L'anticipation visuospatiale en aéronautique : optimisation de l'interaction Homme système des véhicules autonomes

Marianne Jarry, Colin Blättler, Vincent Ferrari

▶ **To cite this version:**

Marianne Jarry, Colin Blättler, Vincent Ferrari. L'anticipation visuospatiale en aéronautique : optimisation de l'interaction Homme système des véhicules autonomes. 33e conférence internationale francophone sur l'Interaction Humain-Machine (IHM'22), AFIHM, Apr 2022, Namur, Belgique. 10.1145/3500866.3516369 . hal-03847546

HAL Id: hal-03847546
https://hal.science/hal-03847546

Submitted on 17 Nov 2022

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.

L'anticipation visuospatiale en aéronautique : optimisation de l'interaction Homme Système des véhicules autonomes

Visuospatial Anticipation in Aeronautic: Human and Autonomous Vehicle System Interaction Optimisation

Marianne Jarry
marianne.jarry@ecole-air.fr
Centre de Recherche de l'Ecole de l'Air
Salon-de-Provence, France

Colin Blättler
colin.blattler@ecole-air.fr
Centre de Recherche de l'Ecole de l'Air
Salon-de-Provence, France

Vincent Ferrari
vincent.ferrari@ecole-air.fr
Centre de Recherche de l'Ecole de l'Air
Salon-de-Provence, France

ABSTRACT

This study is part of the human-artificial system collaboration optimisation for the development of autonomous vehicles, by optimising a cognitive model of anticipation for use by AI specialists. The research question is whether the spatial representation elicited by auditory-verbal stimulation includes the same properties as the visuospatial anticipation modelled in the Vector Addition model. In the present study, non-pilot and pilot trainees listened to stimuli describing the spatial position of aircraft relative to them. They were then asked to place them on a GPS map. This study shows two new results: (1) the spatial representation developed from audio-verbal information has an anticipatory component and (2) this anticipation is time-bound. This is encouraging to consider a more complete model of anticipation whose formalisation can be used by AI specialists in the design of future artificial agents for autonomous vehicles.

CCS CONCEPTS

• **Human-centered computing → Human computer interaction (HCI)**; **User studies**.

KEYWORDS

Visuospatial Anticipation, Autonomous Vehicles, Human System Interaction, Aviation

RÉSUMÉ

Cette étude s'inscrit dans l'optimisation de la collaboration entre l'humain et le système artificiel pour le développement de véhicules autonomes, en optimisant un modèle cognitif d'anticipation à l'usage des spécialistes de l'IA. La question de recherche est de savoir si la représentation spatiale provoquée par une stimulation auditivoverbale comprend les mêmes propriétés que l'anticipation visuospatiale modélisée dans le *Vector Addition model*. Dans la présente étude des participants non-pilotes et pilotes en formation ont écouté des stimuli décrivant la position spatiale d'aéronefs par rapport à eux. Ils devaient ensuite les placer sur une carte GPS. Cette étude montre deux résultats nouveaux (1) la représentation spatiale développée à partir d'informations auditivoverbales comporte une composante anticipatrice et (2) cette anticipation est limitée dans le temps. Ceci est encourageant pour envisager un modèle plus complet de l'anticipation dont la formalisation pourra être utilisable par les spécialistes de l'IA dans la conception des futurs agents artificiels des véhicules autonomes.

MOTS-CLÉS

Anticipation visuospatiale ; véhicules autonomes ; Interaction Homme-Système ; Aviation

ACM Reference Format:
Marianne Jarry, Colin Blättler, and Vincent Ferrari. 2022. L'anticipation visuospatiale en aéronautique : optimisation de l'interaction Homme Système des véhicules autonomes. In *IHM '22: Proceedings of the 33rd Conference on l'Interaction Humain-Machine (IHM '22), April 5–8, 2022, Namur, Belgium.* ACM, New York, NY, USA, 8 pages. https://doi.org/10.1145/3500866.3516369

1 INTRODUCTION

Cette étude s'inscrit dans l'optimisation de la collaboration entre l'agent humain et l'agent artificiel dans le cadre du développement des véhicules autonomes. En effet, l'apparition des systèmes automatisés dans les cockpits d'avion a provoqué l'apparition de nouvelles erreurs humaines [9, 18, 35, 39]. Modéliser le système cognitif humain peut être une solution pour réduire l'incompatibilité du fonctionnement entre l'agent humain et l'agent artificiel. Cependant, modéliser tous les processus aboutirait à ce que l'agent artificiel reproduise les mêmes erreurs que l'agent humain. C'est pourquoi le choix des processus cognitifs à modéliser doit se focaliser sur les processus les plus essentiels : ici l'anticipation. En effet, [38],[7] et [23] montrent que l'anticipation sert aussi à compenser les moments très fréquents où les flux d'informations sont interrompus. Dans la vie de tous les jours, lorsqu'ils conduisent, les individus sont très régulièrement interrompus par d'autres tâches (e.g., regarder devant soi est interrompu par des rétrocontrôles). Ils ne sont pour autant pas pris de cours par l'évolution de la situation qui s'est produite pendant cette interruption. Bien que les véhicules soient et seront équipés de capteurs et d'algorithmes permettant d'anticiper, les résultats du calcul de la machine et celui de l'individu peuvent être différents, puisque basés sur des algorithmes différents [19, 20, 22, 33]. Ceci peut conduire à des ambiguïtés et générer des effets de surprise, amenant l'individu à réagir de façon

IHM '22, April 5–8, 2022, Namur, Belgium
© 2022 Copyright held by the owner/author(s). Publication rights licensed to ACM. This is the author's version of the work. The definitive version of record was published in IHM '22: Proceedings of the 33rd Conference on l'Interaction Humain-Machine.
ACM ISBN 978-1-4503-9189-4/22/04.
https://doi.org/10.1145/3500866.3516369

inappropriée et donc dangereuse. Ainsi, il est possible de s'attendre à de nouvelles erreurs humaines avec l'autonomisation des véhicules aériens ou terrestres (seul le cas de l'aviation est traité) basée sur des agents artificiels [1].

Dans cette veine, la présente étude considère pertinent de développer un modèle de l'anticipation visuospatiale humaine. En effet, cette anticipation est intégrée aux trois piliers du fonctionnement cognitif humain, l'action, la perception et la cognition. [23, 24, 26]. De plus, aucune action ne peut être réalisée avec succès sans une anticipation ou prévisualisation des évènements à venir [21]. Un des modèles d'anticipation visuospatiale particulièrement adapté est le *vector addition model* [25]. En effet, il prend en compte de très nombreux facteurs qui modulent cette anticipation humaine, ce qui apporte une vue d'ensemble du fonctionnement de l'anticipation suffisamment large pour en envisager une opérationnalisation en situation écologique. Cependant ce modèle et la littérature sur laquelle il repose ne prennent pas en compte les situations dans lesquelles l'individu doit se représenter visuospatialement une information auditivoverbale. En effet, une des activités la plus importante pour assurer la sécurité aérienne est le "voir et éviter". Il s'agit de détecter visuellement des éléments pouvant entrer en collision avec l'avion (oiseaux, montagnes, autres avions). Cette activité est déjà entravée par les cockpits automatisés [6, 32]. Si la détection des menaces se fait régulièrement par contact visuel il arrive souvent que ces informations soient communiquées par les contrôleurs aériens (ou d'autres pilotes dans le cadre de vols militaires) par messages radio. Dans ce cas les membres d'équipage doivent se représenter spatio-temporellement la position de la menace potentielle afin de déterminer s'il y a un risque de collision. Le développement de cette représentation spatiale a-t-elle les mêmes caractéristiques anticipatrices que les représentations spatiales issues de données visuelles ? Cette information est essentielle pour concevoir un modèle d'anticipation utile en milieux opérationnels permettant une collaboration efficace entre les agents humains et artificiels. C'est pourquoi il est nécessaire d'inclure dans le *vector addition model* la manière dont l'anticipation visuospatiale est générée sur la base de stimulations auditivoverbales. L'objectif de cette étude est de vérifier si la représentation visuospatiale élaborée à partir de messages auditivoverbaux inclut une propriété dynamique permettant une anticipation visuospatiale.

2 ETAT DE L'ART

Depuis une quarantaine d'années, l'anticipation visuospatiale est étudiée à partir des études sur le *Representational Momentum* (*RM*) [14]. Le *RM* est la tendance qu'a l'humain à mémoriser la position spatiale d'une cible ou d'une scène en mouvement plus loin, dans le sens du mouvement, qu'elle ne l'est en réalité [14]. Ainsi, sur la base des processus sous-jacent au *RM*, la représentation spatiale élaborée est une anticipation d'un état futur de ce que nous voyons. C'est en raison de son rôle majeur dans l'activité humaine que sa modélisation est pertinente pour développer un agent artificiel intégré aux cockpits ou tableaux de bord des véhicules routiers.

Le *RM* a été observé en laboratoire par [14]. Dans cette expérience trois orientations successives différentes d'un même rectangle induisent un mouvement dans le sens des aiguilles d'une montre ou dans le sens inverse (voir figure 1). Après cette phase d'induction de mouvement, une quatrième orientation du rectangle est présentée. Ce quatrième rectangle peut avoir la même orientation que le troisième ou une différente. Si l'orientation est différente elle peut indiquer soit une poursuite de la rotation, soit une orientation indiquant un début de rotation dans le sens inverse du mouvement induit. La tâche des participants est de décider si le troisième et le quatrième rectangles sont identiques ou non. Les résultats ont montré qu'il était plus difficile de répondre que les rectangles n'étaient pas identiques quand les orientations poursuivaient le mouvement de rotation induit par rapport à celles indiquant un début de rotation dans le sens inverse de celui induit. Le résultat fondamental est que les participants mémorisent la position spatiale d'un objet en mouvement comme étant plus loin dans la direction du mouvement : c'est ce qui traduit l'anticipation. Un second résultat fondamental est apporté par [15] : la propriété dynamique de l'anticipation visuospatiale. Ces auteurs ont repris le même protocole que [14] à l'exception du délai entre la fin de la phase d'induction et la présentation de la quatrième orientation du rectangle. Ces délais appelés intervalles de rétentions (IR) entre 10 ms et 900 ms, ont permis d'observer (voir figure 2) que l'anticipation augmentait entre 10 et 300 ms puis diminuait jusqu'à 900 ms. Pour les auteurs, tout se passe comme-ci la représentation visuospatiale évoluait au cours du temps. Ces résultats ont inspiré l'idée du *momentum metaphor* [12–14] suggérant que les principes physiques sont incorporés dans la représentation mentale. Par exemple, un objet lancé ne s'arrête pas brutalement mais conserve une dynamique pendant encore quelques temps si aucune autre force ne lui est appliquée.

Depuis ces premiers résultats de très nombreuses études ont été conduites (voir [23, 28] pour une revue complète) et ont montré une très grande variété de facteurs qui influencent l'anticipation visuospatiale. Cette grande variété a permis de développer plusieurs modèles dont le *vector addition model* [25] est celui qui explique la plus grande partie des influences connues dans la littérature

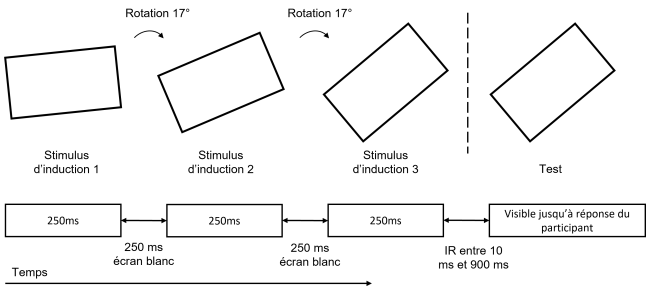

Figure 1: Matériel adapté de [15] inspiré du matériel utilisé par [14]. Le mouvement de rotation du rectangle était d'abord induit par la présentation successive de trois rectangles. Après un intervalle de rétention (IR) les participants devaient répondre si le quatrième rectangle était le même que le troisième. Le quatrième rectangle était, soit le même, soit avait effectué une rotation de +/- 2, 4, 6 ou 8°. Le délai entre la disparition du troisième et l'apparition du quatrième stimulus, l'IR, a été manipulé et variait entre 10ms et 900 ms.

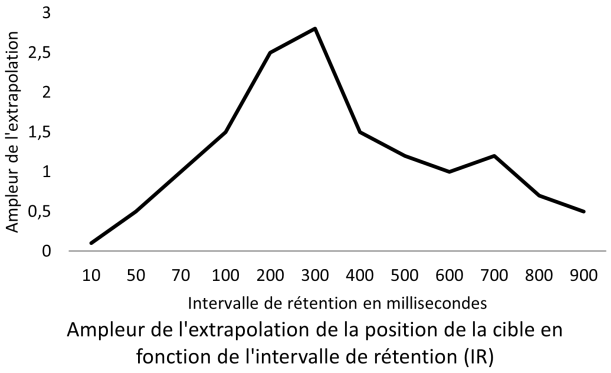

Figure 2: Présentation des résultats adaptés de [15] montrant un effet de l'IR sur l'anticipation de la position de la cible, dans le sens de son mouvement.

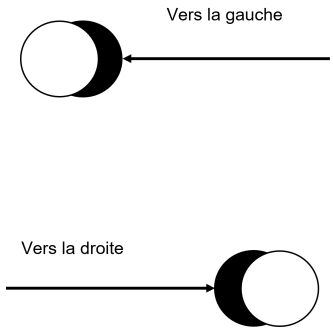

Figure 3: Matériel et résultats adaptés de [29]. Les mouvements des cibles étaient horizontaux de gauche à droite ou de droite à gauche. Les ronds blancs correspondent aux positions de disparition indiquées par les participants en moyenne. Les ronds noirs correspondent aux vrais points de disparition des stimuli.

2.1 Modèle théorique de l'anticipation visuospatiale : *vector addition model*

Le *vector addition model*, considère chaque facteur d'influence comme étant un vecteur influençant l'anticipation visuospatiale dans un sens et une amplitude [25]. En effet, il n'est pas rare d'observer que plusieurs facteurs d'influence interagissent sur l'anticipation visuospatiale. A titre d'exemple, [29] ont montré un effet de la direction du mouvement mais aussi des connaissances implicites de la gravité. Dans cette étude les participants observaient un mouvement continu d'une cible (rond noir). La cible disparaissait de manière impromptue le long de sa trajectoire rectiligne de gauche vers la droite ou de droite vers la gauche. Après la disparition de la cible les participants devaient cliquer, à l'aide d'une souris, à la position à laquelle ils pensaient que la cible avait disparu. Les résultats ont permis d'observer que les participants rappelaient la position des cibles non seulement dans le sens du mouvement de la cible mais aussi vers le bas (figure 3) ce qui suggère l'interaction entre deux facteurs d'influence ou deux vecteurs, selon [25].

Le *vector addition model* (figure 4) offre un avantage : il prend en compte la complexité d'une situation dynamique en intégrant et combinant l'ensemble des facteurs d'influence connus. Cette intégration multifactorielle est un élément clé pour modéliser ce qu'il se passe dans la nature. En effet, les situations naturelles (i.e., hors du laboratoire) combinent systématiquement plusieurs facteurs d'influence. Bien que cet avantage soit indéniable ce modèle n'est pourtant que théorique. En effet, aucune mise à l'épreuve n'a encore été effectuée (ce qui n'est pas l'objectif de cette étude).

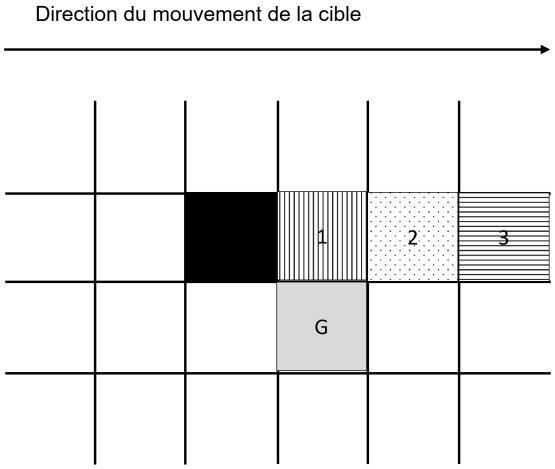

Figure 4: Schéma idéalisé du *vector addition model*. Les carrés 1, 2, 3 et G sont des rappels de la position de la cible en fonction des facteurs d'influence. Le carré noir représente la position de la cible observée lors de sa disparition[29]. La direction du mouvement est indiquée par la flèche. Les carrés 1, 2 et 3 représentent l'amplitude de l'anticipation visuospatiale avec un IR de 50ms, 100ms et 300ms, respectivement. Le carré G représente le résultat de l'anticipation visuospatiale sous la double influence du sens du mouvement de la cible et de la connaissance implicite de la force gravitationnelle.

2.2 Limites du *vector addition model*

En plus de cette absence de mise à l'épreuve, le modèle ne considère aucunement le cas de l'anticipation visuospatiale générée sur la base de stimulations auditivoverbales. C'est cette lacune que vise à combler cette étude. Dans cette veine, la seule étude qui considère cette question est celle de [41]. Ces auteurs montrent que la compréhension du langage implique des simulations perceptives dynamiques. Dans leur expérience, le participant entend une phrase telle que "le lanceur vous envoie la balle" puis un affichage présente une balle, puis un second affichage présente la même balle soit un peu plus grosse (comme si elle avait avancé vers l'observateur) soit un peu plus petite (comme si elle s'était éloignée de l'observateur). Le participant doit donner un jugement de similarité entre les deux affichages de la balle. Les résultats montrent que si le second affichage présentait une balle en concordance avec le message audio (ici la balle avance) alors le participant a plus de difficultés à rejeter cet affichage. Ces résultats sont explicables par le cadre théorique de l'anticipation visuospatiale si l'on considère que les stimulations

auditivoverbales peuvent générer une anticipation visuospatiale comprenant les mêmes propriétés que l'anticipation visuospatiale générée par une stimulation visuelle. La propriété essentielle pour montrer que ce qui est mesuré est bien de l'anticipation visuospatiale est la propriété dynamique des représentations visuospatiales à l'aune de [15]. C'est ce que propose de tester l'étude présentée ici.

2.3 Vue d'ensemble de l'expérience et hypothèses

Dans la présente étude des participants non-pilotes (NP) et pilotes en formation (PF) écoutent des messages radio indiquant la position, la vitesse et le sens de déplacement d'un avion par rapport à eux. Les participants doivent, à la suite de cette écoute, positionner cet avion sur une carte GPS à l'aide du curseur de la souris. Ainsi, ce dispositif expérimental diffère de ceux classiquement employés car l'individu doit transférer une information auditivoverbale, qui est donc décodé sous forme conceptuelle, en une information visuospatiale pour répondre. Cette étude fait varier le délai entre la fin du message audio et le moment où ils peuvent donner leurs réponses : c'est l'IR. Le participant sait qu'il doit tenir compte de ce délai, sans qu'il lui soit explicité que l'IR varie. Les participants pourraient inférer que le mouvement de l'avion ne décline pas avec le temps sous l'influence de ses moteurs. Cependant, même lorsque la cible possède un tel caractère « animé », comme une fusée astronautique, l'anticipation visuospatiale reste limitée spatialement [40]. Considérant les résultats de [40] et ceux de [15] nous nous attendons à ce que la position spatiale rappelée par les participants soit différente selon l'IR. L'ampleur de l'anticipation devrait être la plus importante avec l'IR le plus court de 250 ms, moins importante pour un IR de 1125 ms et encore moins pour l'IR le plus long de 2250 ms.

De plus, comme le modèle doit s'adresser *in fine* à l'interaction avec des individus expérimentés il convient de vérifier si l'effet de l'expérience a aussi un rôle sur l'anticipation visuospatiale provenant de stimulations auditivoverbales. En effet, la littérature indique un effet de l'expérience dans un domaine spécifique qui favorise l'anticipation visuospatiale [3–5, 8]. Si l'influence va dans le même sens dans l'étude présentée ici alors le groupe le plus expérimenté dans ce type de situations anticipera davantage.

3 EXPÉRIENCE

3.1 Participants

32 personnes (28 hommes et 4 femmes) dont 15 non-pilotes (NP) n'ayant aucune expertise en pilotage d'un avion et 17 pilotes en formation (PF) avec en moyenne 54 heures de vol (SD = 56), ont participé à l'expérience. Tous les participants avaient en moyenne 26 ans (SD = 9). Tous les participants avaient une bonne vue ou corrigée (pour les NP) et comprenaient l'anglais.

3.2 Matériel et stimuli

L'expérience était présentée sur un HP 470 G7 Notebook avec un écran de 17,3 pouces de diagonale et une définition de 1920x1080 pixels. Les participants étaient tous assis à une distance de 60 cm de l'écran. Ils écoutaient les messages audios avec des casques Fox Spirit GHS7.1 avec des haut-parleurs de 50mm, les oreillettes permettant d'être isolés des éventuels bruits extérieurs. L'expérience a été codée sur le logiciel E-Prime 3.0. Les données ont été traitées avec le logiciel RStudio. Les fichiers images utilisés ont été créés avec le logiciel de traitement d'image et de graphisme GIMP 2.10.20. Les fichiers audios ont été générés en anglais par un synthétiseur vocal. Les images et messages audio ont été créés en collaboration avec des instructeurs de vol et adaptés pour correspondre au milieu de l'aéronautique militaire mais rester compréhensible pour des personnes sans expérience dans le domaine aéronautique. Les messages audio conservaient la même structure à travers l'expérience. Les indications données étaient toujours dans le même ordre dans un souci de compréhension (tableau 1).

Table 1: Le tableau contient les 4 messages que les participants entendaient durant l'expérience.

N° message Contenu du message
01 Mach 1.2, 0.7 nautical miles, crossing from right to left, in your 5
02 Mach 1.2, 0.7 nautical miles, crossing from left to right, in your 7
03 Mach 1.2, 0.7 nautical miles, crossing from left to right, in your 11
04 Mach 1.2, 0.7 nautical miles, crossing from right to left, in your 1

Les fichiers images créés correspondaient, de manière très épurée, au type de présentation existante sur les systèmes de navigation des avions (e.g., Garmin 1000). Les participants étaient représentés par un plot blanc (stimulus en forme de triangle). L'image de ce stimulus mesurait 0.5*0.5 centimètres et était placée au centre de l'image. Un cercle blanc en pointillés de 10 centimètres de diamètre apparaissait autour du plot représentant le participant pour leur permettre de connaître l'échelle spatiale utilisée, 1cm = 0.1 nautical miles (NM) (figure 5).

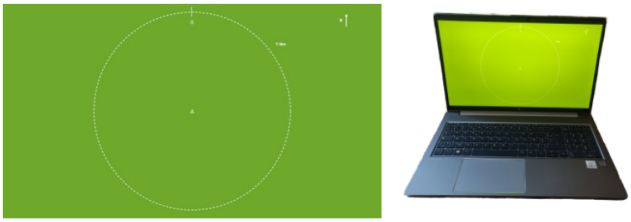

Figure 5: Image présentée après un IR variable. Le cercle représente un périmètre de 1 NM autour du plot blanc. Le cap nord est représenté par le trait et le 0 en haut du cercle. La flèche en haut à droite représente l'orientation de la configuration spatiale de l'image (toujours le nord vers le haut). Cette image était présentée pendant la phase d'exposition des consignes.

3.3 Procédure

Les participants étaient assis devant un ordinateur portable à 60 cm de l'écran. Le contexte communiqué dans l'explication des consignes les plaçait dans un repère spatial dans lequel ils devaient s'imaginer dans un SR20 (avion utilisé à l'escadron d'instruction en vol de la base aérienne de Salon-de-Provence) volant à 120 nœuds soit 220 km/h environ. Autour de l'avion des participants, représenté par un plot blanc, était affiché un cercle de 10 cm de rayon représentant un périmètre de 1 NM (1.8 km environ). Ce cercle indiquait quelle échelle était utilisée. Sur ce cercle était représenté le nord (cap 0). Il était précisé aux participants que le cap nord serait conservé tout le long de l'expérience. Ensuite la structure des messages était détaillée et expliquée pour permettre à des NP de comprendre ce qu'ils allaient entendre. Les participants étaient alors prévenus qu'ils allaient d'abord entendre la vitesse de déplacement de l'aéronef décrit (1.2 Mach soit 1480 km/h), la distance à laquelle il se trouvait par rapport à eux (0.7 nautiques), le sens du mouvement de l'aéronef (de gauche à droite ou de droite à gauche) puis la position au cadran par rapport à eux (dans leur 5 heures, 7 heures, 11 heures ou 1 heure). Les participants devaient à la suite de l'écoute du message positionner l'avion décrit sur une carte GPS à l'aide du curseur de la souris en prenant en compte le délai entre la fin du message et l'apparition de l'image. Avant de commencer l'expérience, les participants effectuaient une phase de familiarisation au cours de laquelle 2 messages par IR (égal à 250 ms, 1125 ms ou 2250 ms) étaient présentés aléatoirement. Chaque participant effectuait alors 6 essais de familiarisation. Un rappel des consignes était présenté puis l'expérience commençait.

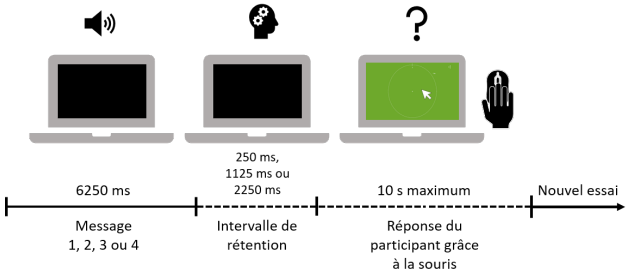

Figure 6: Déroulement d'un essai au cours duquel un message audio de 6250ms était présenté, puis un écran noir de 250, 1125 ou 2250 ms était affiché et enfin l'image était affichée pour que les participants puissent cliquer à l'endroit auquel ils pensaient que l'aéronef se trouvait.

L'expérience se composait de 132 essais comprenant toutes les conditions et tous les messages (11 fois * 4 messages * 3 IR). Un essai se déroulait comme décrit dans la figure 6. Un des sons était d'abord présenté, tous duraient 6250 ms. Après un IR variable une image était affichée. Ils devaient alors cliquer, à l'aide du curseur de la souris, à la position à laquelle ils pensaient que l'aéronef décrit était au moment d'apparition de l'image. Ils devaient répondre à l'aide de la souris et du clic gauche. Les participants avaient au maximum 10 secondes pour répondre, sans réponse au bout de ce délai un autre message était présenté. L'IR pouvait être de 250 ms, 1125 ms ou 2250 ms de manière aléatoire à travers les essais. De cette manière les conditions d'IR apparaissaient le même nombre de fois à travers les essais pour les 4 messages.

3.4 Résultats

Les coordonnées des positions rappelées par les participants sont enregistrées en pixels sur l'axe x et sur l'axe y. Les coordonnées rappelées sont comparées aux coordonnées de références, i.e., la position en pixels x et y exacte en fonctions de l'IR. Par exemple pour le message 1 pour lequel l'avion va de droite à gauche, la mesure pour chaque participant est : M_{diffx} = x référence - x rappelé et M_{diffy} = y référence - y rappelé. Pour les messages dans lesquels l'avion va de gauche à droite le calcul pour M_{diffx} = x rappelé - x référence. Ainsi pour chaque participant est obtenue une mesure de M_{diffx} et de M_{diffy} par conditions d'IR. Ainsi le participant est en avance si M_{diffx} est positif, en retard si M_{diffx} est négatif ou n'a pas de décalage si M_{diffx} n'est pas différent de zéro.

3.4.1 Axe Vertical (y). Une ANOVA à mesures répétées a été réalisée sur les M_{diffy} avec la variable Groupe (Pilotes en formation, PF ; Non Pilotes, NP) comme facteur inter-groupe et la variable IR (250 ms ; 1125 ms ; 2250 ms) comme facteur intra-groupe. Cette analyse ne révèle ni d'effet de Groupe, $F(1, 30) < 1$; $p = .8$, ni d'effet de l'IR $F(1, 30) < 1$; $p = .6$, ni d'interaction $F(2, 60) < 1$; $p = .3$. Des T-test ont été réalisés pour chaque condition d'IR, sans tenir compte des Groupes, pour vérifier si un décalage est présent. Pour 250 ms, 1125ms et 2250 ms aucun T-test n'est significatif, $t(31) < 1$; $p = .5$ (M_{diffy} = -2 px ; SD = 23) ; $t(31) < 1$; $p = .5$ (M_{diffy} = -2 px ; SD = 26) ; et $t(31) < 1$; $p = .3$ (M_{diffy} = -8 px ; SD = 51) respectivement. Ainsi, il n'est pas observé de biais de réponse sur l'axe y.

3.4.2 Axe Horizontal (x). Une ANOVA à mesures répétées a été effectuée sur les M_{diffx} avec la variable Groupe (PF ; NP) comme facteur inter-groupe et la variable IR (250 ms ; 1125 ms ; 2250 ms) comme facteur intra-groupe (voir figure 7). Aucun effet de Groupe n'a été observé ni d'interaction, $F(1, 30) < 1$; $p = .5$, $F(2,60) < 1$; $p = .6$, respectivement. Un effet significatif de l'IR a été observé, $F(2,60) = 37.845$; $p < .001$. Des tests post-hoc de Tukey ont été réalisés. Les post-hoc significatifs sont les suivant : PF 250 vs PF 2250, $p < .0001$; PF 1125 vs PF 2250, $p = .002$; NP 250 vs NP 2250, $p < .0005$; NP 1125 vs NP 2250, $p < .012$. Pour vérifier si les participants sont significativement en avance, en retard ou non, par rapport à zéro, des T-tests pour échantillons uniques ont été réalisés pour chaque condition d'IR, sans tenir compte des Groupes. Pour 250 ms et 2250 ms les résultats sont significatifs, $t(31) = 2.871$; $p < .01$ (M_{diffx}= + 57 px ; SD = 113), $t(31) = -7.816$; $p < .001$, (M_{diffx}= -110 px ; SD = 80) respectivement. Pour 1125 ms le T-test n'est pas significatif, $t(31) <1$; $p = 0.5$, (M_{diffx}= + 16 px, SD = 141).

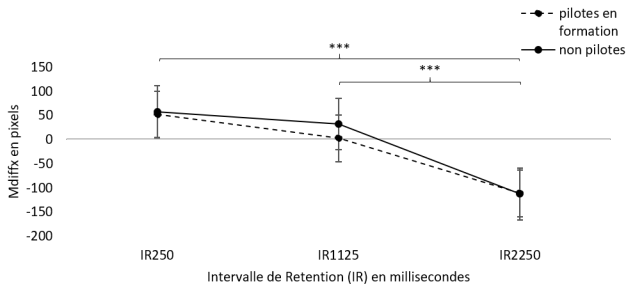

Figure 7: M_{diffx} pour chaque IR (250 ms, 1125 ms et 2250 ms) pour les deux groupes de participants PF et NP. Les barres d'erreurs représentent les erreurs standards.

4 DISCUSSION

L'objectif de cette étude est de vérifier si la représentation visuospatiale élaborée à partir de messages auditivoverbaux inclut une propriété dynamique permettant une anticipation visuospatiale. Une évaluation de l'anticipation visuospatiale a été réalisée en utilisant la méthode d'étude du *Representational Momentum* [14, 29] adaptée avec celle de [41]. Les participants, pilotes en formation (PF) et non pilotes (NP), écoutaient des messages radios indiquant la position d'un avion par rapport à eux, sa direction et sa vitesse. Ils devaient ensuite positionner sur une carte GPS la position de cet avion selon trois IR. Un objectif subsidiaire était de vérifier si l'expérience des PF est suffisante dans ce contexte pour améliorer l'anticipation visuospatiale [3–5, 8]. Aucun effet de groupe n'a été observé. En effet, aucune différence significative n'a été observée entre les positions rappelées par les PF et NP. Cette absence de différence peut être due au fait que les PF avaient peu d'heures de vol (Moyenne des heures de vol = 54 ; *SD* = 56). Dans les précédentes études sur l'effet de l'expertise les différences entre les groupes étaient beaucoup plus marquées. En effet le groupe d'expert cumulait en moyenne 3193 heures de vol [3] ou au moins 10 ans de pratique de conduite [4, 5]. En restant dans le domaine de l'aéronautique il serait pertinent de refaire la même expérience avec des individus cumulant davantage d'heures de vols.

Les résultats montrent qu'en moyenne la position spatiale rappelée par les participants est différente selon l'IR. Pour la condition IR 250 ms, la M_{diffx} est positive, ce qui veut dire que les participants sont en avance : dans cette condition une anticipation est présente. Pour la condition IR 1125 ms, la M_{diffx} n'est pas différente de zéro, ce qui veut dire que les participants ne sont ni en avance, ni en retard. Pour la condition IR 2250 ms, la M_{diffx} est négative, ce qui veut dire que les participants sont en retard. De plus, la position rappelée par les participants est significativement différente entre 250 ms et 2250 ms ainsi que entre 1125 ms et 2250 ms. Pris ensemble, ces résultats sont en accord avec l'idée selon laquelle la représentation visuospatiale élaborée à partir d'informations auditivoverbales comporte une propriété dynamique. En effet, la représentation visuospatiale de la situation élaborée évolue avec le temps. Bien que seule la condition de 250 ms permette d'observer de l'anticipation visuospatiale les résultats dans les deux autres conditions montrent une réduction de ce que les auteurs appellent l'extrapolation [15, 27]. Selon [27] l'extrapolation est un processus qui permet de générer des informations qui n'existent pas encore sur la base de celles qui ont été perçues.

Selon l'approche computationnelle de [34] un processus sert à résoudre un problème. Le problème que l'anticipation visuospatiale peut résoudre est de deux ordres. A la fois il permettrait de compenser les délais neuronaux [7] mais aussi de compenser les moments très fréquents où le flux d'informations est interrompu. Par exemple, en conduite automobile, lorsque le conducteur effectue un rétrocontrôle il perd momentanément le flux d'informations de la route devant lui. Il n'est pourtant pas étonné, quand il revient en contact visuel avec la route, que tout ait évolué. La propriété dynamique de la représentation visuospatiale servirait ainsi à faire perdurer, pendant quelques temps, la dynamique de la situation en cours alors que le flux d'information est interrompu. [15] montrent que ce temps peut aller jusqu'à 900 ms. Dans la présente étude, il est montré que cette compensation peut aller jusqu'à plus d'une seconde (1125 ms) et qu'au contraire à 2250 ms le système n'est plus capable de compenser. C'est le deuxième résultat majeur de cette étude puisque la limite de la compensation de l'anticipation visuospatiale chez l'être humain est montrée. Ainsi dans la conception de futures IA cette limite doit être prise en compte. En effet, en aviation, qui est une situation multitâche dynamique à forte pression temporelle [2, 16, 17], les membres d'équipage sont souvent interrompus par diverses tâches. Ainsi une interruption de plus de deux secondes est trop longue et potentiellement dangereuse. L'approfondissement des recherches sur la limite temporelle de l'anticipation visuospatiale pourrait faire évoluer ou confirmer les prescriptions de l'Aeronautical Information Manual de la Federal Aviation Administration (AIM/FAA). En effet, l'AIM/FAA prescrit que le temps que les pilotes passent sur une tâche visuelle dans la cabine de l'avion ne doit pas représenter plus de 1/4 à 1/3 du temps passé à regarder à l'extérieur de la cabine, ce qui représente en moyenne 2 secondes dans la cabine pour 6 secondes à l'extérieur [11].

5 CONCLUSION

L'étude présentée ici fait partie d'un projet plus vaste qui vise à modéliser l'anticipation visuospatiale humaine, dans le but d'optimiser la collaboration entre les agents artificiels et humains au sein des futurs véhicules autonomes. Dans cet optique le modèle *vector addition* de [25] est un bon candidat mais reste incomplet. Jusqu'à aujourd'hui ce modèle n'intégrait pas les situations dans lesquelles une information auditivoverbale était représentée visuospatialement. Ainsi le *vector addition model* intégrant cette nouvelle dimension devient encore plus séduisant pour en envisager une formalisation à l'usage des spécialistes de l'IA. En effet, ce modèle bien que théorique peut servir de base pour le développement d'un modèle prédictif de l'anticipation humaine. Dans ce sens il conviendrait de déterminer quel est le meilleur modèle d'apprentissage machine pour ce cas de figure. Par exemple [10] parviennent en utilisant un modèle en réseau de neurones à reproduire certains résultats d'anticipation visuospatiale obtenus dans la littérature sur des participants humains. Les avancées plus récentes dans le domaine de l'apprentissage machine sauraient certainement répondre à ces objectifs.

Une des limites de l'approche présentée ici est de ne se concentrer que sur l'anticipation visuospatiale. En effet, l'anticipation visuospatiale est limitée comme nous venons de le montrer, à une très courte période temporelle. Le modèle d'anticipation qui doit être développé devrait intégrer des processus d'anticipation (1) qui se font sur de plus longues périodes temporelles et (2) non uniquement liées à la dimension spatiotemporelle. [26, 27] aborde cette question en décrivant d'autres processus d'anticipation. Par exemple le *behavioral momentum* [36, 37] ainsi que le *psychological momentum* [30, 31] sont des processus d'anticipation s'étalant sur des périodes temporelles beaucoup plus longues (des mois voire des années). L'élément intéressant que propose [27] est que ces processus reposent sur un processus commun : l'extrapolation. Il reste cependant à intégrer ces deux processus d'anticipation au *vector addition model* afin d'aboutir à un modèle générique de l'anticipation humaine. En effet, les résultats obtenus dans cette étude montrent que la composante anticipatrice n'est pas présente uniquement dans les processus perceptifs mais aussi à un niveau d'abstraction plus élevée, c'est-à-dire, plus central du système cognitif. L'obtention d'un tel modèle générique permettrait de développer des solutions adaptées aux problèmes rencontrés dans différents contextes dans lesquels les interactions entre agents humains et artificiels sont à optimiser (e.g., aéronautique, automobile).

RÉFÉRENCES

[1] Wayne Biever, Linda Angell, and Sean Seaman. 2020. Automated Driving System Collisions : Early Lessons. *Human Factors* 62, 2 (2020), 249–259. https://doi.org/10.1177/0018720819872034

[2] Colin Blättler, Ludovic Fabre, and Gregory Froger. 2021. Display on Demand Method Increases Time Spent Looking Outside the Cockpit. *Aviation Psychology and Applied Human Factors* 11, 1 (2021), 54–58. https://doi.org/10.1027/2192-0923/a000201

[3] Colin Blättler, Vincent Ferrari, André Didierjean, and Evelyne Marmèche. 2011. Representational Momentum in Aviation. *Journal of Experimental Psychology : Human Perception and Performance* 37, 5 (2011), 1569–1577. https://doi.org/10.1037/a0023512

[4] Colin Blättler, Vincent Ferrari, André Didierjean, and Evelyne Marmèche. 2012. Role of expertise and action in motion extrapolation from real road scenes. *Visual Cognition* 20, 8 (2012), 988–1001. https://doi.org/10.1080/13506285.2012.716799

[5] Colin Blättler, Vincent Ferrari, André Didierjean, Pierre van Elslande, and Evelyne Marmèche. 2010. Can expertise modulate representational momentum ? *Visual Cognition* 18, 9 (2010), 1253–1273. https://doi.org/10.1080/13506281003737119

[6] Stephen M. Casner. 2006. Mitigating the loss of navigational awareness while flying with GPS and moving map displays under VFR. *International Journal of Applied Aviation Studies* 6, 1 (2006), 121–130.

[7] Russell L. De Valois and Karen K. De Valois. 1991. Vernier acuity with stationary moving Gabors. *Vision Research* 31, 9 (jan 1991), 1619–1626. https://doi.org/10.1016/0042-6989(91)90138-U

[8] André Didierjean, Vincent Ferrari, and Colin Blättler. 2014. Role of Knowledge in Motion Extrapolation : The Relevance of an Approach Contrasting Experts and Novices. *Psychology of Learning and Motivation - Advances in Research and Theory* 61 (2014), 215–235. https://doi.org/10.1016/B978-0-12-800283-4.00006-X

[9] M. R. Endsley and E. O. Kiris. 1995. The out-of-the-loop performance problem and level of control in automation. *Human Factors* 37, 2 (1995), 381–394. https://doi.org/10.1518/001872095779064555

[10] Wolfram Erlhagen and Dirk Jancke. 2004. The role of action plans and other cognitive factors in motion extrapolation : A modelling study. *Visual Cognition* 11, 2-3 (2004), 315–340. https://doi.org/10.1080/13506280344000293

[11] Aviation Administration Federal. 2017. Aeronautical information manual : Official guide to basic flight information and ATC procedures. *Washington, DC : U.S. Department of Transportation.* (2017).

[12] Ronald A. Finke and Jennifer J. Freyd. 1985. Transformations of Visual Memory Induced by Implied Motions of Pattern Elements. *Journal of Experimental Psychology : Learning, Memory, and Cognition* 11, 4 (1985), 780–794. https://doi.org/10.1037/0278-7393.11.1-4.780

[13] Ronald A. Finke, Jennifer J. Freyd, and Gary C.W. Shyi. 1986. Implied Velocity and Acceleration Induce Transformations of Visual Memory. *Journal of Experimental Psychology : General* 115, 2 (1986), 175–188. https://doi.org/10.1037/0096-3445.115.2.175

[14] Jennifer J. Freyd and Ronald A. Finke. 1984. Representational momentum. *Journal of Experimental Psychology : Learning, Memory, and Cognition* 10, 1 (1984), 126–132. https://doi.org/10.1037/0278-7393.10.1.126

[15] Jennifer J. Freyd and J. Q. Johnson. 1987. Probing the Time Course of Representational Momentum. *Journal of Experimental Psychology : Learning, Memory, and Cognition* 13, 2 (1987), 259–268. https://doi.org/10.1037/0278-7393.13.2.259

[16] Grégory Froger, Colin Blatter, Emilien Dubois, Cyril Camachon, and Nathalie Bonnardel. 2017. L'apprentissage d'un séquençage visuo-attentionnel pour lutter contre la distraction en contexte aéronautique de double tâche. (2017). https://hal.archives-ouvertes.fr/hal-01778727

[17] Grégory Froger, Colin Blättler, Emilien Dubois, Cyril Camachon, and Nathalie Bonnardel. 2018. Time-Interval Emphasis in an Aeronautical Dual-Task Context : A Countermeasure to Task Absorption. *Human Factors* 60, 7 (2018), 936–946. https://doi.org/10.1177/0018720818783946

[18] Ken Funk, Fenth Lyall, Jennifer Wilson, Rebekah Vint, and Niemczyk Mary. 2009. The International Journal of Aviation Psychology : Preface. *International Journal of Aviation Psychology* 19, 1 (2009), 1–7. https://doi.org/10.1080/10508410802597382

[19] Guy Hoffman and Cynthia Breazeal. 2007. Effects of anticipatory action on human-robot teamwork efficiency, fluency, and perception of team. *HRI 2007 - Proceedings of the 2007 ACM/IEEE Conference on Human-Robot Interaction - Robot as Team Member* (2007), 1–8. https://doi.org/10.1145/1228716.1228718

[20] Guy Hoffman and Cynthia Breazeal. 2010. Effects of anticipatory perceptual simulation on practiced human-robot tasks. *Autonomous Robots* 28, 4 (2010), 403–423. https://doi.org/10.1007/s10514-009-9166-3

[21] Bernhard Hommel, Jochen Müsseler, Gisa Aschersleben, and Wolfgang Prinz. 2001. The theory of event coding (TEC) : A framework for perception and action planning. *Behavioral and brain sciences* 24, 5 (2001), 849–878.

[22] Chien-ming Huang. [n.d.]. Huang2016Hri. Section V ([n. d.]).

[23] Timothy L. Hubbard. 2005. Representational momentum and related displacements in spatial memory : A review of the findings. *Psychonomic Bulletin and Review* 12, 5 (2005), 822–851. https://doi.org/10.3758/BF03196775

[24] Timothy L. Hubbard. 2006. Bridging the gap : Possible roles and contributions of representational momentum. *Psicologica* 27, 1 (2006), 1–34.

[25] Timothy L. Hubbard. 2010. Approaches to representational momentum : Theories and models. *Space and Time in Perception and Action* June (2010), 338–365. https://doi.org/10.1017/CBO9780511750540.020

[26] Timothy L. Hubbard. 2015. The varieties of momentum-like experience. *Psychological Bulletin* 141, 6 (2015), 1081–1119. https://doi.org/10.1037/bul0000016

[27] Timothy L. Hubbard. 2019. Momentum-like effects and the dynamics of perception, cognition, and action. *Attention, Perception, and Psychophysics* 81, 7 (2019), 2155–2170. https://doi.org/10.3758/s13414-019-01770-z

[28] Timothy L Hubbard. 2020. Representational gravity : Empirical findings and theoretical implications. *Psychonomic bulletin & review* 27, 1 (2020), 36–55.

[29] Timothy L. Hubbard and Jamshed J. Bharucha. 1988. Judged displacement in apparent vertical and horizontal motion. *Perception & Psychophysics* 44, 3 (1988), 211–221. https://doi.org/10.3758/BF03206290

[30] Seppo E. Iso-Ahola and Charles O. Dotson. 2014. Psychological momentum : Why success breeds success. *Review of General Psychology* 18, 1 (2014), 19–33. https://doi.org/10.1037/a0036406

[31] Seppo E. Iso-Ahola and Ken Mobily. 1970. "Psychological Momentum" : a Phenomenon a N D a N Empirical (Unobtrusive) Validation of Its Influence I N a Competitive Sport Tournament. (1970).

[32] Nicholas Johnson, Douglas Wiegmann, and Christopher Wickens. 2006. Effects of advanced cockpit displays on general aviation pilots' decisions to continue visual flight rules flight into instrument meteorological conditions. *Proceedings of the Human Factors and Ergonomics Society* (2006), 30–34. https://doi.org/10.1177/154193120605000107

[33] Harmanpreet Kaur and Alex C Williams. 2019. Building Shared Mental Models between Humans and AI for Effective Collaboration. *Chi* May 2019 (2019). https://michae.lv/ai-hci-workshop/#accepted-papers

[34] David Marr. [n.d.]. *Vision : A computational investigation into the human representation and processing of visual information.*

[35] National Transportation Safety Board. 2010. Introduction of Glass Cockpit Avionics into Light Aircraft. (2010), 91. http://app.ntsb.gov/doclib/safetystudies/SS1001.pdf

[36] John A. Nevin. 2015. Behavioral Momentum : A Scientific Metaphor. *Cognitive Science* (2015).

[37] John A. Nevin, Charlotte Mandell, and Jean R. Atak. 1983. the Analysis of Behavioral Momentum. *Journal of the Experimental Analysis of Behavior* 39, 1 (1983), 49–59. https://doi.org/10.1901/jeab.1983.39-49

[38] Romi Nijhawan. [n.d.]. Motion extrapolation un catching. ([n. d.]).

[39] Raja Parasuraman and Victor Riley. 1997. Humans and automation : Use, misuse, disuse, abuse. *Human Factors* 39, 2 (1997), 230–253. https://doi.org/10.1518/001872097778543886

[40] Norman G Vinson and Catherine L Reed. 2002. Sources of object-specific effects in representational momentum. *Visual Cognition* 9, 1-2 (2002), 41–65.

[41] Rolf A. Zwaan, Carol J. Madden, Richard H. Yaxley, and Mark E. Aveyard. 2004. Moving words : Dynamic representations in language comprehension. *Cognitive Science* 28, 4 (2004), 611–619. https://doi.org/10.1016/j.cogsci.2004.03.004