** De plus, dans les cas où nos corps-

[16] Par exemple, un point de rendez-vous consistait à se retrouver dans un des coins de l'espace d'exploration via une instruction verbale.

images sont superposés, la forme est « masquée » (les lignes constituant la forme étant de taille inférieure à celle du corps-image du partenaire). Malgré des tentatives de la part de **C.** d'anticiper mes mouvements d'exploration pour me devancer sur la forme, chaque contact avec son corps-image s'est avéré déstabilisant et m'ont, la plupart du temps, induit en erreur, la conséquence étant une perte de la forme à explorer. Avec cette configuration, je n'ai pu identifier qu'une figure sur les trois explorées, mais au prix d'une faible proportion d'interaction. Dans l'incapacité de suivre le corps-image de **C.** tout en maintenant le contact avec la forme, j'ai alors exploré cette forme individuellement alors que je recherchais **C.** dans l'espace.

Lors du scénario de guidage, les mêmes difficultés sont survenues. Il ne m'a pas été possible de suivre efficacement le corps-image de **C.**, particulièrement lorsque celui-ci choisissait une direction à notre arrivée sur un embranchement. Malgré mes difficultés, il s'est avéré que **C.** était capable pour sa part d'explorer la forme tout en essayant de garder le contact avec moi, ou du moins tout en ayant conscience de ma localisation dans l'espace. Ainsi, le niveau d'expertise de **C.** semble pouvoir lui permettre de réussir ce type de tâche.

Pour diminuer l'effet d'occlusion résultant de la taille supérieure des corps-images par rapport à la taille des lignes à explorer, nous avons opté pour une réduction de la taille des corps-images tout en conservant la taille d'origine pour les corps-percevant (voir Figure 24).

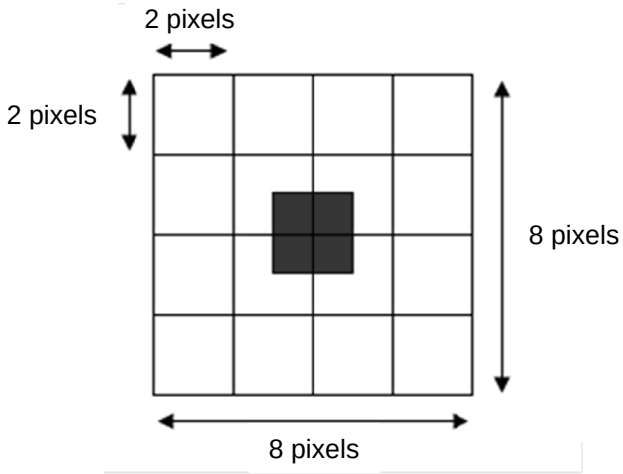


Figure 24 : Matrice des champs récepteurs élémentaires et corps-image (en noir au centre) de 2x2 pixels situé au centre de la matrice

Avec cette configuration, tout contact entre le corps-percevant d'un utilisateur et le corps-image de l'autre utilisateur implique un toucher réciproque, tout comme la configuration précédente. Toutefois, dans ce cas, la probabilité de se rencontrer dans l'espace

est réduite. En effet, la surface totale de toucher réciproque est maintenant de 63x3 pixels, soit la surface cumulée du corps-percevant d'un utilisateur et la surface du corps-image de l'autre utilisateur moins un pixel en commun. De cette façon, la rencontre dans l'espace a nécessité davantage d'instructions verbales, notamment pour convenir de points de rendez-vous. Une fois le contact établi, la qualité du guidage vers la forme à explorer n'a pas été affectée. En dehors de toute stimulation due à la rencontre de la forme, il a en effet été possible de suivre **C.** en se basant sur les stimulations résultant de notre rencontre. Comme lors de la situation précédente, la stratégie mise en œuvre par **C.** a consisté à aller et venir successivement et de façon rapide entre la forme à explorer et mon corps-image, de façon à prendre connaissance de la direction et de la distance de celle-ci par rapport à moi. Ainsi, lorsque je me dirige dans la bonne direction, **C.** maintient le contact tout en anticipant mon déplacement : de cette façon, il me devance toujours dans l'espace et ne me quitte que pour contrôler la localisation de la forme.

Une fois la forme rencontrée, la stratégie de **C.** a consisté à anticiper systématiquement mon activité d'exploration de façon à réduire au maximum les interférences entre les stimulations résultant de la forme et les stimulations résultant du toucher réciproque. Des phénomènes d'occlusions perceptives ont eu lieu, mais de façon moins fréquente, et celles-ci ont entraîné moins de pertes de la forme. En effet, compte tenu de la taille du corps-percevant et de la diminution de la taille des corps-image, il est désormais plus facile d'être en contact simultanément avec la forme à explorer et le corps-image du partenaire, ce qui a favorisé pour ma part le maintien du contact avec la forme. Dans ce type d'interaction, nous pouvons isoler deux situations qui nous semblent particulièrement intéressantes :

- Mon corps-percevant est en contact avec la forme à explorer et avec le corps-image de l'expert, mais le corps-percevant de ce dernier n'est pas en contact avec la forme (voir Figure 25). L'expert se focalise sur mes déplacements, tout en m'indiquant une direction en fonction de sa connaissance de la forme à explorer et de sa localisation. Nous pouvons qualifier ce type d'interaction d' « ***exploration supervisée*** ».
- Nos deux corps-percevant sont en contact avec la forme tout en maintenant un toucher réciproque (voir Figure 26). Dans ce cas, nous avons une forme d' « ***exploration mutuelle*** ». En anticipant l'exploration de la forme, l'utilisateur expert contrôle la survenue de la fin de la ligne explorée et peut prévoir l'arrivée d'un changement de direction, tout en contrôlant ma position et la qualité de mon exploration de la forme.

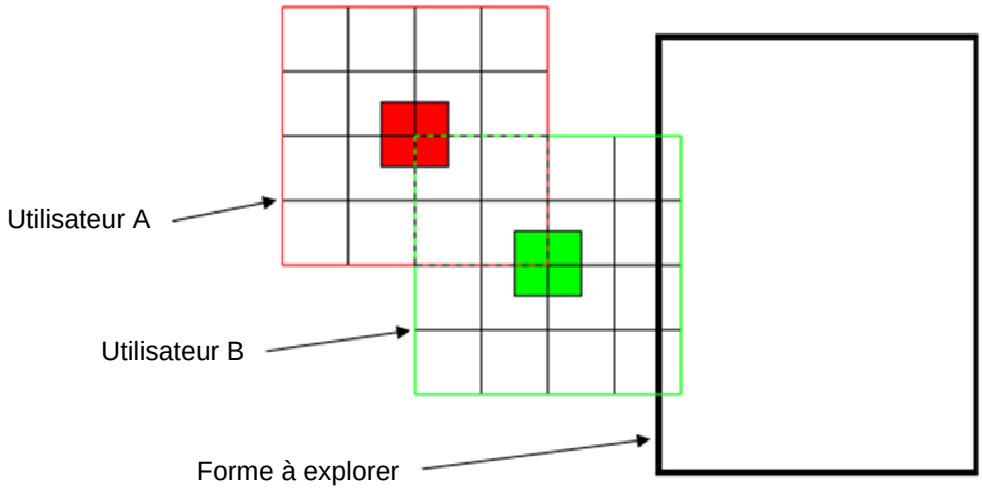


Figure 25 : Situation d' « exploration supervisée » : l'utilisateur A se focalise sur l'activité de l'utilisateur B

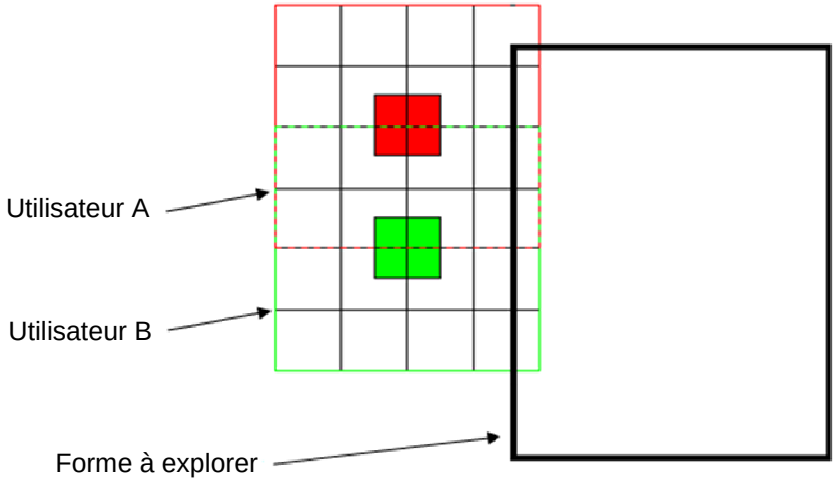


Figure 26 : Situation d' « exploration mutuelle ». Chaque utilisateur perçoit le corps-image du partenaire et la forme à explorer de façon simultanée

De par sa plus grande expertise, **C.** semble contrôler l'alternance dynamique entre ces deux situations : il peut passer successivement de l'une à l'autre selon le contexte de notre interaction, soit pour vérifier la présence de la forme explorée ou pour se focaliser exclusivement sur mon activité d'exploration. Par ailleurs, **C.** ayant compris que le toucher réciproque lors de l'exploration d'une forme était problématique pour moi, sa stratégie a consisté à « tourner » autour de moi, m'indiquant sa présence sans pour autant interférer avec

mon exploration, et intervenant – tactilement ou verbalement – lorsqu'il le jugeait nécessaire. Lorsque l'intervention de **C.** est tactile, sa stratégie consiste à établir un contact puis à tenter de m'entraîner dans une direction donnée, en contrôlant la présence de la forme. De cette façon, **C.** pouvait exercer un guidage de mon activité tout en contrôlant de façon succincte la présence de la forme et de mon corps-image, ainsi que l'adéquation de ma trajectoire en fonction de la forme. De mon point de vue, il était plus simple de se focaliser sur la forme tout en bénéficiant ponctuellement de stimulations tactiles supplémentaires résultant de la présence du corps-image de **C.** dans mon champ perceptif. En cas de doute ou de problème, il me suffisait d'attendre quelques secondes pour que **C.** entre en situation de toucher réciproque, puis qu'il essaie d'amorcer un mouvement dans une direction donnée dans le but de me faire continuer l'exploration de la forme. Ainsi, j'ai été capable de reconnaître les trois formes explorées. Cependant, il est difficile de conclure sur la raison de cette performance. En effet, il semble que la configuration des corps-image/corps percevant ait facilité nos interactions, tout en rendant possible l'exploration de la forme. Mais il est également possible que la stratégie de **C.** ait évolué par rapport aux essais de la première configuration, en tenant compte de mes difficultés propres, tout comme il est possible que ma propre efficacité dans ce type de tâche ait évolué.

Lors du scénario de guidage, la réalisation de la tâche s'est avérée plus complexe, ce qui s'est traduit par un échec. En effet, comme nous l'avons vu précédemment, lors de l'exploration d'une forme géométrique simple, un seul « chemin » est disponible : celui qui est tracé par la forme elle-même, et qui guide l'exploration grâce à la possibilité d'anticipation stable des stimulations en fonction de l'activité de l'utilisateur (Lenay, 1997 ; Lenay et al., 1997). Les stimulations supplémentaires, qui résultent du toucher réciproque, ont donc présenté un support additionnel pour ma propre activité d'exploration. En revanche, lorsque plusieurs chemins sont disponibles, il ne suffit pas de suivre une ligne correctement, mais de suivre la ligne qui a été choisie par l'utilisateur expert. De cette façon, la tâche est double : il s'agit d'organiser son mouvement d'exploration de la forme, tout en suivant une source de stimulations supplémentaires. Dans cette situation, de nombreuses pertes de contact sont survenues, et celles-ci se révélaient souvent définitives. Pour se retrouver, il était donc nécessaire de se donner un point de rendez-vous en dehors des lignes, le plus souvent dans un coin de l'écran, pour éviter les interférences liées aux stimulations résultant du contact avec une ligne. Il est à noter que de son côté, **C.** parvient à faire la différence aisément entre les deux types de stimulations : il est alors capable de me retrouver même si je suis en train

d’explorer une partie du parcours. Il semble donc que le niveau d’expertise joue un rôle crucial dans les situations d’interaction.

3.1.2. Configurations asymétriques

Une dernière configuration a été testée. Cette fois, les propriétés des corps-percevant et des corps-image de chacun sont définies de façon asymétrique. Ainsi, nous déplaçons tous deux un corps-percevant (matrice de 4x4 champs récepteurs élémentaires identique à celle utilisée lors des situations précédentes), mais cette fois, je suis le seul à posséder un corps-image (surface carrée de 2x2 pixels située au centre de son corps-percevant). Par conséquent, nous pouvons percevoir la forme à explorer, mais seul l’utilisateur expert peut me percevoir. De cette façon, **C.** peut savoir où je suis et de quelle manière j’explore la forme sans interférer avec elle. En contrepartie, en cas de difficultés, seules les consignes orales données par **C.** peuvent me guider.

Ici, les difficultés sont inversées : du fait d’une absence de toucher réciproque et de guidage dans l’espace, la localisation de la forme est plus difficile. Il a donc été nécessaire que **C.** intervienne verbalement pour me guider dans l’espace. En revanche, une fois la forme atteinte, l’exploration est facilitée grâce à l’absence d’interférences tactiles : mon activité d’exploration a pu se focaliser sur les stimulations résultant du contact avec la forme. A l’inverse, les interventions de **C.**, limitées aux instructions verbales, conduisent à plus d’imprécisions et d’approximations. Il ne peut en effet m’indiquer que des directions générales ou des conseils plus ou moins efficaces (comme par exemple, « bouge moins vite » ou « déplace-toi de gauche à droite sans quitter la ligne ») sans jamais pouvoir en faire une démonstration. Quoi qu’il en soit, dans ce mode d’interaction, j’ai pu identifier les formes et lors du scénario de guidage, j’ai pu réussir à parcourir le chemin choisi par **C.** grâce à ses instructions verbales.

3.1.3. Synthèse des observations préliminaires

Ces observations préliminaires conduisent à plusieurs résultats contradictoires. En effet, il apparaît que le toucher réciproque présente des avantages pour le guidage d’un utilisateur novice par un utilisateur expert. En effet, il semble nécessaire pour guider l’utilisateur novice vers la forme à explorer. De plus, il apporte un soutien lors des incidents géométriques (comme les angles) ou lors des pertes ponctuelles de la forme à explorer qui conduisent à une désorientation de l’utilisateur novice. Enfin, il permet à l’utilisateur expert d’attirer l’attention

de l'utilisateur novice sur une particularité de la forme à explorer. Néanmoins, il apparaît que la présence systématique du corps-image de son partenaire est une source de difficulté supplémentaire qui perturbe la tâche d'exploration elle-même. En effet, lors des observations préliminaires, nous avons pu identifier des difficultés d'interaction liées aux configurations symétriques. Celles-ci sont principalement dues à des effets d'interférences entre les stimulations dues à la rencontre de la forme et les stimulations dues au toucher réciproque. Ces interférences produisent même dans certains cas des occlusions perceptives qui perturbent voire empêchent la résolution de la tâche par le sujet novice. En effet, les stimulations dues à la rencontre du sujet expert doivent être discriminées des stimulations dues à la rencontre de la forme, ce qui nécessite une certaine expérience qui fait défaut pour le sujet novice. En effet, de son côté, le sujet expert n'exprime pas de difficultés particulières, et ne semble pas perturbé par les deux types de stimulations. Une piste de solution serait d'aider le sujet novice à la discrimination des deux sources d'information, par exemple en associant un son au corps-image de l'utilisateur expert.

Avec la configuration asymétrique, au contraire, l'exploration semble favorisée par l'absence totale d'interférences. Ce mode d'interaction semble donc indiqué lorsque le sujet novice explore une forme, ou s'engage dans le suivi d'une ligne. Dans ce cas, le sujet expert peut se placer en position d'observateur extérieur, ce qui le renseigne sur l'activité du sujet novice sans le perturber. En revanche, le guidage repose alors exclusivement sur les instructions verbales du sujet expert. Si le sujet novice se trompe ou perd le contact avec la forme, l'impossibilité d'interagir tactilement fait alors défaut.

De façon à conjuguer les avantages liés au deux configurations, il serait intéressant de donner la possibilité au sujet expert d'observer ou d'intervenir en temps réel. Par l'intermédiaire d'un bouton situé sur le module d'interaction tactile, le sujet expert pourrait alors décider de faire apparaître ou non son corps-image, selon la nécessité de la tâche ou selon les besoins spécifiques des utilisateurs.

3.2. Les rencontres interpersonnelles : une problématique d'usage

Comme nous l'avons vu, les espaces d'interaction tactile posent la question de la rencontre de plusieurs utilisateurs distants dans des espaces numériques partagés. Il est alors nécessaire que ces utilisateurs aient la possibilité de se discriminer les uns des autres, voire de se reconnaître. En effet, si plusieurs utilisateurs sont connectés simultanément dans le même espace numérique, chacun peut rencontrer le corps-image des autres utilisateurs par l'intermédiaire de son corps-percevant. L'intérêt de ces rencontres tactiles numériques,

comme nous l'avons déjà mentionné, est de rendre possible l'émergence d'une communauté de pratique dans un contexte d'exploration mutuelle de contenus numériques (pédagogiques, pratiques ou ludiques), de communication, d'entraide et/ou de guidage. Dans ce cadre, la possibilité de discrimination interpersonnelle des utilisateurs est cruciale.

Bien entendu, les espaces numériques partagés peuvent être pensés comme les réseaux numériques existants. Dans ceux-ci, l'identité de l'utilisateur, au moins à travers son pseudonyme, est révélée, si bien que chaque utilisateur peut savoir qui est connecté en même temps que lui ou qui a participé au réseau. Il n'en reste pas moins que l'interaction avec un des utilisateurs connectés au sein de l'espace partagé ne doit pas être ambiguë.

Lors des analyses d'usage préliminaires, le test des émoticônes tactiles, ou tact-icônes, avait conduit les utilisateurs à suggérer la possibilité de définir la forme de son corps-image, en tant que configuration personnalisée individualisant ce que l'utilisateur laisse à percevoir aux autres. Cette possibilité nous semble pertinente, car de cette façon, les utilisateurs pourraient se discriminer les uns des autres par la forme prédéterminée de leur corps-image. Pour cela, la surface de capteurs de pression (FSR) du module d'interaction tactile (MIT) est une solution possible, dans la mesure où cette interface d'entrée multipoint présente une surface carrée analogique au corps-image par défaut. Ainsi, chaque utilisateur pourrait définir par pression (simultanée ou non) quelles zones de cette surface seraient perceptibles par les autres utilisateurs lors d'un contact. Le nombre de configurations possibles différentes, étant donné les 16 picots de la matrice de stimulateurs tactiles, est conséquent (16! soit plus de 200 milliards configurations différentes). Seulement, un nombre très important de possibilités pourraient alors conduire à des ambiguïtés (voir Figure 27), obligeant l'utilisateur à focaliser son attention sur la forme exacte des picots, en procédant à des micro-mouvements avec le doigt de la main non-dominante pour explorer la matrice de stimulateurs.

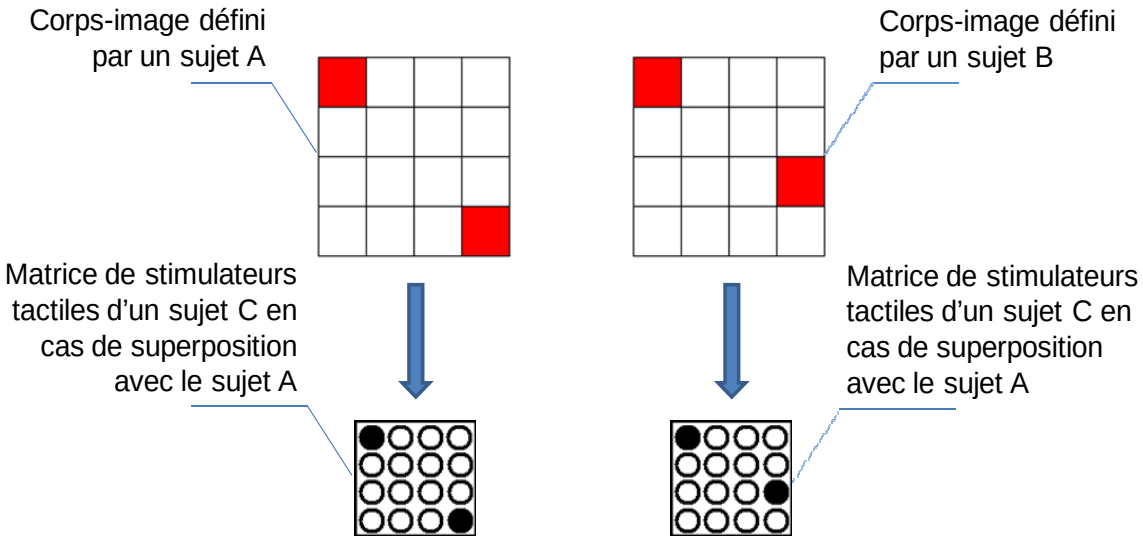


Figure 27 : Ambiguïté des corps-image des sujets A et B pour un sujet C

De plus, cette ambiguïté se trouverait démultipliée par le caractère non-statique des rencontres dans l'espace. En effet, les déplacements communs effectués par les sujets peuvent conduire à des phases de superposition partielle, pouvant rendre ambiguës certaines stimulations résultant des rencontres (voir Figure 28).

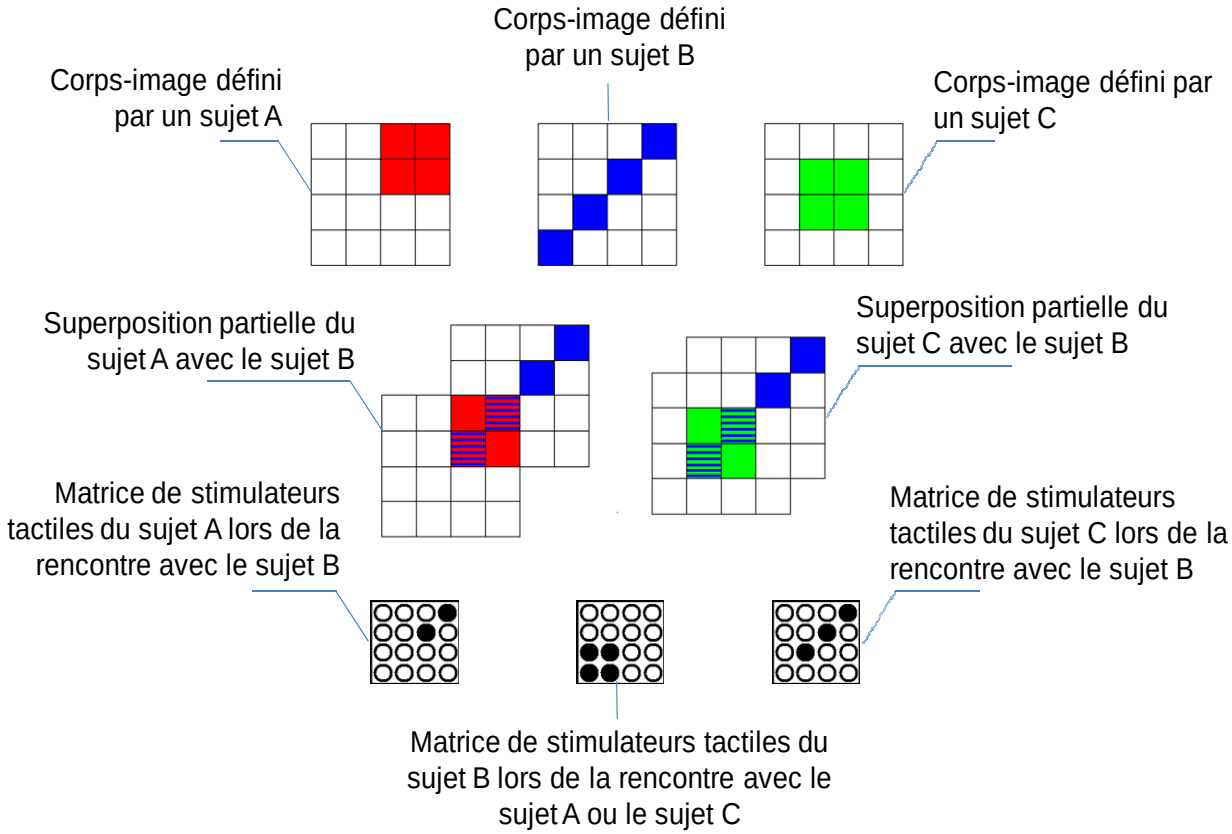


Figure 28 : Ambiguïté des superpositions partielles des sujets A et C avec le sujet B, qui conduisent à des stimulations tactiles identiques à un instant t

Par conséquent, deux utilisateurs qui auraient choisi la même configuration de leur corps-image, ou des configurations trop proches, ne pourraient pas être discriminés par les autres utilisateurs par l'intermédiaire de cette seule information.

Au-delà de la définition statique des corps-image, d'autres possibilités s'offrent aux sujets pour discriminer les autres utilisateurs. En effet, la discrimination et la reconnaissance interpersonnelle dans les espaces d'interaction tactile pourraient être basées sur la dynamique de l'activité perceptive déployée par chaque utilisateur. Lors des analyses d'usage du dispositif Tactos, tout comme lors des études expérimentales concernant la reconnaissance de formes, nous avons par exemple observé à de multiples reprises que les dynamiques d'exploration sont propres aux utilisateurs (en termes de vitesse, d'amplitude, de réaction aux stimulations etc.), même si nous avons pu dégager des stratégies globales que nous pouvons facilement identifier (suivi de contours, micro-balayage orthogonal au contour etc.). Dans le cadre des rencontres tactiles dans les espaces numériques partagés, nous pouvons alors poser l'hypothèse que la dynamique propre de l'activité perceptive de chaque utilisateur pourrait être discriminée voire reconnue par les autres. En d'autres termes, la discrimination interpersonnelle pour être basée sur la détection des « signatures » propres de l'activité perceptive déployée par chaque utilisateur.

Par ailleurs, les utilisateurs pourraient également recourir à des stratégies plus spécifiques. En effet, l'interaction tactile nous semble propice à l'émergence de signes intentionnellement déployés par les utilisateurs pour se faire reconnaître par les autres. En effet, les utilisateurs pourraient tenter de se différencier des autres en effectuant un pattern de mouvement spécifique lors des rencontres tactiles. Par exemple, un utilisateur donné pourrait procéder à des mouvements circulaires de faible amplitude lors d'un contact avec un autre utilisateur. Nous pourrions alors assister à une émergence sémiotique basée sur des signes intentionnels, au départ plus ou moins arbitraires, mais qui prendraient sens au sein de l'interaction. Notons qu'il en est de même dans la communication naturelle, qu'elle soit verbale ou non verbale. En effet, le sens n'est pas circonscrit dans le signe (linguistique ou gestuel), mais il émerge d'une perpétuelle (re)négociation au sein de l'interaction (Maturana & Varela, 1994). Ce processus de négociation nécessite une organisation temporelle très précise de l'interaction, de façon à mettre en place une alternance de « tours de parole » propice à un échange structuré d'information. En effet, si deux utilisateurs déploient en même temps leur signe distinctif (par exemple, un mouvement circulaire de faible amplitude et une oscillation latérale sur le plan horizontal), alors il y a peu de chances que les sujets reconnaissent le signe et donc l'identité du sujet rencontré.

Enfin, la nappe de capteurs de pression implémentée dans le MIT pourrait ici aussi apporter une solution. Au-delà de la définition statique d'un corps-image par les utilisateurs, cette surface pourrait être utilisée de façon dynamique lors des rencontres interpersonnelles. Pour cela, les sujets pourraient exercer des pressions sur le FSR lors des interactions tactiles de façon à modifier dynamiquement la configuration de leur corps-image. Il serait également possible de se placer en situation de face à face (soit en s'immobilisant l'un sur l'autre dans l'espace, soit en basculant dans un mode dédié via une commande-utilisateur spécifique), et utiliser les capteurs de pression pour se mettre en situation de « caresse tactile distale ». Dans cette optique, toutes les pressions exercées sur la surface activent les picots correspondants des stimulateurs tactiles de l'autre participant (voir Figure 29).

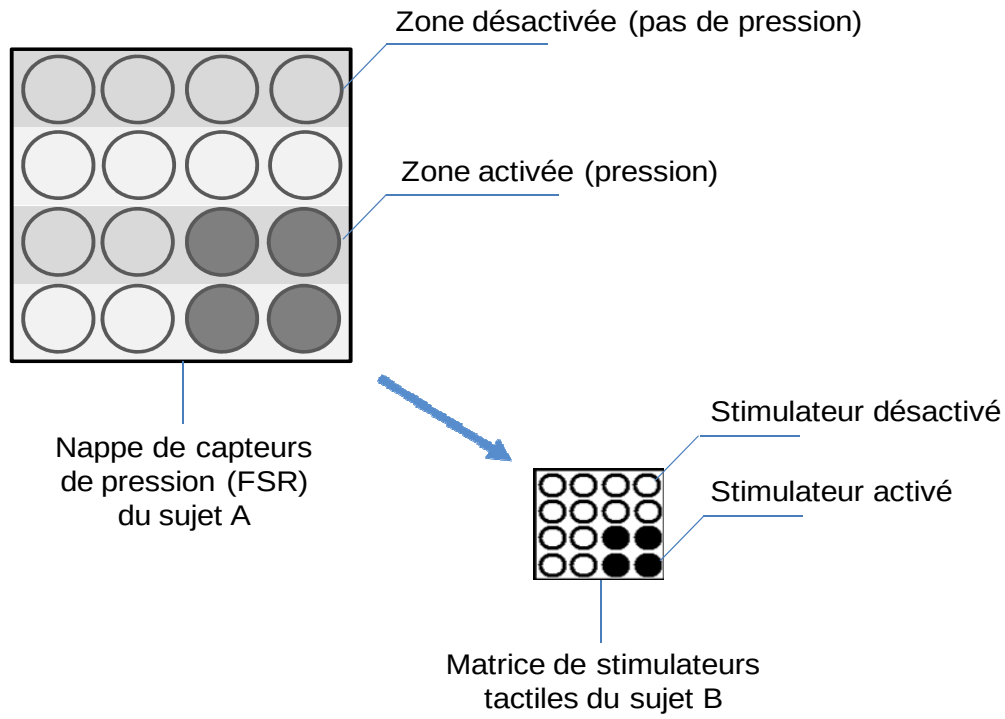


Figure 29 : La « caresse tactile distale ». La nappe de capteurs de pression (FSR) est divisée en 16 zones, chacune d'elles étant reliée au picot correspondant de la matrice de stimulateurs tactiles de l'autre sujet

De façon à apporter des éléments objectifs à ces scénarii possibles, il a été alors question de procéder à des tests expérimentaux pour étudier différentes conditions de possibilité de discrimination et d'identification interpersonnelle.

3.3. Discrimination interpersonnelle (1)

Dans un premier temps, nous avons choisi d'expérimenter la discrimination interpersonnelle à travers la dynamique de l'activité perceptive déployée par chaque

utilisateur. Pour cela, nous avons proposé une tâche dans laquelle les sujets rencontreront d'autres sujets et devront apprendre à les reconnaître sans pouvoir modifier la configuration de leur corps-image. Notons que dans cette situation, les utilisateurs ont la possibilité de déployer intentionnellement des signes distinctifs produits au sein de l'interaction, même si aucune indication de ce type ne leur est fournie. Nous nous demandons donc si les utilisateurs auront recours à cette éventualité de façon spontanée.

3.3.1. Population

24 sujets, répartis en 6 groupes de 4 sujets, ont participé à cette expérience. Les sujets sont des étudiants de l'Université de Technologie de Compiègne âgés entre 18 et 25 ans, recrutés au moyen de petites annonces dans l'établissement. Aucun sujet n'a de connaissance préalable du dispositif.

3.3.2. Matériel

L'expérience se déroule dans deux salles adjacentes, et a donc nécessité la présence de deux expérimentateurs[17]. Chaque salle contient respectivement deux postes de travail destinés aux sujets. Dans une des deux salles, un ordinateur serveur permet de connecter les sujets dans un espace partagé, et ainsi piloter les différentes phases de l'expérience.

Chaque poste de travail destiné aux sujets comprend :

- un ordinateur portable sur lequel est installé le logiciel Tactos ;
- une souris qui permet le déplacement d'un curseur dans l'espace numérique[18] ;
- un boîtier de stimulateurs tactiles composé de deux cellules Braille piézoélectriques juxtaposées (soit 2 x 8 picots) ;
- un casque audio qui permet de délivrer des consignes individuellement.

L'espace partagé est un espace bidimensionnel de 160 x 160 pixels, non-visible sur l'écran de l'ordinateur. Il s'agit en réalité d'un tore géométrique, si bien que les sujets n'ont pas de conscience de limites dans l'espace d'action : le franchissement d'un des bords de l'écran entraîne la réapparition simultanée au bord opposé.

[17] Nous remercions vivement Camille Barot et Cécile Laborde pour avoir participé à la réalisation de cette expérience ainsi qu'au traitement des résultats.

[18] Lorsque cette expérimentation a eu lieu, nous n'avions pas encore à disposition le premier prototype de module d'interaction tactile (voir Partie 1, Section 2.1.1, p.43). Quoi qu'il en soit, nous avons préféré un effecteur relatif (souris) à un effecteur absolu (tablette/stylet), parce que nous nous intéressons ici à la dynamique du déplacement du sujet plus qu'à la succession de ses positions dans l'espace.

A l'aide de la main dominante sur la souris, les sujets déplacent un curseur composé d'un corps-percevant et d'un corps-image. Le corps-percevant est une matrice de 16 champs récepteurs élémentaires de 4 x 4 pixels chacun. Lorsqu'au moins un pixel de couleur croise un des champs récepteurs élémentaires, cela active le picot correspondant de la matrice de stimulateurs tactiles, situé sous l'index de la main non-dominante. Nous avons donc un parallélisme entre les champs récepteurs et les stimulateurs tactiles.

Le corps-image est une surface carrée de 4 x 4 pixels attachée au corps-percevant (voir Figure 30).

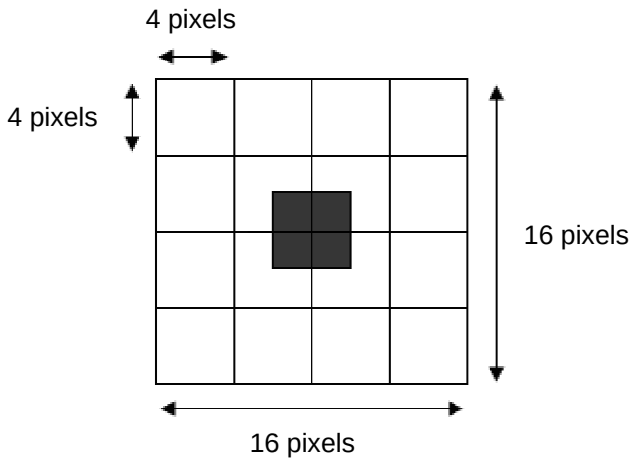


Figure 30 : Curseur déplacé par les sujets, composé d'un corps-percevant de 4 x 4 champs récepteurs élémentaires de 4 x4 pixels chacun, et d'un corps-image de 4 x 4 pixels (en noir au centre) attaché au centre du corps-percevant.

Pour les besoins de la tâche expérimentale, chaque sujet se voit attribuer un corps-image avec une des quatre couleurs suivantes : Bleu, Rouge, Jaune, Vert[19]. La Figure 31 ci-dessous présente un schéma détaillé du dispositif expérimental.

[19] Notons que ces couleurs ne sont pas perceptibles par les sujets, il s'agit uniquement d'un attribut qui leur permettra de se nommer les uns les autres pour la suite de l'expérience. Par ailleurs, ils permettent à l'expérimentateur de visualiser les interactions en temps réel.

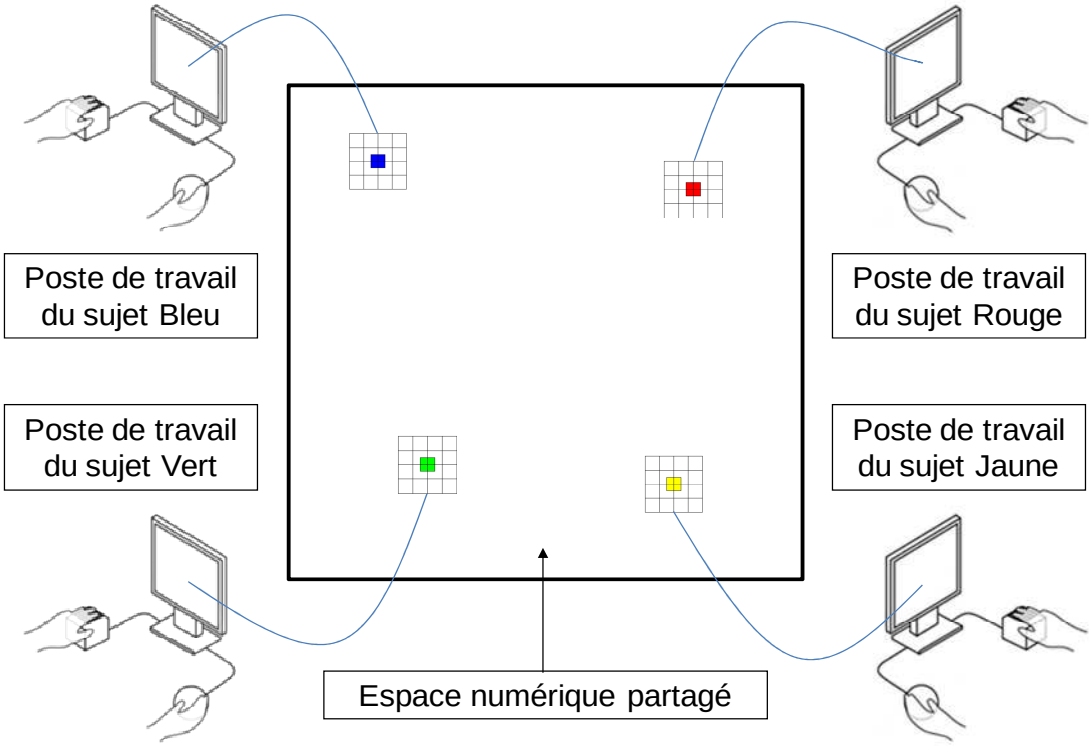


Figure 31 : Schéma du dispositif expérimental

3.3.3. Procédure

Lorsque nous accueillons les sujets, nous leur expliquons le principe général du dispositif, ainsi que la nature de la tâche à laquelle ils vont participer. Les sujets sont ensuite répartis dans les deux salles, chacun s'installant devant un des postes de travail disponibles.

Une première phase, dite de « familiarisation », leur est alors proposée. Dans celle-ci, les sujets sont connectés dans des espaces privés respectifs où un rectangle noir est inscrit. L'objectif pour les sujets consiste à déplacer leur curseur via la souris de façon à rencontrer la forme. Une explication détaillée du rapport entre l'activation des stimulateurs tactiles et la rencontre du curseur avec une partie de la forme est fournie par l'expérimentateur présent dans chaque salle. Il leur est alors demandé d'essayer de garder le contact avec la forme tout en essayant de la parcourir. Dans cette première phase, les sujets sont informés qu'ils peuvent poser des questions sur le fonctionnement du dispositif. Quand les sujets n'ont plus de questions, et que les expérimentateurs constatent que le principe est bien compris, une deuxième tâche leur est proposée. Dans celle-ci, les sujets sont connectés dans un deuxième espace privé dans lequel 4 carrés noirs de 4 x 4 pixels chacun sont présents. Les sujets sont informés que chaque carré noir a la même taille et la même forme que le curseur qu'ils déplacent : lors des phases ultérieures, quand les sujets seront connectés dans des espaces

partagés, ils pourront « toucher » les autres participants. L'objectif de cette tâche consiste pour chaque sujet à essayer de rencontrer plusieurs de ces carrés, et d'essayer de localiser leur position les uns par rapport aux autres. Aussi, il est question d'expliquer la nature particulière de l'espace, qui est en réalité infini : si les sujets déplacent leur souris toujours dans la même direction, ils parcourent alors l'espace plusieurs fois. Les expérimentateurs s'assurent alors que les sujets ont bien compris le dispositif et n'ont plus de questions concernant le principe de fonctionnement, avant de lancer la phase suivante.

Dans la phase 2, dite de « rencontre », les 4 sujets sont connectés dans un espace numérique partagé. A partir de cette phase, il est demandé aux sujets de ne plus parler. Chaque sujet peut alors rencontrer tactilement chacun des trois autres sujets. Les sujets sont informés que cette fois, aucune forme n'est inscrite dans l'espace. Par conséquent, les stimulations tactiles qu'ils reçoivent proviennent nécessairement de la rencontre de leur curseur avec le curseur d'un des trois autres participants. Quand une stimulation survient, les sujets ont la possibilité de cliquer de façon à déclencher la lecture de l'identité du curseur rencontré au moyen du casque audio. Chaque sujet est alors désigné par une couleur (Rouge, Bleu, Vert et Jaune) qui sera identique pendant toute l'expérience. Notons que la lecture de l'identité du sujet rencontré n'est déclenchée que si le clic survient au plus tard deux secondes après la fin de la stimulation. De plus, si le sujet clique alors qu'il est en contact avec plusieurs participants simultanément, il entend successivement toutes les couleurs concernées (par exemple, « Rouge, Vert »). Durant cette phase, il est demandé aux sujets de faire bien attention à la façon dont se déplacent chacun des trois autres sujets, car dans les phases suivantes, ils devront essayer de les identifier. Cette phase dure deux minutes.

Dans la phase 3, dite de « recherche », le principe est le même, mais cette fois, chaque sujet reçoit pour consigne de cliquer sur un des trois autres sujets (Par exemple, Bleu reçoit pour consigne de cliquer sur Rouge). Comme lors de la phase précédente, chaque clic entraîne la lecture de l'identité du sujet rencontré, et aucune pénalité n'est distribuée en cas de mauvais clic. Pendant cette phase, les sujets reçoivent successivement trois consignes, de façon à ce que chaque sujet ait pour objectif de rechercher chacun des trois autres. L'expérimentateur rappelle aux sujets qu'ils doivent encore être particulièrement attentifs aux interactions, car par la suite, ils devront essayer d'identifier chaque participant. Chaque consigne est suivie d'une phase de recherche de 1 minute et 30 secondes, si bien que cette phase dure au total 4 minutes et 30 secondes.

Dans la phase 4, dite de « décalage », les conditions sont en tous points identiques à la phase de « recherche », mais cette fois, chaque clic correct d'un sujet entraîne un déplacement

de son curseur dans l'espace. Le déplacement est de l'ordre de 50 pixels dans une direction aléatoire. L'objectif de ce décalage consiste à empêcher les sujets de rester en interaction avec sa cible et donc de cliquer plusieurs fois successivement. Au contraire, le décalage oblige les sujets à interagir avec un des trois autres sujets de façon à cliquer uniquement s'ils pensent avoir identifié la cible. Comme lors de la phase précédente, chaque consigne est suivie d'une phase de recherche de 1 minute et 30 secondes, si bien que cette phase dure au total 4 minutes et 30 secondes.

La phase de test à proprement parler commence alors. Les conditions sont identiques à la phase de « décalage », mais cette fois, un score est comptabilisé pour chaque clic effectué. Un clic correct sur la cible entraîne le gain d'un point, tandis qu'un clic incorrect entraîne la perte d'un point. Chaque sujet commence avec un capital de 50 points, et la tâche leur est présentée comme un concours dans lequel il s'agit de gagner un maximum de points. Ici, chaque clic déclenche la lecture de l'évolution du score, et n'indique plus par conséquent la lecture de l'identité du sujet rencontré. 5 essais de 2 minutes chacun sont alors proposés successivement. Pour chaque essai, chaque sujet reçoit pour consigne de cliquer sur un des trois autres sujets. Deux types de consignes sont alors alternés à chaque essai : les consignes symétriques et les consignes asymétriques. Lors des consignes symétriques, chaque sujet a pour cible le sujet pour qui il est la cible. Par exemple, si Bleu reçoit pour consigne de cliquer sur Rouge, Rouge reçoit pour consigne de cliquer sur Bleu. Lors des consignes asymétriques, chaque sujet a pour cible un sujet qui lui-même a une autre cible. Par exemple, Bleu reçoit pour consigne de cliquer sur Rouge qui reçoit pour consigne de cliquer sur Vert… Le premier essai est toujours une consigne symétrique, si bien que pour les 5 essais, les sujets recevront 3 consignes symétriques et 2 consignes asymétriques. Notons que ces conditions expérimentales ne sont pas expliquées aux sujets.

Lorsque les 5 essais expérimentaux sont effectués, une pause de quelques minutes est proposée aux sujets, pendant laquelle ils n'ont pas l'autorisation de se parler. Ensuite, les sujets passent l'ensemble de l'expérimentation une seconde fois, depuis la phase de rencontre jusqu'à la phase de test. A la fin de l'expérience, les 4 sujets sont réunis dans la salle de l'ordinateur-serveur, et un débriefing est animé par les expérimentateurs, de façon à recueillir les impressions des sujets sur la tâche ainsi que l'explicitation de leurs stratégies éventuelles.

Les résultats de cette expérience seront examinés à la lumière de trois hypothèses :

- Notre première hypothèse est que la dynamique de déplacement des sujets va leur permettre de réussir la tâche. Dans une visée exploratoire, nous cherchons à savoir si

cette réussite peut être basée sur une signature « naturelle » de l'activité perceptive déployée par les sujets, ou si un signe intentionnel est nécessaire. En effet, il est possible que les sujets cherchent à se démarquer les uns des autres en déployant de façon intentionnelle un pattern d'action caractéristique lors des rencontres.

- Selon une seconde hypothèse, nous nous attendons à ce que les sujets obtiennent de meilleurs scores d'identification lors de la deuxième session expérimentale que lors de la première. Ce résultat pourra refléter un apprentissage progressif de la discrimination interpersonnelle basé sur les rencontres successives, ainsi que la prise de conscience de la nécessité de déployer un signe intentionnel pour se démarquer les uns des autres.
- Enfin, nous posons l'hypothèse que l'identification des sujets est favorisée par une interaction réciproque, les deux activités perceptives étant alors dirigées l'une vers l'autre. En d'autres termes, les consignes symétriques entraîneront de meilleurs scores d'identification que les consignes asymétriques.

3.3.4. Résultats[20]

Dans un premier temps, nous observons qu'une grande majorité de clics a été effectuée lorsque les sujets ont effectivement rencontré un autre sujet dans une fenêtre temporelle de deux secondes (92,92%). Les autres clics (7,08%) ne sont pas considérés dans l'analyse, car ils correspondent soit à des clics survenus trop tard, et donc difficilement attribuables à une rencontre, soit à des clics non-intentionnels, effectués par erreur par les sujets alors qu'ils manipulaient la souris.

Nous calculons donc le pourcentage de clics corrects par rapport au nombre total de clics attribués à la rencontre d'un sujet dans une fenêtre temporelle de deux secondes. Globalement, nous observons que le pourcentage de clics corrects obtenu par les sujets est peu élevé (voir Figure 32).

[20] Pour des détails concernant la méthodologie d'analyse statistique, veuillez vous reporter à l'Annexe A, p. 221.

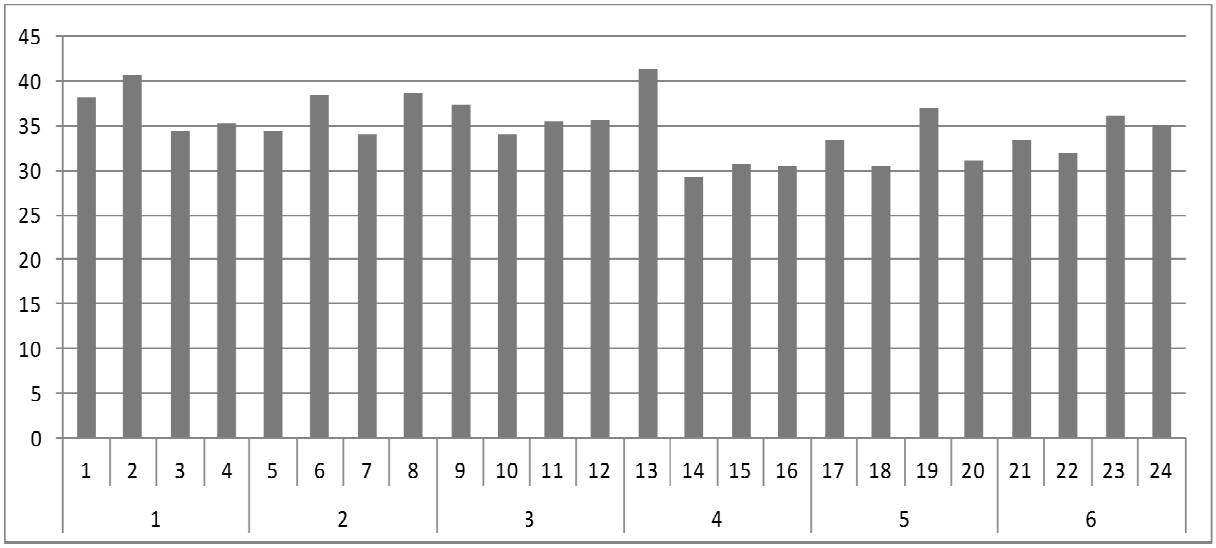


Figure 32 : Pourcentages de clics corrects obtenus pour chaque sujet (de 1 à 24) regroupés en fonction des 6 passations expérimentales

Chaque sujet pouvant rencontrer trois autres sujets, le sujet a 33,33% de chances de cliquer sur la bonne cible. Or, la moyenne globale de clics corrects s'élève à 34,84%, avec une dispersion relativement faible (s = 3,25). Les résultats obtenus par l'ensemble des sujets ne sont pas significativement différents du hasard (test de Student : $t_{(23)} = 1.7909$; $p = 0.0865$). Malgré tout, une analyse bayésienne révèle que cette différence ne peut pas être considérée comme négligeable ($Pr\ |\delta/\sigma| < .33 = .45$).

Si nous nous intéressons à l'évolution des scores obtenus par chaque sujet, nous observons que globalement, plus les sujets cliquent, plus ils commettent d'erreurs (voir Figure 33). En effet, nous observons une co-variation linéaire entre le nombre de clics effectués et le score final ($r^2 = .87$).

Par ailleurs, nous observons que les sujets obtiennent un meilleur pourcentage de clics corrects dans la première session expérimentale que dans la seconde (d = 4,91) (voir Tableau 5). Cette différence n'est pas significative (test de Student : $t_{(23)} = 1.9855$; $p = 0.0591$), même si elle ne peut pas être considérée comme négligeable ($Pr\ |\delta/\sigma| < .33 = .37$).

Enfin, nous observons que les sujets obtiennent de meilleurs pourcentages de clics corrects lors des consignes asymétriques que lors des consignes symétriques (d = 2,37) (voir Tableau 5). Toutefois, cette différence n'est pas significative (test de Student : $t_{(23)} = 0.9996$; $p = 0.3279$), même si là encore, elle ne peut pas être considérée comme négligeable ($Pr\ |\delta/\sigma| < .33 = .73$).

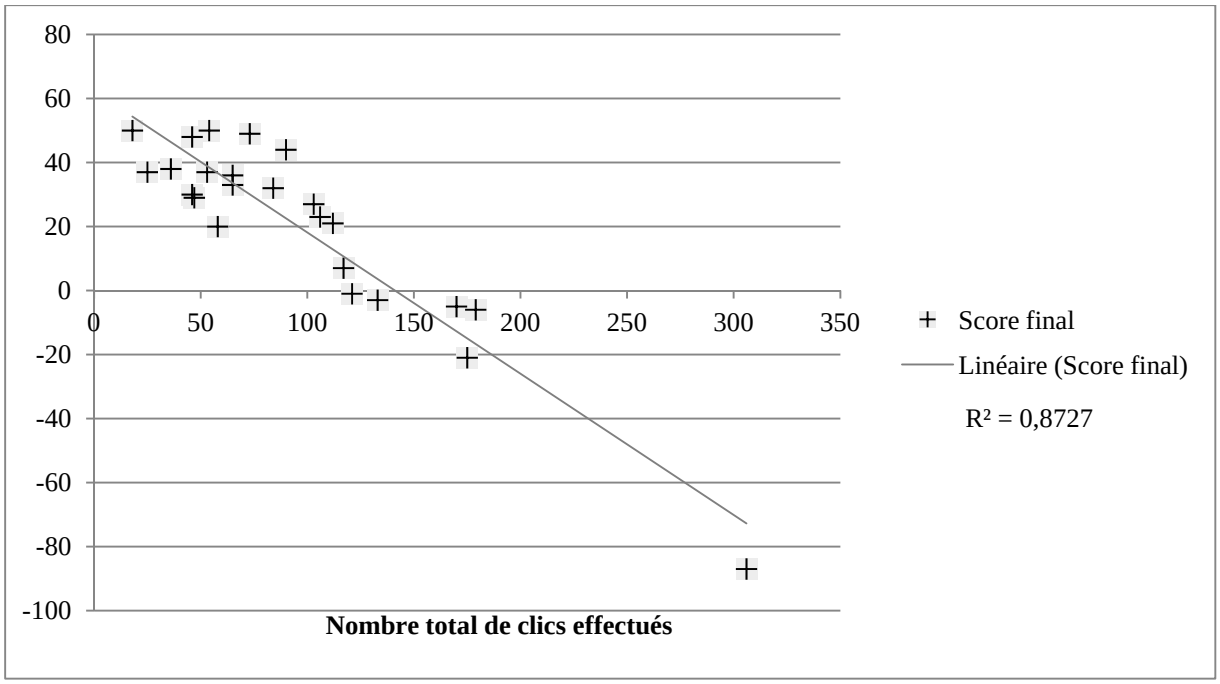


Figure 33 : Scores finaux obtenus pour chaque sujet en fonction du nombre total de clics effectués

	MOYENNE	ECART-TYPE
% de clics corrects	34,84%	3,25186328
Session 1	37,30%	6,11031184
Session 2	32,39%	7,55335489
Consignes symétriques	33,89%	6,21891671
Consignes asymétriques	36,27%	7,04420224

Tableau 5 : Récapitulatif descriptif des résultats obtenus

3.3.5. Discussion

Nous pensions que les sujets pouvaient parvenir à s'identifier les uns les autres sur la seule base de la dynamique de leurs activités perceptives. Mais les résultats obtenus ne nous permettent pas de conclure dans ce sens. En effet, les pourcentages de clics corrects obtenus ne sont pas statistiquement différents du hasard. Par ailleurs, nous avons observé une co-variation du nombre de clics effectués avec les scores finaux obtenus. Ce résultat laisse à penser que les sujets peuvent être regroupés selon deux catégories de profils individuels. Quand les premiers cliquent peu, et obtiennent les scores finaux les plus élevés, les seconds cliquent beaucoup et voient leur score final chuter. Nous interprétons ce résultat par le fait que lorsque le sujet prend le temps d'interagir avec l'autre avant de cliquer, il commet mois d'erreurs. Toutefois, nous avons observé que pour ces sujets, le score final ne dépasse jamais le nombre de points alloué à chaque sujet au début de l'expérience (soit 50 points). En d'autres termes, ceux-ci font tout de même des erreurs d'identification.

Par ailleurs, contrairement à ce que nous attendions, nous ne pouvons pas conclure à une amélioration des scores d'identification entre les deux sessions expérimentales. Ce résultat peut être expliqué par la fatigue des sujets devant une tâche jugée trop longue. En effet, lors du débriefing, une grande majorité des sujets a fait part d'un désintérêt progressif pour la tâche, qui leur a semblé interminable. Pour les sujets qui n'ont pas exprimé cette fatigue, la trop grande difficulté de la tâche a été mise en avant. Devant l'impossibilité de trouver une stratégie adaptée, ils ont alors subi la tâche en cliquant au hasard, en se basant sur l'évolution des scores plus que sur la dynamique des activités perceptives rencontrées.

Enfin, contre toute attente, nous ne pouvons pas conclure à un effet du type de consigne (symétrique vs asymétrique) sur les scores d'identification. De façon générale, les sujets n'ont pas spontanément mis en œuvre de signes distinctifs, et ont déclaré avoir avant tout tenté de rester en contact avec le premier sujet rencontré, quelle que soit la cible énoncée par la consigne. Il semble donc que tous les essais ont été appréhendés de la même manière par les sujets. Au contraire, pour les sujets qui n'ont pas exprimé de fatigue ni de difficulté spécifique, les essais sont jugés trop courts pour pouvoir mettre en place de réelles stratégies.

D'une façon générale, il semble que ce premier protocole expérimental ait souffert de limites importantes, notamment au niveau méthodologique. En effet, il était demandé aux sujets d'essayer d'identifier les autres sujets au fur et à mesure des essais. Ainsi, chaque sujet était placé dans une attitude d'observateur, ce qui a pu contrarier la nature même de la tâche. En effet, chaque sujet s'est alors retrouvé à essayer d'observer le comportement des différents sujets rencontrés, ceux-ci étant également engagés dans la même activité d'observation. Au final, les sujets devaient interagir avec des observateurs peu actifs, ce qui a pu entraver la discrimination interpersonnelle, en masquant les dynamiques propres des activités perceptives déployées dans l'espace partagé.

En outre, certaines rencontres ont pu être parasitées par la présence simultanée d'autres utilisateurs dans le même espace. En effet, à de nombreuses reprises, des sujets se sont retrouvés à interagir à trois ou à quatre simultanément, chacun pensant interagir avec un des trois autres. Dans ce contexte, les possibilités de discrimination interpersonnelle ont été compromises, les dynamiques déployées individuellement par les sujets ayant pu interférer les unes avec les autres.

3.4. Discrimination interpersonnelle (2)

Suite aux remarques précédentes, nous avons élaboré un second protocole destiné à surmonter ces difficultés.

Premièrement, nous avons opté pour des interactions deux à deux plus propices à une éventuelle différenciation des rencontres successives. En conséquence, il n'est plus demandé aux sujets de cliquer sur un sujet-cible préalablement défini, mais d'interagir avec un des trois autres sujets pendant la totalité de l'essai. En conséquence, les sujets ne doivent fournir qu'une seule réponse à la fin de chaque essai, basée sur la totalité de l'interaction.

Deuxièmement, nous demandons aux sujets de reconnaître les autres sujets, mais aussi de se faire reconnaître par eux. De cette manière, nous tentons de favoriser un engagement mutuel des sujets dans l'interaction, de façon à rendre possible l'émergence d'une dynamique interpersonnelle.

Enfin, nous proposons un nouvel effecteur aux sujets, spécialement conçu pour les interactions tactiles distales : il s'agit du prototype de module d'interaction tactile (MIT), que nous élaboré suite aux analyses d'usage précédemment citées (voir Partie 1, section 2.1.1, p. 43). Contrairement à la première étude expérimentale, nous postulons que l'utilisation d'un effecteur absolu pourrait faciliter le déploiement des dynamiques individuelles. Ce type d'effecteur permet en effet aux sujets de connaître à chaque instant la position exacte de leur curseur dans l'environnement partagé, ce qui facilite la compréhension de l'espace par l'accès à des connaissances proprioceptives tirées du geste d'exploration (Auvray et al., 2005). En outre, cette meilleure gestion des déplacements pourrait favoriser le déploiement de signes intentionnels, qui pourraient permettre aux sujets de s'identifier. De plus, le module d'interaction tactile (MIT) offre une nouvelle possibilité d'action aux sujets. En effet, la nappe de capteurs de pression (FSR) implémentée dans le MIT, pourra être utilisée ici pour modifier dynamiquement la forme de son corps-image.

3.4.1. Population

16 sujets ont participé à cette deuxième expérience. Ils sont tous étudiants de l'Université de Technologie de Compiègne et sont âgés entre 18 et 25 ans. Comme pour l'étude précédente, chaque passation concerne 4 sujets. Cette fois, deux groupes expérimentaux ont été constitués en fonction des possibilités d'action fournies aux sujets pour réussir la tâche (voir section 3.4.3 ci-après).

3.4.2. Matériel

Le dispositif expérimental est identique à celui de l'étude précédente[21]. Toutefois, des différences sont à noter, notamment en ce qui concerne les postes de travail destinés aux sujets. En effet, ici, chaque poste de travail comprend :

- un ordinateur portable sur lequel est installé le logiciel Tactos ;
- un module d'interaction tactile, qui comprend sur une des deux faces un effecteur Touchpad configuré en absolu, et sur la face opposée une nappe de capteurs de pression (FSR). Sur chaque tranche du module sont disposées deux cellules Braille piézoélectriques (soit 2 x 8 picots) ;
- un casque audio qui permet de délivrer des consignes individuellement.

Dans cette expérience, l'espace partagé, qui n'est pas visible à l'écran, est un espace fini de 300 x 400 pixels. L'effecteur utilisé étant configuré en absolu, les bords de l'espace correspondent aux bords du Touchpad, et la position du doigt sur celui-ci définit la position du curseur sur l'écran. Nous avons préconisé pour cette tâche une tenue verticale du module d'interaction tactile, ce qui a conduit à utiliser un espace orienté verticalement.

A l'aide de l'index de la main dominante sur le Touchpad, les sujets déplacent un curseur composé d'un corps-percevant et d'un corps-image. Le corps-percevant est une matrice de 16 champs récepteurs élémentaires de 8 x 8 pixels chacun. Le corps-image est une surface carrée correspondante de 8 x 8 pixels attachée au corps-percevant (voir Figure 34).

Tout comme lors de l'étude précédente, lorsqu'au moins un pixel de couleur croise un des champs récepteurs élémentaires, cela active le picot correspondant de la matrice de stimulateurs tactiles, situé sous l'index de la main non-dominante. Nous avons donc ici encore un parallélisme entre les champs récepteurs et les stimulateurs tactiles.

De la même façon, pour les besoins de la tâche expérimentale, chaque sujet se voit attribuer un corps-image avec une des quatre couleurs suivantes : Bleu, Rouge, Jaune, Vert.

[21] Nous remercions vivement Pierre-Alexandre Kofron et Caroline Laroche pour avoir participé à la réalisation de cette expérience ainsi qu'au traitement des résultats.

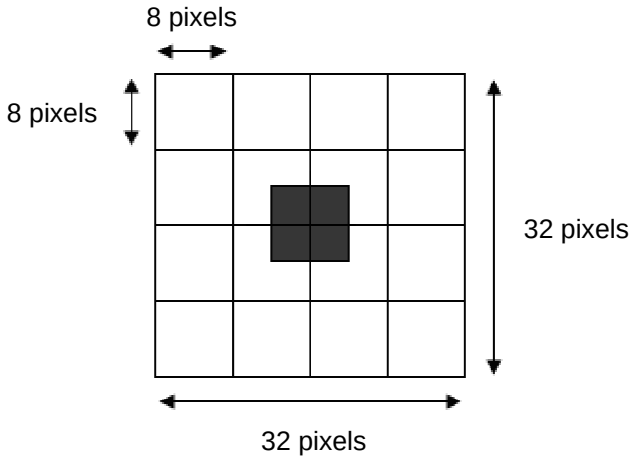


Figure 34 : Curseur déplacé par les sujets, composé d'un corps-percevant de 4 x 4 champs récepteurs élémentaires de 8 x8 pixels chacun, et d'un corps-image de 8 x 8 pixels (en rouge) attaché au centre du corps-percevant.

3.4.3. Procédure

Chaque groupe de 4 sujets constituant une passation est attribué à un des deux groupes expérimentaux :

- le premier groupe expérimental est caractérisé par l'utilisation d'un effecteur unique : le Touchpad du module d'interaction tactile, qui permet de déplacer le curseur dans l'espace partagé bidimensionnel. Pour les sujets composant ce groupe, aucune mention n'est faite à propos de la nappe de capteurs de pression, qui est par ailleurs désactivée. Le corps-image n'est donc pas déformable.
- le deuxième groupe expérimental est caractérisé par l'utilisation conjointe de deux effecteurs, qui fournissent des possibilités d'action complémentaires. Ainsi, le Touchpad du module d'interaction tactile permet, comme pour le groupe 1, de déplacer le curseur dans l'espace partagé bidimensionnel. Sur l'autre face du module, la nappe de capteurs de pression (FSR) permet de déformer en temps réel le corps-image. Lorsqu'une pression est appliquée sur une partie du FSR, la forme du corps-image se modifie en conséquence. Par exemple, si un sujet A applique une pression sur deux points du FSR, les deux points correspondants du corps-image seront activés : la rencontre de ce corps-image modifié avec le corps-percevant d'un sujet B active les picots associés à ces points sur sa matrice de stimulateurs tactiles. De la même façon, si le sujet B applique une pression sur son FSR, le sujet A pourra percevoir cette déformation à l'aide de ses propres stimulateurs tactiles (voir Figure 35). Si aucune

pression n'est exercée sur le FSR, le corps-image par défaut est perceptible par les autres sujets.

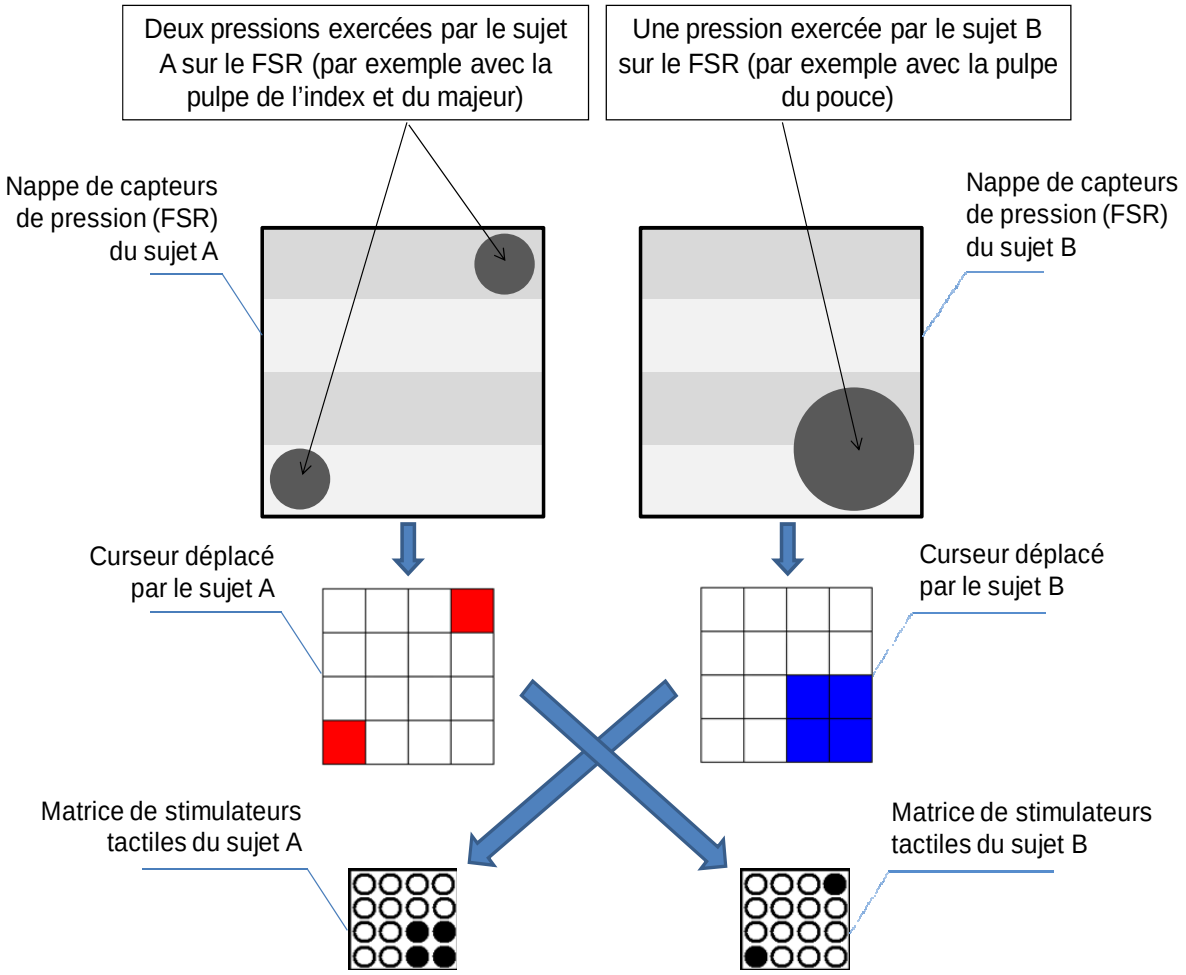


Figure 35 : Schéma du principe de déformation des corps-image via le FSR, partant du principe que le curseur des sujets A et B sont superposés dans l'espace. Sur la partie consacrée au curseur de chaque participant, les carrés de couleur (respectivement rouge et bleu) indiquent quels parties du corps-image sont actives relativement à la pression exercée sur le FSR. Les parties restantes ne sont pas actives et donc non-perceptibles par les corps-percevant.

Pour tous les sujets, quel que soit leur groupe expérimental, le principe général du dispositif est expliqué, ainsi que la nature de la tâche à laquelle ils vont participer. Les sujets sont ensuite répartis dans les deux salles, chacun s'installant devant un des postes de travail disponibles.

Une première phase, dite de « familiarisation », est proposée à tous les sujets. Dans celle-ci, chaque sujet est connecté dans un espace privé dans lequel un carré noir est inscrit. L'objectif des sujets est de tenter de localiser ce carré en déplaçant leur curseur dans l'espace à l'aide du Touchpad. Les sujets sont informés que la rencontre de ce carré par leur curseur active la montée des stimulateurs tactiles.

Les sujets du groupe 1 sont informés que le carré possède la même taille et la même forme que les corps-image des autres participants.

Les sujets du groupe 2 sont informés que le carré possède la même taille et la même forme que les corps-image par défaut des autres participants, mais qu'ils peuvent modifier cette forme en exerçant une pression sur le FSR. Lors de cette phase, les sujets peuvent constater l'effet de leur propre déformation sur le carré qu'ils explorent. En d'autres termes, le carré exploré est en réalité le reflet statique de leur propre corps-image : si les sujets exercent une pression sur leur FSR tout en étant en contact avec le carré, ils peuvent ressentir cette déformation sur leur propre matrice de stimulateurs.

Dans la deuxième phase, dite de « rencontre », tous les sujets ont l'occasion d'interagir avec chacun des trois autres sujets successivement pendant 1 minute et 30 secondes chacun. Au début de chaque essai, l'identité de la personne rencontrée est précisée dans la consigne (Par exemple : « Vous allez interagir avec Bleu »). L'ordre des rencontres est fixé préalablement par l'expérimentateur et est identique pour toutes les passations. Lors de cette phase, il est précisé aux sujets qu'ils doivent faire très attention aux autres sujets, car ils devront essayer de les identifier dans la phase suivante. Par ailleurs, et contrairement à l'étude expérimentale précédente, il leur est précisé que chacun des trois autres sujets devront être capables de les identifier en retour.

Ensuite, la phase de test à proprement parler débute. Ici, les conditions sont en tous points identiques à la phase précédente, hormis le fait que l'identité des sujets n'est plus révélée dans la consigne. De même, l'ordre des rencontres est modifié. Chaque sujet rencontre donc successivement chacun des trois autres sujets pendant 1 minute et 30 secondes. A la fin de chaque essai, les sujets doivent indiquer sur une feuille de réponse l'identité du sujet avec lequel ils pensent avoir interagi. Aucun feedback concernant la qualité de leur réponse n'est fourni.

Les sujets sont informés qu'une réponse correcte entraîne le gain d'un point, qui est doublé si le partenaire de l'interaction donne également une réponse correcte.

A la fin de cette phase, une pause de quelques minutes est proposée aux sujets, pendant laquelle ils ne sont pas autorisés à parler. Ensuite, les sujets passent l'ensemble de l'expérimentation une seconde fois, en procédant de nouveau à la phase de rencontre puis à la phase de test. A la fin de l'expérience, les 4 sujets sont réunis dans la salle de l'ordinateur-serveur, et un débriefing est animé par l'expérimentateur, de façon à recueillir les impressions des sujets sur la tâche ainsi que l'explicitation de leurs stratégies éventuelles.

Nous pensons que la simplification du protocole expérimental va permettre aux sujets d'être plus attentifs aux interactions tactiles. En effet, les interactions deux à deux dans des espaces partagés va permettre aux sujets de se concentrer sur les particularités de chacun. Comme lors de la première étude expérimentale, nous posons plusieurs hypothèses :

- Tout d'abord, l'utilisation du MIT, spécialement conçu pour les interactions tactiles, devra favoriser la compréhension de l'espace et les actions des sujets. De cette façon, les sujets pourront discriminer les autres sujets et les identifier sur la base de la dynamique propre de leurs activités perceptives.
- Nous nous attendons à ce que les sujets du groupe 2, qui ont la possibilité de modifier dynamiquement la forme de leur corps-image, obtiendront de meilleurs pourcentages d'identification correcte que les sujets du groupe 1.
- Par ailleurs, quel que soit le groupe expérimental, les performances d'identification seront supérieures lors de la deuxième session expérimentale que lors de la première. En effet, la répétition des rencontres aura pour conséquence de stabiliser les dynamiques d'interaction et de rendre saillantes les différences interindividuelles.

3.4.4. Résultats

De façon globale, nous observons que les pourcentages d'identifications correctes sont plus élevés que lors de la première étude expérimentale. Ici encore, les sujets ont une chance sur trois de donner des réponses correctes : le score du hasard est donc fixé à 33.33%. Or, de façon globale, les sujets identifient correctement les autres sujets à hauteur de 53.13%, même si nous notons cette fois une dispersion des scores d'identification relativement élevée (s = 26,68) (voir Figure 36). Ce résultat est statistiquement différent du hasard (test de Student : $t_{(15)} = 3.0173$; $p = 0.0087$), même si la différence ne peut pas être considérée comme notable ($Pr\ |\delta/\sigma| > .66 = .60$).

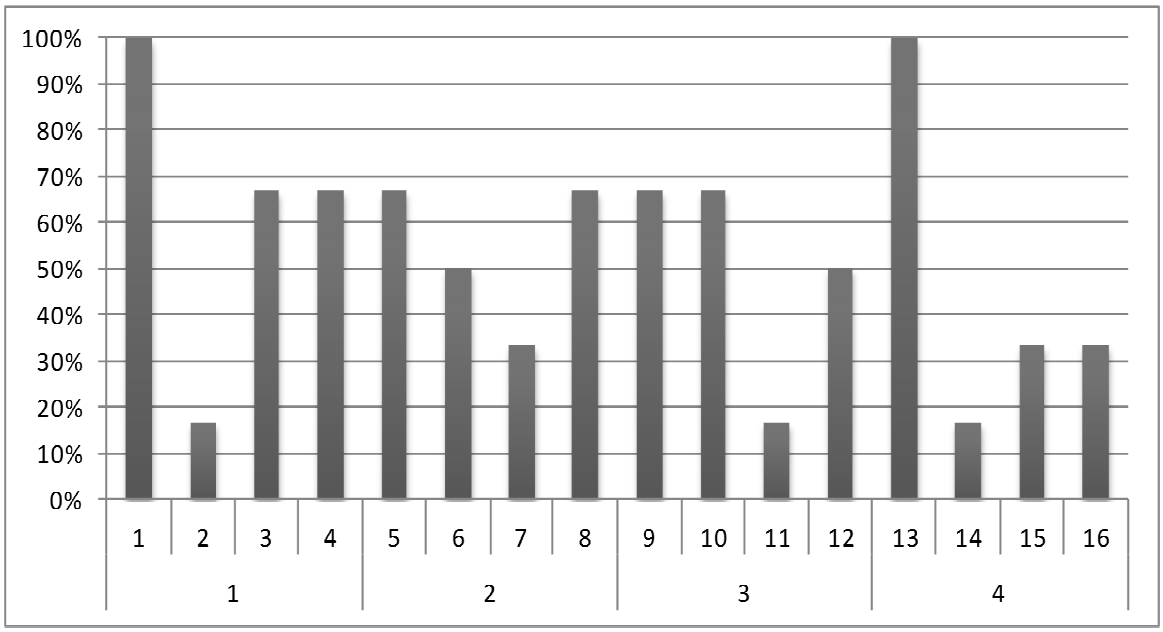


Figure 36 : Pourcentages d'identifications correctes pour chaque sujet (de 1 à 16) regroupés en fonction des 4 passations expérimentales

Parallèlement, nous observons que de façon globale, 35,42% des rencontres ont donné lieu à une identification mutuelle (les deux sujets se sont mutuellement identifiés), alors que 39,57% des rencontres ont entraîné une identification unilatérale (seul un des deux sujets a correctement identifié l'autre), et 25,01% des rencontres ont conduit à une non-identification mutuelle (aucun des deux sujets n'a identifié l'autre) (voir Figure 37). Toutefois, la différence globale entre les trois modalités de reconnaissance n'est pas significative (test de Kruskall-Wallis : $H_{(2,48)} = 4{,}046303$; $p = 0.1322$).

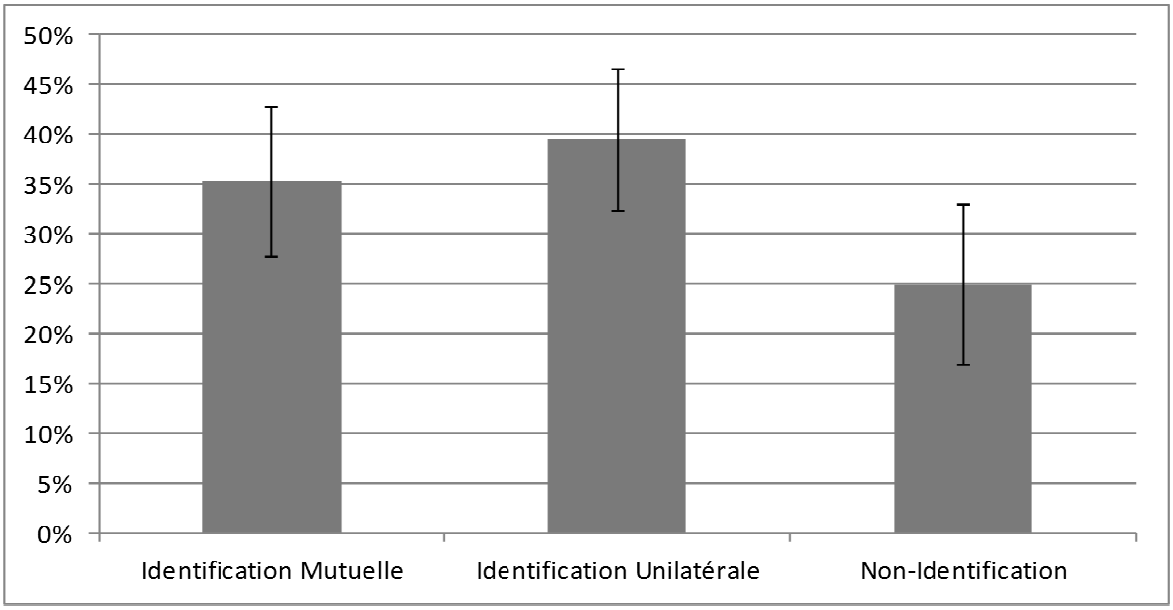


Figure 37 : Pourcentages d'identification mutuelle, d'identification unilatérale et de non-identification

Concernant les différences entre les deux groupes, nous observons que les sujets du groupe 1 obtiennent de meilleurs pourcentages d'identifications correctes que le groupe 2 (d = 10.42) (voir Figure 38). Cette différence n'est pas significative (test de Student : $t_{(14)}$ = 0.7702 ; p = 0.4540), mais elle ne peut pas être considérée comme négligeable (Pr |δ/σ| < .33 = .39).

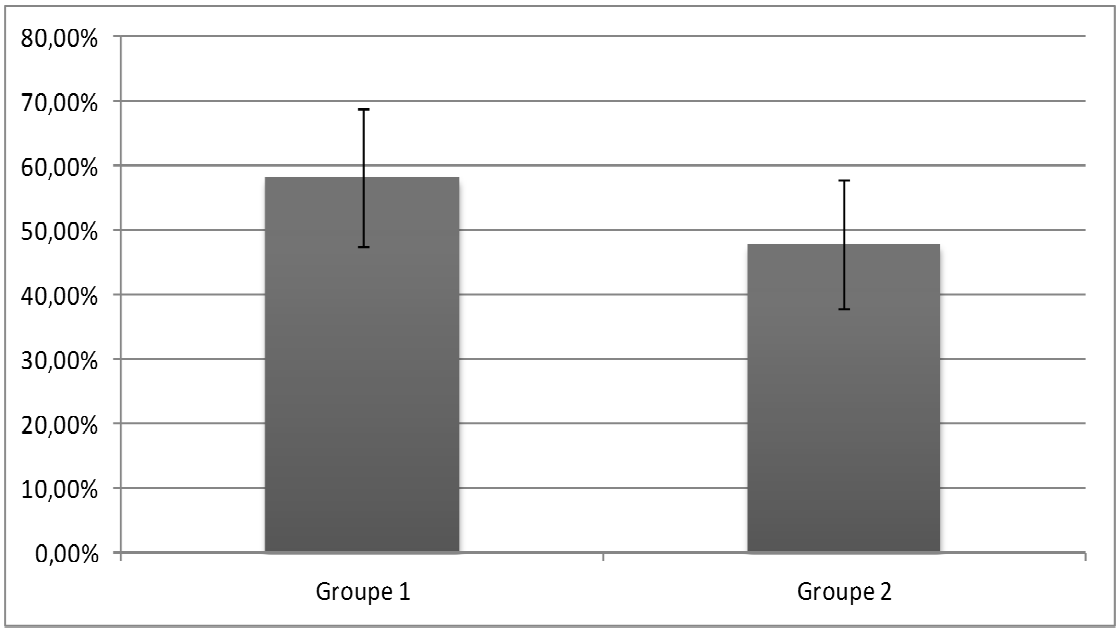


Figure 38 : Pourcentages d'identifications correctes en fonction des groupes expérimentaux (Groupe 1 = déplacement seul ; Groupe 2 = déplacement et déformation des corps-image)

Quel que soit le groupe expérimental, les sujets obtiennent de meilleures performances dans la deuxième session expérimentale que dans la première session (d = 14.58). Mais cette différence n'est pas significative (test de Student : $t_{(15)}$ = 1.8155 ; p = 0.0895), même si elle ne peut pas être considérée comme négligeable (Pr |δ/σ| < .33 = .33). Toutefois, nous observons que la différence entre les deux sessions est plus importante pour le groupe 2 (d = 20.83) que pour le groupe 1 (d = 8.33) (voir Figure 39). Pourtant, l'effet d'interaction entre le facteur « Groupe expérimental » et le facteur « Session expérimentale » n'est pas significatif (test de Student : $t_{(14)}$ = 0.7673 ; p = 0.4556), même s'il ne peut pas être considéré comme négligeable (Pr |δ/σ| < .33 = .39).

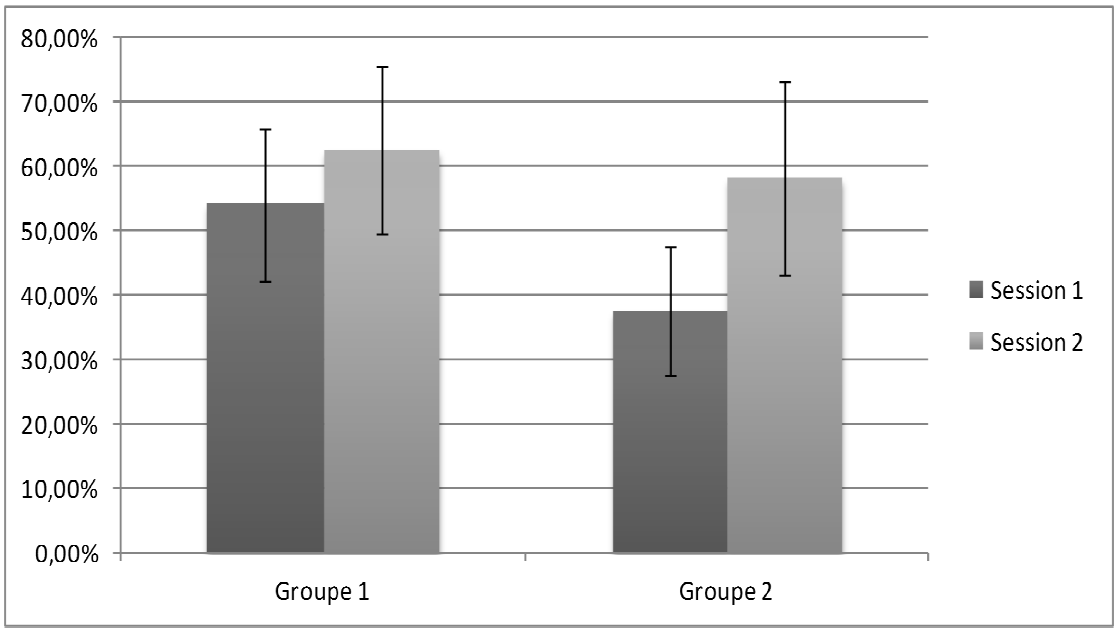


Figure 39 : Pourcentages d'identifications correctes en fonction du groupe expérimental (Groupe 1 = déplacement seul ; Groupe 2 = déplacement et déformation des corps-image) et des sessions expérimentales (Session 1 = première phase de test ; Session 2 = deuxième phase de test)

Si nous nous intéressons maintenant à la fréquence d'utilisation du FSR (chez les sujets du groupe 2), nous observons que globalement, celle-ci est peu élevée. En effet, les sujets ont modifié leur corps-image pendant 17.31% du temps total alloué aux interactions. De plus, nous observons que le FSR est de moins en moins utilisé à mesure que les phases expérimentales se succèdent (voir Figure 40). La différence globale d'utilisation du FSR en fonction des quatre phases expérimentales est significative (test de Kruskall-Wallis : $H_{(3,96)} = 20.01107$; $p = 0.0002$). Il est également à noter que les sujets utilisent moins le FSR pendant les phases de test des deux sessions expérimentales que pendant la phase de rencontre qui les précède (phases durant lesquelles les sujets rencontrent chacun des trois autres et où leur identité est révélée).

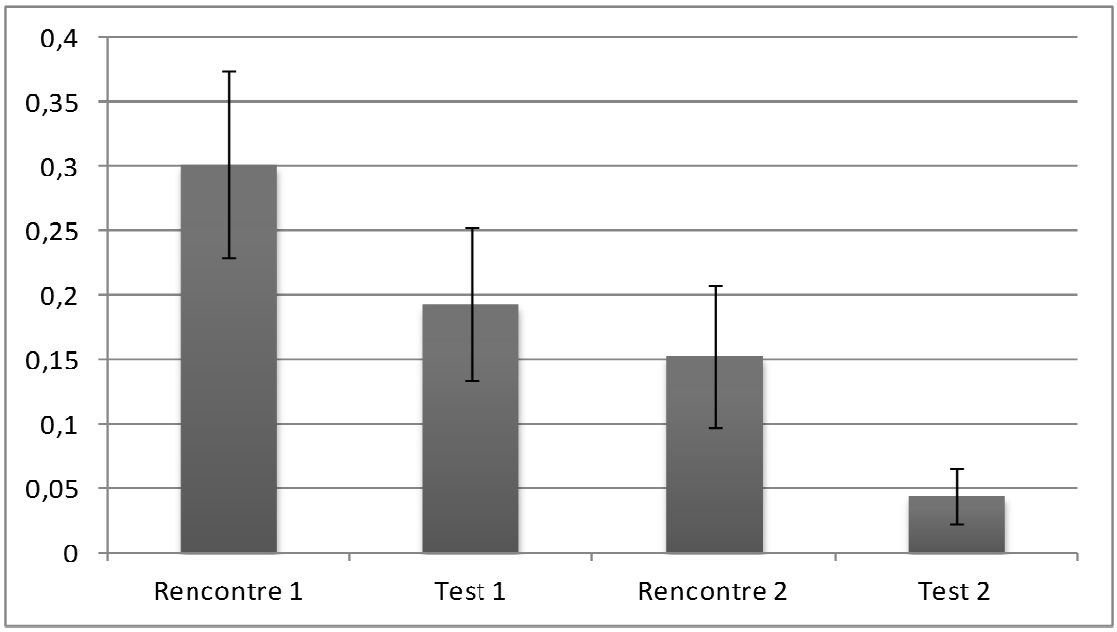


Figure 40 : Fréquences d'utilisation du FSR en fonction des phases expérimentales

De façon à préciser ces observations, nous nous penchons sur les profils individuels d'utilisation du FSR. Nous constatons alors que seuls deux sujets sur les huit ont utilisé systématiquement le FSR, mais en maintenant une pression statique, souvent déclenchée après le premier contact et relâchée après une période prononcée de perte de contact. Si la plupart des sujets (6 sur 8) ont systématiquement utilisé le FSR lors de la première phase de rencontre, seuls quatre d'entre eux ont continué à l'utiliser dans la première phase de test. Lors de la deuxième phase de rencontre, seuls deux sujets ont systématiquement utilisé le FSR, contre deux qui ne l'ont pas du tout utilisé et quatre qui l'ont utilisé dans au moins un essai sur les trois. Enfin, lors de la deuxième phase de test, seuls deux sujets ont systématiquement utilisé le FSR, mais de façon très brève lors des contacts, tandis que trois sujets ne l'ont pas du tout utilisé et que trois autres sujets l'ont utilisé seulement pendant un essai sur les trois.

De façon à vérifier l'impact de l'utilisation du FSR sur les capacités de discrimination interpersonnelle, nous regroupons d'un côté les scores obtenus par les sujets lors des essais où le FSR n'a pas été utilisé, et de l'autre les scores obtenus par les sujets lors des essais où le FSR a été utilisé.

Nous observons que le pourcentage moyen d'identifications correctes obtenu dans les essais où les sujets n'ont pas utilisé le FSR est supérieur au pourcentage obtenu dans les essais où les sujets ont utilisé le FSR (d = 12.49). Cette différence est significative (test de Mann-Whitney : U = 4 ; $Z_{ajusté}$ = 2,2536 ; p = 0.02419), même si l'effet peut être considéré comme faible (g de Hedge = 0.17). Par ailleurs, nous constatons que les essais où les sujets ont utilisé

le FSR n'ont pas entraîné de meilleurs pourcentages d'identification correcte pour les sujets qui ont interagi avec eux par rapport aux essais où les sujets n'ont pas utilisé le FSR (d = 9.72). Cette différence n'est pas significative (test de Mann-Whitney : U = 12 ; $Z_{ajusté}$ = 0.9607 ; p = 0.3089).

3.4.5. Analyse des trajectoires

Nous nous intéressons maintenant aux trajectoires perceptives déployées par les sujets lors des essais. Nous avons pour cela mené une analyse qualitative des trajectoires obtenues pour chaque essai, à partir de leur représentation graphique. Parallèlement, les commentaires des sujets obtenus lors du débriefing ont également apporté des pistes qui ont permis de guider nos choix. Cette analyse qualitative a permis de définir un ensemble d'indicateurs qui nous a semblé pertinent, et qui a donné lieu à une analyse quantitative des données[22]. L'analyse de ces traces numériques a révélé principalement deux indicateurs qui semblent pouvoir expliquer les résultats.

Tout d'abord, notre attention s'est portée sur la durée totale des stimulations reçues lors des essais en fonction de la qualité des réponses fournies par les sujets. Nous rappelons que puisque l'espace numérique partagé ne contient pas d'objet, les seules sources de stimulation résultent de la rencontre entre les deux sujets.

Nous observons que quel que soit le groupe expérimental, lorsque le sujet identifie correctement son partenaire, la durée totale moyenne des stimulations est plus élevée que lorsque le sujet se trompe (d = 11412.56 ms) (voir Figure 41). Cette différence est significative (test de Student : $t_{(26)}$ = 2.1859 ; p = 0.0380), mais elle ne peut pas être considérée comme notable (Pr $|\delta/\sigma| > .66 = .66$).

[22] Pour des détails concernant la méthodologie d'analyse des trajectoires, veuillez vous reporter à l'Annexe B, p. 259

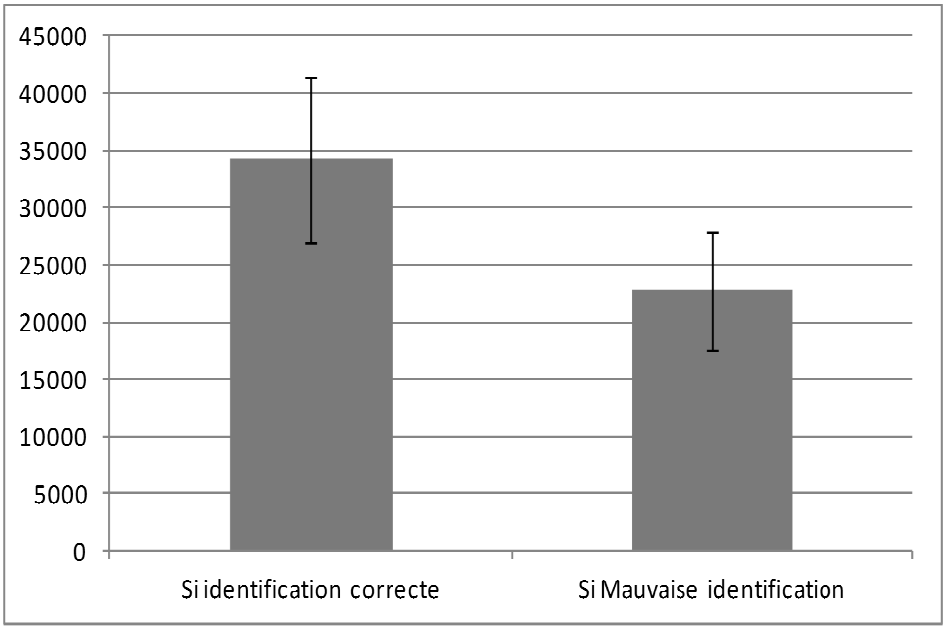


Figure 41 : Durée totale moyenne des stimulations en fonction de la qualité des réponses fournies par les sujets (Identification correcte versus Mauvaise identification)

Parallèlement, nous nous intéressons à la dispersion des sujets dans l'espace au cours des essais. La dispersion est un indicateur calculé à partir du nombre total de positions différentes occupées par un sujet pendant un essai donné divisé par le nombre total de positions possibles. Plus le résultat tend vers 1, plus grande est la dispersion, c'est-à-dire que le sujet a occupé un grand nombre de positions différentes tout au long de l'essai. Par exemple, une dispersion de 1 indiquerait que le sujet a occupé toutes les positions possibles de l'espace. Au contraire, un résultat faible (qui tend vers 0) indique que l'activité du sujet s'est concentrée dans une petite partie de l'espace.

Nous observons que lorsque le sujet identifie correctement son partenaire, la dispersion de son activité dans l'espace est plus faible que lorsqu'il fournit une mauvaise réponse ($d = -0.0018$) (voir Figure 42). Cette différence est significative (test de Student : $t_{(26)} = 5.5177$; $p < 0.0001$), et peut même être considérée comme notable ($Pr\ |\delta/\sigma| > .66 = .98$).

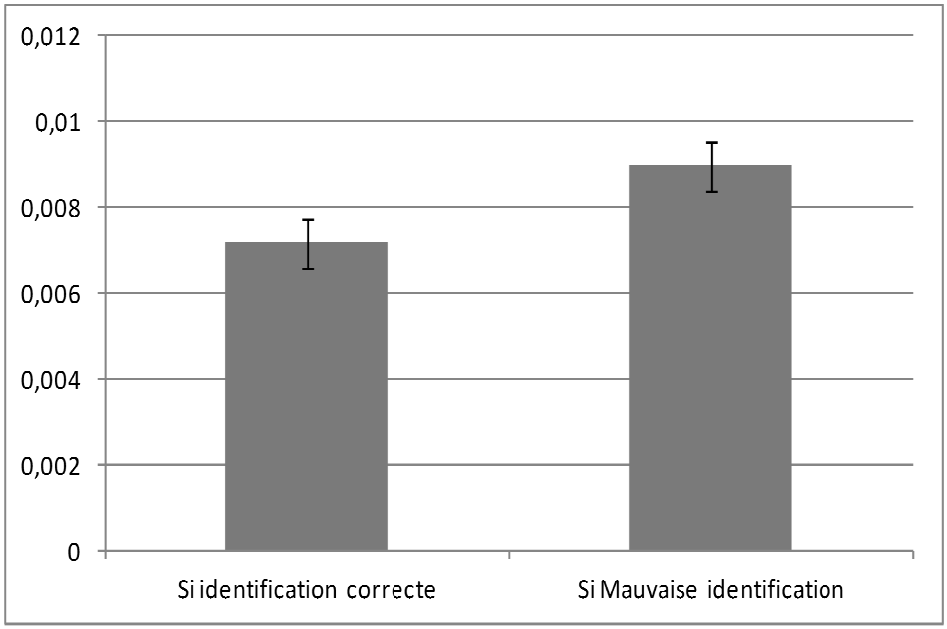


Figure 42 : Dispersion moyenne des sujets en fonction de la qualité de la réponse (Identification correcte versus Mauvaise identification)

De façon à préciser ces résultats, nous nous intéressons à l'évolution de ces indicateurs en fonction des modalités d'identification collective des binômes. En effet, il apparaît que la durée totale moyenne des stimulations diffère en fonction des modalités d'identification (voir Figure 43). Cette différence est significative (test de Fisher : $F_{(2,34)} = 9.2877$; $p = 0.0006$).

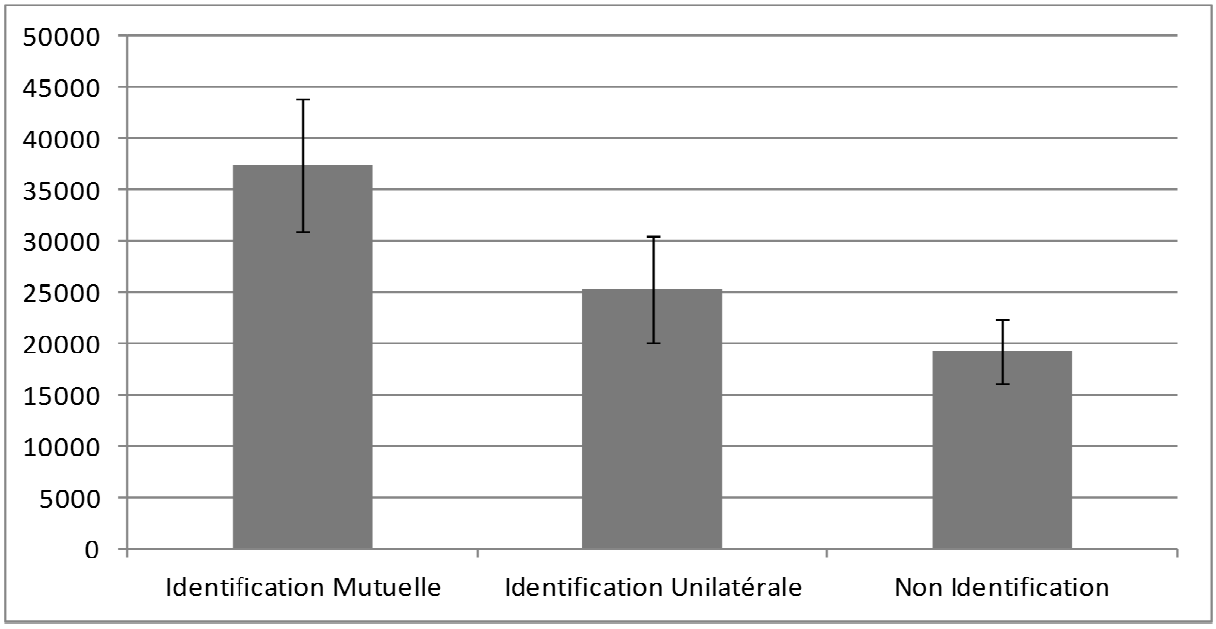


Figure 43 : Durée totale moyenne des stimulations en fonction des modalités d'identification

Nous observons que la durée totale moyenne des stimulations est plus élevée au cours des essais ayant conduit à une identification mutuelle comparativement aux essais ayant

conduit à une identification unilatérale (d = 12154.78 ms). Cette différence est significative (test de Student : $t_{(25)}$ = 2.7173 ; p = 0.0118), mais elle ne peut pas être considérée comme notable (Pr |δ/σ| > .66 = .82). De la même façon, la durée totale moyenne des stimulations est plus élevée durant les essais ayant conduit à une identification mutuelle par rapport aux essais ayant conduit à une non-identification (d = 18219.92 ms). Cette différence est significative (test de Student : $t_{(21)}$ = 4.1201 ; p = 0.0005), et peut même être considérée comme notable (Pr |δ/σ| > .66 = .95). Enfin, la durée totale moyenne des stimulations est plus élevée lors des essais ayant conduit à une identification unilatérale par rapport aux essais ayant conduit à une non-identification (d = 6065.14 ms). Cette fois, la différence n'est pas significative (test de Student : $t_{(22)}$ = 1.6440 ; p = 0.1144), même si elle ne peut pas être considérée comme négligeable (Pr |δ/σ| < .33 = .20).

De la même façon, nous nous intéressons à la dispersion moyenne en fonction des trois modalités d'identification. En effet, nous observons que d'une façon globale, la dispersion est plus importante dans le cas des essais ayant conduit à une non identification (voir Figure 44) que dans le cas des essais ayant conduit à une identification unilatérale (voir Figure 45). Par ailleurs, les essais ayant conduit à une identification mutuelle semblent conduire à la dispersion moyenne la plus faible (voir Figure 46). Ainsi, nous observons que la dispersion varie en fonction des réponses collectives fournies par les binômes (voir Figure 47). La différence globale est significative (test de Fisher : $F_{(2,44)}$ = 13.2216 ; p < 0.0001).

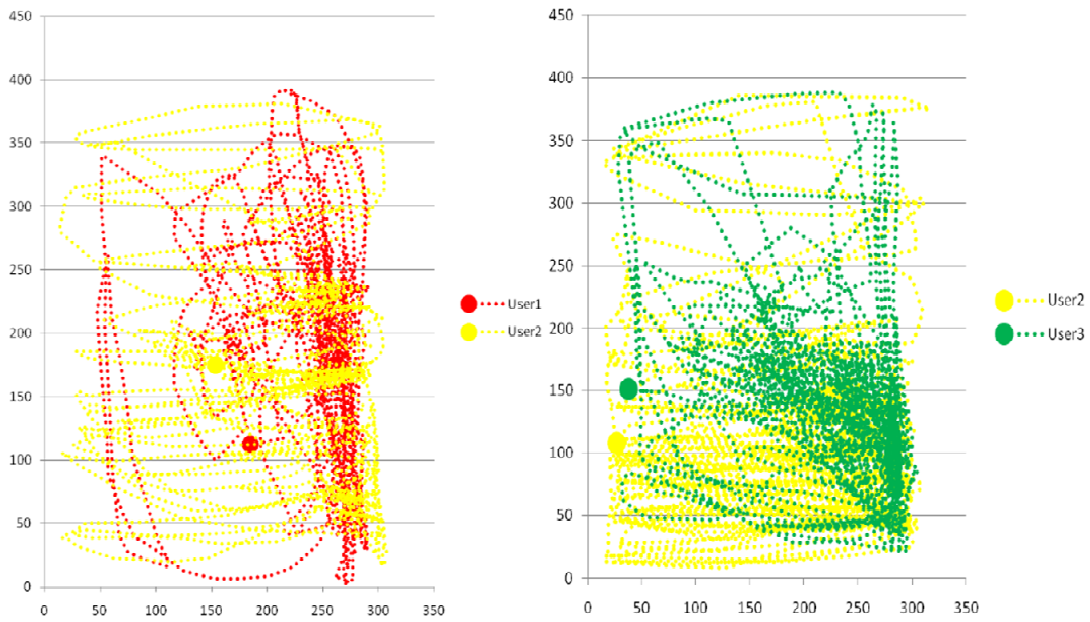


Figure 44 : Exemples de trajectoires obtenues lors d'essais ayant conduit à une non-identification

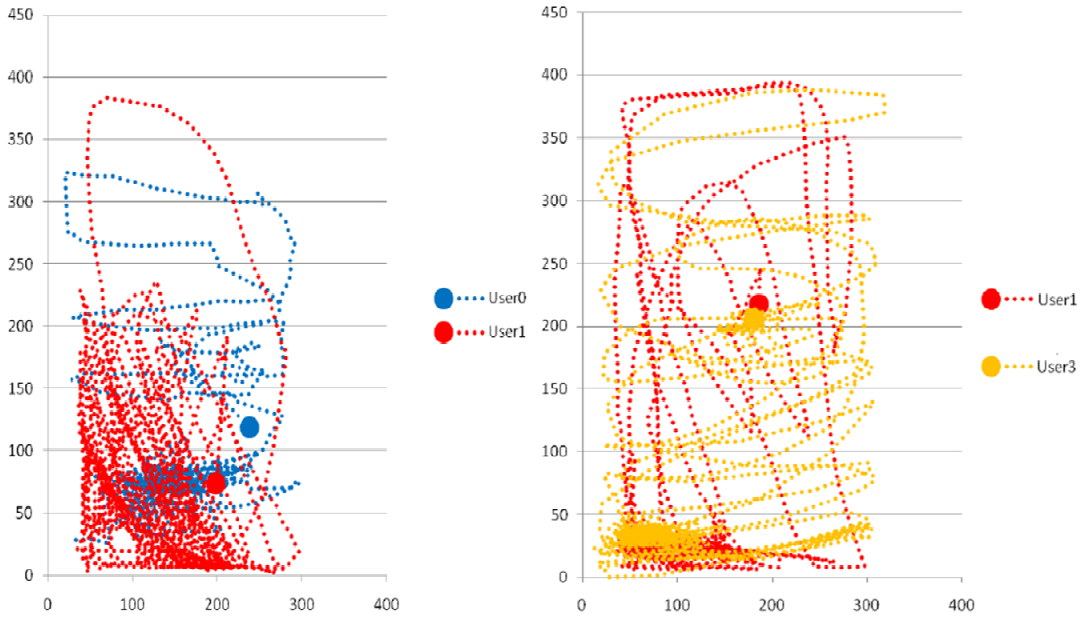


Figure 45 : Exemples de trajectoires obtenues lors d'essais ayant conduit à une identification unilatérale

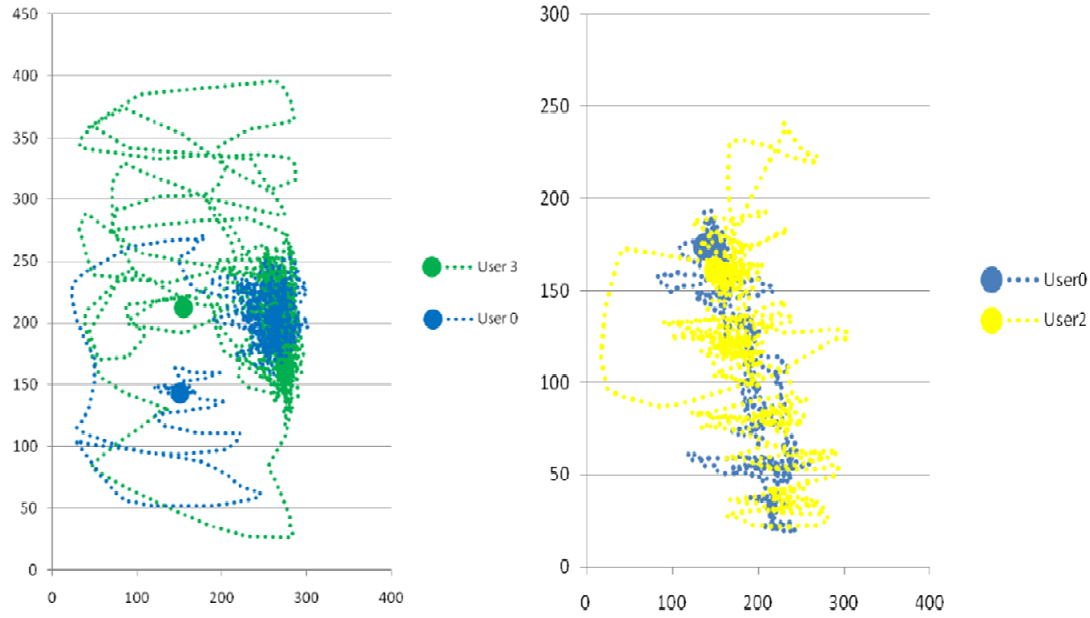


Figure 46 : Exemples de trajectoires obtenues lors d'essais ayant conduit à une identification mutuelle

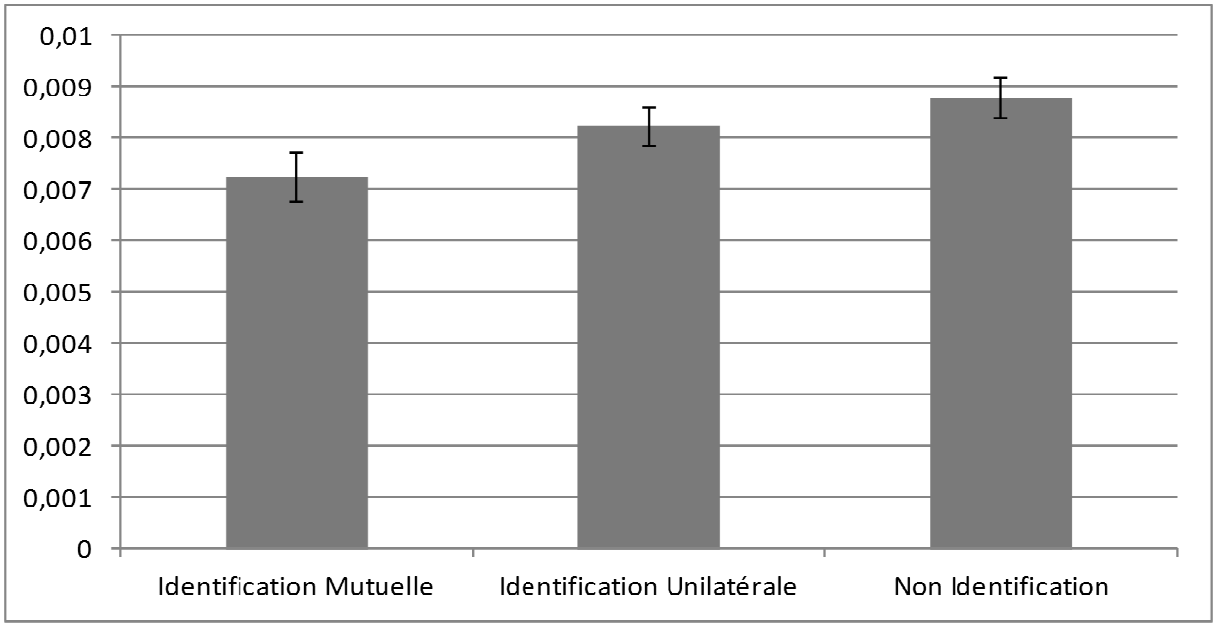


Figure 47 : Dispersion moyenne des sujets en fonction des trois modalités d'identification

Si nous comparons les modalités d'identification deux à deux, nous observons que la dispersion moyenne est plus faible dans le cas des essais ayant conduit à une identification mutuelle par rapport aux essais ayant entraîné une identification unilatérale (d = - 0.0010). Cette différence est significative (test de Student : $t_{(30)} = 3.1895$; $p = 0.0033$), mais elle ne peut pas être considérée comme notable (Pr $|\delta/\sigma| > .66 = .85$). De la même façon, la dispersion moyenne est plus faible dans les cas des essais ayant conduit à une identification mutuelle par rapport aux essais ayant entraîné une non-identification (d = - 0.0015). Cette différence est significative (test de Student : $t_{(29)} = 4.7950$; $p < 0.0001$), et peut même être considérée comme notable (Pr $|\delta/\sigma| > .66 = .97$). Enfin, la dispersion moyenne est plus faible durant les essais ayant conduit à une identification unilatérale par rapport aux essais ayant entraîné une non-identification (d = - 0.0006). Cette différence est significative (test de Student : $t_{(29)} = 2.0483$; $p = 0.0497$), mais ne peut pas être considérée comme notable (Pr $|\delta/\sigma| > .66 = .57$).

3.4.6. Discussion

Notre hypothèse globale était basée sur le fait que les sujets allaient démontrer des capacités de discrimination interpersonnelle, grâce notamment à un protocole simplifié qui favorise les interactions entre les sujets. Cette hypothèse est vérifiée, puisque quel que soit le groupe expérimental, nous observons que les pourcentages d'identification correcte sont significativement différents du hasard. Par conséquent, nous pouvons conclure que dans les

conditions de l'expérience, les sujets sont parvenus à s'identifier les uns les autres lors des rencontres successives, et ce, à un degré supérieur à celui du hasard.

Pour autant, les résultats ne suggèrent pas que cette capacité implique une identification nécessairement réciproque. En effet, nous ne pouvons pas conclure à une différence entre les trois modalités d'identification, les pourcentages d'identification mutuelle, d'identification unilatérale et de non- identification ne présentant pas de différence significative.

Pour tenter de savoir si les conditions expérimentales ont eu un effet sur les pourcentages d'identification correcte, nous avons comparé les résultats obtenus par les deux groupes de sujets. Pour rappel, alors que le groupe 1 ne disposait que d'une seule possibilité d'action (le déplacement dans l'espace bidimensionnel par l'intermédiaire du Touchpad), le second bénéficiait d'une possibilité d'action supplémentaire (la déformation dynamique du corps-image via le FSR). Contrairement à notre hypothèse, nous avons constaté que les sujets du groupe 1 ont obtenu de meilleurs pourcentages d'identification correcte que les sujets du groupe 2. Par ailleurs, si nous ne pouvons pas conclure à un effet global d'apprentissage entre les deux sessions expérimentales, nous avons observé que la différence entre les deux sessions est plus importante pour le groupe 2. Malgré cela, il est important de noter que les résultats du groupe 2 dans la deuxième session n'égalent pas les résultats du groupe 1 dans cette session. L'évolution des performances du groupe 2 entre les deux sessions peut alors être expliquée par l'abandon progressif du FSR au fur et à mesure de l'expérience.

En effet, comme nous l'avons vu, les pourcentages d'utilisation du FSR par les sujets du groupe 2 se dégradent au fur et à mesure des phases expérimentales. En effet, si les sujets tentent de l'utiliser lors de la première phase de rencontre, la fréquence d'utilisation du FSR diminue dès la phase de test suivante. Il en est de même pour les deux phases de la deuxième session expérimentale. Lors du débriefing, les sujets ont en effet exprimé une grande difficulté à utiliser conjointement les deux effecteurs, ce qui les a conduit à progressivement n'utiliser que la possibilité de déplacement dans l'espace bidimensionnel. Pour les sujets qui ont le plus utilisé la possibilité de déformation, il s'avère qu'ils ont opté pour une pression statique, de façon à imprimer une forme particulière à leur corps-image. Toutefois, pour ces sujets, nous ne pouvons pas conclure à un effet de cette « démarcation physique » sur les pourcentages d'identification correcte de leurs partenaires lors des interactions. En effet, au lieu de favoriser leur identification par les autres sujets, l'utilisation du FSR influence paradoxalement les pourcentages d'identification de ceux qui l'utilisent.

Ce résultat nous semble particulièrement intéressant, dans la mesure où nous pouvons interpréter ce résultat au regard des trajectoires déployées dans l'espace numérique. Le fait de

modifier son corps-image, même de façon statique, semble avoir une influence sur le processus de l'interaction, ou du moins sur le comportement interactif du partenaire. En effet, la forme du corps-image rencontré par un sujet peut influencer la dynamique de son activité perceptive, en l'engageant dans une exploration active d'une source de stimulation qui diffère des autres sources rencontrées. La rencontre d'un corps-image différent des autres peut alors modifier le comportement du sujet en faisant varier la dynamique de son mouvement d'exploration. De cette façon, le sujet qui a modifié son corps-image se trouve stimulé par une activité perceptive plus active, ce qui a pu rendre saillant sa différence par rapport aux autres activités perceptives rencontrées. Ce résultat renforce alors notre conviction selon laquelle la dynamique propre de l'activité perceptive des sujets peut rendre possible la discrimination interpersonnelle. Par ailleurs, lors des débriefings, seuls 31% des sujets (5 sujets sur 16) ont exprimé avoir essayé de déployer un signe distinctif. Lors de l'analyse des trajectoires, nous avons observé que ces signes ont la plupart du temps consisté en un déploiement intentionnel d'une oscillation unidimensionnelle rapide, la plupart du temps sur un plan horizontal (3 sur 5), et sur un plan vertical (1 sur 5). Le dernier sujet a déployé un mouvement circulaire de faible amplitude. Pour autant, ces sujets n'ont pas plus été identifiés par les autres sujets. De ce fait, il semble que les résultats soient expliqués en premier lieu par la reconnaissance d'une dynamique spécifique de l'activité perceptive, qui se manifeste dans l'interaction en tant que signature non-intentionnelle plutôt qu'en tant que signe intentionnellement déployé.

L'analyse des trajectoires perceptives nous semble apporter des éléments complémentaires. En effet, les deux indicateurs de trajectoire que nous avons isolés varient selon que les essais ont conduit à une identification correcte ou non : la durée des stimulations est importante et la dispersion est faible lorsque le sujet identifie correctement son partenaire. Plus encore, ces deux indicateurs semblent encore plus sensibles aux trois modalités d'identification collective. En effet, lorsque l'identification est mutuelle, la durée des stimulations est la plus importante et la dispersion est la plus faible. Cela signifie que lorsque les deux sujets se sont mutuellement identifiés, c'est à la suite d'un essai où les deux activités perceptives se sont organisées l'une par rapport à l'autre. Cette organisation semble se manifester via un contrôle réciproque des sources de stimulations, qui ont conduit les sujets à s'accrocher perceptivement et à démultiplier les stimulations dans une zone restreinte de l'espace. Cette organisation conduit à l'émergence d'une dynamique spécifique d'interaction, qui semble être une condition nécessaire et suffisante pour permettre une discrimination interpersonnelle. En effet, l'accroche perceptive mutuelle permet au sujet de percevoir la dynamique de l'activité perceptive de son partenaire, mais également d'organiser sa propre

activité perceptive par rapport à elle. De cette façon, les sujets sont en mesure d'identifier dans les patterns d'interaction les caractéristiques nécessaires à l'identification de l'interactant, tout en manifestant en retour une dynamique perceptive qui garantit l'identification mutuelle. La discrimination interpersonnelle implique donc minimalement une coordination des activités perceptives.

4. Vers des questionnements théoriques

4.1. L'interaction sociale comme objet d'étude

Notre travail est avant tout guidé par des préoccupations pragmatiques : il s'agit d'élaborer des espaces d'interaction tactile, rendus possible par la connexion en réseau de plusieurs dispositifs de suppléance perceptive (Tactos). Pour cela, notre objectif consiste à définir les conditions techniques et matérielles pour que des utilisateurs puissent interagir dans de tels espaces, et pour que cette interaction ait du sens. Par conséquent, la réflexion sur l'usage d'un tel dispositif, outre les problématiques de conception que nous avons relaté dans la première partie de ce travail, font émerger un ensemble de questions fondamentales. Nous souhaitons donc articuler une recherche appliquée, centrée sur la conception d'un dispositif pertinent à partir de l'analyse de son usage, à une recherche fondamentale, centrée sur les questions théoriques soulevées par l'usage lui-même.

Par exemple, les résultats que nous avons obtenus lors des études préliminaires concernant la discrimination interpersonnelle suggèrent que les activités perceptives des participants s'organisent d'une façon particulière. C'est cette organisation qui semble expliquer comment les utilisateurs peuvent d'identifier les uns des autres dans les espaces numériques partagés. Ce constat pose la question du ou des mécanismes qui rendent possible la coordination interpersonnelle et l'émergence de sens dans l'interaction sociale.

D'une façon générale, l'interaction sociale désigne la relation interpersonnelle entre au moins deux individus. Cette relation présuppose une co-présence des individus, mais aussi une influence réciproque de leurs comportements. Classiquement, cette inter-influence est décrite de façon séquentielle, en définissant l'interaction comme un cadre général de communication : un émetteur fournit un message, un récepteur en est le destinataire, et l'ensemble du processus est réduit à une simple transmission (Shannon, 1948). De ce point de vue, le succès de l'interaction est soumis aux aptitudes individuelles des partenaires sociaux à traiter l'information issue de l'environnement social. En psychologie, cette aptitude réfère à la

cognition sociale. D'une façon générale, celle-ci désigne l'ensemble des processus cognitifs impliqués dans les interactions sociales chez l'homme.

Tout comme l'ensemble des compétences humaines, la cognition sociale a été particulièrement étudiée dans le cadre du développement de l'enfant. Chez l'adulte, l'étude des mécanismes de la cognition sociale peut en effet se révéler problématique, dans la mesure où ceux-ci se trouvent intégrés dans la globalité du comportement perceptif, cognitif et social. De ce fait, l'adulte manifeste des capacités de compréhension et d'adaptation sociale composites, où les fondements ontogénétiques se sont trouvés combinés avec des construits cognitifs plus complexes. L'étude de la genèse de ces capacités nous semble donc particulièrement pertinente pour en comprendre ses modes d'organisation et de fonctionnement, tout comme ses liens éventuels avec d'autres phénomènes. En effet, « (…) l'enfant explique l'homme autant, et souvent plus, que l'homme n'explique l'enfant » (Piaget & Inhelder, 1966, p.6). Dans notre cas, cette posture génétique nous paraît d'autant plus justifiée que nous offrons aux utilisateurs un mode d'interaction inédit. En effet, nous avons une occasion originale d'observer la constitution même des interactions sociales et de la compréhension interpersonnelle, quand les moyens disponibles aux utilisateurs ne sont pas déjà constitués à priori.

Mais dans cette entreprise, nous souhaitons éviter un « individualisme méthodologique » qui s'est imposé dans l'étude de la cognition sociale (Lenay & Stewart, 2012 ; Froese & Di Paolo, 2011) : les protocoles expérimentaux visant à élucider la compréhension interpersonnelle ne permettent pas de réelles interactions entre les sujets. La plupart du temps, en effet, des sujets passifs observent des situations sociales extérieures, puis sont invités à porter un jugement après-coup (voir par exemple Frith, 2008 ; Van Overwalle, 2009 ; Decety, Jackson, Sommerville, Chaminade & Meltzoff, 2004). Les recherches concernant la cognition sociale restent alors focalisées sur les processus socio-cognitifs à la troisième personne (Frith, 2008), qui impliquent une nécessaire « attitude téléologique ». Comprendre les situations sociales revient à se représenter les actions d'autrui comme le produit d'états mentaux intentionnellement dirigés vers un but rationnel (Csibra & Gergely, 1998 ; Gergely & Csibra, 2003). L'interaction sociale est alors considérée comme le but à atteindre via une mentalisation explicite ou implicite (Frith & Frith, 2003), ou encore comme un simple input sur lequel s'opèrent des processus internes (Gergely & Watson, 1996 ; 1999 ; Tomasello, 1995 ; 1999).

En réaction à la prédominance de cette conception individualiste, un renversement épistémologique est proposé : plutôt que d'expliquer la compréhension sociale par des

opérations internes, il s'agit de réhabiliter la participation active des individus à la constitution de sens (Thompson, 2007 ; Varela, Thompson & Rosch, 1991). Ainsi, le monde d'est pas pré-donné, mais est le produit d'interactions pragmatiques et dynamiques entre l'agent et son environnement (y compris l'environnement social) (Varela et al., 1991). Par conséquent, l'unité d'analyse de la cognition sociale ne doit pas être l'individu, mais le système de couplage lui-même, c'est-à-dire le processus de l'interaction qui implique les participants, leurs interactions dynamiques et le contexte dans lequel ces interactions s'inscrivent (De Bruin & De Haan, 2012). De cette façon est rendue possible une conception de la cognition sociale qui réhabilite d'une part l'expressivité corporelle et l'action, et d'autre part l'engagement mutuel pour la compréhension interpersonnelle (De Jaegher, 2006 ; Krueger, 2011 ; Reddy & Morris, 2004). Il s'agit donc de prendre en compte plus sérieusement le processus même de l'interaction, qui implique un jeu complexe de coordinations spatio-temporelles (Fogel, 1993 ; Stern, 1977). Cette posture pleinement interactionniste, issue de la conception micro-sociologique des processus sociaux (Mead, 1934 ; Goffman, 1959), postule que l'interaction elle-même conduit à l'émergence sémiotique et à l'organisation des comportements individuels en société. Ainsi, le processus même de l'interaction est considéré comme une alternative en soi aux mécanismes individuels, en considérant les dynamiques interpersonnelles comme constitutives de la cognition sociale (De Jaegher, Di Paolo & Gallagher, 2010).

4.2. Le minimalisme comme cadre d'analyse

Pour se donner les moyens d'étudier concrètement le processus de l'interaction, nous avons opté pour une méthodologie minimaliste (Lenay, 2006c). Celle-ci est rendue possible par la mobilisation de médiations techniques telles que notre dispositif de suppléance perceptive. Nous avons suivi ce principe dans ses deux dimensions complémentaires : le ***minimalisme expérimental*** et le ***minimalisme fondamental***.

Dans sa dimension expérimentale, le minimalisme désigne une forme de réduction expérimentale particulière, dont elle reprend les principes de base mais en poussant davantage le raisonnement. La réduction expérimentale consiste à réduire la complexité d'une situation ou d'un phénomène en neutralisant certaines variables, et en contrôlant celles qui restent. Il s'agit donc de simplifier pour se donner les moyens de comprendre ce que l'on étudie (Ali Ammar, 2005). Avec le minimalisme expérimental, il s'agit de contrôler jusqu'à l'activité même du sujet, notamment en contraignant à l'extrême ses entrées sensorielles et ses moyens d'action, de façon à pouvoir observer concrètement leur déploiement dans l'espace et dans le

temps. En d'autres termes, il s'agit d'externaliser l'ensemble des activités perceptives et cognitives de façon à les rendre accessibles à l'analyse. Il est alors question d'identifier les conditions minimales, nécessaires et suffisantes pour qu'un phénomène donné émerge. Lorsque ces conditions sont établies, nous pouvons alors complexifier progressivement la situation, les entrées sensorielles ou les moyens d'action sur la base de ses dimensions les plus élémentaires.

Le minimalisme expérimental va de paire avec le minimalisme fondamental. En effet, cette méthode nous semble particulièrement appropriée dans le cadre de l'étude des mécanismes perceptifs. Elle reprend dans ce sens une certaine tradition philosophique concernant la compréhension de la perception et de l'expérience perceptive, en la décomposant notamment à travers le prisme de la technique. Par exemple, nous pouvons citer Merleau-Ponty qui pose les bases d'une étude phénoménologique de la perception : « Considérons-la [la perception] avant toute élaboration notionnelle. (…) Il faut nous adresser à quelque cas exceptionnel, où elle se défasse et se refasse sous nos yeux, par exemple, aux cas de vision sans inversion rétinienne[23]. » (Merleau-Ponty, 1945, p. 282). Dans le cadre du dispositif Tactos, en effet, nous offrons aux utilisateurs un mode de perception inédit, qui ne repose pas sur des connaissances constituées à priori (Lenay et al., 2000)[24]. L'utilisateur doit alors *exercer* un nouveau mode de couplage, de façon à extraire de nouvelles régularités sensori-motrices. En tant que médiation technique originale, le dispositif Tactos a alors permis l'étude d'une forme de ***néo-genèse*** de l'acte perceptif : la constitution même de l'acte de percevoir est rendue disponible à l'analyse, permettant ainsi l'étude du déploiement de cette activité d'un point de vue génétique.

Dans notre cas, nous posons l'hypothèse que la mise en réseau de ce dispositif ouvre la voie vers l'étude de la ***néo-genèse de l'interaction***. Notre dispositif présente en effet une occasion originale pour explorer les processus même de la rencontre interpersonnelle, des modes d'engagement individuels et collectifs, des dynamiques d'interaction et de la co-constitution d'un monde commun, empreint de significations et de valeurs partagées. De cette façon, il sera intéressant d'être particulièrement attentif à ce qui caractérise l'échange interindividuel et l'organisation de l'interaction à travers la coordination des activités

[23] Merleau-Ponty fait ici référence aux « lunettes de Stratton », qui, par un jeu de miroirs réfléchissants, permet d'inverser le haut et le bas, ou la droite et la gauche dans l'expérience perceptive du sujet (Stratton, 1896, cité par Merleau-Ponty, 1945).

[24] Bien entendu, nous pouvons considérer que la constitution de cette nouvelle modalité repose sur un savoir-faire perceptif déjà présent pour l'utilisateur. Toutefois, une réorganisation fonctionnelle est nécessaire, dans la mesure où i) la stimulation tactile n'a rien à voir avec une sensation visuelle, et ii) la manipulation d'un effecteur n'a rien à voir avec les mouvements oculaires ou les mouvements de la tête (Lenay & Stewart, 2012).

perceptives individuelles, que ce soit dans des contextes d'interaction strictement dyadique (sujet-sujet) (voir Partie 2, p.115), ou dans des contextes d'interaction dyadique en présence d'objets partagés (sujet-sujet-objet) (voir Partie 3, p.183).

PARTIE 2 : COGNITION SOCIALE ET INTERACTION

1. Des positions classiques

Rencontrer autrui dans un espace numérique d'interaction tactile soulève de nombreuses questions pratiques. Par exemple, nous nous demandons quelles sont les conditions techniques minimales pour qu'une telle interaction soit possible. Parallèlement, rendre possible ces interactions ne suffit pas. Il est nécessaire que celles-ci aient du sens pour les individus impliqués, ce qui pose des questions fondamentales sur la nature de la cognition sociale, et sur les liens entre celle-ci et l'interaction entre les individus. Si nous assumons que notre dispositif rend possible une étude de la néo-genèse de l'interaction chez l'adulte, les études concernant les interactions précoces et la genèse de la cognition sociale nous semblent pertinentes pour trouver des éléments de réponse.

De façon classique, la cognition sociale est décrite comme une aptitude strictement individuelle, formulée en termes mentalistes. Puisque des états mentaux internes sont nécessairement à l'origine du comportement, et que ces états mentaux ne sont pas observables, une « théorie de l'esprit » est nécessaire pour expliquer et prédire le comportement d'autrui (Premack & Woodruff, 1978). Deux positions classiques se sont opposées concernant la nature de cette théorie. Il s'agit pour les uns d'un ensemble de connaissances sur les états mentaux organisé de façon proto-scientifique (la ***théorie de la théorie***, ou TT) ; tandis que pour les autres, il s'agit d'une capacité à simuler les états mentaux d'autrui, c'est-à-dire à les reproduire ou à les imiter de façon interne, de manière à les comprendre via ses propres états mentaux (la ***théorie de la simulation***, ou ST).

1.1. La théorie de la théorie (TT)

Le principe de la théorie de la théorie (TT) repose sur l'idée que les connaissances de l'individu concernant les états mentaux sont organisées sous forme de règles et constituent un ensemble théorique, ou *folk psychology* (Sellars, 1956). La théorie de l'esprit serait donc un ensemble structuré de connaissances sur l'esprit, qui concerne notamment les intentions, les désirs et les croyances (Astington & Gopnick, 1991). Lorsque l'individu est confronté à une situation sociale ou interagit avec autrui, un système d'inférences permet de mettre en relation les informations disponibles avec la théorie de l'esprit stockée en mémoire.

Ce principe général étant communément admis, le débat a porté sur la nature des mécanismes impliqués (Leslie, 1987 ; Fodor, 1992 ; Baron-Cohen, 1995 ; Scholl & Leslie, 1999). Par exemple, selon Baron-Cohen, deux mécanismes innés sont à l'œuvre dans la compréhension sociale précoce : un mécanisme de détection de l'intentionnalité (ID) et un mécanisme de détection du regard d'autrui (EDD) (Baron-Cohen, 1995). Ces deux mécanismes traitent les relations intentionnelles entre le soi et autrui, en permettant notamment l'élaboration de représentations dyadiques. L'exercice précoce de ces relations permet l'élaboration progressive d'un mécanisme d'attention partagée (SAM), cette fois dirigé vers un troisième terme lors de l'interaction. Cette étape cruciale permet l'élaboration de représentations triadiques, par l'intermédiaire de nouvelles conduites telles que l'attention conjointe ou le pointage. Au final, ID, EDD et SAM permettent la lecture des états mentaux volitionnels (désirs et buts) et des états mentaux perceptifs (voir), puis de partager ces états au travers de l'interaction sociale. L'ensemble des connaissances qui va découler de cette « lecture mentale » va se trouver intégré dans un dernier module, dit de théorie de l'esprit (TOMM). Celui-ci organise les connaissances sur l'esprit d'une façon proto-scientifique, ce qui permet à l'individu la lecture des états mentaux épistémiques (tels que le faire-semblant, la pensée, le savoir ou la croyance) (ibid.).

D'autres chercheurs, comme Leslie, postulent que la compréhension des états mentaux serait intimement liée à la capacité de prétendre : se représenter un état mental serait équivalent à se représenter une représentation. La théorie de l'esprit serait donc rendue possible par un mécanisme méta-représentationnel (Leslie, 1987). Dans un premier temps, l'enfant élabore un ensemble de représentations primaires, relevant de connaissances « vraies » concernant le monde objectif. Puis, vers l'âge de deux ans, un mécanisme de « découplage » permet à l'enfant de se détacher de ses propres représentations de façon à élaborer des représentations de ces représentations, en y intégrant de nouvelles propriétés (ibid.). Ce mécanisme est marqué de façon essentielle par l'apparition de nouvelles conduites chez l'enfant, notamment en ce qui concerne le jeu symbolique et le jeu de « faire-semblant » (Leslie, 1988).

Notons que quelle que soit la nature de ses mécanismes, les théories mettent l'accent sur les processus développementaux qui président à la constitution de la théorie de l'esprit : celle-ci serait acquise à travers l'expérience des interactions sociales (Garfield, Peterson & Perry, 2001). Les connaissances sociales sont alors mises à l'épreuve et modifiées au fur et à mesure des expériences intersubjectives de l'enfant (Perner, 1991 ; Gopnick & Wellman, 1992, Gopnik & Meltzoff, 1997). Malgré cela, la pleine maîtrise de la théorie de l'esprit ne serait

pas atteinte avant quatre ans, âge où l'enfant est capable de réussir le test classique des fausses croyances (Wimmer & Perner, 1983), d'ailleurs utilisé depuis comme une épreuve critique pour la détection des troubles de nature autistique (Baron-Cohen, Leslie & Frith, 1985).

1.2. La théorie de la simulation (ST)

Selon la théorie de la simulation (ST), concéder un état mental à autrui nécessiterait une capacité à simuler son comportement pour lui attribuer intentions, désirs et croyances (Goldman, 1995 ; Gordon, 1986). Dans l'acception la plus générale de ST, pour comprendre et prédire les états mentaux d'autrui, le sujet est conduit à produire une analogie, plutôt qu'une inférence, et ce, par rapport à ses propres états mentaux.

Pour Goldman, l'individu attribue des états mentaux à autrui via une projection de ses propres états mentaux dans autrui, comme s'il prenait littéralement sa place (Goldman, 1995). La simulation consiste alors ici à générer des états mentaux dans une situation hypothétique, puis à les interpréter de façon à prédire les états mentaux d'autrui. Mais cette nécessité d'interprétation conduit Goldman à assumer que TT constitue une alternative intéressante de ST (ibid.). En effet, le processus de simulation ne semble pas toujours adéquat, dans la mesure où l'individu peut manquer d'informations concernant autrui, notamment en ce qui concerne ses connaissances et ses compétences spécifiques. Ainsi, ces lacunes peuvent être complétées par des mécanismes d'induction, faisant appel aux connaissances sur les états mentaux stockées en mémoire. Par conséquent, même si le processus de simulation reste l'heuristique privilégiée pour la compréhension d'autrui, les connaissances que le sujet acquiert sur les états mentaux peuvent donner lieu à des généralisations, permettant de constituer des structures de connaissances (ibid.).

Une version plus radicale de ST est proposée par Gordon : nous avons une capacité relativement efficace à prédire notre comportement dans des situations hypothétiques et avant d'agir, nous avons fréquemment recours à une simulation anticipatoire de la situation et de notre comportement probable dans celle-ci (Gordon, 1986). Cette compétence nous serait également utile pour prédire le comportement et les états mentaux d'autrui, cette fois en nous conduisant à simuler les actions des autres. De cette façon, nous n'avons pas besoin de stocker une théorie des états mentaux, puisque nous avons l'expérience de nos propres états mentaux et que ceux-ci peuvent nous servir de guide pour la compréhension d'autrui. De cette façon, Gordon défend plus radicalement le processus de la simulation en tant qu'analogie, en considérant que la position de Goldman ne consiste ni plus ni moins qu'à décrire un mécanisme d'inférence implicite reposant sur l'introspection (Gordon, 1995). En effet, pour

Gordon, l'introspection et l'inférence ne sont pas nécessaires, dans la mesure où le sujet est capable via la simulation d'opérer une véritable « transformation », plus qu'un simple « transfert » (ibid.). Cette transformation est décrite comme un déplacement égocentrique : lorsque je perçois autrui, je ne me mets pas à sa place pour prédire son comportement sur la base de mes propres états mentaux, mais je *deviens* autrui en simulant *l'entièreté* de la situation, mes états mentaux et mes processus cognitifs étant postulés comme nécessairement analogiques à ceux d'autrui. Pour Gordon, cette forme de simulation permet de transformer une simple attribution d'états mentaux en un questionnement général sur le monde (ibid.).

Aujourd'hui, ST trouve principalement des appuis empiriques et expérimentaux dans le domaine des neurosciences, via l'étude des neurones-miroir. En effet, puisque les mêmes neurones sont activés à la fois quand un observateur perçoit une action donnée et quand il réalise lui-même cette action, ces neurones seraient le substrat neurophysiologique de la théorie de l'esprit (Gallese & Goldman, 1998). Toutefois, cette approche cérébro-centrée conduit à de nombreux scepticismes : par exemple, il est argumenté que l'activité des neurones-miroir ne pourrait constituer qu'un épiphénomène n'apportant pas d'explications particulières sur la nature de la cognition sociale (Gallagher, 2007).

1.3. Une détection de contingences

Qu'elle soit relative à un mécanisme d'inférence ou d'analogie, la cognition sociale n'en reste pas moins basée sur la perception et le traitement des indices sociaux issus de l'interaction. En effet, dans les situations d'interaction sociale, l'individu perçoit les manifestations comportementales d'autrui (les mouvements, les expressions corporelles et faciales, comme les sourires), et les interprète à la lumière d'une théorie de l'esprit préalablement constituée. Dans notre situation, cela reviendrait à dire que les utilisateurs de notre dispositif interprètent les manifestations comportementales d'autrui (soit les stimulations tactiles résultant de la rencontre des corps-image/corps-percevant) grâce à des connaissances stockées en mémoire (TT) ou grâce à leur propre expérience des états mentaux (ST), et ce, de façon à produire une réponse appropriée. La perception des indices sociaux serait cruciale, parce qu'elle rendrait possible la synchronisation, la symétrie et le caractère *contingent* des comportements (Brazelton & Cramer, 1991).

Une contingence se définit par la présence d'une relation temporelle entre deux évènements (Moran, Dumas & Symons, 1992)[25]. Ainsi, les contingences sociales impliquent

[25] Notons que cette relation peut être causale ou non : la présence d'une contingence est une condition nécessaire mais pas suffisante pour parler de causalité (Desrochers, 2003 ; Moran et al., 1992).

une relation temporelle entre les actions ou comportements d'individus en situation de co-présence (Bigelow & Birch, 1999). D'une façon générale, il est reconnu que la présence de contingences sociales dans les interactions précoces favorise le développement social, émotionnel et cognitif (Bigelow, 1998 ; Dunham & Dunham, 1990). Le développement d'une sensibilité perceptive au rythme mutuel et aux contingences sociales semble donc être une habileté cruciale pour permettre un engagement dans les interactions sociales coordonnées (Crown, Feldstein, Jasnow, Beebe & Jaffe, 2002 ; Stern, 1977). Une micro-analyse d'interactions dyadiques mère-bébé montre que bien souvent, l'adulte est l'initiateur et le principal organisateur des échanges (Stern, 1977). Il présente donc un modèle organisé à l'enfant, souvent dans le cadre d'un jeu « caractérisé par des contours exagérés, des changements de tempo marqués et des répétitions systématiques » (Rochat, Querido & Striano, 1999, p. 951, *traduction personnelle*). Ces échanges présentent l'occasion pour l'enfant d'apprendre à détecter des ensembles de patterns expressifs, et à les exploiter de façon à produire une « synchronie interactionnelle » (Beebe & Gerstman, 1984). Celle-ci est caractérisée par une coordination temporelle très précise et rapide, de l'ordre du dixième de seconde (Trevarthen, 1977). Chez l'adulte, de nombreuses études suggèrent en effet que les agents coordonnent leurs mouvements et leurs énoncés, ce qui suppose de la part de chaque interactant une sensibilité très précise au rythme et au timing de l'interaction (voir notamment Condon, 1986 ; Davis, 1982 ; De Jaegher, 2006 ; Issartel, Marin, et Cadopi, 2007).

D'un point de vue classique, les contingences sociales seraient traitées par un module inné spécifique (Gergely & Watson, 1996 ; 1999). Dès la naissance, le module de détection des contingences permettrait au bébé d'analyser les relations existantes entre un comportement donné et une réponse de l'environnement (ibid.). Grâce à l'activation de ce module, le bébé peut, dès la première année, savoir quand ses propres actions sont responsables du mouvement d'un mobile suspendu (Rovee & Rovee, 1969), ou quand la succion d'une tétine fait apparaître une scène visuelle intéressante (Siqueland & De Lucia, 1969). Cette capacité précoce amènerait progressivement l'enfant à détecter des degrés variés de contingence (Gergely, 2001). En effet, jusqu'à 3 mois, le module de détection des contingences serait génétiquement programmé de façon à orienter le bébé vers les stimulations parfaitement contingentes à ses comportements (ibid.). Par exemple, il a été montré que les bébés de 3 mois regardent plus longtemps leur propre reflet dans un miroir que l'image d'un autre bébé dont les comportements ne sont pas contingents (Field, 1979). Mais au-delà de 3 mois, la valeur cible du module change au profit d'une préférence pour les contingences élevées mais imparfaites (Gergely, 2001). Par exemple, les bébés de 5 mois

regardent significativement plus longtemps un écran leur présentant l'image pré-enregistrée de leurs jambes (situation peu contingente) qu'un écran leur présentant le mouvement de leurs jambes en direct (situation parfaitement contingente) (Bahrick & Watson, 1985). Les mêmes résultats sont obtenus lorsque le stimulus manipulé est le visage (Papousek & Papousek, 1974), ou concerne l'exploration manuelle d'un objet (Schmukler, 1996). La préférence précoce pour les contingences imparfaites se révèle être adaptative, car elle permet à l'enfant de s'ouvrir sur le monde social, où les réponses d'autrui ne sont pas parfaitement contingentes au comportement de l'enfant (Rochat & Striano, 1999). En outre, ces contingences imparfaites pourraient servir de support à l'expérience précoce de la différenciation soi/autrui (Neisser, 1991), et à l'établissement des liens d'attachement en rendant possible l'apparition de comportements d'affiliation (Bigelow, 2001).

Par ailleurs, l'expérience précoce des interactions contingentes avec un partenaire donné semble avoir une influence sur la préférence attentionnelle des bébés de 4 et 5 mois lors d'interactions ultérieures (Bigelow & Birch, 1999). Dans cette étude, lors d'une première phase, les bébés sont confrontés simultanément à deux partenaires via deux écrans côte à côte. Conformément à des études antérieures, les bébés regardent significativement plus longtemps le partenaire agissant de façon contingente que le partenaire agissant de façon non-contingente (Bigelow, MacLean & MacDonald, 1996 ; Hains & Muir, 1996). Dans une deuxième phase, après un délai de 6 jours, les bébés sont confrontés simultanément à deux partenaires agissant de façon contingente à travers le même dispositif. Parmi eux, l'un est le partenaire avec lequel le bébé a interagi en phase 1, l'autre est un étranger. Les résultats montrent que les bébés ont une préférence attentionnelle pour les partenaires avec qui ils ont préalablement eu une histoire interactive, même si celle-ci a été brève (Bigelow & Birch, 1999). De façon plus cruciale, il est important de noter que dans cette expérience, les phases d'interaction sont précédées d'une minute de « still face » (soit une minute pendant laquelle le visage des adultes reste impassible, voir Tronick, Als & Adamson, 1979). Lors de celle-ci, les bébés ne manifestent aucune préférence, et ce, que ce soit dans la phase 1 ou dans la phase 2. Dans la phase 1, ce résultat dénote que la préférence pour le partenaire contingent est fonction de l'interaction elle-même et non de caractéristiques physiques particulières. Dans la phase 2, cela montre que cette préférence ne se manifeste que dans la dynamique actuelle de l'interaction (Bigelow & Birch, 1999).

D'une façon générale, nous reconnaissons que la compréhension des manifestations comportementales d'autrui peut être basée sur la perception de telles contingences. En revanche, nous ne partageons pas l'idée selon laquelle ces contingences ne seraient qu'une

donnée de l'environnement social traitée par des processus ou des modules internes. Si tel était le cas, alors la perception elle-même ne serait que le résultat d'une réception passive des informations issues de l'environnement, qui seraient soumises à un traitement interne pour accéder à la connaissance. Or, du point de vue de la psychologie écologique, toute connaissance est extraite de la perception (Neisser, 1991). En effet, dès le début de la vie, les systèmes perceptifs ont pour fonction d'extraire l'information qui spécifie l'environnement et sa relation avec l'observateur (Gibson, 1966). Crucialement, cette information est une invariance détectée à travers l'activité perceptive de l'observateur. Ainsi, le mouvement, ou plutôt les modifications du flux sensoriel qu'il engendre, permet au sujet d'extraire les invariants perceptifs qui spécifient les propriétés de l'environnement (ibid.). En tant que phénomènes observables, les interactions sociales impliquent effectivement qu'elles fournissent des informations à l'observateur. Mais contrairement aux conceptions classiques, ces informations sociales sont perçues *directement* à travers l'activité perceptive, dans l'immédiateté de l'interaction et non après-coup via une opération sur des états internes (Neisser, 1988). Ainsi, parallèlement à l'émergence d'un soi écologique, relatif à la situation immédiate de l'individu dans l'environnement, un « soi interpersonnel » émerge de la perception des situations sociales dans lesquelles l'individu est engagé (Neisser, 1991).

1.4. Vers une alternative : la théorie de l'interaction

Dans ce contexte, Gallagher propose la « théorie de l'interaction », en formulant un argument phénoménologique à l'encontre des théories classiques concernant la cognition sociale (Gallagher, 2004). De façon générale, il approuve les deux premiers mécanismes proposés par Baron-Cohen (1995) : dès la naissance, l'enfant est en effet capable de percevoir les intentions d'autrui, ce qui se manifeste lors du stade de l'intersubjectivité primaire (Trevarthen, 1979). De façon très précoce, l'enfant semble spontanément orienté vers la communication, le partage émotionnel et intentionnel. Dans les dyades mère-bébé, le bébé est naturellement sensible aux comportements expressifs de sa mère et à leur rythme, et parvient à se synchroniser avec elle pour échanger des proto-conversations et pour s'exprimer en alternance (Bateson, 1979). L'enfant aurait donc une appréciation consciente, active et immédiate des intentions communicatives de l'adulte (Trevarthen, 1979). Selon la théorie de l'interaction, cette capacité est suffisante pour expliquer la cognition sociale : nous percevons *directement* les intentions à travers les mouvements corporels, les mimiques faciales et la direction du regard d'autrui. De cette façon, « nous ne voyons pas des distorsions faciales et faisons l'inférence que [l'individu] ressent de la joie, de la tristesse, de l'ennui. Nous

décrivons un visage immédiatement comme triste, radieux, ennuyé, même si nous sommes incapables de donner une autre description des caractéristiques » (Wittgenstein, 1980, §580, *traduction personnelle*). La base des interactions humaines et de la compréhension d'autrui peut déjà être trouvée dans l'enfance précoce à travers les pratiques incarnées (émotionnelles, sensori-motrices, perceptives et non-conceptuelles) (Gallagher, 2004). Celles-ci constituent notre accès premier à la compréhension d'autrui, et ce mode d'accès reste privilégié même après avoir élaboré une théorie de l'esprit de type troisième personne (ibid.).

De cette façon, Gallagher ne nie pas qu'il puisse exister un mécanisme d'inférence à propos des états mentaux d'autrui, mais celui-ci est une construction conceptuelle : il ne se manifeste que dans de rares occurrences cartésiennes, notamment lors du test classique des fausses croyances (Wimmer et Perner, 1983). Ce test ne semble donc pas adapté pour rendre compte de la cognition sociale (Gallagher, 2004), car contrairement aux situations quotidiennes, il ne permet pas à l'enfant de s'impliquer dans la situation. L'engagement du sujet y est secondaire, car celui-ci est placé dans une situation de spectateur à la troisième personne (Hutto, 2004). Cette mise à distance par rapport à la situation exige une attitude théorique qui n'a pas lieu d'être dans l'expérience des situations sociales quotidiennes. Au contraire, dès les interactions précoces, l'enfant adopte plutôt une attitude pragmatique à la seconde personne[26] (Reddy & Morris, 2004). Dans la relation interpersonnelle, en effet, l'autre n'est pas perçu comme un objet, mais comme une personne douée d'intentions, d'émotions et de pensées (Macmurray, 1991). L'enfant a donc l'expérience d'autrui comme un être intentionnel en relation avec sa propre intentionnalité : l'interaction peut être mieux comprise comme un processus qui engage les deux participants (Reddy, 1996). La cognition sociale est donc essentiellement composée de capacités sensori-motrices incarnées, qui permettent la détection et la compréhension des intentions d'autrui dans les situations quotidiennes, contextualisées et pragmatiques (Gallagher, 2004).

Cette conception écologique de la cognition sociale semble tout à fait appropriée pour rendre compte du stress éprouvé par le bébé lorsque sa mère, pourtant pleine d'attention jusque-là, cesse soudainement toute activité et fige son visage dans une attitude neutre (le « still face » de Tronick et al., 1979). En effet, dans cette situation, la détresse manifestée par le bébé peut être attribuée au fait que le bébé ne parvient plus à détecter les intentions

[26] L'approche à la seconde personne permet une voie moyenne entre une phénoménologie des situations sociales à la première personne et une approche purement conceptuelle à la troisième personne. De cette façon, l'approche à la seconde personne permet de prendre en compte l'engagement actif des individus dans des relations sociales pragmatiques et contextualisées (Reddy & Morris, 2004).

communicatives de la mère, précisément parce qu'elle n'est plus expressive et ne manifeste plus ses intentions à travers ses comportements.

Au contraire, cette position ne semble pas pouvoir expliquer pourquoi le bébé, alors qu'il interagit avec sa mère via un dispositif vidéo synchronisé, éprouve de la détresse lorsqu'il se trouve soudainement confronté à l'enregistrement d'un comportement antérieur de sa mère (Murray & Trevarthen, 1985), et ce même si celui-ci est tiré de la même séquence interactive (Nadel, Carchon, Kervella, Marcelli & Réserbat-Plantey, 1999). Par conséquent, nous pouvons en conclure que le bébé est sensible non seulement à l'expressivité et à la qualité émotionnelle du comportement de la mère, mais aussi et de façon plus fondamentale au timing et aux contingences des comportements expressifs (ibid.).

La perception des contingences semble donc être une habileté cruciale pour la compréhension sociale, et celle-ci est mise en évidence dès les interactions précoces. Dans le contexte d'une conception écologique de l'interaction, les contingences seraient perçues directement, sans passer par la médiation d'un module cognitif assurant un traitement interne (Gergely & Watson, 1996 ; 1999). Mais en admettant cette idée, il reste à savoir par quel moyen ces contingences sont perçues. D'une façon générale, le développement de la compréhension sociale est massivement étudié chez les enfants tout venants et voyants (voir par exemple Brazelton, Tronick, Adamson, Als & Wise, 1975 ; Bullowa, 1979 ; Hobson, 2002). Cette constatation pose la question de la perception de ces contingences chez l'aveugle. Si la vision est absente ou perturbée, par quels moyens le sujet peut-il détecter les relations entre le comportement d'autrui et son propre comportement ? Cette question nous apparaît essentielle, dans la mesure où nous souhaitons concevoir un dispositif destiné à des utilisateurs déficients visuels, et qui ne repose donc pas sur la modalité visuelle. Par ailleurs, comme nous allons le voir, le cas de l'interaction et de la compréhension des situations sociales chez l'aveugle nous apporte des éléments essentiels pour comprendre le lien entre interaction et cognition sociale.

2. La compréhension sociale en l'absence de vision

Le fait que la compréhension sociale soit étudiée à travers la vision n'est pas étonnant. Lorsque la modalité visuelle est disponible, elle constitue en effet la première modalité d'interaction entre l'enfant et l'adulte, à travers laquelle les deux partenaires ont un niveau de contrôle similaire (D'odorico & Levorato, 1990). Par ailleurs, les enfants étudiés sont généralement issus d'une culture occidentale où le regard occupe une place particulièrement

importante, ce qui n'est pas forcément le cas dans d'autres cultures (Akhtar & Gernsbacher, 2008). Dans un contexte occidental, le regard mutuel possède en effet un statut particulier (Argyle & Cook, 1976), et sa manifestation entre l'adulte et l'enfant est considérée comme le premier signe d'un engagement interpersonnel (Csibra & Gergely, 2006). Dès la naissance, le regard mutuel semble être une situation privilégiée par l'enfant voyant. Par exemple, il a été montré que les nouveau-nés regardent plus longtemps des photographies représentant des visages qui les regardent que les mêmes photographies avec les yeux orientés ailleurs (Farroni, Csibra, Simion & Johnson, 2002). Ce résultat est classiquement interprété comme une préférence innée pour le regard mutuel, mais celle-ci est mise en évidence dans des contextes statiques où l'interaction est absente (Akhtar & Gernsbacher, 2008). Dans les situations d'interaction quotidiennes, la préférence des enfants pour le regard mutuel est plus complexe. Par exemple, si le regard mutuel semble calmer les enfants stressés (Blass, Lumeng & Patil, 2007), les enfants manifestant un état d'excitation trop important évitent activement le regard de l'adulte (Field, 1981).

Quoi qu'il en soit, la cognition sociale semble reposer sur la perception visuelle des manifestations comportementales d'autrui (les mouvements, les expressions corporelles et faciales, ou encore les sourires). La vision semble alors cruciale pour l'élaboration d'un espace intersubjectif (Trevarthen, 1979), et constitue une source permanente d'apprentissage social, notamment via l'imitation (Meltzoff, 1999). En effet, dans les situations d'interaction sociale, l'enfant voyant perçoit les manifestations comportementales de l'adulte, ce qui lui permet d'accéder à ses intentions communicatives (Trevarthen, 1979). L'enfant s'engage alors dans des contacts visuels mutuels, et répond aux sollicitations sociales de l'adulte. De la même façon, la capacité de réponse de l'adulte repose sur sa sensibilité visuelle aux indices comportementaux manifestés par l'enfant. La disponibilité de la vision facilite donc l'implication des deux partenaires, en offrant à chacun une source d'informations perceptives réciproque. Mais qu'en est-il de l'enfant déficient visuel ? Quelles sont les caractéristiques de l'interaction précoce lorsque la vision est absente ou perturbée ?

2.1. Les conséquences de la cécité sur l'interaction sociale

Sans suffisamment de vision pour percevoir les indices comportementaux de leurs partenaires sociaux, et sans habiletés adaptées pour compenser ce manque d'informations, les interactions sociales semblent compromises. Par exemple, il a été montré que la cécité a un impact négatif sur les interactions sociales précoces et sur les compétences communicatives de l'enfant (Fraiberg, 1977). Dans cette optique, un parallèle entre cécité congénitale et

trouble autistique a même été mis en avant (voir par exemple Chess, 1971 ; Gense & Gense, 1994). Toutefois, si les connaissances du soi interpersonnel sont effectivement perturbées chez l'aveugle complet (Neisser, 1991 ; Bigelow, 2003), il apparaît que celui-ci ne manifeste pas nécessairement les déficits socio-émotionnels caractéristiques du spectre autistique (Brown, Hobson, Lee & Stevenson, 1997). Si l'établissement de l'interaction et son maintien dans le temps sont perturbés chez l'enfant déficient visuel (Sacks & Silberman, 2000), c'est probablement en raison de difficultés perceptives à détecter des patterns signifiants dans les interactions sociales, alors qu'ils sont habituellement facilement détectés via la vision (Hobson, 1993). Si l'absence ou la perturbation sévère de la vision pose certaines difficultés, celles-ci peuvent pourtant être dépassées par l'intermédiaire de modes d'interaction alternatifs (Pérez-Pereira & Conti-Ramsden, 1999).

Pour comprendre comment peuvent s'organiser ces modes d'interaction alternatifs, il est important de situer les conséquences de la cécité, particulièrement lorsqu'elle est congénitale. Tout d'abord, il faut souligner que le comportement de l'enfant non-voyant est caractérisé par une tendance à l'inactivité motrice, particulièrement en ce qui concerne le visage. De ce fait, les expressions faciales sont moins prononcées et moins nombreuses que chez l'enfant voyant (Fraiberg, 1977). D'un point de vue corporel, les comportements sont caractérisés par des maniérismes (maladresses gestuelles) et des blindismes (comportements moteurs stéréotypés) (Hatwell, 2003). Ces traits comportementaux[27] perturbent sensiblement les interactions de l'enfant déficient visuel avec son environnement social, car ils provoquent bien souvent un sentiment d'étrangeté aux yeux des autres (Fraiberg, 1977). Par exemple, l'absence de contact visuel dans les interactions avec un enfant non-voyant conduit à une diminution de l'attention visuelle de la mère envers son enfant, et donc à une diminution des réponses vocales qu'elle pourrait apporter à ses comportements (Rogers & Puchalski, 1984 ; Rowland, 1984). Si certaines mères regardent beaucoup le visage de leur enfant, et sont particulièrement attentives à ses manifestations expressives (Urwin, 1983), elles éprouvent toutefois de grandes difficultés à reconnaître les intentions communicatives de l'enfant (Kekelis & Andersen, 1984). Pourtant, au moyen de micro-analyses des interactions mère/enfant déficient visuel, il est observé que différents signaux moteurs expressifs, comme la posture corporelle ou le mouvement des mains, sont effectivement mobilisés par l'enfant, mais que bien souvent, les mères ne les remarquent pas ou n'y prêtent pas attention (Rowland, 1984). De même, certains

[27] Notons que nous n'avons pas mentionné ici les conséquences spécifiques de la cécité sur les connaissances spatiales, la locomotion et l'atteinte des objets. Nous développerons ce point particulièrement lorsque nous aborderons l'introduction d'objets dans la dyade (voir Partie 3, Section 5, p. 160).

comportements attentionnels de l'enfant, comme le fait de stopper son activité en cas de sollicitation sociale, sont interprétés à tort comme la manifestation d'un manque d'intérêt ou d'un désengagement. En conséquence, les mères ont tendance à réduire leurs initiatives, ce qui réduit par conséquent les opportunités d'interaction sociale de l'enfant (Pérez-Pereira & Conti-Ramsden, 1999 ; Preisler, 1991).

2.2. Des modes d'interaction alternatifs

Ces difficultés interactives peuvent être interprétées par l'absence de réciprocité dans les canaux de communication habituellement utilisés (Baird, Mayfield & Baker., 1997), ce qui nuit au développement d'un espace intersubjectif propice aux interactions réciproques (Pérez-Pereira & Conti-Ramsden, 1999). En effet, dans le cas où la mère est voyante, elle privilégie spontanément des informations visuelles, ce qui n'est pas compatible avec les modalités perceptives disponibles chez l'enfant. Malgré cela, il est noté que la capacité de la mère à répondre aux comportements de l'enfant est adaptative. Dans une étude longitudinale de deux enfants aveugles de naissance, il est observé que lorsque la mère découvre certains indices pertinents dans le comportement de l'enfant, ou certains comportements qui provoquent des réponses de la part de l'enfant, elle les utilise de façon répétée (Urwin, 1978). Ainsi, l'établissement d'une communication efficiente implique que les participants puissent maintenir l'échange en fournissant des réponses séquentielles appropriées, ce qui conduit à une coordination interpersonnelle effective (Rattray & Zeedyk, 2005). Parallèlement, il est à noter qu'en l'absence de vision, les enfants investissent les autres modalités perceptives, comme le toucher ou l'audition.

A l'instar de la vision, l'audition est une modalité perceptive qui fournit constamment des informations. Les enfants déficients visuels utilisent donc beaucoup les informations auditives, comme les changements de sons, mais aussi des flux d'information dérivés comme l'écholocation (Millar, 1988). Par exemple, le fait que l'enfant fige son activité en cas de sollicitation sociale pourrait refléter une concentration particulière sur l'environnement sonore, de façon à détecter des indices sur la situation ou le comportement du partenaire social. Par ailleurs, les enfants déficients visuels répondent de façon précoce aux sollicitations verbales de la mère, notamment en effectuant des mouvements articulatoires avec la bouche, accompagnés ou non de vocalisations (Preisler, 1991). Quand celles-ci surviennent, elles sont parfois apparentées à des tentatives d'imitation verbale [28]. En outre, dans son étude

[28] Ces comportements sont habituellement observés chez les enfants voyants, et ont été attribués à un processus imitatif principalement basé sur des indices visuels, qui déclenche un processus transmodal

longitudinale, Preisler remarque qu'un des enfants produit beaucoup de mouvements des mains, et ce, de façon coordonnée avec les énoncés verbaux de la mère. Selon Trevarthen (1986), les mouvements des bras et des mains participent à l'expression de l'enfant et à la régulation de son engagement émotionnel. Nous pouvons poser l'hypothèse que chez l'enfant aveugle, il s'agit d'une réponse aux sollicitations sociales de la mère. Ces réponses témoignent d'une motivation innée pour la communication interpersonnelle, qui passe par la coordination des comportements, et ce, par tous les moyens disponibles (ibid.). La participation active de l'enfant, grâce à l'étayage de la mère, contribue alors à l'établissement d'une relation positive et confiante (« core-relatedness », Stern, 1985).

En règle générale, la fréquence de vocalisations est plus faible chez les enfants aveugles que chez les enfants voyants (Rogers & Puchalski, 1984). En effet, si la cécité ne semble pas avoir d'incidence sur l'apparition et l'évolution des premières vocalisations (« babillages » comparables à ceux des enfants voyants), celles-ci sont utilisées essentiellement en réponse à une sollicitation de l'adulte et non pour provoquer une interaction (Fraiberg, 1977 ; Rowland, 1984). Parallèlement, les mères d'enfants déficients visuels sont plus directives que les mères d'enfants voyants, et sont plus susceptibles d'introduire une conversation que de laisser cette opportunité à l'enfant. Cette tendance se reflète par ailleurs dans le langage utilisé par les mères, qui mobilise plus de tournures impératives que de tournures déclaratives, ce qui a une influence négative sur le développement du langage chez l'enfant déficient visuel (Kekelis & Andersen, 1984). Pourtant, il est noté que lorsque la mère utilise des intonations vocales pour traduire ses états émotionnels ou ses expressions faciales, la communication avec l'enfant est plus efficiente (Urwin, 1983). De la même façon, il est noté que l'imitation verbale, les réponses aux vocalisations de l'enfant et le fait de solliciter explicitement des réponses sont des facteurs qui favorisent l'engagement social du jeune aveugle (ibid.).

Par ailleurs, il est observé que les mères répondent aux vocalisations de l'enfant plus par le toucher que par le discours (Rowland, 1984). En effet, le toucher ou les mouvements corporels sont particulièrement utilisés par les mères d'enfants déficients visuels, et ce notamment pour attirer l'attention de l'enfant et établir un cadre de communication (Pérez-Pereira & Conti-Ramsden, 1999 ; Rattray & Zeedyk, 2005). Dans cette perspective, il est important de noter que le contact physique pourrait jouer le même rôle sur l'engagement social que le contact visuel (Akhtar & Gernsbacher, 2008), ces deux formes de contact étant par ailleurs souvent associées chez l'enfant voyant dans les interactions sociales précoces

(Meltzoff, 1999). L'observation de ces comportements chez l'enfant déficient visuel montre que les indices perçus ne sont pas nécessairement visuels.

(Stack, 2001). Par exemple, il a été montré que les effets négatifs du still-face sont atténués si la mère continue de toucher l'enfant lorsqu'elle fige son visage (Stack & Muir, 1992). Plus encore, ce n'est pas seulement la présence ou l'absence de toucher qui est déterminante, mais la qualité de celui-ci (Stack & Arnold, 1998). Chez l'enfant voyant, le toucher semble alors signaler à l'enfant que la mère est toujours engagée avec lui, même si elle n'est pas répondante de façon visuelle. En cas de déficience visuelle, le toucher est particulièrement utilisé, comme en témoigne les patterns d'interaction tactile spécifiques qui émerge lors des interactions précoces (Rattray & Zeedyk, 2005). Outre les vocalisations, le toucher semble alors être une alternative privilégiée pour garantir une interaction signifiante dans les dyades. Les auteurs distinguent deux formes de toucher, le toucher passif et le toucher actif. Le toucher passif implique un contact tactile entre les deux partenaires ou entre un des partenaires et un objet, mais sans manipulation. Le toucher actif implique un contact tactile accompagné d'une manipulation active (ibid.). Le toucher actif semble alors favoriser l'interaction et le partage des expériences, car il permet à l'enfant déficient visuel de sentir les effets de son comportement sur les comportements d'autrui, tout comme les contingences du comportement d'autrui par rapport à son propre comportement (Bigelow, 2003).

2.3. L'interaction comme support de la compréhension sociale

Comme nous l'avons vu, des conduites adaptées semblent pouvoir amener la dyade à dépasser certaines contraintes liées à la cécité. Plus encore, les observations de dyades mère / enfant déficient visuel montrent que les interactions précoces semblent se développer typiquement avant le premier anniversaire (Preisler, 1991). En utilisant des flux d'information alternatifs à la vision, les deux partenaires parviennent à synchroniser leurs comportements, et à partager des expériences sociales de façon dyadique. Si les mères d'enfants déficients visuels ont appris à détecter la variété des expressions communicatives de leur enfant et si elles lui permettent d'adopter un rôle actif dans le maintien des proto-conversations dont elle est l'initiatrice (ibid.), un contexte d'accordage affectif est rendu possible (Stern, 1985). Comme le souligne Hobson, la participation active de l'enfant dans des interactions affectives est le facteur le plus important pour le développement social (Hobson, 1993). Ainsi, la déficience visuelle de l'enfant peut être compensée par le support de l'adulte, qui prend en compte les limitations perceptives de l'enfant et l'amène à donner du sens aux autres canaux de communication, par l'expérience de situations sociales adaptées. Mais si ce soutien de la part de l'adulte est nécessaire, celui-ci doit être ajusté continuellement aux progrès de l'enfant, en lui laissant un rôle actif quand cela est possible. Sans ces ajustements, l'adulte

risque de surprotéger l'enfant en contrôlant la totalité des interactions, ce qui peut entraîner à long terme une sur-dépendance et donc un isolement social (Zebehazy & Smith, 2011). Dans cette optique, il faut comprendre la déficience visuelle d'un point de vue social, ce qui permet de déplacer ses conséquences de l'individu au système social dans lequel il s'insère et où ses actions prennent du sens. Comme l'ont fait remarquer à juste titre Rogers et Puchalski (1984), lorsqu'un enfant est déficient visuel, c'est la dyade toute entière qui est déficiente[29]. C'est avant tout le processus de l'interaction qui assure des échanges signifiants, et non seulement les caractéristiques individuelles des partenaires impliqués.

Ainsi, même en l'absence de vision, l'interaction sociale précoce semble reposer sur l'alternance cyclique des comportements et sur la qualité émotionnelle des échanges (Trevarthen, 2001). Par exemple, l'absence d'alternance de regard peut être compensée par l'alternance de contacts physiques ou de vocalisations, qui signalent à chaque participant l'engagement de l'autre. De cette façon, quelle que soit la modalité perceptive impliquée de façon privilégiée, la structure et la qualité des interactions sociales précoces, mais aussi l'engagement actif et mutuel des deux participants sont cruciales pour la co-construction d'un cadre de référence partagé.

3. L'interaction comme engagement mutuel

3.1. Une critique fondamentale des positions classiques

Comme nous l'avons vu, les théories classiques, fondamentalement individuelles, ne permettent pas de rendre compte de la complexité de la compréhension sociale ni de la capacité d'engagement mutuel qui semble être pourtant au cœur du processus de l'interaction. Cette conception individualiste conduit à négliger le processus même de l'interaction (De Jaegher, 2009 ; De Jaegher et al., 2010), le considérant comme un simple input à traiter sur la base de connaissances ou de savoir-faire déjà constitués. Or la prise en compte du processus de l'interaction permet de dépasser la focalisation sur les processus individuels internes, et de mettre l'accent sur l'engagement actif et situé des partenaires de l'interaction (Neisser, 1991 ; Gallagher, 2008). Le cas des interactions chez le sujet déficient visuel illustre avec pertinence la nécessité de prendre en compte l'ensemble de la dyade, qui semble pouvoir dépasser certaines contraintes individuelles par l'élaboration de modes de communication alternatifs et

[29] « Les deux partenaires de la dyade sont en carence, l'enfant par rapport à l'information visuelle, mais aussi la mère par rapport au manque d'information affective en retour. » (Rogers & Puchalski, 1984, p.55, *traduction personnelle*).

nécessairement partagés (Pérez-Pereira & Conti-Ramsden, 1999). Ceux-ci semblent reposer crucialement sur une coordination des acteurs de la dyade, quels que soient les moyens disponibles. Si la théorie de l'interaction propose que les manifestations comportementales d'autrui soient perçues directement, sans la médiation de représentations mentales (Gallagher, 2004), elle ne semble pas pouvoir dépasser l'idée d'une saisie individuelle de l'expressivité et des comportements d'autrui (De Jaegher, 2009). Au contraire, les théories de la cognition sociale à la deuxième personne ouvrent la voie à la réhabilitation de l'action et de l'engagement des participants dans les situations sociales (Marsh, Richardson, Baron & Schmidt, 2006 ; Reddy & Morris, 2004). C'est en effet à travers l'engagement dans le monde et l'engagement interpersonnel que les individus prennent conscience qu'autrui est un être intentionnel (Hobson, 2002 ; Trevarthen, 1993). L'information extraite de l'engagement mutuel précoce suffit au développement d'une conscience de soi et des autres, sur une base émotionnelle plutôt que représentationnelle (Reddy, 2003).

L'idée de prendre en compte un nécessaire engagement mutuel dans l'interaction nous paraît essentielle. Dans notre situation, les utilisateurs ne sont pas des « observateurs » passifs. Si tel était le cas, nous considèrerions la perception comme une réception passive des informations issues de l'environnement, qui justifierait par là même un « individualisme méthodologique » pour la compréhension de la cognition sociale (Froese & Di Paolo, 2011 ; Lenay & Stewart, 2012). L'objectif de l'individu pourrait alors n'être que de détecter dans le comportement d'autrui – ici, dans les séquences de stimulations tactiles résultant de la rencontre des corps-image/corps-percevant – des indices objectifs à mettre en relation avec des connaissances stockées en mémoire (TT) ou avec sa propre expérience des états mentaux (ST).

Au contraire, les utilisateurs déploient nécessairement une activité perceptive en fonction des retours sensoriels qui résultent de cette activité (Lenay et al., 2000). Par exemple, la reconnaissance d'une forme numérique via le dispositif Tactos consiste à découvrir et à exercer des lois de contingence sensori-motrices, ce qui souligne l'impossibilité de séparer la perception et l'action (O'Regan & Noé, 2001). Partant de ce principe, il nous semble essentiel de réhabiliter la participation active des sujets pour la perception des informations sociales. Mais plus crucialement, dans le cadre d'une interaction, les deux utilisateurs déploient une activité perceptive en fonction des stimulations tactiles qui résultent de la rencontre de l'activité perceptive de l'autre. Ainsi, les deux sujets sont nécessairement engagés dans l'interaction, ne serait-ce que pour percevoir l'autre.

3.2. L’ancrage de la compréhension sociale dans l’action

Si nous réhabilitons la participation active et réciproque des partenaires de l’interaction, alors la réponse classique de l’enfant dans la situation du still-face (Tronick et al., 1979) et du dispositif vidéo synchronisé (Murray & Trevarthen, 1985 ; Nadel et al., 1999), pourrait être expliquée par une perturbation de l’engagement mutuel qui nuit au processus même de l’interaction. En effet, si la réaction de l’enfant est expliquée par un module interne de détection des contingences, l’interaction ne peut être qu’un input voire un simple facteur contextuel, et l’essentiel de son activité peut être résumée en une opération cognitive (De Jaegher et al., 2010). Cette interprétation peut être questionnée à travers certaines expériences qui ne sont pas parvenues à répliquer les résultats initiaux alors même que les enfants testés ne présentaient pas de troubles particuliers dans le domaine social (voir par exemple Rochat, Neisser & Marian, 1998). Ici, il se peut que l’échec soit du au fait que les expérimentateurs n’ont pas permis aux dyades mère-enfant de suffisamment s’engager dans l’interaction avant d’introduire le replay (De Jaegher et al., 2010). Par ailleurs, dans une variante de cette expérience, il a été montré que la mère est autant perturbée que l’enfant par la distorsion de l’interaction (Murray & Trevarthen, 1986). Il semble donc nécessaire de prendre en compte un engagement mutuel des participants dans l’interaction, plutôt que de considérer uniquement des capacités individuelles de perception des contingences (McGann & De Jaegher, 2009). De cette façon, il ne suffit pas seulement que la mère soit engagée dans l’interaction pour que l’enfant puisse détecter les contingences de son comportement par rapport au sien, mais il faut aussi et crucialement que l’enfant soit engagé en tant que partenaire actif de l’interaction pour que de telles contingences puissent émerger et avoir du sens.

La compréhension sociale semble par conséquent ancrée dans les actions des individus qui interagissent (Knoblich & Sebanz, 2008). Plus précisément, le contexte social semble structurer et organiser la dynamique des actions individuelles. Dans une série d’expériences, il a été montré que les mouvements d’atteinte et de saisie d’un objet sont modifiés lorsque le contexte de l’action est social (Becchio, Sartori, Bulgheroni & Castiello, 2008a ; Becchio, Sartori & Castiello, 2010). Les patterns cinétiques des mouvements individuels sont en effet influencés par la finalité du geste (positionner un objet pour qu’un partenaire puisse le saisir), mais aussi par l’attitude de ce partenaire, qui peut être coopératif ou compétitif (Becchio, Sartori, Bulgheroni & Castiello, 2008b). Mais si la perception des indices sociaux joue un rôle

important dans les interactions, en influençant la dynamique des actions individuelles[30], il ne faut pas perdre de vue le fait que toute réaction à l'action d'autrui est susceptible d'influencer ses actions ultérieures. En effet, « l'action agissant sur des êtres qui son personnellement capables d'actions, la réaction, outre qu'elle est une réponse, est toujours une action nouvelle qui crée à son tour et affecte autrui » (Arendt, 1983, p. 214).

La réhabilitation de l'action dans la cognition sociale en fait une forme de négociation collective et de gestion interactive de l'espace commun (Krueger, 2011). L'espace commun est un espace pratique, ancré dans le corps et ses possibilités d'action. Les interactions reconfigurent cet espace en un espace social, où se jouent des coordinations intersubjectives et émotionnelles dynamiquement structurées par l'engagement mutuel des partenaires de l'interaction. De cette façon, les actions expressives (les regards, les expressions faciales et corporelles, les gestes, ou encore les vocalisations et les énoncés verbaux) sont cruciales pour la compréhension interpersonnelle, en ce qu'elles permettent la co-régulation de l'espace commun (ibid.). Cette co-régulation est effective quand les activités conjointes des sujets en interaction s'influencent mutuellement (Fogel, 1993). Les gestes expressifs sont alors conçus comme des affordances sociales, au sens où ils structurent activement l'espace de l'interaction : ils sont autant de contraintes et de possibilités d'action disponibles pour chaque sujet engagé dans l'interaction (Krueger, 2011). La saisie de ces possibilités d'action par l'interactant permet d'élaborer un système de couplage étroit dans lequel les deux partenaires jouent un rôle actif dans la régulation de l'échange. La cessation des activités expressives de la mère lors de l'épisode de « still face » rompt brutalement les activités de co-régulation. En effet, si l'un des deux partenaires cesse toute activité, ou s'il ne s'engage pas du tout dans la régulation de l'échange, alors l'interaction ne peut avoir lieu ni se maintenir.

3.3. L'interaction comme "*participatory sense-making*"

La pleine prise en compte du caractère incarné et situé des agents ouvre la voie à une perspective énactive de la compréhension des situations sociales. D'une façon générale, l'énaction permet de décrire la cognition comme une activité incarnée et située fondamentalement ancrée dans les processus vitaux, les dynamiques auto-organisationnelles et l'expérience vécue du corps animé (Varela et al., 1991 ; Thompson, 2007 ; Stewart, Gapenne & Di Paolo, 2010). Selon cette approche, tout agent est constamment engagé dans une activité de construction de sens (« *sense-making* ») qui co-détermine l'agent lui-même et

[30] Notons que cette dynamique peut être interprétée comme le résultat d'un alignement interpersonnel qui n'est pas nécessairement intentionnel (Marsh, Richardson & Schmidt, 2009 ; Issartel et al., 2007).

l'environnement. La cognition est alors caractérisée par la génération et la transformation de significations que le sujet produit (ou plutôt « *énacte* ») au travers de son couplage avec l'environnement (De Jaegher & Di Paolo, 2007). Dans ce contexte, l'interaction sociale peut être définie comme un couplage co-régulé entre au moins deux agents autonomes, où la co-régulation et le couplage affecte mutuellement chaque participant de façon à constituer une auto-organisation autonome en termes de dynamique relationnelle et où l'autonomie des agents impliqués n'est pas détruite (ibid.). La notion d'autonomie est ici cruciale : elle désigne l'auto-organisation d'un réseau de processus qui maintient son *domaine de viabilité* malgré des conditions précaires (Varela et al., 1991 ; Thompson, 2007 ; Di Paolo, 2009).

La compréhension interpersonnelle consiste alors en la coordination des activités individuelles de construction de sens, exprimées notamment dans les mouvements, les attitudes et dans les énoncés. Il s'agit du « ***participatory sense-making*** », qui caractérise la coordination des activités intentionnelles dans l'interaction. (ibid.). Selon De Jaegher, pour comprendre l'intersubjectivité, il ne faut pas se focaliser sur les capacités individuelles, mais sur le processus même de l'interaction, qui assure cette coordination. Il est donc crucial de replacer les interactions sociales au centre même de l'étude de la cognition sociale, de façon à impliquer l'acte même d'interagir dans la compréhension sociale (De Jaegher, 2009).

De cette façon, la cognition sociale émerge dans l'immédiateté de l'interaction et de son processus dynamique, en tant que construction collective. Tout comme la cognition émerge d'une participation active de l'agent dans la construction de sens (Varela et al, 1991 ; Thompson, 2007), la cognition sociale émerge d'une action spécifique : l'interaction sociale (De Jaegher et Di Paolo, 2007, Fuchs et De Jaegher, 2009). Dans cette optique, l'interaction sociale peut être décrite comme le résultat d'un échange entre les activités de sense-making d'un agent et les réponses de l'environnement, qui comprennent les activités de sense-making d'autrui. Le couplage des boucles perception-action des interactants se manifeste via une alternance de phases de synchronisation et de désynchronisation, de co-variation rythmique des gestes, des expressions faciales ou verbales etc. D'un point de vue développemental, la coordination est d'abord rythmique, musicale (Trevarthen, 2008), mais rapidement, elle devient le « format » par lequel l'enfant va réguler l'ensemble de ces échanges sociaux (Bruner, 1983). Les actions mutuelles coordonnées et co-régulées participent à une co-construction de sens et à une coordination des significations (De Jaegher & Di Paolo, 2007 ; Fuchs & De Jaegher, 2009). Loin d'être le résultat d'un processus inférentiel isolé, l'interaction sociale est le fruit de la rencontre d'agents incarnés et situés, dont l'engagement

mutuel fait émerger un processus dynamique d'ajustement continu des actions et des intentions (Fogel & Garvey, 2007 ; De Jaegher et Di Paolo, 2007).

4. Croisement perceptif et dynamique collective

De façon à illustrer cette idée, nous allons maintenant détailler une expérience qui concerne les croisements perceptifs dans un environnement minimaliste (Auvray, Lenay & Stewart, 2009). Cette expérience est cruciale relativement à notre travail, dans la mesure où elle concerne l'interaction via un dispositif sensiblement similaire à celui que nous avons utilisé (voir Partie 2, section 5.2, p. 141 pour plus de détails). Dans cette expérience, deux utilisateurs interagissent de façon tactile dans un environnement numérique simplifié à l'extrême (ibid.). Il s'agit d'un espace numérique unidimensionnel, où les sujets déplacent latéralement un curseur au moyen d'une souris. Ce curseur est « sensible » aux objets présents dans cet espace, dans la mesure où la rencontre de ce curseur avec les objets permet d'activer un retour tactile en tout ou rien délivré au sujet. Chaque sujet peut rencontrer trois types d'objets, chacun lui délivrant une stimulation tactile identique : un objet fixe, un objet mobile, et le curseur de l'autre participant. Il est important de noter que d'une part, l'objet fixe est un objet « privé », perceptible par un des participants, l'autre participant ayant son propre objet fixe à une position distincte ; et d'autre part, de façon plus intéressante, l'objet mobile est en réalité un leurre attaché par un lien virtuel rigide au curseur de chaque participant. Par conséquent, l'objet mobile se déplace exactement comme se déplace le curseur auquel il est attaché (voir Figure 48).

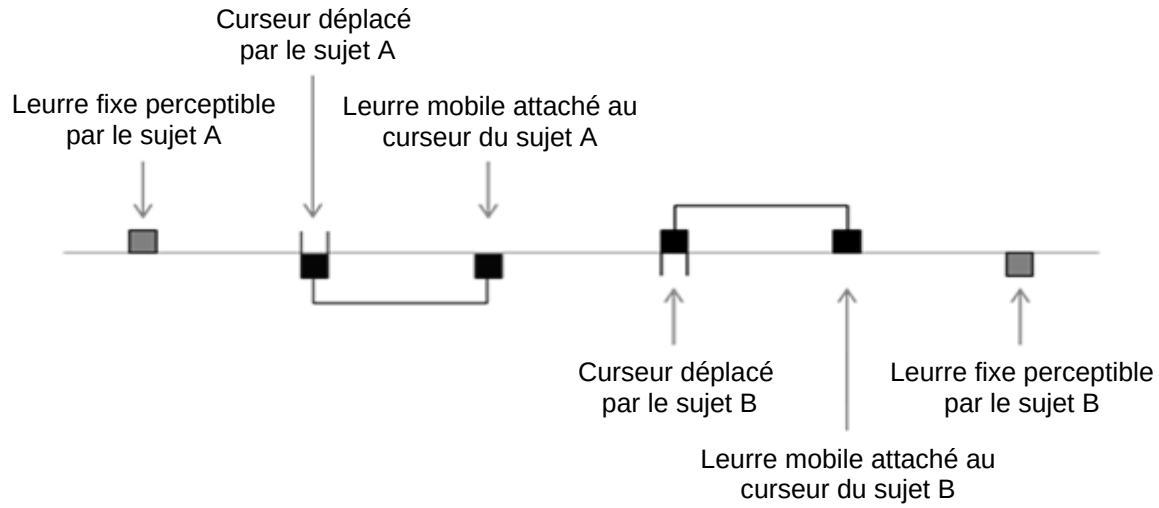


Figure 48 : Schéma du dispositif expérimental utilisé dans l'expérience de Auvray et al. (2009)

Les sujets sont invités à interagir via le dispositif pendant 3 sessions de 5 minutes, et doivent cliquer avec la souris lorsqu'ils estiment que la stimulation tactile résulte de la rencontre avec l'autre participant. A première vue, les résultats semblent indiquer que les sujets réussissent la tâche. En effet, la fréquence de clics est très élevée lorsque la distance entre les deux participants est faible (voir Figure 49), ce qui laisse à penser que chaque sujet parvient à reconnaître le mouvement du curseur de l'autre participant comme étant intentionnel. Mais une analyse plus fine de ces résultats dévoile qu'en réalité, le pourcentage de stimulations reçues par chaque sujet en fonction du type d' « objet rencontré » est plus important pour le curseur de l'autre participant (52,2%) que pour l'objet fixe (32,7%) et pour l'objet mobile (15,2%).

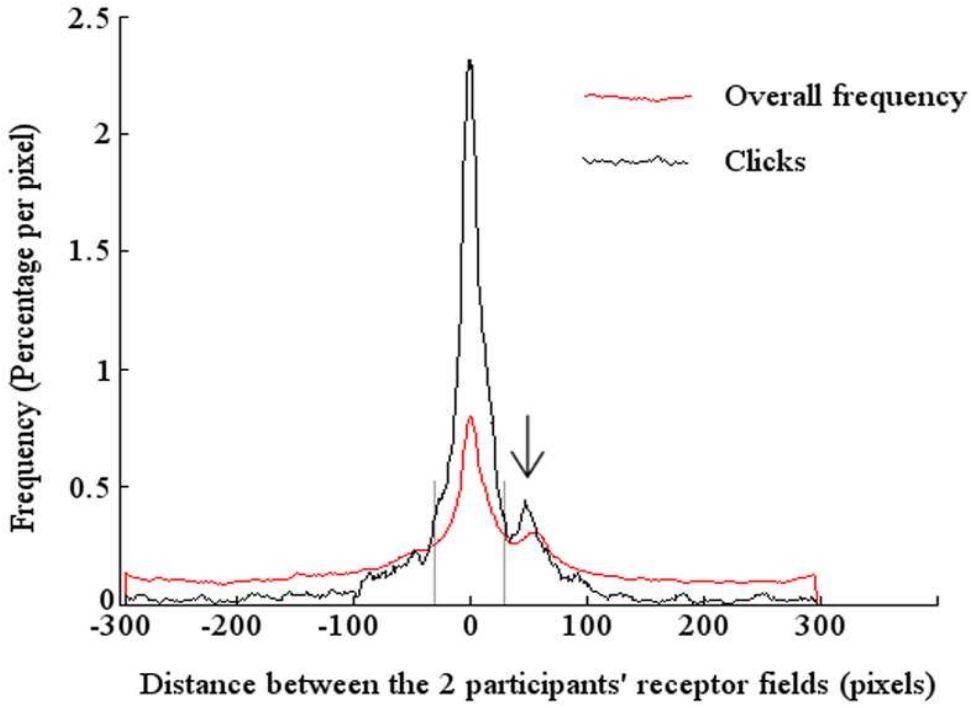


Figure 49 : Distribution des fréquences de clics et de stimulations en fonction de la distance entre les curseurs des deux participants. La courbe noire représente la distribution de la fréquence des clics : 62% des clics sont effectués alors que la distance entre les deux sujets est de ±30 pixels. La courbe rouge représente la distribution de la fréquence des stimulations reçues par les sujets : 28% des stimulations sont reçues alors que la distance entre les deux sujets est de ±30 pixels. Dans les deux cas, les distributions font apparaître un pic important lorsque la distance entre les deux sujets est nulle, ainsi qu'un pic mineur (indiqué par une flèche) lorsque la distance est de 50 pixels, soit la distance entre le curseur de chaque participant et le leurre attaché (tiré de Lenay & Stewart, 2012).

En conséquence, même si effectivement, les sujets cliquent plus fréquemment sur le curseur de l'autre participant (65,9% des clics) que sur l'objet mobile (23% des clics), le ratio clic/stimulation entre le curseur de l'autre participant et l'objet mobile ne présentent pas de différence significative (respectivement 1.26 contre 1.51).

Malgré l'apparente déception que peuvent suggérer ces résultats, il apparaît qu'ils masquent en réalité un phénomène très intéressant. En effet, une analyse des trajectoires perceptives de chaque participant démontre une corrélation négative globale (- .72) entre la vitesse du mouvement avant la stimulation et l'accélération du sujet suite à cette stimulation. Nous retrouvons donc ici un résultat classique que nous observons dans le cas de l'exploration de formes numériques avec le dispositif Tactos : la recherche d'une source de stimulation conduit l'utilisateur à déployer une activité perceptive dans l'espace de l'écran, et quand une première stimulation survient, l'utilisateur la perd instantanément. Trois raisons expliquent ce fait :

i) la source de stimulation est spatialement déterminée (la forme à explorer et ses éléments constitutifs sont inscrits dans l'espace de l'écran),
ii) le champ perceptif de l'utilisateur est local (le champ récepteur déplacé par l'utilisateur ne détecte que des coïncidences spatiales avec les pixels de la forme),
iii) le mouvement d'exploration a une certaine vitesse.

Ainsi, la rencontre d'une source de stimulation est un évènement local qui échappe dans un premier temps à l'utilisateur. La stratégie de l'utilisateur consiste alors à inverser son mouvement d'exploration (soit s'arrêter et produire une accélération inverse) de façon à essayer de retrouver la source de stimulation. S'il y parvient de nouveau, il peut progressivement organiser son activité perceptive de façon à tenter de contrôler la survenue des stimulations tactiles en fonction du mouvement d'exploration déployé. En d'autres termes, l'utilisateur peut progressivement anticiper la source de stimulation, ce qui rend possible la reconnaissance de la forme elle-même. Tout comme la perception d'un objet fixe se réalise dans la constitution d'un invariant sensori-moteur qui attire et organise mon activité (O'Regan & Noé, 2001 ; Lenay, 2005), la rencontre de deux activités perceptives intentionnellement dirigées l'une vers l'autre fait émerger un attracteur dynamique collectif (Auvray et al., 2009). Par conséquent, le pourcentage de stimulations est plus élevé dans le cas où le curseur de chaque sujet rencontre le curseur de l'autre participant précisément parce que l'activité perceptive de chaque participant attire l'autre activité (Auvray et al., 2009 ; Lenay, 2010). En effet, puisque les deux sujets sont engagés dans la recherche des sources de stimulation présentes dans l'espace, la survenue d'une stimulation tactile résultant de la rencontre de leurs curseurs respectifs entraînent les sujets dans « une sorte de danse » (ibid., p. 13) : chaque sujet inverse son mouvement et l'autre en fait autant, ce qui fait que les sujets ne

peuvent se déterminer spatialement comme dans le cas de la rencontre d'un objet fixe. Notons que dans le cas de l'objet fixe, la survenue des stimulations est *parfaitement contingente* avec le mouvement d'exploration de l'utilisateur. De son côté, le leurre mobile ne peut pas non plus être déterminé, puisqu'il reflète à distance l'activité du curseur auquel il est attaché : la rencontre du leurre mobile n'est *pas contingente* avec l'activité de celui qui le perçoit, puisque son mouvement n'est pas organisé en fonction des objets qu'il rencontre. La rencontre des deux curseurs est un évènement unique, puisqu'elle fait émerger un jeu subtil dans lequel chaque utilisateur tente de reconnaître l'autre, en essayant d'anticiper la stimulation sans pouvoir la localiser précisément, bien qu'elle soit toujours présente autour de chaque sujet (ibid.). En conséquence, chacun reçoit périodiquement une stimulation *imparfaitement contingente* avec son propre mouvement d'exploration.

Mais cette contingence imparfaite n'est précisément pas une donnée de l'environnement qu'il s'agit de percevoir. Je ne reconnais pas autrui parce qu'il me donne à percevoir un comportement plus ou moins contingent, mais je reconnais autrui parce que je suis engagé avec lui dans une activité collective. C'est cette activité qui fait émerger ces contingences mais aucun de nous ne contrôle réellement leur survenue. La différence dans les fréquences de clics est alors attribuable à l'émergence d'un couplage interpersonnel stable à partir de la rencontre des stratégies individuelles, bien qu'aucune d'elles ne soit parfaitement prévisible pour les participants (Di Paolo, Rohde & Iizuka, 2008). Ainsi, le processus de reconnaissance d'autrui est intrinsèque à l'activité perceptive partagée et à la dynamique mutuelle qu'elle engendre (Auvray et al., 2009). Ces résultats suggèrent donc que le processus de l'interaction est autonome, et qu'il prime sur la reconnaissance individuelle de l'intentionnalité des interactants. Ici, l'interaction ne joue pas seulement un rôle contextuel ou déclenchant, elle *constitue* la cognition sociale (De Jaegher et al., 2010), notamment en rendant possible l'émergence de contingences sociales, entendues à la fois comme l'origine et la conséquence de la coordination des comportements.

Des critiques ont été récemment formulées quant aux conclusions de cette expérience. Par exemple, des chercheurs montrent un certain scepticisme à considérer le rôle constitutif de l'interaction sociale (Michael & Overgaard, 2012). Selon les auteurs, la cognition sociale concerne la formation de jugements sociaux de la part des sujets. Le contexte de l'interaction fournirait donc uniquement un support sur lequel les sujets baseraient leur évaluation subjective de l'intentionnalité. Les sujets feraient alors appel à une heuristique simple : si la source de stimulation bouge, alors il s'agit d'autrui, la dynamique de l'interaction m'assurant que la plupart des rencontres résultent du croisement perceptif avec l'autre participant (ibid.).

Effectivement, la dynamique interpersonnelle est expliquée par la rencontre des stratégies individuelles, et il existe une part de clics non-négligeable sur le leurre mobile (Auvray et al., 2009). Pourtant, il nous semble qu'une explication individuelle ne suffit pas. L'intentionnalité perceptive déployée par chacun des deux sujets vise à se stabiliser autour d'une singularité, de façon à la constituer spatialement. Or, autrui résiste à cette constitution, malgré le fait que sa présence s'impose à ma perception et se maintient dans le temps (Lenay, 2010). Quelque chose de l'interaction m'échappe et seule la dynamique de la rencontre interpersonnelle vient stabiliser l'attracteur de la dynamique collective, alors que précisément, l'individu ignore tout de cette dynamique (ibid.). Il apparaît donc, conformément à une perspective énactive de l'interaction sociale, qu'une nouvelle signification émerge dans l'interaction alors qu'elle faisait défaut à l'individu isolé (De Jaegher & Di Paolo, 2007). L'interaction joue donc un rôle constitutif, en ce qu'elle rend présent ce qui n'est pas disponible à des individus isolés (De Jaegher et al., 2010). Mais pour autant, il n'est pas question de dissoudre le rôle des individus : si le comportement des participants est principalement guidé par la dynamique globale de l'interaction, la construction de signification demeure qualitativement non-affectée par la rencontre avec une autre activité perceptive dirigée vers elle (Froese & Di Paolo, 2011). L'individu doit donc re-saisir dans le processus de l'interaction des éléments qui vont faire sens (Lenay & Stewart, 2012).

Pour tenter d'apporter des éléments dans ce sens, une nouvelle expérience a été conduite (ibid.). Dans celle-ci, le protocole est identique, hormis le fait qu'un son est attribué à chacun des trois objets. Après chaque essai de 2 minutes, il est demandé aux sujets d'associer chacun des trois sons avec les objets correspondants. Les résultats montrent des scores élevés de catégorisation correcte pour les trois types d'objets (ibid.). Ici, ce qui explique la réussite individuelle est la capacité des sujets à distinguer la dynamique du croisement perceptif de la dynamique de la rencontre avec les autres objets, en particulier avec le leurre mobile. Les sons attribués à chaque objet ont permis de lever l'ambigüité dans les différentes situations d'interaction, et le croisement perceptif avec le partenaire est alors reconnu comme tel. La dynamique du croisement perceptif est la même, mais cette fois, elle peut être saisie par les participants en associant les propriétés intrinsèques des objets avec les propriétés caractéristiques de la dynamique de l'interaction. En d'autres termes, les résultats sont expliqués par une re-saisie individuelle de la dynamique collective (ibid.). Cette dynamique est perçue directement, à travers l'engagement mutuel des sujets, et permet en retour un ajustement dynamique et immédiat des individus relativement à elle.

5. Etude expérimentale : test de Turing minimaliste non-verbal[31]

De façon à tenter de caractériser cette dynamique, nous avons élaboré une étude expérimentale basée sur le célèbre « jeu d'imitation » de Turing (Turing, 1950), plus connu en intelligence artificielle sous le terme de « test de Turing ». A l'origine, il s'agissait d'une épreuve de pensée destinée à décider si une machine suffisamment complexe pouvait se faire passer pour un humain. Dans le test décrit par Turing, un sujet s'engage dans une conversation verbale à l'aide d'un clavier avec deux interlocuteurs placés dans deux pièces séparées : un autre sujet humain et un ordinateur. La tâche pour le premier humain consiste à essayer de savoir dans quelle pièce est la machine[32]. Si le sujet humain ne peut pas répondre, alors la machine aura réussi le test, et on pourra considérer qu'elle est intelligente, ou du moins qu'elle est capable de simuler l'intelligence (Searle, 1980). Le principe de ce test a motivé les partisans de l'intelligence artificielle classique, basé sur le langage et la manipulation de symboles, en prenant comme point de départ que le langage est nécessairement la preuve ultime de l'intelligence. Or, conformément aux positions théoriques adoptées ici, il apparaît que l'intelligence peut être mieux comprise comme une propriété qui émerge d'un agent incarné et situé capable de s'adapter à son environnement (Brooks, 1991). Cette capacité d'adaptation n'est pas nécessairement la capacité de verbalisation des raisons de l'action, mais plutôt l'habileté pragmatique à agir de façon appropriée dans des situations données (De Jaegher, et al., 2010). Ainsi, certains chercheurs ont élaboré des tests de Turing non-verbaux, que ce soit pour tester la plausibilité d'avatars virtuels (Ventrella, Seif El-Nasr, Aghabeigi & Overington, 2010), ou encore pour tester l'influence des contingences sociales sur l'interaction visuelle dans une tâche d'attention conjointe (Pfeiffer, Timmermans, Bente, Vogeley & Schilbach, 2011). Dans notre cas, l'utilisation d'un test de Turing non-verbal nous semble appropriée pour offrir un cadre d'analyse aux dynamiques interpersonnelles.

Nous proposons donc une étude expérimentale dans le contexte de la discrimination humain/robot, avec pour objectif de déterminer avec précision les caractéristiques essentielles à la dynamique interpersonnelle, notamment pour que celle-ci soit comprise par les sujets comme un engagement mutuel intentionnel. Dans cette optique, il est question de contrôler le

[31] Cette étude expérimentale a donné lieu à deux communications orales avec actes :

Deschamps, L., Lenay, C., Rovira, K., Stewart, J. & Aubert, D., (2011), Interagir pour construire du sens : une opportunité pour les dispositifs d'aide aux personnes en situation de déficience visuelle, *Colloque Jeunes Chercheurs et Jeunes Chercheuses de l'IFRATH, 23 juin 2011*, INJS Paris, p.33-43

Deschamps, L., Le Bihan, G., Lenay, C., Rovira, K., Stewart, J. & Aubert, D. (2012), Interpersonal recognition through mediated tactile interaction, *Proceedings of IEEE Haptics Symposium 2012*, March 4-7, 2012, Vancouver, BC, Canada, p. 239-245

[32] Il est important de noter que l'humain et la machine font tout pour être reconnu comme un humain.

type de dynamique avec laquelle le sujet humain va s'engager, et d'observer l'effet de la variation de cette dynamique sur la reconnaissance d'un engagement mutuel de la part du sujet. Plus précisément, nous nous intéressons à l'effet de différents niveaux de contingence dans l'interaction, de façon à les comparer à ceux que nous pourrons constater dans les interactions impliquant deux sujets humains. Contrairement à l'expérience princeps sur le croisement perceptif minimaliste (Auvray et al., 2009), les sujets ne pourront rencontrer qu'une seule entité dans un environnement unidimensionnel : un autre sujet humain ou un robot. Nous nous plaçons donc dans le cadre d'interactions strictement dyadiques. Les robots joueront le rôle de « leurres mobiles », mais cette fois, leur comportement ne sera pas le reflet à distance du comportement d'un sujet humain. Chaque robot est défini soit par la manifestation d'une plus ou moins grande régularité dans leur comportement, soit par leur capacité à produire de la contingence en fonction de la rencontre des sujets humains. Dans ce cadre, nous postulons que malgré la capacité de certains robots à produire des réponses imparfaitement contingentes, les sujets humains parviendront à se reconnaître mutuellement en vertu d'une co-régulation dynamique de leurs activités perceptives, qui ne se réduit pas à la détection unilatérale de contingences imparfaites. Malgré cela, si certains robots, en particulier les robots destinés à interagir de façon contingente, induiront les sujets ponctuellement en erreur, ceux-ci seront capables de produire de moins en moins d'erreurs au fur et à mesure de l'expérience. En effet, la répétition des essais se traduira par une stabilisation progressive de la dynamique interpersonnelle, cette dernière pouvant ainsi servir de support de plus en plus pertinent pour la reconnaissance mutuelle des interactants. Si ces résultats sont avérés, alors l'analyse précise des trajectoires perceptives de chaque participant, ainsi que de la dynamique de leurs rencontres successives, permettra de dégager la signature objective de la reconnaissance intersubjective des participants. En d'autres termes, il est question de tenter de caractériser la dynamique des rencontres interpersonnelles.

5.1. Population

54 sujets, répartis en 3 groupes expérimentaux de 18 sujets chacun, ont participé à cette expérience. Les sujets sont majoritairement des étudiants de l'Université de Rouen et de l'Université de Technologie de Compiègne âgés entre 18 et 26 ans, recrutés par un système de petites annonces au sein de leurs établissements respectifs. Notons que les autres sujets sont des enseignants-chercheurs de l'Université de Rouen. Ils n'ont pas de connaissance particulière du dispositif et n'ont pas connaissance des hypothèses théoriques sous-jacentes à l'expérience.

Les groupes expérimentaux sont constitués en fonction du niveau de complexité des robots.

5.2. Matériel

Les sujets sont assis devant un ordinateur portable équipé du dispositif Tactos. Ce dispositif permet au sujet de déplacer un curseur sur un écran d'ordinateur via un effecteur, en l'occurrence une souris d'ordinateur. Le sujet manipule la souris avec sa main dominante, tandis que l'index de sa main non-dominante est posé sur un boîtier de stimulateurs tactiles (en l'occurrence deux cellules Braille piézo-électriques configurées en tout ou rien). Lorsque le curseur déplacé à l'écran rencontre un autre curseur, déplacé par un autre sujet ou contrôlé par un robot, cela active la montée simultanée des 16 picots du boîtier de stimulateurs tactiles.

Puisque pour un essai donné, les trois sujets humains peuvent être en face de robots, le dispositif expérimental comporte 7 ordinateurs, soit un serveur qui va gérer les connexions, 3 ordinateurs-client destinés aux sujets humains et 3 ordinateurs-client destinés au contrôle des robots. Tous ces ordinateurs sont connectés via un réseau local. De cette façon, les délais inhérents au transfert des informations à travers le réseau sont maintenus constants pour tous les essais.

L'espace dans lequel se déplacent les sujets est un espace horizontal unidimensionnel de 200 pixels qui se boucle sur lui-même. En d'autres termes, il s'agit d'un tore géométrique dans lequel le sujet n'a pas la conscience de bords ou de limites. Du point de vue d'un observateur, le franchissement du bord gauche par un curseur se traduit par une réapparition simultanée à droite, et inversement.

A l'aide de la souris, les sujets déplacent un champ récepteur élémentaire de 2 pixels de long (le corps percevant, qui permet de percevoir le curseur des autres sujets ou des robots) auquel correspond topologiquement un avatar de 2 pixels de long (le corps-image, perceptible par les autres sujets). Les déplacements s'effectuant dans un espace 1-D horizontal, les déplacements verticaux ne sont pas pris en compte.

Les robots contrôlent un avatar de 2 pixels de long. De cette façon, les avatars déplacés par les sujets et les avatars-robots possèdent des caractéristiques semblables (même taille, même forme). En revanche, les avatars-robots sont déplacés par l'intermédiaire d'algorithmes informatiques (scripts LUA), produisant des comportements déterminés en fonction de trois niveaux de complexité. Ceux-ci sont définis par la manipulation de trois paramètres : la vitesse de déplacement, l'amplitude du mouvement et la possibilité d'arrêt. En revanche, tous les robots partagent les mêmes règles initiales. Premièrement, tous les robots effectuent des

oscillations latérales qui débutent à une position aléatoire sur le tore. Deuxièmement, chaque changement de direction décrit une courbe sinusoïdale, soit une décélération suivie d'une ré-accélération inverse : l'aller-retour ainsi constitué forme un cycle de mouvement.

Les robots du groupe 1 (***robots simples***) ont une amplitude de mouvement fixe (35 pixels pour un cycle) et une vitesse maximale fixe (120 pixels / seconde), se traduisant par une oscillation latérale rythmique ; les robots du groupe 2 (***robots intermédiaires***) ont une amplitude de mouvement et une vitesse qui varient périodiquement (+/- 40% des valeurs initiales appliquées à chaque demi-cycle); les robots du groupe 3 (***robots complexes***) sont identiques aux robots intermédiaires, mais présentent une probabilité d'arrêt (5% de chances de s'arrêter à chaque demi-cycle pour une durée comprise entre 1 et 5 secondes) (voir Figure 50). Pour chaque niveau de complexité, deux types de robots sont constitués : des robots élaborés en termes de ***programme clos*** (les cycles de mouvement sont déployés dans l'espace 1-D indépendamment des mouvements du sujet humain), et des robots élaborés en termes de ***programme d'interaction*** (les cycles de mouvement sont provoqués par la rencontre du sujet humain dans l'espace 1-D). De façon à ce que les programmes clos ne soient pas identifiés uniquement par une zone d'oscillation persistante tout au long de l'essai, une légère translation (5 pixels dans une direction aléatoire) est appliquée à chaque début de cycle (voir Figure 50).

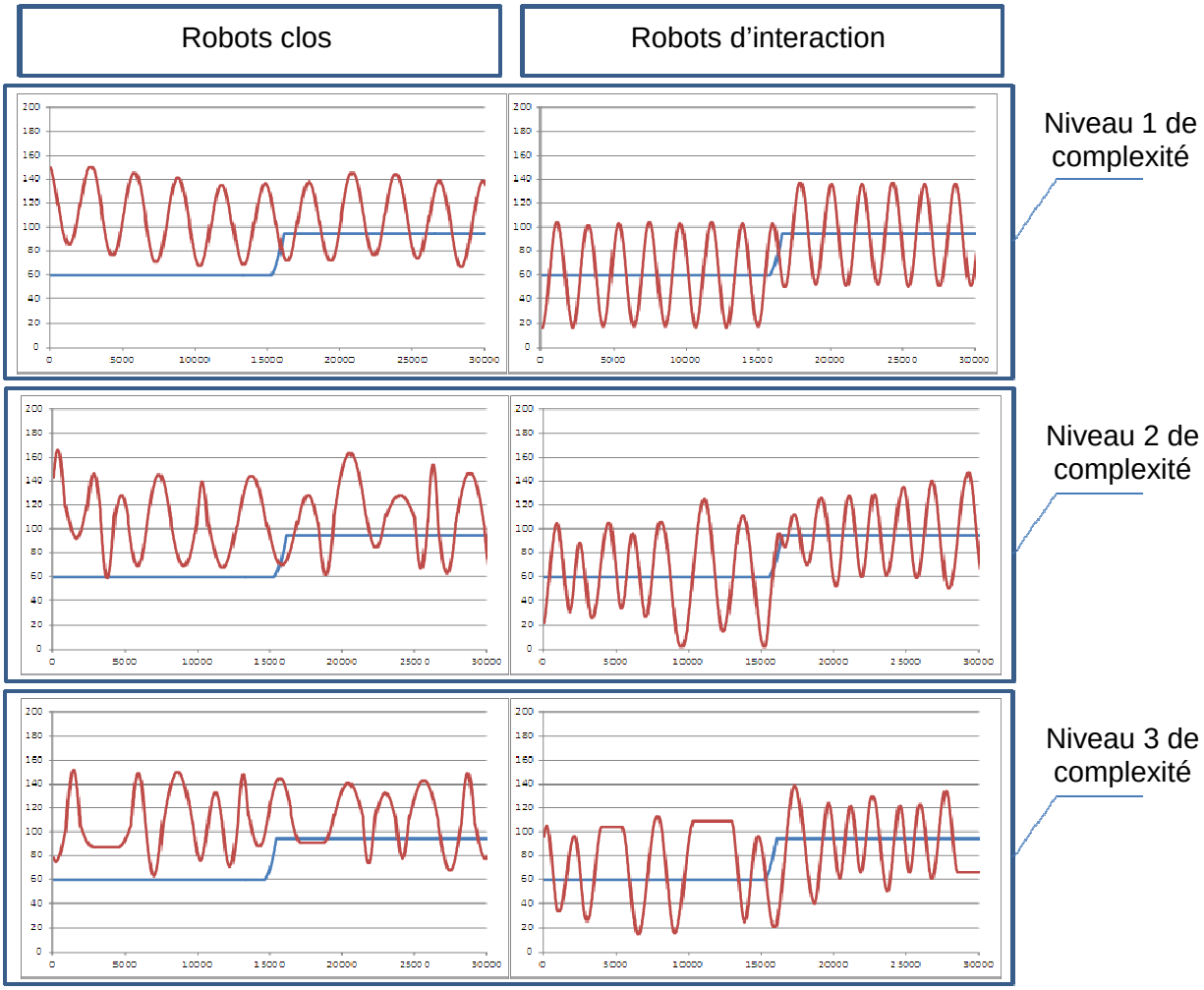


Figure 50 : Trajectoires déployées par les robots clos et les robots d'interaction en fonction des 3 niveaux de complexité. Pour chaque graphique, l'axe des abscisses représente le temps (en ms), et l'axe des ordonnées représente la position en x (en pixels). Les trajectoires rouges représentent le comportement des robots, tandis que les trajectoires bleues représentent la trajectoire hypothétique d'un sujet humain qui occuperait dans un premier temps une position fixe, avant de se déplacer légèrement et de s'arrêter à une nouvelle position fixe.

5.3. Procédure

Trois groupes indépendants sont définis en fonction de chaque niveau de complexité défini ci-dessus. Dans chaque groupe expérimental, 18 sujets sont répartis en 6 trinômes, chaque trinôme donnant lieu à une passation. Les sujets, regroupés dans une même salle n'ont pas les yeux bandés pendant l'expérience, mais aucune information n'est visible sur leur écran. De plus, des précautions ont été prises de façon à ce que les sujets ne puissent ni se voir ni s'entendre : la seule interaction permise étant une interaction tactile à travers le dispositif décrit ci-dessous.

Pour chaque groupe, un robot en programme clos et un robot en programme d'interaction d'un même niveau de complexité sont utilisés. De cette façon, le premier groupe est confronté aux robots simples, le deuxième aux robots intermédiaires, et le troisième aux

robots complexes. Chaque passation concerne donc 3 sujets confrontés à deux robots. De cette façon, chaque sujet peut interagir selon les essais avec un des deux autres sujets (condition « Humain ») ou avec un des deux robots d'un même niveau de complexité (condition « Robot »).

Chaque trinôme passe 17 essais : un essai de familiarisation et 16 essais expérimentaux. Chaque essai dure 1 minute. Chaque sujet rencontre 4 fois chacun des deux autres sujets et 4 fois chacun des deux robots dans un ordre fixé préalablement. De façon à contrôler un éventuel effet d'ordre des essais, quatre scénarii sont élaborés et sont contrebalancés lors des différentes passations. Notons que lorsque deux sujets se rencontrent lors d'un essai donné, le troisième rencontre nécessairement un des deux robots. Toutefois, dans certains essais, les trois sujets rencontrent simultanément et respectivement un des deux robots.

Au début de l'expérience, les sujets sont informés qu'à chaque essai, ils pourront interagir tactilement soit avec un robot, soit avec un humain. A la fin de chaque essai, ils devront indiquer sur une feuille de réponse s'ils pensent avoir interagi avec un robot ou avec un humain. Aucune information supplémentaire quant à la nature des robots, à leur nombre ou à la façon dont ils se déplacent n'est délivrée. De plus, aucun feedback n'est donné aux sujets pendant la durée de l'expérience.

Par ailleurs, une pause de 5 minutes est proposée après l'essai n° 8, ce qui nous permet d'obtenir deux blocs de 8 essais. Notons que durant la pause, les sujets ne peuvent pas communiquer entre eux.

Pour tous les groupes, un premier essai, dit de familiarisation, est proposé pendant 2 minutes. Il s'agit d'explorer l'espace et d'y rencontrer un objet fixe de deux pixels de long. Cette étape est nécessaire pour prendre connaissance du dispositif. Il s'agit pour chaque sujet de comprendre que la montée des picots résulte de la rencontre de son avatar, qu'il déplace le long du tore à l'aide de la souris, avec un objet numérique dans l'espace. D'autre part, cela permet au sujet d'évaluer de façon subjective la longueur du tore, c'est-à-dire la distance qui sépare deux rencontres successives avec l'objet fixe s'il se déplace toujours dans la même direction. A la fin des deux minutes, l'expérimentateur s'assure que le principe du dispositif est bien compris, et lance la phase de test par l'intermédiaire d'une consigne générale : à la fin de chaque essai, chaque sujet doit se prononcer sur la nature de l'avatar avec lequel il a interagi (humain ou robot). En outre, durant chaque essai, l'activité des sujets est enregistrée sous forme de traces numériques.

5.4. Hypothèses opérationnelles

- Les sujets humains donneront plus de réponses « Humain » face aux robots complexes que face aux robots intermédiaires et aux robots simples. Le niveau de complexité des robots pourra donc induire les sujets en erreur.
- Les programmes d'interaction entraineront plus d'erreurs que les programmes clos, et ce d'autant plus que le niveau de complexité des robots sera élevé. Les programmes d'interaction intermédiaires et complexes sont en effet destinés à produire des réponses imparfaitement contingentes, ce qui pourrait entraîner les sujets dans une dynamique d'interaction caractéristique d'un engagement mutuel intentionnel.
- Malgré cela, les sujets humains seront à même de reconnaître les sujets humains en vertu de l'émergence d'un processus dynamique de co-ajustements et de co-régulation du processus de l'interaction, qui ne se réduit pas seulement à une détection de contingences. Par ailleurs, les sujets humains reconnaîtront davantage les autres sujets humains dans le deuxième bloc d'essais (essais 9 à 16), en vertu de la stabilisation progressive de cette dynamique d'interaction.
- Enfin, les sujets se reconnaîtront davantage de façon mutuelle que de façon unilatérale, la reconnaissance elle-même émergeant du processus de l'interaction, et non d'un jugement individuel unilatéral.

5.5. Résultats

Pour l'analyse des résultats, nous avons comparé les fréquences de réponses « Humain » en fonction du facteur « Niveau de complexité des robots » et du facteur « Type d'interactant » (Humain vs Robot), puis en fonction du facteur « Niveau de complexité des robots » et du facteur « Type de robot » (Programme clos vs Programme d'interaction). Ensuite, nous nous sommes intéressés plus précisément aux essais de la condition « Humain » en comparant les modalités de reconnaissance (Reconnaissance mutuelle, Reconnaissance unilatérale et Non-reconnaissance) en fonction des deux blocs d'essais.

5.5.1. Type d'interactant

Pour rappel, chaque sujet est confronté à 4 interactants (rencontrés chacun 4 fois tout au long des 16 essais expérimentaux) : deux sujets humains et deux robots, ces derniers étant constitués d'un robot élaboré en termes de programme clos et d'un robot élaboré en termes de programme d'interaction d'un même niveau de complexité.

Nous regardons tout d'abord s'il y a un effet du type d'interactant (Humain vs Robot) sur les fréquences de réponse « Humain » quel que soit le niveau de complexité des robots.

De façon générale et tous groupes confondus, nous observons que la fréquence de réponses « Humain » est supérieure lorsque le sujet se trouve face à un humain que lorsqu'il se trouve face à un robot (d = 0.28) (voir Figure 51).

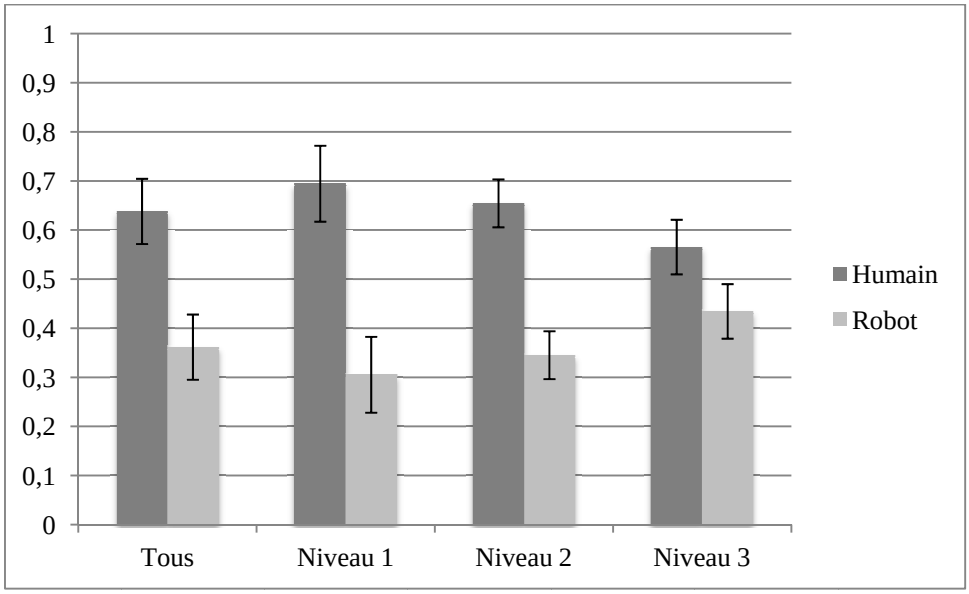


Figure 51 : Fréquences de réponses « Humain » en fonction du type d'interactant (condition « Humain » vs condition « Robot ») et des niveaux de complexité de robots

En d'autres termes, les sujets répondent plus fréquemment « Humain » lorsqu'ils ont effectivement interagi avec les humains que lorsqu'ils ont interagi avec les robots. En outre, les robots sont plus souvent identifiés par erreur comme des humains que les humains comme des robots. Cette différence est statistiquement significative (test de Student : $t_{(51)} = 8.2230$; $p < 0.0001$), et peut même être considérée comme notable d'après une analyse bayésienne ($Pr |\delta/\sigma| > .66 = .99$). Nous observons un effet d'interaction significatif entre le facteur Niveau de complexité et le facteur Type d'interactant (test de Fisher : $F_{(2,51)} = 5.1495$; $p = 0.0092$). En effet, la différence entre la condition « Humain » et la condition « Robot » est élevée pour le premier niveau de complexité (d = 0.39), modérée pour le second niveau (d = 0.31), et chute pour le troisième niveau (d = 0.13) (voir Figure 51). Néanmoins, si la différence entre les deux conditions est significative pour chaque niveau de complexité, les analyses bayésiennes révèlent que l'effet est notable en ce qui concerne le niveau 1 et 2, mais pas pour le niveau 3 (voir Tableau 6).

Complexité	Différence	Test	Probabilité	Taille d'effet
Niveau 1	0.3891	$t_{(17)} = 5.3376$	$p < 0.0001$	Pr $\|\delta/\sigma\| > .66 = .97$ (notable)
Niveau 2	0.3092	$t_{(17)} = 6.7294$	$p < 0.0001$	Pr $\|\delta/\sigma\| > .66 = .98$ (notable)
Niveau 3	0.1310	$t_{(17)} = 2.5002$	$p = 0.0229$	Pr $\|\delta/\sigma\| > .66 = .37$ (non notable)

Tableau 6 : Récapitulatif des résultats statistiques concernant les fréquences de réponse « Humain » dans les conditions « Humain » et « Robot » en fonction des niveaux de complexité.

Il semble donc que le niveau de complexité des robots exerce une influence relative sur la fréquence de réponses « Humain » correctes. En effet, même dans le cas des robots les plus complexes, la fréquence de réponses « Humain » est significativement plus élevée lorsque les sujets ont effectivement interagi avec les autres sujets que lorsqu'ils ont interagi avec les robots. Pourtant, nous observons que la fréquence de réponses « Humain » correctes diminue à mesure que le niveau de complexité grandit, ce qui implique que les robots les plus complexes sont plus souvent identifiés par erreur comme des humains que les robots intermédiaires et les robots simples.

5.5.2. Apprentissage

D'une façon globale, nous observons que quel que soit le type d'interactant, les fréquences de bonnes réponses évoluent en fonction des deux blocs d'essais (essais 1 à 8 versus essais 9 à 16) (voir Figure 52).

Les fréquences de bonnes réponses sont plus élevées dans le deuxième bloc d'essais que dans le premier bloc d'essais ($d = 0.18$). L'effet d'apprentissage est significatif (test de Student : $t_{(51)} = 5.1107$; $p < 0.0001$), mais ne peut pas être considéré comme notable (Pr $|\delta/\sigma| > .66 = .57$). Cet effet est observé quel que soit le niveau de complexité des robots (voir Figure 52) : l'effet d'interaction entre le facteur « Niveau de complexité » et le facteur « Apprentissage » n'est pas significatif (test de Fisher : $F_{(2 ; 51)} = 1.2863$; $p = 0.2851$).

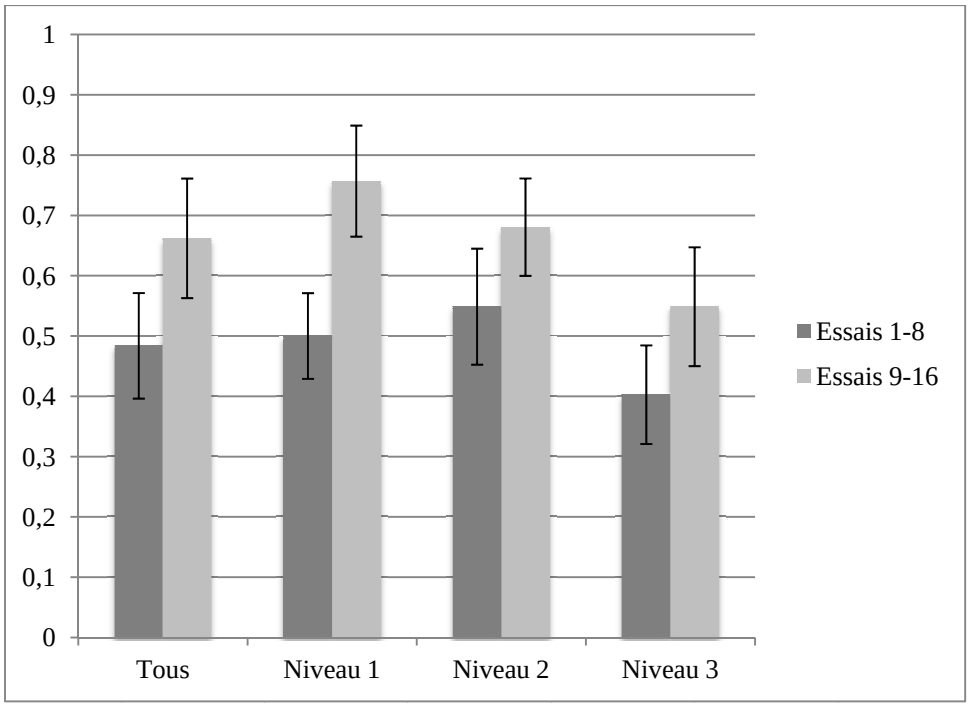


Figure 52 : Fréquences de bonnes réponses quel que soit le type d'interactant en fonction du niveau de complexité des robots (Niveau 1, Niveau 2 et Niveau 3) et des deux blocs d'essais (Essais 1-8 et Essais 9-16

5.5.3. Type de robots

Pour poursuivre cette analyse, nous nous penchons désormais sur les fréquences de réponses « Humain » obtenues lors de la condition « Robot » en fonction du type de robot (Programme clos vs Programme d'interaction). En d'autres termes, nous nous intéressons aux erreurs d'identification commises par les sujets lorsqu'ils ont interagi avec les deux types de robots.

Contrairement à ce que nous attendions, les programmes clos sont plus souvent identifiés par erreur comme des humains que les programmes d'interaction (d = 0.11) (voir Figure 53). Toutefois, cette différence n'est pas significative (test de Student : $t_{(51)} = 2.0040$; $p = 0.0504$), même si elle ne peut pas être considérée comme négligeable ($Pr\ |\delta/\sigma| < .33 = .67$).

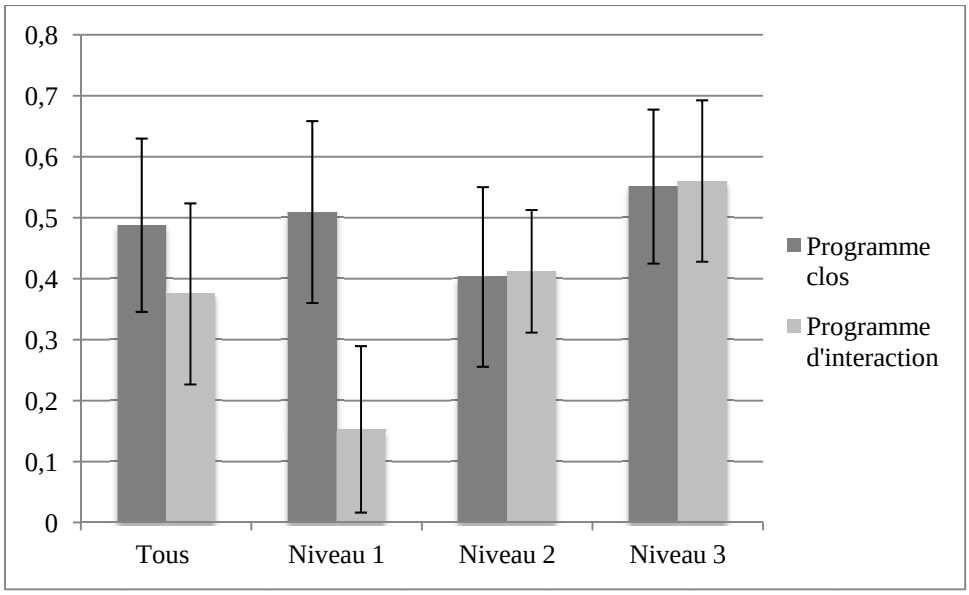


Figure 53 : Fréquences de réponses « Humain » en fonction du type de robot (Programme clos vs Programme d'interaction) et des niveaux de complexité de robots (Niveau 1, Niveau 2 et Niveau 3)

Par ailleurs, nous constatons que cette différence n'est observée que pour le groupe confronté aux robots les plus simples (d = 0.36), tandis que les groupes confrontés aux robots intermédiaires et aux robots complexes présentent très peu de différences. L'effet d'interaction entre le facteur « Niveau de complexité » et le facteur « Type de robot » est statistiquement significatif (test de Fisher : $F_{(2 ; 51)} = 4.7035$; $p = 0.0133$).

5.5.4. Reconnaissance mutuelle

Enfin, une dernière analyse concerne plus spécifiquement la condition « Humain ». Nous nous intéressons ici aux modalités de reconnaissance, définies selon trois cas de figure : la reconnaissance mutuelle (les deux sujets s'identifient mutuellement comme des humains), la reconnaissance unilatérale (seul un des deux identifie l'autre comme un humain), et la non-reconnaissance (aucun des deux n'identifie l'autre comme un humain). Ainsi, nous comparons les fréquences d'occurrence des différentes modalités de reconnaissance en fonction du niveau de complexité des robots.

Globalement, nous observons que la fréquence de reconnaissance mutuelle est plus élevée que la fréquence de reconnaissance unilatérale (d = 0.10). De plus, cette dernière est elle-même plus élevée que la fréquence de non-reconnaissance (d = 0.17). Il apparaît que la fréquence de reconnaissance mutuelle n'est pas statistiquement différente de la fréquence de reconnaissance unilatérale (test de Student : $t_{(15)} = 1.3370$, $p = 0.2011$). En revanche, la reconnaissance mutuelle est statistiquement plus fréquente que la non-reconnaissance (test de Student : $t_{(15)} = 4.1167$, $p = 0.0009$), même si cet effet ne peut pas être considéré comme

notable (Pr |δ/σ| >.66 = .84). De la même façon, la reconnaissance unilatérale est statistiquement plus fréquente que la non-reconnaissance (test de Student : $t_{(15)} = 2.8131$, $p = 0.0131$), même si cet effet ne peut pas être considéré comme notable (Pr |δ/σ| > .66 = .48). Si cette hiérarchie est retrouvée pour le niveau 1 et le niveau 3, il apparaît que la fréquence de reconnaissance mutuelle et la fréquence de reconnaissance unilatérale sont quasiment équivalents pour le niveau 2 (d = 0.01) (voir Figure 54). Toutefois, nous ne pouvons pas conclure a un effet d'interaction entre le facteur « Niveau de complexité » et les modalités de reconnaissance (test de Fisher : $F_{(4,30)} = 1.2847$; $p = 0.2981$).

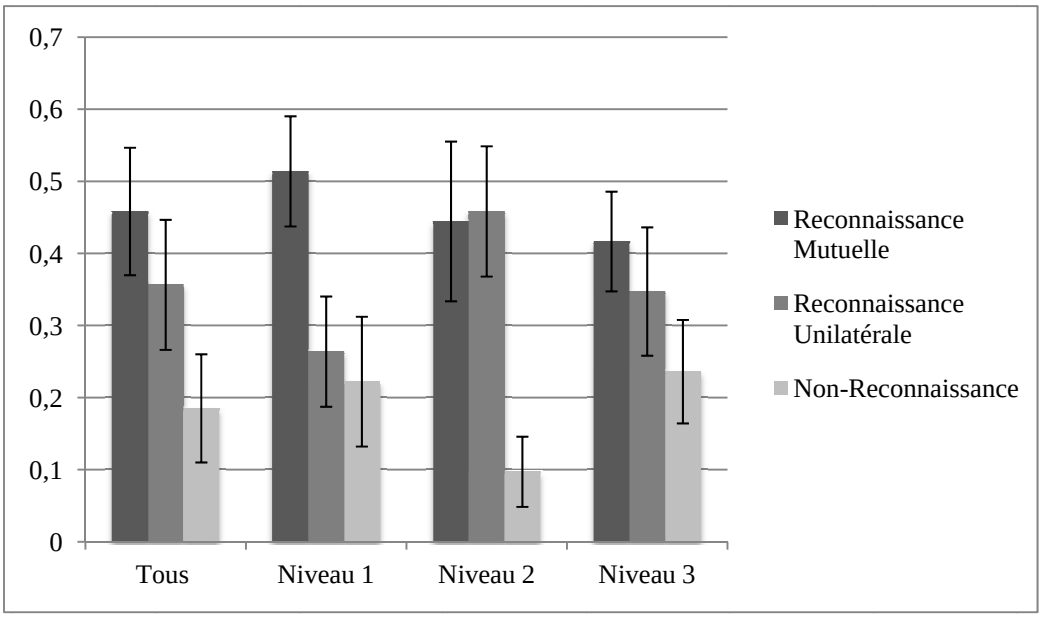


Figure 54 : Fréquences de reconnaissance mutuelle, de reconnaissance unilatérale et de non-reconnaissance pour la condition « Humain » en fonction des niveaux de complexité des robots (Niveaux 1, 2 et 3)

Par ailleurs, nous ne pouvons pas conclure à un effet du facteur « Niveau de complexité » pour aucune des modalités de reconnaissance, que ce soit pour la reconnaissance mutuelle (test de Fisher : $F_{(2,15)} = 0.4503$; $p = 0.6458$), pour la reconnaissance unilatérale (test de Fisher : $F_{(2,15)} = 1.9526$; $p = 0.1763$), et pour la non-reconnaissance (test de Fisher : $F_{(2,15)} = 1.6855$; $p = 0.2186$).

En ce qui concerne les modalités de reconnaissance en fonction des deux blocs d'essais, nous observons que la fréquence de reconnaissance mutuelle est plus élevée dans le deuxième bloc d'essais que dans le premier bloc d'essais (d = 0.28) (voir Figure 55). Cette différence est statistiquement significative (test de Student : $t_{(15)} = 4.7730$; $p = 0.0002$), et peut être considérée comme notable (Pr |δ/σ|>.66 = .92). Parallèlement, la fréquence de reconnaissance unilatérale est moins élevée dans le deuxième bloc d'essais que dans le premier bloc d'essais

(d = 0.20). Cette différence est statistiquement significative (test de Student : $t_{(15)} = 2.7020$; p = 0.0164), mais ne peut pas être considérée comme notable (Pr $|\delta/\sigma|>.66 = .44$). Enfin, la fréquence de non-reconnaissance est moins élevée dans le dernier bloc d'essais que dans le premier bloc d'essais (d = 0.10). Cette différence n'est pas statistiquement significative (test de Student : $t_{(15)} = 1.7438$; p = 0.1017), mais ne peut pas être considérée comme négligeable (Pr $|\delta/\sigma|>.33 = .39$).

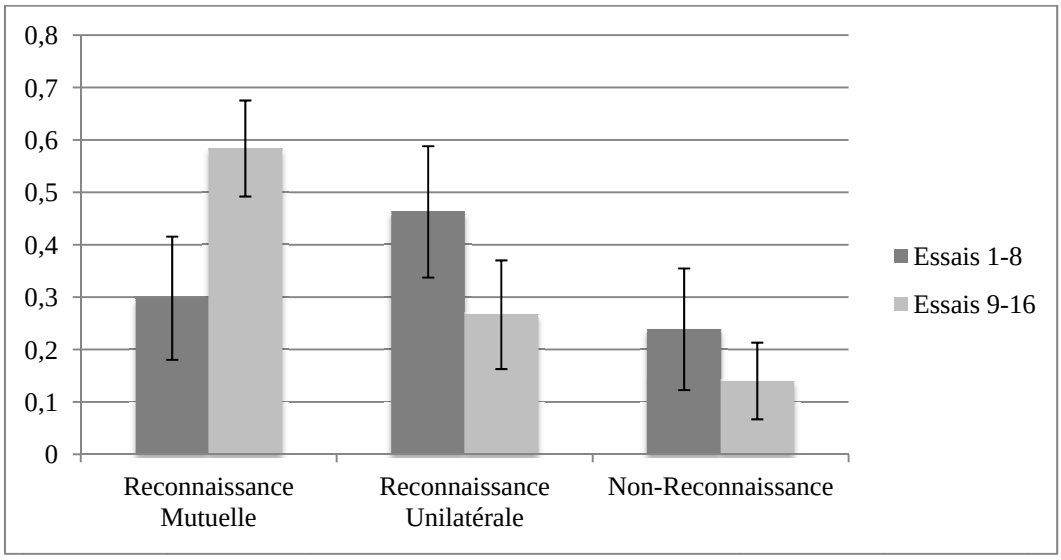


Figure 55 : Fréquences de reconnaissance mutuelle, de reconnaissance unilatérale et de non-reconnaissance en fonction des deux blocs d'essais (essais 1-8 et essais 9-16)

Enfin, il est important de noter que les résultats observés en ce qui concerne la reconnaissance mutuelle ne peuvent pas être expliqués par la probabilité que les sujets répondent « Humain » mutuellement par hasard. En effet, en ce qui concerne l'ensemble des essais, la fréquence de reconnaissance mutuelle est de 45,83%, alors que la probabilité conditionnelle que les deux sujets répondent mutuellement « Humain » est de 25,34%[33]. En ce qui concerne le deuxième bloc d'essais, la fréquence de reconnaissance mutuelle est de 58,38%, alors que la probabilité conditionnelle que les deux sujets répondent mutuellement « Humain » est de 28,59%.

[33] Nous calculons la probabilité conditionnelle en déterminant la probabilité de réponse « Humain » donnée par un sujet lorsque son partenaire a répondu « Humain ».

5.6. Analyses de trajectoires

Parallèlement à l'analyse des réponses des sujets, nous avons procédé à une analyse des trajectoires perceptives déployées par les sujets lors de chaque essai[34]. Ces analyses sont rendues possibles par la méthodologie minimaliste à laquelle nous avons adhérée, ainsi qu'à la médiation technique qui a permis l'enregistrement de ces trajectoires.

D'une façon générale, nous avons analysé une multitude d'indicateurs, de façon à tenter de comprendre les réponses des sujets et de caractériser leurs trajectoires perceptives. D'un point de vue méthodologique, nous sommes partis d'une analyse qualitative des trajectoires, à partir de leur représentation graphique (voir des exemples en Annexe C, p. 292). Celle-ci nous a permis de poser un certain nombre d'hypothèses, de façon à guider une analyse quantitative des indicateurs qui nous semblaient pertinents. Ainsi, nous avons repéré que l'interaction s'organisait bien souvent en séquences distinctes. Par souci de concision, nous ne présenterons ici que les indicateurs qui concernent ces séquences d'interaction, à la fois d'un point de vue microscopique (qualité de l'accroche mutuelle entre les interactants) et macroscopique (alternance de séquences d'interaction et de ruptures de l'accroche mutuelle au cours des essais).

Dans cette analyse, nous nous intéressons spécifiquement aux réponses « Humain », selon trois axes : la qualité de la réponse « Humain » (Réponse correcte vs Réponse incorrecte) ; les erreurs concernant les réponses « Humain » (Réponse « Humain » alors que le sujet a interagi avec un programme clos vs Réponse « Humain » alors que le sujet a interagi avec un programme d'interaction) ; et les modalités collectives de réponse dans la condition « Humain » (Reconnaissance Mutuelle, Reconnaissance Unilatérale, et Non Reconnaissance).

5.6.1. Proportion d'accroche mutuelle

Tout d'abord, nous nous sommes intéressés à la « proportion d'accroche mutuelle ». Pour chaque essai, nous avons calculé la durée de l'ensemble des séquences d'interaction pendant lesquelles les écarts entre les stimulations sont inférieurs à deux secondes. Nous calculons alors une proportion d'accroche mutuelle en additionnant ces durées avec les durées des stimulations correspondantes, et en divisant cette somme par la durée totale d'un essai. Notons que dans ce calcul, nous ne prenons pas en compte la durée des arrêts mutuels (cas où

[34] Pour plus de détails concernant la méthode d'analyse utilisée pour l'étude des trajectoires, veuillez vous reporter à l'Annexe B, p. 247.

les deux interactants s'arrêtent l'un sur l'autre) lorsque ceux-ci sont supérieurs à deux secondes.

Que ce soit de façon globale (tous niveaux de complexité confondus) ou de façon spécifique (dans chacun des trois niveaux de complexité), la proportion d'accroche mutuelle ne présente pas de différence significative entre les essais où le sujet répond « Humain » correctement et les essais où le sujet répond « Humain » par erreur (voir Tableau 7).

Niveaux	Différence	Test	Probabilité	Taille d'effet
Tous	- 0.30%	U = 1333 ; $Z_{ajusté}$ = - 0.610619	p = 0.5414	x
1 (simples)	- 2.11%	$t_{(33)}$ = 0.5931	p = 0.5571	B = .60 (non nég.)
2 (interm.)	- 3.26%	U = 134 ; $Z_{ajusté}$ = - 0.885881	p = 0.3756	x
3 (complexes)	4.50%	U = 153 ; $Z_{ajusté}$ = - 0.284747	p = 0.7758	x

Tableau 7 : Récapitulatif des résultats statistiques concernant la proportion d'accroche mutuelle en fonction de la qualité de la réponse « Humain » (Bonne réponse vs Mauvaise réponse) et des niveaux de complexité des robots.

Par ailleurs, nous observons que dans le cas des réponses « Humain » fournies par erreur, la proportion d'accroche mutuelle est toujours statistiquement inférieure lorsque le robot est un programme clos que lorsqu'il s'agit d'un programme d'interaction (voir Figure 56 et Tableau 8). Notons cependant que la proportion d'accroche mutuelle avec les programmes clos ne varie pas statistiquement en fonction des niveaux de complexité (test de Kruskall-Wallis : $H_{(2,48)}$ = 0.3197279 ; p = 0.8523) ; alors qu'elle décline progressivement dans le cas des programmes d'interaction (test de Kruskall-Wallis : $H_{(2,48)}$ = 6.688262 ; p = 0.0353).

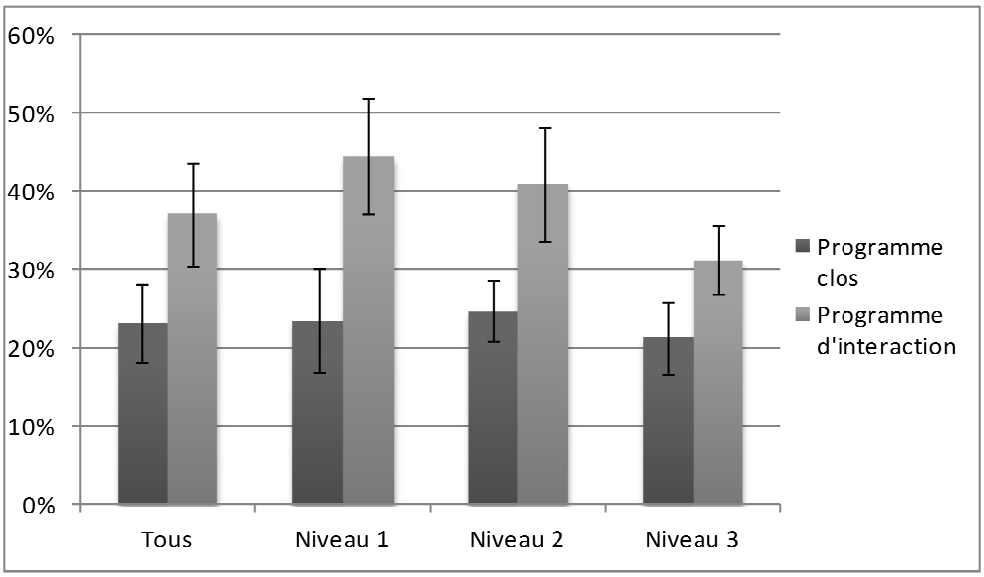


Figure 56 : Proportion d'accroche mutuelle lorsque le sujet répond « Humain » par erreur en fonction du type de robot (Programme clos vs Programme d'interaction) et des niveaux de complexité des robots.

Niveaux	Différence	Test	Probabilité	Taille d'effet
Tous	- 13.92%	$t_{(86)} = 5.6066$	$p < 0.0001$	B = .99 (notable)
1 (simples)	- 21.01%	$t_{(20)} = 3.2594$	$p = 0.0039$	B = .95 (notable)
2 (interm.)	- 16.05%	$U = 35$; $Z_{ajusté} = - 3.35992$	$p = 0.0007$	g = 1.36 (fort)
3 (complexes)	- 9.88%	$t_{(33)} = 3.2561$	$p = 0.0026$	B = .88 (non notable)

Tableau 8 : Récapitulatif des résultats statistiques concernant la proportion d'accroche mutuelle en fonction du type de robot (Programme clos vs Programme d'interaction) et des niveaux de complexité des robots.

Enfin, en ce qui concerne la condition « Humain », nous observons que la proportion d'accroche mutuelle est plus élevée lorsque les sujets se reconnaissent mutuellement que lorsque la reconnaissance est unilatérale (d = 1.50), où lorsqu'aucun des deux sujets ne reconnaît l'autre (d = 7.55) (voir Figure 57). La différence globale est significative (test de Kruskall-Wallis : $H_{(2,50)} = 6.071933$; $p = 0.0480$).

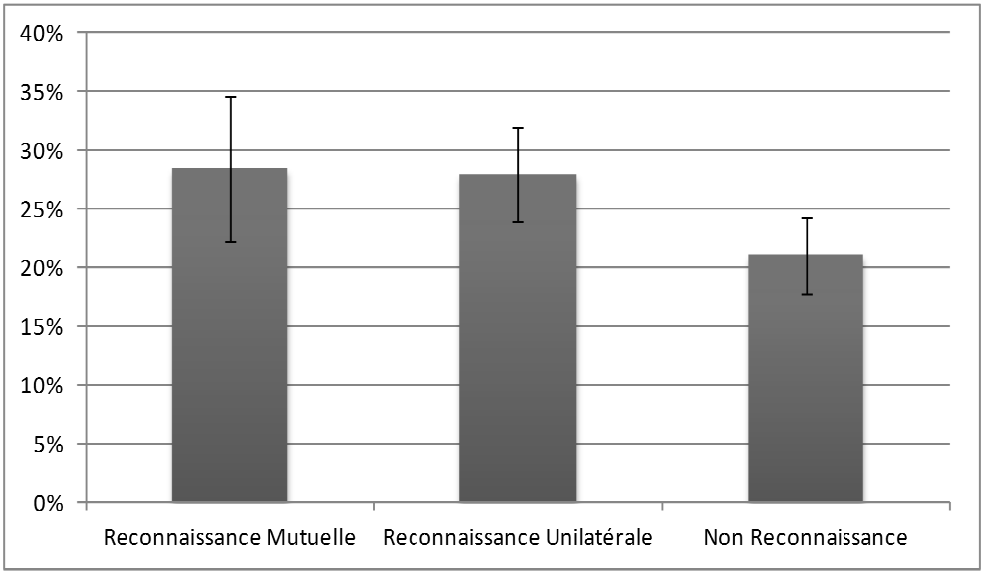


Figure 57 : Proportion d'accroche mutuelle dans la condition « Humain » en fonction des trois modalités collectives de réponse (Reconnaissance Mutuelle, Reconnaissance Unilatérale et Non Reconnaissance).

Toutefois, de façon plus spécifique, la différence entre « Reconnaissance Mutuelle » et « Reconnaissance Unilatérale » n'est pas significative (test de Mann & Withney : U = 157 ; $Z_{ajusté}$ = 0.158193 ; p = 0.8743). En revanche, la différence entre « Reconnaissance Mutuelle » et « Non-Reconnaissance » est significative (test de Mann & Withney : U = 71 ; $Z_{ajusté}$ = 2.089277 ; p = 0.0366), tout comme la différence entre « Reconnaissance Unilatérale » et « Non-Reconnaissance » (test de Mann & Withney : U = 67 ; $Z_{ajusté}$ = 2.241225 ; p = 0.0251). Notons que ces deux effets peuvent être considérés comme modérés (respectivement g = 0.72 et g = 0.68).

5.6.2. Variabilité de l'accroche mutuelle

De façon à caractériser les séquences d'accroche que nous avons identifiées, nous avons calculé la variabilité des écarts temporels entre les stimulations, en prenant comme base de données les séquences où ces écarts sont inférieurs à deux secondes. Pour cela, nous avons calculé l'écart-type des écarts entre les stimulations dans les séquences d'accroche mutuelle pour chaque interaction. Ceci nous permet alors de caractériser la dispersion des écarts temporels entre les stimulations dans un continuum allant d'une accroche mutuelle modérée (écarts de maximum deux secondes) à une accroche mutuelle forte (écarts tendant vers zéro).

Tout d'abord, nous ne trouvons pas de différence significative concernant la variabilité de l'accroche en fonction de la qualité de la réponse « Humain », et ce que ce soit de façon globale (tous niveaux de complexité confondus) ou de façon spécifique (dans chacun des trois niveaux de complexité) (voir Tableau 9).

Niveaux	Différence	Test	Probabilité	Taille d'effet
Tous	1.89	U = 1317 ; $Z_{ajusté}$ = 0.710311	p = 0.4775	x
1 (simples)	43.48	U = 122 ; $Z_{ajusté}$ = 1.023152	p = 0.3062	x
2 (interm.)	- 10.09	U = 142 ; $Z_{ajusté}$ = - 0.632772	p = 0.5268	x
3 (complexes)	- 27.03	U = 107 ; $Z_{ajusté}$ = - 1.74012	p = 0.0818	x

Tableau 9 : Récapitulatif des résultats statistiques concernant la variabilité de l'accroche mutuelle en fonction de la qualité de la réponse « Humain » (Bonne réponse vs Mauvaise réponse) et des niveaux de complexité des robots.

En ce qui concerne les réponses « Humain » fournies lors de la condition « Robot », nous observons que lorsque le sujet a interagi avec un programme clos, la variabilité de l'accroche est plus élevée que lorsqu'il a interagi avec un programme d'interaction (d = 99.69) (voir Figure 58). Cette différence est significative (test de Mann & Withney : U = 405 ; $Z_{ajusté}$ = 4.650911 ; p < 0.0001), et l'effet peut même être considéré comme fort (g de Hedge = 0.95). Ce résultat est retrouvé dans chacun des trois niveaux de complexité (voir Tableau 10). En outre, nous ne trouvons pas de différence significative concernant la variabilité de l'accroche en fonction des niveaux de complexité, et ce que ce soit pour les programmes clos (test de Kruskall-Wallis : $H_{(2,48)}$ = 2.968816 ; p = 0.2266), ou pour les programmes d'interaction (test de Kruskall-Wallis : $H_{(2,40)}$ = 4.616717 ; p = 0.0994).

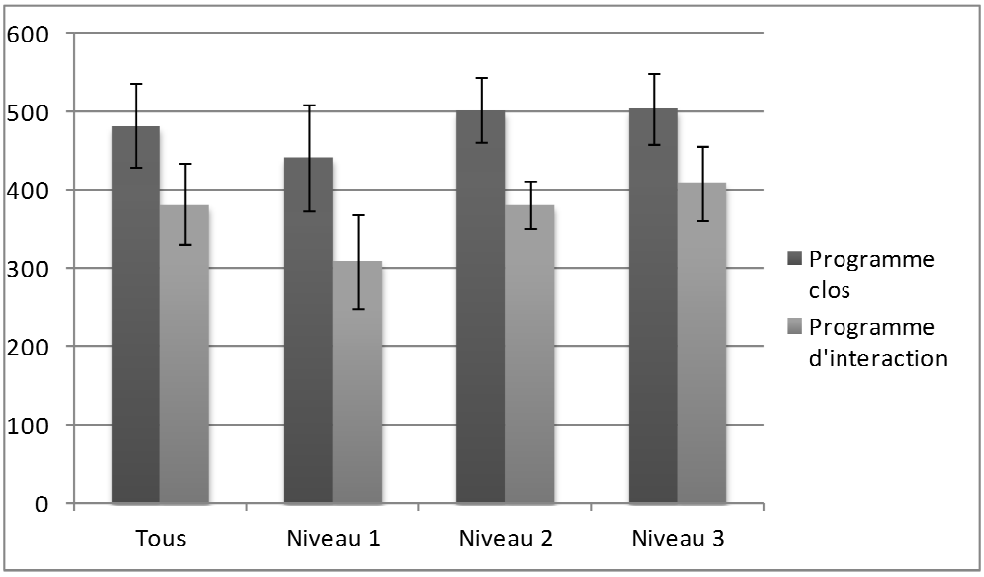


Figure 58 : Variabilité de l'accroche mutuelle lorsque le sujet répond « Humain » par erreur en fonction du type de robot (Programme clos vs Programme d'interaction) et des niveaux de complexité des robots.

Niveaux	Différence	Test	Probabilité	Taille d'effet
1 (simples)	131.74	U = 17 ; $Z_{ajusté}$ = 2.285350	p = 0.0222	g = 0.95 (fort)
2 (interm.)	120.66	U = 32 ; $Z_{ajusté}$ = 3.478505	p = 0.0005	g = 1.65 (fort)
3 (complexes)	95.73	U = 69 ; $Z_{ajusté}$ = 2.772413	p = 0.0055	g = 1.03 (fort)

Tableau 10 : Récapitulatif des résultats statistiques concernant la variabilité de l'accroche mutuelle lorsque le sujet répond « Humain » par erreur en fonction du type de robot (Programme clos vs Programme d'interaction) et des niveaux de complexité des robots.

Enfin, nous observons que globalement, la variabilité de l'accroche mutuelle est plus élevée lorsque les deux sujets se reconnaissent mutuellement que lorsque la reconnaissance est unilatérale (d = 62.54), ou lorsqu'aucun des deux ne reconnaît l'autre (d = 66.14). La différence globale est significative (test de Kruskall-Wallis : $H_{(2,50)}$ = 13.98596 ; p = 0.0009).

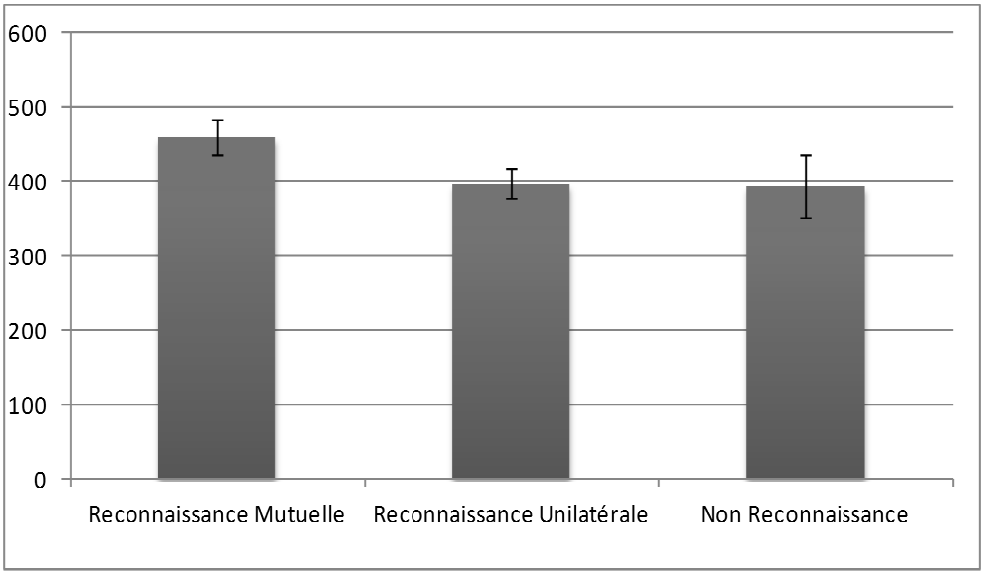


Figure 59 : Variabilité de l'accroche mutuelle dans la condition « Humain » en fonction des trois modalités collectives de réponse (Reconnaissance Mutuelle, Reconnaissance Unilatérale et Non Reconnaissance).

D'une façon plus spécifique, la reconnaissance mutuelle est statistiquement plus élevée que la reconnaissance unilatérale (test de Mann & Withney : U = 40 ; $Z_{ajusté}$ = 3.859909 ; p = 0.0001) et que la non reconnaissance (test de Mann & Withney : U = 63 ; $Z_{ajusté}$ = 2.393172 ; p = 0.0167), ces deux effets pouvant même être considérés comme forts (respectivement g = 1.46 et g = 1.02). En revanche, la reconnaissance unilatérale n'est pas statistiquement différente de la non-reconnaissance (test de Mann & Withney : U = 111 ; $Z_{ajusté}$ = 0.569803 ; p = 0.5688).

5.6.3. Désengagements de l'accroche mutuelle

Enfin, nous avons analysé les interactions de façon plus macroscopique, de façon à rendre compte de l'organisation de la succession des séquences d'interaction lors des différents essais. En particulier, nous nous sommes intéressés au nombre de « désengagements », qui signent le nombre de ruptures des séquences d'interaction à l'intérieur de chaque essai. De façon plus précise, nous considérons comme une rupture les cas où une séquence d'accroche mutuelle se termine, soit lorsque l'écart entre les stimulations augmente au-delà de deux secondes, soit parce qu'aucune autre stimulation ne suit cette séquence. Notons que nous ne comptons pas comme un désengagement les cas où une séquence d'interaction est en cours alors que l'essai prend fin.

Tous niveaux de complexité confondus, nous observons que lorsque le sujet répond « Humain » correctement, l'interaction a présenté plus de désengagements que lorsqu'il a répondu « Humain » par erreur (d = 0.65) (voir Figure 60). Cette différence est significative (test de Mann & Withney : U = 1004 ; $Z_{ajusté}$ = 2.666659 ; p = 0.0076), et l'effet peut être considéré comme modéré (g de Hedge = 0.59). Mais si cette différence est observée dans chacun des trois niveaux de complexité, elle n'est significative que pour le niveau 2 (robots intermédiaires) (voir Tableau 11).

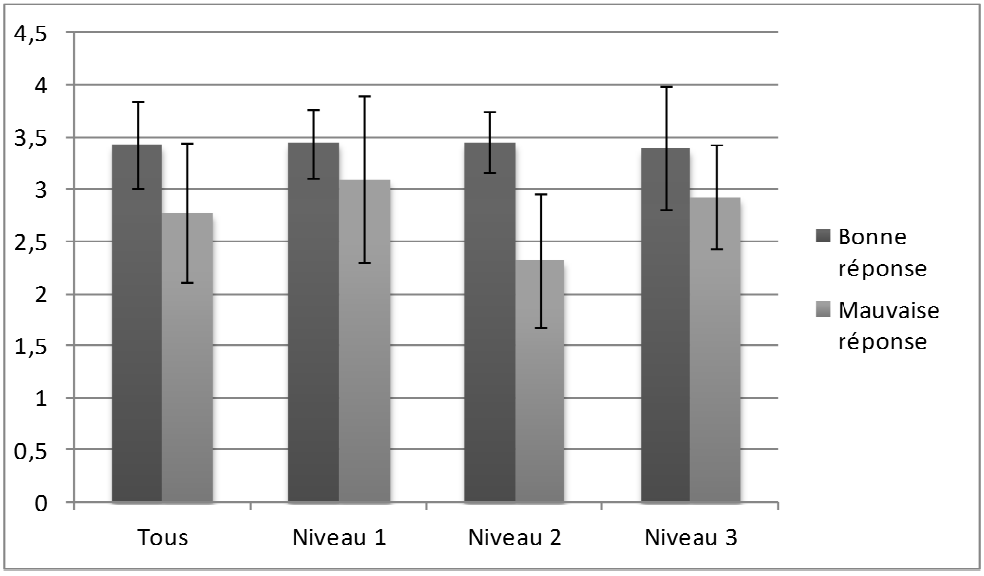


Figure 60 : Nombre moyen de désengagements en fonction de la qualité de la réponse « Humain » (Bonne réponse vs Mauvaise réponse) et des niveaux de complexité des robots.

Niveaux	Différence	Test	Probabilité	Taille d'effet
1 (simples)	0.34	U = 141.5 ; $Z_{ajusté}$ = 0.381539	p = 0.7028	x
2 (interm.)	1.14	U = 64.5 ; $Z_{ajusté}$ = 3.089140	p = 0.0020	g = 1.13 (fort)
3 (complexes)	0.47	U = 134.5 ; $Z_{ajusté}$ = 0.872931	p = 0.3827	x

Tableau 11 : Récapitulatif des résultats statistiques concernant le nombre moyen de désengagements en fonction de la qualité de la réponse « Humain » (Bonne réponse vs Mauvaise réponse) et des niveaux de complexité des robots.

En ce qui concerne la condition « Robot », nous observons que lorsque le sujet répond « Humain » alors qu'il a interagi avec un programme clos, l'interaction présente plus de désengagements que lorsqu'il a répondu « Humain » alors qu'il a interagi avec un programme d'interaction (d = 1.84) (voir Figure). Cette différence est significative (test de Mann & Withney : U = 348 ; $Z_{ajusté}$ = 5.177866 ; p < 0.0001), et l'effet peut même être considéré comme fort (g de Hedge = 1.31). Ce résultat est retrouvé dans chacun des trois niveaux de complexité (voir Tableau 12). De plus, nous constatons que le nombre de désengagements du sujet avec les programmes clos ne varie pas significativement en fonction des niveaux de complexité (test de Kruskall-Wallis : $H_{(2,48)}$ = 0.5495392 ; p = 0.7597). En revanche, le nombre de désengagements augmente à mesure que la complexité des programmes d'interaction augmente (test de Kruskall-Wallis : $H_{(2,40)}$ = 6.641211 ; p = 0.0361).

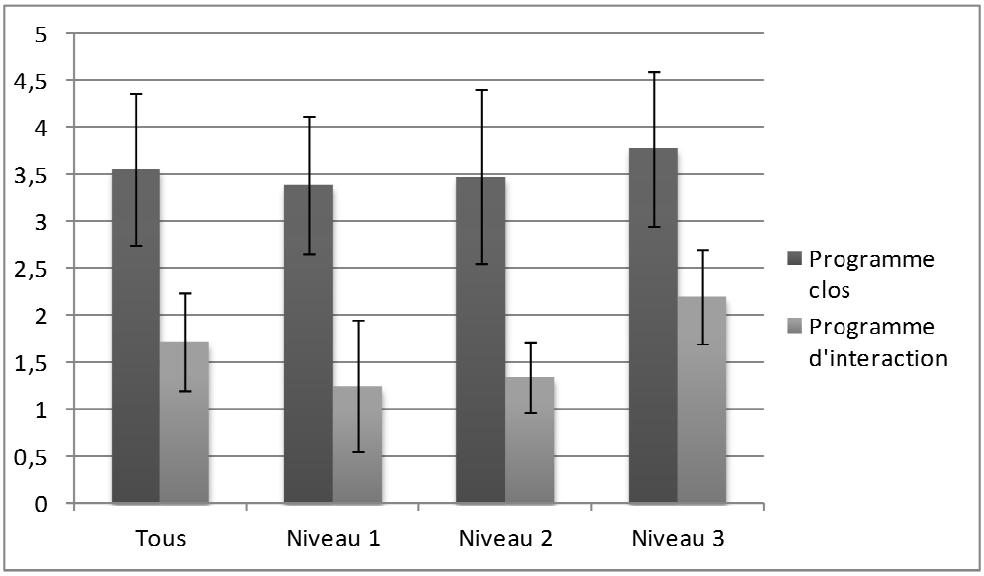


Figure 61 : Nombre moyen de désengagements lorsque le sujet répond « Humain » par erreur en fonction du type de robot (Programme clos vs Programme d'interaction) et des niveaux de complexité des robots.

Niveaux	Différence	Test	Probabilité	Taille d'effet
1 (simples)	2.14	U = 16 ; $Z_{ajusté}$ = 2.382048	p = 0.0172	g = 1.49 (fort)
2 (interm.)	2.13	U = 32.5 ; $Z_{ajusté}$ = 3.521053	p = 0.0004	g = 1.52 (fort)
3 (complexes)	1.58	U = 65 ; $Z_{ajusté}$ = 2.934591	p = 0.0033	g = 1.16 (fort)

Tableau 12 : Récapitulatif des résultats statistiques concernant le nombre moyen de désengagements lorsque le sujet répond « Humain » par erreur en fonction du type de robot (Programme clos vs Programme d'interaction) et des niveaux de complexité des robots.

Dans la condition « Humain », nous observons que lorsque les deux sujets se sont reconnus mutuellement, l'interaction a présenté plus de désengagements que lorsque la reconnaissance est unilatérale (d = 1.13), où lorsqu'aucun des deux sujets n'a reconnu l'autre (d = 0.94) (voir Figure 62). La différence globale est significative (test de Fisher : $F_{(2,47)}$ = 6.0989 ; p = 0.0044). D'un point de vue spécifique, nous constatons que la différence entre « Reconnaissance Mutuelle » et « Reconnaissance Unilatérale » d'une part, et la différence entre « Reconnaissance Mutuelle » et « Non Reconnaissance » d'autre part sont significatives (respectivement $t_{(34)}$ = 3.6677 ; p= 0.0008 et $t_{(30)}$ = 2.6045 ; p = 0.0142). Toutefois, seule la différence entre « Reconnaissance Mutuelle » et « Reconnaissance Unilatérale » peut être considérée comme notable (Pr |δ/σ| >.66 = .93). En revanche, la différence entre « Reconnaissance Unilatérale » et « Non Reconnaissance » n'est pas significative (test de Student : $t_{(30)}$ = 0.4634 ; p = 0.6464), même si l'effet ne peut pas être considéré comme négligeable (Pr |δ/σ| <.33 =.60).

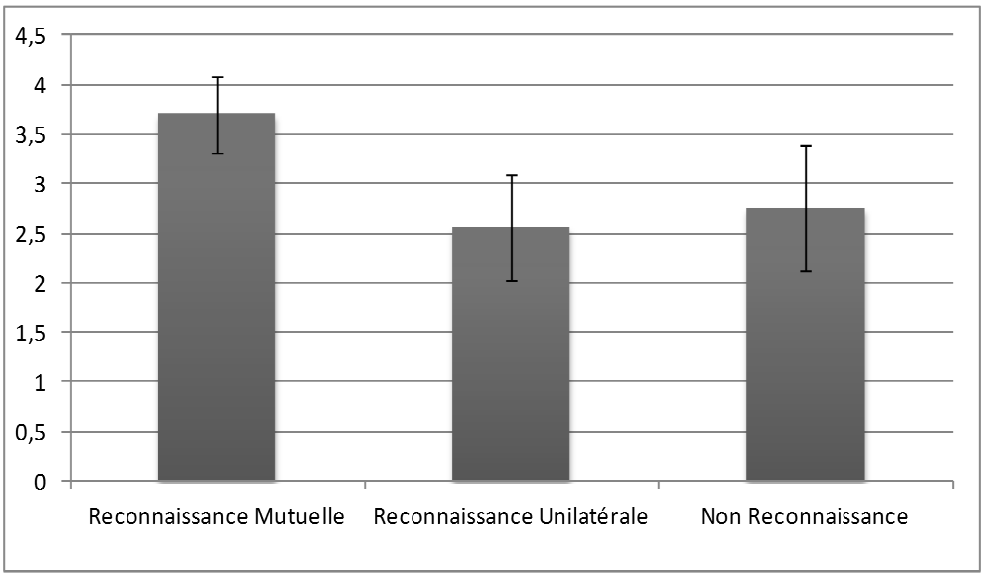


Figure 62 : Nombre moyen de désengagements dans la condition « Humain » en fonction des trois modalités collectives de réponse (Reconnaissance Mutuelle, Reconnaissance Unilatérale et Non Reconnaissance).

5.7. Discussion

Nous avions posé l'hypothèse que les interactions entre les sujets humains font émerger une dynamique particulière, cette dynamique pouvant servir de support aux participants pour la discrimination entre les humains et les robots, quel que soit leur niveau de complexité. En même temps, il était attendu que la complexité de certains robots puisse induire les sujets en erreur, dans la mesure où cette complexité pourrait les entraîner dans une telle dynamique. En particulier, nous pensions que les robots destinés à interagir de façon contingente avec les actions du sujet induiraient davantage d'erreurs. En outre, la répétition des essais devait se traduire par une stabilisation progressive de la dynamique interpersonnelle, ce qui devait avoir pour conséquence une meilleure discrimination des humains et des robots au fur et à mesure de l'expérience.

Les résultats généraux tendent à montrer que l'implémentation de variabilité et de probabilité d'arrêt dans les mouvements du robot peut suffire à tromper les sujets. En effet, la fréquence de réponses « Humain » obtenue dans la condition « Robot » augmente en fonction du niveau de complexité des robots. Mais cet effet de la complexité des robots ne se réduit pas au seul fait d'interagir de façon contingente. Si c'était le cas, nous aurions observé des fréquences de réponses « Humain » plus importantes lors des rencontres avec les robots d'interaction. Au contraire, nous observons que les robots d'interaction sont plus souvent identifiés comme des robots que les robots élaborés en termes de programmes clos. Mais ce résultat n'est significatif que pour le niveau 1 (robots simples). Plus précisément, nous observons que les robots interactifs du niveau de complexité le plus bas n'entraînent que très peu d'erreurs d'identification, contrairement aux robots interactifs des niveaux de complexité supérieure (robots intermédiaires et robots complexes).

Or, les robots interactifs simples ne présentent aucune variation rythmique. Quel que soit le déplacement du sujet, ces robots « traquent » la position de l'interactant, et inversent leur trajectoire après l'avoir croisé, et ce à un rythme constant. En ce sens, ils manifestent une contingence parfaite, au sens où leur comportement traduit une recherche rythmique de contact avec le curseur déplacé par le sujet humain. Comme nous l'avons vu, les contingences sociales jouent un rôle crucial dans la compréhension interpersonnelle (Crown & al., 2002 ; Stern, 1977). De façon très précoce, la rupture de la contingence est détectée par le bébé lors du « still face » (Tronick & al., 1979), tout comme la perturbation de celle-ci (Murray & Trevarthen, 1985 ; Nadel & al., 1999). Mais si le rythme des comportements expressifs est de prime importance, une trop grande régularité de ce rythme ne semble pas caractéristique d'un

comportement humain lors d'une interaction. Notre résultat est donc conforme aux études antérieures, qui soulignent la préférence pour les contingences élevées mais imparfaites dans un contexte social, et ce, dès la première année (Gergely & Watson, 1996 ; 1999 ; Bigelow & al., 1996). Les contingences imparfaites seraient donc caractéristiques des interactions sociales humaines (Rochat & Striano, 1999).

Pourtant, les résultats montrent que les robots destinés à interagir de façon imparfaitement contingente avec les humains ne conduisent pas à une fréquence de réponses « Humain » différente de celle obtenue avec les robots non-contingents. De plus, la fréquence de réponses « Humain » obtenue dans la condition « Humain » reste significativement supérieure à la fréquence de réponses « Humain » obtenue dans la condition « Robot », quel que soit le niveau de complexité de ces derniers. Deux hypothèses peuvent alors être avancées pour expliquer ce résultat.

D'une part, nous pourrions considérer que la variabilité implémentée dans les robots intermédiaires et complexes s'est révélée insuffisamment élevée par rapport à la variabilité qui caractérise les interactions entre sujets humains. Dans cette optique, il serait intéressant de poursuivre cette expérience, en implémentant progressivement dans des robots de complexité supérieure un plus haut degré de variabilité dans leurs oscillations. Nous nous demandons alors si les sujets humains pourront toujours faire la différence entre les humains et les robots, et jusqu'à quel point cette variabilité pourrait les tromper systématiquement, ou du moins les conduire à une indécision.

D'autre part, nous pouvons supposer que la dynamique des interactions humaines présente une organisation spécifique, qui ne se réduit pas à de simples relations contingentes, fussent-elles imparfaites. L'analyse des trajectoires perceptives apporte alors des éléments objectifs pour tenter de comprendre ce qui singularise cette organisation.

De façon minimale, nous pourrions caractériser l'interaction humaine par l'émergence d'une accroche perceptive mutuelle. C'est en effet cette accroche qui semble caractériser la reconnaissance de l'autre dans les croisements perceptifs minimalistes, qui est alors décrite comme le produit d'un attracteur de la dynamique collective (Auvray et al., 2009 ; Lenay, 2010). Dans notre cas, en effet, nous ne pouvons pas conclure à une différence de la proportion d'accroche mutuelle quand le sujet répond « Humain » entre la condition « Humain » et la condition « Robot ». Mais dans cette optique, nous devrions conclure que l'effet du niveau de complexité des robots est expliqué par une accroche perceptive de plus en plus forte.

Cette hypothèse n'est pas confirmée par l'analyse de la proportion d'accroche mutuelle lorsque le sujet répond « Humain » dans la condition « Robot ». Au contraire, nous pouvons même conclure à une diminution de l'accroche mutuelle en fonction de la complexité des robots, en tout cas pour ce qui est des programmes d'interaction. En revanche, nous ne pouvons pas conclure à une différence de proportion d'accroche en fonction des niveaux de complexité en ce qui concerne les programmes clos. Notons toutefois que dans ces erreurs de reconnaissance, la proportion d'accroche mutuelle est plus élevée lorsque le sujet a interagi avec un programme d'interaction qu'avec un programme clos. Mais comme nous l'avons vu, dans le niveau 1 (robots simples) les programmes d'interaction sont plus souvent reconnus comme des robots que les programmes clos.

Ce résultat pourrait alors être expliqué par une trop grande régularité des écarts temporels entre les stimulations dans le cas des robots interactifs simples, ce qui semble être reconnu par les sujets comme une accroche perceptive non naturelle[35]. En revanche, nous pouvons expliquer les réponses « Humain » données par erreur face aux programmes d'interaction de niveau 2 et 3 (robots intermédiaires et robots complexes), par l'effet combiné d'une proportion d'accroche mutuelle élevée avec une variabilité temporelle de cette accroche. Par conséquent, l'irrégularité de ces robots semble alors avoir entraîné les sujets dans une dynamique d'interaction qui les a induits en erreur.

En ce sens, la variabilité de l'accroche pourrait jouer un rôle crucial dans l'explication des réponses des sujets. En effet, lorsque le sujet répond « Humain », nous ne pouvons pas conclure à une différence de variabilité temporelle de l'accroche mutuelle en fonction de l'interactant (Humain vs Robot). En d'autres termes, plus le couplage présente des variabilités temporelles, plus les contingences sociales peuvent être considérées comme imparfaites, et plus le sujet répond « Humain ». Toutefois, dans les cas où le sujet répond « Humain » par erreur, la variabilité de l'accroche est significativement plus élevée lorsqu'il a interagi avec un programme clos que lorsqu'il a interagi avec un programme d'interaction. Pour comprendre ce résultat, il nous faut souligner un point essentiel. Dans le cas des programmes d'interaction, la qualité de l'accroche semble être intrinsèque à leur comportement : ils sont programmés pour produire *unilatéralement* une certaine structure d'interaction, qu'ils imposent aux sujets quels que soient les déplacements de ce dernier. Dans le cas des robots simples, cette structure est facilement identifiée parce qu'elle est trop régulière. Dans le cas des robots de complexité supérieure, cette structure est plus difficilement identifiable parce qu'elle présente des

[35] Les commentaires laissés par les sujets au cours de l'expérience vont dans ce sens. Lorsqu'ils sont confrontés aux robots interactifs simples, ils qualifient l'interaction comme étant « *mécanique* » ou « *robotique* ».

irrégularités, ce qui conduit à plus d'erreurs. Dans le niveau 1 (robots simples), ce résultat pourrait expliquer pourquoi les programmes clos sont plus souvent identifiés par erreur que les programmes d'interaction, malgré une proportion d'accroche plus faible. Comme le robot n'organise pas son oscillation autour d'une source de stimulation, la succession des rencontres avec ces robots a plus de chance de présenter des écarts temporels variables. Dans ce cas, l'accroche mutuelle et sa variabilité sont le fait unilatéral du sujet lui-même, qui cherche activement une source de stimulation et qui organise sa propre activité autour d'elle.

De cette façon, l'interaction avec un programme clos peut donner lieu à une illusion : je reçois des stimulations non régulières, soit parce que je bouge dans le champ de l'oscillation du robot, soit parce que le robot lui-même présente une oscillation irrégulière autour de moi. Je peux parfois perdre le contact un certain temps, et le retrouver, parce qu'à un moment donné j'ai bougé à l'extérieur du champ d'oscillation du robot, ou que celui-ci s'est arrêté alors que nous interagissions (robots complexes). Je suis alors entraîné dans une dynamique que je reconnais comme un engagement mutuel intentionnel. Mais je peux aussi détecter une certaine régularité dans l'oscillation du robot, qu'elle soit parfaite (robots simples) ou imparfaite (robots intermédiaires ou robots complexes), et ce, particulièrement si je m'arrête et que je porte mon attention sur la fréquence des stimulations. Je peux aussi quitter le champ d'oscillation du robot et y revenir intentionnellement, parce que j'ai repéré une certaine zone d'activité du robot. Je contrôle alors l'interaction et peux même parfois la prévoir : j'interagis avec un robot.

Que je sois dans l'illusion ou dans le contrôle, le caractère non interactif des programmes clos rend possible une alternance d'engagements et de désengagements. De cette façon, parallèlement à la qualité de la micro-structure de l'interaction (accroche perceptive et variabilité des écarts entre les stimulations faisant émerger des séquences d'interaction), la macro-structure (succession temporelle des séquences d'interaction) semble avoir également une influence sur la performance des sujets. Pris isolément, le nombre de désengagements ne permet pas à lui seul d'expliquer les réponses « Humain », du moins pour ce qui est du groupe expérimental confronté aux robots intermédiaires. En revanche, pour les robots simples et les robots complexes, nous ne pouvons pas conclure à une fréquence de désengagements différente lorsque le sujet répond « Humain ». Dans la condition « Robot », quand le sujet répond « Humain » par erreur, le nombre de désengagements est significativement plus élevé dans le cas des programmes clos que dans celui des programmes d'interaction. Les erreurs dues aux programmes clos pourraient alors être expliquées par un nombre élevé de désengagements, qui comme nous l'avons déjà mentionné, semble être plutôt le fait d'une

rupture de l'interaction de la part de l'humain. Rappelons tout de même que ces désengagements peuvent aussi être le fait d'un arrêt exécuté par un robot complexe, alors que le sujet « interagissait » avec lui. Au contraire, les programmes d'interaction, en tant qu'ils imposent une structure d'interaction qui peut être régulière (robots simples) ou non (robots intermédiaires et robots complexes), rendent le désengagement plus difficile. Pour cela, il faut que le sujet aille au moins aussi vite que le robot et dans la même direction, ou que celui-ci s'arrête dans le cas des robots complexes.

Ces différents résultats nous permettent d'avancer des hypothèses concernant les caractéristiques de l'interaction entre des sujets humains. En effet, nous avons constaté que les humains sont plus souvent identifiés comme des humains que ne le sont les robots, quelle que soit la complexité de ces derniers. L'interaction perceptive entre deux sujets humains semble donc qualitativement différente, et conduit à une meilleure identification. Conformément aux cas des réponses « Humain » dans la condition « Robot », nous pouvons alors envisager que la qualité de l'interaction repose sur une combinaison de caractéristiques microscopiques (comme la variabilité de l'accroche) et de caractéristiques macroscopiques (comme l'alternance d'engagements et de désengagements).

L'analyse des trajectoires perceptives en fonction des modalités de reconnaissance nous permet en effet d'affirmer que la reconnaissance mutuelle a lieu non seulement lorsque l'accroche perceptive mutuelle est élevée, mais surtout lorsqu'elle est variable. En effet, nous ne pouvons pas conclure à un effet de la proportion d'accroche mutuelle seule sur la reconnaissance mutuelle, tandis que nous pouvons conclure à un effet de la variabilité de cette accroche. Mais contrairement aux cas des programmes d'interaction intermédiaires et complexes, nous postulons que la variabilité de cette accroche est intrinsèque à un jeu de co-ajustements dynamiques qui résulte de l'engagement mutuel des sujets dans les séquences d'interaction (De Jaegher et Di Paolo, 2007 ; Fogel, 1993 ; Fogel & Garvey, 2007 ; Hobson, 2002 ; Trevarthen, 1993). Dans le cas d'un croisement perceptif entre deux sujets, en effet, chaque partenaire est engagé dans la régulation de son propre mouvement d'exploration en fonction des retours sensoriels qu'il reçoit. Puisque les retours sensoriels résultent de la rencontre de l'activité perceptive de l'autre, l'engagement de l'un régule l'engagement de l'autre. C'est cette co-régulation dynamique qui produit une structure temporelle propre au processus interactif, ce qui témoigne de l'émergence d'une dynamique relationnelle autonome (De Jaegher et al., 2010) La rythmicité de l'interaction sociale renvoie donc à la variabilité des formes de coordinations temporelles sur lesquelles l'interaction est construite (De Jaegher, 2006 ; De Jaegher & Di Paolo, 2007). Ainsi, la variabilité de l'accroche mutuelle ne

semble pas pouvoir être réduite à une simple production de contingences qui seraient détectées par l'autre : dans la mesure où les deux sujets sont activement engagés dans l'interaction, ils s'entraînent mutuellement dans une accroche perceptive qui émerge de leur volonté réciproque à établir un contact, à le maintenir ou à le retrouver lorsque celui-ci est perdu. Se faisant, ils participent activement à la co-régulation de l'interaction, en saisissant les opportunités d'action disponibles localement (Krueger, 2011, Lenay & Stewart, 2012).

En outre, cette co-régulation ne fait pas seulement émerger une rythmicité spécifique du couplage à un niveau microscopique, mais aussi une organisation spécifique de l'interaction à un niveau macroscopique. En effet, nous pouvons conclure à un effet du nombre de désengagements sur la reconnaissance mutuelle. La plupart du temps, les ruptures de l'interaction peuvent être interprétées par une incompréhension individuelle de la dynamique d'interaction. L'un des deux sujets a pu essayer d'entraîner son interactant dans une direction donnée, pour voir si celui-ci était capable de le suivre dans l'espace, tandis que l'autre n'a pas perçu ce déplacement. Parallèlement, il est également possible que l'un des deux sujets ait perçu à tort une dérive de l'interaction, et l'a donc anticipée dans l'espace, ce qui a conduit à une perte de contact. Mais l'engagement mutuel des deux participants dans la recherche active de stimulations rend alors possible une restauration de la dynamique, qui peut ensuite conduire à de nouveaux désengagements. C'est précisément cette capacité de réengagement qui nous semble cruciale. Les sujets ont l'expérience d'une activité perceptive intentionnelle dirigée vers leur propre activité perceptive, qui s'accroche et se désengage périodiquement. Ils peuvent alors tenter de moduler cette expérience, en provoquant intentionnellement la rupture pour tester la capacité de réengagement de l'autre. Les deux sujets entrent alors dans une stratégie active de vérification, qui leur permet de donner du sens à la dynamique d'interaction et de la comprendre non comme un engagement unilatéral (comme dans le cas des robots), mais comme le résultat d'un engagement mutuel intentionnel. Tout comme l'accroche mutuelle microscopique se stabilise en tant qu'attracteur de la dynamique collective, l'alternance de désengagements et de réengagements se stabilise en tant que structure macroscopique qui organise les séquences d'interaction et leur donne du sens. En d'autres termes, nous postulons que l'organisation macro-structurelle de l'interaction peut être progressivement re-saisie par les interactants comme un indice objectif d'un engagement mutuel.

Par ailleurs, nous pouvons également mentionner ici que cette organisation ne se réduit pas seulement à une alternance maîtrisée de rupture et de rétablissement de l'interaction. Nous avons parfois observé des séquences d'interaction très intéressantes, où les activités

perceptives des deux partenaires se sont organisées en termes de « tours de parole » (turn-taking) (voir exemple en Annexe C, section a, p. 292) : un des sujets s'arrête tandis que l'autre oscille autour de lui, puis les rôles s'inversent. La maîtrise d'une telle organisation macroscopique semble résulter d'une stratégie intentionnelle commune visant à vérifier la stabilité de l'interaction et l'engagement de son partenaire. De futures analyses devront être conduites pour analyser ces séquences, et analyser l'impact d'une telle organisation sur les réponses des sujets.

Ainsi, la rythmicité microscopique et macroscopique de l'interaction émerge d'un processus dynamique de co-ajustements mutuels qui pourrait progressivement servir de support pour la reconnaissance interpersonnelle. En effet, les sujets améliorent leur performance entre les deux blocs d'essais, même s'ils ne reçoivent aucun feedback sur la qualité de leurs réponses au fur et à mesure des essais. Nous postulons que les sujets apprennent d'abord à discriminer les humains et les robots au cours de l'expérience, puis ils reconnaissent les autres sujets humains via la stabilisation d'une structure d'interaction spécifique. Cette hypothèse est confortée par le fait que dans les derniers essais de la condition « Humain », les sujets se reconnaissent davantage de façon mutuelle que de façon unilatérale, et ce, de façon supérieure à la seule propension des sujets à répondre « Humain » quel que soit le type d'interactant. Nous pouvons donc conclure à une évolution différenciée des modalités de reconnaissance durant la tâche. L'apprentissage se traduit par une augmentation de la reconnaissance mutuelle et par une diminution de la reconnaissance unilatérale entre les deux blocs d'essais. Ainsi, la dynamique de l'interaction semble progressivement se stabiliser (Di Paolo & al., 2008). Par conséquent, elle devient de plus en plus saillante, et les individus peuvent alors la re-saisir en tant qu'opportunité d'interaction (Lenay & Stewart, 2012). Cette re-saisie peut alors conduire au déploiement de stratégies intentionnelles visant à moduler la dynamique d'interaction, et à en observer les effets sur l'engagement de l'autre. La reconnaissance interpersonnelle semble donc basée sur la dynamique de l'interaction, qui devient le support de la co-construction de sens (De Jaegher & Di Paolo, 2007 ; Fuchs & De Jaegher, 2009).

Cette étude expérimentale nous a permis d'isoler des éléments caractéristiques de l'interaction interpersonnelle dans un environnement minimaliste, réduit à une dimension. La prise en compte de ces éléments est cruciale, car ils permettent de comprendre comment les interactants organisent leurs activités perceptives, et comment cette organisation peut les conduire à faire sens de l'engagement de leur partenaire. Mais dans notre situation, les espaces numériques partagés ont vocation à héberger des contenus typiquement organisés en

deux dimensions (comme des cartes, des plans, des diagrammes etc.). Nous nous demandons alors si les utilisateurs pourront coordonner leurs activités perceptives dans un espace bidimensionnel. Plus précisément, nous nous intéressons à la dynamique de l'interaction en tant que support au guidage mutuel de deux utilisateurs. Suite aux suivis longitudinaux systématiques qui ont permis une observation sur le terrain de l'appropriation et de l'« utilisabilité » de Tactos par des utilisateurs non-voyants, il s'agit donc maintenant de mettre à l'épreuve les interactions tactiles médiatisées dans des contextes d'usage écologiques.

6. Retour sur les usages : le cas des interactions dyadiques

Après les observations préliminaires, qui ont consisté à interagir avec le sujet expert via le dispositif dans une visée exploratoire, nous avons procédé à des ***observations contrôlées***. Nous avons alors systématisé l'observation de séquences d'interaction à travers des situations prédéfinies dans lesquelles le sujet expert (« **C.** », déjà mobilisé lors des suivis longitudinaux concernant l'appropriation du dispositif Tactos et lors des analyses d'usage préliminaires, voir Partie 1, sections 1.3, p. 27 et 2.1, p. 41) a interagi avec un sujet novice (désigné « **L.** »[36]). Chaque séquence a été filmée, de façon à recueillir les verbalisations, et les trajectoires ont été enregistrées pour dégager les stratégies utilisées et obtenir les taux de contact.

Lors de ces séquences, nous avons placé nos deux utilisateurs dans des situations d'interaction dyadique au sein d'un espace bidimensionnel de 800 x 600 pixels. Nous nous sommes intéressés dans ces situations à la possibilité d'un guidage interpersonnel dans l'espace, à travers deux scénarii : un premier scénario exploratoire, visant à observer l'organisation d'une forme de guidage élémentaire avec aucun autre but que de se suivre dans un espace vide, et un scénario où le guidage est cette fois finalisé : les rôles sont définis à l'avance et l'objectif est de se guider vers une cible située dans l'espace. Pour cela, **C.** a utilisé le prototype de module d'interaction tactile, autorisé par son niveau d'expertise et par son expérience avec cet effecteur, alors que **L.** a utilisé l'effecteur tablette/stylet, plus conforme à son niveau actuel de compétence avec le dispositif Tactos[37].

[36] L. est une adulte de 36 ans atteinte de cécité complète depuis l'âge de 14 ans, suite à une atteinte orpheline du nerf optique couplée à des spasmes de convergence. Elle prend part à des analyses d'usage du dispositif depuis septembre 2011, lorsque l'APICADEV dont elle est membre devient partenaire du projet. L'APICADEV (Association PICarde des Aveugles et DEficients Visuels) vise l'autonomie des personnes déficientes visuelles en Picardie, en menant des actions visant l'accessibilité des personnes en situation de déficience visuelle parallèlement à la sensibilisation du public.

[37] Lorsque nous effectuons ces analyses d'usage, L. a bénéficié de 4 mois de pratique de Tactos, à raison de deux séances par semaine.

Lors de ces scénarii, une attention particulière a été portée sur la configuration des corps-image/corps-percevant, de façon à évaluer leurs avantages et inconvénients. Trois configurations ont été alternativement testées pour chaque scénario (voir Figure 63) :

- **Configuration « Mono »** : les corps-percevant sont constitués d'un unique champ récepteur élémentaire de 32 x 32 pixels. Les corps-image sont des surfaces carrées de 32 x 32 pixels. De cette façon, chaque pixel noir rencontré par la surface du corps-percevant entraîne la levée simultanée des 16 picots du boitier de stimulateurs tactiles.
- **Parallélisme de type I** : les deux sujets ont un corps-percevant constitué de 16 champs récepteurs élémentaires de 8 pixels chacun. La surface totale du corps-percevant est donc de 32 x 32 pixels. Les corps-image sont des surfaces carrées de 32 x 32 pixels. Lorsque chacun des champs récepteurs élémentaires rencontre au moins un pixel noir, cela active le picot correspondant du boitier de stimulateurs tactiles. Il y a donc un parallélisme entre les champs récepteurs et les stimulateurs tactiles.
- **Parallélisme de type II** : les deux sujets ont un corps-percevant constitué de 16 champs récepteurs élémentaires de 8 pixels chacun. La surface totale du corps-percevant est donc de 32 x 32 pixels. Les corps-images sont des surfaces carrées de 8 x 8 pixels situées au centre du corps-percevant de chaque sujet. Le parallélisme entre les champs récepteurs et les stimulateurs tactiles est donc identique au parallélisme de type I, mais cette fois, la surface de contact est réduite.

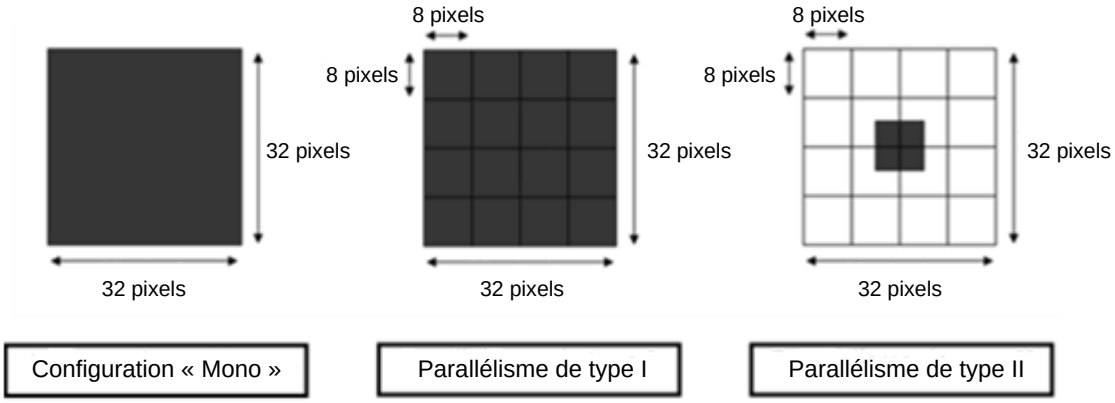


Figure 63 : Configuration des corps-image/corps-percevant utilisées lors des observations contrôlées

6.1. Scénario 1 : rencontre et suivi dans un espace vide

Dans un premier temps, nous proposons des séquences de rencontres dans un espace vide. Les seules sources de stimulations pour chaque sujet sont donc le corps-image de l'autre participant. Chaque essai dure deux minutes, et nous procédons à deux essais consécutifs pour

participant. Chaque essai dure deux minutes, et nous procédons à deux essais consécutifs pour chaque configuration de corps-image/corps-percevant. La consigne consiste simplement à essayer de se trouver, puis de se suivre en essayant de ne pas se perdre. De plus, il est demandé aux sujets d'éviter de se parler pendant les essais, sauf s'ils le jugent absolument nécessaire. Nous n'attribuons pas de rôle particulier aux deux sujets : nous voulons tout d'abord observer la rencontre et voir comment s'organise l'interaction de façon spontanée.

6.1.1. Configuration « Mono »

Dans un premier temps, la configuration « Mono » est utilisée. Spontanément, les deux sujets recherchent le contact avec l'autre en déployant des stratégies très différentes. En effet, tandis que **C.** effectue des mouvements amples et rapides à travers l'espace, **L.** peine à s'aventurer dans celui-ci. Elle effectue globalement des mouvements lents et de faible amplitude, en ne s'éloignant jamais des bords de l'espace. De cette façon, le premier contact est très difficile à établir. Lors du premier essai, 1 minute et 36 secondes se sont écoulées avant la première rencontre. Les deux sujets restent alors immobiles quelques secondes, avant que **C.** ne prenne l'initiative d'un déplacement. De cette façon, **C.** endosse spontanément le rôle de guideur, en essayant d'emmener **L.** dans une direction donnée. Il semble que son expertise avec la manipulation de Tactos lui permet de s'engager dans l'espace avec plus d'aisance. Mais cette première tentative de guidage reste infructueuse, ce qui conduit à une perte de contact. **C.** retrouve alors quasi-instantanément **L.**, en explorant systématiquement la zone du dernier contact. Après quelques tentatives, les deux utilisateurs commencent à se déplacer ensemble. Dans ce cas, le mouvement de guidage de **C.** est très lent, et s'exécute jusqu'à la perte de **L.**, avant de retourner immédiatement au contact et de réessayer. Dans ces phases, **L.** reste le plus souvent immobile, avant d'amorcer un mouvement hésitant dans une direction donnée, qui n'est pas forcément celle que tente de lui indiquer **C.** Dans ces conditions, à la fin du premier essai, le taux de contact est très faible (11%), et nous n'avons pas réellement assisté à une coordination de leurs activités perceptives. Les deux sujets se montrent alors sceptiques quant à la possibilité de se guider dans l'espace avec cette configuration (voir Verbatim 1).

C. : « *Je pense que c'est pas possible comme ça, ou alors il faut vraiment bouger doucement et avoir de la chance (…). On sait pas du tout où va l'autre (…).* »

L. : « *J'ai l'impression d'avoir passé mon temps à chercher où était* ***C.*** *(…). Quand on s'est trouvé, on s'est vite perdu et je ne savais pas où aller après.* »

Verbatim 1 : Commentaires des sujets à la fin du premier essai effectué avec la configuration « Mono »

Au début du deuxième essai, les deux utilisateurs peinent une nouvelle fois à se retrouver dans l'espace. Cette fois, ils décident de se donner une instruction verbale directe, de façon à convenir d'un point de rendez-vous explicite dans l'espace. Le premier contact a été alors favorisé, et est intervenu beaucoup plus rapidement que lors du premier essai (34 secondes). **C.** reprend alors spontanément le rôle de guideur, et tente comme lors de l'essai précédent d'emmener **L.** dans une direction donnée. Après quelques tentatives infructueuses, **C.** demande alors explicitement à **L.** de le suivre en lui indiquant la direction à suivre. Sinon, **L.** reste soit immobile, soit s'engage dans une mauvaise direction, ce qui conduit nécessairement à une perte de contact. Lorsque les sujets parviennent finalement à se déplacer ensemble dans la même direction, c'est au prix d'une anticipation du mouvement de **L.** de la part de **C.** En effet, celui-ci se déplace plus vite que **L.**, et la devance systématiquement dans l'espace. Il se déplace alors lentement dans une direction donnée et procède à de rapides contrôles vers l'arrière dès que le contact est perdu. Au final, lors de l'essai, les deux sujets sont parfois parvenus à se déplacer ensemble, mais souvent au prix d'une instruction verbale explicite. A la fin de l'essai, le taux de contact s'élève à 37%. Ici encore, il est difficile de parler de coordination des activités perceptives. La configuration « Mono » semble donc problématique. En effet, les stimulations en tout ou rien ne permettent pas de percevoir la direction du mouvement du partenaire. Il s'agit alors de choisir une direction, choix qui est conforté s'il conduit rapidement à une stimulation, mais qui présente le risque d'une rupture prolongée de contact si la direction choisie n'est pas celle du guideur. Lorsque le mouvement est amorcé, les sujets se déplacent par anticipation, le guideur devançant le suiveur : le mouvement est asynchrone, et le guidage semble s'effectuer par une « dérive » quasi-involontaire de l'interaction. Après l'essai, les avis des deux sujets sont mitigés (voir Verbatim 2).

C. : « *En fait ça paraît faisable, même si c'est pas facile (…). Faut pas aller trop vite et faut pas aller trop loin. En fait j'essayais d'avancer un peu et de revenir en arrière tout de suite pour voir si elle me suivait. Et au début elle me suivait vraiment*»

L. : « *Ben moi je suis pas très sure… La plupart du temps j'ai eu de la chance (…). C'est pas facile, moi je fais confiance à **C.** et j'essaie de le suivre, mais parfois on sait plus bien qui suit qui (…).*»

Verbatim 2 : Commentaires des sujets à la fin du deuxième essai de la configuration « Mono »

6.1.2. Parallélisme de type I

Nous procédons à la même tâche, en utilisant cette fois le parallélisme de type I. Les sujets sont informés que la configuration des corps-image/corps-percevant a changé, mais ils ne savent pas dans quelle mesure. Pour rappel, les corps-percevant sont maintenant des matrices de 16 champs récepteurs élémentaires de 8x8 pixels chacun. Les corps-images sont des surfaces carrées correspondantes de 32x32 pixels. Nous avons donc maintenant un parallélisme des champs récepteurs avec les stimulateurs tactiles.

Au début du premier essai, **L.** attend au point de rendez-vous fixé lors de l'essai précédent, pendant que **C.** « scanne » l'espace, avec des mouvements rapides et de grande amplitude. Très vite, il retrouve **L.** Immédiatement, les deux sujets réagissent au premier contact, en commentant les changements perçus (voir Verbatim 3).

C. : « *Ah ouais, ça change carrément ! Là par exemple, je sais que **L.** est à gauche.*»

L. : « *C'est clair, ça va être plus facile comme ça (…). J'arrive à lire sur les picots où est **C.** par rapport à moi.* »

Verbatim 3 : Commentaires des sujets en réaction au 1er contact lors du premier essai effectué avec le parallélisme de type I

Dès ce premier contact, l'interaction s'organise. **C.** reprend naturellement le rôle de guideur : il choisit une direction, puis amorce un mouvement lent, mais cette fois, il ne va pas jusqu'à la perte de contact. Pour faire comprendre la direction choisie à **L.**, **C.** tente immédiatement de tirer parti du parallélisme, en se plaçant juste à la limite de la perte de contact du côté où il veut emmener **L.** Cette dernière comprend correctement la direction, en détectant la disposition des picots de sa matrice de stimulateurs, et un suivi fluide et rapide s'ensuit alors. Parfois, au gré des mouvements, le contact est rompu, mais **C.** parvient à

retrouver quasi-immédiatement le contact. Dans ce cas, **L.** s'arrête, et **C.** se remet en contact statique sur le côté de **L.**, en attendant qu'elle se déplace. Lorsque le déplacement est détecté par **C.**, le suivi reprend. Ici, les déplacements sont mutuels et synchronisés. Même les changements de directions opérés par le guideur n'entraînent pas de perte de contact prolongée. Au fur et à mesure de l'interaction, la vitesse de déplacement semble augmenter, sans que cela ne perturbe le guidage. Lors du deuxième essai, **L.** prend même le rôle de guideur, et ce, sans en avoir averti **C.** Elle s'arrête, attend que **C.** se place en contact puis se déplace lentement dans une direction. Après quelques secondes d'hésitations, **C.** semble comprendre que les rôles sont inversés et se met à suivre **L.** dans l'espace, sans perte de contact, mais avec toutefois une vitesse moindre.

Avec cette configuration, la qualité du guidage est très nettement supérieure à celle que nous avons observée lors de la configuration « Mono ». Les sujets sont parvenus à se suivre de façon très fluide, en procédant à des déplacements mutuels et synchronisés, témoignant d'une coordination efficace des activités perceptives. Les directions sont maintenant clairement perceptibles via les matrices de stimulateurs, et cette information est correctement utilisée pour garantir un guidage efficace. Lors des deux essais, quelques pertes de contact sont survenues, mais celles-ci ont rapidement conduit à une reprise de l'interaction. De même, quelques verbalisations concomitantes ont eu lieu, mais celles-ci n'ont pas consisté en des instructions verbales. Il s'agissait plutôt de constatations, qui accompagnaient un guidage en cours déjà effectif. Au final, pour l'essai 1, le taux de contact s'est élevé à 78%, tandis que pour l'essai 2, celui-ci a atteint les 81%. Lorsque le deuxième essai s'est arrêté, les deux sujets se sont avérés satisfaits de leur interaction (voir Verbatim 4).

C. : « *Ben oui, là maintenant, c'est beaucoup plus simple (…). Même quand on est pas sûr, on se retrouve tout de suite et dans la bonne direction (…). J'ai aussi trouvé plus simple de suivre* ***L.*** *que d'être guideur, parce que là, il y a juste à suivre les stimulations, et on sait où elle va…*»

L. : « *C'est plus facile. Mais je préfère passer devant en fait, parce que comme* ***C.*** *est toujours là, je me pose pas de question et j'avance.* »

Verbatim 4 : Commentaires des deux sujets à la fin du deuxième essai effectué avec le parallélisme de type I

6.1.3. Parallélisme de type II

Nous passons maintenant au parallélisme de type II, toujours avec la même tâche, et toujours sans informer les sujets de la nature du changement. Pour rappel, cette configuration permet aussi un parallélisme des champs récepteurs avec les stimulateurs tactiles, mais cette fois, le corps-image est plus petit que le corps-percevant. Le corps-image est une surface carrée de 8x8 pixels situé au centre d'une matrice de 16 champs récepteurs élémentaires de 8x8 pixels chacun.

Avec cette configuration, le début du premier essai est laborieux. Alors que **L.** attend **C.** au point de rendez-vous habituel, **C.** ne parvient pas à la rencontrer. En effet, le corps-image étant cette fois plus petit que lors de la configuration 2, la surface de contact est réduite. Après 30 secondes de recherche infructueuse, **L.** rappelle verbalement le lieu du rendez-vous à **C.** Après une exploration plus systématique de ce lieu de rendez-vous, le premier contact survient alors. Comme lors de la configuration précédente, ce premier contact entraîne immédiatement une constatation de la part des deux sujets (voir Verbatim 5).

C. : « *Ok, j'ai compris ce que tu as fait (…), c'est beaucoup plus sensible… euh… enfin, c'est précis en fait (…), mais on a du mal à se trouver.* »

L. : « *On est tout petit (rires), il va falloir être concentré pour se suivre (…)* »

Verbatim 5 : Commentaires des sujets en réaction au 1er contact du premier essai effectué avec le parallélisme de type II

Très vite, **C.** reprend alors le rôle de guideur, et amorce un mouvement dans une direction donnée, qui est immédiatement suivi par **L.** Comme avec le parallélisme de type I, le suivi est fluide et rapide, et les déplacements sont synchronisés. **C.** parvient à guider **L.** dans l'espace sans rompre le contact, en utilisant la même stratégie que précédemment : **C.** devance légèrement **L.**, se place sur le côté de la direction souhaitée et se déplace à mesure que **L.** avance dans cette direction. La différence est que cette fois, avec la réduction de la taille du corps-image par rapport à la taille du corps-percevant, le guidage est plus fin et suit un nombre de direction plus important. Par exemple, pendant une phase de guidage lors de l'essai, **C.** emmène **L.** vers la droite, mais se décale en même temps et involontairement vers le bas. Comme **L.** cherche à maintenir la stimulation au centre de sa matrice de champs récepteurs, le mouvement des deux sujets suit alors une diagonale, ce qui fait réagir les deux sujets (voir Verbatim 6).

C. : « *C'est un peu plus dur, mais c'est plus précis. On peut maintenant se guider dans toutes les directions.* »

L. : « *Oui mais c'est vraiment dur, j'ai l'impression qu'on va se perdre d'un moment à l'autre.* »

Verbatim 6 : Commentaires des deux sujets lors du premier essai effectué avec le parallélisme de type II

Malgré ces inquiétudes, le suivi est très efficace : les deux sujets ne se perdent quasiment jamais, et les changements de direction ne posent aucun problème. L'interaction semble plus fluide. En outre, les deux sujets alternent leurs rôles sans se prévenir, et sans que cela ne nuise au suivi mutuel. Pour cela, le suiveur amorce un changement de direction pendant le guidage, ce qui entraîne une perte de contact, puis une alternance de ruptures de l'interaction et de réengagements. Si l'initiative est maintenue par le suiveur, le guideur se retrouve à suivre ce premier, ce qui conforte son initiative et lui fait accélérer la cadence. A la fin du premier essai, le taux de contact ne dépasse pas les 67%, ce qui est expliqué par la difficulté initiale de la rencontre. Mais à la fin de l'essai 2, le taux de contact s'élève à 85%.

Si la rencontre initiale a semblé plus difficile, de par la réduction de la surface de contact, le guidage en lui-même n'en a pas été moins efficace. En effet, les déplacements sont là encore mutuels et synchronisés, et n'entraînent que très peu de pertes de contact. De plus, les sujets sont parvenus à alterner leur rôle sans rompre l'interaction, avec une fluidité et une vitesse remarquable. Il semble que la réduction de la surface de contact est un atout lorsque l'interaction est amorcée, car elle permet plus de souplesse lors des changements de direction. Ainsi, les deux utilisateurs parviennent aisément à coordonner leurs activités perceptives de façon à se guider dans l'espace. A la fin du deuxième essai, les sujets ont exprimé une réticence initiale pour cette configuration, mais ils admettent une précision plus fine (voir Verbatim 7).

L. : « *Au début, on sait pas comment on va faire, on a moins d'informations. Mais en fait on oublie assez vite et on s'y fait (…) J'arrive à le suivre assez facilement sans vraiment savoir comment…* »

C. : « *C'est super précis comme ça. Je trouve que pour se guider c'est vraiment bien (…), mais faut quand même se trouver et là on a eu du mal.* »

Verbatim 7 : Commentaires des deux sujets à la fin du deuxième essai effectué avec le parallélisme de type II

6.2. Scénario 2 : rencontre et suivi finalisé

Cette fois, nous proposons une tâche plus précise à nos deux sujets, en leur attribuant explicitement et alternativement les rôles de guideur et de suiveur pour chaque essai. Le guideur est informé qu'un objet sonore est disposé dans l'espace. Cet objet sonore a une dimension de 32x32 pixels. Lorsqu'il le rencontre, celui-ci émet un son que lui seul entend. L'objectif du guideur consiste alors à localiser le son, puis à rechercher le suiveur pour l'emmener vers le son. Nous procédons à deux essais de 2 minutes et 30 secondes, toujours dans un espace bidimensionnel de 800 x 600 pixels. Notons que l'objet sonore occupe une place différente pour chaque essai. La tâche est réussie lorsque les deux sujets sont en contact du son à la fin de l'essai, ou lorsque les deux sujets restent en contact du son pendant 20 secondes durant l'essai. Comme lors de notre premier scénario, les trois configurations sont successivement testées.

6.2.1. Configuration « Mono »

Pour le premier essai de la configuration « Mono », **L.** se voit attribuer le rôle de guideur. Comme lors du premier scénario, **L.** peine à s'éloigner des bords de l'espace, ce qui rend la recherche de l'objet sonore laborieuse. Après 1 minute de 37 secondes, elle finit par le rencontrer, puis reste quelques secondes en contact avec lui. Elle part ensuite au point de rendez-vous expérimenté lors du premier scénario, mais ne rencontre pas de stimulation. **L.** fait alors appel explicitement à **C.** pour établir le contact. Après 1 minute et 54 secondes, le premier contact survient enfin. **L.** part alors en direction de l'objet avec un mouvement lent et continu, sans se soucier si **C.** la suit réellement. Ce dernier parvient à la suivre tant bien que mal, mais au prix de ruptures de contact incessantes. A chaque perte de contact, **C.** procède alors à des mouvements circulaires avec une amplitude croissante dans la zone du dernier contact, de façon à localiser **L.** Les deux sujets progressent alors très lentement. Lorsque l'essai s'arrête, l'objet sonore n'a pas été retrouvé, et le taux de contact entre les deux sujets est très faible (17%). Celui-ci reflète à la fois la difficulté initiale pour **L.** à localiser l'objet sonore, mais aussi la difficulté des deux sujets à se suivre dans l'espace.

Pour le deuxième essai, les rôles sont inversés. De par sa stratégie efficace d'exploration de l'espace, **C.** parvient à localiser l'objet sonore très rapidement (23 secondes). Il retrouve alors **L.** au point de rendez-vous (premier contact à 35 secondes). Comme lors du premier essai, le suivi dans l'espace est très difficile, malgré toute l'attention que porte **C.** à la position et au mouvement de **L.** En effet, comme lors du premier scénario, **C.** maintient un contact,

puis avance doucement dans une direction jusqu'à la perte de contact, puis revient immédiatement au contact. Malgré toutes ces précautions, **C.** ne parvient pas à emmener **L.** dans la bonne direction. En effet, lorsque **L.** s'engage dans une autre direction, elle continue de recevoir les stimulations qui la confortent dans ce choix, alors que **C.** tente de l'emmener ailleurs sans rompre définitivement le contact. De cette façon, l'interaction dérive dans le sens de l'anticipation du suiveur plutôt que dans celui du mouvement du guideur. De façon à tenter de corriger cette trajectoire, **C.** tente alors d'éviter le contact en restant très proche de **L.** dans la direction souhaitée. Cette stratégie amène alors **L.** à s'arrêter, du fait de l'absence de stimulations, et de tenter d'autres directions pour retrouver **C.** Après quelques essais, **L.** part dans la bonne direction et **C.** tente de la précéder de près. Lorsque les deux sujets arrivent dans la zone de l'objet sonore, **C.** se rend compte que sa trajectoire de guidage n'est pas optimale, et qu'il risque de manquer l'objet sonore. Il effectue alors de rapides allers-retours entre l'objet et **L.** de façon à tenter de corriger la trajectoire de **L.** pour aller plus directement vers l'objectif. Mais puisque **L.** continue dans la mauvaise direction, il tente d'appliquer la même stratégie qu'au début de l'essai. Cette fois, **L.** semble interpréter l'absence de stimulations comme une perte définitive, et fait machine arrière pour tenter de retrouver **C.** Au final, même si le guideur n'est pas parvenu à emmener le suiveur jusqu'à l'objet, les deux sujets sont parvenus à se suivre sur une plus longue distance. Le taux de contact est par conséquent plus élevé (51%). Il semble que l'expertise de **C.** lui permette de varier ses stratégies et de s'adapter à la stratégie de **L.** De cette façon, si les activités perceptives semblent coordonnées, c'est au prix de gros efforts de la part de **C.**, qui déploie beaucoup d'énergie pour tenter de maintenir l'interaction. A la fin de ce deuxième essai, les deux utilisateurs réaffirment les difficultés liées à cette configuration (voir Verbatim 8).

C. : « *Ben c'est pareil qu'avant, ça me paraît tendu sans la direction (…). J'essaie des trucs, mais ça marche pas tout le temps* »

L. : « *J'avais quand même l'impression de mieux y arriver, mais c'était peut-être de la chance.* »

Verbatim 8 : Commentaires des deux sujets à la fin du deuxième essai effectué avec la configuration « Mono »

6.2.2. Parallélisme de type I

Pour ce premier essai dans cette configuration, **C.** obtient le rôle de guideur. Cette fois encore, **C.** localise l'objet sonore très rapidement (20 secondes), et retrouve **L.** en haut à

gauche de l'écran (25 secondes). Dès le premier contact, **C.** reste sur la droite de **L.**, qui commence à se déplacer dans cette direction. **C.** devance alors doucement **L.** en se déplaçant vers la droite tout en maintenant **L.** sur sa gauche, et ce, en ne perdant quasiment jamais le contact. Le suivi est fluide et rapide jusqu'au son. Arrivé à celui-ci, **C.** s'arrête et attend quelques secondes au contact de **L.** et du son. Puis il entreprend une oscillation verticale rapide au contact de **L.**, pour essayer de lui faire comprendre qu'ils sont arrivés. Après 20 secondes de contact, l'essai s'arrête. Le taux de contact est élevé (85%), et les deux sujets semblent être satisfaits de l'essai (voir Verbatim 9).

C. : « *Pas de problème pour celui-là, quand on sent les directions (…) Mais après, arrivé au son, je savais pas trop quoi faire, à part rester là (…). [à **L.**] : T'as compris quand j'ai bougé sur le son ?*»

L. : « *Pas vraiment, mais bon, j'avais déjà compris quand tu t'étais arrêté, je me suis dit que c'était bon.* »

Verbatim 9 : Commentaires des deux sujets à la fin du premier essai effectué avec le parallélisme de type I

Pour le deuxième essai, **L.** se voir attribuer le rôle de guideur. Cette fois encore, elle rencontre des difficultés pour localiser le son. L'exploration de l'espace se limite aux bords de l'écran, même si après quelques secondes, elle parvient à s'en détacher légèrement. L'objet sonore n'est trouvé qu'après 1 minute et 10 secondes. Elle retrouve **C.** rapidement en haut à gauche (1 minute et 18 secondes), et dès le premier contact, elle amorce un déplacement continu vers le bas. Contrairement à **C.**, elle n'attend pas spécialement le mouvement du suiveur pour continuer, ce qui oblige **C.** à la rechercher autour de lui lorsque le contact est rompu. **L.** parvient quand même à emmener **C.** vers le bas, même si plusieurs pertes de contact sont survenues. L'objet sonore est rencontré par **L.** in extremis avant la fin de l'essai, alors que **C.** la suivait de près. Le taux de contact s'élève à 48%. Ici, le guidage a reposé essentiellement sur l'expertise de **C.**, qui est capable de retrouver rapidement **L.** dans l'espace lorsque le contact est rompu (voir Verbatim 10).

C. (à **L.**) : « *C'était encore loin ?* »

L. (à **C.**) : « *Non, non, on y était ! Mais j'ai eu encore du mal à le trouver (rires).* »

C. (à **L.**) : « *En tous cas, je te suivais bien (…) J'ai cru t'avoir perdu un moment, mais en fait non, tu avais continué tout droit…* »

L. (à **C.**) : « *Ah ouais ? Pourtant je te sentais toujours là…* »

Verbatim 10 : Commentaires des deux sujets à la fin du deuxième essai effectué avec le parallélisme de type I

6.2.3. Parallélisme de type II

Nous changeons une dernière fois de configuration. Pour le premier essai, **C.** est guideur. Cette fois, l'objet est plus difficile à trouver pour **C.** Après avoir exploré les quatre coins puis avoir effectué des mouvements de grande amplitude au centre de l'espace, il reprend l'exploration en partant du coin supérieur gauche et longe les bords de l'espace avec un mouvement sinusoïdal d'amplitude moyenne. L'objet est trouvé en 34 secondes, et **L.** est retrouvée rapidement après (42 secondes). Le suivi commence, et les mouvements sont fluides. Il semble que **C.** anticipe les déplacements de **L.** de façon à être toujours en contact, mais uniquement du côté de la direction souhaitée. Jusqu'alors, il emmène **L.** vers la droite. A mi-parcours, il part rechercher l'objet, puis retrouve **L.** qui s'était arrêtée en l'absence de contact. Les deux sujets reprennent leur parcours, mais **C.** se place un peu plus bas que **L.** de façon à corriger sa trajectoire. Il emmène alors **L.** au niveau de l'objet, puis effectue une oscillation verticale rapide, ce qui provoque le rire de **L.** 20 secondes après, l'essai s'arrête. Le taux de contact atteint 64% (voir Verbatim 11).

C. : « *Cette fois c'est moi qui ai eu du mal… Mais on s'en est bien sorti.* »

L. (à **C.**) : « *…et j'ai bien compris quand on est arrivé !* »

Verbatim 11 : Commentaires des deux sujets à la fin du premier essai effectué avec el parallélisme de type II

Pour le deuxième essai, **L.** est guideur. Elle commence à explorer l'espace, toujours pas très loin du bord, mais cette fois en procédant à des oscillations de faible amplitude de façon à couvrir un espace plus étendu qu'auparavant. De plus, sa vitesse de recherche est plus élevée. L'objet est trouvé en 37 secondes. **L.** retourne en haut à gauche mais ne trouve pas tout de suite **C.** Elle explore alors plus systématiquement le coin puis rencontre finalement **C.** (57 secondes). Immédiatement, la rencontre est suivie par un départ de **L.** vers le bas, suivi de près par **C.** Comme lors des essais précédents où **L.** était guideur, il semble que la qualité du

suivi repose essentiellement sur la capacité de **C.** à ne pas la perdre, plutôt que sur celle de **L.** à le guider efficacement. En effet, **L.** se contente d'avancer lentement mais sûrement vers l'objectif, même lorsque le contact est perdu. Quoi qu'il en soit, l'objet sonore est rencontré par **L.**, et celle-ci s'y arrête. **C.** bouge alors autour d'elle, tout en gardant le contact, puis finit par s'immobiliser au contact de l'objet. Lorsque les 2 minutes et 30 secondes allouées à l'essai sont écoulées, les deux sujets sont au contact de l'objet sonore. Le taux de contact s'élève à 73% (voir Verbatim 12).

L. : « *Pour une fois, j'ai trouvé le son facilement ! Après pour le suivi je sais pas, mais on y est quand même arrivés.* »

C. (à **L.**) : « *Oui je me suis dit aussi que si tu ne bougeais plus, c'est qu'on y était… J'aime bien comme ça, avec les picots comme ça (…) je trouve ça vraiment plus facile pour savoir où va l'autre…* »

Verbatim 12 : Commentaires des sujets à la fin du deuxième essai effectué avec le parallélisme de type II

6.3. Synthèse des observations contrôlées

Comme nous l'avons vu, la configuration « Mono » a posé de gros problèmes de guidage. En effet, cette configuration ne permet pas aux sujets de détecter la direction du mouvement de l'autre, ce qui complique la coordination des activités perceptives dans un environnement bidimensionnel. Par conséquent, les sujets sont obligés d'anticiper une direction et d'attendre que cette anticipation soit confortée par une stimulation. Lorsque cette anticipation est satisfaite, la direction du mouvement effectué rend possible un suivi, mais celui-ci se présente comme une alternance de contacts et de pertes de contact. Ici, cette alternance se présente comme une contrainte sur la dynamique d'interaction, dans la mesure où la tâche consiste pour les deux sujets à se synchroniser pour se déplacer dans la même direction. Le guideur doit alors devancer sans trop s'éloigner, sous peine de contrarier l'anticipation du suiveur et de le perdre définitivement. Mais il est à noter que lorsque les deux sujets parviennent à synchroniser leurs déplacements, le suiveur peut alors s'engager dans la direction du guideur via la maîtrise progressive d'une structure d'anticipation. Ainsi, si la dynamique d'interaction peine à se mettre en place, elle peut ponctuellement résulter en un guidage efficace.

Par ailleurs, les changements de direction du guideur se sont également révélés problématiques. Lors des phases de suivi, le suiveur est entraîné dans une dynamique

d'interaction qui conforte son engagement dans l'espace. Lorsque le guideur tente d'emmener le suiveur dans une nouvelle direction, ses tentatives ne font que maintenir la dynamique d'interaction dans le sens de son anticipation initiale. L'interaction dérive alors, les stimulations sont de plus en plus espacées, et conduisent à la perte de contact. Avec cette configuration, les verbalisations concomitantes ont été souvent nécessaires pour soutenir le guidage dans l'espace : le guideur donne alors des instructions explicites au suiveur, de façon à lui indiquer la direction en cours ou le changement de direction.

De leur côté, le parallélisme de type I et II ont été plus convaincants. Dans le cas du parallélisme des champs récepteurs avec les stimulateurs tactiles, les sujets ont pu facilement détecter la direction du mouvement de l'autre, ce qui s'est traduit par un suivi efficace et un guidage fluide. Les taux de contact sont alors plus élevés, les deux sujets parvenant à se déplacer ensemble tout en restant en contact. La dynamique d'interaction est alors soutenue par les informations spatialisées des stimulateurs tactiles, qui permettent une détection de la position de l'autre par rapport à sa propre position. L'anticipation du mouvement du guideur est alors facilitée. Par ailleurs, les sujets ne ressentent plus la nécessité de se donner des instructions verbales. De plus, les changements de direction posent moins de problème. Dans le cas du parallélisme de type II, ceux-ci sont même plus précis, le guideur pouvant emmener le suiveur dans toutes les directions possibles. En revanche, la surface de contact étant réduite, la rencontre initiale est parfois problématique.

Une solution possible pour optimiser le guidage dans les espaces numériques partagés serait de donner la possibilité aux utilisateurs de modifier la taille de leur corps-image au cours des interactions. En effet, lorsque le corps-image est de même taille que le corps-percevant, la rencontre dans l'espace est favorisée, de par la surface de contact accrue. Dans cette optique, il serait intéressant de pouvoir augmenter la taille du corps-image/corps-percevant au-delà de la taille prévue initialement, de façon à localiser l'autre participant plus facilement. En effet, des travaux antérieurs ont déjà démontré l'intérêt d'implémenter un « zoom haptique » pour rendre possible une localisation fine des éléments présents dans l'espace (Ziat et al., 2006). Ce zoom pourrait être graduel ou dynamique, et serait activé par l'utilisateur en pressant un bouton prévu à cet effet (Ziat et al., 2007). En revanche, une fois que le contact est établi, la réduction de la taille du corps-image semble optimiser le guidage, en rendant possible des changements de directions plus précis. Ainsi, parallèlement à une modification de la taille globale du couple corps-image/corps-percevant, une possibilité de modifier la taille du corps-image indépendamment de la taille du corps-percevant pourrait favoriser les interactions tactiles, et notamment le guidage au sein des espaces partagés. Pour

autant, il s’agit de ne pas trop compliquer les interactions par l’intermédiaire d’une multiplication des commandes-utilisateurs.

En effet, comme nous l’avons vu au travers de ces deux scénarii d’usage, le niveau d’expertise semble déterminant pour une utilisation optimale du dispositif. Le sujet expert a fait preuve d’une flexibilité au niveau de ses stratégies d’exploration, au contraire du sujet novice. La maîtrise du dispositif par le sujet expert semble alors lui fournir une liberté d’action qui ne nuit pas à la qualité de son exploration et de son interaction avec autrui.

Dans cette partie, nous avons étudié le processus d’interaction dans des situations strictement dyadiques. Nous en retenons que la coordination interpersonnelle fait émerger une dynamique d’interaction spécifique, qui sert de support à la compréhension de l’engagement de son partenaire, et qui organise l’engagement du sujet. Dans la partie suivante, nous souhaitons explorer l’émergence de cette dynamique et de sa signification lorsque celle-ci a lieu dans un monde commun constitué d’objets partagés.

PARTIE 3 : LA PERCEPTION MUTUELLE

1. Introduction

Comme nous l'avons vu, le dispositif Intertact permet aux utilisateurs de rencontrer tactilement d'autres participants à travers le réseau. Mais les espaces numériques partagés que nous souhaitons élaborer ne permettent pas uniquement des rencontres interpersonnelles et une coordination des activités perceptives dans un contexte dyadique. Des objets partagés y sont également et crucialement inscrits, et c'est autour d'eux que doivent pouvoir s'organiser des interactions signifiantes. En effet, un des objectifs de ce dispositif consiste à rendre possible l'exploration mutuelle de contenus numériques, de façon à permettre la collaboration d'utilisateurs distants. Celle-ci suppose une coordination interpersonnelle des activités perceptives individuelles autour d'objets communs. Les utilisateurs doivent donc être en mesure de percevoir les autres participants, et en même temps percevoir les objets qui occupent les espaces numériques partagés. Mais de façon plus capitale, il est nécessaire que les participants puissent reconnaître le fait que les objets sont partagés, et qu'ils peuvent donc être perçus par les autres utilisateurs. Ceci semble poser un problème : d'un point de vue individuel, « regarder » un objet nécessite de placer son corps-percevant sur l'objet perçu, de façon à activer les stimulateurs tactiles. Dans le contexte d'un espace d'interaction tactile, « regarder » simultanément un même objet suppose de la part de deux observateurs une coïncidence spatiale de leurs positions respectives avec la position de l'objet perçu. Mais dans ce cas, les deux utilisateurs reçoivent simultanément : i) une stimulation tactile résultant de la rencontre de leur corps-percevant avec l'objet, et ii) une stimulation tactile résultant de la rencontre de leur corps-percevant avec le corps-image de l'autre participant. Or, les deux types de rencontres conduisant à des stimulations tactiles identiques, chaque participant reçoit alors une stimulation tactile indifférenciée. Le corps-image de l'autre participant est alors nécessairement confondu avec l'objet perçu. Une solution technique consisterait à paramétrer le système de façon à délivrer des informations tactiles différenciées (par exemple, utiliser des plages de stimulateurs tactiles différentes, ou en décidant que certains stimulateurs s'activent lors de la rencontre d'objets, et que les autres s'activent lors de la rencontre d'autres participants).

Mais sans cet artefact technique supplémentaire, la perception mutuelle d'un même objet semble toutefois possible. En effet, les deux utilisateurs peuvent également aller et venir

autour de l'objet, de façon à rencontrer successivement deux sources de stimulation distinctes qu'il reste à discriminer : le croisement perceptif résultant de la rencontre des corps-percevant avec les corps-image respectifs des deux participants, et la rencontre du corps-percevant de chacun avec l'objet. A l'instar des rencontres interpersonnelles précédemment étudiées, nous postulons que la perception mutuelle d'un objet peut se réaliser au travers de la dynamique collective des activités perceptives. Nous nous intéressons donc aux mécanismes rendant possible la perception mutuelle d'un objet, et qui signent la reconnaissance interindividuelle qu'un objet du monde est partagé.

2. Position du problème

Dans la littérature, les interactions sociales impliquant l'individu, autrui et un objet partagé sont classiquement étudiées à travers le phénomène de ***l'attention conjointe*** (pour des revues, voir Bakeman & Adamson, 1984 ; Eilan, Hoerl, McCormack & Roessler, 2005 ; Moore & Dunham, 1995). D'une façon générale, l'attention conjointe fait référence à la capacité d'un sujet donné à coordonner son attention avec un partenaire social, de façon à partager avec lui un point de référence, que ce soit un objet, un évènement ou un autre sujet (Mundy & Newell, 2007). Cette coordination prend place dans une relation triangulaire entre deux sujets et un objet, où autrui n'est plus la cible du focus attentionnel, mais un point de passage vers une autre cible (Conein, 1998).

Une des situations exemplaires du phénomène d'attention conjointe est le suivi du regard d'autrui : l'objet de l'attention n'est plus le regard de l'autre, comme dans le cas des interactions en face-à-face impliquant un regard mutuel (Argyle & Cook, 1976), mais la cible du regard de l'autre agent : on parle de regard déictique (Gomez, 1991). Si le suivi du regard est parfois suffisant en soi, la coordination des focus attentionnels peut également reposer sur des gestes déictiques, comme le pointage (Bates, Camaioni & Volterra, 1975). Le pointage est une activité complexe, qui est considérée parfois comme le meilleur indicateur de la compréhension sociale chez l'enfant (ibid.). Il présente plusieurs fonctions, selon le contexte dans lequel il survient. Le pointage peut être proto-impératif, et dans ce cas, autrui est instrumentalisé par l'enfant de façon à atteindre un objet ou à produire une action spécifique. Mais c'est dans sa fonction proto-déclarative qu'il est le plus représentatif de l'attention conjointe. Le pointage est alors utilisé pour diriger l'attention d'autrui de façon à partager avec lui un point de référence (Bates et al., 1975 ; Liszkowski, Carpenter, Striano & Tomasello, 2006). Le fait que l'enfant accomplisse de tels actes proto-déclaratifs a soulevé

beaucoup d'intérêt, parce que la communication gestuelle serait un marqueur essentiel d'une théorie de l'esprit implicite (Bretherton, 1991), et permettrait de soutenir le processus de référenciation sociale dans le développement de l'enfant (Campos & Sternberg, 1981). Par exemple, son absence dans le cours du développement pourrait être prédictive de troubles infantiles sévères comme l'autisme (voir par exemple Carpenter, Pennington & Rogers, 2002).

D'une façon générale, il est communément admis que l'attention conjointe est une étape cruciale du développement social (Bakeman & Adamson, 1984 ; Mundy & Crowson, 1997), et un précurseur nécessaire pour le développement d'habiletés sociales complexes comme le langage (voir par exemple Morales, Mundy, Delgado, Yale, Messinger, Neal, & Schwartz, 2000 ; Bruner, 1983). Mais si l'importance de l'attention conjointe est reconnue, il n'en reste pas moins que ce phénomène soulève de nombreux débats, que ce soit en ce qui concerne l'âge d'apparition de cette compétence dans le développement de l'enfant (voir par exemple Bakeman & Adamson, 1984 ; Butterworth & Jarrett, 1991 ; Carpenter, Nagell & Tomasello, 1998 ; D'entremont, Hains & Muir, 1997 ; Moore & Dunham, 1995), ou plus fondamentalement en ce qui concerne la nature de celle-ci et de ses implications sur la compréhension des situations sociales (Butterworth, 1991 ; Racine & Carpendale, 2007 ; Tomasello, 1995).

De façon classique, l'attention conjointe requiert en premier lieu la conscience que les deux participants dirigent un focus attentionnel sur le monde, et crucialement, que celui-ci peut être partagé (Baldwin, 1995 ; Tomasello, 1995). La perception mutuelle d'un objet serait alors le résultat d'une capacité réflexive des individus impliqués dans l'interaction, qui rend possible la compréhension du fait qu'autrui a un focus attentionnel dirigé sur le même objet que soi (Tomasello, 1995). Les capacités d'attention conjointe chez l'enfant seraient alors dépendantes d'un certain niveau de développement cognitif (Adamson, Bakeman & Smith, 1990). Pour tester cette hypothèse, l'interrelation parmi un ensemble de tâches sociales et non-sociales a été évaluée chez des enfants de 9 à 15 mois (Carpenter et al., 1998). Il en ressort que les performances obtenues lors des tâches impliquant une coordination de l'attention sont systématiquement corrélées entre elles chez les enfants testés, ce qui tend à montrer que les habiletés sociales liées à l'attention conjointe émergeraient en même temps dans le développement (ibid). L'attention conjointe reposerait donc sur une compétence sous-jacente unique, qui n'émergerait qu'entre 9 et 12 mois.

Cette interprétation est corroborée par de nombreuses expériences, qui semblent attester de l'apparition des capacités d'attention conjointe autour du premier anniversaire. Par

exemple, il a été montré que les enfants d'un an sont capables de comprendre que le regard connecte l'observateur à l'objet qui est observé (Woodward, 2003). De la même façon, ils peuvent suivre le regard d'autrui pour localiser un objet intéressant, puis, un peu plus tard, utiliser le regard de l'adulte pour déterminer la direction de son focus attentionnel parmi deux objets proches (Baldwin, 1993a ; 1993b). Le premier anniversaire est alors considéré comme une véritable révolution socio-cognitive (Tomasello, 1995), à travers laquelle se déroule une transition développementale majeure : le passage de l'intersubjectivité primaire, qui caractérise les interactions dyadiques précoces (Trevarthen, 1979) à l'intersubjectivité secondaire, qui ouvre l'enfant à la conscience d'un troisième terme dans l'interaction (Trevarthen & Hubley, 1978 ; Fogel & DeKoeyer-Laros, 2007).

Le caractère tardif de cette compétence est classiquement expliqué par le fait qu'elle requière de la part de l'enfant la compréhension qu'il est, au même titre qu'autrui, un agent intentionnel (Dennett, 1987 ; Tomasello, 1995). Or, cette compréhension nécessite préalablement la distinction des relations moyens-buts, qui, selon une perspective piagétienne, n'émerge qu'à partir de 9 mois[38] (Piaget, 1936). Cette distinction semble nécessaire, dans la mesure où l'enfant reconnait alors que les buts qu'il poursuit sont séparés des moyens comportementaux destinés à les atteindre (Tomasello, 1999). Parallèlement, il observe qu'autrui manifeste lui aussi un ensemble de comportements dirigés vers l'environnement (Tomasello, Carpenter, Call, Behne & Moll, 2005). A travers un mécanisme d'analogie, il fait alors l'hypothèse que les actions d'autrui sont elles aussi dépendantes de buts ou d'intentions sous-jacents (Meltzoff, 2005). L'attention conjointe nécessiterait donc conjointement une attitude intentionnelle (Dennett, 1987 ; Csibra & Gergely, 1998 ; Gergely & Csibra, 2003), et une capacité d'attribution de cette intentionnalité à autrui (Baldwin & Moses, 1997 ; Meltzoff, 2005).

Mais, ainsi posée, l'attention conjointe désigne un mécanisme individuel destiné à la compréhension de l'*état* attentionnel de l'autre, ce qui rend possible le partage d'une représentation de l'objet de leur attention (ce partage étant lui-même représenté chez l'individu). L'attention conjointe repose ainsi sur une conception mentaliste de l'esprit (Baron-Cohen, 1995 ; Meltzoff, 1995 ; Tomasello, 1995 ; 1999) : un monde interne privé est nécessairement à l'origine du comportement, et la compréhension du comportement d'autrui serait médiée par une psychologie de sens commun, classiquement organisée en une théorie

[38] soit lors du sous-stade 4 du stade sensori-moteur (Piaget, 1936). Après avoir élaboré un ensemble de schèmes d'action lors des réactions circulaires secondaires (qui consistent à tenter de reproduire des résultats intéressants lors des activités tournées vers les objets), l'enfant cherche maintenant à atteindre des buts fixés à l'avance, en coordonnant les schèmes entre eux pour explorer des situations nouvelles (ibid).

de l'esprit (Astington & Gopnick, 1991 ; Premack & Woodruff, 1978). Dans cette perspective, l'attention conjointe serait alors le résultat d'une coordination d'états mentaux internes qui précède l'***action conjointe.*** Minimalement, l'action conjointe désigne toute forme d'interaction sociale où au moins deux individus coordonnent leurs actions dans l'espace et dans le temps de façon à produire un changement dans l'environnement (Sebanz, Bekkering & Knoblich, 2006). Ici, la coordination des actions ne pourrait se réaliser que si les interactants partagent *à priori* – et ont conscience qu'ils partagent – un même référent (Tomasello, 1995 ; 1999).

Or, comme nous l'avons déjà évoqué, cette conception mentaliste de l'esprit est largement questionnée par une approche incarnée et située de la cognition, qui permet notamment de réhabiliter l'action et l'engagement mutuel pour rendre compte de la compréhension sociale (Gallagher, 2004 ; Reddy & Morris, 2004 ; De Jaegher & Di Paolo, 2007). Ainsi, plus qu'une coordination interne d'états mentaux, la coordination qui a lieu dans le phénomène d'attention conjointe serait d'abord perceptive, et donc sensori-motrice (Gallagher, 2011). De ce point de vue, l'attention conjointe prendrait place au cœur d'une activité perceptive partagée, qui rend possible la co-construction d'un cadre de référence pragmatique. Pour replacer le phénomène d'attention conjointe dans une telle approche, la première étape consiste à considérer l'attention non comme un *état* qui rend possible l'engagement de l'individu dans le monde, mais comme un *processus* qui résulte de cet engagement (Reddy, 2011). Dans cette perspective, l'attention conjointe ne se déploie pas dans un monde interne de représentations et d'états mentaux (Tomasello, 1995 ; 1999), mais est dirigée dans le monde à travers l'engagement actif du sujet. En ce sens, la coordination interpersonnelle précède et informe, plutôt qu'elle ne dérive de représentations conceptuelles du soi et d'autrui (Reddy, 2003). Par conséquent, être conscient de l'attention d'autrui, c'est être conscient de son engagement dans le monde, celui-ci étant directement perceptible au travers de son action (Gallagher, 2004 ; Reddy, 2011 ; Robertson & Johnson, 2009).

3. L'attention conjointe comme coordination perceptive

Si l'engagement interpersonnel est premier, alors il ne peut pas être expliqué par une attitude réflexive. Dans cette optique, notamment, il a été démontré que la nature intentionnelle des mouvements est perçue par les enfants bien avant qu'ils ne développent une connaissance réflexive des états mentaux. En effet, des indices intentionnels primitifs semblent être détectés à même le mouvement, et ce, sans que cela ne requière une capacité à

lire et à attribuer des états mentaux (Premack & Premack, 1997). Dans la même veine, malgré un modèle modulaire de la compréhension sociale critiqué (Gallagher, 2008 ; Racine & Carpendale, 2007 ; Reddy & Morris, 2004), des mécanismes perceptifs sont à la base de l'élaboration d'une théorie de l'esprit (Baron-Cohen, 1995). En effet, des modules de détection de l'intentionnalité et de la direction du regard d'autrui (ID et EDD) précèdent l'élaboration d'un module destiné aux représentations triadiques (SAM), dont l'exercice conduira à l'élaboration d'un dernier module destiné à l'analyse des situations sociales (ToMM). Ainsi, la théorie de l'esprit elle-même ne serait que la dernière étape d'un processus essentiellement basé sur des mécanismes perceptifs primitifs.

Ces modèles semblent remettre en question le caractère unidimensionnel de l'attention conjointe (Carpenter et al., 1998). En effet, contrairement à l'étude de Carpenter et collaborateurs (ibid.), peu de corrélations sont trouvées parmi les performances obtenues par des enfants de 8 à 14 mois dans un ensemble de tâches sociales et non-sociales, même si leur nombre est plus restreint (Slaughter & McConnell, 2003). Ces résultats ont été interprétés comme reflétant une multiplicité des conduites d'attention conjointe, qui sont difficilement attribuables à un facteur commun (Striano & Bertin, 2005). Dans cette optique, par exemple, une distinction est faite entre les conduites relevant de *réponses* à l'attention conjointe, comme par exemple le suivi de regard, et les conduites *initiatrices* de l'attention conjointe, qui consistent à diriger activement l'attention d'autrui (Mundy & Acra, 2006). Dans le premier cas, l'enfant utilise les indices comportementaux d'autrui pour détecter la direction de son focus attentionnel (Mundy, 1995), ce qui donne lieu à une forme d'engagement mutuel « passif » (Adamson & Bakeman, 1985). Dans le deuxième cas, l'enfant utilise intentionnellement l'orientation de son focus attentionnel ou des gestes de pointage pour diriger le focus attentionnel d'autrui, de façon à partager une expérience (Mundy, 1995). Ces derniers comportements sont alors caractéristiques d'un engagement mutuel coordonné (Adamson & Bakeman, 1985). Bien souvent, ces deux niveaux dans les capacités d'attention conjointe sont confondus, ce qui explique la cohabitation de définitions contradictoires (Sigman & Kasari, 1995).

Ce qu'il est important de noter, c'est que ces deux ensembles de comportements présenteraient des trajectoires développementales différentes (Mundy, Block, Delgado, Pomares, Van Hecke & Parlade, 2007), mais complémentaires (Striano & Bertin, 2005). L'enfant commencerait d'abord par répondre à l'orientation attentionnelle de l'adulte (voir par exemple Bakeman & Adamson, 1984), avant d'être capable d'être l'initiateur des épisodes d'attention conjointe (Murray, Creaghead, Manning-Courtney, Shear, Bean & Prendeville,

2008). Mais si la capacité à répondre à l'attention conjointe est première, celle-ci n'apparaîtrait pas de façon abrupte (Bakeman & Adamson, 1984).

A partir de 6 mois, l'enfant est déjà capable de suivre la direction du focus attentionnel de la mère, même si cela n'aboutit pas nécessairement à la focalisation sur l'objet de son attention (Butterworth & Cochran, 1980 ; Butterworth & Jarrett, 1991). L'interprétation de cette capacité précoce est ici purement perceptive, et ne repose pas sur une nécessaire compréhension de l'intentionnalité, ni sur une capacité de « lecture mentale » de haut niveau (Tomasello, 1995 ; Baron-Cohen, 1995). Au contraire, l'attention conjointe présuppose ici un monde d'objets, pouvant être perçus depuis plusieurs points de vue, et existant dans un espace commun (Butterworth & Jarrett, 1991). L'enfant, assis devant sa mère, perçoit la direction de son focus attentionnel via un mécanisme « écologique », et engage sa propre activité perceptive dans cette direction (ibid.). Dès cette étape, l'enfant passe d'une interaction strictement dyadique (attention mutuelle), à une interaction sociale impliquant un objet. Ce mécanisme attentionnel met déjà en jeu une capacité de coordination complexe où l'autre n'est plus une cible mais une source perceptuelle qui permet d'organiser son propre comportement (Conein, 1998 ; Gomez, 1991). Les signaux émis par la mère (les mouvements des yeux et de la tête) sont informatifs, dans la mesure où ils signalent la présence potentielle d'un objet intéressant dans l'espace. Ils agissent donc comme des affordances sociales, qui offrent à l'enfant des possibilités d'action dans le contexte de l'interaction (Krueger, 2011). Dans une optique constructiviste, cette capacité précoce de suivi du regard formerait la base sensori-motrice de la participation de l'enfant à des contextes d'interaction de plus en plus complexes (Racine & Carpendale, 2007). Cette idée est soutenue par de nombreux exemples témoignant d'une progression graduelle des capacités d'attention conjointe (Striano & Bertin, 2005 ; Striano & Rochat, 1999), les premières étapes étant même observées dès 3 mois (Scaife & Bruner, 1975 ; D'entremont et al., 1997 ; Tremblay & Rovira, 2007).

Selon Butterworth et Jarrett, cependant, le mécanisme écologique ne serait réellement consolidé que vers 12 mois : l'enfant manifeste alors une capacité à suivre le regard d'autrui en tant que focalisation attentionnelle sur un objet du monde partagé (Butterworth & Jarrett, 1991). Cette fois, la cible est correctement localisée, même si celle-ci n'est pas la première cible rencontrée par l'enfant lors de l'orientation de son regard dans la direction du focus attentionnel de la mère. Le mécanisme est donc maintenant « géométrique » : l'enfant peut extrapoler de l'orientation du corps et du regard de la mère l'intersection de son regard avec une zone précise de son propre champ visuel. Toutefois, il est à noter que cette localisation n'est possible que si la cible est présente dans le champ visuel direct de l'enfant (ibid.). A

partir de 18 mois, l'enfant est capable de suivre le regard de la mère vers une cible située derrière lui, mais à la condition qu'aucune cible ne soit présente dans son champ visuel direct (Butterworth & Jarrett, 1991). Les enfants de 18 mois sont donc capables d'interpréter l'attention visuelle de l'adulte comme référant potentiellement à quelque chose situé hors de leur champ visuel. Ainsi, les enfants peuvent accéder à un espace représenté uniquement s'il n'y a pas de compétition avec un espace directement perçu. Cette transformation de la compréhension de l'orientation attentionnelle de la mère permettrait à l'enfant de comprendre qu'il peut lui aussi « contrôler » la direction du regard d'autrui au moyen de son propre regard de façon à l'orienter vers une cible intéressante (Butterworth, 1991 ; Mundy & Acra, 2006). Notons que le pointage proto-déclaratif apparaît également à cet âge (Bates et al., 1975), celui-ci étant alors nécessairement précédé dans le développement par des capacités à répondre à l'orientation attentionnelle de l'autre.

Avec ce modèle, l'attention conjointe est alors définie comme un mécanisme fondamentalement perceptif (Butterworth, 1991 ; Butterworth & Jarrett, 1991), qui acquiert du sens au fur et à mesure des expériences sociales où il est impliqué (Racine & Carpendale, 2007). Cette conception perceptive et constructiviste nous paraît plus satisfaisante qu'une approche mentaliste, dans la mesure où elle permet une interprétation énactive du phénomène d'attention conjointe (voir par exemple Gallagher, 2011 ; Varela, et al., 1991). Dans cette perspective, nous pouvons rappeler ici que la perception est conçue comme une activité relationnelle permettant d'établir un couplage entre l'agent et son environnement (Lenay, 2005 ; 2006a). Cette activité relationnelle s'exerce via une extraction des invariants de la sensation (Gibson, 1966 ; 1979), ou plutôt via une constitution de lois de contingences sensori-motrices qui structurent la relation entre les actions et les sensations qui en résultent (O'Regan & Noé, 2001). De plus, cette activité n'est pas neutre d'un point de vue cognitif, dans la mesure où son rôle est de spécifier les propriétés de l'environnement (Gibson, 1966), et plus fondamentalement les propriétés de la relation entre l'agent et cet environnement (O'Regan & Noé, 2001). En ce sens, la perception est une activité de construction de sens (« sense-making ») (De Jaegher & Di Paolo, 2007), car elle permet à l'agent à maintenir son autonomie, en régulant ses interactions avec un environnement viable, parce que signifiant (Di Paolo, 2009). Dans ce contexte, la perception mutuelle d'un objet, et la conscience de cette perception seraient le produit de la rencontre des activités perceptives, qui établissent un *participatory sense-making* dans lequel la connaissance serait co-construite (De Jaegher & Di Paolo, 2007).

Ce déplacement épistémologique rend possible une interprétation plus mesurée des précurseurs nécessaires aux interactions impliquant des objets, basée sur une connaissance du fait qu'autrui porte son attention sur certains aspects du monde (Bruner, 1995), ou encore sur une reconnaissance du fait que les focus attentionnels sont partagés (Baldwin, 1995). Si cette connaissance ou cette reconnaissance est construite sur la base de mécanismes perceptifs, alors il y a tout lieu de penser que notre dispositif rende possible l'établissement d'épisodes d'attention conjointe, à travers la coordination des activités perceptives dans un environnement partagé. Mais il nous reste à comprendre précisément ce qui est perçu dans le cadre d'une telle coordination entre la mère et l'enfant.

4. Le rôle de la dynamique du mouvement

Selon le modèle de Baron-Cohen, l'enfant serait à même de détecter la direction du focus attentionnel de façon très précoce sur la seule base de la direction des yeux (Baron-Cohen, 1995). En effet, il est montré que dès le 2e mois, les enfants sont attirés de façon préférentielle par la région des yeux lorsqu'ils regardent le visage d'autrui (Maurer, 1985), et qu'ils discriminent deux visages caractérisés par une direction des yeux différente (Vecera & Johnson, 1995). Pourtant, un certain nombre de recherches suggère que la direction du regard en tant que telle ne serait détectée par l'enfant qu'après l'orientation du corps (Meltzoff & Brooks, 2007), ou de la tête (Corkum & Moore, 1995). En effet, il semble que les enfants de 12 mois engagent leur regard dans la direction du mouvement de la tête d'un adulte, même si le regard mutuel est maintenu. En revanche, le regard ne suit pas la direction du mouvement des yeux de l'adulte si la tête reste immobile face à l'enfant (ibid.). Ces résultats pourraient être expliqués par la difficulté de ces enfants à se désengager de la saillance du visage de l'adulte dans leur expérience perceptive du face à face (Hood, Willen & Driver, 1998). Il est alors possible que ce désengagement de la fixation sur le visage soit favorisé lors d'un mouvement de la tête parce que les yeux sont par conséquent moins visibles (Maurer, 1985). Mais ces résultats peuvent également être expliqués par le fait que les mouvements de la tête ou du corps fournissent un indice du mouvement plus saillant que le seul mouvement des yeux (Anstis, Mayhew & Morley, 1969, cité par Hood et al., 1998). Par exemple, dans l'expérience de Butterworth et Jarrett, le focus attentionnel de la mère est dirigé de façon globale : l'enfant pourrait alors s'engager dans la direction du focus attentionnel parce qu'il a assisté à la combinaison du mouvement d'orientation des yeux avec celui de l'orientation de la tête et du corps (Butterworth & Jarrett, 1991). En effet, plus que la direction du regard en

soi, la détection du focus attentionnel d'autrui pourrait dériver de l'intégration de la direction du regard et de l'orientation corporelle, parallèlement à différents indices verbaux (Vaughan Van Hecke & Mundy, 2007). De cette façon, l'enfant pourrait s'engager dans le suivi de regard via les aspects dynamiques du mouvement d'orientation attentionnelle (Lempers, 1979 ; Moore, Angelopoulos & Bennett, 1997).

Dans une expérience princeps concernant cette question, des enfants de 9, 12 et 14 mois sont placés en situation de face-à-face avec leur mère. Lorsque le contact visuel est établi, l'expérimentateur distrait l'enfant, pendant que la mère oriente sa tête et ses yeux en direction d'une cible située sur le côté. De cette façon, lorsque l'enfant regarde à nouveau sa mère, il n'aperçoit que la position finale de la tête et des yeux et pas le mouvement d'orientation. Dans ce cas de figure, aucun des enfants de 9 mois, et seulement la moitié des enfants de 12 et 14 mois parviennent à localiser la cible (Lempers, 1979). Plus récemment, une expérience similaire a été effectuée, de façon à comparer les résultats obtenus entre les conditions « mouvement et orientation finale » et « orientation finale seule » chez des enfants de 9 mois (Moore et al., 1997). De plus, la capacité initiale de ces enfants à s'engager dans le suivi du regard d'un partenaire interactif a été contrôlée. Les résultats montrent que quelle que soit la capacité initiale de l'enfant à suivre le regard, les résultats sont meilleurs si l'enfant a assisté au mouvement d'orientation que s'il a été confronté uniquement à l'orientation finale de la tête du partenaire. De façon à confirmer le rôle crucial du mouvement dans la capacité des enfants de 9 mois à suivre le regard d'un partenaire, une variante de cette expérience a consisté à ne présenter à l'enfant que le mouvement d'orientation, en cachant la position finale de la tête. Il apparaît que le mouvement seul permet aux enfants de s'orienter dans la bonne direction (ibid.). Enfin, si le mouvement est crucial, n'importe quel mouvement n'est pas suffisant : il faut que ce mouvement d'orientation soit naturel. En effet, il a été montré que les enfants réussissaient davantage la tâche de suivi du focus attentionnel dans les cas où le mouvement de la tête du modèle s'oriente sur un plan horizontal, et ce, même si le mouvement n'aboutit pas jusqu'à une position finale où les yeux atteignent la cible (ibid.). Lorsque le mouvement d'orientation s'exécute sur un plan vertical (l'oreille du modèle s'approche vers l'épaule dans la direction de la cible), la majorité des enfants ne réussit pas la tâche.

Toutefois, dans ces expérimentations, il est à noter que les enfants qui ont manifesté spontanément la capacité à suivre le regard avant l'expérimentation ont obtenus des résultats particuliers. En effet, ils ont été capables i) de s'engager dans la bonne direction sur la base de la seule orientation finale de la tête ; et ii) d'apprendre la signification du mouvement de la

tête sur un plan vertical. Ainsi, quand la capacité à suivre le regard est déjà établie, il semble que l'enfant puisse non seulement inférer la direction du focus attentionnel sans assister au mouvement d'orientation, mais aussi apprendre à utiliser d'autres formes de mouvements pour guider leur attention vers la cible (Moore et al., 1997). En revanche, quand cette capacité n'est pas encore établie, le mouvement d'orientation semble nécessaire pour apprendre la signification de la direction du focus attentionnel. Ainsi, le développement de la capacité à suivre le regard serait expliqué par un processus de conditionnement (Moore & Corkum, 1994 ; Corkum & Moore, 1995), se réalisant au travers des situations sociales naturelles auxquelles l'enfant est confronté. La capacité à suivre le focus attentionnel d'autrui pourrait alors être construite à travers l'extraction de régularités dans le comportement manifeste des agents lors des situations sociales naturelles. En effet, dans celles-ci, l'enfant est rarement confronté à l'orientation finale seule du focus attentionnel. Au contraire, l'enfant peut percevoir l'ensemble des comportements de l'agent interactif, qui s'oriente dynamiquement dans un monde d'objets. Les enfants sont donc entraînés dans ces situations, en apprenant notamment que lorsque l'agent produit un mouvement d'orientation dans une direction donnée, il y a souvent un objet ou un évènement à percevoir dans cette direction. Les capacités précoces d'attention conjointe émergeraient donc de l'expérience des situations d'interaction, qui permettent à l'enfant de construire un ensemble de connaissances pragmatiques et contextualisées (Carpendale & Lewis, 2004 ; Racine & Carpendale, 2007).

Mais pour l'heure, l'élaboration de ces connaissances semble exclusivement dépendante de la modalité visuelle, qui rend possible la perception distale de la dynamique des mouvements dans les situations d'interaction. Dans cette optique, notre dispositif ne semblerait pas remplir les conditions nécessaires à l'établissement d'une coordination perceptive autour d'objets communs. En effet, nous pouvons rappeler ici que notre dispositif ne permet pas une perception visuelle, et n'en est pas non plus un substitut. Au contraire, il rend possible la constitution d'une nouvelle modalité perceptive, basée sur l'extraction des lois de couplage entre le mouvement d'exploration et les retours sensoriels qui en résultent (Lenay et al., 2000). De cette façon, si l'utilisateur a effectivement accès à des informations proprioceptives (les informations kinesthésiques de son propre mouvement d'exploration, voir Auvray et al., 2005), il ne peut pas avoir accès aux informations kinesthésiques liées au mouvement de l'autre utilisateur. Il lui reste alors à relier son propre mouvement avec des retours tactiles séquentiels, qui résultent de la rencontre de son corps-percevant avec un objet de l'espace commun, qu'il soit fixe ou mobile (comme dans le cas d'un objet en mouvement ou du corps-image d'un autre utilisateur). De cette façon, la perception du mouvement en tant

que telle n'est pas possible : l'utilisateur n'accède qu'aux positions successives des objets rencontrés, en fonction de son propre mouvement d'exploration. De façon à savoir si la coordination des activités perceptives peut se réaliser dans ces conditions, le cas de l'attention conjointe chez le sujet déficient visuel peut alors nous apporter des pistes intéressantes.

5. L'attention conjointe sans la vision ?

Dans la littérature, il existe peu de documentation concernant l'émergence de l'attention conjointe chez les enfants aveugles, ce qui peut être expliqué par la prégnance de la modalité visuelle dans la compréhension des interactions sociales précoces (voir par exemple Brazelton et al., 1975), tout comme par la rareté des sujets aveugles ne souffrant pas de troubles cognitifs associés (Bigelow, 2003). En effet, nous avons vu que la cécité a un impact négatif sur les interactions sociales précoces, réduisant à la fois l'engagement social du jeune aveugle, mais aussi les sollicitations sociales de son entourage (Pérez-Pereira & Conti-Ramsden, 1999). Ainsi, c'est l'ensemble de la dyade qui est affectée par la déficience visuelle (Rogers & Puchalski, 1984), à moins que des conduites adaptées ne se développent pour garantir une coordination interpersonnelle propice à l'établissement d'un espace intersubjectif entre l'enfant et ses partenaires sociaux. Ce qui nous intéresse plus particulièrement ici, c'est l'introduction d'objets dans la dyade, et comment l'interaction s'organise pour que les partenaires puissent coordonner leurs perspectives attentionnelles autour d'objets partagés quand au moins un des deux partenaires est déficient visuel.

Ceci nous amène à compléter les conséquences de la cécité sur le comportement de l'enfant, pour nous focaliser plus particulièrement sur ses conduites envers le monde physique et les objets. En effet, parallèlement à la perturbation des connaissances du soi interpersonnel, la privation visuelle a une incidence notable sur les connaissances tirées du soi écologique (Neisser, 1991). Chez l'enfant aveugle, en effet, l'absence de valeur d'appel de la part de l'environnement induit une maîtrise de l'espace plus tardive que chez les enfants bénéficiant de la vision, même lorsque celle-ci est résiduelle (Hatwell, 2003). En effet, en l'absence totale de vision, la perception du monde physique et des relations spatiales (entre les objets et entre le soi et les objets) ne peut se réaliser qu'au travers de rencontres séquentielles avec l'environnement, via la modalité tactile. La perception tactile est en effet une modalité proximale, rendant nécessaire un contact entre le corps et l'objet perçu. De cette façon, la perception des objets est déterminée par la position du corps dans l'espace. En conséquence, les connaissances spatiales sont essentiellement égocentrées (ibid.). De plus, l'absence de

perception distale induit une difficulté à comprendre que des objets potentiellement intéressants sont situés au-delà de l'espace péri-personnel, et que ceux-ci bénéficient d'une permanence propre indépendamment du soi (Bigelow, 1995). Les possibilités d'interaction de ces enfants avec l'environnement et avec les objets qui y sont présents sont alors réduites. De cette façon, la difficulté d'accès aux informations spatiales, couplée à un défaut d'incitation perceptive, a un impact négatif sur les comportements d'atteinte et sur la locomotion (Adelson & Fraiberg, 1974 ; Bigelow, 1992), et plus généralement sur l'engagement corporel autonome dans l'espace (Bigelow, 1995). La compréhension de l'espace pourrait alors nécessiter des pré-requis cognitifs ou des efforts d'interprétation supplémentaires par rapport aux enfants voyants (Adamson et al., 1990 ; Bigelow, 1995 ; 2003), à moins que l'adulte n'intervienne efficacement pour offrir aux enfants aveugles des opportunités d'expériences pertinentes avec le monde physique (Hobson, 1993 ; Bigelow, 2003). L'intervention de l'adulte, au travers de séquences d'interaction avec l'enfant déficient visuel, pourrait alors lui servir de support pour s'engager dans l'espace, et ainsi acquérir des connaissances sur l'environnement et sur les objets qui s'y trouvent. Mais cette intervention de l'adulte n'est elle-même pas facilitée par la déficience visuelle, dans la mesure où ce type de guidage implique une focalisation attentionnelle partagée. Or, il est difficile pour l'enfant aveugle de détecter la direction de l'attention d'autrui, et donc de comprendre ses manifestations comportementales et ses réactions émotionnelles envers les objets en l'absence d'une référence perceptive commune (Brown et al., 1997 ; Hobson, 1993). Parallèlement, il est difficile pour leurs partenaires interactifs de détecter sur quoi l'attention de l'aveugle est dirigée : il n'y a pas d'orientation du regard ou du visage, ni de gestes de pointage qui pourraient manifester clairement la direction de son focus attentionnel (Fraiberg, 1979). Toutefois, sur ce dernier point, il est important de noter que certains comportements de l'enfant déficient visuel pourraient véhiculer ce type d'information pour autrui (Fraiberg, 1979 ; Rowland, 1984). Par exemple, dans son observation longitudinale de dyades mère/enfant déficient visuel, Preisler (1991) remarque que lorsque l'enfant entend un son particulier de l'environnement qui attire son attention, il semble figer sa posture et ses expressions faciales. Mais les analyses détaillées de la posture de l'enfant révèlent qu'en réalité, il se penche légèrement vers celui-ci, signalant potentiellement à un entourage averti qu'il focalise son attention dans cette direction. Ce type de comportement, s'il est correctement détecté et interprété par la mère, pourrait alors servir à l'établissement d'un cadre de référence mutuel, en étant comparable à une forme de « pointage corporel ». Malheureusement, ces indices ne sont bien souvent pas remarqués par

la mère et ne sont donc pas utilisés, et les difficultés subsistent (Fraiberg, 1979 ; Rowland, 1984).

Malgré tout, il apparaît que des capacités d'attention conjointe peuvent tout de même se développer chez l'enfant aveugle, même s'il est à noter que ces capacités émergent plus tardivement que chez les enfants voyants et malvoyants. (Bigelow, 2003 ; Hobson, 1993 ; Ingsholt, 2000). Ceci confirme de nouveau le caractère facilitant de la modalité visuelle pour l'interaction en général, et plus particulièrement en ce qui concerne l'engagement dans les interactions sociales entre les partenaires et un objet (Bigelow, 2003 ; Ingsholt, 2000). Dans cette optique, les difficultés rencontrées par les enfants aveugles pourraient résulter de difficultés strictement perceptives pour coordonner des orientations subjectives avec autrui (Hobson, 2002). Mais comme dans le cas de l'établissement d'interactions dyadiques, des moyens alternatifs peuvent compenser l'absence de vision (Pérez-Pereira & Conti-Ramsden, 1999). Par exemple, dans le cadre d'une étude longitudinale concernant l'impact de la déficience visuelle sur les interactions communicatives précoces, des auteurs ont mené une analyse qualitative des échanges entre la mère et l'enfant, en fonction de leurs statuts visuels respectifs (Rattray & Zeedyk, 2005). Cette étude est exemplaire à bien des égards, dans la mesure où à notre connaissance, l'analyse de séquences d'interactions impliquant deux partenaires déficients visuels et un objet partagé est absente de la littérature. Les résultats montrent que lorsqu'au moins un des deux participants est déficient visuel, le processus de d'interaction repose sur le toucher, les vocalisations et même sur l'orientation faciale et/ou posturale des deux partenaires (ibid.). Dans ce contexte, un bel exemple de coordination interpersonnelle nous est rapporté : il s'agit de la transcription d'une séquence d'interaction dans laquelle une mère non-voyante joue avec son enfant non-voyant de 16 mois :

> « La mère et le bébé sont assis sur le sol. La mère est assise derrière le bébé, de telle façon qu'ils sont tous les deux orientés dans la même direction. La mère tient une grosse pomme plate en bois, qui comporte des trous sur l'extérieur et une grande baguette en bois attachée par une ficelle, qui est destinée à s'insérer dans les trous de la pomme. Le bébé tient la baguette dans sa main, et la mère dit : 'Mets-la dans le trou'. La mère retire alors une de ses mains de la pomme et aide l'enfant à insérer la baguette dans le premier trou.
>
> La mère dit alors : 'Tu fais le suivant' Le bébé essaye d'insérer la baguette dans le prochain trou de la pomme, mais a des difficultés pour la retirer complètement. La mère

retire sa main du bord de la pomme et aide le bébé à retirer complètement la baguette. Le bébé dit 'Yeh !' » (Rattray & Zeedyk, 2005, p.293, *traduction personnelle*).

Cet exemple nous montre que malgré l'absence de vision, les comportements des deux partenaires sont coordonnés efficacement, en utilisant conjointement l'ensemble des modalités perceptives disponibles. Si celles-ci soutiennent généralement la communication dans la dyade, elles sont mobilisées notamment pour signaler la direction de leurs focus attentionnels respectifs, et ce, dans le cours même de l'action (Rattray & Zeedyk, 2005). Ainsi, si l'attention conjointe se réfère typiquement à l'établissement d'une focalisation attentionnelle mutuelle pour l'engagement interpersonnel dans une séquence d'interaction, cet exemple souligne le fait que la coordination des focus attentionnels est également nécessaire pour le maintien de cette interaction dans le temps. En effet, le maintien de l'engagement mutuel est soumis à un ensemble de coordinations attentionnelles locales, qui s'inscrivent dans une routine d'interaction (Bruner, 1983). De cette façon, l'attention conjointe n'est pas une simple procédure de désignation, mais renvoie à une structure d'échange qui permet non seulement de coordonner les focus attentionnels autour d'un même objet ou d'une même activité, mais aussi et crucialement de maintenir cette structure dans le temps (ibid.). Cette extension du concept d'attention conjointe est également soutenue dans le cadre de l'observation d'interactions entre des enfants voyants dans un contexte écologique, où l'attention conjointe est décrite comme un processus qui concerne autant la production ou la réponse à une orientation attentionnelle initiale que la production ou la réponse à une transformation de cette orientation attentionnelle au cours d'une activité conjointe (Kidwell & Zimmerman, 2007).

Par ailleurs, si l'attention conjointe se réfère typiquement à la coordination des focus attentionnels via la modalité visuelle, celle-ci peut également dériver d'informations multimodales (Baron-Cohen, 1995), comme les informations tactiles ou auditives (Sarria, Gomez & Tamarit, 1996). Par exemple, dans ce contexte, une fréquence élevée de vocalisations de la part de la mère est observée, et ce, quel que soit le statut visuel des participants (Rattray & Zeedyk, 2005). Celles-ci fournissent non seulement à l'enfant déficient visuel une présence rassurante, mais elles encouragent ou restreignent l'action elle-même, en renseignant l'enfant sur son déroulement et sur sa signification en fonction de l'état du monde. En outre, elles rendent possible la constitution de patterns d'interaction synchronisés et réciproques (proto-conversations), étape nécessaire à l'acquisition du langage par la consolidation de « formats » structurants et signifiants (Bruner, 1983). En effet,

« l'enfant n'apprend pas d'abord à parler ; il apprend d'abord les usages du langage dans son commerce quotidien avec le monde, en particulier avec le monde social. » (ibid., p. 25). Plus tard, lorsque le langage est suffisamment maîtrisé par l'enfant lui-même, il est alors un moyen privilégié par la mère pour guider explicitement l'attention de l'enfant sur un aspect de l'environnement (Rattray & Zeedyk, 2005). Parallèlement, il est un outil pertinent pour l'enfant de façon à diriger l'attention d'autrui sur son propre focus attentionnel (Bigelow, 2003).

En ce qui concerne le toucher, il est souligné que son utilisation est principalement influencée par le statut visuel de l'enfant (Rattray & Zeedyk, 2005). En effet, les mères d'enfants déficients visuels, qu'elles soient elles-mêmes déficientes visuelles ou non, utilisent le toucher actif pour assister le processus d'interaction, et ce, notamment dans le cas des interactions impliquant des objets. Les mères apportent alors les objets à l'enfant et l'assistent dans leur manipulation de façon à compenser l'absence de feedback visuel. La mère « montre » alors à l'enfant déficient visuel ce qu'il y a à faire et à percevoir via le toucher actif, quand les mères d'enfants voyants peuvent donner des indications à l'enfant via le regard ou le pointage. Dans tous les cas, l'adulte structure le contexte de l'interaction, en présentant à l'enfant les affordances disponibles et en assistant ses actions envers les objets. Chez l'aveugle, cette structuration de l'interaction, via l'étayage de la mère (Bruner, 1983), permet de favoriser l'acquisition de connaissances spatiales, tout comme la découverte du fait que leur expérience avec les objets peut être partagée[39] (Bigelow, 2003). Ainsi, la déficience visuelle de l'un des partenaires ou même des deux ne réduit pas en soi la qualité des échanges, mais rend nécessaire la co-construction de modes de communications alternatifs (Pérez-Pereira & Conti-Ramsden, 1999).

Or, ces modes d'interaction alternatifs nous semblent particulièrement révélateurs d'un fait essentiel : les caractéristiques inhérentes à ces modalités perceptives forcent le déploiement des activités perceptives. Par exemple, la modalité tactile étant nécessairement proximale, elle rend nécessaire le déploiement spatio-temporel d'un geste d'exploration nécessaire à la perception. On parle alors de perception haptique, ou tactilo-kinesthésique (Hatwell, 2003). Dans le cadre de l'interaction, le toucher actif désigne alors la coordination de ces activités déployées dans l'espace et dans le temps. Dans le cas de l'audition, nous avons affaire à un flux d'informations sonores ou verbales, dont le déploiement est alors cette

[39] Notons que ceci est également vrai pour les enfants voyants, dont les parents contrôlent bien souvent l'ensemble des variables qui doivent coïncider pour que l'enfant puisse participer activement aux interactions sociales impliquant des objets (Carpenter et al., 1998).

fois purement temporel (Gibson, 1966). L'échange de ces informations nécessite alors une coordination temporelle entre les participants, via la mise en place d'une alternance de « tours de parole ». L'usage de ces modalités perceptives met alors en évidence le fait que la focalisation attentionnelle commune découle d'une coordination spatio-temporelle des activités perceptives. Dans ce sens, la coordination qui a lieu dans l'attention conjointe est d'abord motrice et affective (Gallagher, 2011), et peut même être décrite comme une action conjointe basique[40] (Fiebich & Gallagher, 2012). L'attention conjointe est alors entendue comme une séquence spécifique d'activités coordonnées, qui prennent place bien souvent dans le cadre d'un contexte interactif plus global (Kidwell & Zimmerman, 2007). Enfin, si le toucher ou l'audition rend plus facilement observable le déploiement de ces activités, ce déploiement est également nécessaire en ce qui concerne la vision : si la multiplicité des champs récepteurs de la rétine permet de capturer à chaque instant un grand nombre d'informations spatialisées, la perception visuelle ne se réalise que lorsque le sujet maîtrise les lois de couplage qui relient son action (le mouvement de l'œil, de la tête ou du corps dans l'espace) avec la variation des stimulations visuelles captées par la rétine (O'Regan & Noé, 2001). Dans le cadre d'une attention conjointe visuelle, le suivi du regard est en soi un *engagement* de l'enfant en réponse à l'engagement de la mère (Butterworth & Jarrett, 1991 ; Reddy, 2011). Par ailleurs, il est souvent noté que l'attention conjointe visuelle est établie lorsque les individus impliqués alternent leur regard entre l'objet de l'attention et autrui (Carpenter et al., 1998 ; Mundy et al., 2007). En ce sens, l'engagement attentionnel des partenaires est englobé dans un jeu de co-ajustements sensori-moteurs qui dépendent des modalités perceptives disponibles (soit des retours sensoriels couplés à des possibilités d'action spécifiques), et qui définissent la signification même des comportements en fonction du contexte et des objets présents.

6. L'attention conjointe comme pratique sociale incarnée

Si la coordination des focus attentionnels est en soi une action conjointe, celle-ci n'est pas un évènement isolé. Elle résulte d'un processus d'engagement mutuel qui ne peut pas être désolidarisé du flux continu de l'activité sociale et du contexte dans lequel cette activité se déroule (Kidwell & Zimmerman, 2007). Dans cette optique, la coordination des focus attentionnels ne peut faire sens qu'à l'intérieur d'un contexte pragmatique qui offre une

[40] Une action est dite basique lorsqu'elle ne requière pas d'autres actions pour sa réalisation (Chisholm, 1969, cité par Fiebich & Gallagher, 2012)

structure à l'orientation attentionnelle de l'individu et à celle de son partenaire. L'attention conjointe est alors enracinée dans une pratique partagée qui fait sens pour les participants en fonction du contexte dans lequel l'interaction se déroule (Carpendale & Lewis, 2004).

Par exemple, dans une expérience sur la sensibilité à la direction du regard lors d'interactions dyadiques médiatisées par un dispositif de double lien vidéo, des enfants de 3 à 6 mois ne semblent pas discriminer les situations où l'adulte regarde ailleurs en déplaçant seulement les yeux sur le côté ou en bas, ou en déplaçant la tête et les yeux sur le côté. La durée des regards est comparable entre ces trois conditions et une condition contrôle, où l'adulte maintient un regard mutuel avec l'enfant (Hains & Muir, 1996). Toutefois, il est noté que les enfants sourient moins lors des conditions où l'adulte regarde ailleurs que lorsque le contact visuel direct est maintenu (ibid). Cette expérience montre que l'enfant est sensible à la direction de l'attention d'autrui lorsqu'elle est dirigée vers lui, mais ne discrimine pas les cas où l'attention d'autrui est portée ailleurs. Ce résultat semble aller dans le sens d'une forme de conscience précoce du fait que le soi est un objet potentiel d'attention (Reddy, 2003). De façon intéressante, nous notons que l'enfant continue de regarder le modèle même si celui-ci regarde ailleurs : il ne s'engage pas dans un suivi de son regard, malgré la présence des indices dynamiques de son mouvement d'orientation attentionnelle (Lempers, 1979 ; Moore, Angelopoulos & Bennett, 1997). Mais lorsque le regard de l'adulte n'est pas orienté vers l'enfant, il est important de noter que celui-ci s'oriente vers un ailleurs indéterminé, et pas vers un objet que l'enfant pourrait lui-même percevoir. Or, la présence même d'un objet est une condition nécessaire à l'établissement d'une attention conjointe via une coordination des regards (Corkum & Moore, 1995). A partir de cette constatation, des auteurs ont élaboré une expérience similaire, en plaçant cette fois l'enfant dans le contexte d'une interaction en présence d'objets (Striano & Stahl, 2005).

Dans cette expérience, des enfants de 3, 6 et 9 mois font face à un expérimentateur (E). Un objet est placé à la gauche de l'enfant, à égale distance entre les participants. Pendant des sessions de 5 minutes, E maintient un contact visuel avec l'enfant et interagit avec lui lors des minutes 1, 3 et 5. Pendant les minutes 2 et 4, E rompt l'interaction en face à face pour regarder l'objet, et ce, en fonction de deux conditions expérimentales. Dans la condition 1, qualifiée de « séquence d'attention conjointe », E regarde l'objet, sourit, parle de façon positive puis alterne son regard entre l'objet et l'enfant. Dans la condition 2, E procède de la même manière, mais cette fois, il n'alterne pas son regard entre l'objet et l'enfant. Les résultats montrent que dès 6 mois, les enfants s'engagent dans le suivi du regard dans les deux conditions expérimentales et focalisent leur attention alternativement sur l'objet et sur

l'adulte. De plus, si les enfants de 3 mois ne s'engagent pas dans le suivi du regard, ils ne manifestent pas l'effet classique du still-face dans la condition où E alterne son regard entre l'objet et l'enfant, ce qui peut être interprété par le fait que bien que l'interaction dyadique soit rompue, ces enfants maintiennent leur engagement dans une interaction qui implique désormais l'objet de l'attention du modèle. En revanche, dans la condition où E n'alterne pas son regard, la réponse classique du still-face est observée, même si E continue de manifester des sourires et un ton de voix positif (ibid.).

Ces résultats suggèrent que les capacités d'engagement de l'enfant dans le contexte d'une interaction impliquant un objet, qu'elles conduisent ou non à un suivi de regard, se manifestent dans des contextes pragmatiques où la présence même d'objets donne du sens à l'engagement de chacun des interactants. D'un point de vue individuel, en effet, l'objet est une affordance qui attire le focus attentionnel et organise l'action. Dans le contexte d'une interaction, l'objet présent dans le monde partagé peut être conçu comme une *affordance simultanée* (Marsh et al., 2006) : il attire les focus attentionnels de façon réciproque, et fournit une opportunité d'action identique pour les interactants. En ce sens, l'objet commun lui-même peut être conçu comme l'attracteur des activités perceptives, qui organise la rencontre et qui donne lieu à une coordination interpersonnelle. Comme dans le cadre des interactions dyadiques, nous postulons que l'établissement et le maintien de cette coordination se réalise via l'émergence d'une dynamique relationnelle autonome (De Jaegher et al., 2010 ; Lenay & Stewart, 2012). La *conscience* ou la *connaissance* du fait que l'objet de l'attention est partagé découle de cette dynamique, car c'est à travers elle que peut se co-construire le sens via un participatory sense-making (De Jaegher & Di Paolo, 2007). En d'autres termes, c'est la rencontre de nos engagements dans le monde qui permet de donner du sens aux contributions respectives des interactants, en même temps qu'elle structure ces engagements. En effet, les comportements individuels s'inter-influencent nécessairement, et transforment ainsi dynamiquement les réponses de chacun (voir par exemple Richardson, Marsh, Isenhower, Goodman & Schmidt, 2007). Ici, la coordination interpersonnelle ne concerne pas seulement un co-ajustement des activités individuelles dirigées l'une vers l'autre, mais elle concerne un co-ajustement des activités individuelles dirigées successivement l'une vers l'autre et vers un objet partagé. Par conséquent, si nous souhaitons mettre l'accent sur la dynamique de la rencontre des activités perceptives, il nous semble nécessaire d'inclure l'influence de l'objet lui-même sur le processus de coordination interpersonnelle. De cette façon, l'attention conjointe va au-delà de compétences individuelles et même d'une simple relation en face-à-face. Il s'agit donc d'élargir le cadre conceptuel de l'interaction pour y inclure dès le départ

un monde de relations interpersonnelles et des objets partagés, autour desquels peuvent se constituer des communautés de pratique (Kummer, 1967, cité par Conein, 1998). Dans cette optique, Chapman propose d'adopter le « triangle épistémique » en tant que cadre d'analyse des interactions (Chapman, 1991, p.211). Il s'agit alors de considérer les relations mutuelles entre le sujet, l'objet de sa connaissance et un autre sujet, qu'il soit présent ou non (ibid.). En d'autres termes, la compréhension de l'attention conjointe requière la pleine prise en compte du contexte de l'interaction, des actions individuelles respectives se déployant dans ce contexte, et du processus de l'interaction qui émerge de la rencontre des activités individuelles.

Dans ce cadre, nous proposons une expérimentation qui vise à étudier l'engagement interpersonnel dans le contexte d'un espace numérique partagé. Cette expérimentation présente plusieurs objectifs complémentaires. Premièrement, il s'agit de tenter de démontrer expérimentalement que la coordination interpersonnelle des activités perceptives (dirigées vers l'autre et vers les objets) est une condition nécessaire et suffisante pour l'attention conjointe. Ici, nous entendons l'attention conjointe en tant que processus qui permet d'aboutir à une connaissance du fait que les focus attentionnels des interactants sont dirigés vers le même objet. Cette connaissance émerge de la rencontre interpersonnelle autour d'objets communs, qui fait émerger une dynamique relationnelle entre les sujets et les objets, et qui sert de support à la co-construction d'un cadre de référence partagé. Pour forcer le déploiement de cette coordination, nous partirons d'un cadre minimaliste strict. Deuxièmement, il s'agit de comparer différentes configurations de corps-image/corps-percevant, de façon à observer l'effet du dispositif de couplage lui-même sur l'émergence d'une telle dynamique. Si cet axe peut se révéler pertinent d'un point de vue théorique pour tenter d'isoler les éléments fondamentaux de la coordination interpersonnelle autour d'objets partagés, il permettra également d'explorer des solutions techniques en vue de mettre à disposition des utilisateurs un dispositif qui favorise au mieux l'interaction dans les espaces numériques partagés.

7. Etude expérimentale : objets communs et objets privés[41]

7.1. Population

24 sujets, répartis en 3 groupes expérimentaux de 8 sujets chacun (soit 4 binômes dans chaque groupe), ont participé à cette expérience. Les sujets sont des étudiants de l'Université de Technologie de Compiègne âgés entre 18 et 25 ans. Ils n'ont pas de connaissance particulière du dispositif et n'ont pas connaissance des éléments théoriques mobilisés dans l'expérience.

Les groupes expérimentaux sont définis en fonction de l'ordre de passation des conditions expérimentales décrites ci-après.

7.2. Matériel

Comme lors de l'étude expérimentale précédente, le dispositif Tactos est utilisé. Chaque utilisateur est donc installé devant un poste de travail qui comprend :

- un ordinateur-client avec le logiciel Tactos ;
- un effecteur souris pour contrôler le déplacement du curseur et avec une accélération logicielle contrôlée et identique pour tous les sujets ;
- un stimulateur tactile composé de deux cellules Braille piézoélectriques (soit 16 picots), configurés adéquatement selon les conditions expérimentales décrites ci-après ;
- un casque audio, qui permet de délivrer des consignes simultanément aux deux sujets participant à une passation expérimentale.

Ici, le dispositif expérimental comporte trois ordinateurs : un serveur qui va gérer les connexions et deux ordinateurs-clients destinés à chacun des deux sujets.

L'espace sur lequel se déplacent les sujets est un espace horizontal unidimensionnel de 400 pixels qui se boucle sur lui-même. Il s'agit ici encore d'un tore géométrique dans lequel le sujet n'a pas la conscience de bords ou de limites.

[41] Nous remercions vivement Gabrielle Le Bihan pour avoir participé à la réalisation de cette expérience.
Par ailleurs, cette étude a donné lieu à une communication orale :
Deschamps, L., (2013), Minimalist perceptual crossing: The case of triadic interactions, *Séminaire Phiteco (Philosophie, Technologie, Cognition), From Perceptual Interactions to Extended Cognition*, Université de Technologie de Compiègne, Compiègne, France, 21-25 janvier 2013

A l'aide de l'effecteur souris, les sujets déplacent un avatar de 2 pixels de long (le corps-image), situé au centre d'un corps-percevant d'une longueur de 16 pixels dont la sensibilité et la configuration varient en fonction des conditions expérimentales.

Pour chaque essai, trois objets fixes d'une longueur de 2 pixels sont disposés le long de l'espace unidimensionnel. La position de ces objets est différente pour chacun des 8 essais des 3 conditions expérimentales. Parmi ces trois objets, un seul est perceptible conjointement et simultanément par les deux sujets : il s'agit d'un objet « public ». Les deux autres objets sont des objets « privés », c'est-à-dire que chacun d'eux n'est perceptible respectivement que par un des sujets évoluant dans l'espace partagé. En d'autres termes, pour chaque sujet, deux objets sont présents à chaque essai : un objet public et un objet privé (voir Figure 64).

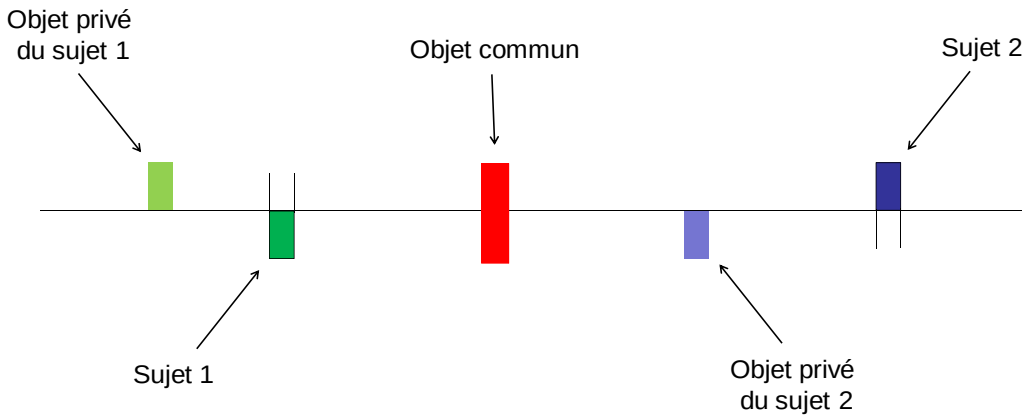


Figure 64 : Schéma de la situation expérimentale

7.3. Procédure

Dans chaque groupe expérimental, 8 sujets sont répartis en 4 binômes. Chaque binôme passe 3 conditions expérimentales, chacune des conditions expérimentales étant définie par une configuration différente des corps-percevant et de leur sensibilité associée.

Dans la condition « Mono », le corps-percevant est composé d'un unique champ récepteur élémentaire d'une longueur de 16 pixels. La rencontre de ce champ récepteur avec les objets ou le corps-image de l'autre participant active la montée simultanée des 16 picots de la matrice de stimulateurs tactiles. En ce sens, la perception des objets et du corps-image de l'autre conduisent à des stimulations tactiles indifférenciées.

Dans la condition « Différenciée », le corps-percevant est composé d'un unique champ récepteur élémentaire d'une longueur de 16 pixels, comme pour la condition « Mono ». En revanche, la rencontre de ce champ récepteur avec les objets ou avec le corps-image de l'autre participant produit une activation différenciée des stimulateurs tactiles. En effet, la rencontre

de ce champ récepteur avec un objet (qu'il soit public ou privé) active la montée simultanée des 4 picots supérieurs de la matrice de 16 stimulateurs. A l'inverse, la rencontre de ce champ récepteur avec le corps-image de l'autre participant active la montée simultanée des 4 picots inférieurs de la matrice de 16 stimulateurs. De cette façon, il est possible de percevoir simultanément un objet et le corps-image de l'autre participant : dans ce cas, les 4 picots supérieurs et les 4 picots inférieurs sont activés simultanément.

Dans la condition « Parallélisme », le corps-percevant est composé de 4 champs récepteurs élémentaires de 4 pixels chacun (soit une longueur totale de 16 pixels). Chaque champ récepteur élémentaire est couplé de façon topologique à chaque rangée de stimulateurs tactiles. Ainsi, lorsque le champ récepteur situé le plus à gauche du corps-percevant rencontre un objet (qu'il soit public ou privé) ou le corps-image de l'autre participant, les quatre stimulateurs tactiles situés le plus à gauche de la matrice de stimulateurs (soit la première rangée) s'activent simultanément. De cette façon, cette condition expérimentale implique un parallélisme des champs récepteurs avec la matrice de stimulateurs tactiles. Ici, les stimulations ne sont plus différenciées, mais il est possible de percevoir simultanément un objet et le corps-image de l'autre participant, à condition que les deux sources de stimulation ne se trouvent pas au même endroit et que la distance qui les sépare soit inférieure à la taille du corps-percevant (soit 16 pixels).

Tous les binômes passent successivement les trois conditions expérimentales, mais dans un ordre différent : les groupes expérimentaux sont ainsi constitués en fonction de l'ordre de passation des conditions. Les binômes du premier groupe expérimental commencent par la condition « Mono », puis passent la condition « Différenciée » et enfin la condition « Parallélisme ». Les binômes du second groupe expérimental commencent par la condition « Différenciée », puis passent la condition « Parallélisme » et enfin la condition « Mono ». Enfin, les binômes du troisième groupe expérimental commencent par la condition « Parallélisme », puis passent la condition « Mono » et enfin la condition « Différenciée ».

Chaque condition expérimentale est constituée de 13 essais, dont 5 essais de familiarisation et 8 essais expérimentaux.

Les passations se font en binôme. Les sujets sont accueillis dans une salle, et le dispositif leur est brièvement présenté, ainsi qu'une consigne globale décrivant l'expérience à laquelle ils vont participer. Dans cette consigne, les sujets sont informés que l'expérience va consister à interagir tactilement avec l'autre participant dans un espace partagé où sont disposés des objets fixes. Le but est de deviner lequel des objets est un objet commun, c'est-à-dire lequel des objets perçus est également perceptible par l'autre participant. Il est également

annoncé que l'expérience est divisée en trois phases, chacune de ces phases étant décomposée en deux parties distinctes : une phase de familiarisation et une phase de test. Après cette consigne, les sujets sont installés dans deux salles adjacentes, de façon à isoler les participants et à éviter tout indice perceptif autre que les stimulations tactiles résultant de la rencontre entre les corps-image dans l'environnement numérique (tels que les bruits de la souris sur la table, ceux des stimulateurs tactiles ou encore les réactions des sujets lors de la rencontre,…). Lorsque les sujets sont assis devant leur dispositif Tactos respectif, nous leur demandons d'installer leur casque audio sur leurs oreilles de façon à recevoir des informations orales concernant chaque essai (le début de l'essai, la fin de l'essai, et le moment de chaque essai expérimental où les sujets pourront cliquer de façon à désigner l'objet commun). Les sujets n'ont pas les yeux bandés pendant l'expérience, mais aucune information n'est visible sur leur écran. Lorsque les sujets sont prêts le premier essai de la première condition de chaque groupe expérimental commence.

Le premier essai est un essai de familiarisation avec le dispositif et avec la condition expérimentale en cours. Les sujets sont connectés dans des espaces privés respectifs où est disposé un objet fixe de 2 pixels de long. La consigne pour cet essai indique au sujet qu'en déplaçant la souris sur un axe horizontal, ils peuvent rencontrer un objet fixe. Cette rencontre produit l'activation des picots selon la condition expérimentale en cours, et qu'il faut bien faire attention à la façon dont les picots se lèvent, car ils se lèveront de façon différente lors de chaque condition. Pour le premier essai de la première condition passée par les sujets, il est également demandé à ce que les sujets se déplacent dans une direction donnée jusqu'à obtenir deux stimulations successives. Cette étape consiste à insister sur le fait que les deux stimulations n'indiquent pas la présence de deux objets, mais qu'un seul et même objet a été croisé deux fois. Ainsi, la distance entre deux stimulations équivaut à la taille totale de l'espace. Il est enfin indiqué que cette taille sera identique pour tous les essais de l'ensemble de l'expérience. Cet essai n'est pas chronométré, et dure tant que les sujets le jugent utile. L'expérimentateur s'assure que les sujets ont bien compris, et leur indique le cas échéant que l'essai suivant va commencer.

Le deuxième essai est un essai de familiarisation avec la possibilité d'interagir avec l'autre participant de l'expérience. Les sujets sont alors connectés dans un espace partagé où aucun objet n'est présent. La seule source de stimulation tactile résulte donc de la rencontre entre les deux participants. Les sujets sont informés que deux rôles leur seront attribués alternativement par l'intermédiaire du casque audio au début de l'essai. Si le sujet A est guideur, le sujet B est suiveur et inversement. Le sujet guideur est informé qu'un évènement

sonore est disposé à un endroit de l'espace et que lui seul peut l'entendre. Son objectif est de localiser cet évènement, et de tenter de guider l'autre participant à cet endroit. Le sujet suiveur est informé que l'autre participant entend un son à un endroit de l'espace, et que son objectif est de suivre le guideur jusqu'à ce son. Cet essai n'est pas chronométré, et s'arrête lorsque l'expérimentateur constate que les deux sujets sont arrêtés sur l'évènement sonore ou oscillent autour de lui. Lorsque c'est le cas, l'expérimentateur lance l'essai suivant, qui est identique si ce n'est que l'évènement sonore est déplacé et que les rôles s'inversent : le guideur devient suiveur et vice-versa. Cet essai est alors reproduit quatre fois, ce qui implique que chaque sujet est alternativement deux fois guideur et deux fois suiveur.

Le sixième essai est le premier essai de test. Les sujets sont alors informés que l'expérience à proprement parler commence. L'objectif consiste alors à explorer un espace partagé ou chacun des deux sujets pourra sentir la présence de l'autre, ainsi que la présence de deux objets fixes, suivant les mêmes conditions expérimentales que les 5 essais de familiarisation précédents. Parmi ces deux objets, un objet est commun, l'autre est privé. Le but du jeu pour chaque sujet est d'interagir avec l'autre participant pendant 1 minute, et d'essayer de deviner lequel des deux objets présents est l'objet commun. A la fin de cette minute, une consigne délivrée via le casque audio indique aux sujets qu'ils ont 15 secondes pour cliquer sur l'objet commun. A la fin de ces 15 secondes, l'expérimentateur lance l'essai suivant, et ce, jusqu'à la réalisation de 8 essais consécutifs.

A la fin des 8 premiers essais expérimentaux, une pause de deux minutes est proposée aux sujets, durant laquelle ils ne peuvent pas interagir. Ensuite, la deuxième phase de l'expérience peut commencer. La deuxième condition expérimentale commence alors, suivant le même scénario que la première condition passée (soit 5 essais de familiarisation suivis de 8 essais expérimentaux). Les sujets sont invités à faire attention aux modifications du type de stimulations reçues lors des essais de familiarisation, de façon à s'assurer que la nouvelle condition expérimentale est constatée par les deux sujets.

Cette deuxième phase est suivie d'une pause de deux minutes, qui fait suite à une troisième phase où la troisième condition expérimentale est menée. A la fin de l'expérience, nous réunissons les sujets de façon à procéder à un débriefing, où les sujets peuvent nous faire part de leurs difficultés éventuelles ou expliciter leur stratégie.

7.4. Hypothèses opérationnelles

- Quelle que soit la configuration des corps-image/corps-percevant, les sujets cliqueront davantage sur l'objet commun que sur les objets privés. Nous nous attendons à ce que

cette réussite soit rendue possible en vertu de l'émergence et de la stabilisation d'une dynamique interpersonnelle autour de l'objet commun. En ce sens, nous nous attendons à ce que la réussite soit collective : la fréquence des cas où les deux sujets cliqueront mutuellement sur l'objet commun sera plus élevée que la fréquence des cas où un seul des deux cliquera sur l'objet commun.

- Malgré cette réussite globale, nous nous attendons à ce que les différentes configurations de corps-image/corps-percevant aient pour effet de faire varier la nature de cette dynamique, ce qui aura une incidence sur la réussite de la tâche. En outre, la condition « Différenciée » et la condition « Parallélisme », en tant qu'elles ont vocation à surmonter des difficultés inhérentes à la condition « Mono » (indifférenciation autre/objet et impossibilité de percevoir les deux en même temps), devront conduire à plus de clics sur l'objet commun que de mauvais clics.

- Toutefois, la condition « Mono » étant la condition la plus minimaliste, elle aura pour effet de forcer le déploiement des activités perceptives, ce qui pourra entraîner un nombre de clics sur l'objet commun supérieur au nombre de mauvais clics. En outre, l'analyse des trajectoires perceptives issues de cette condition et dans les cas de réussite devrait nous permettre d'isoler des paramètres objectifs pour tenter d'expliquer la nature de la réussite en termes de coordination des activités perceptives.

- Enfin, quelle que soit la condition expérimentale, nous nous attendons à ce que le nombre de clics sur l'objet commun soit supérieur dans le deuxième bloc d'essais (essais 5 à 8 de chaque condition) que dans le premier bloc d'essais (essais 1 à 4 de chaque condition). Ce résultat pourra être expliqué conjointement par la stabilisation de la dynamique interpersonnelle et par l'appropriation individuelle du système de couplage lui-même, qui permettra un meilleur engagement au fur et à mesure des essais.

7.5. Résultats

7.5.1. Analyse des réponses individuelles[42]

Dans un premier temps, nous nous intéressons aux types de clics indépendamment des conditions expérimentales et de l'ordre de passation. Les clics peuvent être répartis en quatre catégories : les clics sur l'objet commun, les clics sur l'objet privé du sujet, les clics sur le

[42] Par « Analyse des réponses individuelles », nous n'entendons pas une analyse de profils individuels de réponse, mais une analyse des réponses fournies par les sujets indépendamment des réponses de l'autre.

corps-image de l'autre sujet[43] et les clics « vides », c'est à-dire les clics situés ni sur un objet ni sur le corps-image de l'autre sujet. Nous considérons ici le nombre moyen de clics pour chaque catégorie et pour l'ensemble des sujets.

Globalement, nous observons que le nombre moyen de clics sur l'objet commun est supérieur au nombre moyen de clics sur l'objet privé du sujet (d = 9), au nombre moyen de clics sur le corps-image de l'autre sujet (d = 11.87) et au nombre moyen de clics sur des emplacements vides (d = 12.16). Par ailleurs, nous observons que le nombre moyen de clics sur l'objet commun est également supérieur au nombre moyen d'occurrences où les sujets ne cliquent pas (d = 6.91) (voir Figure 65). La différence globale est significative (test de Kruskall-Wallis : $H_{(4,120)} = 87{,}42915$; $p < 0.0001$).

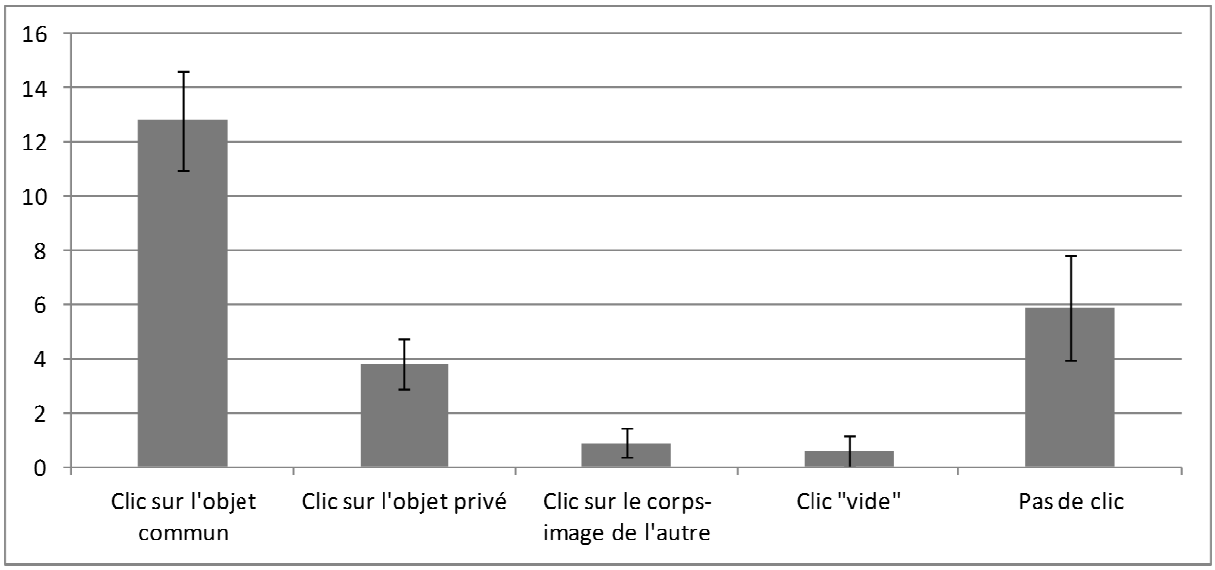


Figure 65 : Nombre moyen de clics par type de clic

De façon à simplifier les analyses, nous avons exclu les essais où les sujets n'ont pas cliqué[44], et nous regroupons les clics qui ne sont pas situés sur l'objet commun en une catégorie globale « Mauvais clics ». Ici encore, le nombre moyen de clics sur l'objet commun est supérieur au nombre moyen de mauvais clics (d = 7.45). Cette différence est significative ($t_{(23)} = 7.8369$; $p < 0.0001$) et peut même être considérée comme notable d'après une analyse bayésienne ($Pr\ |\delta/\sigma| > .66 = .99$).

[43] D'une façon générale, ces clics sont situés aux alentours de l'objet privé de l'autre participant : par exemple, si le sujet 1 explore son objet privé et que dans le même temps, les deux sujets interagissent dans cette zone, les stimulations reçues par le sujet 2 ont pu lui donner l'illusion de la présence d'un objet à cet endroit.

[44] Cette exclusion nous paraît justifiée dans la mesure où une absence de clic ne nous permet pas de savoir si le sujet n'a pas souhaité fournir de réponse, ou s'il n'a tout simplement pas eu le temps de cliquer dans le temps imparti. Par ailleurs, il nous semble plus opportun de comparer les réponses effectives : soit la différence entre les clics sur l'objet commun et les mauvais clics.

Toutefois, nous constatons que l'effet d'interaction entre la qualité de la réponse (bons clics vs mauvais clics) et les conditions expérimentales est statistiquement significatif ($F_{(2,69)}$ = 31.0327 ; p < 0.0001). En effet, si la différence globale entre le nombre moyen de clics sur l'objet commun et le nombre moyen de mauvais clics suit la même évolution dans chacune des trois conditions expérimentales, celle-ci se manifeste dans des proportions différentes : la différence observée est faible pour la condition « Mono » (d = 0.71) et pour la condition « Parallélisme » (d = 1.25), alors quelle est plus importante pour la condition « Différenciée » (d = 5.5) (voir Figure 66).

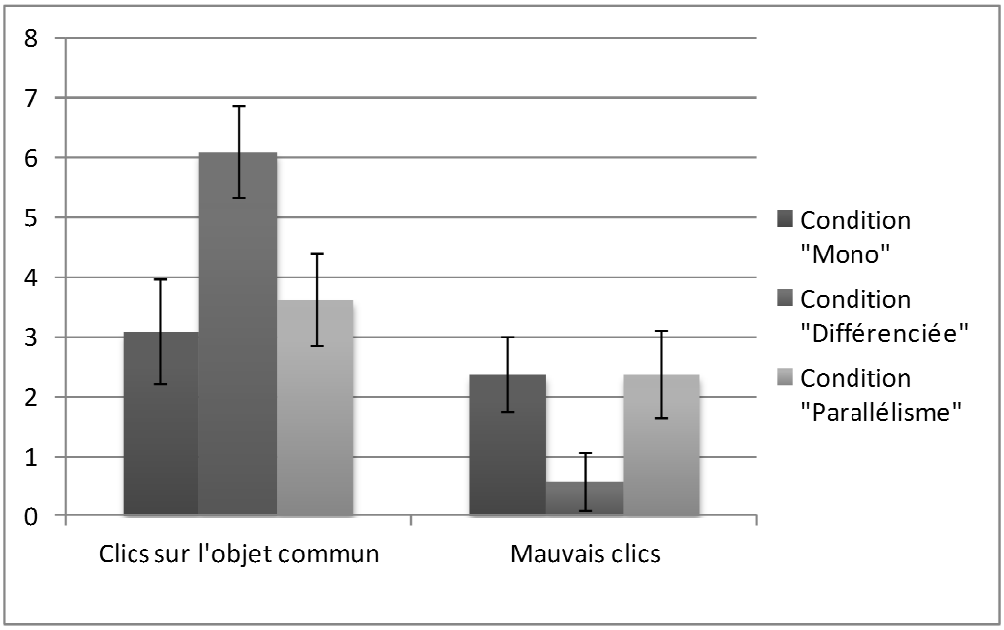


Figure 66 : Nombre moyen de clics sur l'objet commun et de mauvais clics en fonction des 3 conditions expérimentales

D'un point de vue statistique, la différence n'est pas significative pour la condition « Mono » ($t_{(23)}$ = 1.4433 ; p = 0.1624), même si elle ne peut pas être considérée comme négligeable (Pr |δ/σ| <.33 = .58) ; la différence est significative pour la condition « Différenciée » ($t_{(23)}$ = 11.9978 ; p < 0.0001), et peut même être considérée comme notable (Pr |δ/σ| >.66 = .99) ; enfin, la différence est significative pour la condition « Parallélisme » ($t_{(23)}$ = 2.6979 ; p = 0.0128), mais elle ne peut pas être considérée comme notable (Pr |δ/σ| >.66 = .29).

Si nous centrons notre analyse sur le nombre moyen de clics sur l'objet commun, nous constatons un effet global des conditions expérimentales (test de Kruskall-Wallis : $H_{(2,72)}$ = 28,16017 ; p < 0.0001). De façon plus spécifique, nous observons que le nombre moyen de clics sur l'objet commun est plus faible dans la condition « Mono » que dans la condition

« Différenciée » (d = 3) (voir Figure 66). Cette différence est significative (test de Mann & Withney : U = 60,5 ; $Z_{ajusté}$ = - 4.74722 ; p < 0.0001), et l'effet peut même être considéré comme fort (g de Hedge = 1.82). De la même façon, le nombre moyen de clics sur l'objet commun est plus faible dans la condition « Parallélisme » que dans la condition « Différenciée » (d = 2.45) (voir Figure 66). Cette différence est significative (test de Mann & Withney : U = 81 ; $Z_{ajusté}$ = - 4.32365 ; p = 0.00015), et l'effet peut même être considéré comme fort (g de Hedge = 1.60). Enfin, le nombre moyen de clics sur l'objet commun est plus faible dans la condition « Mono » que dans la condition « Parallélisme » (d = 0.54) (voir Figure 66). Toutefois, cette différence n'est pas significative (test de Mann & Withney : U = 242 ; $Z_{ajusté}$ = - 0.963283 ; p = 0.3354). Il est à noter que cet effet des conditions expérimentales sur le nombre moyen de clics sur l'objet commun se retrouve quel que soit l'ordre de passation. En effet, la différence globale entre les ordres de passation n'est pas significative ($F_{(2,69)}$ = 0.3518 ; p = 0.7047).

D'une façon générale, nous observons un effet d'apprentissage à l'intérieur des conditions expérimentales. En effet, toutes conditions confondues, le nombre moyen de clics sur l'objet commun est supérieur au cours des quatre derniers essais comparativement aux quatre premiers essais (d = 1.20). Cette différence est significative ($t_{(23)}$ = 2.3499 ; p = 0.0277), mais elle ne peut pas être considérée comme notable (Pr $|\delta/\sigma|$ >.66 = .19). Parallèlement, le nombre moyen de mauvais clics est plus important dans le premier bloc d'essais que dans le second (d = 1.08). Cette différence est significative ($t_{(23)}$ = 2.8487 ; p = 0.0091), mais l'effet ne peut pas être considéré comme notable (Pr $|\delta/\sigma|$ >.66 = .34) (voir Figure 67).

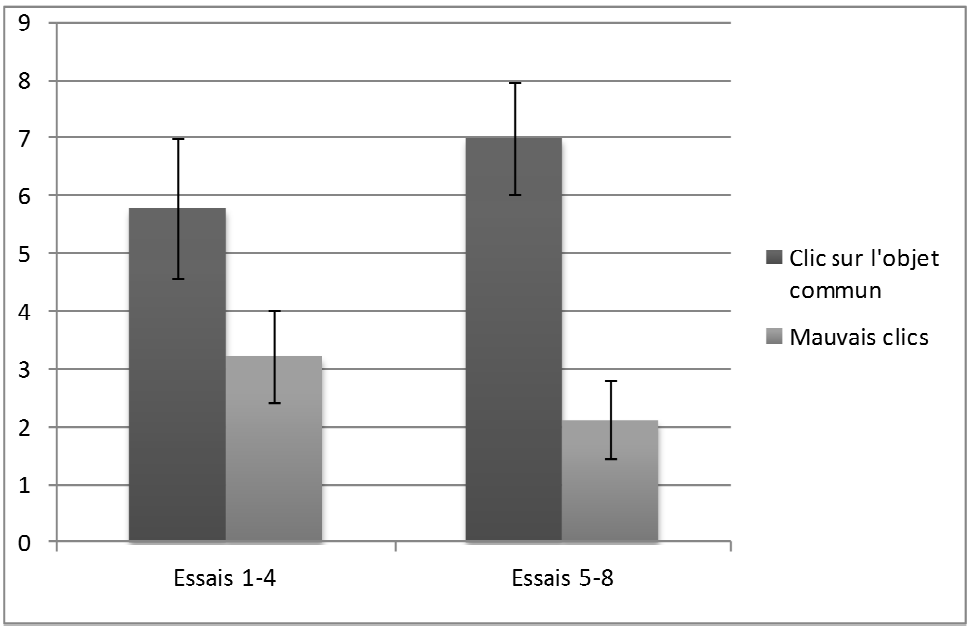


Figure 67 : Nombre moyen de clics sur l'objet commun et de mauvais clics en fonction des deux blocs d'essais (Essais 1 à 4 versus Essais 5 à 8)

De façon descriptive, l'effet de l'apprentissage est très faible en condition « Différenciée » (d = 0.08), plus important pour la condition « Mono » (d = 0.33) ainsi que pour la condition « Parallélisme » (d = 0.79). Toutefois, nous ne pouvons pas conclure à un effet d'interaction entre le facteur « Apprentissage » et le facteur « Conditions expérimentales » ($F_{(2,69)} = 1.4932$; $p = 0.2318$).

7.5.2. Analyse des réponses collectives

Après avoir analysé les réponses individuelles des sujets indépendamment des réponses de leurs partenaires, nous nous intéressons maintenant aux données collectives, c'est-à-dire à la qualité de la réponse du binôme pour chaque essai. Dans ce cadre, nous considérons trois cas de figure, correspondant à trois modalités de réponse par binôme : les clics mutuels sur l'objet commun (c'est-à-dire les cas où les deux sujets ont cliqué sur l'objet commun), les clics solitaires (c'est-à-dire les cas où seul un des deux sujets a cliqué sur l'objet commun, l'autre ayant cliqué ailleurs ou n'ayant pas cliqué), et les autres cas (c'est-à-dire les cas où les deux sujets ont cliqué ailleurs que sur l'objet commun, ou un des deux a cliqué ailleurs, l'autre n'ayant pas cliqué). Notons que dans ce dernier cas de figure, les cas où les deux sujets n'ont pas cliqué n'ont pas été pris en compte[45].

[45] Nous avons exclu ce cas de figure pour les mêmes raisons que lors de l'analyse des réponses individuelles (voir note 44, p. 180)

Globalement, nous observons que la fréquence de clics mutuels est supérieure à la fréquence de clics solitaires (d = 0.16), et à la fréquence des autres cas (d = 0.11). De plus, la fréquence de clics solitaires est inférieure à la fréquence des autres cas (d = 0.04) (voir Figure 68). La différence globale est significative ($F_{(2,33)}$ = 6.7480 ; p = 0.0035).

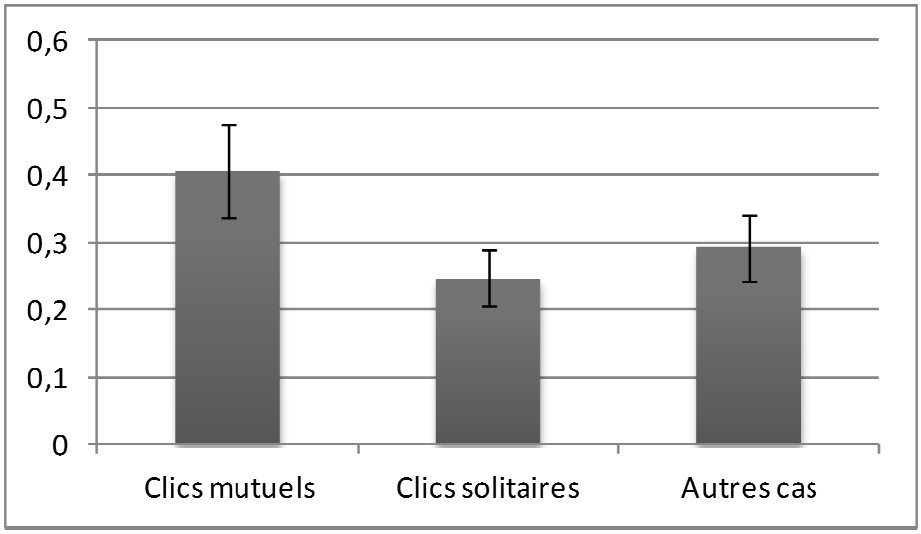


Figure 68 : Fréquence des trois modalités de réponse pour l'ensemble des binômes

De façon plus spécifique, la différence entre la fréquence de clics mutuels et la fréquence de clics solitaires est significative ($t_{(22)}$ = 3.3663 ; p = 0.0028), et peut même être considérée comme notable (Pr |δ/σ| >.66 = .94). De la même façon, la différence entre la fréquence de clics mutuels et la fréquence des autres cas est significative ($t_{(22)}$ = 2.3276 ; p = 0.0295), mais cette fois, elle ne peut pas être considérée comme notable (Pr |δ/σ| >.66 = .74). En revanche, la différence entre la fréquence de clics solitaires et la fréquence des autres cas n'est pas significative ($t_{(22)}$ = 1.2269 ; p = 0.2328), même si elle ne peut pas être considérée comme négligeable (Pr |δ/σ| <.33 = .33).

De façon descriptive, nous observons que les différences entre les trois modalités de réponse s'expriment de façon différente selon les conditions expérimentales (voir Figure 69). En effet, l'effet d'interaction entre les modalités de réponse des binômes et le facteur « Conditions expérimentales » est statistiquement significatif ($F_{(4,66)}$ = 12.9981 ; p < 0.0001).

En ce qui concerne la condition « Mono », nous observons que la fréquence de clics mutuels est inférieure à la fréquence de clics solitaires (d = 0.08), elle-même inférieure à la fréquence des autres cas (d = 0.08) (voir Figure 69). Toutefois, la différence globale n'est pas significative ($F_{(2,22)}$ = 2.0230 ; p = 0.1561).

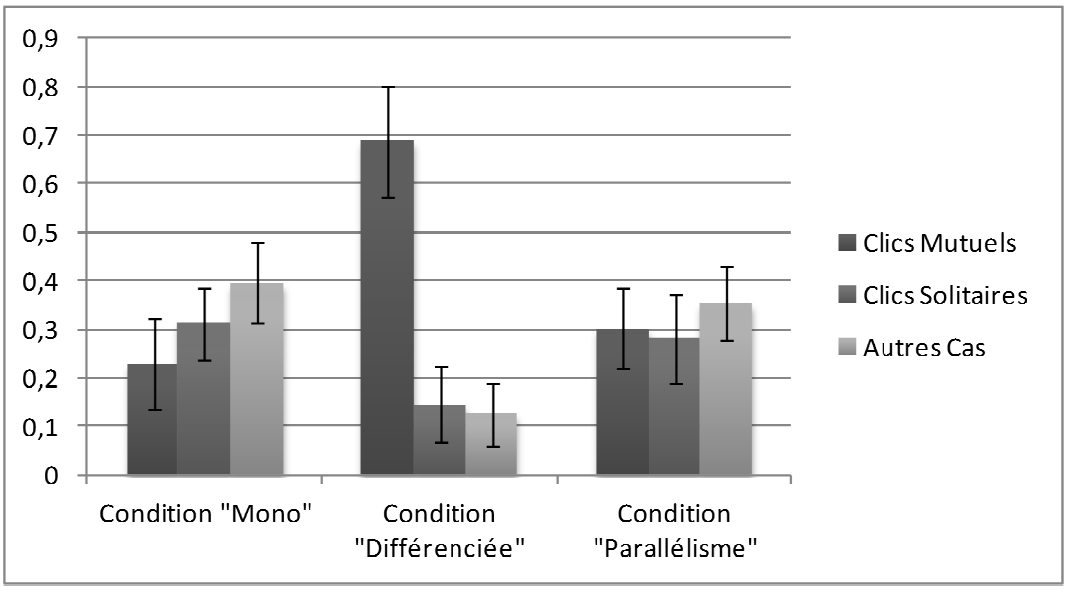


Figure 69 : Fréquence des trois modalités de réponse en fonction des conditions expérimentales

Concernant la condition « Différenciée », la fréquence de clics mutuels est supérieure à la fréquence de clics solitaires (d = 0.54), et à la fréquence des autres cas (d = 0.56). Enfin, la fréquence de clics solitaires est supérieure à la fréquence des autres cas (d = 0.02). La différence globale est significative ($F_{(2,22)}$ = 26.1250 ; p < 0.0001). De façon plus spécifique, la différence entre la fréquence de clics mutuels et la fréquence de clics solitaires d'une part, et la fréquence des autres cas d'autre part sont toutes deux significatives (respectivement $t_{(11)}$ = 5.1719 ; p = 0.0003 et $t_{(11)}$ = 6.0512 ; p < 0.0001), ces effets pouvant même être considérés comme notables (respectivement Pr |δ/σ| >.66 = .97 et Pr |δ/σ| >.66 = .99). En revanche, la différence entre la fréquence de clics solitaires et la fréquence des autres clics n'est pas significative ($t_{(11)}$ = 0.3403 ; p = 0.7401), même si l'effet ne peut pas être considéré comme négligeable (Pr |δ/σ| <.33 = .72).

Enfin, en ce qui concerne la condition « Parallélisme », la fréquence de clics mutuels est supérieure à la fréquence de clics solitaires (d = 0.02), mais est inférieure à la fréquence des autres cas (d = 0.05). Enfin, la fréquence de clics solitaires est inférieure à la fréquence des autres cas (d = 0.07). Toutefois, la différence globale n'est pas significative ($F_{(2,22)}$ = 0.4658 ; p = 0.6337).

Si nous nous centrons sur la fréquence de clics mutuels en fonction des conditions expérimentales, nous observons que la fréquence de clics mutuels est plus importante en condition « Différenciée » qu'en condition « Mono » (d = 0.46), et « Parallélisme » (d = 0.38). Par ailleurs, la fréquence de clics mutuels est plus importante en condition « Parallélisme » qu'en condition « Mono » (d = 0.07) (voir Figure 69). D'une façon

statistique, nous pouvons conclure à un effet des conditions expérimentales sur la fréquence de clics mutuels ($F_{(2,33)} = 18.8427$; $p < 0.0001$). De façon plus spécifique, la différence est significative entre la condition « Mono » et la condition « Différenciée » d'une part ($t_{(22)} = 5.3216$; $p < 0.0001$), et entre la condition « Différenciée » et la condition « Parallélisme » d'autre part ($t_{(22)} = 4.7374$; $p < 0.0001$). Notons que ces différences peuvent même être considérées comme notables (respectivement Pr $|\delta/\sigma| > .66 = .98$ et Pr $|\delta/\sigma| > .66 = .98$). Par ailleurs, la différence entre la condition « Mono » et la condition « Parallélisme » n'est pas significative ($t_{(22)} = 1.0037$; $p = 0.3264$), même si elle ne peut pas être considérée comme négligeable (Pr $|\delta/\sigma| < .33 = .39$).

Ces différences étant constatées, nous souhaitons ici souligner une remarque importante. Paradoxalement, les clics solitaires nous semblent en effet révélateurs d'une dynamique collective. Nous nous intéressons au type de réponse d'un sujet B lorsque le sujet A a cliqué correctement. D'une façon globale, nous observons que la proportion de mauvais clics de B est très faible (5.90%). Par conséquent, B n'a pas cliqué dans une grande majorité des cas (94.10%). En outre, en ce qui concerne les cas où ni A ni B n'ont cliqué sur l'objet commun, nous observons que lorsque A a effectué un mauvais clic, B n'a pas cliqué dans la majorité des cas (89.93% des cas). En outre, dans les 10.07% des cas où les deux sujets ont cliqué ailleurs que sur l'objet commun, ils ont cliqué au même endroit dans 75.86% des cas (soit sur l'objet privé de l'un des deux partenaires, soit dans de plus rares cas sur un emplacement vide). Ces observations se retrouvent dans chacune des trois conditions expérimentales (voir Tableau 13).

Cas	Réponse du sujet A	Réponse du sujet B	Condition « Mono »	Condition « Différenciée »	Condition « Parallélisme »
Clics Solitaires	Clic correct	Pas de clic	90.62%	98.96%	92.71%
	Clic correct	Mauvais clic	9.38%	1.04%	7.29%
Autres cas	Mauvais clic	Pas de clic	88.54%	98.96%	82.39%
	Mauvais clic	Mauvais clic	11.46%	1.04%	17.71%

Tableau 13 : Proportion des types de réponse collective dans les cas des clics solitaires et des autres cas pour chaque condition expérimentale.

Par ailleurs, d'une façon globale, nous observons que la fréquence de clics mutuels est plus élevée dans le deuxième bloc d'essais que dans le premier bloc d'essais ($d = 0.04$) (voir Figure 70). Cette différence n'est pas significative ($t_{(22)} = 0.5519$; $p = 0.5866$), même si elle ne peut pas être considérée comme négligeable (Pr $|\delta/\sigma| < .33 = .52$). Par ailleurs, la fréquence

de clics solitaires est plus élevée dans le deuxième bloc d'essais que dans le premier bloc d'essais (d = 0.12). Cette différence est cette fois significative ($t_{(22)}$ = 2.8591 ; p = 0.0091), mais elle ne peut pas être considérée comme notable (Pr |δ/σ| >.66 = .86). En revanche, la fréquence des autres cas est moins élevée dans le deuxième bloc d'essais que dans le premier bloc d'essais (d = 0.13). Cette différence est significative ($t_{(22)}$ = 2.3307 ; p = 0.0293), mais elle ne peut pas être considérée comme notable (Pr |δ/σ| >.66 = .74).

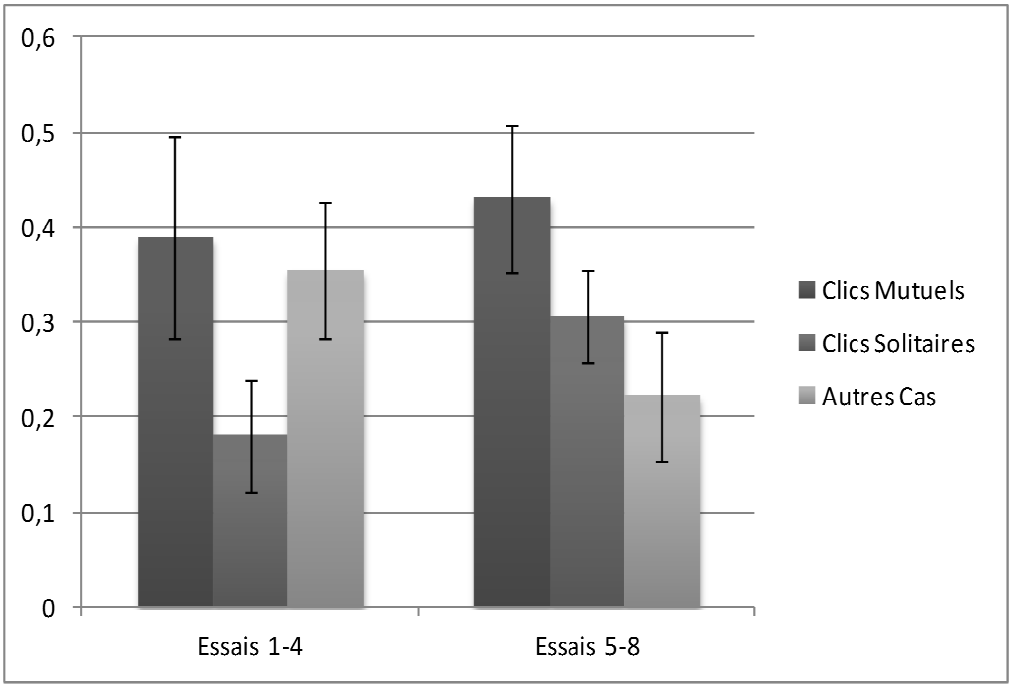


Figure 70 : Fréquence des trois modalités de réponse en fonction des deux blocs d'essais

Pour finir, nous regardons si les conditions expérimentales ont un effet sur la fréquence de clics mutuels en fonction des deux blocs d'essais. De façon descriptive, nous observons que les fréquences de clics mutuels ne présentent que très peu de différences en ce qui concerne la condition « Mono » et la condition « Différenciée » en fonction des deux blocs d'essais (respectivement d = 0.02 et d = 0.04) (voir Figure 71). En revanche, la fréquence de clics mutuels est plus élevée dans le deuxième bloc d'essai que dans le premier bloc d'essais en ce qui concerne la condition « Parallélisme » (d = 0.10). Toutefois, l'effet d'interaction entre les conditions expérimentales et les deux blocs d'essais n'est pas significatif ($F_{(2,33)}$ = 0.4304 ; p = 0.6538).

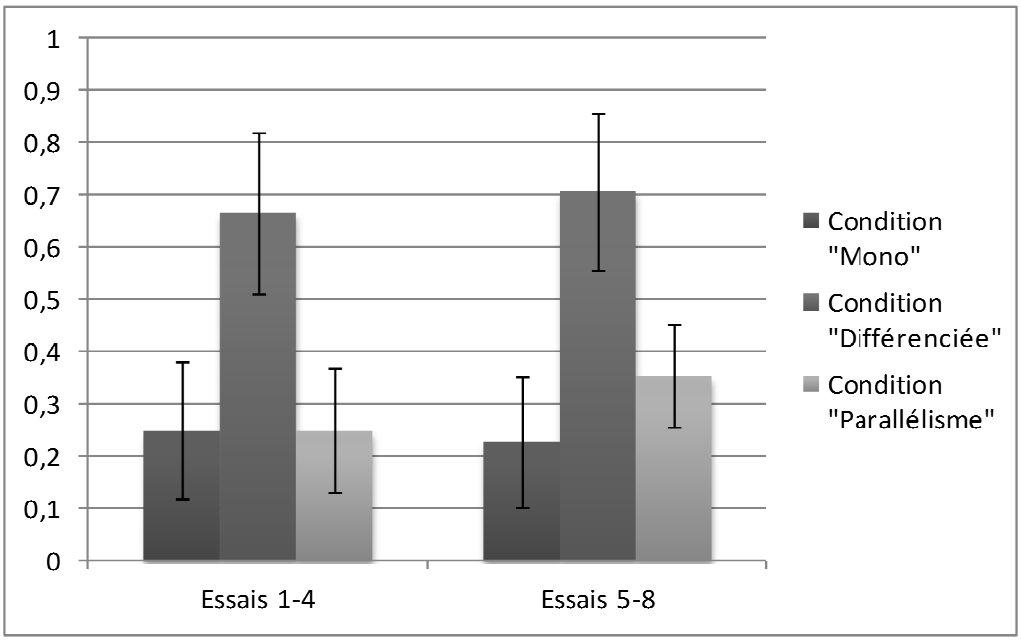


Figure 71 : Fréquence de clics mutuels en fonction des deux blocs d'essais et des conditions expérimentales

7.6. Analyse des trajectoires

Comme pour notre expérience concernant les interactions dyadiques dans le contexte de la discrimination Humain/Robot, une analyse des trajectoires perceptives déployées par les sujets a été menée[46]. Ici encore, nous avons guidé nos choix d'indicateurs via les informations fournies par les sujets lors du débriefing de l'expérience, couplée à une analyse qualitative des graphiques représentant les différents essais (voir des exemples en Annexe D, p. 294). Le choix de différents indicateurs permet alors de procéder à une analyse quantitative systématique. La richesse des données a conduit à analyser de très nombreux indicateurs, mais par souci de clarté, nous ne relaterons ici que les résultats principaux.

Trois axes distincts ont guidé notre analyse : l'interaction des sujets avec les deux types d'objets (Objet commun vs Objet privé) ; l'interaction entre les sujets autour des deux types d'objets ; et l'interaction entre les sujets en fonction de la qualité de leur réponse individuelle (Clics sur l'objet commun vs Mauvais clics), et collective (Clics mutuels, Clics Solitaires et Autres cas).

7.6.1. Interaction des sujets avec les deux types d'objets

Dans un premier temps, nous avons isolé un premier groupe d'indicateurs, de façon à décrire l'interaction individuelle des sujets avec les deux types d'objets (Objet commun vs

[46] Pour plus de détails concernant la méthode d'analyse utilisée pour l'étude des trajectoires, veuillez vous reporter à l'Annexe B, p. 247.

Objet privé), et ce, quelle que soit leur réponse à la fin des essais. Nous nous intéressons ici à la durée totale pendant laquelle les sujets se sont trouvés dans un intervalle de 40 pixels autour de chacun des deux types d'objets, ainsi qu'au nombre de stimulations[47] dues à la rencontre de chacun d'eux.

D'une façon globale, nous observons que les sujets ont en moyenne passé plus de temps autour de l'objet commun qu'autour des objets privés (d = 15438.81 ms) (voir Figure 72). Cette différence est significative (test de Mann & Withney : U = 618.5 ; $Z_{ajusté}$ = 7.885277 ; p < 0.0001), et l'effet peut même être considéré comme fort (g de Hedge = 1.49). Par ailleurs, cet effet se retrouve dans chacune des trois conditions expérimentales (voir Figure 72 et Tableau 14).

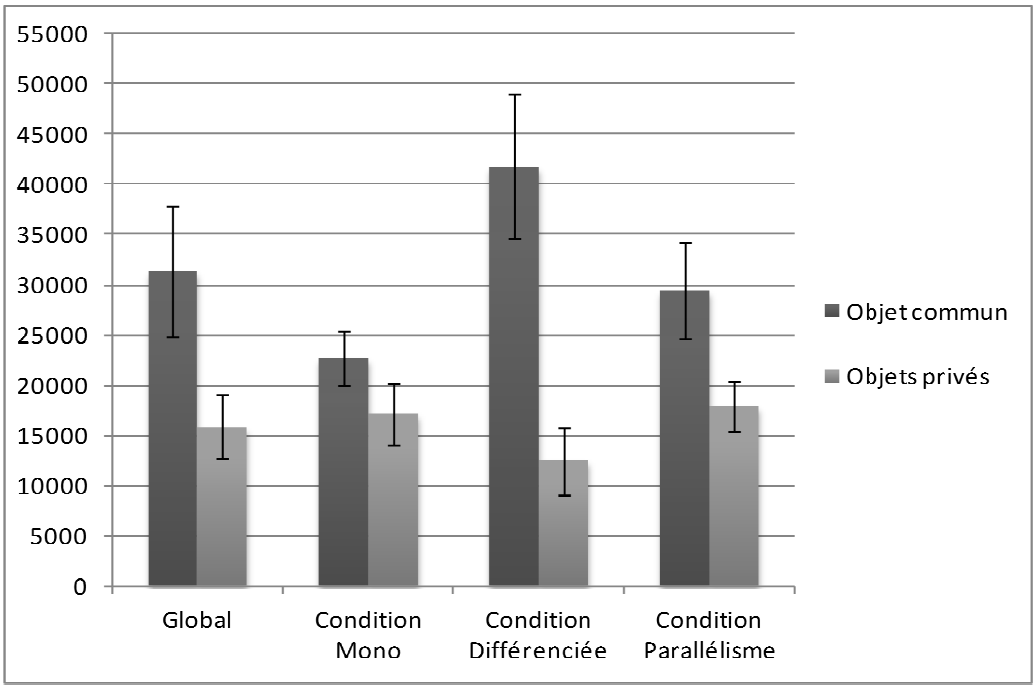


Figure 72 : Durée totale moyenne (en ms) pendant laquelle les sujets se sont trouvés dans un intervalle de 40 pixels autour de chacun des deux types d'objets (Objet commun vs Objets privés) en fonction des conditions expérimentales.

Conditions	Différence	Test	Probabilité	Taille d'effet
« Mono »	5504.06	U = 139 ; $Z_{ajusté}$ = 3.072328	p = 0.0021	g = 0.94 (fort)
« Différenciée »	29326.87	U = 35 ; $Z_{ajusté}$ = 5.216772	p < 0.0001	g = 1.98 (fort)
« Parallélisme »	11485.52	U = 72 ; $Z_{ajusté}$ = 4.453845	p < 0.0001	g = 1.50 (fort)

Tableau 14 : Résultats statistiques concernant les durées totales moyennes passées dans un intervalle de 40 pixels autour des deux types d'objets (Objet commun vs Objets privés) pour chaque condition expérimentale.

[47] Notons que nous nous intéressons au nombre de stimulations différentes. En effet, les stimulations sont parfois étendues dans le temps, comme par exemple dans le cas d'un arrêt sur un des objets.

Pour autant, nous constatons des différences pour chaque type d'objet en fonction des conditions expérimentales. En effet, les sujets passent plus de temps autour de l'objet commun dans la condition « Différenciée » que dans la condition « Parallélisme » (d = 12318.02 ms) et dans la condition « Mono » (d = 19069.58 ms) (voir Figure 72). La différence globale est significative (test de Kruskall-Wallis : $H_{(2,72)} = 24.76057$; $p < 0.0001$). De plus, chaque comparaison deux à deux conduit à un résultat significatif (voir Tableau 15).

Comparaison	Différence	Test	Probabilité	Taille d'effet
« Mono » / « Différenciée »	- 19069.58 ms	U = 67 ; $Z_{ajusté}$ = - 4.55719	p < 0.0001	g = 1.74 (fort)
« Mono » / « Parallélisme »	- 6751.56 ms	U = 156 ; $Z_{ajusté}$ = - 2.72194	p = 0.0064	g = 0.76 (modéré)
« Différenciée » / « Parallélisme »	12318.02 ms	U = 141 ; $Z_{ajusté}$ = 3.031253	p = 0.0024	g = 1.01 (fort)

Tableau 15 : Résultats statistiques concernant les comparaisons deux à deux des conditions expérimentales pour le temps total passé par chaque sujet dans un intervalle de 40 pixels autour de l'objet commun.

Parallèlement, les sujets passent plus de temps autour des objets privés en condition « Parallélisme » que dans la condition « Mono » (d = 770.11 ms) et dans la condition « Différenciée » (d = 5523.34 ms) (voir Figure 72). La différence globale est significative (test de Kruskall-Wallis : $H_{(2,72)} = 24.76057$; $p < 0.0001$). Mais lorsque nous comparons les conditions expérimentales deux à deux, nous constatons que la différence réside entre la condition « Différenciée » et les deux autres conditions (voir Tableau 16).

Comparaison	Différence	Test	Probabilité	Taille d'effet
« Mono » / « Différenciée »	4753.23 ms	U = 167 ; $Z_{ajusté}$ = 2.495113	p = 0.0125	g = 0.73 (modéré)
« Mono » / « Parallélisme »	- 770.11 ms	U = 250 ; $Z_{ajusté}$ = - 0.78358	p = 0.4332	x
« Différenciée » / « Parallélisme »	- 5523.34 ms	U = 154 ; $Z_{ajusté}$ = - 2.76318	p = 0.0057	g = 0.91 (fort)

Tableau 16 : Résultats statistiques concernant les comparaisons deux à deux des conditions expérimentales pour le temps total passé par chaque sujet dans un intervalle de 40 pixels autour des objets privés.

En outre, toutes conditions confondues, le nombre de stimulations résultant de la rencontre de l'objet commun est plus élevé que le nombre de stimulations résultant de la rencontre des objets privés (d = 3.64) (voir Figure 73). Cette différence est significative (test

de Mann & Withney : U = 1924 ; $Z_{ajusté}$ = 2.669012 ; p = 0.0076), et l'effet peut être considéré comme modéré (g de Hedge = 0.44).

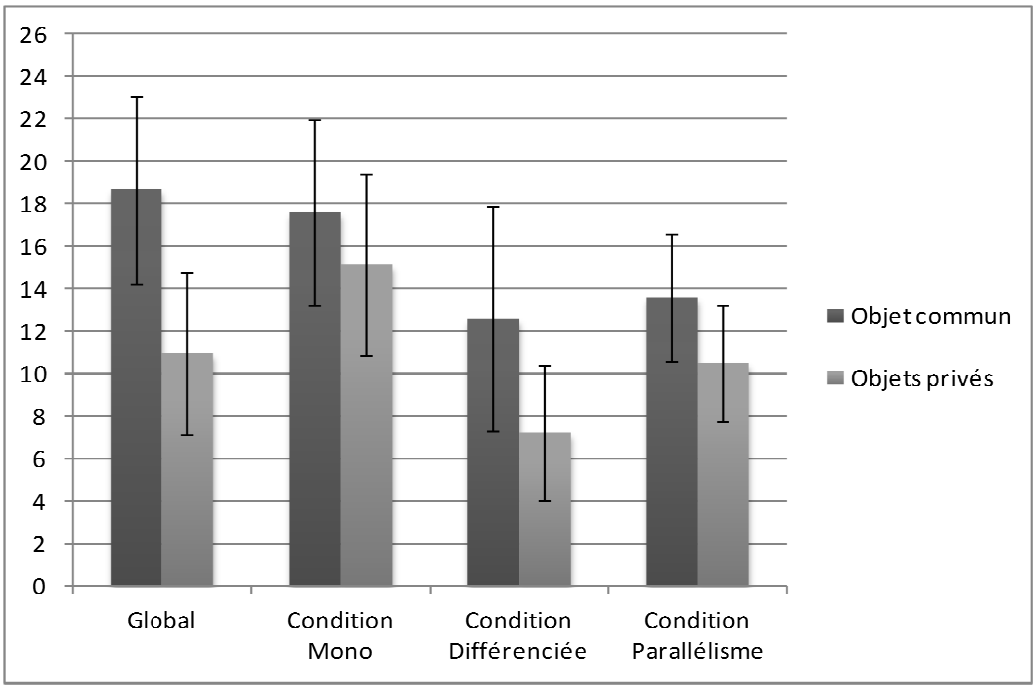


Figure 73 : Nombre moyen de stimulations résultant de la rencontre des deux types d'objets (Objet commun vs Objets privés) en fonction des conditions expérimentales.

En ce qui concerne cette différence en fonction des conditions expérimentales, nous remarquons au préalable que le nombre de stimulations résultant de la rencontre des objets (qu'ils soient privés ou commun) est plus important en condition « Mono » qu'en condition « Parallélisme » (d = 8.69) et « Différenciée » (d = 12.93). La différence globale est significative (test de Kruskall-Wallis : $H_{(2,72)}$ = 10.65433 ; p = 0.0049), tout comme l'ensemble des comparaisons deux à deux des conditions expérimentales (voir Tableau 17).

Comparaison	Différence	Test	Probabilité	Taille d'effet
« Mono » / « Différenciée »	12.93	U = 138.5 ; $Z_{ajusté}$ = 3.083224	p = 0.0020	g = 0.83 (fort)
« Mono » / « Parallélisme »	8.69	U = 190.5 ; $Z_{ajusté}$ = 2.010634	p = 0.0443	g = 0.38 (modéré)
« Différenciée » / « Parallélisme »	- 4.24	U = 189 ; $Z_{ajusté}$ = - 2.04151	p = 0.0412	g = 0.40 (modéré)

Tableau 17 : Résultats statistiques concernant les comparaisons deux à deux des conditions expérimentales pour le nombre moyen de stimulations résultant de la rencontre des objets.

Quoi qu'il en soit, nous observons que dans chacune des trois conditions expérimentales, le nombre moyen de stimulations dues à la rencontre de l'objet commun est plus élevé que celui qui résulte de la rencontre des objets privés (voir Figure 73). Toutefois, nous pouvons noter que la différence n'est pas significative en ce qui concerne la condition « Mono » et la condition « Parallélisme » (voir Tableau 18).

Conditions	Différence	Test	Probabilité	Taille d'effet
« Mono »	2.4635	U = 233.5 ; $Z_{ajusté}$ = 1.124015	p = 0.2610	x
« Différenciée »	5.3489	U = 184 ; $Z_{ajusté}$ = 2.144793	p = 0.0319	g = 0.61 (modéré)
« Parallélisme »	3.1093	U = 204.5 ; $Z_{ajusté}$ = 1.721928	p = 0.0850	x

Tableau 18 : Résultats statistiques concernant le nombre moyen de stimulations résultant de la rencontre des sujets avec les deux types d'objets (Objet commun vs Objets privés) pour chaque condition expérimentale.

Mais là encore, nous constatons des différences pour chaque type d'objet en fonction des conditions expérimentales. En effet, le nombre de stimulations résultant de la rencontre de l'objet commun est plus élevé dans la condition « Mono » que dans la condition « Parallélisme » (d = 4.02) et dans la condition « Différenciée » (d = 5.02) (voir Figure 73). La différence globale est significative (test de Kruskall-Wallis : $H_{(2,72)}$ = 6.621082 ; p = 0.0365). Toutefois, les comparaisons deux à deux révèlent que la différence n'est significative qu'entre la condition « Mono » et la condition « Différenciée » (voir Tableau 19).

Comparaison	Différence	Test	Probabilité	Taille d'effet
« Mono » / « Différenciée »	5.02	U = 169 ; $Z_{ajusté}$ = 2.453938	p = 0.0141	g = 0.51 (modéré)
« Mono » / « Parallélisme »	4.02	U = 222.5 ; $Z_{ajusté}$ = 1.350771	p = 0.1767	x
« Différenciée » / « Parallélisme »	1.00	U = 218.5 ; $Z_{ajusté}$ = - 1.43322	p = 0.1517	x

Tableau 19 : Résultats statistiques concernant les comparaisons deux à deux des conditions expérimentales pour le nombre moyen de stimulations résultant de la rencontre de chaque sujet avec l'objet commun.

Par ailleurs, le nombre de stimulations résultant de la rencontre des objets privés est plus élevé dans la condition « Mono » que dans la condition « Parallélisme » (d = 4.67) et dans la condition « Différenciée » (d = 7.91) (voir Figure 73). La différence globale est significative (test de Kruskall-Wallis : $H_{(2,72)}$ = 14.95307 ; p = 0.0006). Lorsque nous comparons les conditions expérimentales deux à deux, nous constatons que la différence réside entre la condition « Différenciée » et les deux autres conditions (voir Tableau 20).

Comparaison	Différence	Test	Probabilité	Taille d'effet
« Mono » / « Différenciée »	7.91	U = 112.5 ; $Z_{ajusté}$ = 3.619437	p = 0.0002	g = 1.04 (fort)
« Mono » / « Parallélisme »	4.67	U = 200 ; $Z_{ajusté}$ = 1.814529	p = 0.0694	x
« Différenciée » / « Parallélisme »	- 3.24	U = 169 ; $Z_{ajusté}$ = 2.45401	p = 0.0141	g = 0.54 (modéré)

Tableau 20 : Résultats statistiques concernant les comparaisons deux à deux des conditions expérimentales pour le nombre moyen de stimulations résultant de la rencontre de chaque sujet avec les objets privés.

7.6.2. Interaction entre les sujets autour des deux types d'objets

Dans un second temps, nous avons isolé une série d'indicateurs qui concernent l'interaction entre les deux sujets autour des deux types d'objets, et ce, indépendamment de la réponse que les sujets ont fourni à la fin des essais. Nous avons choisi de délimiter un intervalle de 50 pixels autour de chacun des deux types d'objets. Nous nous intéressons dans ces intervalles au nombre de stimulations résultant de la rencontre des deux sujets, ainsi qu'au nombre d'arrêts mutuels[48] et à leur durée totale.

Globalement, nous observons que le nombre moyen de stimulations interpersonnelles autour de l'objet commun est supérieur à celui obtenu autour des objets privés (d = 7,90) (voir Figure 74). Cette différence est significative (test de Mann & Withney : U = 509 ; $Z_{ajusté}$ = 5.130746 ; p < 0.0001), et l'effet peut même être considéré comme fort (g de Hedge = 0.88). Par ailleurs, cet effet se retrouve dans chacune des trois conditions expérimentales (voir Figure 74 et Tableau 21).

[48] Un arrêt mutuel est défini par un arrêt simultané des deux sujets alors qu'ils sont en contact.

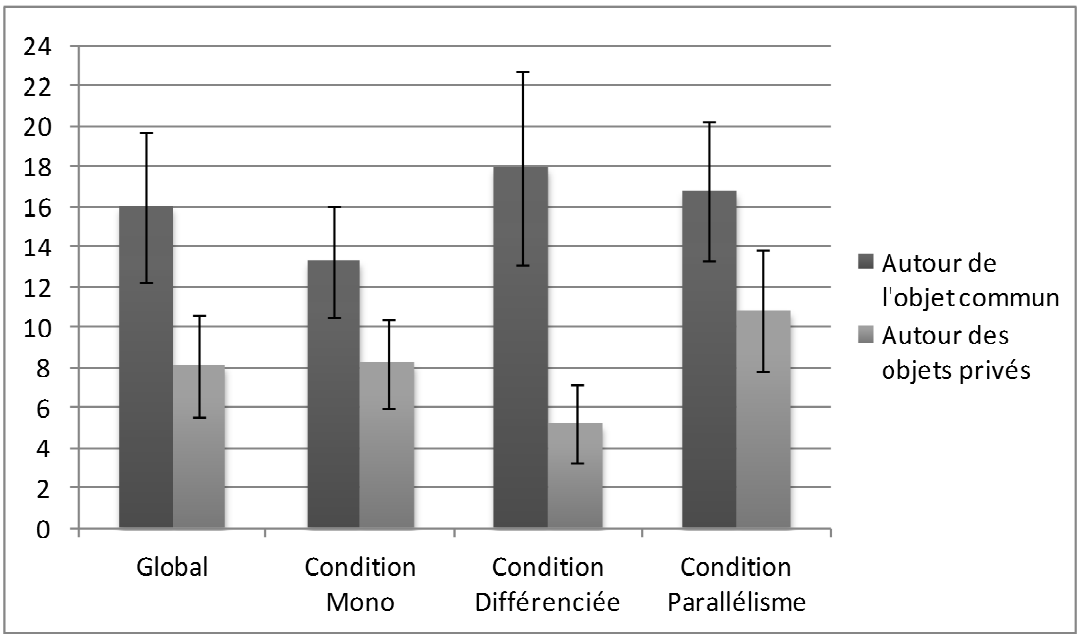


Figure 74 : Nombre moyen de stimulations interpersonnelles autour des deux types d'objet (Autour de l'objet commun vs Autour des objets privés) en fonction des conditions expérimentales.

Conditions	Différence	Test	Probabilité	Taille d'effet
« Mono »	5.0632	U = 68 ; $Z_{ajusté}$ = 2.55105	p = 0.0107	g = 0.69 (modéré)
« Différenciée »	12.663	U = 22 ; $Z_{ajusté}$ = 4.10436	p < 0.0001	g = 1.32 (fort)
« Parallélisme »	5.9769	U = 81 ; $Z_{ajusté}$ = 2.11468	p = 0.0344	g = 0.63 (modéré)

Tableau 21 : Résultats statistiques concernant le nombre moyen de stimulations interpersonnelles survenues dans un intervalle de 50 pixels autour des deux types d'objets (Objet commun vs Objets privés) pour chaque condition expérimentale.

De plus, nous constatons que le nombre de stimulations interpersonnelles obtenues autour de l'objet commun est moins élevé en condition « Mono » qu'en condition « Parallélisme » (d = - 3.51) et en condition « Différenciée » (d = - 4.65). Mais la différence globale n'est pas significative (test de Kruskall-Wallis : $H_{(2,36)}$ = 0.9774775 ; p = 0.6134).

Par ailleurs, le nombre de stimulations interpersonnelles obtenues autour des objets privés est moins élevé en condition « Différenciée » qu'en condition « Mono » (d = - 0.84) et en condition « Parallélisme » (d = - 3.19). Cette fois, la différence globale est significative (test de Kruskall-Wallis : $H_{(2,72)}$ = 9.769922 ; p = 0.0076). Toutefois, de façon spécifique, cette différence n'est significative qu'entre la condition « Différenciée » et la condition « Parallélisme » (voir Tableau 22).

Comparaison	Différence	Test	Probabilité	Taille d'effet
« Mono » / « Différenciée »	0.84	U = 195.5 ; $Z_{ajusté}$ = 1.907318	p = 0.0556	x
« Mono » / « Parallélisme »	- 2.96	U = 242 ; $Z_{ajusté}$ = - 0.948787	p = 0.3427	x
« Différenciée » / « Parallélisme »	- 3.19	U = 134.5 ; $Z_{ajusté}$ = - 3.16873	p = 0.0015	g = 0.73 (modéré)

Tableau 22 : Résultats statistiques concernant les comparaisons deux à deux des conditions expérimentales pour le nombre moyen de stimulations interpersonnelles autour des objets privés.

Globalement, nous observons que le nombre moyen d'arrêts mutuels est plus élevé autour de l'objet commun qu'autour des objets privés (d = 7.71) (voir Figure 75). Cette différence est significative (test de Mann & Withney : U = 392 ; $Z_{ajusté}$ = 5.891533 ; p < 0.0001), et l'effet peut même être considéré comme fort (g de Hedge = 1.34).

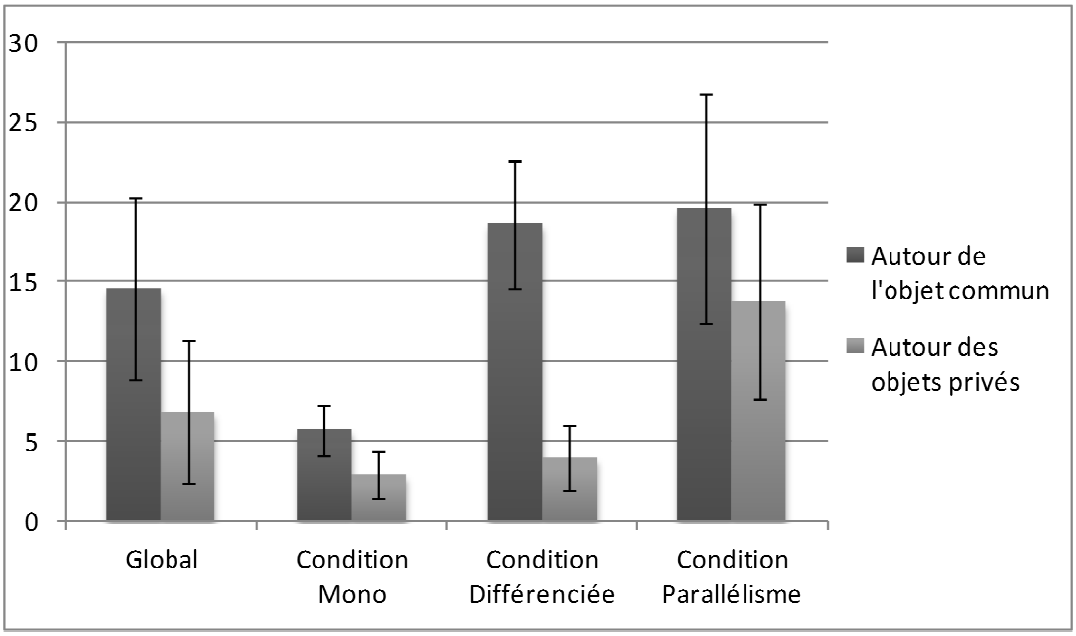


Figure 75 : Nombre moyen d'arrêts mutuels autour des deux types d'objet (Autour de l'objet commun vs Autour des objets privés) en fonction des conditions expérimentales.

En ce qui concerne cette différence en fonction des conditions expérimentales, nous remarquons au préalable que le nombre d'arrêts mutuels autour des objets (qu'ils soient privés ou commun) est plus important en condition « Parallélisme » qu'en condition « Différenciée » (d = 10.42) et « Mono » (d = 24.07). La différence globale est significative (test de Kruskall-Wallis : $H_{(2,36)}$ = 16.57808 ; p = 0.0003). Mais les comparaisons des conditions expérimentales deux à deux révèlent que la différence n'est pas significative entre la condition « Différenciée » et la condition « Parallélisme » (voir Tableau 23).

Comparaison	Différence	Test	Probabilité	Taille d'effet
« Mono » / « Différenciée »	- 13.65	U = 13 ; $Z_{ajusté}$ = - 3.40637	p = 0.0006	g = 1.86 (fort)
« Mono » / « Parallélisme »	- 24.07	U = 12 ; $Z_{ajusté}$ = - 3.46410	p = 0.0005	g = 1.59 (fort)
« Différenciée » / « Parallélisme »	- 10.42	U = 52 ; $Z_{ajusté}$ = - 1.15470	p = 0.2482	x

Tableau 23 : Résultats statistiques concernant les comparaisons deux à deux des conditions expérimentales pour le nombre moyen d'arrêts mutuels survenus autour des deux objets

Quoi qu'il en soit, nous observons que dans chacune de trois conditions expérimentales, le nombre moyen d'arrêts mutuels autour de l'objet commun est plus élevé que le nombre moyen d'arrêts mutuels autour des objets privés (voir Figure 75). Mais la différence n'est pas significative en ce qui concerne la condition « Parallélisme » (voir Figure 75 et Tableau 24).

Conditions	Différence	Test	Probabilité	Taille d'effet
« Mono »	2.7902	U = 55 ; $Z_{ajusté}$ = 2.98664	p = 0.0027	g = 0.69 (modéré)
« Différenciée »	14.547	U = 11.5 ; $Z_{ajusté}$ = 4.45243	p < 0.0001	g = 2.56 (fort)
« Parallélisme »	5.7978	U = 92 ; $Z_{ajusté}$ = 1.745343	p = 0.0809	x

Tableau 24 : Résultats statistiques concernant le nombre moyen des arrêts mutuels survenus dans un intervalle de 50 pixels autour des deux types d'objets (Objet commun vs Objets privés) pour chaque condition expérimentale.

De plus, nous observons que le nombre moyen d'arrêts mutuels autour de l'objet commun est nettement moins élevé en condition « Mono » qu'en condition « Différenciée » (d = - 12.85) et « Parallélisme » (d = - 13.82). La différence globale est significative (test de Kruskall-Wallis : $H_{(2,36)}$ = 15.85388 ; p = 0.0004). Mais plus précisément, la différence n'est significative qu'entre la condition « Mono » et « Différenciée » (test de Mann & Withney : U = 7.5 ; $Z_{ajusté}$ = - 3.72634 ; p < 0.0001), et entre la condition « Mono » et « Parallélisme » (test de Mann & Withney : U = 18.5 ; $Z_{ajusté}$ = - 3.0895 ; p = 0.0020). Dans les deux cas, l'effet peut être considéré comme fort (respectivement g de Hedge = 2.09 et 1.32).

Par ailleurs, le nombre moyen d'arrêts mutuels autour des objets privés est nettement plus élevé en condition « Parallélisme » qu'en condition « Différenciée » (d = - 9.71) et « Mono » (d = - 10.81). Cette différence globale est significative (test de Kruskall-Wallis : $H_{(2,72)}$ = 24.37691 ; p < 0.0001). Plus précisément, la différence n'est significative qu'entre la condition « Parallélisme » et « Différenciée » (test de Mann & Withney : U = 73.5 ; $Z_{ajusté}$ = 4.42688 ; p < 0.0001), et entre la condition « Parallélisme » et « Mono » (test de Mann &

Withney : U = 97.5 ; $Z_{ajusté}$ = 3.93199 ; p < 0.0001). Dans les deux cas, l'effet peut même être considéré comme fort (respectivement g de Hedge = 1.22 et 1.07).

Enfin, la durée totale moyenne de ces arrêts mutuels est plus élevée autour de l'objet commun qu'autour des objets privés (d = 10424.55 ms) (voir Figure 76). Cette différence est significative (test de Mann & Withney : U = 417 ; $Z_{ajusté}$ = 5.457165 ; p < 0.0001), et l'effet peut même être considéré comme fort (g de Hedge = 1.21).

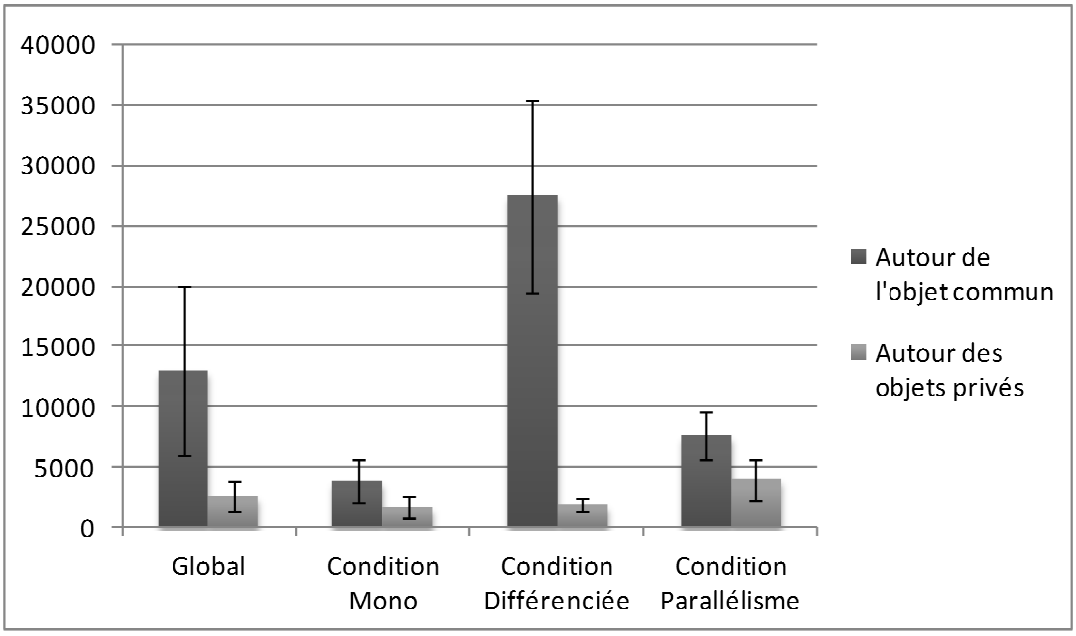


Figure 76 : Durée totale moyenne des arrêts mutuels autour des deux types d'objet (Autour de l'objet commun vs Autour des objets privés) en fonction des conditions expérimentales.

Concernant la différence entre les conditions expérimentales, nous remarquons tout d'abord que la durée totale des arrêts mutuels autour des objets (qu'ils soient privés ou commun) est plus important en condition « Différenciée » qu'en condition « Parallélisme » (d = 17647.26 ms) et « Mono » (d = 23536.11 ms). La différence globale est significative (test de Kruskall-Wallis : $H_{(2,36)}$ = 20.40691 ; p < 0.0001). En outre, cette différence est significative pour chacune des comparaisons deux à deux (voir Tableau 25).

Comparaison	Différence	Test	Probabilité	Taille d'effet
« Mono » / « Différenciée »	- 23536.11 ms	U = 5 ; $Z_{ajusté}$ = - 3.86825	p = 0.0001	g = 2.06 (fort)
« Mono » / « Parallélisme »	- 5888.85 ms	U = 25 ; $Z_{ajusté}$ = - 2.71355	p = 0.0066	g = 1.20 (fort)
« Différenciée » / « Parallélisme »	17647.26 ms	U = 20 ; $Z_{ajusté}$ = 3.002221	p = 0.0026	g = 1.49 (fort)

Tableau 25 : Résultats statistiques concernant les comparaisons deux à deux des conditions expérimentales pour la durée totale moyenne des arrêts mutuels autour des objets

Malgré tout, la durée totale moyenne des arrêts mutuels autour de l'objet commun est significativement plus élevée que la durée totale moyenne des arrêts mutuels autour des objets privés pour chacune des trois conditions expérimentales (voir Figure 76 et Tableau 26).

Conditions	Différence	Test	Probabilité	Taille d'effet
« Mono »	2.7902	U = 67 ; $Z_{ajusté}$ = 2.20802	p = 0.0272	g = 0.83 (fort)
« Différenciée »	14.547	U = 0 ; $Z_{ajusté}$ = 4.75695	p < 0.0001	g = 2.73 (fort)
« Parallélisme »	5.7978	U = 69 ; $Z_{ajusté}$ = 2.516835	p = 0.0118	g = 1.01 (fort)

Tableau 26 : Résultats statistiques concernant la durée totale moyenne des arrêts mutuels survenus dans un intervalle de 50 pixels autour des deux types d'objets (Objet commun vs Objets privés) pour chaque condition expérimentale.

En outre, nous observons que la durée totale des arrêts mutuels autour de l'objet commun est nettement plus élevée en condition « Différenciée » qu'en condition « Parallélisme » (d = 19758.61 ms) et « Mono » (d = 23546.40 ms). La différence globale est significative (test de Kruskall-Wallis : $H_{(2,36)}$ = 21.2177 ; p < 0.0001). De façon plus spécifique, si la différence est bien entendue significative entre la condition « Différenciée » et « Parallélisme » (test de Mann & Withney : U = 13 ; $Z_{ajusté}$ = 3.406367 ; p = 0.0006), et entre la condition « Différenciée » et « Mono » (test de Mann & Withney : U = 4 ; $Z_{ajusté}$ = 3.92598 ; p < 0.0001), nous la retrouvons également entre la condition « Parallélisme » et « Mono » (test de Mann & Withney : U = 31 ; $Z_{ajusté}$ = 2.36714 ; p = 0.0179). Dans tous les cas, l'effet peut même être considéré comme fort (respectivement g de Hedge = 1.70 ; 2.04 et 1.01).

Par ailleurs, la durée totale des arrêts mutuels autour des objets privés est plus élevée en condition « Parallélisme » qu'en condition « Différenciée » (d = 2072.38 ms) et « Mono » (d = 2249.04 ms). La différence globale est également significative (test de Kruskall-Wallis : $H_{(2,67)}$ = 9,47412 ; p = 0.0088). Mais plus précisément, la différence n'est significative

qu'entre la condition « Parallélisme » et « Différenciée » (test de Mann & Withney : U = 165 ; $Z_{ajusté}$ = 2.17701 ; p = 0.0294), et entre la condition « Parallélisme » et « Mono » (test de Mann & Withney : U = 132 ; $Z_{ajusté}$ = 2.73027 ; p = 0.0063). Dans les deux cas, l'effet peut être considéré comme modéré (respectivement g de Hedge = 0.78 et 0.76).

7.6.3. Interaction entre les sujets en fonction de la qualité des clics

Finalement, nous nous intéressons à la qualité de l'interaction entre les sujets en fonction de la qualité de leurs clics individuels (Clics sur l'objet commun vs Mauvais clics), puis en fonction de leur modalité de réponse collective (Clics mutuels vs Clics solitaires vs Autres cas).

Pour cela, nous avons calculé la proportion d'accroche mutuelle. Nous avons alors isolé les séquences d'interaction pendant lesquelles les écarts entre les stimulations sont proches. Nous calculons la proportion d'accroche mutuelle en additionnant les durées totales pendant lesquelles les stimulations résultant des rencontres sont espacées de moins de deux secondes et la durée des stimulations correspondantes, et en divisant cette somme par la durée totale d'un essai. Notons que nous ne prenons pas en compte la durée des arrêts mutuels lorsque ceux-ci sont supérieurs à deux secondes.

Nous observons que globalement, la proportion d'accroche mutuelle est plus élevée lors des essais qui ont conduit à un clic sur l'objet commun par rapport aux essais ayant conduit à un mauvais clic (d = 5.12%) (voir Figure 77). Cette différence est significative (test de Mann & Withney : U = 1348 ; $Z_{ajusté}$ = 2.549637 ; p = 0.0107), et l'effet peut être considéré comme modéré (g de Hedge = 0,43).Toutefois, nous observons des différences en fonction des conditions expérimentales. Dans les conditions « Mono » et « Parallélisme », la proportion d'accroche mutuelle est supérieure quand le sujet a cliqué sur l'objet commun que lorsqu'il a cliqué ailleurs. Ces différences sont significatives, mais nous notons que l'effet est notable pour la condition « Parallélisme » contrairement à la condition « Mono ». Dans la condition « Différenciée », la proportion d'accroche mutuelle est moins élevée lorsque le sujet a cliqué sur l'objet commun que lorsqu'il a cliqué ailleurs. Toutefois, cette différence n'est pas significative (voir Figure 77 et Tableau 27).

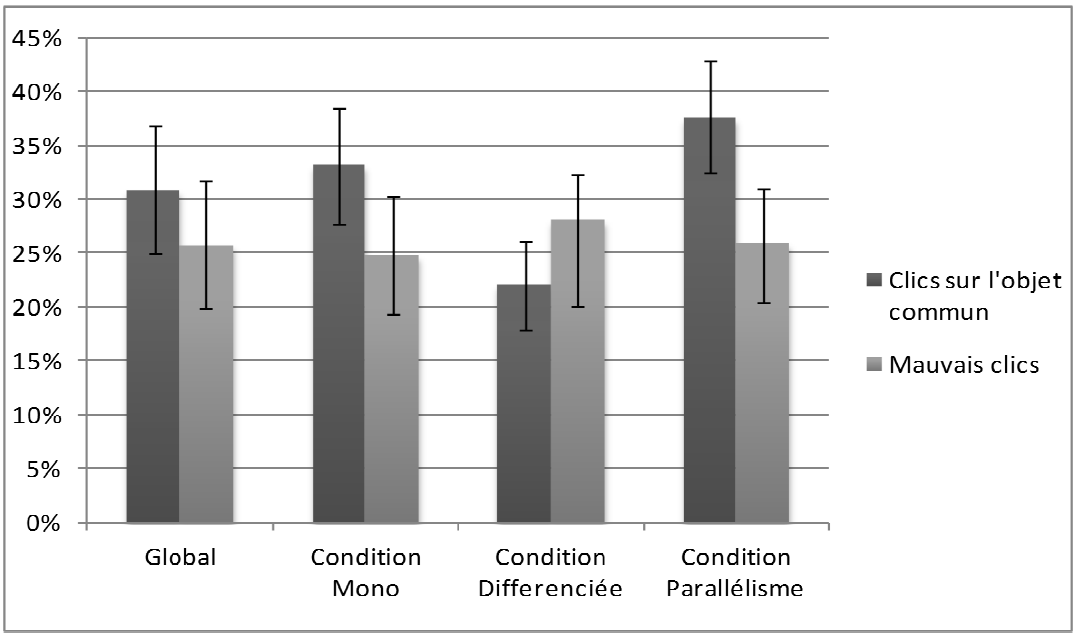


Figure 77 : Proportion d'accroche mutuelle entre les deux sujets en fonction de la qualité de leurs clics (clics sur l'objet commun vs mauvais clics) en fonction des conditions expérimentales.

Conditions	Différence	Test	Probabilité	Taille d'effet
« Mono »	8.25%	$t_{(43)} = 2.5423$	p = 0.0147	B = 0.61 (non notable)
« Différenciée »	- 6.13%	U = 80 ; $Z_{ajusté} = - 0.6966$	p = 0.4860	x
« Parallélisme »	11.82%	$t_{(44)} = 3.7988$	p = 0.0004	B = 0.92 (notable)

Tableau 27 : Résultats statistiques concernant la proportion d'accroche mutuelle entre les deux sujets en fonction de la qualité de leurs clics (clics sur l'objet commun vs mauvais clics) pour chaque condition expérimentale.

Par ailleurs, si nous ne considérons que les clics sur l'objet commun, nous observons que la proportion d'accroche mutuelle est plus élevée dans la condition « Parallélisme » que dans la condition « Mono » (d = 4.51%) et « Différenciée » (d = 15.61%) (voir Figure 77). La différence globale est significative (test de Kruskall-Wallis : $H_{(2,71)} = 23.28953$; $p < 0.0001$). D'une façon plus spécifique, nous pouvons conclure à une différence significative entre la condition « Mono » et la condition « Différenciée » (test de Mann & Withney : U = 116 ; $Z_{ajusté} = 3.405518$; p = 0.0006), et entre la condition « Différenciée » et la condition « Parallélisme » (test de Mann & Withney : U = 66 ; $Z_{ajusté} = - 4.57756$; $p < 0.0001$). Notons que ces deux effets peuvent être considérés comme forts (respectivement g = 1.16 et g = 1.68). En revanche, la différence entre les conditions « Mono » et « Parallélisme » n'est pas significative (test de Mann & Withney : U = 210 ; $Z_{ajusté} = - 1.40461$; p = 0.1601).

Enfin, de façon globale, nous observons que la proportion d'accroche mutuelle est plus élevée dans le cas des clics mutuels que dans celui des clics solitaires (d = 6.11%) et des

autres cas (d = 10.94%) (voir Figure 78). La différence globale est significative (test de Kruskall-Wallis : $H_{(2,95)}$ = 13.23588 ; p = 0.0013). De façon plus spécifique, la proportion d'accroche est significativement plus élevée dans le cas des clics mutuels que dans le cas des clics unilatéraux (test de Mann & Whitney : U = 355 ; $Z_{ajusté}$ = 2.085286 ; p = 0.0370) et dans les autres cas (test de Mann & Whitney : U = 254 ; $Z_{ajusté}$ = 3.585593 ; p = 0.0003). Si dans le premier cas, l'effet est modéré (g de Hedge = 0.47), il peut être considéré comme fort dans le second cas (g de Hedge = 0.97). En revanche, la différence n'est pas significative entre le cas des clics solitaires et celui des autres cas (test de Mann & Whitney : U = 361 ; $Z_{ajusté}$ = 1.500329 ; p = 0.1335).

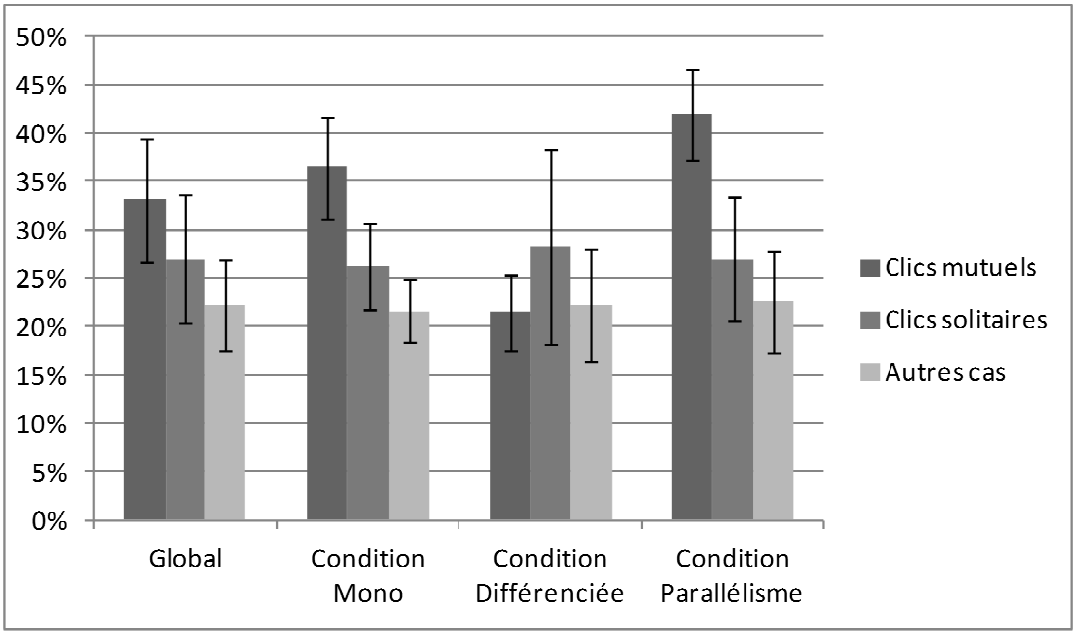


Figure 78 : Proportion d'accroche mutuelle entre les deux sujets en fonction de la modalité de réponse collective (Clics Mutuels, Clics Solitaires et Autres cas) en fonction des conditions expérimentales.

De façon plus spécifique, la différence globale entre les modalités de réponse collective n'est pas significative pour la condition « Différenciée » (test de Kruskall-Wallis : $H_{(2,27)}$ = 0.2544407 ; p = 0.8805). Mais en ce qui concerne les conditions « Mono » et Parallélisme », nous retrouvons les mêmes résultats que ceux que nous avons obtenus toutes conditions confondues (voir Figure 78, Tableau 28 et Tableau 29).

Comparaison	Différence	Test	Probabilité	Taille d'effet
Globale	x	$H_{(2,33)} = 10.39331$	p = 0.0055	x
Clics Mutuels / Clics Solitaires	10.24%	U = 29 ; $Z_{ajusté} = 2.044079$	p = 0.0147	g = 1.05 (fort)
Clics Mutuels / Autres cas	14.89%	U = 10 ; $Z_{ajusté} = 3.168805$	p = 0.0015	g = 1.71 (fort)
Clics Solitaires / Autres cas	4.65%	U = 46 ; $Z_{ajusté} = 1.230915$	p = 0.2183	x

Tableau 28 : Récapitulatif des résultats statistiques concernant la proportion d'accroche mutuelle en fonction des modalités de réponse collective (Clics Mutuels, Clics Solitaires et Autres cas) dans la condition « Mono ».

Comparaison	Différence	Test	Probabilité	Taille d'effet
Globale	x	$H_{(2,35)} = 12.54711$	p = 0.0019	x
Clics Mutuels / Clics Solitaires	14.92%	U = 25 ; $Z_{ajusté} = 2.523376$	p = 0.0116	g = 1.32 (fort)
Clics Mutuels / Autres cas	19.37%	U = 10 ; $Z_{ajusté} = 3.168805$	p = 0.0015	g = 1.91 (fort)
Clics Solitaires / Autres cas	4.45%	U = 56 ; $Z_{ajusté} = 0.615457$	p = 0.5382	x

Tableau 29 : Récapitulatif des résultats statistiques concernant la proportion d'accroche mutuelle en fonction des modalités de réponse collective (Clics Mutuels, Clics Solitaires et Autres cas) dans la condition « Parallélisme ».

Pour finir, nous comparons la proportion d'accroche mutuelle obtenue dans chaque modalité de réponse collective en fonction des conditions expérimentales. Concernant les cas de clics mutuels, la proportion de l'accroche est moins élevée en condition « Différenciée » qu'en condition « Mono » (d = - 15.08%), et « Parallélisme » (d = - 20.48%) (voir Figure 78). La différence globale est significative (test de Kruskall-Wallis : $H_{(2,34)} = 17.79261$; p = 0.0001). Plus précisément, nous constatons que la différence n'est significative qu'entre la condition « Différenciée » et « Mono » (test de Mann & Whitney : U = 11 ; $Z_{ajusté} = -3.230964$; p = 0.0012), et entre la condition « Différenciée » et « Parallélisme « (test de Mann & Whitney : U = 6 ; $Z_{ajusté} = - 3.81051$; p = 0.0001). Dans les deux cas, l'effet peut être considéré comme fort (respectivement g de Hedge = 1.66 et 2.35). En revanche, la différence n'est pas significative entre la condition « Mono » et la condition « Parallélisme » (test de Mann & Whitney : U = 45 ; $Z_{ajusté} = - 0.989071$; p = 0.3226).

En ce qui concerne les cas de clics unilatéraux et ceux des autres cas, nous ne trouvons pas de différence significative en fonction des conditions expérimentales (respectivement $H_{(2,30)} = 0.1205139$; p = 0.9415 et $H_{(2,31)} = 0.0066441$; p = 0.9967).

7.7. Discussion

Notre expérience avait pour vocation de tester la possibilité pour deux sujets engagés dans un environnement minimaliste de percevoir qu'un des objets de cet environnement est également perçu par l'autre. Pour cela, dans les conditions de notre situation, mais aussi conformément à notre position théorique concernant l'émergence d'une signification partagée au cœur d'une coordination des activités perceptives, nous postulions que l'organisation du processus d'interaction entre les sujets et les objets d'une part, et entre les sujets d'autre part, pouvait conduire à une fréquence de clics sur l'objet commun plus élevée que la fréquence de mauvais clics, qui témoignerait de capacités d'attention conjointe. Par ailleurs, la coordination des focus attentionnels se réalisant via un engagement mutuel des participants, la fréquence de clics mutuels devait être plus importante que la fréquence des clics solitaires. Pour autant, les conditions expérimentales devaient avoir un effet sur la réussite de la tâche, puisque la nature du dispositif de couplage entre les actions et les sensations individuelles devaient conduire à des modes d'organisation de l'interaction différents. Dans cette optique, si la condition « Mono » devait conduire à une fréquence de clics sur l'objet commun plus faible que dans les autres conditions, cette configuration devait en même temps être particulièrement révélatrice d'une dynamique interpersonnelle, le dispositif de couplage exigeant en soi un déploiement plus important des activités perceptives et donc un niveau plus élevé de coordination de ces activités. En outre, tout comme lors de notre première étude expérimentale, nous nous attendions à ce que la réussite soit meilleure dans le deuxième bloc d'essais de chaque condition expérimentale, témoignant d'une stabilisation de la dynamique collective et d'une appropriation individuelle du dispositif de couplage.

D'une façon, globale, nous pouvons conclure à un succès de la tâche, que ce soit à un niveau individuel (clics sur l'objet commun), ou à un niveau collectif (clics mutuels). En effet, toutes conditions confondues, le nombre moyen de clics sur l'objet commun est supérieur au nombre moyen de mauvais clics. De plus, nous pouvons conclure à un effet d'apprentissage individuel global, qui se traduit par une augmentation de clics sur l'objet commun et à une diminution des mauvais clics entre les deux blocs d'essais. Dans les conditions de l'expérience, les sujets ont donc pu percevoir qu'un des deux objets est également perçu par l'autre : ils ont pu coordonner leurs activités perceptives de façon à établir un cadre de référence partagé, comme dans le cas du suivi de regard (Butterworth & Jarrett, 1991 ; Fiebich & Gallagher, 2012).

D’un point de vue collectif, la fréquence des clics mutuels est supérieure à la fréquence des clics solitaires et des autres cas. Ce résultat semble montrer que la réussite de la tâche est attribuable à un engagement interpersonnel (Hobson, 2002 ; Reddy, 2011), plus qu’à une simple « lecture » unilatérale de la dynamique du mouvement de l’autre (Lempers, 1979 ; Moore et al., 1997), ou encore à une attribution individuelle d’une intentionnalité chez autrui (Tomasello, 1995 ; 1999). Ainsi, nous pouvons poser l’hypothèse que la connaissance du fait qu’un objet de l’environnement est partagé émerge de la dynamique relationnelle, en tant que processus autonome qui influence les contributions individuelles et qui permet aux interactants de leur attribuer du sens (De Jaegher & Di Paolo, 2007 ; De Jaegher et al., 2010). En revanche, dans la mesure où nous ne pouvons pas conclure à un effet d’apprentissage collectif global, qui se traduirait minimalement par une augmentation des clics mutuels et une diminution des clics solitaires entre les deux blocs d’essais, nous ne pouvons pas conclure à une stabilisation de cette dynamique qui aurait été profitable aux deux sujets. Au contraire, l’effet d’apprentissage collectif ne se traduit que par une diminution des autres cas. De cette façon, l’évolution de la performance ne se vérifie qu’à un niveau individuel, ce qui ne veut pas dire qu’elle ne peut pas être basée sur la dynamique collective. En effet, nous pouvons rappeler ici que le fait de mettre l’accent sur le couplage interpersonnel ne revient pas à dissoudre le rôle de l’individu : celui-ci doit pouvoir re-saisir cette dynamique individuellement, de façon à construire du sens à partir d’elle (Lenay & Stewart, 2012).

En outre, si nous pouvons attester d’une réussite globale, nous pouvons conclure que celle-ci est dépendante des conditions expérimentales. A un niveau individuel, si les trois conditions amènent plus de clics sur l’objet commun que de mauvais clics, nous ne pouvons pas conclure à la réussite de la tâche dans la condition « Mono ». Par ailleurs, nous pouvons conclure que les sujets sont plus efficaces dans la condition « Différenciée » que dans les conditions « Mono » et « Parallélisme ». De cette façon, la condition « Différenciée » semble favoriser la perception mutuelle d’un objet commun, et surtout la connaissance du fait que cet objet est partagé, ce qui a amené les sujets à cliquer davantage sur l’objet commun. A un niveau collectif, en revanche, nous ne pouvons pas conclure à une fréquence de clics mutuels plus élevée que les autres modalités de réponse collective dans la condition « Mono » et dans la condition « Parallélisme ». Seule la condition « Différenciée » amène significativement plus de clics mutuels que de clics solitaires et des autres cas. Mais malgré cet apparent échec collectif, nous avons souligné dans la majorité des cas de clics solitaires, lorsqu’un des deux sujets clique sur l’objet commun, l’autre ne clique pas. De la même façon, dans les autres cas, lorsqu’un des deux sujets clique ailleurs que sur l’objet commun, l’autre ne clique pas. Or,

comme nous l'avons vu, il est difficile d'attribuer à une absence de réponse le même statut qu'une mauvaise réponse. Si nous ne pouvons pas attester d'une réussite collective, nous ne pouvons pas pour autant affirmer que celle-ci serait purement individuelle. En effet, nous avons montré que la réussite de l'un des deux sujets n'est pas liée à l'échec de l'autre, mais seulement à une absence de réponse de celui-ci.

L'analyse des trajectoires perceptives nous apporte des éléments cruciaux pour tenter de comprendre ce qui amène les sujets à cliquer sur l'objet commun. Nous nous intéressons ici précisément au déploiement des activités perceptives et à la coordination interpersonnelle, telles qu'elles sont rendues accessibles à l'analyse via la méthode minimaliste (Lenay, 2006c).

Tout d'abord, nous avons constaté que les sujets passent significativement plus de temps autour de l'objet commun qu'autour des objets privés, et ce que ce soit à un niveau global (toutes conditions confondues), qu'à un niveau plus spécifique (dans chacune des trois conditions expérimentales). Dans le cadre de notre dispositif, nous avons vu que l'objet attire l'activité perceptive de l'utilisateur, car il constitue une source de stimulations qui permet d'organiser le geste d'exploration (Lenay et al., 2000). Ici, bien que les deux types d'objet possèdent exactement les mêmes propriétés objectives (même taille, même forme), les activités perceptives déployées par les sujets sont plus attirées par l'objet commun que par l'objet privé. Toutefois, nous pouvons conclure que l'objet commun attire de façon plus importante l'activité dans la condition « Différenciée » que dans les autres conditions. En outre, les sujets sont plus fortement attirés par l'objet commun en condition « Parallélisme » qu'en condition « Mono ». Parallèlement, les objets privés attirent moins l'activité des sujets dans la condition « Différenciée » que dans les autres conditions expérimentales, alors que nous ne pouvons pas conclure à une différence entre les conditions « Parallélisme » et « Mono ».

Mais si cette attirance se traduit globalement par un nombre plus élevé de stimulations résultant de la rencontre de l'objet commun par rapport au nombre de stimulations résultant de la rencontre de l'objet privé, cette différence n'est vérifiée qu'en condition « Différenciée ». Par conséquent, nous pouvons conclure que dans les conditions « Mono » et « Parallélisme », les activités perceptives déployées par les sujets sont attirées par une zone de l'espace située autour de l'objet plutôt que par l'objet lui-même. Par ailleurs, étant entendu que nous avons constaté un nombre de stimulations résultant de la rencontre des objets plus élevé dans la condition « Mono » que dans les autres conditions expérimentales, nous pouvons tout de même conclure que la condition « Mono » amène plus de stimulations résultant de la rencontre de l'objet commun que la condition « Différenciée ».

L'analyse des indicateurs concernant l'interaction entre les sujets autour des deux types d'objets permet d'expliquer l'attirance plus forte de l'objet commun. En effet, nous pouvons conclure à un nombre plus élevé de stimulations résultant du croisement perceptif autour de l'objet commun qu'autour des objets privés, et ce quelle que soit la condition expérimentale. En d'autres termes, si les deux types d'objet attirent les activités perceptives individuelles, l'attirance pour l'objet commun est plus forte parce que celui-ci attire réciproquement l'activité du partenaire. En effet, l'objet commun présente une affordance simultanée (Marsh et al., 2006), qui fournit une opportunité d'action identique pour les deux participants. Dans cette optique, il est probable que la rencontre entre les deux activités perceptives soit d'abord fortuite : l'objet commun conditionne un nombre élevé de rencontres autour de lui par son caractère partagé et sa disponibilité pour les deux interactants. Par conséquent, dans notre situation, la connaissance du fait que cet objet est partagé ne provient pas nécessairement d'une conduite intentionnelle de la part d'un des deux sujets pour diriger activement l'attention de l'autre vers les objets présents (Adamson & Bakeman, 1985 ; Mundy, 1995 ; Mundy & Acra, 2006). Si effectivement, l'analyse qualitative des trajectoires révèle que la plupart des sujets déploient une stratégie orientée vers les objets (ils recherchent activement les objets et cherchent à détecter la présence de l'autre autour d'eux), il est à noter que dans certains cas, les sujets cherchent d'abord à se rencontrer pour explorer l'environnement de façon conjointe (stratégie orientée sujet ; voir exemples en Annexe D, p. 294). Mais dans tous les cas, nous ne pouvons pas réduire l'émergence de la connaissance du caractère commun de l'objet à une simple rencontre interpersonnelle inopinée autour de cet objet. Puisque les sujets n'accèdent pas aux stimulations reçues par leur partenaire, il faut encore qu'ils puissent se coordonner efficacement autour d'un objet pour que les actions de l'un puissent avoir du sens pour l'autre et réciproquement. De cette façon, l'affordance simultanée de l'objet commun doit pouvoir se combiner avec une affordance sociale (Krueger, 2011), de façon à constituer une dynamique relationnelle autonome organisée autour de l'objet commun (De Jaegher & Di Paolo, 2007 ; De Jaegher et al., 2010). Rappelons que c'est dans cette dynamique que peut avoir lieu un *participatory sense-making* pouvant expliquer la co-construction d'une connaissance partagée (ibid.).

Dans ce sens, l'analyse de l'accroche perceptive mutuelle entre les deux sujets révèle que toutes conditions confondues, plus la coordination entre les deux sujets est grande, plus les sujets cliquent sur l'objet commun. En outre, cette accroche ne semble pas déterminer uniquement les réponses individuelles, puisque crucialement, cette coordination amène les sujets à cliquer sur l'objet commun de façon mutuelle. En effet, l'accroche mutuelle est

significativement plus élevée dans les cas de clics mutuels que dans celui des clics solitaires et des autres cas. En revanche, de façon plus spécifique, si ces résultats sont retrouvés dans les conditions « Mono » et « Parallélisme », nous ne les retrouvons pas dans la condition « Différenciée ». Par ailleurs, si les cas des clics sur l'objet commun et des clics mutuels de la condition « Différenciée » présentent significativement moins d'accroche mutuelle que dans la condition « Mono » et dans la condition « Parallélisme », ces deux dernières ne présentent pas de différence significative.

Il semble donc que si la condition « Différenciée » amène les meilleurs résultats, ceux-ci ne peuvent pas être expliqués par un haut degré d'accroche perceptive mutuelle, du moins pas si nous la définissons comme une coordination des activités perceptives déployées dans l'espace et dans le temps[49]. Ainsi, les résultats obtenus dans cette condition semblent à première vue remettre en cause l'idée selon laquelle l'attention conjointe est une action conjointe basique (Fiebich & Gallagher, 2012), se définissant par une coordination interpersonnelle sensori-motrice (Gallagher, 2011). Mais le cas des arrêts mutuels nous apportent des éléments de compréhension supplémentaires.

En effet, nous avons vu que d'une façon globale, les sujets s'arrêtent mutuellement plus souvent autour de l'objet commun qu'autour des objets privés. Or cette différence est très nette en condition « Différenciée ». De plus, la durée de ces arrêts mutuels est plus importante autour de l'objet commun qu'autour de l'objet privé. Là encore, nous avons observé une très nette différence dans la condition « Différenciée ». Nous pouvons expliquer ce résultat par le fait que dans cette condition, la différence entre le partenaire et les objets est déjà donnée au sujet dès la stimulation sensorielle. En d'autres termes, contrairement aux conditions « Mono » et « Parallélisme », il n'est pas nécessaire de les discriminer à partir de l'activité perceptive. De plus, étant donné que le sujet peut percevoir en même temps l'objet et le partenaire, l'arrêt du sujet sur l'objet a plus de sens que dans les autres conditions. Nous pourrions alors résumer la stratégie adoptée majoritairement en condition « Différenciée » de cette manière :

- Je cherche une source de stimulation résultant de la rencontre d'un objet. Quand j'en ai trouvé une, je m'arrête et j'attends.

[49] Rappelons ici que nous avons défini l'accroche perceptive mutuelle comme une séquence d'interaction pendant laquelle les écarts entre les stimulations sont inférieurs à deux secondes.

- Si mon partenaire vient à ma rencontre, je reçois alors deux stimulations différenciées de façon simultanée. Mais le partenaire est peut-être alors simplement en train de me « regarder ».
- Si mon partenaire repart, c'est que je suis en train de regarder l'objet privé. En revanche, s'il maintient sa position, c'est que lui aussi perçoit l'objet que je suis en train de percevoir. En d'autres termes, je comprends son arrêt comme je comprends le mien : nous sommes tous les deux stimulés par un objet et le corps-image de l'autre.

De cette façon, le maintien de l'arrêt du partenaire permet au sujet de savoir que l'objet de son attention est partagé : il fait sens du maintien de l'arrêt de l'autre en fonction d'un objet que lui-même perçoit. Ainsi, si la connaissance du fait que l'objet est partagé ne résulte pas en soi d'un déploiement d'une activité perceptive conjointe autour de cet objet (telle qu'elle est entendue par notre définition d'une accroche perceptive mutuelle), elle résulte tout de même d'un engagement actif de la part du sujet (la recherche d'un objet). Par ailleurs, nous pouvons également considérer que le fait de maintenir un arrêt est en soi la manifestation d'un engagement actif de la part du sujet[50], qui signe une accroche perceptive mutuelle autour d'un objet.

Nous pourrions considérer que l'arrêt, et le fait de rester arrêté est une situation privilégiée et favorisée par le dispositif de couplage lui-même. Dans ce cadre, les différences observées en fonction des conditions expérimentales appuient un fait essentiel : la perception n'est pas une réception passive d'informations sensorielles mais est le résultat de l'appropriation d'un dispositif de couplage sensori-moteur. En ce sens, la médiation technique offerte par les dispositifs de suppléance perceptive, comme les dispositifs de couplage naturels, rend possible la perception, en mettant à disposition un répertoire de retours sensoriels possibles qui contraint en retour les actions utiles (Lenay et al., 2000). Dans la condition « Différenciée », par exemple, le système de couplage induit les arrêts mutuels, parce que les types de stimulations reçues suffisent pour comprendre l'engagement du partenaire en fonction des objets présents. Ainsi, le sujet peut extraire de la situation du maintien de l'arrêt mutuel l'information selon laquelle l'autre perçoit l'objet qu'il perçoit lui-même, parce que l'arrêt de l'autre n'a de sens qu'en tant qu'activité perceptive dirigée simultanément vers soi et vers l'objet.

[50] Pour illustrer cette idée, nous pouvons mentionner par exemple l' « inattention civile », qui illustre le fait qu'il est impossible de ne pas interagir en situation de co-présence (Goffman, 1959), ou encore certaines théories de la communication qui décrivent l'absence d'émissions verbales de la part d'un interlocuteur comme un acte de communication en soi (Searle, 1972).

En condition « Mono », au contraire, la différenciation autrui/objet n'est pas pré-donnée aux sujets, mais elle doit se constituer au travers de l'interaction. De plus, il est impossible de percevoir l'autre et l'objet en même temps. La perception mutuelle nécessite alors une alternance dynamique des « regards » et un fort niveau de coordination interpersonnelle, que ce soit à un niveau microscopique (accroche mutuelle et dynamique de l'activité perceptive autour de l'objet) qu'à un niveau macroscopique (alternance structurée des regards vers l'autre et vers l'objet, comme produisant un processus continu d'engagement et de désengagement interpersonnel). Par ailleurs, nous avons remarqué que le nombre de stimulations résultant de la rencontre des objets est plus élevé en condition « Mono » que dans les autres conditions. Ce résultat témoigne d'un déploiement de l'activité perceptive plus élevé que dans les autres conditions.

La condition « Parallélisme » peut être considérée comme une condition « Mono » augmentée. En effet, la différenciation autrui/objet n'est toujours pas pré-donnée aux sujets, mais les stimulations sensorielles présentent une spatialité intrinsèque qui produit une économie de mémoire et de mouvement de la part des individus. En outre, il est possible de percevoir simultanément l'autre et l'objet si les activités perceptives s'organisent autour de l'objet. La condition « Parallélisme » implique alors une coordination plus fine des regards, qui est rendue possible via un déploiement de l'activité plus restreint qu'en condition « Mono ». Mais de façon intéressante, nous avons également noté que cette coordination seule ne suffit pas. En effet, la condition « Parallélisme » est la condition qui présente le plus d'arrêts mutuels, même si ceux-ci sont nettement moins longs que dans la condition « Différenciée » et qu'ils ont lieu autour des deux types d'objet. Il semble donc que les arrêts mutuels soient également une action utile dans la condition « Parallélisme » pour faire sens de l'activité de son partenaire. Ainsi, si les meilleures performances sont obtenues avec la condition « Différenciée », nous pensons que la condition « Parallélisme » est la plus encourageante, car elle offre un répertoire d'actions plus complet, notamment en rendant possible le déploiement des activités perceptives nécessaire à une coordination interpersonnelle spatio-temporelle. Rappelons qu'il s'agit également de la condition dans laquelle les résultats ont le plus progressé entre les deux blocs d'essais.

Malgré l'apparente stabilité des stratégies observées en fonction des conditions expérimentales, une importante limite peut être soulignée. En effet, dans cette expérience, les sujets ont effectué les trois conditions expérimentales. Si nous avons contrebalancé l'ordre de passation de ces conditions, de façon à neutraliser un éventuel effet d'ordre, il est possible que les stratégies déployées par les sujets dans une condition donnée aient été influencées par les

conditions expérimentales précédentes. Par exemple, le fait de commencer avec la condition « Différenciée » a pu induire un type de stratégie qui s'est révélée incompatible avec les conditions suivantes, ce qui a pu influencer les résultats. De futures analyses auront alors vocation à tester un éventuel effet d'interaction des conditions expérimentales sur les stratégies individuelles et collectives. Par ailleurs, il serait intéressant de séparer ces conditions expérimentales selon des groupes expérimentaux indépendants. De cette façon, nous pourrions proposer plus d'essais, de façon à favoriser l'appropriation du système de couplage par les utilisateurs, et ainsi permettre une meilleure stabilisation de la dynamique d'interaction pouvant conduire à de meilleurs résultats.

Quoi qu'il en soit, ces résultats soulèvent la question de la configuration optimale des corps-image/corps-percevant pour que les utilisateurs de notre dispositif puissent coordonner leurs activités perceptives autour d'objets communs. Nous avons vu que dans les conditions de l'expérimentation, la différenciation des stimulations objet/autrui favorisait la reconnaissance collective qu'un objet du monde est partagé. Cette reconnaissance est rendue possible par un arrêt de l'utilisateur sur l'objet, qui peut alors faire sens de l'arrêt de son partenaire comme l'indice d'une activité perceptive intentionnelle dirigée vers soi et vers l'objet. Or, dans le cadre des espaces numériques bidimensionnels, les utilisateurs ne doivent pas seulement reconnaître que l'autre perçoit un objet, mais ils doivent coordonner leurs activités perceptives pour explorer ensemble le contenu partagé. De cette façon, la configuration « Différenciée » ne nous semble pas transposable telle quelle dans de tels espaces, dans la mesure où les stimulations, si elles sont différenciées, sont ici délivrées en tout ou rien. Ceci ne nous semble pas propice à l'exploration mutuelle de contenus spatialisés, et ce pour plusieurs raisons :

- Il a été souligné que le parallélisme du corps-percevant favorisait l'exploration et la reconnaissance de formes numériques bidimensionnelles, en permettant d'accéder à des informations spatiales dès l'entrée sensorielle (voir Sribunruangrit, 2004, et Partie 1, section 1.2.3, p. 25).
- Nous avons souligné que les stimulations en tout ou rien ne favorisaient pas la coordination des activités perceptives dans les espaces numériques partagés bidimensionnels, en rendant le guidage interpersonnel difficile (voir Partie 2, section 6, p. 168).

De façon à tirer parti de la différenciation des stimulations objet/autrui, et du parallélisme des champs récepteurs, une solution consisterait à utiliser deux plages de deux cellules Braille piézoélectriques, comme sur notre module d'interaction tactile. Mais comme nous l'avons vu, l'accès simultané à ces deux plages n'est pas forcément confortable, et l'utilisation simultanée de ces deux informations pourrait compliquer la tâche du sujet. Une deuxième solution consisterait à utiliser des modalités sensorielles différentes, en signalant par exemple la présence de l'autre par des retours sonores. Mais ces retours pourraient alors interférer avec les informations sonores ou verbales relatives au contenu partagé lui-même. La condition « Parallélisme » nous semble être la configuration la plus indiquée pour l'exploration mutuelle de contenus bidimensionnels.

Afin d'évaluer la pertinence de la condition « Parallélisme » dans un tel contexte, nous avons donc procédé à de nouvelles analyses d'usage en situation écologique.

8. Retour sur les usages : exploration mutuelle d'un contenu partagé

De façon à analyser l'usage du dispositif dans un contexte d'exploration mutuelle d'un contenu partagé, nous avons constitué de nouveaux scénarii d'usage. De façon à mener ces observations, nous faisons de nouveau appel à **C.** (l'utilisateur expert mobilisé dans l'ensemble des analyses d'usage précédentes) et **L.** (l'utilisateur novice déjà mobilisé dans les analyses d'usage du dispositif dans le cadre des interactions dyadiques ; voir Partie 2, section 6, p. 168). Dans ces situations d'exploration mutuelle, nous nous sommes particulièrement intéressés aux configurations des corps-image/corps-percevant qui offrent un parallélisme des champs récepteurs avec les stimulateurs tactiles (soit les configurations renvoyant à un parallélisme de type I et à un parallélisme de type II). De cette façon, nous avons testé deux des trois configurations évaluées dans le cadre des analyses d'usage des interactions dyadiques (voir Figure 79). Pour rappel, voici le détail de ces configurations :

- **Parallélisme de type I** : les deux sujets ont un corps-percevant constitué de 16 champs récepteurs élémentaires de 8 pixels chacun. La surface totale du corps-percevant est donc de 32 x 32 pixels. Les corps-image sont des surfaces carrées de 32 x 32 pixels.
- **Parallélisme de type II** : les deux sujets ont un corps-percevant constitué de 16 champs récepteurs élémentaires de 8 pixels chacun. La surface totale du corps-percevant est donc de 32 x 32 pixels. Les corps-images sont des surfaces carrées de 8 x 8 pixels

situées au centre du corps-percevant de chaque sujet : la surface de contact est par conséquent plus petite que dans le parallélisme de type I.

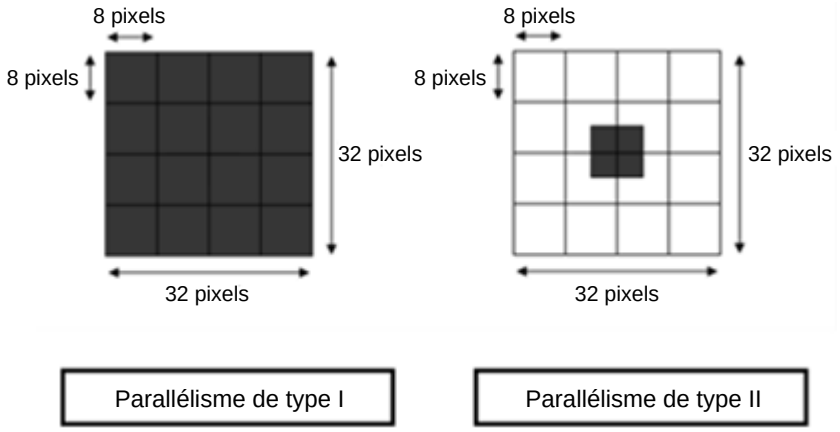


Figure 79 : Configurations des corps-image/corps-percevant utilisées dans les observations contrôlées concernant l'exploration mutuelle d'un contenu partagé

8.1. Scénario 1 : exploration mutuelle simple

Nous élaborons alors un premier scénario d'usage dans lequel les sujets ont pour objectif d'explorer ensemble une ligne horizontale située dans un espace partagé de 800 x 600 pixels. Nous n'attribuons pas de rôle à priori aux deux utilisateurs, et nous leur demandons de privilégier les interactions tactiles aux instructions verbales, sauf quand cela leur paraît nécessaire. Chaque configuration donne lieu à un essai de 3 minutes.

8.1.1. Parallélisme de type I

Dès le début de l'essai, **C.** semble prendre le rôle de guideur, alors qu'aucune consigne spécifique n'était donnée dans ce sens. Il part à la recherche de la ligne, qu'il trouve très rapidement, pendant que **L.** attend dans le coin supérieur gauche. Le point de rendez-vous négocié lors des analyses d'usage en situation dyadique semble alors être conservé, malgré les deux mois qui séparent ces deux sessions. **C.** vient alors la chercher et l'emmène vers la ligne. Comme lors des scénarii de guidage, le suivi des deux utilisateurs dans l'espace est très fluide et ils ne perdent que très peu le contact. Quand les deux sujets arrivent à la ligne, **L.** la détecte tout de suite, et commence à l'explorer, tandis que **C.** la parcoure très rapidement pour en localiser les extrémités. Il revient alors au contact avec **L.** et tente de l'emmener vers la droite. Pour cela, la même stratégie que dans les phases de guidage est utilisée : il se place à côté de **L.**, à la limite de la perte de contact, et attend qu'elle amorce un mouvement dans cette direction. Lorsque **L.** s'engage dans cette direction, il la devance le long de la ligne. La

présence simultanée des stimulations résultant de la ligne avec les stimulations résultant du croisement perceptif ne semblent pas affecter les deux sujets. En revanche, lorsque **C.** arrive au bout de la ligne, il s'arrête, ce qui provoque l'arrêt de **L.** à ses côtés. **C.** comprend alors que **L.** ne perçoit pas l'extrémité elle-même. Il s'éclipse alors, se plaçant à proximité mais sans toucher ni la ligne ni **L.** Cette dernière reprend alors l'exploration autonome de la ligne jusqu'à l'extrémité. Après quelques secondes, **C.** revient au contact et il décide d'emmener **L.** jusqu'à l'autre extrémité de la ligne. Le suivi sur la ligne est très fluide et rapide. **C.** devance **L.**, parfois trop vite, mais en revenant au contact lorsqu'il la perd. Arrivés au bout, **C.** s'éclipse de nouveau et laisse **L.** explorer l'extrémité. Au bout de l'essai, la ligne a été explorée en totalité par les deux sujets à plusieurs reprises. Le contact a été parfois perdu, mais ils ont pu se retrouver systématiquement le long de la ligne, ce qui leur a permis de poursuivre leur exploration mutuelle. Au final, le taux de contact entre les deux sujets est élevé (75%), et les taux de contact respectifs entre les deux sujets et la ligne sont globalement similaires (64% pour **L.** et 69% pour **C.**). Les deux sujets commentent alors cet essai (voir Verbatim 13).

L. : « *(…) on l'a vraiment fait ensemble. Je sentais où il était, parce que la ligne est plus petite (…). En fait, comme il est à côté de moi, je sens d'un côté la ligne et de l'autre **C.**, qui couvre tous les picots. J'essayais de garder la ligne au milieu des picots, pendant que je le suivais de l'autre côté et avec tous les picots.* »

C. : « *C'était facile. Je sentais aussi les deux, et je savais où était la ligne et où était **L.** A chaque fois, je la sentais aller dans ma direction, et on y arrivait.* »

Verbatim 13 : Commentaires des deux sujets à la fin de l'essai effectué avec le parallélisme de type I

8.1.2. Parallélisme de type II

Nous procédons une nouvelle fois à cet exercice, en utilisant maintenant le parallélisme de type II, caractérisé par un corps-image plus petit que le corps-percevant. Le scénario global reste identique, et les deux sujets arrivent ensemble sur la ligne très rapidement. Le suivi a été particulièrement efficace, et **L.** détecte la ligne dès le premier contact avec celle-ci. Immédiatement, **C.** parcoure rapidement la ligne en localisant les deux extrémités, mais il ne semble pas en mesure de retrouver **L.** Il explore alors les alentours, puis parcoure à nouveau la ligne en maintenant cette fois le contact sur le bas de ses stimulateurs. De son côté, **L.** reste accrochée à la ligne, en maintenant les stimulations au centre de sa matrice de stimulateurs. **C.**

entreprend alors l'exploration de la ligne à l'aide d'un balayage vertical. Il parvient alors à localiser **L.**, mais les deux sujets ne parviennent pas à se suivre de façon fluide. **C.** donne alors quelques conseils à **L.** (voir Verbatim 14).

> **C.** (à **L.**) : « *Il faut être l'un à côté de l'autre, sinon on se sent moins bien (..). Essaie de garder la ligne sur les picots que tu veux, mais pas trop près, pour l'explorer. Si d'autres picots se lèvent quand tu bouges pas, c'est que c'est moi qui passe.* »

Verbatim 14 : Commentaire de **C.** adressé à **L.** lors de l'essai effectué avec le parallélisme de type II

Petit à petit, **L.** se détache de la ligne en continuant de l'explorer. Elle se décale vers le haut de façon à maintenir les stimulateurs du bas activés lorsqu'elle se déplace le long de la ligne. **C.** la détecte de plus en plus facilement, et à partir du moment où le corps-image de **L.** n'est plus en contact avec la ligne, les deux sujets ne se perdent quasiment plus. Ils explorent alors plusieurs fois la ligne d'un bout à l'autre, et **C.** n'a même plus à s'éclipser pour la détection des extrémités. **L.** semble anticiper la position de la ligne, en procédant à un suivi continu classique, mais uniquement avec les picots du bas. Toutes les autres stimulations reçues sont donc attribuables à la présence de **C.**, qui la devance légèrement pour l'emmener dans la direction souhaitée. Au final, le taux de contact entre les deux sujets est élevé (77%), tout comme le taux de contact des deux sujets avec la ligne (78% pour **C.** et 81% pour **L.**). Les commentaires recueillis à la fin de l'essai sont par ailleurs très encourageants (voir Verbatim 15).

> **L.** : « *J'essaie de rester accrochée à la ligne, et je sais que tout ce qui bouge sans moi, c'est **C.**, et je n'ai plus qu'à le suivre en faisant attention à la ligne.* »
>
> **C.** : « *Le contact, c'est seulement 1 ou 2 picots, mais si on fait attention, on sait où est l'autre quand il n'est pas directement sur la ligne. En explorant, je savais tout le temps ce que faisait **L.*** »

Verbatim 15 : Commentaires des deux sujets à la fin de l'essai effectué avec le parallélisme de type II

8.2. Scénario 2 : exploration mutuelle complexe

Pour ce deuxième scénario, nous nous intéressons cette fois à l'exploration mutuelle d'un contenu complexe, caractérisé par un enchevêtrement de lignes orthogonales présentant différents croisements (voir Figure 80).

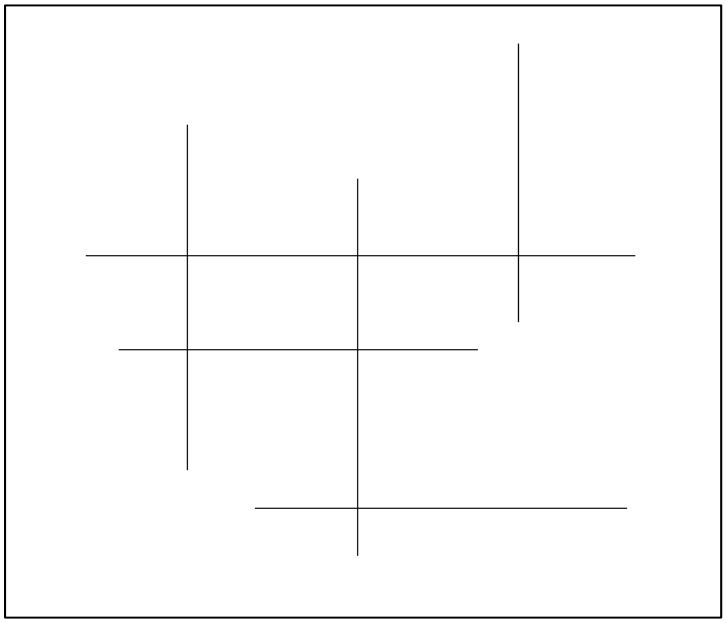


Figure 80 : Image utilisée pour l'exploration mutuelle complexe. L'espace numérique partagé a une dimension totale de 800 x 600 pixels

Les sujets sont informés de la nature de la forme à explorer. Comme lors du scénario précédent, nous demandons aux sujets d'explorer ensemble la forme, en essayant de ne pas se perdre lors des croisements qu'ils vont rencontrer. Nous n'attribuons pas de rôles aux sujets, et leur demandons d'éviter dans la mesure du possible de se donner des indications verbales lors de l'exploration. Comme lors du scénario précédent, le parallélisme de type I et II sont successivement testés lors d'essais de 3 minutes.

Outre les indicateurs retenus lors du scénario précédent (taux de contact entre les deux sujets et taux de contact respectifs des sujets avec la forme), nous nous intéressons ici à trois nouveaux indicateurs :

- le nombre de pertes prolongées de contact des sujets avec la forme (une perte de contact prolongée étant définie comme une absence de stimulations résultant du contact avec la forme pendant une durée d'au moins 3 secondes) ;
- le nombre de pertes prolongées de contact entre les deux sujets (une perte de contact étant définie comme une absence de stimulations résultant du contact entre les deux sujets pendant une durée d'au moins 3 secondes) ;
- et le nombre de croisements correctement négociés : les deux sujets ont franchi ensemble un croisement, ou se sont présentés ensemble à un croisement et se sont

retrouvés après le croisement dans la bonne direction au maximum 3 secondes après la perte de contact. Notons qu'un même croisement peut être compté plusieurs fois.

8.2.1. Parallélisme de type I

Nous commençons ce scénario avec le parallélisme de type I, caractérisé par un corps-percevant et un corps-image de taille identique pour les deux participants.

Comme lors des scénarii précédents, **C.** semble une nouvelle fois prendre le rôle de guideur : il explore rapidement l'espace, croisant plusieurs fois les lignes mais sans s'y attarder, puis part à la recherche de **L.** dans le coin supérieur gauche, qui semble avoir de nouveau naturellement endossé le rôle de suiveur. Le premier contact entre les deux sujets survient rapidement (21 secondes). **C.** guide **L.** vers la droite, où ils croisent une première ligne. Dès le premier contact, **L.** semble vouloir explorer la ligne vers le haut, délaissant **C.** qui essayait de la guider vers le bas. Mais très rapidement, **L.** s'arrête, **C.** se place à la limite de la perte de contact dans la direction souhaitée, et les deux sujets s'engagent dans la même direction. **C.** devance légèrement **L.**, qui semble essayer de garder la ligne au centre de ces champs récepteurs, de façon à maintenir la stimulation résultant du contact avec la ligne au centre de sa matrice de stimulateurs. De cette façon, elle se laisse guider par **C.**, qui la devance tout en effectuant des petites oscillations latérales à mesure de leur progression le long de la ligne, comme s'il tentait d'anticiper les croisements. Comme lors des observations antérieures, le suivi dans l'espace est fluide, les déplacements sont simultanés et donnent lieu à très peu de pertes de contact : la présence de la ligne ne semble pas affecter le suivi, mais offre au contraire un repère commun aux deux utilisateurs. Arrivés au premier croisement, **C.** quitte **L.** pour explorer les possibilités de direction, tandis que **L.** semble explorer le croisement en tant que tel. Lorsque **C.** revient au contact, **L.** masque l'intersection par son corps-image, ce qui conduit **C.** à douter de la présence du croisement. Il repart explorer les différentes directions, pour vérifier qu'elles conduisent toutes à l'endroit où est placée **L.** Là encore, la taille du corps-image semble poser problème, dans la mesure où il masque la ligne lors des contacts entre les deux sujets.

Malgré cette difficulté, une stratégie de franchissement des intersections se met vite en place : **C.** part en « éclaireur », explore les possibilités de direction, tandis que **L.** attend au croisement. Lorsque le choix est fait, **C.** revient au contact de **L.**, et se place à a la limite de la perte de contact dans la direction souhaitée. Alors qu'elle est quasiment immobile et contrôle de ce fait les stimulations reçues par le contact avec le croisement, l'arrivée de stimulations supplémentaires invite **L.** à reprendre l'exploration dans la direction indiquée par ce

« surplus » de stimulations. Parfois, **L.** semble délaisser l'interaction, de façon à explorer quelques embranchements. Dans ce cas, il arrive que les sujets se perdent de façon prolongée, **C.** recherchant **L.** au niveau de l'intersection, avant de partir à sa recherche en parcourant les différentes directions au hasard, mais sans s'éloigner de l'intersection. Par conséquent, le taux de contact entre les deux sujets s'élève à 58%. Par ailleurs, le taux de contact de **C.** avec la forme est de 68%, tandis que celui de **L.** est de 75%.

Au final, sur 5 intersections rencontrées, 3 ont été franchies conjointement par les deux sujets. Toutefois, l'exploration est assez lente, et beaucoup de pertes de contact entre les deux sujets sont à déplorer (5 pertes prolongées de plus de 3 secondes, auxquelles s'ajoutent de très nombreuses pertes de moins de 3 secondes). Notons toutefois que les pertes de contact prolongées entre les deux sujets sont toujours le fait de **C.**, qui cherche à localiser **L.** par rapport à l'intersection : **C.** quitte le contact intentionnellement pour explorer les alentours, et en déduire la position exacte de **L.** et de l'intersection. De façon générale, nous n'avons pas réellement observé de phases de guidage. **L.** semble explorer les lignes de façon individuelle, après avoir détecté la direction globale de **C.**

Dans cette exploration mutuelle, la présence des lignes permet d'organiser l'interaction. En effet, lors de cet essai, les deux sujets n'ont perdu la ligne pendant plus de 3 secondes qu'une seule fois chacun.

8.2.2. Parallélisme de type II

Nous procédons à un nouvel essai de 3 minutes, avec utilisant cette fois le parallélisme de type II, caractérisé par un corps-image plus petit que le corps-percevant.

Dès le début de l'essai, **C.** retrouve **L.** dans le coin supérieur gauche de l'écran, avant même d'explorer la forme. Par conséquent, le premier contact est très rapide (6 secondes), et les deux utilisateurs s'engagent en direction de la forme. Cette fois encore, **C.** assume le rôle de guideur : il devance **L.**, et conformément aux observations antérieures avec cette configuration de corps-image/corps-percevant, le guidage est de très bonne qualité : les deux sujets se déplacent conjointement, et leur interaction dérive en direction de la forme tout en maintenant le contact. Lorsque **C.** arrive au contact de la forme, il est immédiatement suivi par **L.**, et les deux sujets maintiennent un contact statique pendant quelques secondes. Mais lorsque **C.** amorce un mouvement vers le bas le long de la ligne, il n'est pas suivi par **L.** : il revient alors au contact puis effectue des petites oscillations autour de **L.,** qui semble alors étonnée (voir Verbatim 16).

L. (à **C.**) « *Ah ! Tu es là… Je croyais qu'on s'était perdus* »

C. (à **L.**): « *Mais non, on s'est suivis depuis le début !* »

L. (à C.) : « *Ben ça commence bien… J'ai bien senti qu'on était arrivés sur la forme, mais là je ne te sentais plus.* »

C. (à **L.**) : « *On est plus petits qu'avant, donc faut faire attention avec la ligne.. .*»

Verbatim 15 : Commentaires des deux sujets à l'arrivée sur la forme lors de l'essai effectué avec le parallélisme de type II

Suite à cet échange, les deux sujets procèdent à de petites oscillations autour de leurs corps-images respectifs tout en restant en contact avec la forme, comme pour bien discriminer ce qui relève de la ligne de ce qui relève du corps-image de l'autre. Puis **C.** amorce un nouveau mouvement vers le bas, mais cette fois en se décalant légèrement par rapport à la ligne. De cette façon, **L.** peut sentir distinctement le corps-image de **C.** et la ligne, ce qui l'amène à se décaler elle aussi dans la direction de **C.**, de telle façon que son corps-image ne se confond plus avec la ligne. Les deux sujets se déplacent alors dans la même direction, en maintenant un contact mutuel. Lors de ce déplacement commun, nous observons que **C.** alterne entre contacts et absence de contacts avec la ligne, mais essaie toujours de maintenir quoi qu'il arrive le contact avec le corps-image de **L.**, tandis que cette dernière maintient quoi qu'il arrive un contact avec la ligne. Notons que nous assistons à une alternance des deux situations que nous avons identifiées lors des observations préliminaires (voir Partie 1, Section 3.1, p. 68), et que nous avions alors qualifiées d' « exploration supervisée » et d' « exploration mutuelle ». Ainsi, les deux sujets arrivent rapidement au premier embranchement.

Contrairement à la configuration précédente, celui-ci ne semble pas poser de gros soucis. **L.** explore rapidement l'embranchement en tant que tel, tandis que **C.** se positionne tout de suite sur l'un des côtés, en faisant en sorte que son corps-image ne soit pas en contact avec la ligne. Lorsque **L.** arrive au contact, elle semble hésiter mais suit la direction indiquée par **C.** Les deux utilisateurs s'engagent alors dans la poursuite de l'exploration mutuelle, **L.** restant accrochée à la ligne tandis que **C.** la devance, lui indiquant la direction à suivre sans interférer avec son exploration. Les autres embranchements rencontrés ont alors conduit à la même stratégie, et les deux utilisateurs sont parvenus à ne pas se perdre jusqu'à la fin de l'essai.

Au final, sur 6 embranchements rencontrés, tous ont été correctement négociés. Dans tous les cas, **C.** maintient une présence aux environs de la ligne, et tente d'entraîner **L.** dès

qu'elle rentre en contact avec lui. A la fin de l'essai, le taux de contact entre les deux sujets s'élève à 64%. Par ailleurs, il est intéressant de noter que le taux de contact des deux sujets avec la forme reflète la stratégie de guidage adoptée par **C.** : **L.** obtient un taux de contact avec la forme élevé (76%), tandis que celui de **C.** avec la forme est plus faible (58%). Parallèlement, une seule perte de contact prolongée de **L.** avec la forme est à déplorer, tandis que 8 pertes prolongées sont observées chez **C.** Mais de façon cruciale, le nombre de pertes de contact prolongé entre les deux sujets est faible (2), ce qui atteste d'un guidage efficace le long du parcours. Cette efficacité de l'interaction est d'ailleurs dénotée par les deux sujets à la fin de l'essai, qui semblent satisfaits de leur prestation (voir Verbatim 16)

L. (à **C.**) « *Super ! Je t'ai suivi sans problème, et sans perdre la ligne…* »

C. (à **L.**): « *Oui on a bien réussi à se comprendre (…) à chaque fois je me mettais à côté pour que tu puisses bien sentir la forme, mais que tu puisses aussi savoir où j'étais.* »

L. (à C.) : « *Oui, c'était nickel comme ça. Je savais où tu voulais aller sans me poser la question (…). Donc avec un peu d'entraînement, on a fini par y arriver !* »

Verbatim 15 : Commentaires des deux sujets à la fin de l'essai effectué avec le parallélisme de type II

8.3. Synthèse des observations contrôlées

Dans le contexte d'une observation contrôlée, nous avons évalué la possibilité d'exploration mutuelle d'un contenu numérique en comparant deux configurations de corps-image/corps-percevant. Celles-ci sont caractérisées par un parallélisme des champs récepteurs avec les stimulateurs tactiles. D'une façon générale, nous pouvons conclure que les exercices proposés ont été un succès : les deux utilisateurs ont pu coordonner leurs activités perceptives de façon à se guider mutuellement le long d'une ligne (exploration mutuelle simple) ou dans un parcours constitué de lignes enchevêtrées (exploration mutuelle complexe). Dans tous les cas, il semble que la présence de la forme fournit un repère commun aux deux utilisateurs, qui organisent leur interaction autour de lui et qui sert de lieu de rencontre lorsque le contact mutuel est perdu.

Toutefois, dans les deux cas, nous avons relevé quelques limites. Avec le parallélisme de type I, nous avons observé un phénomène d'occlusion perceptive résultant de la taille des corps-image. Lors d'un contact mutuel, le corps-image peut en effet masquer le contenu exploré, ce qui ne favorise pas l'exploration mutuelle, particulièrement dans le cas d'un

croisement entre plusieurs lignes. Mais dans ce cas, nous avons vu que l'expertise de **C.** est encore une fois un atout essentiel. Il parvient à comprendre lorsque son propre corps-image est une entrave à l'exploration de son partenaire, et organise son activité en conséquence, en se retirant de l'interaction lorsqu'il le juge nécessaire.

Si nous nous attendions à ce que cette configuration des corps-image/corps-percevant favorise la coordination des deux utilisateurs, en offrant un corps-image plus saillant qu'avec le parallélisme de type II, nous avons noté que le parallélisme de type I entraîne paradoxalement plus de pertes de contact mutuel et une exploration globalement plus lente. Nous interprétons ce résultat au regard de la stratégie d'exploration déployée par le sujet novice. Celle-ci est peut-être en effet gênée par les stimulations résultant du croisement perceptif, et préfère se focaliser sur la forme elle-même, ce qui ne l'empêche pas de se référer à la position de **C.** lorsqu'elle le juge nécessaire, et notamment lorsqu'il s'agit de choisir une direction dans laquelle s'engager. De cette façon, nous pouvons supposer que l'exploration est le plus souvent individuelle.

De son côté, le parallélisme de type II exige une coordination interpersonnelle plus complexe. Cette fois, ce n'est pas le corps-image qui risque de masquer la ligne, mais l'inverse. Malgré tout, dans ces conditions, les utilisateurs ont pu s'organiser pour dépasser ce problème, même si cela a nécessité au départ une instruction verbale explicite. Nous avons alors assisté à une alternance d'exploration mutuelle et d'exploration supervisée, contrôlée par l'utilisateur expert, qui a permis à l'utilisateur novice d'organiser son activité perceptive en fonction de l'autre et en fonction de la ligne. De cette façon, nous avons noté que si l'interaction présentait moins de pertes de contacts, l'utilisateur expert a pour sa part un nombre plus élevé de pertes de contact avec la forme. En effet, sa stratégie a consisté davantage à assister l'exploration de son partenaire plus qu'à explorer la forme avec lui.

Si ces observations nous semblent encourageantes pour envisager la collaboration d'utilisateurs distants dans des espaces numériques partagés, il nous semble nécessaire de poursuivre ce travail. Comme nous l'avions suggéré lors des analyses d'usage préliminaires, une piste sérieuse consisterait à pouvoir agir sur la taille de son corps-image en temps réel, voire à le masquer dans certains cas. Mais cette possibilité présuppose une expertise de la part de l'utilisateur, qui doit pouvoir comprendre à quel moment sa présence est une entrave et à quel moment sa présence est nécessaire pour l'autre.

De la même façon, nous avons testé l'exploration mutuelle dans des espaces numériques simplifiés. La mise à disposition de contenus riches hébergés via le portail Intertact doit

maintenant conduire à des observations d'usage, de façon à poursuivre l'analyse de la coordination interpersonnelle dans les espaces numériques partagés.

SYNTHESE GENERALE

Dans le cadre de la conception d'un dispositif de suppléance perceptive donnant accès à des espaces numériques partagés, notre objectif a consisté à définir les conditions techniques pour que les utilisateurs puissent interagir à distance via la modalité tactile. En d'autres termes, il a été question d'envisager l'extension de l'usage individuel d'un dispositif de suppléance perceptive existant (Tactos) en un usage en réseau (Intertact). Pour cela, nous avons mené des analyses d'usage préliminaires en collaboration étroite avec de futurs utilisateurs. Ces analyses ont conduit à un ensemble de spécifications techniques et fonctionnelles, qui nous ont permis notamment d'élaborer un premier prototype de module spécialement destiné à l'accessibilité d'espaces numériques partagés.

Malgré nos précautions initiales pour orienter la conception à partir de données d'usage et de tests-utilisateurs, il apparaît que ce premier prototype présente des limites. En effet, nous avons vu que le volume global du module pose des problèmes de saisie et d'accès simultané à ses différents composants, notamment pour les utilisateurs les plus jeunes. Si ces problèmes peuvent être compensés par la multiplicité des prises en main possibles, celle-ci implique un ensemble d'opérations de paramétrage qui ne facilite pas une flexibilité réelle du module en situation d'usage. Pour autant, nous avons pu tout de même valider de façon empirique l'intérêt d'un Touchpad configuré en absolu, qui permet de s'affranchir de la tenue d'un stylet, et qui nous semble plus approprié pour les interactions tactiles distales. Quoi qu'il en soit, dans une optique d'évaluation corrective, l'analyse de l'usage de ce module nous a permis un ensemble de validations empiriques et de nouvelles spécifications. Ces résultats ont conduit à proposer de nouveaux effecteurs, comme les écrans tactiles disponibles directement via les nouvelles technologies de l'information et de la communication (Smartphones et tablettes tactiles). Les avantages de ces nouveaux effecteurs sont nombreux :

- la plupart des utilisateurs potentiels sont déjà équipés de ces dispositifs, et il suffira d'y coupler des matrices de stimulateurs tactiles (disponibles par ailleurs sur des barrettes Braille dont ils sont également équipés) ;
- ils répondent à une miniaturisation du volume global de l'effecteur (spécialement en ce qui concerne les Smartphones) ;
- les avancées technologiques dans le domaine des écrans tactiles en font des effecteurs très précis et agréables au toucher ;

- ils permettent d'associer le déplacement du corps-image avec ses déformations en temps réel à travers un seul et même effecteur (le caractère multipoint de l'écran pouvant capter une surface de pression) ;
- ils permettent de désolidariser l'effecteur du poste de travail, et donc d'ancrer l'usage de Tactos dans un environnement situé, ce qui ouvre la voie à de nouvelles fonctionnalités[51]

Parallèlement, nos observations et notre collaboration avec les utilisateurs ont permis d'aboutir à la constitution d'une première version d'un portail Internet d'interaction tactile, destiné à héberger des espaces numériques partagés. D'une part, nous avons contribué à la définition d'une méthode de navigation tirant parti de l'organisation spatiale de l'information (le « Treemapping »). D'autre part, nous avons défini un ensemble de spécifications fonctionnelles concernant l'accessibilité de contenus riches. Pour cela, nous avons soumis des exemples inspirés des supports traditionnels à deux utilisateurs dans le cadre des suivis longitudinaux, que nous avons pu enrichir progressivement en fonction de nos observations et des commentaires recueillis. Nous avons alors relevé que le recours à la bimodalité était prometteur, dans la mesure où celle-ci permet de délivrer des informations sonores ou verbales situées lors de l'exploration des utilisateurs. En revanche, il serait utile de laisser l'utilisateur le soin de déclencher leur lecture, par l'intermédiaire d'une commande spécifique. Par ailleurs, nous avons également suggéré l'usage de textures tactiles, qui semblent utiles pour spécifier des zones d'intérêt à l'intérieur de ces contenus, et celui du zoom, qui permettrait des changements d'échelle intéressants. Tout cela encourage la poursuite d'analyses d'usage, qui pourront prendre une autre dimension lorsque le portail sera définitivement mis en place et accessible à une communauté d'utilisateurs.

Parallèlement à l'étude des conditions techniques qui les rendent possibles, les espaces numériques partagés soulèvent la question de l'interaction tactile d'utilisateurs distants. Nous avons alors mené des observations préliminaires, qui nous ont permis de vérifier la pertinence de notre dispositif et de définir des premières problématiques d'usage. Nous avons alors remarqué que les configurations de corps-image/corps-percevant ont une importance fondamentale, en ce qu'elles favorisent ou non la collaboration des utilisateurs selon les tâches proposées. Par ailleurs, il a été question de conduire une étude expérimentale visant à

[51] Ces nouveaux usages ont donné lieu au projet ITOIP (Interactions Tactiles pour l'Orientation, l'Information et la Présence), qui fait suite au projet « Portintertact ». Il est financé par la région Picardie de Septembre 2011 à Septembre 2014.

tester les possibilités de discrimination interpersonnelle. Nous en avons conclu que celle-ci est possible, tout en ayant soulevé quelques points critiques :

- La présence simultanée de quatre utilisateurs dans l'espace partagé n'a pas été concluante, la présence des uns pouvant interférer avec la présence des autres, et ainsi masquer une éventuelle signature déployée par chaque utilisateur[52]. Les résultats ont été meilleurs lorsque les sujets se sont rencontrés deux à deux.
- L'utilisation d'effecteurs absolus semble favoriser l'interaction, dans la mesure où ceux-ci permettent une meilleure maîtrise de l'engagement individuel dans l'espace partagé.
- L'utilisation conjointe d'un Touchpad configuré en absolu et d'une nappe de capteurs de pression (FSR) destinée à la déformation du corps-image en temps réel ne s'est pas révélée satisfaisante. L'utilisation de cette déformation n'a pas conduit à une meilleure discrimination interpersonnelle, et les sujets ont même massivement abandonné cette possibilité au cours de la tâche pour se concentrer sur les déplacements.
- Enfin, lorsque les sujets se sont mutuellement identifiés, nous avons constaté que leurs activités perceptives avaient convergé l'une vers l'autre au cours de l'essai, ce qui témoigne d'une organisation spécifique de l'interaction.

Ce dernier résultat a particulièrement motivé notre recherche, en permettant de soulever des questions théoriques fondamentales. En effet, l'organisation spatio-temporelle de l'interaction nous a semblé être le support de la discrimination interpersonnelle. Cette hypothèse nous a conduit à adopter une approche pleinement interactionniste de la cognition sociale : la participation active et l'engagement mutuel des participants font émerger une dynamique relationnelle autonome qui devient le support de la compréhension sociale (De Jaegher et al., 2010).

Prendre en compte sérieusement ce processus passe d'abord par un rejet de l'individualisme méthodologique (Lenay & Stewart, 2012 ; Froese & Di Paolo, 2011). L'étude de la cognition sociale est alors menée à partir de situations d'interaction concrètes et contextualisées, telle que notre dispositif les permet. De plus, il s'agit de ne pas circonscrire la cognition sociale à des mécanismes individuels (voir par exemple Gergely & Watson, 1999 ; Tomasello, 1995 ; 1999), mais d'étudier le système de couplage lui-même, qui implique les

[52] Notons que ce constat soulève la question des interactions triadiques (sujet-sujet-sujet) dans le cadre des espaces numériques partagés.

participants, leurs interactions dynamiques et le contexte dans lequel ces interactions s'inscrivent (Chapman, 1991 ; De Bruin & De Haan, 2012). Pour autant, replacer la cognition sociale au cœur de l'interaction ne revient pas à dissoudre le rôle de l'individu. Au contraire, il s'agit d'identifier ce qui relève des contributions individuelles, notamment en ce qui concerne l'émergence et le maintien du processus de l'interaction, de ce qui résulte de ce processus lui-même et qui échappe aux individus. Parallèlement, il s'agit de comprendre comment ce processus permet aux individus de co-construire du sens, notamment via une re-saisie individuelle de la dynamique collective (Lenay & Stewart, 2012).

Pour cela, le minimalisme nous semble être un cadre d'analyse pertinent, notamment en ce qu'il permet d'observer une néo-genèse de la cognition sociale. L'usage de médiations techniques permet en effet d'analyser la constitution même du processus de l'interaction et de la co-construction de sens, à travers la découverte d'un nouvel espace de couplage entre les individus, où les savoir-faire ne sont pas constitués à priori (Lenay, 2006c). Par ailleurs, en simplifiant à l'extrême le répertoire des actions et des retours sensoriels possibles, cette constitution s'offre à l'analyse précise par l'étude du déploiement des engagements individuels et collectifs (ibid.).

Dans ce contexte, l'étude des croisements perceptifs minimalistes a déjà permis de souligner un point essentiel : c'est la dynamique mutuelle de l'interaction qui amène les sujets à se reconnaître les uns les autres par rapport à des leurres, qu'ils soient fixes ou mobiles (Auvray et al., 2009). En d'autres termes, c'est l'activité perceptive dirigée vers soi qui est reconnue, plus que la structure objective des mouvements du partenaire (Lenay, 2010), même si celle-ci peut offrir des informations locales telles que des relations contingentes par rapport à l'observateur (Gergely, 2001 ; Rochat & Striano, 1999). De façon à mettre ce résultat à l'épreuve, nous avons conduit une expérimentation visant à étudier l'effet de la variation du degré de contingences sociales produites par un robot sur la reconnaissance interpersonnelle, dans le contexte d'un test de Turing minimaliste non-verbal. Nos résultats principaux suggèrent que si la présence de contingences imparfaites a un effet sur la fréquence de réponses « Humain », ces contingences sont avant tout produites par le sujet qui interagit plus que par un robot qui les donnerait à percevoir. En d'autres termes, nous suggérons que dans le cas d'une rencontre interpersonnelle, les contingences sont produites dynamiquement dans le couplage de leurs activités perceptives. En effet, le sujet ne cherche pas à agir de façon contingente, offrant ainsi une information à son observateur, mais il déploie une activité perceptive à la recherche d'une source de stimulation. La rencontre de ces activités est alors caractérisée par une accroche perceptive élevée mais variable, et également sur une alternance

d'engagements et de désengagements qui signe la reconnaissance mutuelle. Ainsi la co-régulation active des deux partenaires fait émerger une dynamique caractérisée par une variabilité tant au niveau microscopique qu'au niveau macroscopique.

Comme nous l'avons suggéré, nous pourrions envisager de complexifier davantage les robots, de façon à poursuivre cette expérience jusqu'à ce que les sujets ne puissent plus faire la différence entre les autres sujets humains et les robots. Pour cela, nous pourrions implémenter dans les robots un plus haut degré de variabilité dans leurs mouvements, ainsi qu'une capacité dynamique à s'engager et à se désengager de l'interaction. Nous nous demandons alors jusqu'à quel point les sujets pourront apprendre à discriminer les humains et les robots au fur et à mesure des essais, et jusqu'à quel point ils pourront développer des stratégies de plus en plus complexes pour se démarquer des robots.

Parallèlement, si cette situation nous a semblée pertinente pour tenter de caractériser la dynamique des rencontres interpersonnelles, la signature objective de l'interaction pourrait se révéler utile dans le cadre de l'élaboration d'espaces tutoriels automatisés, où des robots pourraient assister l'utilisateur dans le processus d'appropriation du dispositif.

Comme nous l'avons vu, en effet, la collaboration d'un utilisateur expert avec un utilisateur novice révèle que le niveau de maîtrise du dispositif favorise le guidage interpersonnel dans l'espace. L'utilisateur expert peut assister le processus de l'interaction par l'intermédiaire de stratégies adaptées à la tâche et au niveau de maîtrise du partenaire. La dynamique de l'interaction elle-même est alors mobilisée par les utilisateurs, de façon à guider leurs activités perceptives pour la poursuite d'un but commun. Par ailleurs, nous avons également souligné que cette dynamique et sa re-saisie par les utilisateurs est dépendante des propriétés de la configuration des corps-image/corps-percevant. Dans un espace bidimensionnel, en effet, les stimulations en tout ou rien ne favorisent pas la maîtrise de la dynamique. L'interaction peut alors dériver en contrariant les intentions respectives des utilisateurs pour se suivre dans l'espace. Au contraire, l'utilisation de configurations assurant un parallélisme des champs récepteurs élémentaires avec les stimulateurs tactiles garantit un guidage efficace, en permettant la détection de la position de son partenaire par rapport à soi, ce qui favorise la synchronisation des activités perceptives en limitant la nécessité d'anticipations hasardeuses. Pour autant, nous avons également souligné que la taille du corps-image avait son importance. Lorsque celui-ci est de même taille que le corps-percevant (Parallélisme de type I), la surface de contact est accrue, et la rencontre interpersonnelle dans l'espace est favorisée. En revanche, lorsque le corps-image est plus petit que le corps-percevant, la coordination des activités perceptives suit plus de degré de liberté, à condition

que la rencontre puisse avoir lieu. En cela, nous avons suggéré l'implémentation d'une nouvelle commande-utilisateur, destinée à agir en temps réel sur la taille de son corps-image et/ou de son corps-percevant, ce qui pourra conduire à de nouveaux tests en situation d'usage.

Parallèlement à l'étude des interactions strictement dyadiques, nous nous sommes intéressés à la coordination interpersonnelle autour d'objets communs. Notre objectif a alors consisté à vérifier si les conditions techniques de l'interaction tactile sur le réseau pouvaient permettre des capacités d'attention conjointe de la part des utilisateurs. Cette question est cruciale, dans la mesure où notre projet consiste à concevoir des espaces numériques partagés dans lesquels les sujets pourront coordonner leurs activités perceptives autour d'objets communs. Conformément à l'approche interactionniste de la cognition sociale que nous avons adoptée, nous avons postulé que l'attention conjointe pouvait se réaliser via la coordination des activités perceptives. Ainsi, la dynamique relationnelle autonome que nous avons observée dans un contexte dyadique serait ici organisée en fonction des objets présents. De cette façon, nous avons élaboré une tâche collaborative, qui nécessite une coordination interpersonnelle de façon à atteindre un but prescrit par la tâche. En ce sens, nous avons pris le parti d'étudier les mécanismes de l'attention conjointe à partir d'une activité conjointe déjà définie. Mais cela ne veut pas dire que la connaissance issue de l'attention conjointe est déjà constituée : les sujets doivent crucialement s'engager dans la tâche, et s'approprier le dispositif pour être à même de donner du sens à l'engagement de leurs partenaires dans l'espace partagé.

Nos résultats principaux suggèrent que la connaissance du fait qu'un objet de l'environnement est partagé résulte d'un engagement mutuel de la part des participants plus qu'une simple « lecture » du comportement attentionnel de l'autre (Lempers, 1979 ; Moore et al., 1997) ou à une attribution individuelle d'intentionnalité (Tomasello, 1995 ; 1999). En effet, nos résultats montrent que la perception mutuelle se réalise dans le cadre du processus de l'interaction (Auvray et al., 2009 ; De Jaegher et al., 2010). La connaissance du fait que l'objet est partagé est donc co-construite à travers un processus actif et dynamique de négociation du sens qui se réalise au sein même du couplage interpersonnel (ibid.). Nous considérons alors l'attention conjointe comme une action conjointe basique (Fiebich & Gallagher, 2012), qui conduit à une modification du statut de l'objet, passant d'un objet privé, que l'individu rencontre dans le déploiement de son activité perceptive, à un objet commun, que les deux individus peuvent rencontrer dans le processus de leur interaction. (Sebanz et al., 2006).

En effet, nous avons vu que l'objet commun fournit une affordance simultanée (Marsh, et al., 2006), qui produit une attirance réciproque et augmente ainsi sensiblement la probabilité de rencontrer l'autre autour de cet objet. Mais percevoir l'autre et percevoir l'objet dans une zone donnée de l'espace ne signifie pas nécessairement que l'autre perçoit cet objet. La combinaison de l'affordance résultant de la rencontre de l'objet et d'une affordance sociale (Krueger, 2011) semble donner lieu à une coordination interpersonnelle autour de l'objet qui permet :

i) de donner du sens à l'activité de l'autre en tant qu'engagement actif dans le croisement perceptif (Auvray et al., 2009 ; Lenay & Stewart, 2012),
ii) de donner du sens à son engagement en tant qu'activité dirigée alternativement vers le soi et vers l'objet (Carpenter et al., 1998; Reddy, 2011), et
iii) de donner du sens à l'objet en tant que source de stimulation mutuellement perçue, ce qui conduit à la reconnaissance que l'objet de son attention est également l'objet de l'attention de l'autre (Baldwin, 1995).

Parallèlement, nous avons noté que la constitution de cette connaissance était influencée par les propriétés du système de couplage. Celui-ci offre en effet un répertoire d'actions dépendamment de la nature des retours sensoriels qu'il rend possible. Ces actions sont cruciales, dans la mesure où elles contraignent la dynamique interpersonnelle, et influencent par la même ce qui fait sens pour les interactants. Dans ce cadre, nous avons noté que lorsque les stimulations résultant du croisement perceptif et les stimulations résultant de la rencontre des objets sont différenciées, les performances sont les meilleures, et témoignent même d'un effet plafond dès le premier bloc d'essais. Les sujets peuvent alors faire sens de l'engagement de l'autre en fonction des objets présents, cet engagement étant ici manifesté par un arrêt de l'individu sur l'objet qu'il perçoit. La condition « Différenciée » semble donc être une solution d'usage pertinente, qu'il convient d'envisager dans le cadre de l'exploration mutuelle de contenus numériques.

Pour autant, nous avons rappelé que les meilleures performances d'exploration et de reconnaissance de formes sont obtenues avec une matrice de 4x4 champs récepteurs élémentaires couplés à autant de stimulateurs tactiles (Sribunruangrit, 2001 ; 2004). La différenciation des stimulations nécessiterait alors l'usage de plusieurs matrices de stimulateurs (comme sur notre prototype de module d'interaction tactile), dont on a souligné les limites. Nos dernières analyses d'usage en situation écologique ont donc consisté à tester

l’exploration mutuelle de contenus avec le parallélisme de type I et le parallélisme de type II. Ces analyses sont encourageantes, dans la mesure où nos deux utilisateurs ont pu coordonner leurs activités perceptives pour organiser leur interaction en fonction des formes inscrites à l’écran. Toutefois, nous avons remarqué les possibles interférences entre les deux sources de stimulations (la forme et le corps-image d’autrui), qui peuvent ponctuellement entraver la dynamique de l’interaction en masquant la forme à explorer (parallélisme de type I) ou le corps-image de son partenaire (parallélisme de type II). Ces deux situations ont permis de souligner une nouvelle fois l’impact de l’expertise sur les stratégies déployées. Ainsi, notre utilisateur expert a pu contourner ces difficultés, en évitant d’entraver l’exploration de son partenaire, ou en organisant son activité de façon à maintenir le contact sans interférer avec elle. De cette façon, nous avons pu observer une alternance d’exploration mutuelle et d’exploration supervisée, qui souligne toutefois la difficulté de réellement explorer ensemble un contenu numérique.

CONCLUSION ET PERSPECTIVES

Dans ce travail, nous avons mené une étude des interactions tactiles distales, telles qu'elles sont rendues possibles par la connexion en réseau d'un dispositif de suppléance perceptive. Nous avons pour cela articulé une recherche théorique fondamentale concernant les mécanismes de la cognition sociale, avec une recherche appliquée, destinée à concevoir un dispositif sociotechnique pertinent.

Notre principale contribution a alors consisté à définir les conditions techniques de l'interaction tactile sur le réseau, tout en questionnant la nature des mécanismes qui permettent aux utilisateurs de faire sens des rencontres interpersonnelles minimalistes. Nous avons alors argumenté en faveur d'une approche énactive et interactionniste de la cognition sociale : la coordination des activités perceptives résulte d'un processus actif d'engagement mutuel, qui devient le support d'une co-construction de sens. Ainsi, d'une problématique de conception et d'usage, nous avons alimenté une recherche fondamentale, qui, en retour, permet de guider l'élaboration d'espaces numériques partagés pertinents.

Cette circularité entre théorie et pratique a été favorisée par le caractère minimaliste de notre dispositif. L'intérêt principal du minimalisme revient à forcer le déploiement des activités individuelles et collectives. Notre dispositif étant numérique, il permet par ailleurs d'enregistrer ces activités, et donc de les rendre accessibles à l'analyse. Dans notre cas, cette analyse nous a permis d'avancer quelques éléments objectifs afin de soutenir nos hypothèses. Mais il faut souligner que nous n'avons utilisé que des indicateurs « statiques », qui résument l'interaction ou la séquence d'interaction en une valeur (la moyenne ou l'écart-type). Si celle-ci peut dans une certaine mesure être représentative de l'interaction, en indiquant une tendance globale par sujet et par essai, elle ne permet pas la plupart du temps de capturer l'essence de la dynamique de l'interaction. Par exemple, dans l'expérience de Turing, nous avons calculé l'écart-type des écarts temporels entre les stimulations pour les séquences où ces stimulations sont espacées de moins de 2 secondes. Nous avons ainsi représenté la fluctuation dynamique des micro-ajustements interpersonnels à travers un indicateur global de dispersion des valeurs obtenues dans ces séquences. Nous pensons alors que des outils d'analyse empruntés à l'approche des systèmes dynamiques permettraient d'affiner nos résultats, en utilisant par exemple des méthodes destinées à l'analyse du signal (analyse par temps-fréquence, analyse spectrale et fractale) de façon à isoler des patterns d'interaction propres à la coordination interpersonnelle.

Une meilleure compréhension de cette coordination interpersonnelle pourrait alimenter la conception de notre dispositif. Rappelons que celui-ci est destiné en premier lieu aux personnes déficientes visuelles. Pour cela, nous avons activement collaboré avec des utilisateurs volontaires, de façon à obtenir des données d'usage et à recueillir leurs attentes. Si cette collaboration a permis de guider les étapes de conception et a rendu possible des analyses d'usage en situation écologique, nous pouvons questionner la généralisation des résultats obtenus, au regard du faible nombre de participants, et des différences interindividuelles que nous avons malgré tout constatées. Nous avons notamment insisté sur le niveau d'expertise des utilisateurs, qui a un impact certain sur la réussite des tâches, y compris dans le cadre de l'interaction avec des utilisateurs novices. En effet, nous avons systématiquement analysé les interactions tactiles dans des relations asymétriques expert/novice. L'analyse d'interactions entre utilisateurs de niveau équivalent pourrait alors apporter de nouveaux éléments pour comprendre l'organisation des activités perceptives.

Paradoxalement, nous avons déployé des expérimentations en situation contrôlée, en faisant appel exclusivement à des sujets voyants. Nous pouvons alors légitimement nous poser la question de l'impact de leur expérience visuelle sur la réussite des tâches que nous leur avons proposé, tout comme nous pouvons nous demander si des sujets déficients visuels auraient obtenus les mêmes résultats. Dans ce cadre, nous pouvons mentionner ici que nous avons procédé à la passation de l'expérimentation « Objets communs et Objets privés » par deux utilisateurs déficients visuels ayant une certaine expérience de notre dispositif (**C.** et **L.** par ailleurs mobilisés dans les analyses d'usage que nous avons présenté). Nous n'avons pas pu exposer leurs résultats dans ce travail, mais nous pouvons noter qu'ils ont particulièrement bien réussi la tâche, et ce, dans les trois conditions expérimentales proposées. Ceci nous encourage à procéder à de nouvelles passations, en mobilisant une population plus étendue de sujets déficients visuels, de façon à vérifier si cette réussite peut être attribuée à une utilisation plus efficiente de la modalité tactile, ou à une expertise de notre dispositif.

Dans ce cadre, nous pouvons également indiquer que cette expérience a donné lieu à une recherche menée par une étudiante de l'université de Rouen, qui a comparé les performances d'adolescents déficients visuels avec les performances d'adolescents voyants. Les résultats montrent cette fois que les adolescents voyants ont obtenu de meilleurs résultats que les adolescents déficients visuels. En outre, les premières analyses de trajectoire témoignent de stratégies de résolution qui diffèrent en fonction des statuts visuels des participants. Quand les sujets voyants déploient préférentiellement une stratégie orientée vers les objets, les sujets déficients visuels déploient pour leur part massivement une stratégie

orientée vers les sujets[53]. La poursuite de cette expérience devra conduire à la caractérisation plus précise de ces stratégies, ainsi qu'à une meilleure compréhension de l'impact du statut visuel.

Parallèlement à cette question, un deuxième axe majeur concerne l'émotion et la communication émotionnelle. En effet, nous avons souligné que l'échec des dispositifs de suppléance perceptive pouvait être attribué à l'absence de *qualia* dans les percepts (Bach-Y-Rita, 2002), qui ne pouvaient émerger que dans une histoire interactive partagée à l'intérieur d'une communauté d'usage (Lenay, 2006b). Dans cette optique, nous avons rapporté le cas des émoticônes tactiles (Maillet, 2010), qui ne pouvaient véhiculer du sens en dehors des utilisateurs qui les avaient produites. Dans notre travail, nous avons défini les conditions techniques nécessaires à l'interaction d'utilisateurs distants dans des espaces numériques partagés. Nous avons également soulevé les questions théoriques qui sont posées par de tels espaces, en mettant notamment l'accent sur la dynamique des interactions en tant que processus constitutif des connaissances sociales. La mise à disposition de ce dispositif à des utilisateurs devra maintenant permettre de constituer une communauté d'usage, ce qui donnera l'opportunité d'une approche minimaliste de l'émergence de valeurs émotionnelles dans le processus de l'interaction.

[53] Ce travail a donné lieu à une communication affichée :
Vallée, A., Deschamps, L. & Rovira, K., (2012), Dispositif d'aide aux jeunes aveugles : Interagir au sein d'un Internet tactile, *54e Congrès National de la Société Française de Psychologie, Dynamique des capacités humaines : du laboratoire au terrain*, 3-5 septembre 2012, Montpellier (Poster)

BIBLIOGRAPHIE

Adamson, L. B. & Bakeman, R., (1985), Affect and attention: Infants observed with mothers and peers. *Child Development*, 56, p. 582-593

Adamson, L. B., Bakeman, R. & Smith, C. B., (1990), Gestures, words, and early object sharing, in V. Volterra & C. Erting (Eds.), *From Gesture to Language in Hearing and Deaf Children*, New York: Springer-Verlag, p. 31-41

Adelson, E. & Fraiberg, S., (1974), Gross motor development in infants blind from birth, *Child Development*, 45, p. 114–116

Akhtar, N. & Gernsbacher, M. A., (2008), On privileging the role of gaze in infant social cognition, *Child Development Perspectives*, 2(2), p. 59-65

Ali Ammar, A., (2005), *Analyse des explorations haptiques de formes pour la conception d'un dispositif de suppléance perceptive dédié aux personnes aveugles*, Doctorat de Sciences et Technologies, Université de Technologie de Compiègne

Ali Ammar, A., Gapenne, O., Blomme, E. & Rovira, K., (2005), Analyse de l'exploration tactile sur support traditionnel chez la personne aveugle et conception de l'interface de lecture Tactos, *Revue d'intelligence artificielle*, 19 (1-2), p.339-354

Ali Ammar, A., Gapenne, O., Lenay, C. & Stewart, J., (2002), Effect of bimodality on the perception of 2D forms by means of a specific assistive technology for blind persons, *Proceedings of the Conference on Assistive Technology for Vision and Hearing Impairment (CVHI'2002*), 6-9 August 2002, Grenade, Espagne, p. 45-52

Anstis, S. M., Mayhew, J. W. & Morley, T., (1969), The perception of where a face or television 'portrait' is looking, *The American Journal of Psychology*, 82, p. 474–489

Arendt, H., (1983), *La Condition de l'Homme Moderne,* Paris : Calmann-Lévy

Argyle, M. & Cook, M., (1976), *Gaze and Mutual Gaze*, Cambridge, England: Cambridge University Press

Astington, J. W. & Gopnik, A., (1991), Theoretical explanations of children's understanding of the mind, *British Journal of Developmental Psychology*, 9, p. 7-31

Auvray, M., (2004), *Immersion et perception spatiale. L'exemple des dispositifs de substitution sensorielle*, Doctorat de Psychologie Cognitive, EHESS, Paris

Auvray, M., Lenay, C., O'Regan, J. K. & Lefèvre, J., (2005), Suppléance perceptive, immersion et informations proprioceptives, *Arobase*, 1, p. 94-114

Auvray, M., Lenay, C. & Stewart, J., (2009), Perceptual interactions in a minimalist virtual environment, *New Ideas in Psychology*, 27, p. 32-47

Bach-Y-Rita, P., (1972), *Brain Mechanisms in Sensory Substitution*, New York: Academic Press

Bach-y-Rita, P., (1982), Sensory substitution in rehabilitation, in L. Illis, M. Sedgwick et H. Granville (Eds.), *Rehabilitation of the Neurological Patient*, Oxford: Blackwell Scientific Publications, p.361-383

Bach-y-Rita, P., (2002), Sensory substitution and qualia, in A. Noe & E. Thompson (Eds.), *Vision and Mind*, Cambridge, Massachussets, London, England: MIT Press, p. 497-514

Bach-y-Rita, P., Collins, C.C., Saunders, F.A., White, B., & Scadden, L., (1969), Vision substitution by tactile image projection, *Nature,* 221(184), p. 963-964

Bach-y-Rita, P. & Kercel, S., (2003), Sensory substitution and the human-machine interface, *Trends in Cognitive Sciences,* 7(12), p. 541-546

Bahrick, L. E. & Watson, J. S., (1985), Detection of intermodal proprioceptive-visual contingency basis of self-perception in infancy, *Developmental Psychology,* 21, p. 963-973

Baird, S. M., Mayfield, P. & Baker, P., (1997), Mother's interpretations of the behavior of their infant with visual and other impairments during interactions, *Journal of Visual Impairment and Blindness,* 91, p. 467-483

Bakeman, R. & Adamson, L. B., (1984), Coordinating attention to people and objects in mother-infant and peer-infant interactions, *Child Development,* 55(4), p. 1278-1289

Baldwin, D. A., (1993a), Early referential understanding: Infants' ability to recognize referential acts for what they are, *Developmental Psychology*, 29, p. 832–843

Baldwin, D. A., (1993b), Infants' ability to consult the speaker for clues to word reference, *Journal of Child Language*, 20, p.395–418

Baldwin, D. A., (1995), Understanding the link between joint attention and language, in C. Moore & P. J. Dunham (Eds.), *Joint Attention: Its Origins and Role in Development,* Hillsdale, N.J: Lawrence Erlbaum Associates, p.131-156

Baldwin, D. A. & Moses, L. J., (1997), The ontogeny of social information gathering, *Child Development*, 67(5), p. 1915–1939

Baron-Cohen, S., (1995), *Mindblindness: an Essay on Autism and Theory of Mind*, MIT Press, Bradford Books

Baron-Cohen, S., Leslie, A.M. & Frith, U., (1985), Does the autistic child have a "theory of mind"?, *Cognition,* 21, p. 37–46

Bates, E., Camaioni, L. & Volterra, V., (1975), The acquisition of performatives prior to speech, *Merrill-Palmer Quarterly – Journal of Developmental Psychology*, 21, p. 205-226

Bateson, M. C., (1979), The epigenesis of conversational interaction: A personal account of research and development, in M. Bullowa (Ed.), *Before Speech: The Beginning of Interpersonal Communication*, Cambridge: Cambridge University Press, p. 63-77

Bationo-Tillon, A., (2007), Pratiques des activités narratives instrumentées : une analyse diachronique et structuro-fonctionnelle en amont de la conception (Résumé du Doctorat d'Ergonomie), *@ctivités*, 4(1), p. 238-242

Becchio, C., Sartori, L., Bulgheroni, M. & Castiello, U., (2008a), The case of Dr Jeckyll and Mr Hyde: A kinematic study on social intention, *Consciousness and Cognition*, 17, p. 557-564

Becchio, C., Sartori, L., Bulgheroni, M. & Castiello, U., (2008b), Both your intention and mine are reflected in the kinematics of my reach grasp movement, *Cognition*, 106, p. 894-912

Becchio, C., Sartori, L. & Castiello, U., (2010), Toward you: The social side of actions, *Current Directions in Psychological Science*, 19, p. 183-188

Beebe, B., & Gerstman, L., (1984), A method of defining 'packages' of maternal stimulation and their functional significance for the infant with mother and stranger, *International Journal of Behavioral Development*, 7(4), p. 423-440

Bigelow, A. E., (1992), Locomotion and search behavior in blind infants, *Infant Behavior and Development*, 15, p. 179-189

Bigelow, A. E., (1995), The effect of blindness on the early development of the self, in P. Rochat (Ed.), *The self in Infancy: Theory and Research*, Amsterdam: Elsevier Science, p. 327-347

Bigelow, A. E., (1998), Infants' sensitivity to familiar imperfect contingencies in social interaction, *Infant Behavior & Development*, 21, p. 149-161

Bigelow, A. E., (2001), Discovering self trough other: Infants' preference for social contingency, *Bulletin of the Menninger Clinic*, 65, p. 335-346

Bigelow, A. E., (2003), The development of joint attention in blind infants, *Development and Psychopathology*, 15, p. 259-275

Bigelow, A. E. & Birch, S. A. J., (1999), The effects of contingency in previous interactions on infants' preference for social partners, *Infant Behavior & Development*, 22, p. 367-382

Bigelow, A. E., MacLean, B. K. & MacDonald, D., (1996), Infants' response to live and replay interactions with self and mother, *Merryl-Palmer Quarterly*, 42(2), p. 596-611

Blass, E. M., Lumeng, J. & Patil, N., (2007), Influence of mutual gaze on human infant affect, in R. Flom, K. Lee, & D. Muir (Eds.), *Gaze-Following: Its Development and Significance*, Mahwah, NJ: Erlbaum, p. 113-141

Brazelton, T. & Cramer, B., (1991), *Les Premiers Liens : L'Attachement Parents-Bébé Vu par un Pédiatre et par un Psychiatre*, Paris, France : Stock

Brazelton, T. B., Tronick, E., Adamson, L., Als H. & Wise, S., (1975), Early mother-infant reciprocity, in M. Hofer (Ed.), *Parent-Infant Interaction*, Amsterdam: Elsevier, p. 137-154

Bretherton, I., (1991), Intentional communication and the development of an understanding of mind, in D. Frye and C. Moore (Eds.), *Children's Theories of Mind*, Hillsdale, New Jersey: Erlbaum, p. 49-75

Brooks, R. A., (1991), Intelligence without representation, *Artificial Intelligence*, 47(1-3), p. 139-159

Brown, R., Hobson, R. P., Lee, A. & Stevenson, J., (1997), Are there "autistic-like" features in congenitally blind children? *Journal of Child Psychology and Psychiatry*, 38, p. 693-703

Bruner, J., (1983), *Le Développement de l'Enfant : Savoir Faire, Savoir Dire*, coll. Psychologie d'aujourd'hui, Presses Universitaires de France

Bruner, J., (1995), From joint attention to the meeting of minds: An introduction, in C. Moore & P. J. Dunham (Eds.), *Joint Attention: Its Origin and Role in Development*, Hillsdale, N. J.: Lawrence Erlbaum, p. 189-203

Buisine, S. & Roussel, B., (2008), Analyse de l'usage, in B. Yannou, H. Christofol, D. Jolly & N. Troussier (Eds.), *La Conception Industrielle de Produits, vol. 3: Ingénierie de l'Evaluation et de la Décision*, Paris : Lavoisier Hermès Science, Chapitre 2, p. 43-60

Bullinger, A. & Mellier, D., (1988), Influence de la cécité congénitale sur les conduites sensorimotrices chez l'enfant, *Cahiers de Psychologie Cognitive*, 2, p. 191-203

Bullowa, M. (Ed.) (1979), *Before Speech: The Beginning of Interpersonal Communication*, Cambridge: Cambridge University Press

Butterworth, G., (1991), The ontogeny and phylogeny of joint visual attention, in A. Whiten (Ed.), *Natural Theories of Mind: Evolution, Development, and Simulation of Everyday Mindreading*, Oxford England: Blackwell, p. 223-232

Butterworth, G. & Cochran, E., (1980), Towards a mechanism of joint visual attention in human infancy, *International Journal of Behavioral Development*, 3, p. 253-272

Butterworth, G. & Jarrett, N., (1991), What minds have in common is space: spatial mechanisms serving joint visual attention in infancy, *British Journal of Developmental Psychology*, 9, p. 55-72

Burlingham, D., (1967), Developmental considerations in the occupations of the blind, *Psychoanalytic Study of the Child*, 30, p. 3-14

Campos, J. & Sternberg, C., (1981), Perception, appraisal, and emotion: The onset of social referencing, in M. Lamb & L. Sherrod (Eds.), *Infant Social Cognition: Empirical and Theoretical Considerations*, Hillsdale, NJ: Lawrence Erlbaum Associates, p. 273-314

Carpendale, J. I. M., & Lewis, C., (2004), Constructing an understanding of mind: The development of children's social understanding within social interaction, *Behavioral and Brain Sciences*, 27, p. 79-151

Carpenter, M., Nagell, K. & Tomasello, M., (1998), Social cognition, joint attention, and communicative competence from 9 to 15 months of age, *Monographs of the Society for Research in Child Development*, 63(4), [Serial No. 255]

Carpenter, M., Pennington, B. E. & Rogers, S. J., (2002), Interrelations among social-cognitive skills in young children with autism, *Journal of Autism & Developmental Disorders*, 32, p. 91-106

Chapman, M., (1991), The epistemic triangle: Operative and communicative components of cognitive development, in M. Chandler & M. Chapman (Eds.), *Criteria for Competence: Controversies in the Conceptualization and Assessment of Children's Abilities*, Hillsdale: Erlbaum, p. 209-228

Chess, S., (1971), Autism in children with congenital rubella, *Journal of Autism and Childhood Schizophrenia*, 1, p. 33-47

Chisholm, R., (1969), Some puzzles about agency, in K. Lambert (Ed.), *The Logical Way of Doing Things*, New Haven, CT: Yale University Press, p. 199-217

Collins, C. C. & Bach-Y-Rita, P., (1973), Transmission of pictorial information through the skin, *Advances in Biological Medecine and Physiology*, 14, p. 285-315

Condon, W. S., (1986), Communication: Rhythm and Structure, in J. R. Evans and M. Clynes (Eds.), *Rhythm in Psychological, Linguistic and Musical Processes*, Springfield, Charles C. Thomas, p. 55-77

Conein, B., (1998), Les sens sociaux : coordination de l'attention et interaction sociale, *Intellectica*, 1-2, p.181-202

Corkum, V. & Moore, C., (1995), Development of joint visual attention in infants, in C. Moore & P.J. Dunham (Eds.), *Joint Attention: Its Origins and Role in Development*, Hillsdale, NJ: Erlbaum, p. 61-83

Crown, C. L., Feldstein, S., Jasnow, M. D., Beebe, B., & Jaffe, J., (2002), The cross-modal coordination of interpersonal timing: Six-week-olds infants' gaze with adults' vocal behavior, *Journal of Psycholinguistic Research*, 31(1), p. 1-23.

Csibra, G., & Gergely, G., (1998), The teleological origins of mentalistic action explanations: A developmental hypothesis, *Developmental Science*, 1, p. 255-259

Csibra, G. & Gergely, G., (2006), Social learning and social cognition: The case for pedagogy, in Y. Munakata & M. H. Johnson (Eds.), *Processes of Change in Brain and Cognitive Development. Attention and Performance XXI*, Oxford, England: Oxford University Press, p. 249-274

Davis, M., (1982), *Interaction Rhythms: Periodicity in Communicative Behavior*, New York, Human Sciences Press

Decety, J., Jackson, P. L., Sommerville, J. A., Chaminade, T., & Meltzoff, A. N., (2004), The neural bases of cooperation and competition: an fMRI study, *NeuroImage*, 23, p. 744-751

Dennett, D., (1987), *The Intentional Stance*, Cambridge, MA: MIT Press

D'Entremont, B., Hains, S. M. J. & Muir, D. W., (1997), A demonstration of gaze following in 3- to 6-month-olds, *Infant Behavior and Development*, 20, p. 569-572

Deschamps, L., Rovira, K. & Lenay, C., (2008), The appropriation of prosthetic devices for reading computer screens and network interactions via tactile perception, *Sciences et Technologies pour le Handicap, n° spécial : Technologie Tactile et ses Applications pour les Handicaps,* 2(2), p. 159-173

Deschamps, L., Rovira, K., Lenay, C. & Gapenne, O., (2009), The conception of a tactile Internet : from analysing the use of a perceptual supplementation device (TACTOS) to the elaboration of the INTERTACT project, *CVHI-09, Conference and workshop on assistive technology for people with Visual and Hearing Impairments*, 20-23 april, Wroclaw, Poland (CD-Rom)

Desrochers, S., (2003), Les balbutiements de la causalité chez le nourrisson : Piaget « accommodé », *Archives de Psychologie*, 70, p. 97-119

De Bruin, L. C. & De Haan, S. E., (2012), Enactivism and social cognition: In search for the whole story, *Journal of Cognitive Semiotics*, IV (1), p. 225-250

De Jaegher, H., (2006), *Social interaction rhythm and participatory sense-making. An embodied, interactional approach to social understanding, with implications for autism*, Unpublished DPhil Thesis, University of Sussex

De Jaegher, H., (2009), Social understanding through direct perception? Yes, by interacting, *Consciousness and Cognition*, 18, p. 535-542

De Jaegher, H., & Di Paolo, E., (2007), Participatory sense-making: An enactive approach to social cognition, *Phenomenology and the Cognitive Sciences*, 6(4), p. 485-507

De Jaegher, H., Di Paolo, E. & Gallagher, S., (2010), Can social interaction constitute social cognition?, *Trends in Cognitive Sciences*, 14(10), p. 441-447

Di Paolo, E., (2009), Extended life, *Topoï*, 28, p. 9-21

Di Paolo, E., Rohde, M. & Iizuka, H., (2008), Sensitivity to social contingency or stability of interaction? Modelling the dynamics of perceptual crossing, *New Ideas in Psychology* 26, p. 278-294

D'Odorico, L. & Levorato, M., (1990), Social and cognitive determinants of mutual gaze between mother and infant, in V. Volterra & C. Erting (Eds.), *From Gesture to Language in Hearing and Deaf Children*, Berlin: Springer-Verlag, p. 1-17

Dunham, P, & Dunham, F., (1990), Effects of mother-infant social interactions on infants' subsequent contingency task performance, *Child Development*, 61, p. 785-793

Eilan, N., Hoerl, C., McCormack, T. & Roessler, J. (Eds.), (2005), *Joint Attention: Communication and Other Minds*, Oxford: Oxford University Press

Farroni, T., Csibra, G., Simion, F. & Johnson, M. H., (2002), Eye contact detection in humans from birth, *Proceedings of the National Academy of Sciences*, 99, p. 9602-9605

Ferrell, K. A., (1980), Can infants use the SonicGuide? Two years experience, *Journal of Visual Impairment and Blindness*, 74, p. 209-220

Fiebich, A. & Gallagher, S., (2012), Joint attention in joint action, *Philosophical Psychology*, disponible en ligne via http://dx.doi.org/10.1080/09515089.2012.690176

Field, T. M., (1979), Differential behavioral and cardiac responses of 3-month-old infants to a mirror and peer, *Infant Behavior and Development*, 2, p. 179-184

Field, T. M., (1981), Infant gaze aversion and heart rate during face-to-face interactions, *Infant Behavior & Development*, 4, p.307-315

Fodor, J., (1992), Discussion: A theory of the child's theory of mind, *Cognition*, 44, p. 283-296

Fogel, A., (1993), *Developing Through Relationships: Origins of Communication, Self and Culture*, Chicago: Chicago University Press

Fogel, A. & Garvey, A., (2007), Alive communication, *Infant Behavior and Development*, 30(2), p. 251–257

Fogel, A. & DeKoeyer-Laros, I., (2007), The developmental transition to secondary intersubjectivity in the second half year: A microgenetic case study, *Journal of Developmental Psychology*, 2(2), p. 63-90

Fraiberg, S., (1977), *Insights from the Blind*, London: Souvenir Press

Fraiberg, S., (1979), Blind infants and their mothers: An examination of the sign system, in M. Bullowa (Ed.), *Before Speech: The Beginning of Interpersonal Communication*, Cambridge: Cambridge University Press, p. 149-169

Frith, C. D., (2008), Social cognition, *Philosophical Transactions of the Royal Society B: Biological Sciences,* 363, p. 2033-2039

Frith, U. & Frith, C. D., (2003), Development and neurophysiology of mentalizing, *Philosophical Transactions of the Royal Society B: Biological Sciences,* 358, p. 459-473

Froese, T. & Di Paolo, E., (2011), The enactive approach: Theoretical sketches from cell to society, *Pragmatics & Cognition,* 19(1), p. 1-36

Fuchs, T. & De Jaegher, H., (2009), Enactive intersubjectivity: Participatory sense-making and mutual incorporation, *Phenomenology and the Cognitive Sciences,* 8, p. 465-486

Gallagher, S., (2004), Understanding interpersonal problems in autism: Interaction theory as an alternative to theory of mind, *Philosophy. Psychiatry and Psychology,* 11(3), p. 199-217

Gallagher, S., (2007), Logical and phenomenological arguments against simulation theory, in D. Hutto & M. Ratcliffe (Eds.), *Folk Psychology Re-assessed,* Springer Publishers, p. 63-78

Gallagher, S., (2008), Direct perception in the intersubjective context, *Consciousness and Cognition,* 17, p. 535-543

Gallagher, S., (2011), Interactive coordination in joint attention, in A. Seeman (Ed.), *Joint Attention: New Developments in Psychology, Philosophy of Mind, and Social Neuroscience,* Cambridge, MA: MIT Press

Gallese, V. & Goldman, A., (1998), Mirror neurons and the simulation theory of mind-reading, *Trends in Cognitive Sciences,* 2(12), p. 493-501

Gapenne, O., Lenay, C., Stewart, J., Bériot, H. & Meidine, D., (2001), Prosthetic device and 2D form perception: The role of increasing degrees of parallelism, *Proceedings of the Conference on assistive technology for Vision and Hearing Impairment (CVHI'2001),* Castelvecchio Pascoli, Italie

Gapenne, O., Rovira, K., Ali Ammar, A. & Lenay, C., (2003), TACTOS: A special computer interface for the reading and writing of 2D forms in blind people, in . C. Stephanidis (Ed), *Universal Access in HCI, Inclusive Design in the Information Society,* Lawrence Erlbaum Associates, London, p.1270-1274

Garfield, J. L., Peterson C.C. & Perry, T., (2001), Social cognition, language acquisition and the development of the theory of mind, *Mind & Langage,* 16(5), p.494-541

Gense, M.H., & Gense, D.J., (1994), Identifying autism in children with blindness and visual impairments, *REview,* 16, p. 55-62

Gergely, G., (2001), The obscure object of desire: 'Nearly, but clearly not, like me': Contingency preference in normal children versus children with autism, *Bulletin of the Menninger Clinic, 65*, p. 411-426

Gergely, G., & Csibra, G., (2003), Teleological reasoning in infancy: the naive theory of rational action, *Trends in Cognitive Sciences,* 7, p. 287-292

Gergely, G., & Watson, J.S., (1996), The social biofeedback theory of parental affect-mirroring: The development of emotional self-awarness and self-control in infancy, *International Journal of Psycho-Analysis, 77*, p. 1181-1212

Gergely, G. & Watson, J.D., (1999), Early socio-emotional development: Contingency perception and the social bio-feedback model, in P. Rochat (Ed.) *Early Social Cognition*, Erlbaum, p. 101-136

Gibson, J. J., (1966), *The Senses Considered as Perceptual Systems*, Boston: Houghton Mifflin

Gibson, J. J., (1979), *The Ecological Approach to Visual Perception*, Hillsdale: Lawrence Erlbaum Associates

Goffman, E., (1959), *The Presentation of Self in Everyday Life,* New York: Doubleday Anchor

Goldman, A. I., (1995), Interpretation Psychologized, in M. Davis & T. Stone (Eds.), *Folk Psychology: The Theory of Mind Debate,* Oxford: Blackwell Publishers, pp 74–99

Gomez J. (1991), Visual behavior as a window for reading the mind of others in primates, in Whiten, A. (ed.), *Natural Theories of Mind*, Blackwell, p. 195-207

Gopnik, A. & Wellman, H., (1992), Why the child's theory of mind really is a theory, *Mind and Language,* 7(1/2) p. 145-172

Gopnik, A. & Meltzoff, A. N., (1997), *Words, Thoughts and Theories*, Cambridge, MA, MIT Press

Gordon, R. M., (1986), Folk Psychology as Simulation, *Mind and Language,* 1, p. 158–171 ; reprinted in M. Davies and T. Stone (Eds**.)**, *Folk Psychology: The Theory of Mind Debate*, Oxford: Blackwell Publishers, 1995

Gordon, R. M., (1995), Simulation without introspection or inference from me to you, in M. Davies & T. Stone (Eds.), *Mental Simulation*, Oxford: Blackwell, p. 53–67

Gregory, R. L., (1990), Recovery From Blindness, in *Eye and brain: the psychology of seeing,* Oxford: University Press, Chapitre 11, p. 191-200

Hains, S. M. J. & Muir, M. W., (1996), Infant sensitivity to adult eye direction, *Child Development,* 67, p. 1940–1951

Hanneton S., Gapenne O., Genouëlle C., Lenay C. & Marque C., (1999), Dynamics of Shape Recognition Through a Minimal Visuo-Tactile Sensory Substitution Interface, *Third International Conference On Cognitive and Neural Systems*, Mai 1999, Boston, p. 26-29

Hatwell, Y., (2003), *Psychologie Cognitive de la Cécité Précoce,* Dunod, Paris

Heller, M. A., (2000), *Touch, Representation and Blindness*, UK: Oxford University Press

Heller, M. A., (2002), Tactile picture perception in sighted and blind people, *Behavioural Brain Research*, 135, p. 65-68

Hill, E. & Ponder, P., (1976), *Orientation and Mobility Techniques: A Guide for the Practitioner*, American Foundation for the Blind

Hobson, R. P., (1993), Through feeling and sight to self symbol, in U. Neisser (Ed.), *The Perceived Self: Ecological and Interpersonal Sources of Self-Knowledge*, Cambridge: Cambridge University Press, p. 254–279

Hobson, R. P., (2002), *The Cradle of Thought,* London: Macmillan

Hood, B. M., Willen, J. D. & Driver, J., (1998), Adult's eye trigger shifts of visual attention in human infants, *Psychological Science*, 9, p. 131–134

Hutto, D.D, (2004), The limits of spectatorial folk psychology, *Mind and Language*, 19, p. 548–573

Ingsholt, A., (2000), Being together, interaction and dialogues in interactions between blind infants and their parents, *Proceedings of the 5th European Conference of ICEVI (International Council for Education of People with Visual Impairment), Visions and strategies for the new century*, 9-13 July 2000, Cracow, Poland (http://www.icevi-europe.org/cracow2000/proceedings/chap01.html#8)

Issartel, J., Marin, L & Cadopi, M., (2007), Unintended interpersonal coordination : "can we march to the beat of our own drum?", *Neurosciences Letters,* 411, p. 174-179

Jouët, J., (1993), Pratiques de communication, figures de la médiation, *Réseaux*, 60, p. 99-120

Kay, L., (1974), A sonar aid to enhance spatial perception of the blind. Engineering design and evaluation, *The Radio and Electronic Engineer,* 44, p. 40-62

Kekelis, L. S. & Andersen, E., (1984), Family communication styles and language development, *Journal of Visual Impairment and Blindness*, 78, p. 54-65

Khatchatourov, A. & Auvray, M., (2005), L'outil modifie-t-il la perception ou la rend-il possible ?, *Arobase,* 1, p. 65-70

Kidwell, M. & Zimmerman, D., (2007), Joint attention as action, *Journal of Pragmatics,* 39(3), p. 592-611

Knoblich, G. & Sebanz, N., (2008), Evolving intentions for social interaction: From entrainment to joint action, *Philosophical Transactions of the Royal Society B*, 363, p. 2021-2031

Krueger, J., (2011), Extended cognition and the space of interaction, *Consciousness and Cognition*, 20, p. 643-657

Kummer, H., (1967), Tripartite relations in hamadryas baboons, in S. Altman (Ed.), *Social Communication among Primates*, Chicago, University of Chicago Press, nouvelle édition in Byrne & Whiten (1988), p. 113-121

Lempers. J. D., (1979), Young children's production and comprehension of nonverbal deictic behaviours, *Journal of Genetic Psychology*, 135, p. 93-l 02

Lenay, C., (1997), Mouvement et perception : médiation technique et constitution de la spatialisation, *Actes de la Sixième école d'été de l'Association pour la Recherche Cognitive (ARCo)*, « *Le mouvement. Des boucles sensori-motrices aux représentations cognitives et langagières* », Juillet 1997, p. 69-80

Lenay, C., (2005), Constitution de l'espace et immersion, *Arobase*, 1, p. 85-93

Lenay, C., (2006a), Enaction, externalisme et suppléance perceptive, *Intellectica*, 2006/1, 43, p. 27-52

Lenay, C., (2006b), Réticences et conditions d'adoption des prothèses perceptives, *Actes de la conférence Handicap'2006, « Nouvelles technologies au service de l'Homme »*, 7-9 juin 2006, Paris

Lenay, C., (2006c), La méthode minimaliste : phénoménologie et science objective, *Intellectica*, 2006/1, 43, p. 125-134

Lenay, C., (2010), C'est très touchant. La valeur émotionnelle du contact, *Intellectica*, 53/54, p. 359-397

Lenay, C., Canu, S. & Villon, P., (1997), Technology and perception: the contribution of sensory substitution systems, *Proceedings of the IEEE Second International Conference on Cognitive Technology*, 25-28 Août 1997, Aizu, Japan, Los Alamitos: IEEE, p. 44-53

Lenay C., Gapenne O., Hanneton S., Marque, C. & Genouëlle, C., (2000), La substitution sensorielle. Limites et perspectives, in Y. Hatwell, A. Streri & E. Gentaz (Eds.), *Toucher pour Connaître. Psychologie Cognitive de la Perception Tactile Manuelle*, Paris, PUF, 2000, p. 287-306

Lenay, C. & Stewart, J., (2012), Minimalist approach to perceptual interactions, *Frontiers in Human Neuroscience*, 6, p. 1-18

Leslie, A. M., (1987), Pretense and representation: The origins of "theory of mind", *Psychological Review*, 94, p. 412-426

Leslie, A., (1988), Some implications of pretence for mechanisms underlying the child's theory of mind, in J.W. Astington, P.L. Harris, & D.R.Olson (Eds.), *Developing Theories of Mind*, New York: Cambridge University Press, p. 19–46

Liszkowski, U., Carpenter, M., Striano, T. & Tomasello, M., (2006), Twelve- and 18-month-olds point to provide information for others, *Journal of Cognition and Development*, 7, p. 173 – 187

Macmurray, J., (1991), *Persons in Relation*, London: Faber & Faber (First edition: 1961)

Maillet, B., (2010), *Design et appropriation d'un dispositif d'interactions haptiques pour l'expression et la perception d'attitudes communicationnelles*, Doctorat de Sciences et Technologies, Université de Technologie de Compiègne

Marsh, K. L., Richardson, M. J., Baron, R. B. & Schmidt, R.C., (2006), Contrasting approaches to perceiving and acting with others, *Ecological Psychology*, 18(1), p. 1–38

Marsh, K. L., Richardson, M. J. & Schmidt, R. C., (2009), Social connection through joint action and interpersonal coordination, *Topics in Cognitive Science*, 1, p. 320-339

Maturana, H. R. & Varela, F. J., (1994), *L'Arbre de la Connaissance : Racines Biologiques de la Compréhension Humaine*, Paris : Addison-Wesley

Maurer, D., (1985), Infants' perception of facedness, in T. M. Field & N. A. Fox (Eds.), *Social Perception in Infants*, Norwood, NJ: Ablex, p. 73-100

Mauris, G., Benoit, E. & Foulloy, L., (1998), Fuzzy linguistic methods for the aggregation of complementary sensor information, in Bouchon-Mounier, B. (eds.), *Aggregation and Fusion of Imperfect Information*, Germany: Physica Verlag, p. 214-230

McGann, M. & De Jaegher, H., (2009), Self–other contingencies: enacting social perception, *Phenomenology and the Cognitive Sciences*, 8, p. 417–437

Mead, G. H., (1934), *Mind, Self and Society*, Chicago: University of Chicago Press

Meijer, P. B. L., (1992), An experimental system for auditory image representations, *IEEE Transactions on Biomedical Engineering*, 39, p. 112–121

Mellier, D., (2005), Pour une approche développementale des suppléances chez l'enfant, *Arobase*, 1, p. 127-131

Meltzoff, A. N., (1999), Origins of theory of mind, cognition, and communication, *Journal of Communication Disorders*, 32, p. 251-269

Meltzoff, A. N., (2005), Imitation and other minds: The 'like me' hypothesis, in S. Hurley & N. Chater (Eds.), *Perspectives on Imitation: From Neuroscience to Social Science*, Vol 2, Cambridge, MA, US: MIT Press, p. 55-78

Meltzoff, A. N. & Brooks, R., (2007), Eyes wide shut: The importance of eyes in infant gaze following and understanding other minds, in R. Flom, K. Lee & D. Muir (Eds.), *Gaze-Following: Its Development and Significance*, Mahwah, N.J: Erlbaum, p. 217–241

Merleau-Ponty, M., (1945), *Phénoménologie de la Perception*, Paris, Gallimard

Michael, J., & Overgaard, S., (2012), Interaction and social cognition: A comment on Auvray et al.'s perceptual crossing paradigm, *New Ideas in Psychology*, 30, p.296-299

Millar, S., (1988), Models of sensory deprivation: The nature/nurture dichotomy and spatial representation in the blind, *International Journal of Behavioral Development*, 11, p. 69-87

Millar, S., (1994), *Understanding and Representing Space: Theory and Evidence from Studies with Blind and Sighted Children*, Oxford: University Press

Moore, C., Angelopoulos, M. & Bennett, P., (1997), The role of movement in the development of joint visual attention, *Infant Behavior and Development*, 20(1), p. 83-92

Moore, C. & Corkum, V., (1994), Social understanding at the end of the first year of life, *Developmental Review* 14, p. 349–72

Moore, C. & Dunham, P. J. (Eds.) (1995), *Joint Attention: Its Origin and Role in Development*. Hillsdale, N. J.: Lawrence Erlbaum

Morales, M., Mundy, P., Delgado, C. E. F., Yale, M., Messinger, D., Neal, R. & Schwartz, H. K., (2000), Responding to joint attention across the 6- to 24-month age period and early language acquisition, *Journal of Applied Developmental Psychology*, 21, p. 283–298

Moran, G., Dumas, J. E. & Symons, D. K., (1992), Approaches to sequential analysis and the description of contingency in behavioral interaction, *Behavioral Assessment*, 14, p. 65-92

Mundy, P., (1995), Joint attention, social-emotional approach in children with autism, *Development and Psychopathology*, 7, p. 63-82

Mundy, P., Block, J., Delgado, C., Pomares, Y., Van Hecke, A. V. & Parlad, M. V., (2007), Individual differences and the development of joint attention in infancy, *Child Development*, 78(3), p. 938-954

Mundy, P. & Acra, F., (2006), Joint attention, social engagement and the development of social competence, in P. Marshall & N. Fox (Eds.), *The Development of Social Engagement Neurobiological Perspectives*, New York: Oxford University Press, p. 81-117

Mundy, P. & Crowson, M., (1997), Joint attention and early social communication: Implications for research on intervention with autism, *Journal of Autism and Developmental Disorders*, 27(6), p. 653-673

Mundy, P. & Newell, L., (2007), Attention, joint attention and social cognition, *Current Directions in Psychological Science*, 16, p. 269–274

Murray, D. S., Creaghead, N. A., Manning-Courtney, P., Shear, P. K., Bean, J. & Prendeville, J., (2008), The relationship between joint attention and language in children with autism spectrum disorders, *Focus on Autism and Other Developmental Disabilities*, 23(1), p. 5-14

Murray, L. & Trevarthen, C., (1985), Emotional regulation of interactions between 2-month-olds and their mothers, in T. M. Field and N. A. Fox (Eds.), *Social Perception in Infants*, Norwood, NJ, Ablex**,** p. 177-197

Murray, L. & Trevarthen, C., (1986), The infant's role in mother-infant communication, *Journal of Child Language*, 13, p. 15-29

Nadel, J., Carchon, I., Kervella, C., Marcelli, D., & Réserbat-Plantey, D., (1999), Expectancies for social contingency in 2-month-olds, *Developmental Science*, 2(2), p. 164–173

Neisser, U., (1988), Five kinds of self-knowledge, *Philosophical Psychology, 1(1),* p. 35-59

Neisser, U., (1991), Two perceptually given aspects of the self and their development, *Developmental Review*, 11, p. 197–209

O'Regan, J.K. & Noe, A., (2001), A sensorimotor account of vision and visual consciousness, *Behavioral and Brain Sciences*, 24, p. 939-1011

Pahl, G. & Beitz, W., (1996), *Engineering Design, A Systematic Approach*, Springer-Verlag GmbH

Paillard, J., (1971), Les déterminants moteurs de l'organisation de l'espace, *Cahiers de Psychologie*, 14, p. 261-316

Papousek, H. & Papousek, M., (1974), Mirror image and self-recognition in young human infants: I. A new method of experimental analysis, *Developmental Psychobiology*, 7, p. 149-157

Pérez-Pereira, M. & Conti-Ramsden, G., (1999), *Language Development and Social Interaction in Blind Children*, Hove, UK: Psychology Press

Perner, J., (1991), *Understanding the Representational Mind*, Cambridge, MA, The MIT Press

Pfeiffer, U.J., Timmermans, B., Bente, G., Vogeley, K. & Schilbach, L., (2011), A Non-Verbal Turing Test: Differentiating Mind from Machine in Gaze-Based Social Interaction, *PLoS ONE*, 6(11), e27591

Piaget, J., (1936), *La Naissance de l'Intelligence chez l'Enfant*, Delachaux et Niestlé

Piaget J. & Inhelder, B., (1966), *La psychologie de l'enfant*, Presses Universitaires de France, Collection « Que sais-je ? », Paris

Ploux, S., (1994), An example of perception and action coupling: handwriting recognition, in C. Faure, P. Keuss, G. Lorette, A. Vinter (eds.), *Advances in Handwriting and Drawing. A Multidisciplinary Approach*, Europia

Preisler, G. M., (1991), Early patterns of interaction between blind infants and their sighted mothers, *Child: Care, Health and Development*, 17(2), p. 65-90

Premack, D. & Premack, A., (1997), Infants attribute value +/− to the goal-directed actions of self-propelled objects, *Journal of Cognitive Neuroscience*, 9, p. 848–856

Premack, D. & Woodruff, G., (1978), Does the chimpanzee have a theory of mind?, *Behavioral and Brain Sciences*, 1, pp 515–526

Racine, T. P. & Carpendale, J. I. M., (2007), The role of shared practice in joint attention, *British Journal of Developmental Psychology*, 25, p. 3-25

Rattray, J. & Zeedyk, M. S., (2005), Early communication in dyads with visual impairment, *Infant and Child Development*, 14, p. 287-309

Reddy, V., (1996), Omitting the second person in social understanding, *Behavioral and Brain Sciences*, 19, p. 140-141

Reddy, V., (2003), On being the object of attention: Implications for self-other consciousness, *Trends in Cognitive Sciences*, 7(9), p. 397-402

Reddy, V., (2011), A gaze at grips with me, in A. Seeman (Ed.), *Joint Attention: New Developments in Psychology, Philosophy of Mind and Social Neuroscience*, Cambridge, MA: MIT Press, p. 137-157

Reddy, V. & Morris, P., (2004), Participants don't need theories: knowing minds in engagement, *Theory & Psychology*, 14(5), p. 647-665

Richardson, M. J., Marsh, K. L., Isenhower, R. W., Goodman, J. R. L. & Schmidt, R. C., (2007), Rocking together: Dynamics of unintentional and intentional interpersonal coordination, *Human Movement Science*, 26, p. 867-891

Robertson, S. S. & Johnson, S. L., (2009), Embodied infant attention, *Developmental Science*, 12, p. 297-304

Rochat, P, Querido, J. G. & Striano, T., (1999), Emerging sensitivity to the timing and structure of protoconversation in early infancy, *Developmental Psychology*, 35(4), p. 950-957

Rochat. P., Neisser, U. & Marian, V., (1998), Are young infants sensitive to interpersonal contingency?, *Infant Behavior and Development*, 21, p. 355-366

Rochat, P. & Striano, T., (1999), Social-cognitive development in the first year, in P. Rochat (Ed.), *Early Social Cognition: Understanding Others in the First Months of Life,* Hillsdale, NJ: Lawrence Erlbaum Associates, p. 3-34

Rogers, S. J. & Puchalski, C. B., (1984), Development of symbolic play in visually impaired young children, *Topics in Early Childhood Special Education*, 3(4), p. 57-63

Rovee, C. K. & Rovee, D. T., (1969), Conjugate reinforcement of infant exploratory behavior, *Journal of Experimental Child Psychology,* 8, p. 33-39

Rovira, K. & Gapenne, O., (2009), Tactile classification on traditional and computerized media in three adolescents who are blind, *Journal of Visual Impairment and Blindness,* 103(7), p. 430-435

Rovira, K., Gapenne, O. & Ali Ammar, A., (2010), Learning to recognize shapes with a sensory substitution system: a longitudinal study with 4 non-sighted adolescents, *Proceedings of the Ninth IEEE International Conference on Development and Learning (ICDL 2010),* Ann Arbor, MI, USA, August 18-21, p. 1-6

Rowland, C., (1984), Preverbal communication of blind infants and their mothers, *Journal of Visual Impairment and Blindness*, 78, p. 297-302

Sacks, S. & Silberman, R., (2000), Social skills, in A. J. Koenig and M. C. Holbrook (Eds.), *Foundations of Education (Volume II): Instructional Strategies for Teaching Children and Youths with Visual Impairments,* New York, NY: American Foundation for the Blind, p. 616–652

Sampaio, E., (1989), Is there a critical age for using the sonicguide with blind infants?, *Journal of Visual Impairment and Blindness*, 83, p. 105-108

Sampaio, E., & Dufier, J. L., (1988), Suppléance sensorielle électronique pour les jeunes enfants aveugles, *Journal Français Ophtalmologique,* 11(2), p. 161-167

Sarria, E., Gomes, J. C. & Tamarit, J., (1996), Joint attention and alternative language intervention in autism: Implications of theory for practice, in S. von Tetzchner & M. H. Jensen (Eds.), *Augmentative and Alternative Communication: European Perspectives,* London: Whurr, p. 49-64

Scaife, M. & Bruner, J., (1975), The capacity of joint attention in the infant, *Nature,* 245, p. 265-266

Schmuckler, M. A., (1996), Visual-proprioceptive intermodal perception in infancy, *Infant Behavior and Development*, 19, p. 221-232

Scholl, B., & Leslie, A., (1999), Modularity, development and 'theory of mind', *Mind & Language,* 14, p. 131-153

Searle, J. R., (1972), *Les Actes de Langage*. *Essai de Philosophie du Langage*, Paris, Hermann

Searle, J. R., (1980), Minds, Brains and Programs, *Behavioral and Brain Sciences*, 3(3), p. 417–457

Sebanz, N., Bekkering, H. & Knoblich, G., (2006), Joint action: Bodies and minds moving together, *Trends in Cognitive Sciences*, 10, p. 70–76

Sellars, W., (1956), Empiricism and the philosophy of mind, *Minnesota Studies in the Philosophy of Science*, 1, p. 253–329

Shannon, C., (1948), A mathematical theory of communication, *The Bell System Technical Journal*, 27, p. 623-656

Sigman, M. & Kasari, C., (1995), Joint attention across contexts in normal and autistic children, in C. Moore & P. Dunham (Eds.), *Joint Attention: Its Origins and Role in Development*, Hillsdale, N.J: Lawrence Erlbaum Associates, p. 189-201

Siqueland, E. R. & DeLucia, C. A., (1969), Visual reinforcement of non-nutritive sucking in human infants, *Science*, 165, p. 1144-l146

Slaughter, V. & McConnell, D., (2003), Emergence of joint attention: Relationships between gaze following, social referencing, imitation, and naming in infancy, *Journal of Genetic Psychology*, 164(1), p. 54-71

Sribunruangrit, N., (2001), *Influence des paramètres de la stimulation tactile sur l'exploration et la reconnaissance de formes géométriques*, Mémoire de D.E.A., Université de Technologie de Compiègne

Sribunruangrit, N., (2004), *Etude et développement de systèmes de suppléance perceptive tactile pour les personnes aveugles*, Doctorat de Sciences et Technologies, Université de Technologie de Compiègne

Sribunruangrit N., Marque C., Lenay C., Gapenne O. & Vanhoutte C., (2002), Braille Box: Analysis of the Parallelism Concept to Access Graphic Information for Blind People, *Proceedings of the Second Joint Meeting of the IEEE Engineering in Medicine and Biology Society and the Biomedical Engineering Society (EMBS-BMES)*, USA, Houston, Texas, 23-26 oct. 2002

Sribunruangrit, N., Marque, C., Lenay, C., Gapenne, O. & Vanhoutte, C., (2004), Speed-accuracy tradeoff during performance of a tracking task without visual feedback, *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 12, no.1/2004, p. 131-13

Stack, D. M., (2001), The salience of touch and physical contact during infancy: Unraveling some of the mysteries of the somesthetic sense, in G. Bremner & A. Fogel (Eds.), *Blackwell Handbook of Infant Development*, Malden, MA: Blackwell, p. 351–378

Stack, D. M. & Arnold, S. L., (1998), Changes in mothers' touch and hand gestures influence infant behavior during face-to-face interchanges, *Infant Behavior & Development*, 21, p.451–468

Stack, D. M. & Muir, D. W., (1992), Adult tactile stimulation during face-to-face interactions modulates five-month-olds' affect and attention, *Child Development*, 63, p. 1509–1525

Stern, D., (1977), *The First Relationship: Infant and Mother*. Cambridge: Harvard University Press

Stern, D., (1985), *The Interpersonal World of the Infant: A View from Psychoanalysis and Developmental Psychology*, New-York: Basic Books

Stewart, J. & Gapenne,, O., (2004), Reciprocal modelling of active perception of 2-D forms in a simple tactile-vision substitution system, *Minds and Machines*, 14, p. 309-330

Stewart, J., Gapenne, O. & Di Paolo, E., (2010), *Enaction. Towards a New Paradigm for Cognitive Science*, Cambridge: MIT Press

Stratton, G. M., (1896), Some preliminary experiments on vision without inversion of the retinal image, *Psychological Review*, 3(6), p. 611-617

Streri, A., (2000), Les coordinations intermodales chez le bébé, in Y. Hatwell & E. Gentaz (Eds.), *Toucher pour Connaître. Psychologie Cognitive de la Perception Tactile Manuelle*, Paris, PUF, p. 193-210

Striano, T. & Bertin, E., (2005), Social-cognitive skills between 5 and 10 months of age, *British Journal of Developmental Psychology*, 23, p. 559-568

Striano, T. & Rochat, P., (1999), Relation between dyadic and triadic social competence in 7- to 10-month-old infants, *British Journal of Developmental Psychology*, 17, p. 551–562

Striano, T. & Stahl, D., (2005), Sensitivity to triadic attention in early infancy, *Developmental Science*, 8, p. 333–343

Tao, L., (2009), *Étude de la discrimination de la densité de textures tactiles via un dispositif de suppléance perceptive (TACTOS)*, Rapport de Master ICI (Innovation, Connaissance, Interaction), Université de Technologie de Compiègne

Tellevik, J. M., (1992), Influence of spatial exploration patterns on cognitive mapping by blindfolded sighted persons, *Journal of Visual Impairment and Blindness*, 86, p. 221-224

Theureau, J. & Pinsky, L., (1984), Paradoxe de l'ergonomie de conception et logiciel informatique, *Revue des Conditions de Travail*, 9, p. 25-31

Tomasello, M., (1995), Joint attention as social cognition, in C. Moore & P. Dunham (Eds.), *Joint Attention: Its Origins and Role in Development*, Hillsdale, NJ: Erlbaum, p. 103–130

Tomasello, M., (1999), *The Cultural Origins of Human Cognition*, Cambridge, MA: Harvard University Press

Tomasello, M., Carpenter, M., Call, J., Behne, T. & Moll, H., (2005), Understanding and sharing intentions: The origins of cultural cognition, *Behavioral and Brain Sciences*, 28, p. 675-735

Thompson, E., (2007), *Mind in Life: Biology, Phenomenology, and the Sciences of Mind*, Harvard University Press

Tremblay, H. & Rovira, K., (2007), Joint visual attention and social triangular engagement at 3 and 6 months, *Infant Behavior and Development*, 30(2), p. 366-379

Trevarthen, C., (1977), Descriptive analyses of infant communicative behavior, in H. Schaffer (Ed.), *Studies in Mother–Infant Interaction*, London: Academic Press, p. 227–270

Trevarthen, C., (1979), Communication and cooperation in early infancy: a description of primary intersubjectivity, in M. Bullowa (ed.) *Before Speech: The Beginning of Interpersonal Communication*, Cambridge: Cambridge University Press, p. 321-347

Trevarthen, C., (1986), Form, significance and psychological potential of hand gestures of infants, in J.-L. Nespoulous, P. Perron, A.R. Lecours (Eds.), *The Biological Foundation of Gestures: Motor and Semiotic Aspects*, Hillsdale, N.J, Erlbaum, p. 149-202

Trevarthen, C., (1993), The self born in intersubjectivity, in U. Neisser (Ed.), *The Perceived Self: Ecological and Interpersonal Sources of Self-Knowledge*, Cambridge University Press, p. 121-173

Trevarthen, C., (2001), Intrinsic motives for companionship in understanding: Their origin, development, and significance for infant mental health, *Infant Mental Health Journal*, 22(1–2), p. 95–131

Trevarthen, C., (2008), The musical art of infant conversation: Narrating in the time of sympathetic experience, without rational interpretation, before words, *Musicae Scientiae, Special Issue*, p. 11-37

Trevarthen, C. & Hubley, P., (1978), Secondary intersubjectivity: confidence, confiding and acts of meaning in the first year, in A. Lock (ed.) *Action, Gesture and Symbol: The emergence of language*, London: Academic Press, p. 183-230

Tronick, E. Z., Als, H., & Adamson, L., (1979), Structure of early face-to-face communicative interactions, in M. Bullowa (Eds.), *Before Speech: The Beginning of Interpersonal Communication*, Cambridge, Cambridge University Press, p. 349-370

Turing, A. M., (1950), Computing machinery and intelligence, *Mind, New Series*, 59, n°236, p. 433-460

Urwin, C., (1978), The development of communication between blind infants and their parents, in A. Lock (Ed.), *Action, Gesture, and Symbol: The Emergence of Language*, London: Academic Press, p.79-108

Urwin, C., (1983), Dialogue and cognitive functioning in the early language development of three blind children: Normal and deficient, in A. E. Mills (Ed.), *Language Acquisition in the Blind Child*, London: Croom Helm, p. 142-161

Van Overwalle, F., (2009), Social cognition and the brain: A meta-analysis, *Human Brain Mapping*, 30, p. 829-858

Varela, F. J., Thompson, E. & Rosch, E., (1991), *The Embodied Mind: Cognitive Science and Human Experience*, 6th edition, Cambridge, MA: MIT Press

Vaughan Van Hecke, A. & Mundy, P., (2007), Neural systems and the development of gaze following and related joint attention skills, in R. Flom, K. Lee, & D. Muir (Eds.), *Gaze Following: Its Development and Significance*, Mahwah, N.J: Lawrence Erlbaum Associates, p.17-51

Vecera, S. P. & Johnson, M. H., (1995), Eye gaze detection and cortical processing of face: evidence from infants and adults, *Visual Cognition*, 2, p. 59–87

Ventrella, J, Seif El-Nasr, M., Aghabeigi, B. & Overington, R., (2010), Gestural Turing test: A motion-capture experiment for exploring believability in artificial nonverbal communication, *Proceedings of the AAMAS 2010 International Workshop on Interacting with ECAs as Virtual Characters*, May 2010, Toronto, Canada

Von Hippel, E., (2005), *Democratizing Innovation*, Cambridge, The MIT Press

Wimmer, H. & Perner, J., (1983), Beliefs about beliefs representation and constraining function of wrong beliefs in young children's understanding of deception, *Cognition*, 13, p. 103-128

Wittgenstein, L., (1980), *Remarks on the Philosophy of Psychology*, Volume II, edited by G. H. von Wright and H. Nyman, translated by C.G. Luckhart and M. A. E. Aue, Oxford: Blackwell

Woodward, A. L., (2003), Infants' developing understanding of the link between looker and object, *Developmental Science*, 6, p.297–311

Zebehazy, K. & Smith, T., (2011), An examination of characteristics related to the social skills of youths with visual impairments, *Journal of Visual Impairment & Blindness*, 105(2), p. 84-95

Ziat M., Gapenne O. Lenay C., Lecolinet E., Mouret G. & Stewart J., (2006), Espace de perception et seuils de confort pour un zoom haptique en 2D, *Revue d'Interaction Homme Machine*, 7, p. 29-54

Ziat M., Gapenne O., Stewart J. & Lenay C., (2007), Haptic recognition of shapes at different scales: A comparison of two methods of interaction, *Interacting with Computers,* 19, p. 121-132

ANNEXES

ANNEXE A : Note sur les analyses statistiques inférentielles 286

a. Tests préliminaires 286

b. Statistiques paramétriques 287

c. Statistiques non-paramétriques 288

d. Synthèse de la méthodologie utilisée pour les analyses statistiques..... 289

ANNEXE B : Note sur l'analyse des trajectoires perceptives..... 290

ANNEXE C : Exemples de trajectoires perceptives (Turing) 292

a. Exemple 1 : Interaction entre deux sujets humains (1) 292

b. Exemple 2 : Interaction entre deux sujets humains (2) 293

ANNEXE D : Exemples de trajectoires perceptives (Public/Privé) 294

a. Exemple 1 : Clic Mutuel en condition « Mono » (1)..... 294

b. Exemple 2 : Clic Mutuel en condition « Mono » (2)..... 295

c. Exemple 3 : Clic Mutuel en condition « Différenciée » (1)..... 296

d. Exemple 4 : Clic Mutuel en condition « Différenciée » (2)..... 297

e. Exemple 5 : Clic Mutuel en condition « Parallélisme » (1) 298

f. Exemple 6 : Clic Mutuel en condition « Parallélisme » (2) 299

ANNEXE A : Note sur les analyses statistiques inférentielles

a. Tests préliminaires

Deux grandes familles de tests statistiques sont utilisées dans ce travail : les tests paramétriques et les tests non-paramétriques. Les tests paramétriques sont le plus couramment utilisés en psychologie, mais ils sont dépendants de conditions d'application spécifiques : les effectifs ne doivent pas être trop restreints (pas d'effectifs inférieurs à 7), la distribution doit être normale, et les variances intergroupes ne doivent pas être significativement hétérogènes. Lorsque nous n'avons pas rencontré ces conditions, nous avons appliqué des tests non-paramétriques, basés sur les rangs des valeurs.

Pour tester la normalité de la distribution, nous avons utilisé le test de Shapiro et Wilk, disponible en ligne via le site d'Anastats[54]. Ce test permet de détecter si la distribution observée est comparable à une distribution dite normale (gaussienne), en étant sensible notamment à l'asymétrie et à l'aplatissement de la distribution observée. Lorsque nous souhaitons comparer deux échantillons indépendants ou plus, nous n'effectuons qu'un seul test de normalité : nous calculons à l'intérieur de chaque échantillon les écarts à la moyenne de l'échantillon en question, puis nous réunissons en une distribution l'ensemble des écarts obtenus. Le fait de ne réaliser qu'un seul test permet de réduire le risque de commettre une erreur de type I ou de type II. Lorsque le résultat de ce test est significatif (i.e. lorsque *p* est inférieur à .05), la distribution ne peut pas être considérée comme normale, et les tests non-paramétriques doivent être utilisés. A l'inverse, lorsque ce test n'est pas significatif (i.e. lorsque *p* est supérieur à .05), la distribution peut être considérée comme normale, ce qui est la première condition d'application des tests paramétriques.

Mais il reste à déterminer si les variances intergroupes sont suffisamment homogènes. En termes statistiques, il s'agit de vérifier l'homoscédasticité des résidus. Pour cela, nous avons utilisé le test de Levene[55], disponible en ligne via le site d'Anastats[54]. Ce test permet de tester l'homogénéité des variances de *k* groupes. Si le résultat de ce test est statistiquement significatif (i.e. si *p* est inférieur à .05), les variances peuvent être considérées comme étant hétérogènes, ce qui ne permet pas d'appliquer des tests paramétriques. En revanche, si le

[54] http://www.anastats.fr/ : Ce site met à disposition un ensemble d'utilitaires statistiques réalisés sous format Excel

[55] Notons que le test de Levene classique n'autorise que des distributions symétriques. Lorsque nous avons testé l'homogénéité des variances de distributions non-symétriques, nous avons appliqué la correction de Brown & Forsythe.

résultat de ce test n'est pas significatif (i.e. si *p* est supérieur à .05), les variances ne sont pas significativement hétérogènes et les tests paramétriques sont autorisés.

b. Statistiques paramétriques

Lorsque les conditions d'application des tests paramétriques se sont trouvées remplies, nous avons utilisé le logiciel LePAC[56], de façon à procéder à des analyses inférentielles de type ANOVA.

Lorsque nous comparons *k* groupes indépendants, le test F de Fisher est utilisé. Si le résultat de ce test n'est pas significatif (i.e. si *p* est supérieur à .05), nous ne pouvons pas conclure à une différence entre les *k* groupes. En revanche, si le résultat de ce test est significatif (i.e. si *p* est inférieur à .05), cela signifie qu'au moins un des *k* groupes est statistiquement différent des autres. Dans ce cas, nous comparons deux à deux les *k* groupes.

Lorsque nous comparons 2 groupes indépendants, le test t de Student est utilisé. Si le résultat de ce test n'est pas significatif (i.e. si *p* est supérieur à .05), nous ne pouvons pas conclure à une différence entre les 2 groupes. Si le résultat de ce test est significatif (i.e. si *p* est inférieur à .05), cela signifie que les 2 groupes sont statistiquement différents.

De façon à estimer la grandeur de l'effet, des procédures fiducio-bayésiennes (Lecoutre, 2000) sont disponibles sur LePAC. Celles-ci permettent d'estimer a posteriori la différence vraie (δ) avec une garantie de .90. Les conclusions sont donc établies sur l'effet calibré (différence observée divisée par l'écart-type). En accord avec Corroyer et Wolff (2003), nous avons retenu comme critère que lorsque les 2 groupes ne sont pas statistiquement différents et que l'on peut estimer avec une garantie de .90 que l'effet calibré vrai sera inférieur à une valeur de 0.33, la différence pouvait être considéré comme négligeable[57]. Parallèlement, nous avons posé que lorsque les deux groupes sont statistiquement différents, et que l'on peut estimer avec une garantie de .90 que l'effet calibré vrai sera supérieur à une valeur de 0.66, la différence pouvait être considéré comme notable[58].

[56] http://www.univ-rouen.fr/LMRS/Persopage/Lecoutre/PAC.htm
[57] soit lorsque (Pr $|\delta/\sigma|<.33) \geq .90$
[58] soit lorsque (Pr $|\delta/\sigma|>.66) \geq .90$

c. Statistiques non-paramétriques

Lorsque les conditions d'application des tests paramétriques ne sont pas réunies, nous avons appliqué des tests non-paramétriques. Pour cela, nous avons utilisé le logiciel Statistica©[59].

Lorsque nous comparons *k* groupes indépendants, le test Kruskal-Wallis est utilisé. Si le résultat de ce test n'est pas significatif (i.e. si *p* est supérieur à .05), nous ne pouvons pas conclure à une différence entre les *k* groupes. En revanche, si le résultat de ce test est significatif (i.e. si *p* est inférieur à .05), cela signifie qu'au moins un des *k* groupes est statistiquement différent des autres. Dans ce cas, nous comparons deux à deux les *k* groupes.

Lorsque nous comparons 2 groupes indépendants, le test de Mann & Withney est utilisé. Si le résultat de ce test n'est pas significatif (i.e. si *p* est supérieur à .05), nous ne pouvons pas conclure à une différence entre les 2 groupes. Si le résultat de ce test est significatif (i.e. si *p* est inférieur à .05), cela signifie que les 2 groupes sont statistiquement différents.

Dans ce dernier cas, il est possible d'estimer à posteriori la taille de l'effet. Nous avons calculé pour cela le g de Hedge, que nous avons réalisé sous un format Excel en utilisant la formule suivante :

$$g = \frac{\overline{x}_1 - \overline{x}_2}{s*}$$

où :

$$s* = \sqrt{\frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}}$$

Nous avons posé que si g est inférieur à .20, l'effet pouvait être considéré comme faible. Si g est supérieur à .20 et inférieur à .80, l'effet pouvait être considéré comme modéré. Enfin, si g est supérieur à .80, l'effet pouvait être considéré comme fort.

[59] Statistica est un logiciel de statistiques édité par Statsoft.

d. Synthèse de la méthodologie utilisée pour les analyses statistiques

Le schéma, ci-dessous, résume la méthodologie statistique que nous avons suivie tout au long de notre travail (voir Figure 81).

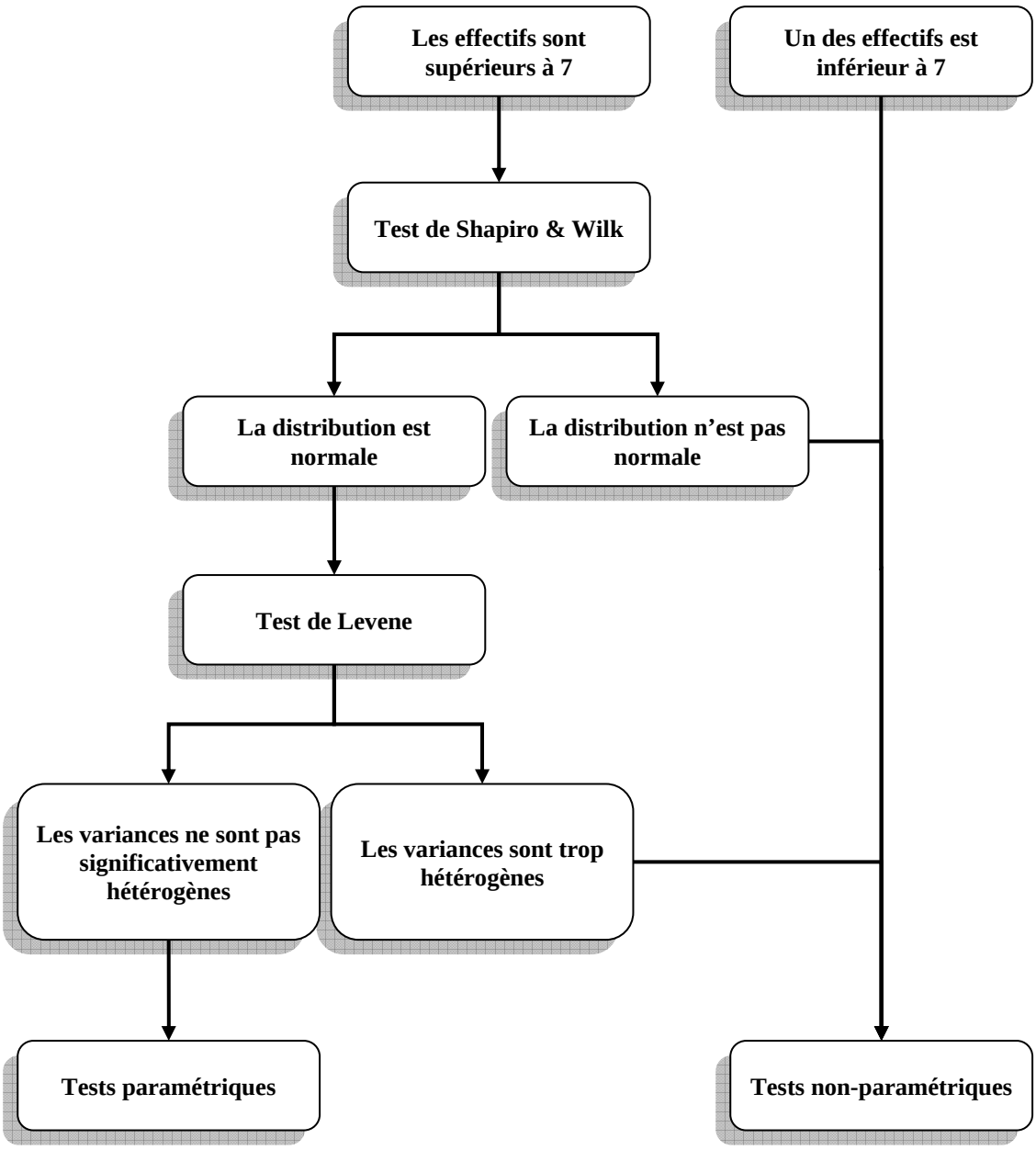


Figure 81 : Schéma récapitulatif de la méthodologie statistique

ANNEXE B : Note sur l'analyse des trajectoires perceptives

Comme nous l'avons déjà mentionné, la méthodologie minimaliste permet de forcer le déploiement des activités perceptives des sujets, de façon à les rendre accessibles à l'analyse. La mobilisation d'une médiation technique, qui plus est lorsqu'elle est numérique, permet d'enregistrer ce déploiement sous formes de fichiers traces, d'où il est possible d'extraire un ensemble d'indicateurs objectifs. Ces indicateurs permettent alors d'expliquer, ou du moins de donner des éléments d'explication concernant les performances des sujets relativement à chacune des tâches dans lesquelles nous les avons placés.

D'une façon générale, le dispositif Tactos permet d'enregistrer la position du ou des sujets dans l'espace numérique (en termes de positions en x/y) en fonction du temps. Dans l'ensemble des expérimentations et analyses d'usage que nous relatons dans ce travail, nous avons recueilli ces positions à raison d'un enregistrement toutes les 20 ms. Ainsi, nous obtenons pour chaque essai d'expérience ou pour chaque exercice d'usage un fichier « trace » daté et comportant tous les renseignements de la situation en question (voir par exemple la Figure 82).

INFO	VERSION : Tactos1V2 2.0.STAR.25 du 19 novembre 2010						
INFO	==========						
INFO	Scenario:	File:	C:\Users\LaboCred\tactos1V2.STAR.TURING\data\G1-Sc1.xml				
INFO	Scenario:	Name:	Passation1				
INFO	Scenario:	Session:	session 1				
INFO	Scenario:	Session:	01 | 16				
INFO	==========						
INFO	Users:	User 0			User 1		
INFO	Cursor:	SensorClassical-1-1-2.xml			SensorClassical-1-1-2.xml		
INFO	Image:	vide.bmp			vide.bmp		
INFO	Robot:	robot-sinus-regulier-centre aléatoire.l			robot-changement-sinus.lua		
INFO	Body:	..\data\Corp	[0;9]		..\data\Corp	[0;9]	
INFO	Groupe:	0			1		
TYPE	Tps(ms)	0.X	0.Y	0.StimMono	1.X	1.Y	1.StimMono
DATA	20	95	100	1	174	100	1
DATA	40	95	100	1	174	100	1
DATA	60	95	100	0	174	100	0
DATA	80	95	100	0	174	100	0
DATA	100	95	100	0	174	100	0
DATA	120	95	100	0	174	100	0

Figure 82 : Exemple de fichier trace. L'encadré bleu (en haut) comporte les informations générales, comme la version de Tactos utilisée ou encore le numéro de passation ou d'essai ; l'encadré rouge comporte les

informations techniques relativement à l'essai enregistré, comme le type de corps-image ou de corps-percevant ; l'encadré orange comporte les données en tant que telles.

A partir des données obtenues (positions en x/y en fonction du temps), il est possible de calculer un ensemble d'indicateurs dérivés. Par exemple, nous pouvons calculer l'écart temporel entre deux stimulations, ou encore la distance parcourue à chaque pas de temps pour chaque sujet. L'ensemble de ces données dérivées sont regroupées dans un fichier d'analyse portant sur l'essai ou la situation d'usage étudiée (voir par exemple la Figure 83).

Tps	Position du sujet 1	Distance parcourue		Position du sujet 2	Distance parcourue	
		Signée	Non signée		Signée	Non signée
20	95			174		
40	95	0	0	174	0	0
60	95	0	0	174	0	0
80	95	0	0	174	0	0
100	95	0	0	174	0	0
120	95	0	0	174	0	0
140	95	0	0	174	0	0
160	95	0	0	174	0	0
180	95	0	0	174	0	0
200	95	0	0	174	0	0

Figure 83 : Exemple d'un fichier d'analyse (extrait).

Ces données dérivées donnent alors lieu à des statistiques descriptives, de façon à obtenir une valeur par sujet, par essai et par indicateur donnés. Ces valeurs sont ensuite regroupées dans un fichier récapitulatif, à partir duquel nous pouvons généraliser les statistiques descriptives en fonction des conditions expérimentales ou encore des types de réponses données par les sujets (Par exemple, la moyenne d'un indicateur donné obtenu par un sujet lorsqu'il a correctement répondu à la tâche) (voir Figure 84).

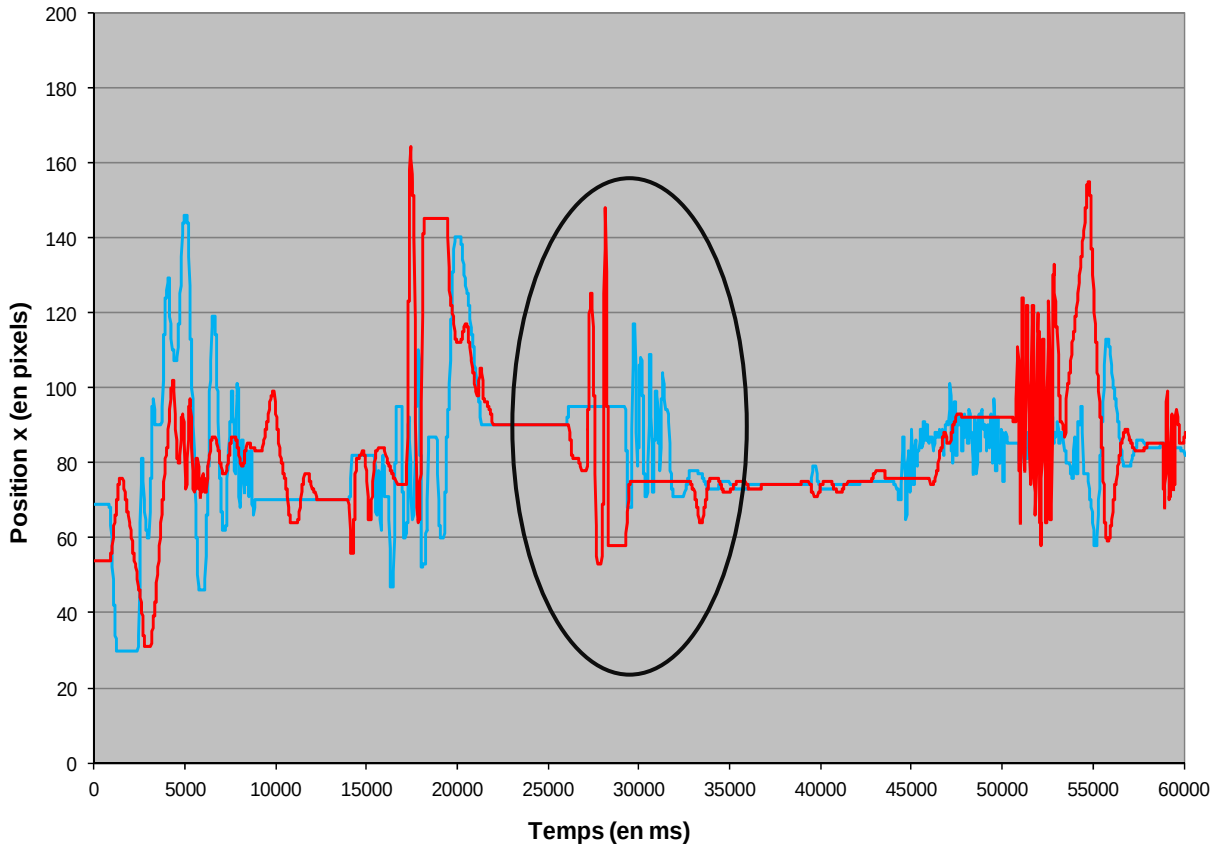
Sujet	Essai	Condition	Bonnes Réponses	Mauvaise réponse	Indicateur 1	Si bonne réponse
1	1	1	1	0	x1	
1	2	1	0	1	x2	
1	3	1	0	1	x3	
1	4	1	1	0	x4	X1
2	1	1	0	1	y1	
2	2	1	1	0	y2	
2	3	1	1	0	y3	
2	4	1	1	0	y4	Y2

Figure 84 : Exemple de fichier récapitulatif (extrait).

ANNEXE C : Exemples de trajectoires perceptives (Turing)

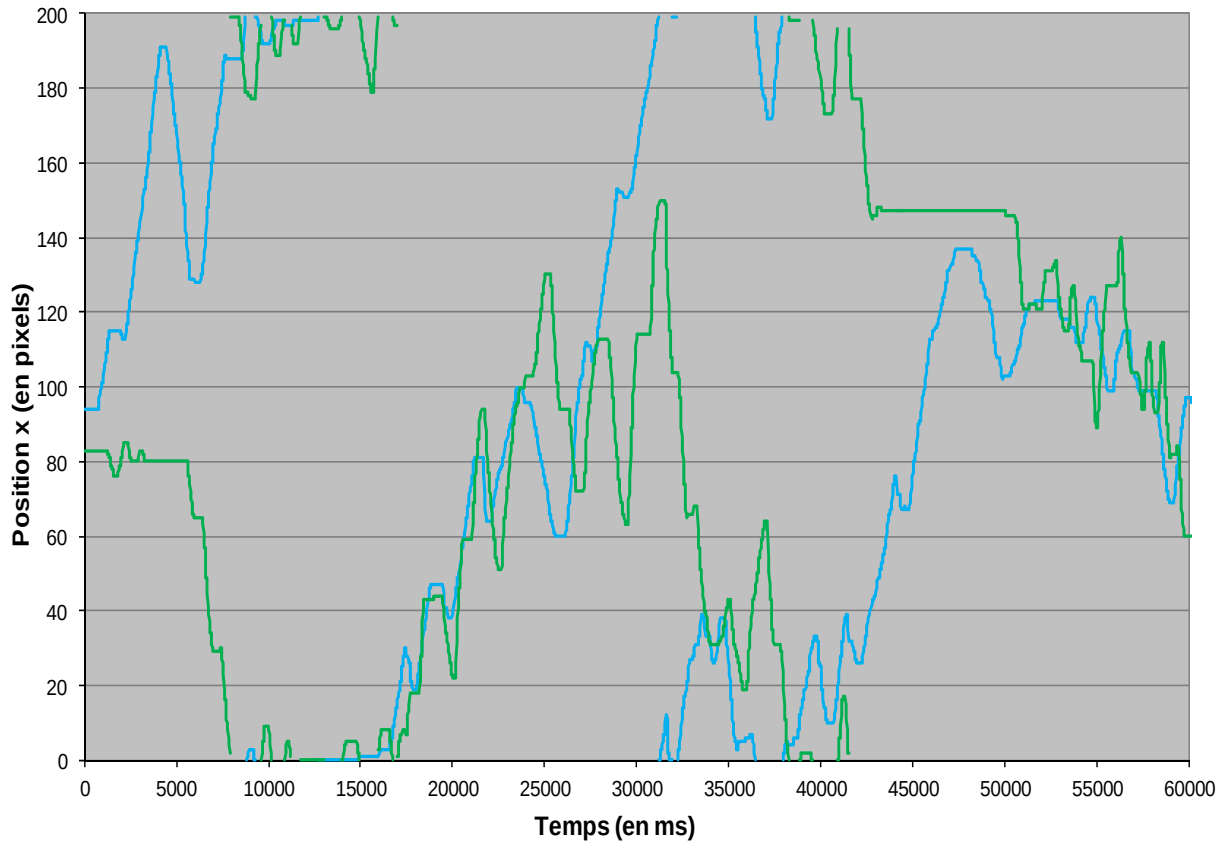
a. Exemple 1 : Interaction entre deux sujets humains (1)

Cette trajectoire correspond à un essai de la condition « Humain », tiré du groupe expérimental confronté au groupe des robots intermédiaires. Cet essai a conduit à une reconnaissance mutuelle, et est caractérisé par un haut degré d'accroche mutuelle, mais aussi par une séquence d'interaction comparable à une alternance de tours de parole (entourée en noir), témoignant d'une organisation de l'interaction à un niveau macroscopique.



b. Exemple 2 : Interaction entre deux sujets humains (2)

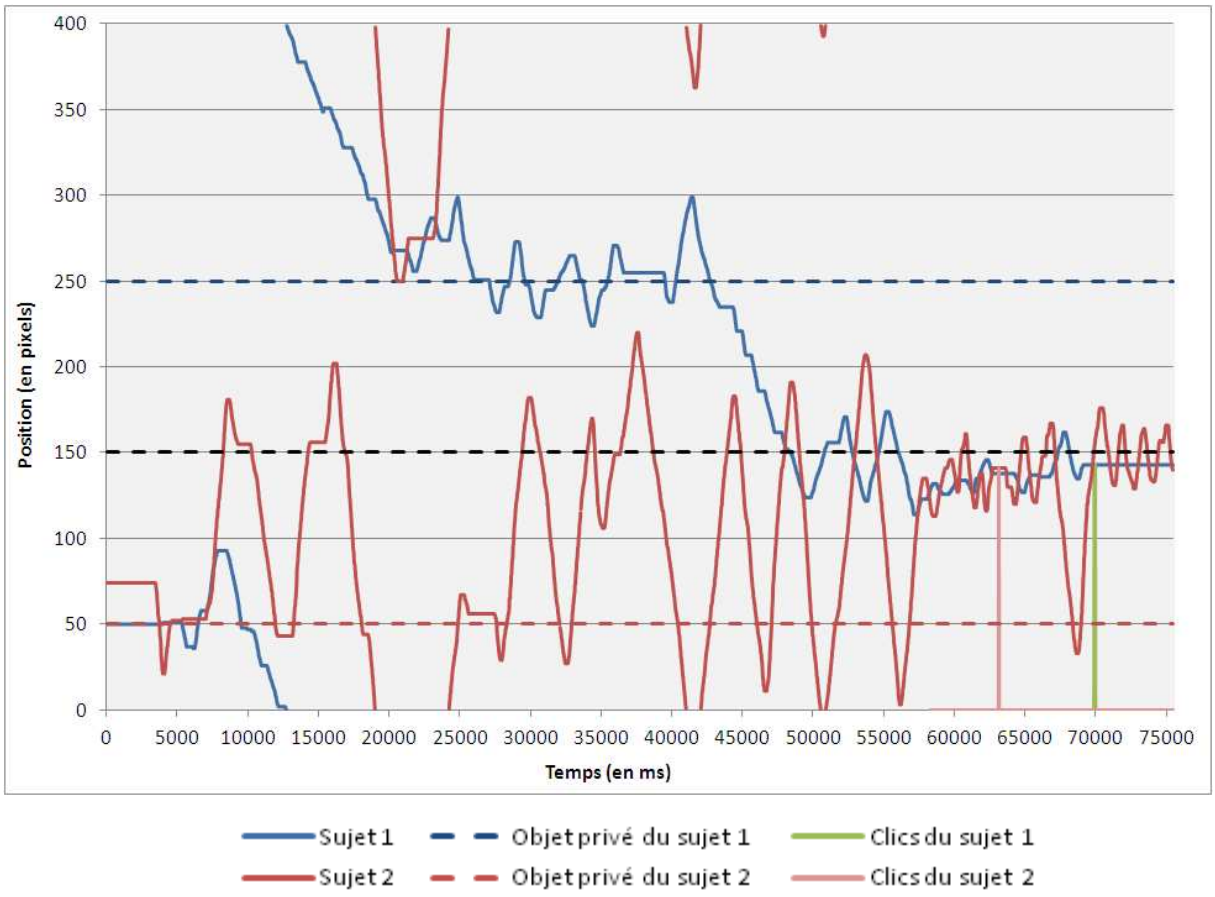
Cette trajectoire correspond à un essai de la condition « Humain », tiré du groupe expérimental confronté au groupe des robots complexes. Cet essai a conduit à une reconnaissance unilatérale (le sujet bleu a reconnu le sujet vert comme étant un humain, tandis que le sujet vert a reconnu le sujet bleu comme étant un robot). Cette interaction est caractérisée par une accroche mutuelle faible, et une variabilité modérée de cette accroche.



ANNEXE D : Exemples de trajectoires perceptives (Public/Privé)

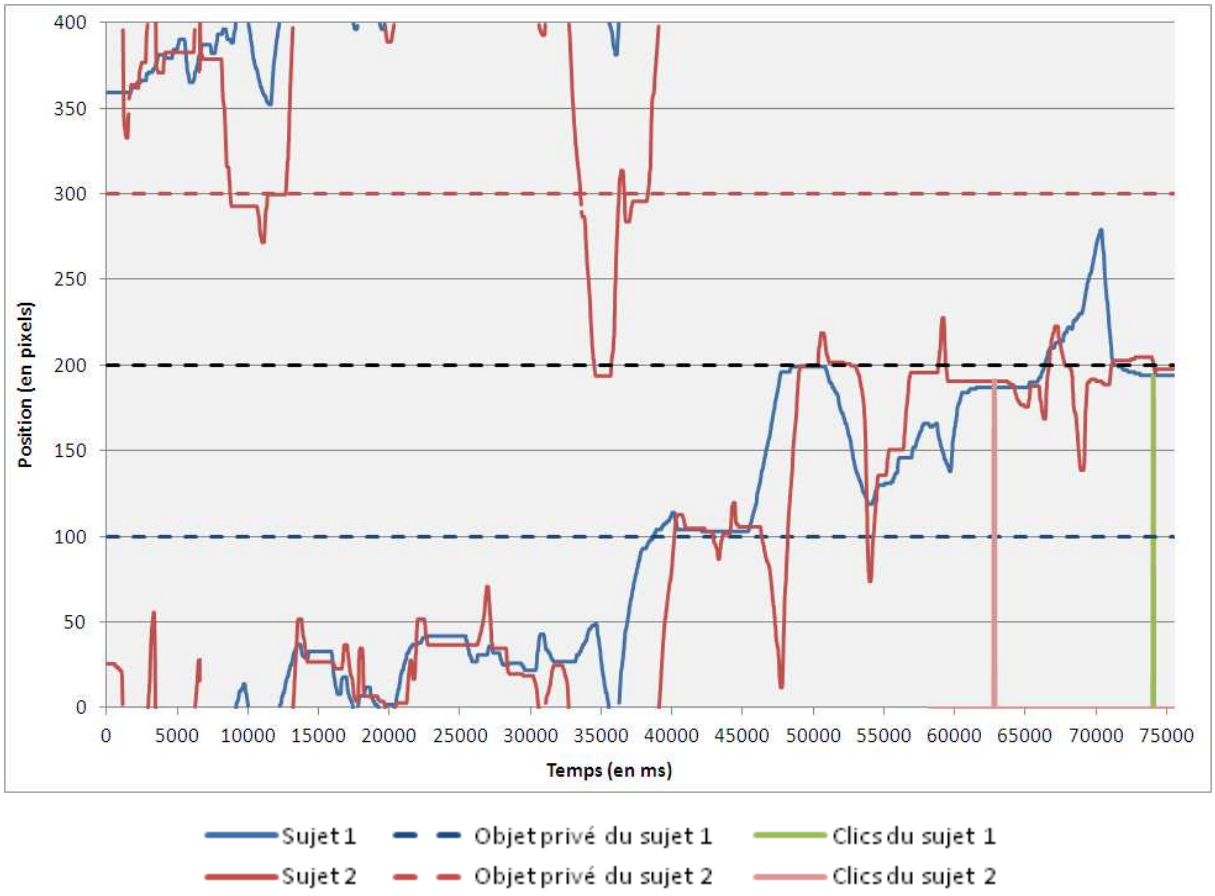
a. Exemple 1 : Clic Mutuel en condition « Mono » (1)

Cette trajectoire perceptive correspond à un essai de la condition « Mono ». Ce qu'il est intéressant de remarquer ici, c'est que les activités perceptives sont d'abord tournées vers la recherche des objets, malgré un premier contact interpersonnel survenu très tôt au début de l'essai. De cette façon, les sujets se rencontrent d'abord respectivement autour de leur objet privé respectif, mais l'interaction ne s'y stabilise pas. En revanche, lorsqu'ils se rencontrent autour de l'objet commun, l'interaction se stabilise et les sujets coordonnent leurs activités perceptives à proximité de celui-ci. Cet essai a conduit à un clic mutuel sur l'objet commun.



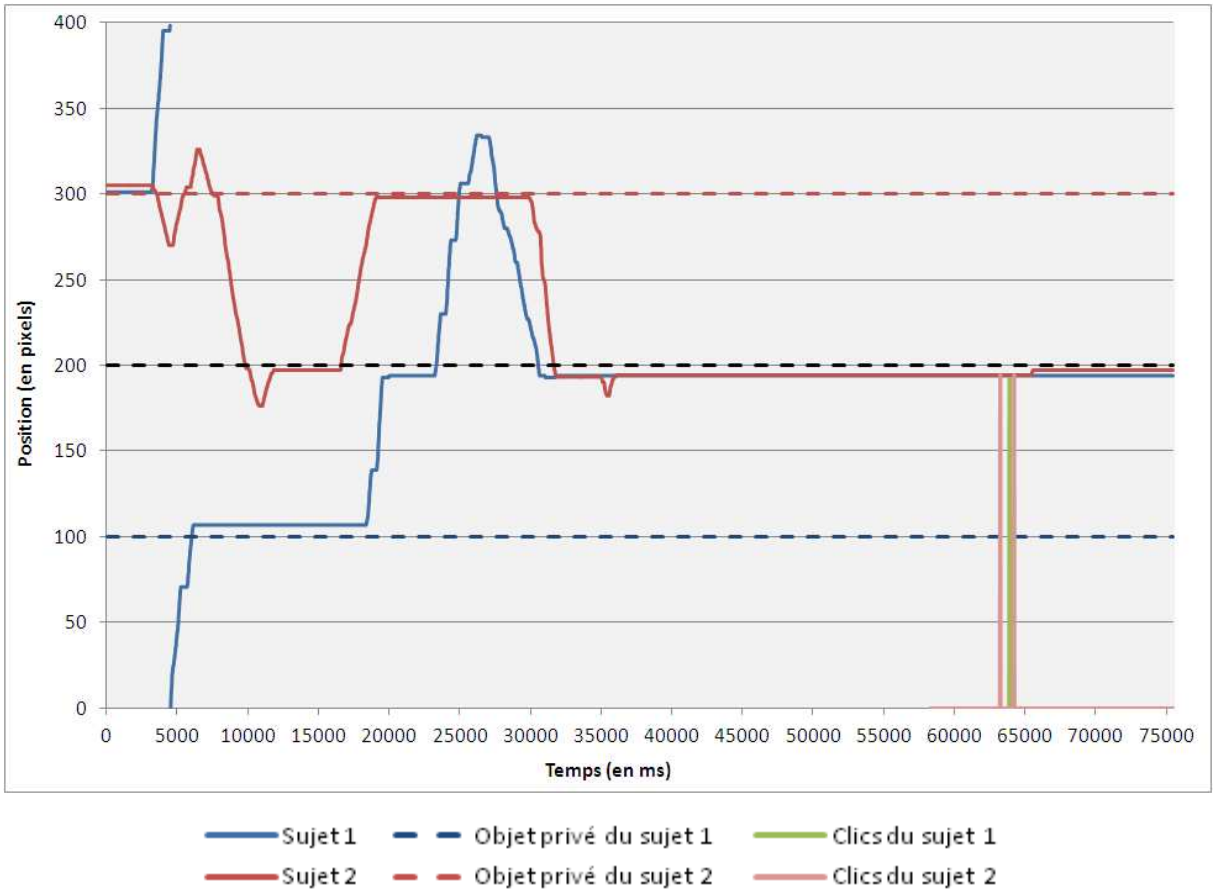
b. Exemple 2 : Clic Mutuel en condition « Mono » (2)

Cette trajectoire perceptive correspond à un essai de la condition « Mono ». Nous notons que cette fois, les activités perceptives déployées par les deux sujets sont d'abord dirigées l'une vers l'autre. Ainsi, ils se suivent dans l'espace et explorent ensemble les objets qui y sont présents, même si ponctuellement, nous observons que le sujet 2 se désengage de l'interaction pour localiser les objets. La rencontre de l'objet privé du sujet 1 produit une courte séquence d'interaction, mais qui ne s'est pas maintenue dans le temps. Cet essai a conduit à un clic mutuel sur l'objet commun.



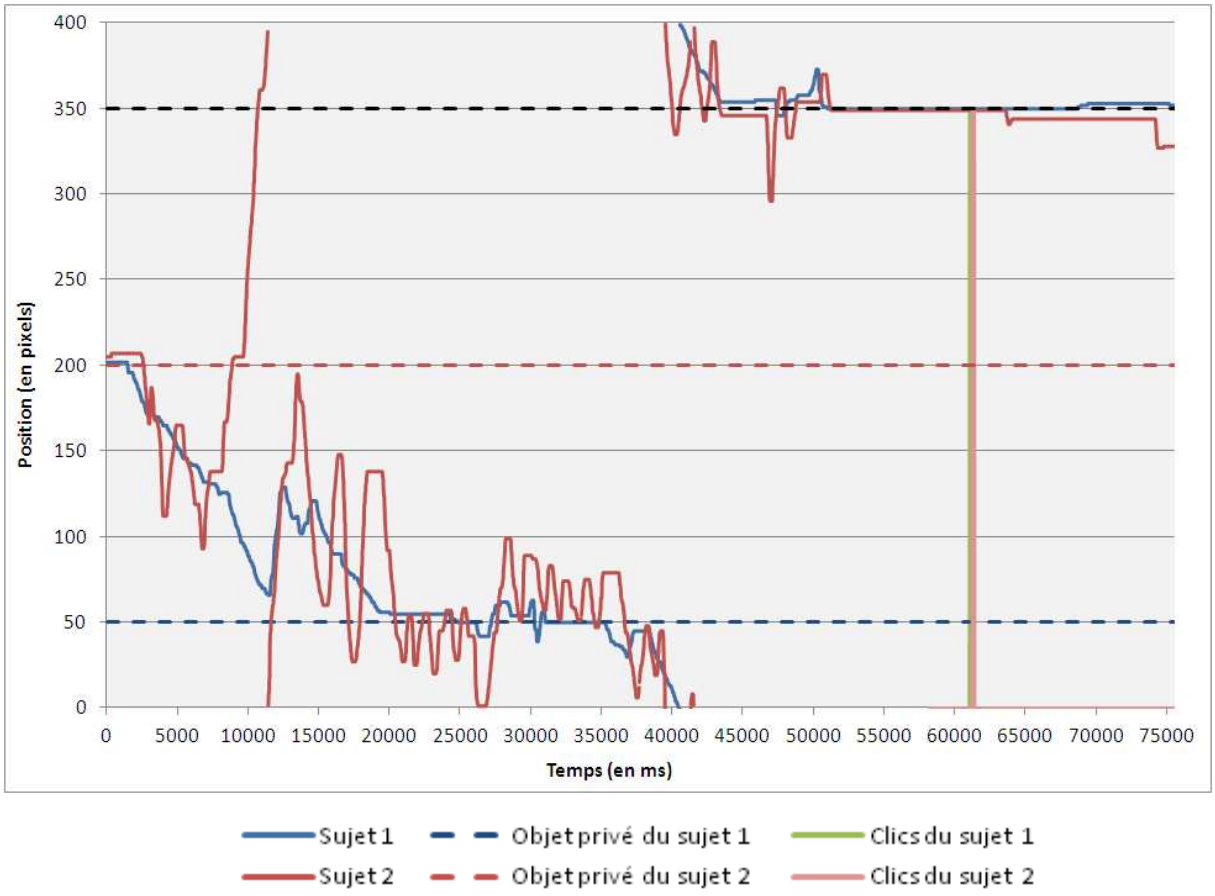
c. Exemple 3 : Clic Mutuel en condition « Différenciée » (1)

Cette trajectoire perceptive est issue d'un essai de la condition « Différenciée ». Ici les stratégies individuelles sont clairement orientées vers les objets : les deux sujets recherchent d'abord les objets, puis s'y arrêtent en attendant l'arrivée de leur partenaire. De cette façon, le déploiement des activités perceptives est caractérisé par une succession d'arrêts sur les objets, qui ne se maintiennent pas dans le cas de la rencontre des objets privés, mais qui est maintenu une bonne partie de l'essai lorsqu'il survient sur l'objet commun. L'essai a conduit à un clic mutuel correct.



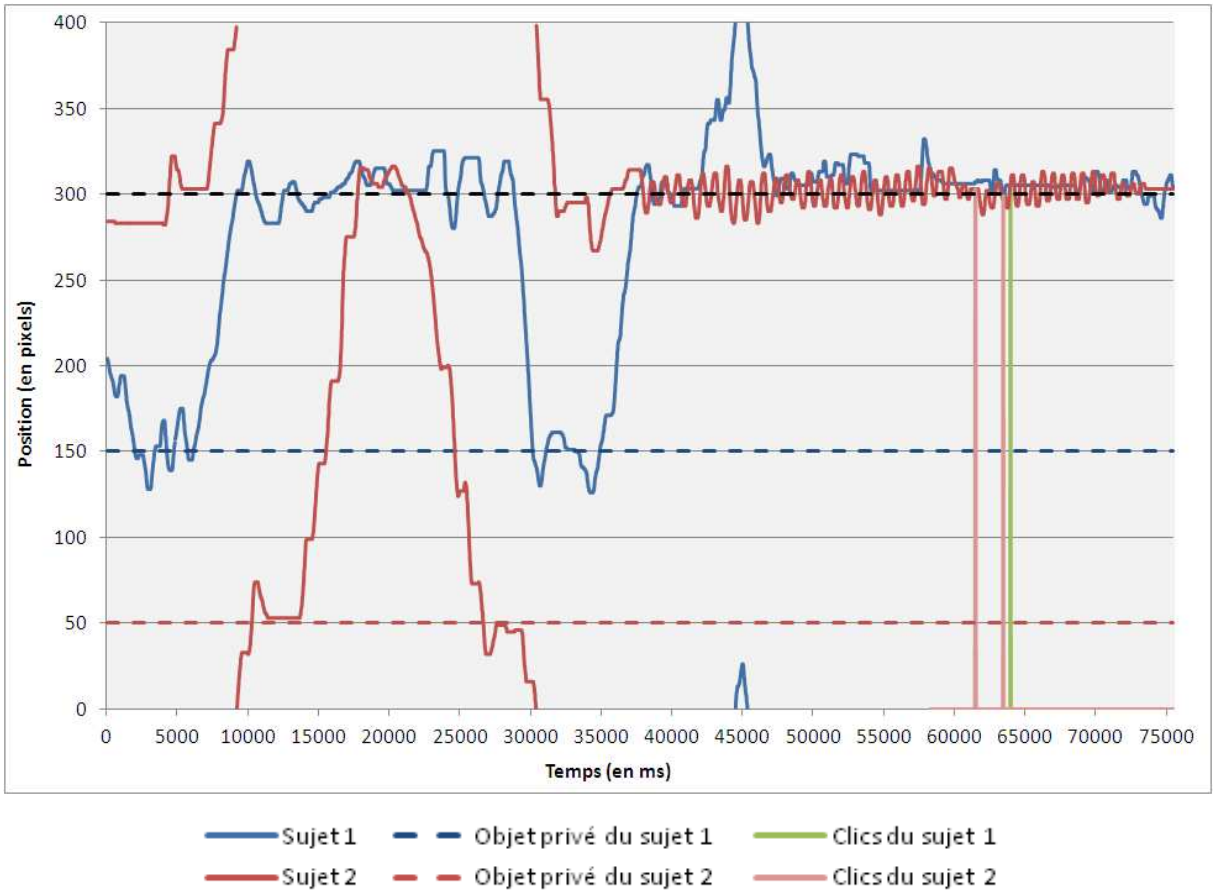
d. Exemple 4 : Clic Mutuel en condition « Différenciée » (2)

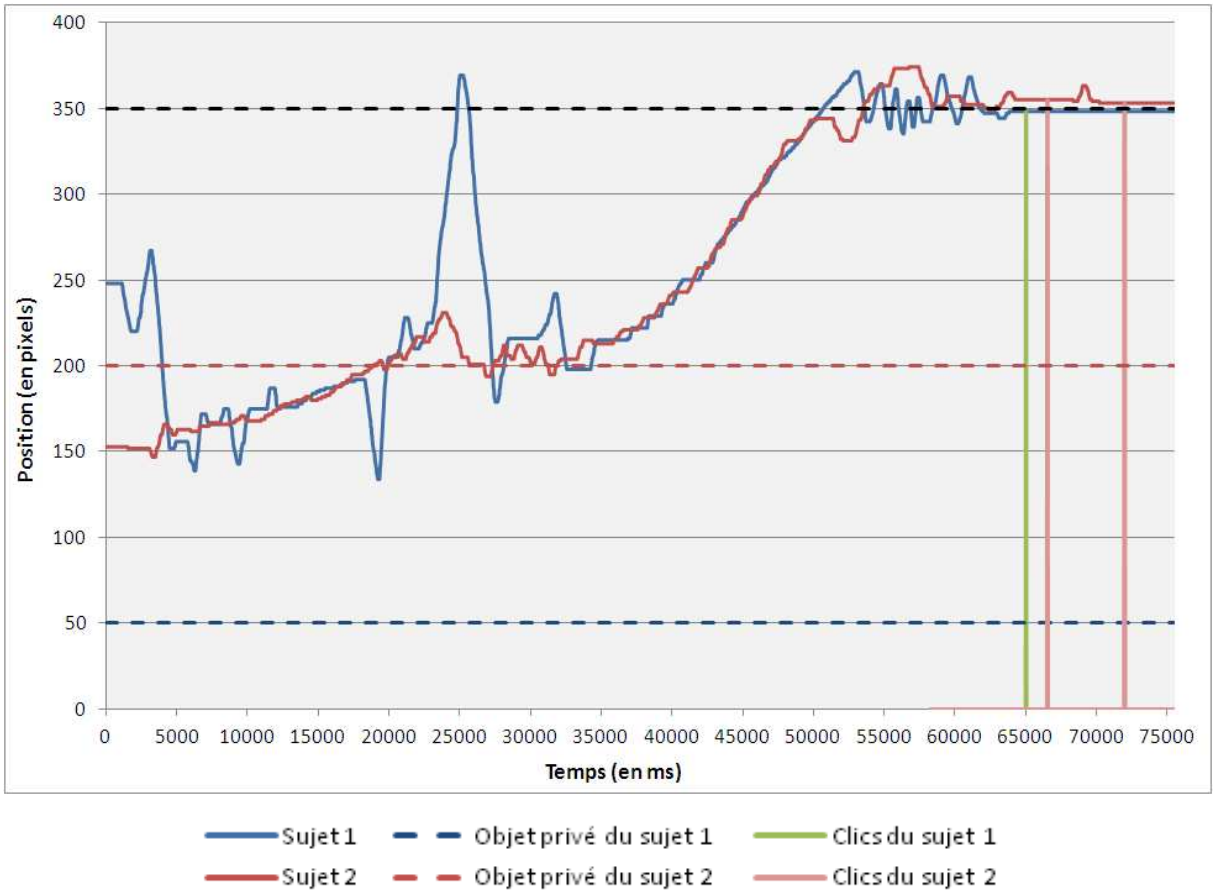
Cette trajectoire perceptive correspond à un essai de la condition « Différenciée ». Contrairement à l'exemple 3, les stratégies sont ici orientées l'une vers l'autre. Les sujets semblent d'abord vouloir maintenir le contact, de façon à explorer ensemble l'espace et les objets présents. Notons que la rencontre de l'objet privé du sujet 1 conduit ce dernier à s'arrêter, mais le sujet 2 continue ses oscillations. En revanche, lorsque les deux sujets arrivent sur l'objet commun, l'interaction se stabilise jusqu'à l'arrêt mutuel, qui conduit alors à un clic mutuel correct en fin d'essai.



e. Exemple 5 : Clic Mutuel en condition « Parallélisme » (1)

Cette trajectoire perceptive est issue d'un essai de la condition « Parallélisme ». Ici, les sujets semblent déployer une stratégie orientée vers les objets. Ils recherchent alors activement les objets présents dans l'espace, en passant de l'un à l'autre. Alors qu'ils se rencontrent une première fois autour de l'objet commun, les sujets poursuivent leur exploration. Au milieu de l'essai, lorsqu'ils se rencontrent de nouveau autour de l'objet commun, le sujet 2 signale sa présence à l'autre en effectuant de rapides oscillations avec une faible amplitude. Le sujet 1, au contraire, déploie une activité qui semble plus chaotique, mais maintient tout de même sa présence autour de l'objet et de l'autre. L'essai conduit à un clic mutuel sur l'objet commun.



f. Exemple 6 : Clic Mutuel en condition « Parallélisme » (2)

Cette trajectoire est issue de la condition « Parallélisme ». Contrairement à l'exemple précédent, les stratégies sont ici clairement orientées l'une vers l'autre. Les sujets cherchent donc d'abord à se rencontrer dans l'espace, avant de l'explorer ensemble. Notons que lorsque le sujet 2 rencontre son objet privé, il suit d'abord le sujet 1, mais revient autour de l'objet. Pendant ce temps, le sujet 1 part à la recherche d'un objet (il n'en a pas encore rencontré), puis revient au contact du sujet 2, qui était resté sur son objet privé. Le sujet 1 parvient alors à emmener le sujet 2 dans la direction de l'objet qu'il a rencontré, et leur interaction s'organise alors autour de l'objet commun. L'essai a conduit à un clic mutuel correct.