

Etude outillée de l'attention visuelle flottante dans la recherche collaborative d'informations sur grands écrans

Damien Marion, Pierre--alexandre Favier, Christophe Bey, Sami Lini, Jean--marc Andre

▶ **To cite this version:**

Damien Marion, Pierre--alexandre Favier, Christophe Bey, Sami Lini, Jean--marc Andre. Etude outillée de l'attention visuelle flottante dans la recherche collaborative d'informations sur grands écrans. EPIQUE 2015, Jul 2015, Aix-en-Provence, France. pp.261-268. hal-01294798

HAL Id: hal-01294798
https://hal.science/hal-01294798

Submitted on 1 Apr 2016

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Etude outillée de l'attention visuelle flottante dans la recherche collaborative d'informations sur grands écrans

Damien MARION
109 avenue Roul, CS40007, 33405 Talence Cedex
damien.marion@ensc.fr

Pierre-Alexandre FAVIER
109 avenue Roul, CS40007, 33405 Talence Cedex
pierre-alexandre.favier@ensc.fr

Christophe BEY
109 avenue Roul, CS40007, 33405 Talence Cedex
christophe.bey@ensc.fr

Sami LINI
109 avenue Roul, CS40007, 33405 Talence Cedex
sami.lini@akiani.fr

Jean-Marc ANDRE
109 avenue Roul, CS40007, 33405 Talence Cedex
jean-marc.andre@ensc.fr

RÉSUMÉ

Dans cet article, nous présentons un système basé sur la détection des mouvements de tête qui permet l'affichage sur un écran de grand format du point d'attention de chaque individu. Ainsi, chaque individu sait où regarde à l'écran chacun de ses partenaires. Nous posons l'hypothèse que ce système améliore la recherche collective d'information en introduisant une collaboration implicite entre les individus.

MOTS-CLÉS
Head Tracking, interface adaptative, public display, point d'attention, collaboration implicite

1 PROBLEMATIQUE ET HYPOTHESES

La démocratisation récente d'écrans de très grands formats et de la « haute définition » change le rapport de l'utilisateur à l'écran. Ces derniers sont de plus en plus en attente des nouvelles technologies, qui doivent leur offrir de plus en plus de fonctionnalités intuitives et naturelles. De ce fait, on voit émerger de nombreux travaux sur les « interfaces intelligentes », qui s'adaptent, se transforment en fonction du nombre d'individus les utilisant ou encore en fonction de l'intérêt que peut porter un utilisateur à l'interface en question (Schiavo, Mencarini, Vovard, & Zancanaro, 2013).

C'est pourquoi, nous nous intéressons ici au partage collaboratif « implicite » d'une surface d'affichage à l'aide du mécanisme de la vision « flottante ». En effet, Nous posons l'hypothèse générale qu'un groupe de personnes à la recherche d'informations sur un écran de grand format sera plus performant dans leur détection si chaque utilisateur est conscient de l'endroit où regardent ses partenaires. Nos travaux se rapprochent sur ce point de ceux de Khan, Matejka, Fitzmaurice et

Les auteurs conservent les droits de leurs publications.

N. Bonnardel, L. Pellegrin & H. Chaudet (Eds.). *Actes du Huitième Colloque de Psychologie Ergonomique (EPIQUE 2015), Aix-en-Provence, France, 8-10 juillet 2015*. Paris, France: Arpege Science Publishing. ISBN 979-10-92329-02-5

Kurtenbach (2005), qui ont exploré une technique permettant de focaliser l'attention d'un groupe de personne en projetant un faisceau lumineux sur la zone de l'écran nécessitant l'attention de l'audience. Cette conscience de l'endroit regardé participe à la construction d'une conscience de groupe (Benali, Bourguin, David, Derycke, & Ferraris, 2002). C'est le mécanisme permettant la perception de l'endroit regardé par ses partenaires que nous nommons « vision flottante », par analogie avec le mécanisme dit « d'écoute flottante » qui permet de percevoir des informations dans une conversation dont on est témoin, même si elle ne nous est pas destinée. L'expérience présentée dans cet article consiste en l'affichage d'images très chargées en informations dans lesquelles trois utilisateurs doivent retrouver une liste d'éléments, à la manière des jeux pour enfant du type « Où est Charlie ?».

La présentation des images se fait selon deux modalités :

- la modalité d'affichage simple (**S**), qui présente la liste des éléments à retrouver et l'image dans laquelle les chercher ;
- la modalité d'affichage avec point d'attention (**PA**), qui présente le même affichage que la modalité **S**, avec en complément la matérialisation du point d'attention de chacun des utilisateurs.

Nous partons du postulat que lors de la recherche collaborative d'informations, la performance du groupe devrait être améliorée et que le temps pour trouver les éléments devrait être moins important lorsque le point d'attention de chaque membre est affiché.

2 MILIEU D'IMPLANTATION

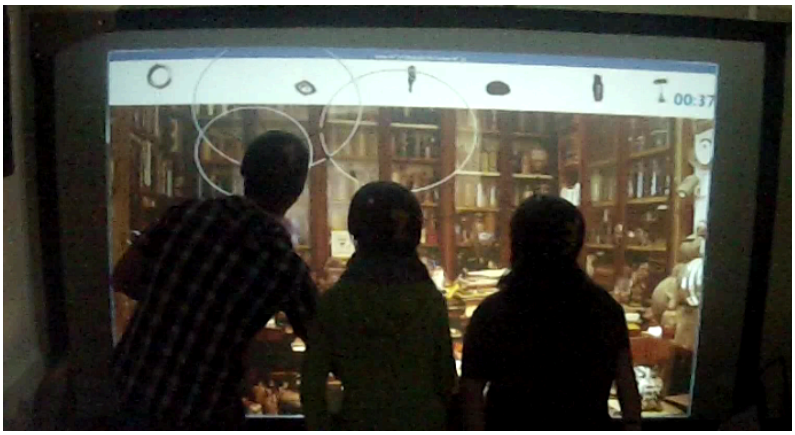


Figure 1 : Photographie de l'expérience

2.1 Méthode

Techniquement, nous utilisons un système de Head Tracking basé sur la détection infra-rouge de marqueurs. La matérialisation du point d'attention est réalisée par l'affichage à l'écran d'une ellipse transparente munie d'une double bordure, une noire et une blanche, afin qu'elle soit toujours visible quelle que soit la couleur du fond des images affichées (figure 1). Les dimensions de l'ellipse sont calculées afin de représenter la zone de discrimination des couleurs, soit 60° en horizontale de part et d'autre de l'axe optique et de 30° en verticale vers le haut et 40° vers le bas (Aw & Rouhet, 2004).

L'expérience se découpe en quatre séquences. La première partie consiste à faire passer le test psychologique informatisé de mesure des styles cognitifs d'indépendance et dépendance au champ, Extended CSA-WA (Peterson & Deary, 2006). Cette mesure des styles cognitifs est essentielle afin de former des groupes de trois utilisateurs les plus homogènes possibles quant aux stratégies d'exploration visuelles mises en place. En effet nous avons formé les groupes d'utilisateurs de sorte que les scores du ratio d'indépendance et dépendance au champ de chaque groupe soient les plus proches possibles.

L'expérience s'est déroulée dans un environnement contrôlé. Les participants sont placés volontairement près de l'écran dans une zone 3,50 mètres de longueur et de 1 mètre de largeur par rapport à l'écran. Cette zone est délimitée physiquement par l'écran et le mobilier de la salle d'expérimentation qui les empêche d'en sortir (figure 2).

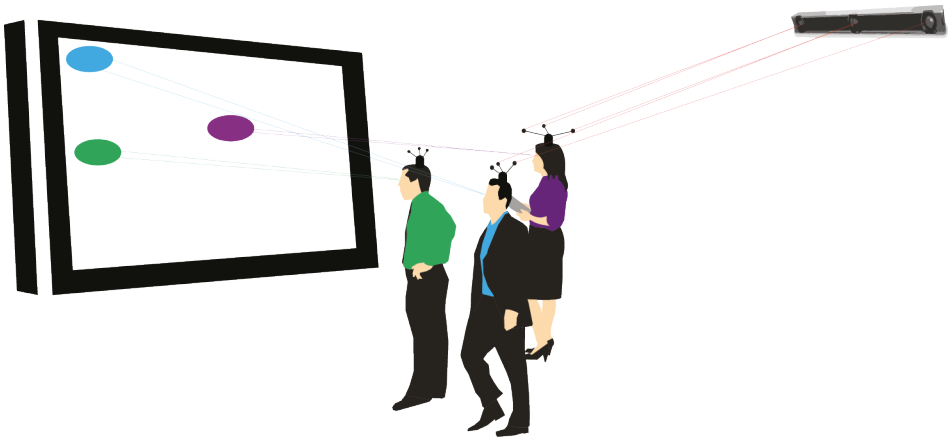


Figure 2 : Schéma du dispositif

Cette contrainte garantit que les sujets restent dans la zone de détection : durant toute cette phase, les mouvements de têtes des 3 sujets sont traqués en temps réel. Aucune consigne les empêchant de se déplacer dans cette zone n'est formulée. Cette contrainte physique est nécessaire afin que le parcours complet de l'image puisse être réalisé « confortablement » seulement en bougeant la tête et non par un simple balayage du regard (tête immobile). Cette proximité à l'écran impose ainsi le recours à des mouvements détectables par notre dispositif de Head Tracking. Des mouvements plus ténus nécessiteraient l'utilisation d'un système d'Eye-Tracking multi-utilisateurs, ce qui serait techniquement complexe et coûteux à mettre en œuvre. D'autre part, nous nous intéressons à la détection de la focalisation sur une zone de l'écran, pas sur un point précis de l'affichage. La précision apportée par notre dispositif est donc suffisante à la mise en évidence de l'impact de la matérialisation de la zone d'attention sur l'élaboration de stratégies de recherche visuelles. Les prises d'information brèves qui engendrent un simple mouvement oculaire sans mouvement de la tête ne correspondent pas à une consultation pertinente dans notre cas d'utilisation. L'exploitation des seuls mouvements de tête intercepte correctement les comportements que nous ciblons.

Les conditions de luminosité sont également maitrisées, du fait de l'utilisation d'une technologie basée sur des émetteurs et capteurs infrarouges, et afin de garantir la stricte équivalence des conditions de passation entre les groupes. Une souris est placée devant les candidats pour qu'ils puissent cliquer sur l'élément trouvé dans l'image. Quand le clic désigne bien un élément à trouver, la zone cliquée correspondante est mise en surbrillance. Simultanément, l'élément trouvé disparaît de la liste des objets à rechercher du bandeau supérieur. Dans le cas contraire, le clic est comptabilisé comme un mauvais clic et il ne se passe rien à l'écran. A chaque clic des participants, une fenêtre évènementielle apparait sur l'écran de l'expérimentateur qui doit alors renseigner le numéro de l'utilisateur ayant cliqué.

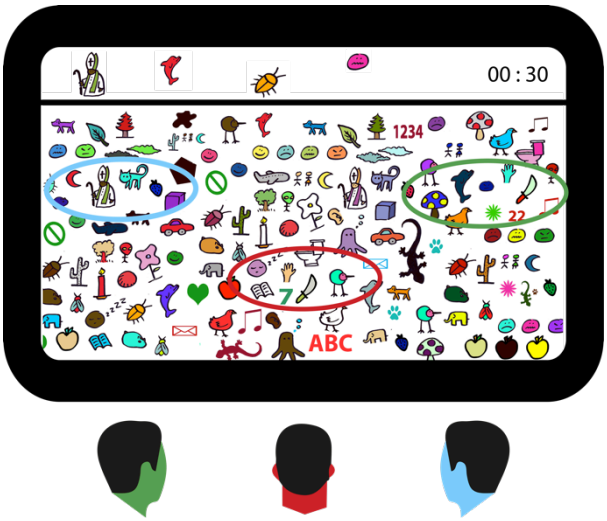


Figure 3 : Schéma de principe de l'expérience

Durant toute l'expérience, les participants ont pour consignes de ne communiquer d'aucune manière que ce soit avec leurs coéquipiers. Cependant, afin de rendre l'expérience plus ludique et de plonger les participants dans la tâche, un score collectif est calculé pendant la troisième partie. Ce score est présenté comme un score d'équipe prenant en compte la date du clic, le nombre d'objets trouvés et le nombre de mauvais clics. Le mauvais clic n'enlève en fait pas de points au score mais la consigne dit explicitement le contraire afin d'éviter les stratégies de clics à l'aveugle pour trouver un maximum d'objets.

Dans la deuxième phase de l'expérience, il est présenté deux images tests en séquence, une image avec la modalité **S** et une image avec la modalité **PA**.

Dans la troisième phase de l'expérience, deux séries de quinze images par groupe sont déroulées. Une série avec la modalité **S** et une série avec la modalité **PA**. L'ordre des séries, et donc des modalités, est contrebalancé pour chaque groupe d'utilisateur. Les numéros de groupe pairs commencent avec la modalité **S** et inversement pour les numéros de groupes impairs. Sur chaque image, les utilisateurs doivent retrouver six éléments en quarante secondes. Les six éléments à retrouver sont affichés au dessus de l'image et placés en nuage pour éviter une répartition implicite des éléments entre les individus. De plus, les utilisateurs sont placés près de l'écran et donc pour éviter les gênes éventuelles, au niveau de la nuque, causées par des mouvements trop répétitifs de haut en bas, les objets à retrouver sont placés dans un bandeau en haut de l'écran ainsi que le compte à rebours de quarante secondes. Sur les six éléments à retrouver, il y a toujours un élément non présent dans l'image, un "leurre", afin d'inciter les participants à chercher pendant toute la durée de présentation d'une image. Entre chaque présentation d'image un écran affichant un compte à rebours de trois secondes s'affiche. De plus, avant le passage d'une série à l'autre, les utilisateurs sont avisés du changement de modalité pour la série suivante, pendant 20 secondes, avec affichage du point d'attention si le groupe a commencé par la modalité S ou disparition du point d'attention si le groupe a commencé avec la modalité **PA**. Concernant, l'affichage des points d'attention, nous avons choisi de les symboliser de la même couleur pour les 3 sujets afin qu'il n'y ait pas de stratégie compensatoire vis à vis d'un seul individu du groupe (en cas d'*a priori* sur la performance des autres membres).

Enfin, dans un dernier temps nous recueillons le ressenti immédiat des participants.

En analysant, le nombre d'objets trouvés selon la modalité d'affichage présentée, nous testons les hypothèses suivantes :

(H1) le nombre d'objets trouvés est supérieur pour les groupes ayant commencé l'expérience avec la modalité **PA** ;

(H2) le temps pour trouver le premier objet est plus faible pour les groupes ayant commencé l'expérience avec la modalité **PA** ;

(H3) le nombre d'objets trouvés par image est plus important pour les groupes commençant en modalité **PA** ;

(H4) les groupes commençant en modalité **PA** réalisent moins de mauvais clics.

En effet, nous souhaitons démontrer l'avantage significatif de l'affichage du point d'attention, lors de la première série d'images, permettant ainsi d'établir une stratégie de recherche individuelle, basée sur l'interprétation personnelle de chacun des participants, concernant la stratégie de recherche mise en place par ses coéquipiers et donc par conséquent la meilleure stratégie à adopter pour améliorer sa propre performance, et la performance collective.

Il est important de noter qu'aucun utilisateur ne s'est déplacé dans la zone plutôt restreinte durant l'expérience malgré qu'aucune consigne n'ait été formulée dans ce sens. De ce fait, les utilisateurs sont restés en ligne et ont conservé l'ordre de placement du début à la fin de l'expérience. L'utilisateur numéro un est toujours à gauche, le numéro deux au centre et le numéro trois sur la droite de l'écran.

2.2 Matériel

Les images sont projetées sur un écran de 2.40 mètres de base horizontal, au format 16/9, permettant une projection par l'arrière. Un capteur Optitrack V3 Trio placé derrière les utilisateurs, au plafond de la salle, détecte les mires infrarouges placées sur les casques portés par chacun des participants. L'expérience est filmée à l'aide d'une caméra GoPro afin de lever les éventuelles ambiguïtés durant le dépouillement des données. Le corpus d'images présentées est de type varié : dessins, photos, infographies, nuages de tags. L'utilisation de différents types d'image dans le corpus est en adéquation avec les guides de conception évoqués par Alt, Schneegaß, Schmidt, Müller et Memarovic (2012). Toutes les images utilisées sont libres de droits et vont servir à constituer une banque de données qui sera mise à disposition.

2.3 Participants

Cinquante-quatre personnes ont participé à l'expérience : dix-huit groupes de trois personnes (n=18), majoritairement des étudiants. Les groupes ont été formés au hasard des disponibilités des individus en veillant tout de même à respecter l'homogénéité des ratios de chaque groupe concernant le style cognitif d'indépendance et dépendance au champ comme évoqué précédemment.

3 RESULTATS

Les résultats décrits ci-après n'exposent qu'une première partie des données. Le dépouillement complet étant encore en cours. Les données présentées ici correspondent aux données relatives aux évènements enregistrés durant l'expérience. Une étude ultérieure présentera les données relatives aux mouvements de tête et au d'attention des utilisateurs (4,5 Go de données consolidées en cours d'analyse).

		Ordre																	
		PA-S									S-PA								
Modalité	PA	51	51	50	51	53	55	57	48	54	49	50	51	41	52	48	55	47	52
	S	39	39	49	38	51	52	49	36	49	51	51	58	45	54	57	54	39	51

Figure 4 : Nombre d'objets trouvés par groupe selon la modalité

Dans la figure 4, la variable « ordre » correspond à l'ordre d'affichage des modalités durant l'expérience. Elle indique qu'un groupe a commencé l'expérience avec la modalité **PA** puis a terminé avec la modalité **S** et inversement. Nous allons comparer le nombre d'objets trouvés selon la modalité de démarrage de l'expérience, soit 9 groupes ayant démarré en **PA** puis terminé en **S** (**PA-S**) et 9 groupes ayant démarré en **S** puis terminé en **PA** (**S-PA**).

Après, un test de Mann-Whitney, nous n'obtenons pas de différence significative entre le nombre d'objets trouvés par les groupes commençant avec le point d'attention dans la première partie de l'expérience et ceux terminant avec cette modalité. Il nous est donc impossible de vérifier statistiquement H1. Cependant, il est important de noter que la moyenne d'objets trouvés si on commence l'expérience avec la modalité **PA** est supérieure de 5,9 objets à celle des groupes ayant commencé l'expérience par la modalité **S**. Ainsi on obtient une moyenne de 101,7 (σ = 5.9, Q1 = 100, Med = 101, Q3 = 105) objets pour les groupes en modalités **PA-S** contre une moyenne de 95,8 (σ = 11.8, Q1 = 90, Med = 100, Q3 = 105) pour les groupes ayant commencé en **S-PA**.

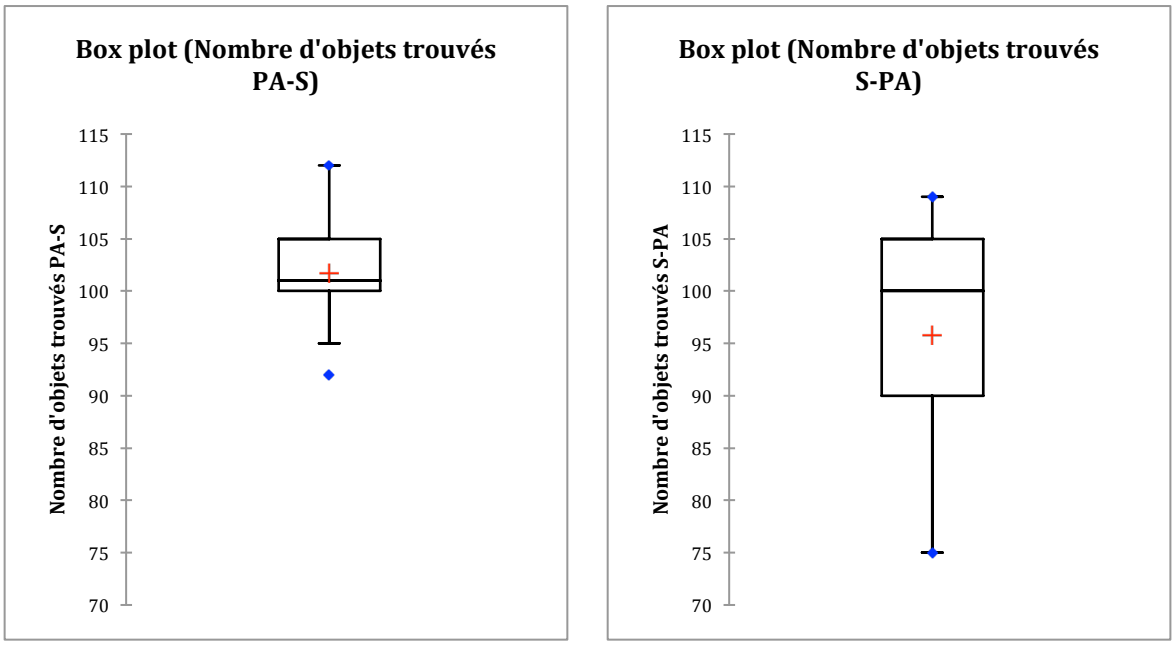


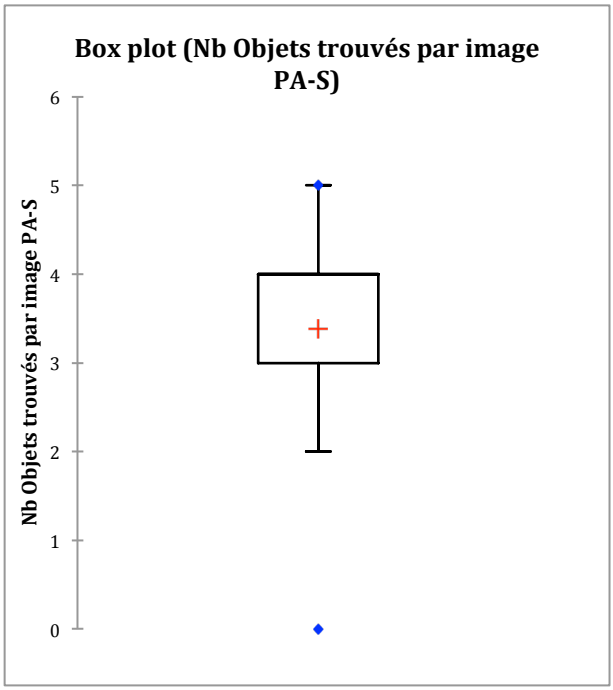
Figure 5 et 6 : Nombre d'objets trouvés en fonction de l'ordre de la modalité de démarrage de l'expérience

Comme le montre les figures 5 et 6 la performance des groupes commençant en modalité **PA** est supérieure mais également moins dispersée que celle des groupes commençant en modalité **S**. Nous constatons donc que l'observation tendancielle conforte H1, mais l'analyse statistique n'autorise pas de conclusion en ce sens.

Concernant l'hypothèse H2, nous étudions le temps pour trouver le premier objet selon l'ordre de réalisation de l'expérience : **PA-S** ou **S-PA**. Dans les résultats qui suivent, la valeur du temps présentée correspond à la valeur du compte à rebours au moment du clic et non au temps mis par les utilisateurs à cliquer sur le premier objet trouvé.

Ainsi, si on veut déterminer le temps écoulé avant le premier clic, il faut soustraire le temps du clic présenté au temps de présentation d'une image qui est de quarante secondes. Les résultats montrent que la moyenne de temps du premier clic sur un objet trouvé pour les groupes en **PA-S** est de 30,8 secondes (σ = 6,1, Q1 = 29, Med = 33, Q3 = 35). C'est-à-dire qu'en moyenne les groupes en **PA-S** mettent 9,2 secondes après le début de la présentation d'une image pour trouver un objet. Pour les groupes en **S-PA**, on observe une moyenne de 30.7 secondes (σ = 6,1 Q1 = 29.75, Med = 33, Q3 = 34.25). Il est donc difficile de conclure concernant H2 tant la différence est minime, ce qu'un test de Mann-Whitney confirme.

L'étude du nombre moyen d'objets trouvés par image selon l'ordre de passage de l'expérience **PA-S** vs **S-PA**, permet d'établir une moyenne de 3.4 objets trouvés en moyenne par image (σ = 1,2, Q1 = 3, Med = 4, Q3 = 4) pour les groupes en **PA-S** contre 3.2 objets trouvés en moyenne par image (σ = 1,2, Q1 = 2, Med = 3, Q3 = 4) pour les groupes en **S-PA**.



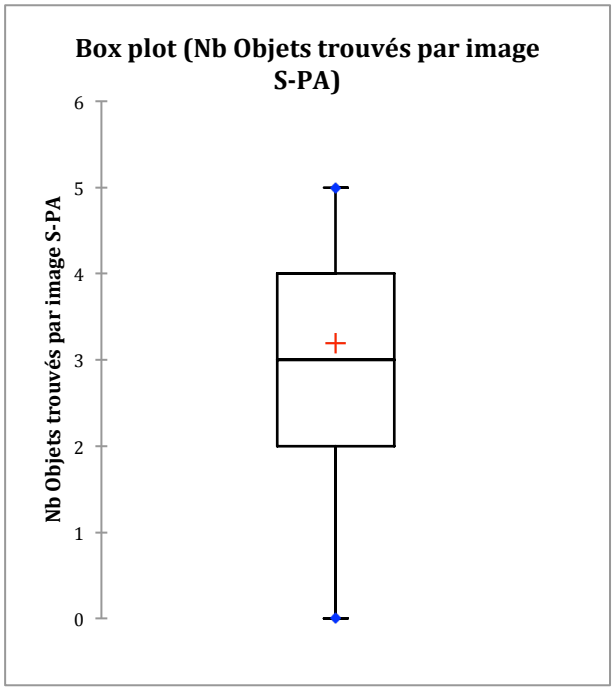


Figure 7 et 8 : Nombre d'objets trouvés par image en fonction de l'ordre de la modalité de démarrage de l'expérience

Les figures 7 et 8 permettent d'observer une performance (en nombre d'objets trouvés) légèrement supérieure pour les groupes en **PA-S,** mais surtout que la performance est moins dispersée.

Concernant l'hypothèse H4, après un test de Mann-Whitney, nous observons une différence significative entre le nombre de mauvais clics entre les groupes ayant réalisé l'expérience en **PA-S** et ceux en **S-PA** ($p < 0,0006$). Nous obtenons une moyenne de 0.5 (σ = 0.9) mauvais clics par image pour les groupes en **PA-S** et 0.9 (σ = 1.4) mauvais clics en moyenne pour les groupes **S-PA**.

4 DISCUSSION ET TRAVAUX FUTURS

Cette expérience préliminaire permet d'appuyer l'hypothèse générale soulignant l'avantage de la construction d'une stratégie performante grâce à la matérialisation du point d'attention. Cette tendance est pour le moment timidement soutenue par nos résultats mais nous sommes encore dans l'analyse de ceux-ci. En effet, les résultats montrent que les groupes ayant eu leurs points d'attention affichés en début d'expérience, et donc la possibilité de prendre en compte l'indication des zones que regardaient leurs partenaires et leur manière de rechercher les images, ont trouvé en

moyenne plus d'objets que les groupes n'ayant eu l'affichage du point d'attention que dans un seconde temps. Ils ont également un meilleur taux de découverte d'objets par image et ils réalisent significativement moins de mauvais clics. L'analyse plus complète de nos résultats nous laisse penser que le mécanisme de « vison flottante » doit être étudié plus largement afin de déterminer si celui-ci est une vraie aide quant à l'utilisation d'écran de grand format et à la recherche d'information sur ceux-ci. D'autre part, le gain en termes de performance est certainement lié à la nature de la tâche, ainsi qu'à l'expertise des utilisateurs. De nombreux travaux ont déjà été menés concernant la capture de l'attention des utilisateurs passant devant un écran (Huang, Koster, & Borchers, 2008; Müller, Walter, Bailly, Nischt, & Alt, 2012 ; Müller, Wilmsmann, Exeler, Buzeck, Schmidt, Jay, & Krüger, 2009 ; Wang, Boring, & Greenberg, 2012). Nous comptons nous appuyer sur ces travaux pour améliorer l'expérience utilisateur liée à l'interaction que nous proposons en matérialisant le point d'attention à l'écran. Dans la suite de nos recherches, nous allons ainsi approfondir le concept de matérialisation du point d'attention dans des tâches de coopération en prenant en compte l'expertise des utilisateurs, leur habitude à coopérer dans la réalisation de tâches comparables et/ou à travailler ensemble.

5 REMERCIEMENTS

Les auteurs remercient la Direction Générale de l'Armement qui finance ce programme de recherche, l'ensemble des volontaires qui on prit part gracieusement à cette expérience, ainsi que Florian TUAL et Yannis MARIE pour leur contribution dans le cadre de leur projet ingénieur.

6 BIBLIOGRAPHIE

Alt, F., Schneegaß, S., Schmidt, A., Müller, J., & Memarovic, N. (2012, June). How to evaluate public displays. In Proceedings of the 2012 International Symposium on Pervasive Displays (p. 17). ACM.

Aw, A., & Rouhet, J. C. (2004). Ergonomie. Système homme-machine.Techniques de l'ingénieur. Informatique industrielle, 2(S7610), S7610-1.

Benali, K., Bourguin, G., David, B., Derycke, A., & Ferraris, C. (2002). Collaboration/Coopération. Information-Interaction-Intelligence, Actes des deuxièmes Assises nationales du GDR I3, J. Le Maître (Ed.), 247-261.

Huang, E. M., Koster, A., & Borchers, J. (2008). Overcoming assumptions and uncovering practices: When does the public really look at public displays?. In Pervasive Computing (pp. 228-243). Springer Berlin Heidelberg.

Khan, A., Matejka, J., Fitzmaurice, G., & Kurtenbach, G. (2005, April). Spotlight: directing users' attention on large displays. In Proceedings of the SIGCHI conference on Human factors in computing systems(pp. 791-798). ACM.

Müller, J., Walter, R., Bailly, G., Nischt, M., & Alt, F. (2012, May). Looking glass: a field study on noticing interactivity of a shop window. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (pp. 297-306). ACM.

Müller, J., Wilmsmann, D., Exeler, J., Buzeck, M., Schmidt, A., Jay, T., & Krüger, A. (2009). Display blindness: The effect of expectations on attention towards digital signage. In Pervasive Computing (pp. 1-8). Springer Berlin Heidelberg.

Peterson, E. R., & Deary, I. J. (2006). Examining wholistic–analytic style using preferences in early information processing. Personality and Individual Differences, 41(1), 3-14.

Schiavo, G., Mencarini, E., Vovard, K., & Zancanaro, M. (2013, April). Sensing and reacting to users' interest: an adaptive public display. In CHI'13 Extended Abstracts on Human Factors in Computing Systems (pp. 1545-1550). ACM.

Wang, M., Boring, S., & Greenberg, S. (2012, June). Proxemic peddler: a public advertising display that captures and preserves the attention of a passerby. In Proceedings of the 2012 international symposium on pervasive displays (p. 3). ACM.