



Université
de Toulouse

THÈSE

En vue de l'obtention du

DOCTORAT DE L'UNIVERSITÉ DE TOULOUSE

Délivré par :

Institut Supérieur de l'Aéronautique et de l'Espace (ISAE)

Présentée et soutenue par :

Jean-Paul IMBERT

le mardi 16 décembre 2014

Titre :

Adaptation du design des visualisations de type supervisions pour optimiser la transmission des notifications classées par niveau d'intérêt

École doctorale et discipline ou spécialité :

ED AA : Informatique et Neurosciences, comportement et cognition

Unité de recherche :

Équipe d'accueil ISAE-ONERA CSDV

Directeur(s) de Thèse :

M. Frédéric DEHAIS (directeur de thèse)

M. Christophe HURTER (co-directeur de thèse)

Jury :

M. Julien CEGARRA - Président du Jury

M. Cyril CAMACHON

M. Philippe CABON - Rapporteur

M. Gilles COPPIN - Rapporteur

M. Frédéric DEHAIS - Directeur de thèse

M. Christophe HURTER - Co-directeur de thèse

M. Sébastien TREMBLAY

IMBERT JP. (2014). ADAPTATION DU DESIGN DES VISUALISATIONS DE TYPE SUPERVISIONS POUR OPTIMISER LA TRANSMISSION DES NOTIFICATIONS CLASSÉES PAR NIVEAU D'INTÉRÊT . Thèse de Doctorat en Informatique, Institut Supérieur de l'Aéronautique et de l'Espace.

Avant-propos et remerciements

Ces travaux de thèse sont l'aboutissement d'une démarche commencée il y a plus de dix ans quand je suis arrivé dans l'équipe PII (Performance et Innovation des IHM) d'Eric Blond du Centre d'Etudes de la Navigation Aérienne (CENA) devenu plus tard le centre de R&D de la Direction Technique de l'Innovation de la Direction des Services de la Navigation Aérienne (DTI/DSNA). Cette équipe multidisciplinaire comprenant ergonomes, graphistes et ingénieurs de recherche, spécialisée dans le domaine de l'interface homme-machine a toujours accueilli en son sein des chercheurs et doctorants provenant de divers domaines comme l'informatique ou l'ergonomie.

Mes activités en tant qu'ingénieur de recherche à PII m'ont amené à participer à de nombreux projets impliquant d'autres équipes du CENA comme architecte intégrateur de composants et monter au cours des années plusieurs plateformes d'expérimentations pour des chercheurs ou doctorants de l'équipe PII. Tout ceci n'aurait pas été possible sans l'excellence et la disponibilité des différents membres de l'équipe ; qu'ils en soient ici globalement remerciés.

Parmi les projets auxquels j'ai eu à collaborer, je citerai ceux qui m'ont donné l'envie de m'intéresser à la fois à des questions de recherche et me poser des questions sur le facteur humain et la problématique de la conception des interfaces homme-machine. Plusieurs collaborations avec les équipes FH (Facteur Humain) de Sylvie Figarol et SAS de Francis Casaux (spécialistes entre autres du TCAS) et expérimentations pendant lesquelles durant des simulations nous placions contrôleurs et pilotes dans des situations de collisions imminentes. Ces expérimentations avaient pour but d'essayer de comprendre les problèmes liés à des représentations mentales différentes dans des situations d'urgence et mesurer le stress associé, elles m'ont amené à m'intéresser au facteur humain et aux moyens de mesure physiologique permettant d'obtenir des indicateurs sur le comportement des opérateurs durant l'activité. Par la suite, grâce à une collaboration avec Philippe Averty, docteur en psychologie dans l'équipe PII, j'ai pu être impliqué dans le projet Creed dont l'objet était d'arriver à estimer la perception du risque de conflit chez les contrôleurs aériens. Plusieurs expérimentations dans différents aérodromes ont été menées dans le cadre de ce projet et m'ont permis d'approfondir mes connaissances métier ainsi que des aspects concernant la perception des informations et la projection de leur état futur sur la visualisation radar nécessaires à une bonne représentation mentale des situations de conflits. À la suite de ce projet, dans le cadre de la thèse de Caroline Martin sur la gestion de la charge mentale des contrôleurs aériens et grâce à l'implication d'autres membres de l'équipe PII nous avons pu franchir une nouvelle étape dans la prise en compte de l'humain dans le processus d'évaluation des interfaces en intégrant un eye tracker à une plateforme de simulation d'étude pour le contrôle aérien. Ces différents projets intégrant le facteur humain ainsi que les différents travaux auxquels j'ai participé pour faire évoluer les outils des contrôleurs aériens m'ont amené à une démarche scientifique plus globale concernant la conception et l'évaluation des interfaces homme-machine.

Ce travail de recherche est l'aboutissement de cette démarche. Il s'est déroulé dans un contexte particulier puisque lors de la réorganisation du centre de R&D en 2012 et le transfert des chercheurs à l'École Nationale de l'Aviation Civile j'ai dû prendre un nouvel emploi comme ingénieur de recherche au sein de l'équipe SDR (Support et Développement pour la Recherche) de Bruno Kriner à l'ENAC. Sans son support, celui de son équipe et le temps qu'ils m'ont permis de dégager pour poursuivre ce travail je n'aurais pu le terminer.

Pour avoir collaboré avec des chercheurs ou des doctorants pendant plusieurs années, je savais qu'une thèse ne se faisait pas seul, je n'imaginai pas à quel point en réalité. Je tiens ici à remercier les différentes personnes qui ont participé à ce projet ou tout simplement m'ont encouragé ou guidé durant ces quatre années :

- Christophe Hurter, co-encadrant de thèse en informatique dont la soutenance de thèse m'a inspiré ce sujet, pour l'énergie semble-t-il inépuisable qu'il peut déverser autour de lui et les encouragements à publier pendant ces quelques années
- Frédéric Dehais, directeur de thèse qui a cru dès le départ à ce travail dans un domaine connexe au sien qui est celui du cockpit et les multiples collaborations très enrichissantes qu'il m'a permis de développer avec les membres de son équipe Neuroergonomie et Facteurs Humains de l'ISAE (Institut Supérieur de l'Aéronautique et de l'Espace) et le laboratoire CoDot de l'Université Laval
- Sébastien Tremblay directeur du Laboratoire CoDot de l'Université Laval, Québec m'a guidé dans le processus difficile de la rédaction d'articles scientifiques et relancé à un moment important pour moi dans ce travail, merci aussi pour m'avoir invité dans ton laboratoire à Québec
- Robert Parise, chercheur en ergonomie, malgré son activité à 100% dédiée à des prestations de service en ergonomie pour la DTI a toujours su trouver le temps de m'en consacrer
- Eric Blond et Bruno Kriner, mes deux chefs durant ces quatre années qui m'ont toujours encouragé et permis d'avancer et gérer les pics de travail associés à mon poste d'ingénieur de recherche à PII ou SDR
- Jean-Marc Alliot directeur du domaine R&D qui m'a encouragé à démarrer cette thèse
- Philippe Palanque professeur à l'université de Toulouse et directeur du Laboratoire ICS qui m'a permis de commencer cette thèse
- Daniel Etienne, Philippe Ribet, François-Régis Colin, Anthony Marion, Alexandre Bustico, Gwenaël Bothorel, François Marais Nicolas Courtel, ingénieurs de recherche avec qui j'ai eu le plaisir de travailler et faire évoluer les outils qui m'ont permis de réaliser ce travail. Sans vous, rien n'aurait pu se faire
- Philippe Averty, Sylvie Figarol, Christelle Pianetti, Jean-Louis Garcia pour m'avoir impliqué dans vos projets et amené à élargir mon horizon professionnel
- Mickaël Causse professeur à l'ISAE et Louise Giraudet doctorante de Mr Causse avec qui j'ai pu collaborer sur la perception des notifications sonores dans la continuité de mes premiers travaux de thèse. Grâce à

vous j'ai découvert un nouveau champ de recherche passionnant, la neuroergonomie

- David Bole Richard, Jean Beidner, Rémi Lesbordes et Géraldine Bebox, contrôleurs aériens et détachés comme experts au centre de R&D dont je faisais partie. Sans eux, la mise au point, le réglage de l'agent de suivi de la tâche réalisé dans le cadre de ces travaux n'aurait pas été possible. Un grand merci à David pour m'avoir aidé et participé à la réalisation de la dernière expérimentation présentée dans ces travaux
- Helen Hodgetts chercheuse à Cardiff Metropolitan University, Royaume-Uni et chercheuse associée du Laboratoire CoDot de l'Université Laval, Québec pour sa contribution aux articles scientifiques issus de ma première expérimentation
- François Vachon professeur au laboratoire CoDot de l'Université Laval, Québec pour sa contribution aux articles scientifiques issus de ma première expérimentation
- Nadine Matton et Ludovic d'Estampes enseignants chercheurs à l'ENAC, pour leur contribution aux analyses statistiques de ma dernière expérimentation
- Bertina Hô-Mock-Hay et l'équipe du projet SESAR 4.7.2 dans lequel j'ai pu insérer ces travaux. Sans votre support, je n'aurais pas pu réaliser la toute dernière expérimentation de cette thèse qui nécessitait des moyens dépassant de très loin ceux que j'aurais pu mettre en place pour évaluer l'agent de suivi de la tâche et le système de notifications associé
- Jean-Luc Vinot enseignant chercheur à l'ENAC et graphiste de formation pour m'avoir inspiré et ouvert sur le monde du graphisme en tant qu'art et sur l'interaction homme-machine. Merci aussi d'avoir commencé ta thèse à un âge encore plus avancé que le mien, ça encourage ☺ et pour ta passion communicative de l'IHM
- Serge Roux et David Paquet, mes collègues de l'équipe SDR pour avoir pris le relais à maintes reprises lorsque je devais travailler sur la rédaction d'un article ou la rédaction de ce manuscrit
- Tous les sujets, contrôleurs aériens ou ingénieurs dans le centre de R&D de la DTI qui ont eu la gentillesse de m'accorder de leur temps pour passer sur mes expérimentations
- Mon épouse et mes deux filles pour leur patience pendant cette période de travail plutôt intense où je menais plusieurs activités de front

Table des matières

Introduction.....	13
--------------------------	-----------

PARTIE I : **CONTEXTE THÉORIQUE**

1. LE SYSTÈME DE GESTION DU CONTRÔLE AÉRIEN ET SES ÉVOLUTIONS	19
--	-----------

1.1. Présentation du système de contrôle aérien.....	19
1.1.1. Organisation du système de contrôle aérien.....	19
1.1.2. Les systèmes d'aide aujourd'hui	23
1.2. Évolutions futures	25
1.3. Erreurs humaines dans le domaine aéronautique.....	28
1.4. Résumé	31

2. CONTRAINTES ET OBJECTIFS DES VISUALISATIONS DE TYPE SUPERVISION	33
---	-----------

2.1. Introduction.....	33
2.2. Perception des informations dans les systèmes de visualisation	34
2.2.1. Variables visuelles.....	34
2.2.2. Perception avec faible charge cognitive.....	36
2.2.3. Approche écologique de la perception et design écologique	37
2.2.4. Design orienté conscience de la situation	40
2.3. Les systèmes de notifications visuelles et sonores, et leurs limites	42
2.3.1. Notifications visuelles.....	42
2.3.2. Notifications sonores.....	44
2.4. Résumé	47

3. AUTOMATISATION ET ADAPTATION	49
3.1. Introduction	49
3.2. Niveaux d'automatisation	49
3.3. Impact de l'automatisation sur l'humain.....	50
3.4. Confiance dans le système	52
3.5. Automatisation et conscience de la situation	53
3.6. Interrompre à quel prix ?	54
3.7. Modalités d'initiative et automatisation adaptative	55
3.8. Résumé	58
4. STRATÉGIE MÉTHODOLOGIQUE	61
4.1. De l'étude de laboratoire à l'opérationnel	61
4.1.1. Etudes de laboratoire	61
4.1.2. Tests de laboratoire appliqués au contexte : les micro-mondes	63
4.1.2.1. Un micro monde aéronautique en contexte multitâche : MATB	66
4.1.2.2. Un micro monde ATC pour l'évaluation du travail collaboratif : CTEAM....	66
4.1.2.3. Un micro-monde pour l'ATC : Laby	67
4.1.3. Environnements de simulations complexes	72
4.2. Évaluation de l'utilisation d'une interface homme-machine.....	72
4.2.1. Utilisation de méthodes d'évaluation subjectives.....	73
4.2.2. Mesures comportementales.....	74
4.2.3. Mesures physiologiques	75
4.2.3.1. Mesures oculométriques	76
4.2.3.2. Mesures neurophysiologiques	81
4.2.3.2.1. Méthode continue	81
4.2.3.2.2. Méthode ERP (Evoked Related Potential).....	83
4.3. Résumé	85
Problématique de la thèse	87

PARTIE II :

DÉVELOPPEMENTS ET VALIDATION

5. CONCEPTION D'UN DESIGN ORIENTÉ CONSCIENCE DE LA SITUATION	93
<hr/>	
5.1. Introduction	93
5.2. Évaluation du design de notifications visuelles	94
5.2.1. Évolutions de Laby vers un environnement de test multitâche.....	94
5.2.2. Conception du labyrinthe de l'expérimentation Saillance.....	95
5.2.3. Types de notifications évaluées	98
5.2.4. Architecture de la plateforme Laby, intégration d'un oculomètre.....	99
5.3. Confrontation de notifications visuelles à des notifications sonores	104
5.4. Réalisation d'un agent de suivi de la tâche de supervision des contrôleurs aériens : SAmonitoring	107
5.4.1. Suivi des clairances de niveau de vol	110
5.4.2. Suivi des consignes de vitesse.....	113
5.4.3. Suivi des consignes de directe	114
5.4.4. Suivi des états de coordination et des états fréquences	114
5.4.5. Suivi du niveau d'entrée dans le secteur	116
5.4.6. Suivi des niveaux de transfert en sortie.....	116
5.5. Intégration de SAmonitoring dans un simulateur d'étude.....	117
5.5.1. Architecture Droïd et intégration d'un dispositif oculométrique	120
5.5.2. Intégration des données de SAmonitoring dans la supervision radar.....	122
5.5.3. Retour d'expérience de l'intégration d'oculomètre dans un simulateur d'étude.....	125
6. ÉVALUATION DE DESIGNS DE NOTIFICATIONS VISUELLES : L'EXPÉRIMENTATION SAILLANCE	127
<hr/>	
6.1. Présentation.....	127
6.2. Protocole.....	127

TABLE DES MATIÈRES

6.2.1. Participants.....	127
6.2.2. Procédure de passation	127
6.2.3. Hypothèses de travail	128
6.2.4. Variables	128
6.3. Résultats	128
6.3.1. Réalisation de la tâche de guidage	129
6.3.2. Réalisation de la tâche de validation des notifications.....	130
6.3.3. Résultats oculométriques.....	131
6.3.4. Résultats qualitatifs.....	133
6.4. Conclusion.....	135
6.4.1. Analyse des résultats.....	135
6.4.2. Discussion et applications.....	137
7. ÉVALUATION DE DESIGNS VISUELS CONFRONTÉS À DES NOTIFICATIONS SONORES : L'EXPÉRIMENTATION BIP	141
<hr/>	
7.1. Présentation.....	141
7.2. Protocole.....	141
7.2.1. Participants.....	141
7.2.2. Procédure de passation	142
7.2.3. Hypothèses de travail	143
7.2.4. Variables	144
7.3. Résultats	144
7.3.1. Réalisation de la tâche de guidage	145
7.3.2. Réalisation de la tâche de validation des notifications.....	146
7.3.3. Réalisation de la tâche auditive.....	147
7.3.4. Résultats physiologiques, analyse de la P300.....	149
7.3.5. Résultats qualitatifs.....	150
7.4. Conclusion.....	151
7.4.1. Analyse des résultats.....	151

7.4.2. Discussion et applications.....	152
8. ÉVALUATION D'UNE VISUALISATION ORIENTÉE CONSCIENCE DE LA SITUATION : LES EXPÉRIMENTATIONS MONITORING & WP4.7.2	155
8.1. Présentation.....	155
8.2. Protocole.....	157
8.2.1. Participants.....	157
8.2.2. Hypothèses de travail	158
8.2.3. Variables	158
8.2.4. Procédure de passation expérimentation Monitoring.....	159
8.2.5. Procédure de passation expérimentation WP4.7.2	160
8.3. Résultats	161
8.3.1. Résultats quantitatifs.....	161
8.3.1.1. Expérimentation Monitoring.....	161
8.3.1.2. Expérimentation WP4.7.2 sans oculométrie.....	164
8.3.2. Résultats qualitatifs.....	165
8.4. Conclusion.....	167
9. CONCLUSION	169
9.1. Bilan et synthèse des résultats	169
9.2. Perspectives de recherche	172
9.2.1. Utilisation dynamique de la détection de l'état cognitif : le projet NINA.....	172
9.2.2. Extension aux notifications haptiques.....	173
9.2.3. Évaluation de nouveaux designs de notifications.....	174
9.3. Contributions	174
9.3.1. Contributions scientifiques.....	174
9.3.2. Réalisations.....	175

Références Bibliographiques	177
Table des illustrations	199
Annexes	202
Conception d'une tâche expérimentale avec Laby.....	203
Questionnaire Monitoring Condition : Sans monitoring.....	206
Questionnaire Monitoring Condition : Avec monitoring	207
Questionnaire alertes visuelles	209

Introduction

Les systèmes utilisés pour réaliser le service de contrôle aérien évoluent relativement lentement ; néanmoins, il y a une quinzaine d'années, le numérique et les réseaux ont fait leur entrée dans ces systèmes. L'interconnexion entre les différents centres de contrôle en France et en Europe est devenue possible, permettant d'échanger des informations concernant la gestion des vols de façon dynamique. Par ailleurs, l'évolution des moyens informatiques a permis le développement de nouveaux outils basés sur ces données dans le but de permettre une augmentation de la capacité des secteurs aériens tout en préservant un niveau de sécurité élevé. Ces outils se situent à différents niveaux d'automatisation, allant de la modification dynamique de l'affichage des informations sur la supervision radar à des outils de détection des conflits et d'aide à la décision.

Dans le futur, les systèmes de gestion du trafic aérien vont évoluer vers toujours plus d'automatisation pour aider les humains dans leur tâche jusqu'à prendre en charge une partie de leur activité. Le projet SESAR (Single European Sky ATM Research - ATM : Air Traffic Management) a pour objectif de moderniser les systèmes existants dans le but de diviser par deux les coûts du contrôle aérien, diminuer de 10% l'impact sur l'environnement, diviser le risque d'accident par 10 et tripler la capacité de l'espace aérien. Pour réaliser ces objectifs, un niveau élevé d'automatisation devra être atteint (SESAR : The ATM Target Concept 2007) tout en maintenant l'humain au centre du système ATM qui jouera le rôle de superviseur et de preneur de décision. Même si cet objectif est particulièrement ambitieux, un certain nombre d'avancées technologiques permettent d'ores et déjà d'améliorer plusieurs de ces critères de façon significative.

Dans le cas du contrôle aérien comme dans celui de tous les systèmes critiques supervisés, la perception des informations à acquérir est cruciale dans le déroulement de la tâche des superviseurs et dans leur prise de décision. Ceux-ci doivent percevoir et mémoriser un grand nombre d'informations dynamiques et interdépendantes ; parmi elles, des notifications (avertissements ou alertes) qui peuvent être délivrées par des systèmes d'aide. La saillance des informations de type notifications (sonores, visuelles ou haptiques) qui peuvent avoir des niveaux de criticité différents doit être équilibrée par rapport aux objectifs de supervision de manière à ne pas interrompre le superviseur trop souvent ou à mauvais escient. Cet aspect est particulièrement critique dans le domaine aéronautique où contrôleurs et pilotes sont soumis au stress, à une forte pression temporelle et aux aspects sécurité de leur tâche. Les interfaces de visualisation de ces supervisions et la qualité de leur couplage avec l'humain et ses capacités de perception sont essentiels pour la bonne réalisation de l'activité de supervision. Le domaine aéronautique a atteint un niveau élevé de sécurité de nos jours ; néanmoins, les différentes études concernant l'analyse des accidents ces dernières années montrent que les erreurs humaines représentent les deux tiers de ceux-ci. Dans ces erreurs humaines, les problèmes de perception et de mémorisation en sont la cause principale ; il reste donc fort à faire dans le domaine de la visualisation pour améliorer le design et l'optimiser par rapport aux capacités limitées de l'humain.

Pour anticiper les problèmes liés à une évolution vers plus d'automatisation dans le contrôle aérien, des leçons peuvent être tirées de plusieurs décennies d'intégration d'automates dans l'aviation. Le développement de l'automatisation a démarré dans le cockpit dans les années 1960 ; à cette époque en raison de la complexité grandissante des appareils et de l'augmentation du trafic, l'automatisation a été développée pour pallier entre autres aux limites de l'humain en termes de capacité d'attention et de ressources. L'introduction de l'automatisation dans le cockpit a abouti à une redéfinition du rôle du pilote dans lequel une nouvelle tâche de supervision et de commande des automates l'aidant dans sa tâche est devenue prépondérante. Les objectifs initiaux de l'automatisation dans l'aviation ont été accomplis, performance et sécurité sont aujourd'hui à un niveau élevé. Toutefois, le couplage entre l'humain et l'automate s'est révélé au cours du temps beaucoup moins simple qu'attendu et les toutes dernières évolutions de ces systèmes avec le glass-cockpit et les FMS (Flight Management System) ont entraîné des changements importants dans la méthode de travail des pilotes, aboutissant à de nouveaux types d'incidents liés à une mauvaise perception de la situation ou de l'état des automates. Les analyses de ces incidents, aussi bien dans le domaine de l'aviation que dans d'autres domaines de supervision de systèmes automatisés et critiques, montrent l'importance du couplage humain-automate(s), et des méthodes de design s'appuyant sur le facteur humain ont émergé ces dernières années comme le design écologique ou les interfaces adaptatives pour essayer de l'améliorer.

Le contrôle aérien entre désormais dans une période d'évolution dans laquelle l'automatisation jouera un rôle considérable. Aussi un des objectifs du projet SESAR est d'étudier l'impact de l'arrivée de ces systèmes automatisés sur le rôle des contrôleurs aériens ainsi que sur leur conscience de la situation. En lien avec cette problématique majeure, nous nous sommes intéressés particulièrement à la perception des informations dynamiques nécessaires à la réalisation des différents objectifs liés à la tâche de contrôle aérien ainsi qu'à la façon de les représenter et de les notifier pour qu'elles soient perçues dans des délais conformes à leur importance pour le contrôleur. Ce sujet est central dans la mesure où l'automatisation va permettre au contrôleur de gérer d'avantages d'avions.

Ainsi, le premier objectif de ces travaux consistait à mesurer l'effet de plusieurs designs de notifications sur la performance du contrôleur. Un enjeu a été de développer un environnement de laboratoire contrôlé mais proche écologique. Nous nous sommes donc appuyés sur un micro-monde de contrôle aérien (Laby), développé dans le centre de R&D de la navigation aérienne qui offre des caractéristiques proches de l'activité tout en étant facilement modifiable et léger à mettre en œuvre. En complément des mesures comportementales d'évaluation de la performance offertes par Laby, nous avons utilisé des capteurs physiologiques afin de mesurer lors de l'activité, l'impact de la perception des notifications pendant son déroulement.

Le second objectif a porté sur la mise en évidence des informations dynamiques nécessaires à l'établissement de la conscience de la situation des contrôleurs. Une analyse de l'activité avec des experts du domaine a permis de mettre en évidence les informations que les contrôleurs doivent détecter pendant leur processus de surveillance. Cette analyse nous a permis de réaliser un automate visant à s'assurer de la réalisation des objectifs associés à la tâche. Cet agent, basé sur l'analyse de la situation aérienne, des actions

des contrôleurs et des règles associées au secteur de contrôle, a pour objet de vérifier en temps réel l'exécution d'un certain nombre d'objectifs de la tâche de contrôle et, le cas échéant, de notifier le contrôleur de façon adaptée lorsque ceux-ci ne sont pas atteints. Les résultats des expérimentations précédentes sur les notifications ont été utilisés lors du développement d'une nouvelle visualisation de supervision montrant les informations détectées par l'agent de suivi de la tâche. Cette position de contrôle a été utilisée dans le cadre d'une expérimentation concernant le projet SESAR et tout particulièrement le work package 4.7.2 qui concerne les nouveaux systèmes pour le contrôle en-route.

Le travail de recherche présenté dans ce document est composé de deux parties. La première décrit le contexte théorique, la seconde présente les développements réalisés ainsi que le processus de validation expérimentale mis en œuvre et les résultats obtenus lors de trois expérimentations. Entre ces deux parties, la problématique décrite précédemment est abordée de façon plus complète en tenant compte des éléments mis en évidence dans le contexte théorique.

La première partie de cette étude est découpée en quatre chapitres. Le premier chapitre décrit le domaine du contrôle aérien qui sert de support à cette thèse ainsi que les enjeux liés à son évolution. Le second a pour but d'expliciter les contraintes et objectifs des visualisations de type supervision qui sont au centre de ces travaux en s'intéressant tout d'abord à la problématique de la perception des informations dans les visualisations ainsi qu'aux modes de transmission des notifications et alertes en provenance des systèmes. Le troisième chapitre aborde la question de l'automatisation, du couplage humain-automate et des solutions mises en œuvre dans la littérature pour améliorer la transmission d'informations à l'humain. Enfin, le quatrième chapitre aborde la démarche expérimentale d'évaluation des interfaces ainsi que les méthodes d'évaluation de la performance utilisant des capteurs physiologiques que nous avons mis en œuvre dans le cadre de ces travaux.

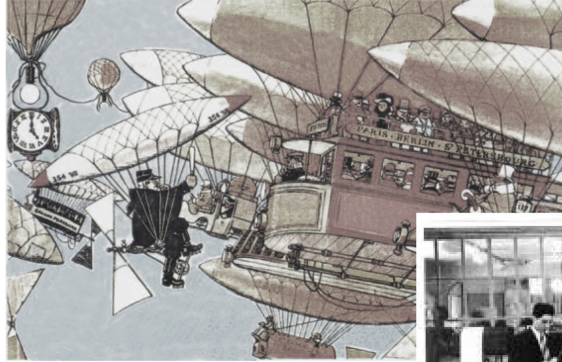
La seconde partie de ce document est composée de quatre chapitres. Le premier, détaille le travail réalisé pour aboutir à une nouvelle position de contrôle dont le design est orienté conscience de la situation. Il décrit les choix méthodologiques faits et les développements réalisés pour concevoir trois expérimentations permettant d'une part d'évaluer plusieurs designs de notification et d'autre part d'évaluer la nouvelle position de contrôle conçue pendant ces travaux. Les deux chapitres suivants portent sur l'évaluation de notifications visuelles ou sonores dans un contexte de double tâche à l'aide d'un micro-monde ATC (Laby). La première expérimentation a pour but de mesurer la saillance de notifications visuelles réalisées avec différents designs. Lors de cette étude, un capteur oculométrique a été utilisé pour la mesure de la performance des designs. La seconde, reprenant deux de ces designs a permis de les confronter à des notifications sonores et analyser leur impact sur la charge de travail à l'aide de mesures électroencéphalographiques. Dans le quatrième et dernier chapitre de la partie expérimentale, des travaux, portant sur la nouvelle position de travail de contrôle aérien dont le design a été réalisé pour améliorer la conscience de la situation des contrôleurs sont présentés ainsi que les résultats préliminaires d'une expérimentation mettant en œuvre cette nouvelle architecture. Pour finir, nous

conclurons sur les résultats obtenus lors de ces travaux et dégagerons de nouvelles perspectives de recherche.

Notre contribution porte sur trois aspects :

- nous avons proposé une méthode d'évaluation de notifications dans le contexte du contrôle aérien et pu mettre en évidence les performances relatives de plusieurs designs utilisant des propriétés différentes à l'aide d'un micro-monde ATC (Laby). Ce micro-monde a été modifié pour prendre en compte ce besoin d'évaluation et permettre l'intégration de capteurs physiologiques comme outils de mesure de la performance.
- nous avons réalisé une nouvelle position de contrôle dont le design a été conçu pour améliorer la conscience de la situation des contrôleurs en s'appuyant sur un agent de suivi de la tâche développé pendant ces travaux ainsi qu'un système de notification optimisé pour prendre en compte la criticité de l'information déterminée par l'agent. Ce système de notification, intégré dans la supervision radar, prend en compte les résultats obtenus dans la phase d'évaluation utilisant le micro-monde.
- la dernière contribution, d'ordre plus général, porte sur l'intégration de capteurs physiologiques dans des architectures systèmes complexes de façon à pouvoir les intégrer comme outils de mesure ou bien dans le but d'utiliser dynamiquement les données physiologiques dans les systèmes. Ce dernier usage permet d'adapter l'interface homme-machine en détectant par exemple le stress ou la fatigue.

PARTIE I : CONTEXTE THÉORIQUE



1. LE SYSTÈME DE GESTION DU CONTRÔLE AÉRIEN ET SES ÉVOLUTIONS

Dans ce chapitre, nous présentons le système de contrôle aérien actuel en France, les systèmes d'aides et d'alerte déjà en place, les erreurs observées dans le contrôle aérien et les évolutions prévues pour le faire progresser vers les objectifs du projet SESAR.

1.1. Présentation du système de contrôle aérien actuel

1.1.1. Organisation du système de contrôle aérien

L'espace aérien contrôlé est découpé en zones gérées par des centres de contrôle en-route pour l'espace supérieur (altitude supérieure à 5800m), de contrôle d'approche pour l'espace inférieur et du contrôle d'aérodrome (Figure 1: du contrôle en-route au contrôle d'aérodrome). Le contrôle d'aérodrome assure les phases de départ et d'arrivée des vols incluant le roulage au sol. Le contrôle d'approche gère la transition entre l'espace aérien supérieur et le contrôle d'aérodrome (phases de montée ou de descente), il assure le séquençage des avions avant l'intégration dans la plateforme aéroportuaire. Le contrôle en route gère les vols dans l'espace supérieur pour des vols se déplaçant à vitesses élevées. La France est découpée en cinq zones pour le trafic en-route (Figure 1: n°1). Le trafic d'approche est assuré par les grandes plateformes aéroportuaires Françaises (Figure 1: n°2). L'espace aérien est parcouru de routes aériennes, sorte d'autoroutes définies par des séries de points à différents niveaux de vol. Le niveau de vol ou FL (Flight-Level), est l'unité d'altitude utilisée dans l'aéronautique. Il vaut 100 pieds soit 30 mètres, les avions sont positionnés à des niveaux séparés de 10 FL soit 1000 pieds (les niveaux finissent toujours par 0, c.-à-d. : 350FL = 35000 pieds soit environ 12km).

Chaque centre de contrôle est découpé en sous-zones appelées secteurs gérées par un binôme de contrôleurs ; un secteur gère une vingtaine d'avions au maximum. Un équipage va donc être en contact avec des contrôleurs différents en fonction de la zone d'espace aérien traversée pendant le déroulement du vol. Dans le binôme de contrôleurs, le contrôleur radariste est en communication radio avec les avions présents dans son secteur. Il supervise le trafic et donne les éventuels ordres nécessaires pour assurer les différents objectifs de sa tâche. Le contrôleur organique a une vision plus en amont du secteur géré par le radariste, il prépare l'entrée des vols dans le secteur traversé en faisant l'intégration des vols et en détectant d'éventuels conflits à venir. Il assure les coordinations (téléphoniques) avec les secteurs adjacents et peut résoudre les conflits en amont en demandant des modifications de trajectoires aux contrôleurs de secteurs adjacents.

Les missions des contrôleurs, telles que définies par ordre de priorité dans la réglementation de la circulation aérienne sont :

- d'assurer la sécurité des aéronefs en maintenant des normes de séparation minimales (5 miles nautiques en horizontal en en-route, 3 Nm en approche, 1000 ft ou 10FL en vertical)
- d'optimiser dans la mesure du possible les trajectoires pour limiter l'impact sur l'environnement (consommation de carburant, trajectoires anti-bruit)
- d'informer les équipages sur leurs conditions de navigation (météo, trafic, régulations)

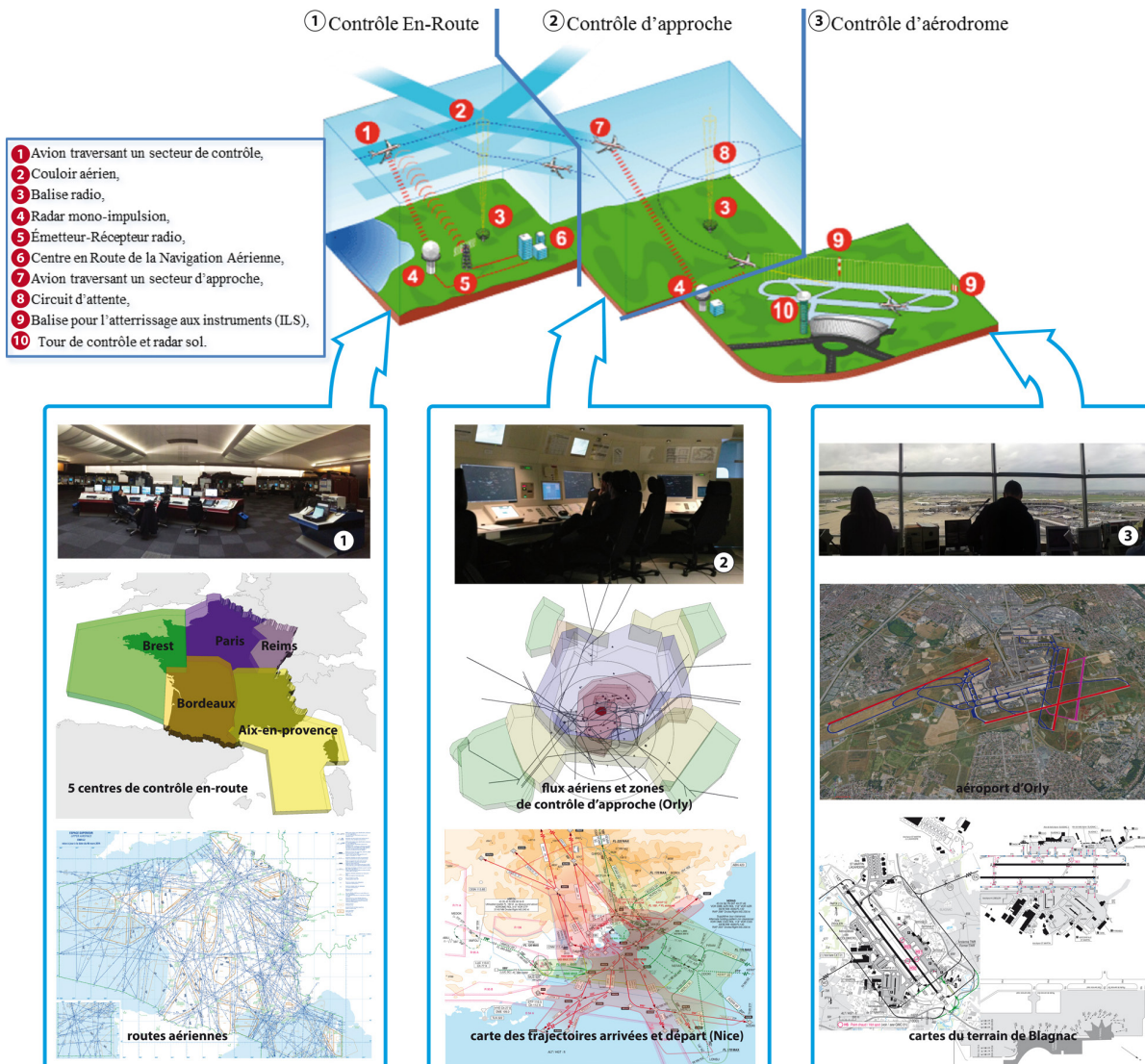


Figure 1: du contrôle en-route au contrôle d'aérodrome

Dans la réalisation de sa tâche, le contrôleur aérien poursuit un grand nombre d'objectifs qui peuvent être classifiés par ordre d'importance ; Endsley (1994) dans une étude sur l'activité des contrôleurs en-route aux États-Unis liste l'ensemble de ces objectifs. Les objectifs et leur ordre de priorité peuvent évoluer en fonction de la situation aérienne ; le maintien des séparations entre aéronefs étant prioritaire, d'autres aspects comme la consommation des aéronefs, les desiderata des compagnies aériennes ou des pilotes sont aussi pris en compte par les contrôleurs dans la mesure du pos-

sible ; le contrôleur doit faire le tri en temps réel entre toutes ces contraintes pour réaliser au mieux sa tâche. La prise d'information lui permet de faire des choix et donner des ordres aux pilotes compte tenu de ces différents objectifs. Le suivi et la mémorisation des informations courantes des avions en charge est une part importante de leur activité.

La figure 2 montre l'évolution des positions de contrôle aérien des années 50 à nos jours. L'acquisition des informations nécessaires à la réalisation de la tâche de contrôle se fait à l'aide de deux outils, la visualisation de supervision radar sur laquelle sont représentés les avions et les informations cartographiques associées au secteur de contrôle, et le tableau de strips papier qui donne sur chaque strip tous les éléments pertinents au secteur concernant le plan de vol pour un aéronef (n° 2 et 3 sur la figure 3). ODS (Operational Display System), la position opérationnelle actuelle dispose d'un écran de 28" carrés (résolution 2000x2000) ou d'un écran 30" 16/9 (résolution 2560x1600) permettant l'affichage d'un grand nombre d'informations. Le plan de vol représenté sur le strip papier est déposé par la compagnie et décrit la trajectoire que doit suivre l'appareil en partant d'un aéroport A pour se rendre à un aéroport B en suivant des procédures de départ et d'arrivée et en passant par des points identifiés nommés « waypoints » se trouvant sur des routes aériennes (voir la carte de routes aériennes sur la figure 1).

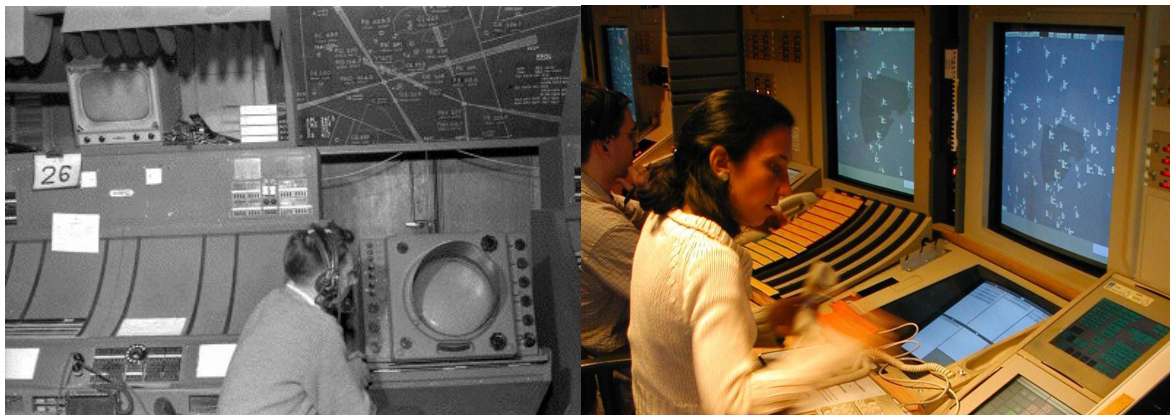


Figure 2: visualisation radar opérationnelle des années 50 à nos jours

La figure 3 illustre les points communs entre la représentation sol et la représentation des informations dans le cockpit. La vision du trafic du contrôleur est géographique dans un repère cartésien centré sur son secteur de contrôle. Les vignettes 1, 4 et 5 montrent cette visualisation jusqu'au zoom sur un avion particulier AF126EK qui vole à 380 knots (environ 700 km/h) de vitesse sol et stable au niveau de vol 350 (l'altitude dans un avion est exprimée en fonction de la pression atmosphérique par rapport à la pression au niveau de la mer dans l'atmosphère standard soit 1013 hPa). Tout l'espace aérien est découpé en secteurs, un secteur de contrôle est un volume d'espace aérien défini par un polygone et un niveau minimum et maximum. Son volume est optimisé en fonction des conditions de trafic dans la zone. L'étiquette radar montre les positions passées de l'appareil (la comète radar), la position courante, le vecteur vitesse ainsi que l'indicatif de vol (ici AF126EK) ; les informations concernant le trafic sont mises à jour toutes les 8 secondes. La vue bord du trafic dans le cockpit est centrée avion et orientée en fonction de la direction du vol (voir vignette 6), les pilotes n'ont aucun détail sur ces autres vols hormis la différence d'altitude et leur distance par rapport à leur appareil. Sur les vignettes 2 et 3 on peut voir

le tableau de strips et le zoom sur le strip de l'AF126EK ; la représentation bord de ce strip est un sous-ensemble des points du plan de vol de l'appareil (les points du plan de vol situés dans le secteur) présenté sur le MCDU (Multifunctional Control and Display Unit, vignette 8). Le MCDU avec le FCU (Flight Control Unit en haut du tableau de bord) sont les deux outils permettant aux pilotes d'interagir avec le système de gestion du vol FMS (Flight Management System).

Contrôleurs et Pilotes: 2 métiers étroitement liés

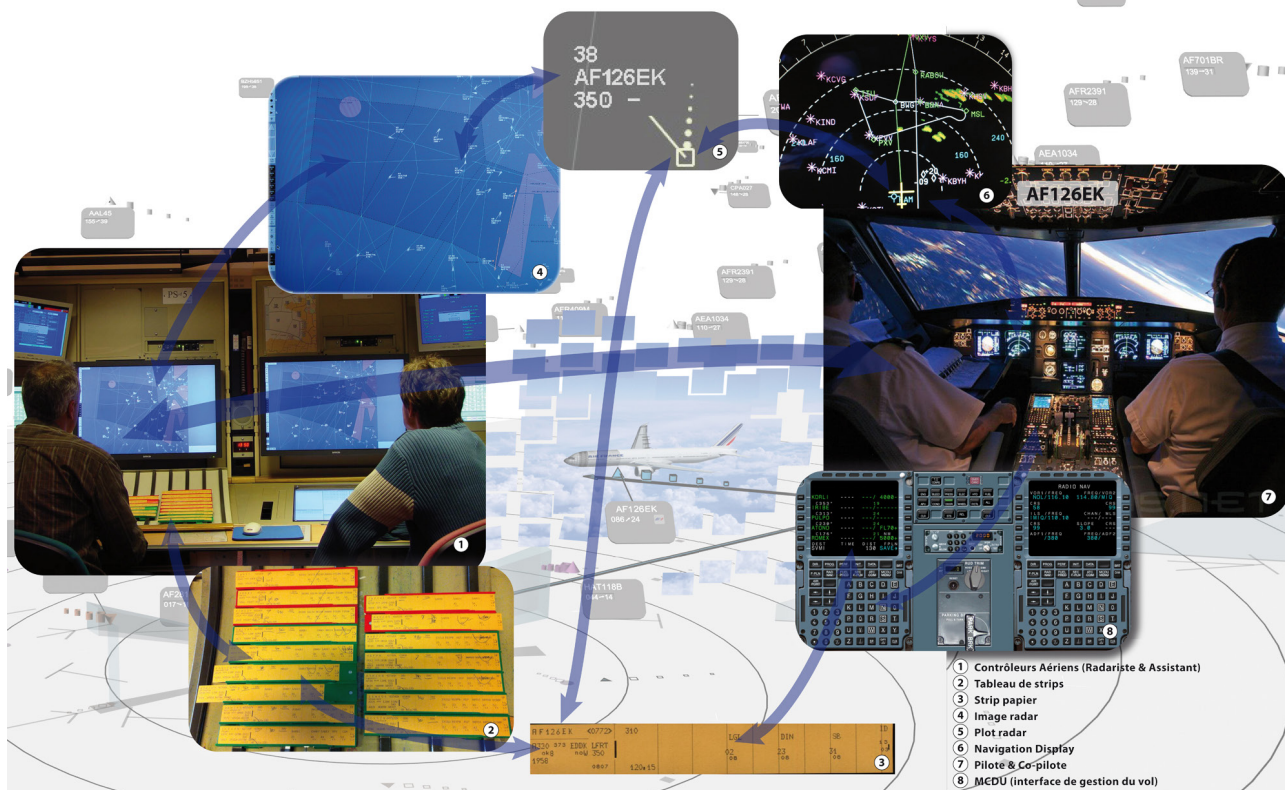


Figure 3: contrôleurs et pilotes ; deux métiers étroitement liés

Le tableau de strips papier (n° 2 sur la figure 3) est encore généralisé en France. Ce n'est pas le cas dans un nombre important de pays qui sont déjà passés à un environnement tout électronique. La disposition et l'orientation des strips sur le tableau, les porte strips utilisés (supports en plastique de couleur) sont autant de moyens pour les contrôleurs d'organiser leur trafic par exemple par conflit ou par point de sortie du secteur. Sur ce strip, le contrôleur radariste dispose d'informations aussi disponibles sur la supervision radar, mais moins accessibles. Il y écrit toutes les consignes données pendant le passage du vol dans le secteur (par exemple un changement d'altitude un ordre de cap etc..). Un strip est imprimé une quinzaine de minutes avant son entrée dans le secteur. Il est pris en compte par le contrôleur organique qui va éventuellement réaliser une coordination avec un secteur adjacent en cas de conflit patent ou noter des informations pertinentes pour le radariste ; à son entrée dans le secteur celui-ci récupère le strip et le place sur son tableau en fonction de son choix de placement (géographique ou autre).

Les informations écrites par les contrôleurs sur ces strips pourraient alimenter de nouvelles fonctions comme par exemple détecter les conflits entre

avions à moyen terme. Elles sont pour l'instant perdues pour le système de gestion du contrôle aérien. À terme de nouvelles interfaces en environnement électronique (sans strip papier) permettront de renseigner le système.

1.1.2. Les systèmes d'aide aujourd'hui

Dans les visualisations de supervision radar opérationnelles en France, assez peu d'automatisation est intégrée et un nombre limité d'informations sont remontées du système de traitement radar pour être notifiées. Néanmoins, plusieurs systèmes d'aide à l'organisation du trafic et d'alerte sont aujourd'hui en exploitation ou en cours de déploiement. Parmi les systèmes d'alerte utilisés, le plus ancien est le filet de sauvegarde ou STCA (Short Term Conflict Alert – années 80) dont le rôle est de détecter les conflits entre aéronefs, c'est-à-dire des trajectoires amenant des avions à passer à une distance inférieure à la norme de séparation minimale (1000 pieds et 5 miles nautiques) et prévenir le contrôleur par une couleur pour alerter l'opérateur lors de conditions anormales ou à risques. Les supervisions radar utilisent de très grands écrans (par exemple écrans de 30" 16/9 ou 2000x2000 28" carrés comme en France). Sur des écrans de cette taille, des notifications risquent d'être montrées aux contrôleurs en vision périphérique. Le choix de la couleur utilisée pour les représenter est donc crucial puisque la perception des couleurs n'est pas uniforme avec l'angle visuel ; par exemple, le vert est très mal perçu à 30 ° du centre de vision, mais le jaune est bien mieux discriminé en vision périphérique (Ancman, 1991).

Lorsque des informations doivent être notifiées concernant un avion particulier comme dans le cas de l'alerte STCA, elles sont représentées sur l'« étiquette radar » (Figure 4 : l'étiquette radar). La position courante de l'avion est affichée par le symbole tête de chaîne, un guideur relie l'étiquette radar à celui-ci. Les positions passées de l'avion sont représentées avec la comète radar qui montre des symboles réduisant en taille en fonction du vieillissement de l'information position ; la comète radar montre 56s de positions passées. Les informations concernant l'avion sont stockées dans l'étiquette radar, elle est composée de plusieurs lignes (L0 à 4) dont certaines affichées de façon optionnelle en fonction du statut de l'avion pour le secteur de contrôle. La vitesse de l'avion est donnée par le vecteur vitesse, qui est tracé entre le symbole tête de chaîne et la position extrapolée de l'avion à X minutes, sa longueur est modifiable par le contrôleur (2 min sur la figure suivante).

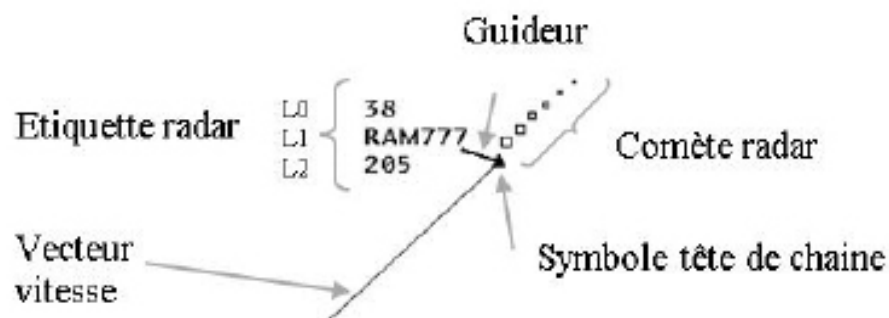
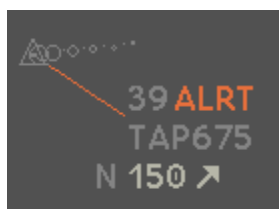


Figure 4 : l'étiquette radar

Deux types de designs sont utilisés pour notifier des informations aux contrôleurs ; ils s'appliquent à des warnings ou des alertes (Figure 5 : alerte filet de sauvegarde). Pour représenter un warning, un champ d'information texte va être montré en rouge orangé en L0 de l'étiquette radar, pour représenter une alerte, le champ d'information texte clignote en utilisant deux couleurs (blanc 200ms vers rouge orangé 800ms) et le guideur reliant l'étiquette radar à la comète est représenté en rouge orangé clignotant également.



Clignotement Texte ALRT: Rouge pendant 800ms alternant avec le Blanc pendant 200ms + clignotement du guideur (Rouge(800ms) et Blanc(200ms))

Figure 5 : alerte filet de sauvegarde

Tous les avions sont équipés d'un transpondeur interrogé périodiquement par des radars secondaires. Celui-ci renvoie un code sur quatre digits en octal. Certains codes permettent au pilote de prévenir le contrôle aérien en cas de détournement (hijack symbolisé par HJCK), de panne radio (radio off : RDOF) ou de détresse (emergency : EMRG). Ce type d'information est affiché dans la première ligne de l'étiquette avec un texte en rouge orangé non clignotant. En cas d'apparition ce type de code nécessite une prise en compte rapide par le contrôleur ; il s'agit d'évènements très rares. Lorsque le contrôleur n'est pas certain de l'identification de l'avion, il peut demander au pilote d'activer le mode IDENT de son transpondeur ce qui se traduira sur l'image radar par le texte IDEN en ligne 0 de l'étiquette. Certains espaces aériens nécessitent des niveaux d'équipement de radio navigation spécifique, par exemple une précision accrue de son système de mesure de l'altitude (pour les zones RVSM, Reduced Vertical Separation Minima). En fonction du niveau d'équipement de l'avion, si celui-ci pénètre une zone nécessitant un niveau d'équipement supérieur au sien, le contrôleur sera de la même façon prévenu avec un texte en rouge orangé en ligne 0. On peut constater que dans ces informations définies comme des Warning, certaines d'entre elles n'ont pas le même niveau de criticité pour le contrôleur, un détournement aérien par exemple étant bien plus grave qu'une notification RVSM, mais ils sont représentés avec le même design.

L'alarme STCA représentée par le texte ALRT clignotant en ligne 0 de l'étiquette, correspond au déclenchement du filet de sauvegarde ; ce système a pour but de détecter les pertes de séparation entre aéronefs et prévenir le contrôleur à temps pour changer leur trajectoire. Dans l'espace aérien, hors zones spéciales, 5 miles nautiques (environ 9 km) en longitudinal et 1000 pieds (300m) en vertical doivent être maintenus entre les aéronefs. Le STCA n'est pas un système anti-abordage comme le TCAS (Trafic Collision Avoidance System) embarqué à bord des appareils qui ne nécessite pas d'action du contrôle aérien ; les contrôleurs sont avertis plusieurs dizaines de secondes avant la perte de séparation. Du point de vue opérationnel, ce type d'alerte est considéré comme très important et nécessite une action rapide du contrôleur ; cependant, dans certaines conditions, le système ne connaissant pas les ordres donnés aux pilotes par le contrôleur, le STCA peut détecter une future perte de séparation sur des avions évolutifs. L'alarme est alors considérée comme intempestive. Il s'agit d'une notification de premier ordre

dans les spécifications de l'image radar, et il y est dit que l'alarme doit être vue en vision périphérique, mais sans contrainte de performance en termes de temps de détection moyen de la notification par les contrôleurs.

D'autres systèmes spécifiques à l'approche sont en place dans ces aéroports comme MSAW (Minimum Safe Altitude Warning) qui détecte les trajectoires dangereuses, car risquant d'amener un avion en descente sur le relief ou APW (Area Proximity Warning), un système de détection d'intrusion dans des zones prohibées (zones militaires, de parachutage, centrales nucléaires etc.). Toujours au niveau de l'approche, un système d'aide MAESTRO (Moyen d'Aide à l'Écoulement Séquence du Trafic avec Recherche d'Optimisation) permet au contrôleur d'approche de préparer la séquence d'arrivée à l'aide d'une représentation temporelle de cette séquence sur un agenda.

Au niveau de la supervision radar en-route opérationnelle, les contrôleurs disposent d'outils très simples pour mettre en évidence des vols au niveau de la visualisation manuellement et quelques notifications les prévenant de situations anormales ainsi que les alertes STCA décrites précédemment. Ils ne disposent actuellement d'aucune aide à la décision.

En Europe d'autres systèmes sont aujourd'hui en service pour les centres en-route. Celui de Maastrich utilise des données en provenance des avions (Downlink Aircraft Parameters) concernant l'altitude sélectionnée sur le système de gestion du vol et les corrèle avec les informations renseignées par les contrôleurs sur leur supervision radar ; en cas d'incohérence, une alarme est remontée au contrôleur. Dans le centre Suisse, un système de détection de conflits à moyen terme (MTCD : Medium Term Conflict Detection) est déployé. Il permet en réalisant une extrapolation horizontale de la position de l'avion de trouver les conflits potentiels et de l'aider à les résoudre en amont du déclenchement du filet de sauvegarde.

1.2. Évolutions futures

Des statistiques en provenance de la Direction Générale de l'Aviation Civile montrent que ces 27 dernières années hormis pendant la période suivant l'attentat du 11 Septembre 2001, le trafic aérien au-dessus de la France a augmenté régulièrement. Il s'établissait à 166M de passagers en 2012 (Figure 6 : évolution du trafic en France entre 1988 et 2012 ; radial : Paris/province, transversal : province/province (source DGAC/DTA/SDE)). Les projections d'évolutions de ce trafic pour les 20 années à venir tablent sur une augmentation de 4.6% par an. Les conséquences pour la gestion du trafic aérien seront considérables.

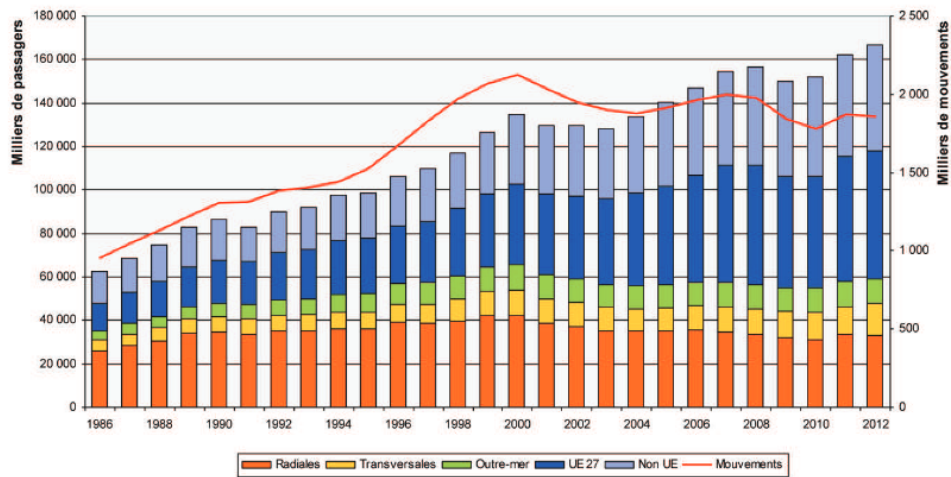


Figure 6 : évolution du trafic en France entre 1988 et 2012 ; radial : Paris/province, transversal : province/province (source DGAC/DTA/SDE)

En France, pour le court et moyen terme, deux projets sont en cours de développement industriel. Le projet EEE (ERATO Environnement Electronique) qui doit être mis en service en 2015, met fin à l'utilisation du strip et propose une interface de supervision (Figure 7 : EEE ERATO Environnement Electronique) disposant de fonction de saisie des clairances de contrôle (consignes données par le contrôleur aux pilotes dans son secteur), de fonctions d'aide à la décision et d'un agenda manuel permettant au contrôleur de créer des étiquettes avec des conflits qu'il a identifié sur une échelle de temps montrant les 10 prochaines minutes de la séquence de contrôle (partie droite de la supervision). Pour aider les contrôleurs dans leur activité de résolution de conflits, un filtrage permet de mettre en évidence (design différent) des avions pertinents par rapport à un vol de référence de façon à aider le contrôleur à analyser plus rapidement la situation. Ce filtrage est basé sur des hypothèses sur le comportement du vol et du contrôleur. Il permet par ailleurs d'avoir une estimation de l'impact d'une consigne sur un vol de référence sur le trafic avoisinant (filtrage simulé).

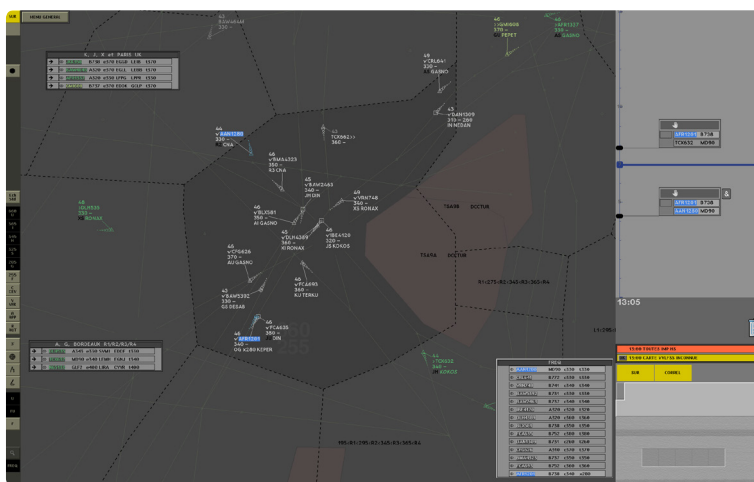


Figure 7 : EEE ERATO Environnement Electronique

Le deuxième, 4Flight, est un partenariat entre le fournisseur de service de contrôle Italien (ENAV) et la France ; c'est la prochaine évolution du système de gestion du trafic aérien. Il comprend une nouvelle supervision radar

et intégrera les outils présents dans EEE ainsi qu'un nouveau système de traitement de plan de vol (CoFlight) qui permettra de faire une prédiction de trajectoire plus précise (prédiction 4D) et d'échanger des informations avec les autres centres et les avions par datalink (canal de données sol-bord). Le développement est en cours pour une cible de mise en tests dans deux centres de contrôle en-route pilotes en 2015, et une généralisation à tous les centres d'ici 2020. Les commissions de sécurité recommandent régulièrement des évolutions du système de notification et particulièrement dans le cas du filet de sauvegarde. Celles-ci sont prises en compte dans le développement de 4-Flight.

Au niveau de l'Europe, pour le moyen et long terme, deux projets sont en cours pour essayer de répondre aux enjeux de cette croissance du trafic, le projet de ciel unique SES (Single European Sky) et le projet SESAR (Single European Sky ATM Research).

Dans le projet de Ciel Unique Européen, l'espace aérien couvrant l'Europe est découpé en blocs fonctionnels (volumes d'espace) regroupant plusieurs pays, la France fait partie du FABEC (Functional Airspace Block Europe Central). La création de ces blocs fonctionnels permettra de s'affranchir des frontières des états et réduire la fragmentation de l'espace aérien.

Le projet SESAR a pour but de moderniser l'infrastructure de contrôle aérien à l'échelle de l'Europe pour les 30 prochaines années. La phase de définition du projet s'est terminée en 2008, la phase de développement est en cours. De nouvelles technologies sont apparues dans le domaine aéronautique permettant de partager des informations entre les différents acteurs (compagnies aériennes, aéroports, services de navigation aérienne). En utilisant ces informations et moyens de communications, SESAR permettra de développer de nouveaux concepts visant à une meilleure intégration des systèmes qui aboutiront entre autres à la réduction des délais d'attente au sol ou en vol. Parmi les concepts testés, le free flight (pas de contrôle aérien centralisé), la business trajectory (trajectoire optimisée en fonction de critères propres à chaque opérateur aérien), la déconfliction des trajectoires (des changements limités de vitesse en amont des conflits pour les faire disparaître), la délégation de responsabilité sol-bord ou bien l'utilisation de nouveaux moyens de navigation comme le satellite (GALILEO) pour concevoir de nouveaux types de procédures. SESAR est découpé en 16 Work Packages (WP), portant sur des aspects spécifiques (par exemple WP4 : en-route, WP6 : approche). SESAR a aussi pour but de favoriser la recherche à long terme et finance un nombre important de projets de recherche dans le cadre du paquet WP E : Innovative Research.

Les travaux présentés dans cette thèse ont été développés dans le cadre du WP4.7.2 portant sur la position de contrôle en-route. Il comporte deux volets, le premier concerne le développement et la mise en œuvre d'un outil permettant de diminuer le nombre de conflits potentiels en agissant sur la vitesse des avions de façon anticipée, le second porte sur une position de contrôle innovante permettant d'intégrer cet outil. Ces légères modifications de vitesse permettent dans un nombre important de cas d'assurer une séparation suffisante sur les points de conflits de façon à limiter la charge de travail induite aux contrôleurs. Le deuxième volet, dans lequel ces travaux sont insérés, concerne la position de travail des contrôleurs aériens et plus spécifiquement la visualisation de supervision radar. L'automate d'aide à la résolution de conflit n'est pas en mesure de résoudre tous les problèmes

détectés ; les conflits résiduels non traités, mais connus de l'agent intelligent doivent être présentés aux contrôleurs de façon à limiter les ressources nécessaires à la détection de ces conflits. Par ailleurs lorsqu'un conflit est géré par l'automate, il est nécessaire d'en informer le contrôleur de façon à éviter une intervention inappropriée qui pourrait complexifier la situation au lieu de l'arranger. Pour assurer un fonctionnement nominal de l'automate ainsi que des outils de prévision de trajectoire des avions, il est nécessaire de renseigner le système avec les ordres donnés par le contrôleur aux avions sous sa responsabilité. Les conséquences de ces changements sont considérables pour la méthode de travail actuelle des contrôleurs. L'intégration d'un agent intelligent en mesure d'agir sur le trafic aérien et d'aider le contrôleur à détecter les situations conflictuelles résiduelles va transformer significativement le rôle du contrôleur pour l'obliger à établir un lien de confiance avec un automate et déléguer une partie de sa responsabilité concernant le maintien des séparations entre aéronefs. L'objectif de la visualisation radar en-route développée dans le cadre du projet SESAR WP4.7.2 est de permettre d'assurer un bon couplage entre humain et automate et faciliter la gestion d'un nombre très important d'avions tout en renseignant le système de façon performante.

1.3. Erreurs humaines dans le domaine aéronautique

Dans l'ATC, les incidents et accidents sont analysés périodiquement par des commissions qui établissent les causes et facteurs contributifs dans le but d'améliorer la sécurité. En France, l'analyse d'un évènement se déroule en quatre étapes : le recueil des faits, l'identification des causes, la classification de l'évènement et l'éventuelle proposition d'actions préventives ou correctives. Les résultats d'analyse sont stockés dans une base de données. Depuis 2011 la grille RAT (Risk Analysis Tool) est utilisée pour classer les incidents. L'Europe, la FAA (Federal Aviation Administration – USA) et ATNS (Air Traffic and Navigation Service – Nouvelle Zélande) utilisent aussi RAT. La méthode d'analyse mise en place en France consiste à réaliser un arbre de causes pour identifier causes et facteurs contributifs. Les évènements sont classés à l'aide de la grille RAT qui liste les cas possibles d'incidents répertoriés par gravité et probabilité de nouvelle occurrence. Les alertes filets de sauvegarde aboutissant à un passage sous les minima de séparation (1000ft 5Nm) font l'objet le plus souvent d'un incident et d'une analyse, les rapprochements de deux vols (nommés HN) sont classés en fonction du degré de franchissement des minima (HN50 = passage sous 50% des minima). Les évènements étudiés en 2011 - 34 cas (Figure 8 : causes et facteurs contributifs aux incidents 2011) font émerger 3 catégories importantes 24 % d'entre eux sont liés à un travail en marge ou en surcharge amenant à une stratégie inadaptée dans 20 % des cas, dans 17 % des cas la gestion de l'attention est mise en cause. Les bilans de l'Instance de Traitement des Evènements de Sécurité (ITES) aboutissent à des recommandations concernant les outils les méthodes de travail consignes et rappels ou des REX (Retour d'EXpérience) sur des analyses d'incidents particuliers. Par exemple, une recommandation de 2011 concernant les outils suggère un renforcement du design de l'alarme STCA sur le secteur de contrôle, car des problèmes liés à sa perception ont été mis en évidence dans certains cas.

Les conséquences d'accidents dans le domaine de l'aviation peuvent être considérables. Les défaillances peuvent provenir de trois sources selon Reason (1990), les défaillances techniques, les défaillances liées à des facteurs organisationnels et les erreurs humaines. Le domaine aéronautique est considéré comme ayant atteint un haut niveau de sécurité, l'essentiel des incidents et accidents sont donc liés à des erreurs humaines (Kinney, et al., 1977; FAA, 1990). Parmi ces erreurs, on trouve les interactions vocales et incompréhensions entre le contrôle aérien et les équipages (Cushing, 1994,1995) spécifiques à cette activité. Toujours selon Reason (1990), le terme d'erreur couvre tous les cas où une séquence planifiée d'activités mentales ou physiques ne parvient pas à ses fins désirées, et quand ces échecs ne peuvent être attribués à l'intervention du hasard. Les erreurs humaines sont inévitables, elles font partie de notre nature ; leur étude permet entre autres d'améliorer la connaissance sur un système dans le but d'optimiser le couplage humain-système.

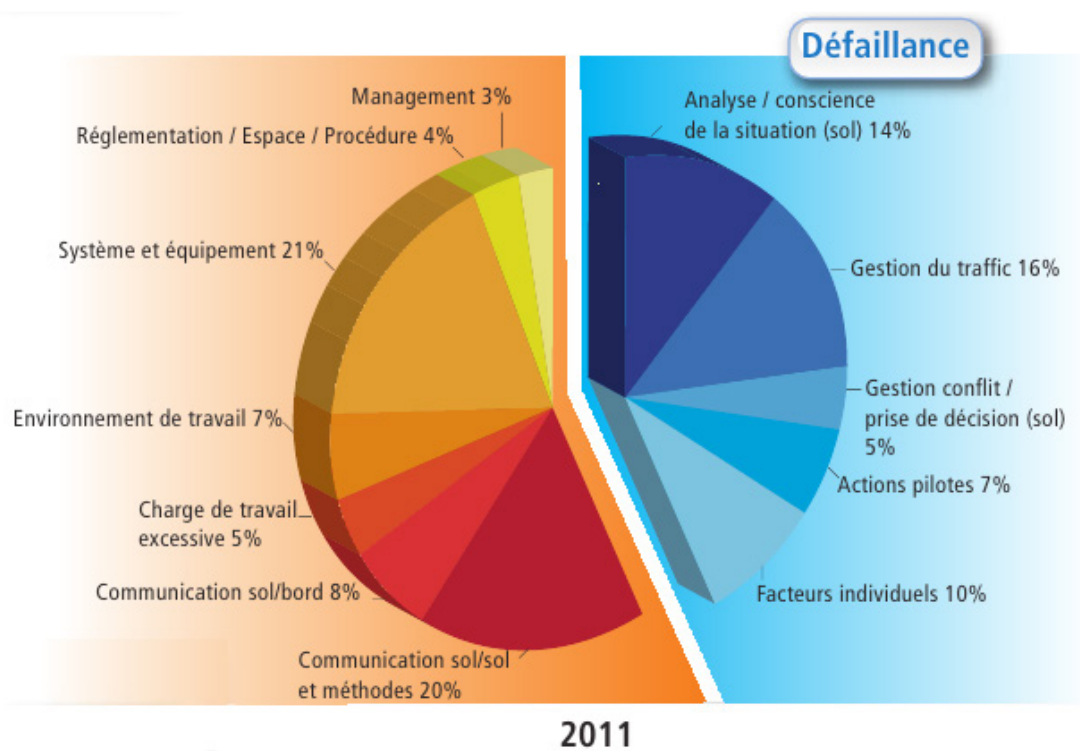


Figure 8 : causes et facteurs contributifs aux incidents 2011- Rapport Sécurité Circulation Aérienne n°52

Une étude de la CAA anglaise (Civil Aviation Authority) de 2009, analysant les accidents (crashes) dans l'aviation commerciale entre 1997 et 2006 dans le monde, montre que l'erreur humaine est la cause de 2/3 de ceux-ci. L'essentiel de ces erreurs est lié à des décisions inappropriées elles-mêmes liées à une mauvaise conscience de la situation. La capacité à développer une représentation mentale de la situation courante et d'arriver à projeter son état futur a été décrite par Endsley (1995) sous le concept de conscience de la situation (Situation Awareness ou SA). La définition généralement employée est :

« La conscience de la situation est :

- *la perception des éléments de l'environnement dans le temps et l'espace (SA1)*

- *la compréhension de leur signification (SA2)*
- *la projection de leur statut dans un futur proche (SA3)*

Les causes de ces erreurs humaines peuvent être catégorisées en fonction du niveau de conscience de la situation (SA1|2|3). Dans le domaine du contrôle aérien, une analyse des erreurs opérationnelles conduite aux États-Unis utilisant cette classification (Endsley & Rodgers, 1998) a montré que (certaines erreurs peuvent porter sur plusieurs catégories) :

- *71 % concernent le niveau SA1 dont :*
 - ▶ *35.7 % dues à l'incapacité à suivre tous les paramètres*
 - ▶ *21.4 % dues à mauvaise perception*
 - ▶ *14.3 % dues à problème de mémoire*
- *21.4 % concernent le niveau SA2*
- *28.6 % concernent le niveau SA3*

Basée sur le modèle de la conscience de la situation, une taxonomie des erreurs en fonction du niveau de conscience de la situation a été établie par Endsley (1995) :

SA1 : informations non perçues correctement

- données non disponibles
- données difficiles à discriminer ou détecter
- défaut de surveillance des données
- perception erronée des données
- perte de mémoire

SA2 : informations mal intégrées ou mal comprises

- Absence ou modèle mental pauvre
- Modèle mental incorrect
- Une dépendance excessive aux valeurs par défaut
- Autres

SA3 : incapacité à projeter l'état du système ou les actions futures

- Absence ou modèle mental pauvre
- Projection des tendances actuelles excessives
- Autres

Cette liste n'est pas exhaustive et d'autres approches classifiant ou modélisant les erreurs sont utilisées. Dans le domaine aéronautique, le modèle de Reason GEMS (Generic Error Modelling System, 1993) est très largement employé. Il catégorise les erreurs suivant trois types : les lapsus, les fautes et les ratés et les classe en fonction des niveaux de contrôle cognitif décrit par Rasmussen (1994) dans son modèle Skill-Rules-Knowledge. (cf. 3.5.1). Hollnagel (1993) propose une autre approche et classe les erreurs en fonction de leurs aspects observables et leurs facteurs contributifs.

1.4. Résumé

Établir une bonne conscience de la situation est la pierre angulaire de la supervision des systèmes critiques. Le contrôle aérien est dans une période de changement majeur, le trafic en augmentation quasi constante depuis les années 60, les prévisions d'évolution tablant sur une croissance supérieure à 4% pour les vingt prochaines années sont un enjeu considérable pour la gestion du trafic aérien. Partout dans le monde, les fournisseurs de service de navigation aérienne font évoluer leurs systèmes pour augmenter sa capacité tout en voulant maintenir un haut niveau de sécurité. En Europe le projet SESAR est l'initiative commune visant à moderniser ces systèmes. La réponse à ces objectifs est le développement de l'automatisation à tous les niveaux du système de gestion du trafic aérien par la mise en œuvre de nouveaux outils de filtrage ou d'aide à la décision, d'outils de régulation du trafic sans action du contrôleur ou du pilote ou de nouveaux moyens d'échange d'informations sol/bord. Dans ce contexte, les interfaces de supervision des contrôleurs aériens auront toujours plus d'informations à afficher qui s'ajouteront ou feront évoluer les interfaces existantes. L'acquisition et la mémorisation des informations présentées sur la supervision radar ainsi que la perception des notifications issues du système sont le prérequis nécessaire à l'établissement de la conscience de la situation des contrôleurs aériens.

Les systèmes dans le domaine aéronautique, aussi bien au niveau du sol que du bord ont atteint un niveau très élevé de fiabilité et de sécurité, mais les analyses d'incidents montrent que l'humain est la source principale d'erreur. Dans ces erreurs humaines, les deux tiers sont liées à des problèmes de perception ou de mémorisation des informations c'est-à-dire le niveau 1 de la situation awareness : la perception des éléments de l'environnement dans le temps et l'espace. Pour maintenir ce haut niveau de sécurité, les nouveaux systèmes devront prendre en compte cette problématique et ses éventuelles interactions avec l'automatisation.

L'interaction entre l'humain et la visualisation ainsi que le système automatisé est au centre de ce travail de recherche ; le prochain chapitre traite donc de la perception des informations dans les visualisations et plus particulièrement des visualisations de type supervision et de leurs caractéristiques.

2. CONTRAINTES ET OBJECTIFS DES VISUALISATIONS DE TYPE SUPERVISION

2.1. Introduction

Les visualisations de type supervision sont des interfaces permettant d'assurer la surveillance et le contrôle d'installations ou de systèmes complexes. Les opérateurs de systèmes de supervision doivent acquérir un grand nombre d'informations dynamiques pouvant avoir des relations d'interdépendance. Ils sont par ailleurs confrontés à des objectifs qui peuvent être antagonistes ; un objectif de performance pouvant être en contradiction avec des exigences de sécurité qui proscrivent les erreurs. Ces opérateurs sont aussi soumis à une forte pression temporelle engendrant un stress important puisque dans certaines situations ils sont amenés à prendre des décisions très rapidement pour maintenir un niveau de sécurité acceptable. On trouve des visualisations de type supervision dans un grand nombre de domaines critiques ou des risques majeurs pour la sécurité du personnel voire la sécurité des habitants dans la zone environnante peut être mise en danger. C'est le cas pour la gestion de centrales nucléaires, les systèmes de commandement et de contrôle, l'aéronautique ou le contrôle aérien.

Dans ce chapitre nous décrivons les spécificités de cette classe de visualisation, de la problématique de la transmission d'informations à l'humain ainsi que des modes de transmission des notifications en provenance de la supervision.

Dans les centrales nucléaires françaises, les postes d'opérateurs sont composés d'instruments de visualisation des différents indicateurs (données et alarmes) ainsi que de panneaux de commande. La conduite normale de la centrale consiste à répondre à un programme de production d'électricité et de surveillance de différents indicateurs, différents systèmes contrôlent la conduite de la centrale et deux types d'alarmes peuvent être émis, celles de type incident et celles de type accident mettant en cause la sécurité de la centrale. En cas de notifications mineures, l'approche événementielle est privilégiée, en cas de notifications majeures une approche par état basée sur plusieurs paramètres donnant l'état du réacteur est utilisée. Le traitement des incidents se fait par application de procédures. Depuis l'incident de Three Mile Island (Kemeny, 1979) une attention toute particulière est apportée à la conception du système de gestion des notifications. Il a été constaté à cette occasion que le système tel qu'il était conçu aboutissait à des surcharges cognitives trop importantes lors de certaines phases, pouvant aboutir à des incidents ou des accidents. Différentes études ont été réalisées pour améliorer la conception de ces postes ; en 2001 Ham et Youn utilisèrent la conception d'interface écologique (Vicente, 1992) ; Bye et al, Hickling (Stanton, 1994) une conception permettant de filtrer et catégoriser les notifications permettant de remonter des alarmes de plus haut niveau donnant une vision centrée fonction plutôt que système.

Dans le domaine de l'aviation civile, contrôleurs aériens et pilotes peuvent être considérés comme des opérateurs d'un système complexe supervisé. Dans le cockpit d'un avion de transport de passagers, quantité d'informations sont présentées aux pilotes de façon dynamique sur plusieurs écrans. Des systèmes d'aide assistent les pilotes dans leur tâche et notifient les informations pertinentes en temps réel en utilisant des designs visuels spécifiques ainsi que des feedback sonores et haptiques. Les problématiques de sécurité et de performance associées font clairement partie des contraintes des pilotes dans la réalisation de leur tâche.

Au sol, les contrôleurs aériens utilisent des visualisations de type supervision pour le suivi des aéronefs dans leur espace de responsabilité. Ces visualisations ont pour origine les écrans radars conçus dans les années 50 (Figure 2: visualisation radar opérationnelle des années 50 à nos jours), elles ont évolué au cours du temps vers des écrans de plus en plus grands sur lesquels sont représentés un nombre considérable d'informations (Figure 1: du contrôle en-route au contrôle d'aérodrome). Comme dans le cockpit différents systèmes d'aide peuvent aider le contrôleur dans la réalisation de sa tâche. Parmi les informations présentées aux contrôleurs se trouvent des notifications provenant de ces systèmes ayant pour but de l'avertir lors de situations accidentogènes, de spécificités de certains aéronefs ou de changements dans l'état du système de traitement de données alimentant son interface de supervision.

Les nouveaux développements dans le domaine des visualisations pour contrôleurs aériens aboutissent à l'utilisation d'écrans multiples et d'écrans de grande taille (30") du fait de l'intégration de systèmes d'aides et de l'augmentation du nombre d'informations à représenter ; cela signifie que des informations seront souvent montrées en vision périphérique. Dans le cockpit aussi, la multiplication des écrans et des instruments rendent cette problématique de perception en périphérie cruciale.

2.2. Perception des informations dans les systèmes de visualisation

2.2.1. Variables visuelles

Les différentes propriétés visuelles ont été définies par Bertin (1988) dans la sémiologie graphique. Il y décrit un ensemble de sept variables visuelles (Figure 9 : variables visuelles de Bertin) qui forment les éléments de plus bas niveau pour la composition graphique : la position (la variable visuelle spatiale), la taille, la valeur (proche de la luminosité), le grain ou texture, la couleur (mélange de teinte et de saturation), l'orientation, la forme. Ces variables visuelles sont appliquées à trois types d'implantation dans le plan : le point, la ligne, la zone.

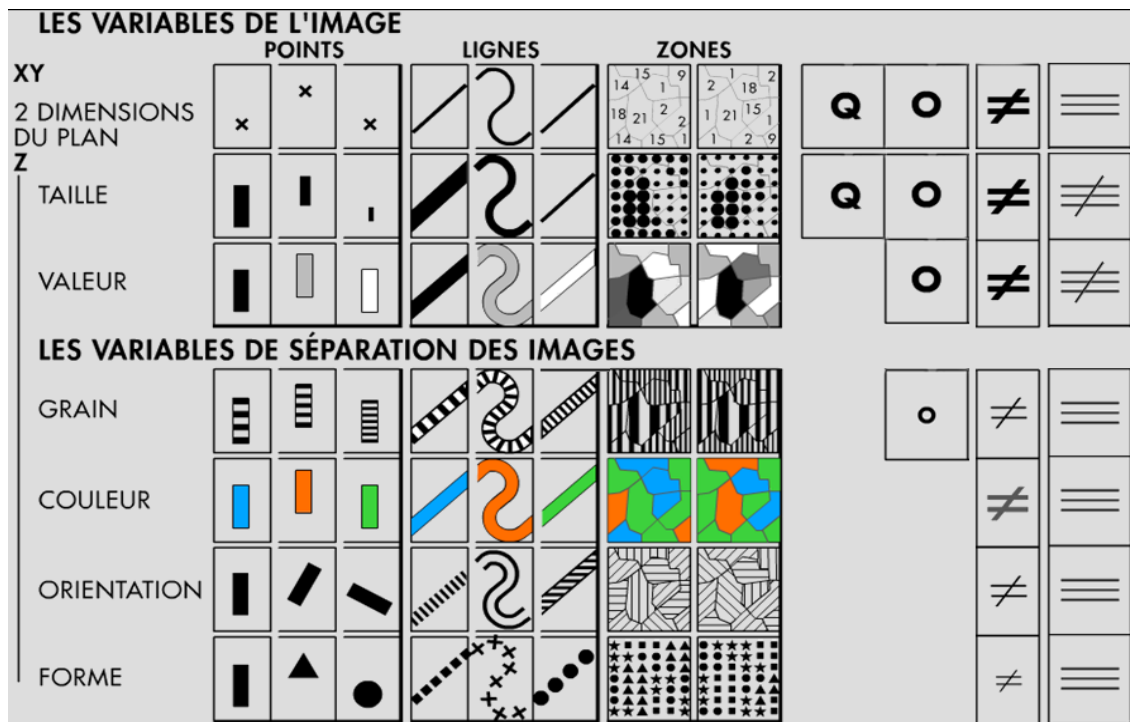


Figure 9 : variables visuelles de Bertin

Les variables visuelles ont des échelles de perception différentes appelées longueur. La longueur d'une variable visuelle est le nombre d'éléments différenciables avec cette variable visuelle. Par exemple, la variable de position possède l'échelle de perception la plus longue et offre même une infinie possibilité de variation.

Une variable visuelle est dite sélective si elle permet d'effectuer un tri dans les informations représentées. Elle permet d'isoler sans effort cognitif toutes les occurrences d'une catégorie d'objets, on peut alors parler de la sélection préattentive (Treisman, 1985). Toutes les variables visuelles sont sélectives, mais à différents degrés. La figure 10 donne, par ordre décroissant, le niveau de sélectivité des variables visuelles défini par Bertin.

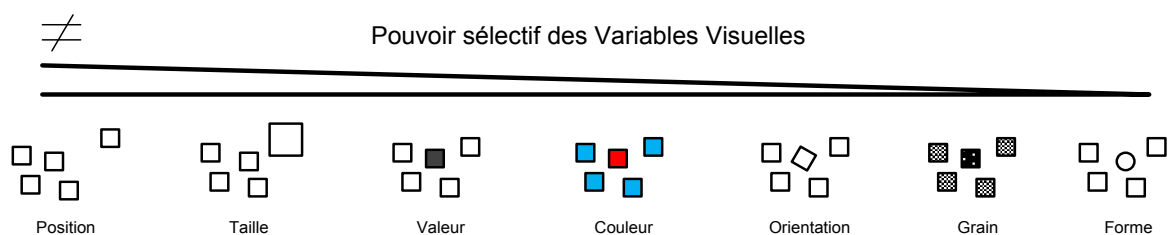


Figure 10: pouvoir sélectif des variables visuelles

Wilkinson (1999) a étendu l'ensemble des variables visuelles en ajoutant la saturation, la transparence, le flou. Il a par ailleurs étudié de nouveaux vecteurs de transmission d'informations comme l'animation ou le son ayant eux aussi des variables spécifiques. Le tableau suivant (Figure 11 : vecteurs de transmission d'informations) présente les choix de design, ordonnés du plus efficace au moins efficace en fonction du vecteur de transmission d'informations (la forme, la surface, l'animation, le son et le texte). À ces vecteurs, l'odorat ou le toucher pourraient être ajoutés.

Forme	Surface	Animation	Son	Texte
Position	Couleur Saturation	Direction Vitesse	Hauteur	Texte
Taille Formes	Luminosité	Accélération	Volume	
Polygone	Teinte		Rythme	
Glyphe Image	Texture		Voix	
Rotation	Motif			
Résolution	Granularité			
	Orientation			
	Flou			
	Transparence			

Figure 11 : vecteurs de transmission d'informations

2.2.2. Perception avec faible charge cognitive

La théorie de la perception préattentive (Treisman, 1985) explique comment certaines propriétés visuelles peuvent être perçues très rapidement (< 250 ms) et sans charge cognitive. La couleur, l'orientation, la longueur, l'épaisseur, la courbure, et d'autres propriétés graphiques peuvent aider à la perception préattentive (Healey, 1992). Si pour une propriété donnée, par exemple la couleur, une représentation utilise trop de couleurs différentes, cette sélection préattentive ne fonctionnera plus. La perception contrôlée mobilise de façon importante les ressources cognitives. Lire est une tâche de perception contrôlée, elle est lente et coûteuse au niveau cognitif. La perception des objets est réalisée par la synthèse mentale des formes en une « surforme », c'est la théorie de la Gestalt (Kofka, 1935). Elle regroupe plusieurs principes de la perception : la loi de la bonne forme, la loi de continuité, la loi de proximité, la loi de similitude, la loi du destin commun et de clôture. La loi de bonne forme explique qu'un ensemble de parties (comme des groupements aléatoires de points) tend à être perçu au premier abord comme une forme avec des propriétés de stabilité et de forme. La loi de proximité définit le regroupement d'objets similaires proches les uns des autres. La loi de continuité statue qu'il est plus facile de construire des entités visuelles à partir d'éléments qui sont lissés et continus plutôt qu'avec des éléments qui comportent des angles, ou des changements de direction abrupts. La loi de similitude indique que les éléments visuels qui partagent les mêmes propriétés se regroupent. La loi du destin commun définit que des parties en mouvement ayant la même trajectoire sont perçues comme faisant partie de la même forme. La loi de clôture montre qu'une forme fermée est plus facilement identifiée comme une figure (ou comme une forme) qu'une forme ouverte.

Dans l'illustration suivante (Figure 12 : exemple de la Gestalt sur un scatterplot), la Gestalt permet d'expliquer les processus perceptifs mis en œuvre dans la perception des informations contenues dans un scatterplot. Le système visuel humain permet de percevoir sans surcharge cognitive des tendances (loi de la continuité), des patterns (loi de la similitude), des exceptions, des groupes (loi de la proximité).

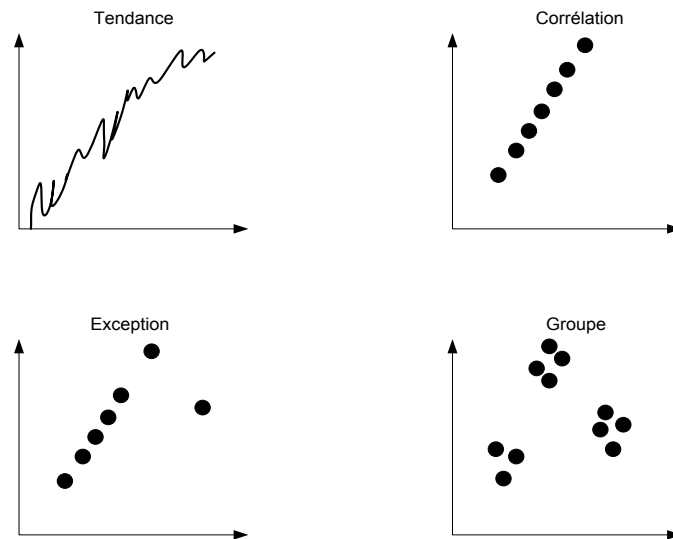


Figure 12 : exemple de la gelstat sur un scatterplot

2.2.3. Approche écologique de la perception et design écologique

L'approche écologique de la perception est une explication du processus de capture d'informations de l'environnement que l'on doit à J.J. Gibson, (1977). Elle consiste à définir ce qu'il y a à percevoir en fonction des capacités humaines et de son environnement. L'approche cognitive consiste à réduire le problème de la perception à l'œil au cerveau et la lumière, au contraire de l'approche de Gibson qui considère l'humain dans son environnement, seule façon pour comprendre comment les choses sont perçues. Pour Gibson, la perception ne correspond pas à une série d'images, mais à un flux continu dont les changements mettent en évidence des invariants de structure. Le rôle du système perceptif visuel ne serait donc que de capturer les informations présentes dans l'environnement de l'homme. Gibson définit l'occlusion comme les surfaces non visibles en un point d'observation donné. Bien que non visible une image mentale de l'objet sous forme tridimensionnelle est néanmoins construite par le cerveau. Elle peut être affinée par un déplacement de l'observateur. Dans les propriétés écologiques de la perception décrites par Gibson, l'affordance décrit les propriétés d'un objet qui peuvent être utiles à un animal. La perception de la composition et de la disposition des surfaces conduit à la perception de leurs affordances et donc à la façon de les utiliser.

Deux approches utilisant les théories de Gibson sur la perception écologique ont été appliquées au design. La première concerne l'affordance que Donald Norman (Norman D. A., 1988) dans *The Design of Everyday Things* a appliqué pour les affordances perçues dans le domaine du design graphique. Dans l'environnement, on peut parler concernant les objets d'affordances qui peuvent être réelles et/ou perçues alors que dans les interfaces graphiques le designer n'a le contrôle que sur les affordances perçues par l'utilisateur. L'affordance dans ce cas est la manifestation de la compréhension intuitive de l'utilisateur qu'il peut par exemple interagir avec l'application en cliquant dans une zone.

La seconde approche date de 1992, Rasmussen et Vicente ont décrit une méthode de conception (EID : Ecological Interface Design) visant à améliorer la perception et la compréhension des informations adaptée à des systèmes de plus en plus complexes. La conception écologique des interfaces homme-machine est une démarche proposée pour faire face à la conception d'interfaces pour des systèmes techniques complexes et temps réel et est basée sur le cadre de Rasmussen SRK (Rasmussen, 1983) et la hiérarchie d'abstraction (AH). SRK décrit trois types de comportements basés sur les compétences (Skills), les règles (Rules) et les connaissances (Knowledge) (Vicente & Rasmussen, 1992). Un comportement basé sur les compétences ou habiletés nécessite très peu de contrôle conscient, libérant des ressources cognitives pour des fonctions de plus haut niveau. Il s'agit d'actions quasi automatiques de la part de l'opérateur. Un comportement basé sur les règles consiste à exécuter des actions définies dans des procédures et qui correspondent à des situations familières, ces procédures ou situations habituelles étant apprises lors de l'activité ou décrites dans des manuels d'opération. Un comportement basé sur les connaissances est adapté à une situation inhabituelle, elle nécessite des ressources cognitives plus importantes que les deux premières puisque l'opérateur doit avoir une idée précise de l'état du système pour prendre une décision. La hiérarchie d'abstraction est définie par une analyse du domaine de travail, et permet de décrire un système complexe à différents niveaux d'abstraction. Dans la méthode EID, la hiérarchie d'abstraction est utilisée pour décider de ce qui doit être affiché et comment sur l'interface de l'opérateur.

Le but des interfaces écologiques est d'apporter un soutien spécifique à l'opérateur, lui fournissant les outils et les informations nécessaires pour résoudre des problèmes dynamiques et en particulier au cours de l'évolution d'événements imprévus (Pawlak, 96). Ce type de démarche a été utilisé dans différents domaines comme les interfaces de supervision de centrales nucléaires (Rasmussen & Vicente, 1992), de l'aviation (Ho, 2003) ou des systèmes de contrôle et de commande militaire (Burns, 2004). L'adoption d'interfaces écologiques, en facilitant la mise en œuvre des comportements basés sur les compétences et les règles permettrait de réduire la charge mentale lorsqu'il s'agit d'événements inconnus et inattendus sous pression psychologique accrue.

Dans le cas du contrôle aérien, les compétences de base sont apprises lors de la formation initiale des contrôleurs. Elles permettent à ceux-ci de maintenir des niveaux de sécurité élevés entre aéronefs, ainsi que d'acquérir les automatismes leur permettant par exemple de donner les ordres de cap adéquats pour résoudre un conflit, ou bien d'assurer la norme de séparation entre aéronefs, compte tenu des performances de ceux-ci de la météo et de leur altitude. Ces apprentissages sont renforcés lors des phases de qualification des contrôleurs sur leur lieu d'affectation ; ils y développent par ailleurs des compétences basées sur des règles spécifiques à leur centre de contrôle ou leur secteur ainsi que la connaissance spécifique sur le trafic évoluant dans leur secteur de contrôle.

Des designs écologiques visant à faciliter la compréhension de l'état du système et donc de limiter la charge cognitive en mode de travail nominal, ont été proposés par Hurter et al. en (2009) et Imbert & Hurter (en préparation). Ces travaux sont basés sur une caractérisation des données affichées sur les visualisations radar opérationnelles suivant la méthode de Card et

Mackinlay (1999) et l'utilisation de deux techniques mises en évidence lors de cette analyse permettant l'émergence d'information grâce à leur représentation en utilisant la rémanence et l'animation (Hurter C. thèse 2010).

Les visualisations qui utilisent la perception écologique ont donc entre autres pour but de limiter l'apprentissage de leurs fonctions, faciliter la perception préattentive des informations et utiliser le principe de la perception des flux continus pour favoriser la détection des changements et l'émergence d'invariants. Le résultat attendu de ce type de visualisation est une charge cognitive pour la perception d'information diminuée ainsi qu'une meilleure compréhension des informations à traiter. Il n'existe pas de méthode systématique permettant aux concepteurs de réaliser des 'designs écologiques', mais la connaissance des propriétés de la perception écologique de l'information permet aux designers d'interfaces et d'interaction homme-système d'optimiser la transmission d'informations.

En ce qui concerne la rémanence, la comète radar (Figure 13 : la comète radar des visualisations radar) présente plusieurs symboles successifs montrant l'avion à des instants différents, le symbole de plus grosse taille correspond à la position la plus récente de l'avion. L'écart de position entre deux symboles permet pour un expert d'avoir une idée de la vitesse de l'avion, si cet écart varie sur plusieurs positions il est possible de déduire l'accélération de l'avion ; ceci illustre le mécanisme de la rémanence (plusieurs positions passées) et de l'émergence d'information (vitesse et accélération).

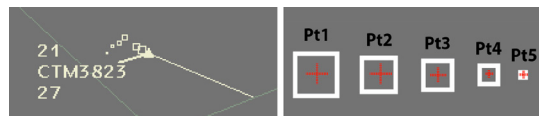


Figure 13 : la comète radar des visualisations radar

Des informations de position sont donc présentées à l'aide de la comète radar, la vitesse à l'aide d'un vecteur présentant la position de l'avion 3 minutes plus tard accroché à la position la plus récente (ligne blanche sur la figure précédente). Différents éléments complémentaires sont affichés dans une étiquette sous forme de texte ou de symboles. Par exemple sur la figure précédente, la flèche présentée à droite du chiffre 27 (l'altitude en niveau de vol, un niveau de vol équivaut à 100 pieds-30m) indique que l'avion est en montée. Pour avoir une idée du taux de montée, le contrôleur doit estimer la variation du niveau de vol entre deux mises à jour de position et faire un rapide calcul mental.

Le design suivant basé sur les théories de Gibson (Figure 14 : design écologique pour la représentation de la vitesse verticale) a été proposé par un contrôleur aérien et étendu en utilisant des variables visuelles supplémentaires ainsi que le mécanisme d'animation évoqué précédemment. Il utilise le vecteur vitesse pour représenter des cercles montrant où l'avion se trouvera au prochain niveau de vol. Dans l'exemple précédent où on peut voir la comète radar opérationnelle, l'avion est au niveau 270 en montée (27 et flèche verticale en ligne 3 de l'étiquette). Si l'on utilise ce design, des cercles seraient dessinés sur le vecteur vitesse montrant sa position X,Y lorsqu'il passerait les prochains niveaux de vol (280, 290 ...). Dans la figure 14, plusieurs designs sont exposés utilisant la couleur ou la taille pour montrer que le niveau augmente. L'écart entre ces points donne en un coup d'œil une idée de la vitesse verticale de l'avion. Plus ils sont serrés plus l'avion monte vite,

l'animation a aussi été utilisée pour faire émerger l'accélération verticale. Lorsque la vitesse verticale évolue, les points sont écartés ou rapprochés en utilisant une animation. De cette façon il est possible pour le contrôleur de percevoir le changement d'attitude. Lorsque les points se rapprochent, l'avion est en phase d'accélération (la vitesse verticale augmente), lorsqu'ils s'éloignent l'avion réduit sa vitesse verticale.

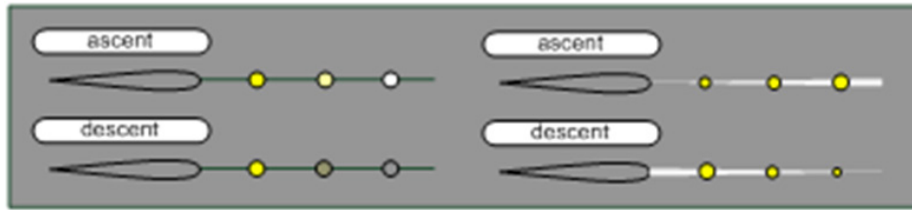


Figure 14 : design écologique pour la représentation de la vitesse verticale

Ce type de design peut être utilisé pour représenter d'autres informations aujourd'hui présentées dans l'étiquette radar. La perception préattentive des informations est rendue possible par ce design. Il n'est plus nécessaire de mémoriser et d'interpréter des informations, tout ceci ayant pour finalité de limiter la charge de travail en mode nominal pour augmenter la capacité à réagir lors de situations inattendues.

2.2.4. Design orienté conscience de la situation

Avant de pouvoir réaliser une interface homme-machine présentant des caractéristiques améliorant la conscience de la situation, il faut au préalable avoir une idée précise des objectifs qu'ont les opérateurs pour réaliser leur travail. Les informations nécessaires et la façon dont elles sont représentées doivent servir de support à la réalisation de ces objectifs. La connaissance de ces objectifs peut être obtenue au travers d'une analyse de la tâche.

Différentes approches existent pour modéliser la tâche ; celle du contrôleur est très complexe, car il poursuit un grand nombre d'objectifs dont les priorités peuvent évoluer en fonction de la situation aérienne. Endsley et Rodgers (1994) ont réalisé une analyse de la tâche très complète concernant les contrôleurs en-route aux États-Unis. La méthode GDTA (Goal Directed Task Analysis) développée par Endsley est une technique d'analyse cognitive de la tâche dont l'objet est de déterminer quelles sont les exigences nécessaires pour assurer celle-ci. GDTA a pour objectif de déterminer de quoi l'opérateur a besoin pour réaliser chaque objectif de sa tâche même si l'information n'est pas forcément disponible compte tenu des limitations de la technologie utilisée. Le processus d'analyse est réalisé à l'aide des documents décrivant la tâche, des observations en situation et des interviews structurées des opérateurs ; ces dernières sont la source principale d'information. Une analyse GDTA permet de mettre en évidence les informations nécessaires au processus de décision dans les environnements complexes comme le contrôle aérien. Elle se focalise sur les informations qui changent dynamiquement durant le déroulement de la tâche et dont les opérateurs ont besoin pour accomplir leur travail. Ces informations dynamiques correspondent aux besoins en termes de conscience de la situation (Situation Awareness : SA) des opérateurs. En définissant les besoins SA des opérateurs, le processus de design de l'interface peut être réalisé en privilégiant la représentation des informations pour maintenir une bonne SA ; on parle alors de SA Oriented

Design. Les interviews doivent être réalisées avec les experts du domaine (Subject Matter Experts), dans notre cas les contrôleurs aériens. La vue synthétique de la GDTA se présente sous forme d'un graphe dont chaque branche est un sous-objectif de la tâche globale. Ce sous objectif entraîne une décision dont le prérequis est une exigence de SA découpée par niveau (niveau 1 : perception, niveau 2 : compréhension, niveau 3 : projection de l'état futur). Au préalable des interviews, l'analyste doit avoir une solide connaissance de l'activité pour faciliter un dialogue technique tout en étant circonspect pour ne pas développer une préconception des objectifs des opérateurs.

La façon dont sont présentées les informations va impacter la conscience de la situation. Elles ne doivent pas être présentées suivant une approche technocentrée, mais plutôt suivant une approche centrée sur l'humain et ses besoins. Basés sur les facteurs impactant la conscience de la situation dans les systèmes complexes, Endsley (2012) a proposé plusieurs principes pour réaliser un design orienté conscience de la situation :

- organiser les informations en fonction des objectifs de la tâche de façon à ce que les informations nécessaires à un objectif soient colocalisées. Les informations nécessaires à l'établissement de la conscience de la situation sont issues de l'analyse GDTA,
- présenter les informations concernant le SA2 pour améliorer et limiter le coût cognitif de la compréhension de l'information. Par exemple dans le contrôle aérien, la tendance de la vitesse verticale peut être suffisante et mieux appropriée que montrer sa valeur numérique,
- donner des outils facilitant la projection de l'état futur du système (SA3),
- donner l'accès à toutes les informations permettant d'établir la conscience de la situation globale (pas d'accès par des interactions compliquées, des fenêtres masquables etc..),
- trouver le meilleur compromis entre un design basé sur les objectifs des opérateurs (goal-driven) et celui basé sur la conscience de la situation globale (data-driven) qui va diriger l'attention des opérateurs sur les objectifs les plus prioritaires. La seconde approche va favoriser l'utilisation de notifications (visuels, sonores, haptique) qui vont détourner l'attention possiblement au détriment de la réalisation de la tâche globale (à n'utiliser que pour activer des objectifs critiques),
- utiliser des notifications avec un bon niveau de saillance pour activer des schémas mentaux associés à la tâche,
- tirer profit des capacités de traitement parallèle des informations des opérateurs (par exemple utiliser des notifications basées sur le visuel, l'haptique ou le son),
- utiliser avec précaution le filtrage d'information. En cas d'un trop grand nombre d'informations à traiter par l'opérateur, les filtrer et les traiter pour faire émerger les informations associées à la SA2 peut être profitable à la conscience de la situation et la charge de travail, mais, en contrepartie, cela peut limiter la réalisation de la SA3 en diminuant la SA globale (voir règle 4). Privilégier donc l'affichage des informations si leur utilité pour la tâche est avérée.

2.3. Les systèmes de notifications visuelles et sonores, et leurs limites

Les supervisions de systèmes critiques ont en commun de montrer des informations permettant de contrôler et surveiller des systèmes. Des informations de type notifications qui peuvent avoir des niveaux de criticité différents sont présentées aux opérateurs pour les aider dans leur tâche. Le terme alarme est le plus souvent utilisé pour les notifications critiques. Il n'en reste pas moins que c'est une classe particulière de notification. Ces notifications peuvent se traduire par des designs visuels spécifiques, des sons ou des feedbacks haptiques. Des systèmes multimodaux utilisant ces trois types de systèmes de notifications pour des supervisions de systèmes critiques ont déjà été mis en service dans des systèmes de transport. Boeing utilise les trois modalités dans ses appareils modernes. C'est aussi le cas dans les voitures haut de gamme chez plusieurs constructeurs. Dans le domaine du contrôle aérien, à ce jour seules les notifications visuelles et sonores sont mises en œuvre dans des systèmes opérationnels. La problématique du partage de l'attention et des capacités des opérateurs humains sont les points essentiels à prendre en compte dans le design de ce type de système. Dans ces travaux, nous nous sommes intéressés aux notifications visuelles et sonores qui sont utilisées aujourd'hui dans le contrôle aérien ; nous traiterons dans les paragraphes suivants l'utilisation de ces deux types de notifications dans les systèmes de supervision en général et dans le domaine aéronautique (aviation et contrôle aérien).

2.3.1. Notifications visuelles

Les notifications visuelles se caractérisent par la présentation d'informations utilisant des designs spécifiques pour attirer l'attention. Ces designs vont utiliser les différentes variables visuelles pour discriminer les notifications du reste des informations affichées.

Un des principaux problèmes dans les systèmes de notification est que des objets apparemment de premier plan dans le champ visuel peuvent parfois échapper à l'attention malgré leur pertinence et leur importance pour la tâche principale (Drew, Vö, & Wolfe, 2013). Cet effet a par ailleurs été mis en évidence dans une expérience très connue de Dan Simons (gorilla effect) sur l'attention sélective dans laquelle on demande au sujet d'analyser dans une vidéo combien de passes avec un ballon sont effectuées dans un groupe, un gorille traverse la scène, il n'est que très rarement perçu par les sujets (The invisible gorilla, Simons & Chabris 2011). Ce phénomène de cécité attentionnelle (Mack & Rock, 1998 ; Most, Scholl, Clifford, & Simons, 2005) peut se produire même si le stimulus est saillant en termes de couleur ou de mouvement, posant la problématique de la conception de la représentation des alertes dans des environnements de systèmes critiques (Mack & Rock, 1998). L'occurrence de la cécité attentionnelle est augmentée par la demande attentionnelle de la tâche (Simons & Chabris, 1999), la charge de la mémoire de travail (Fougnie & Marois, 2007), la nécessité de maintenir l'information dans la mémoire de travail visuospatiale (Todd, Fougnie, & Marois, 2005), et pendant les périodes de forte activité la concurrence attentionnelle des différentes informations présentées (Nikolic & Sarter, 2001 ; Sarter & Woods, 1994). Ainsi, il peut arriver que des objets ou des modifications à une scène

visuelle n'atteignent pas la conscience dans des environnements complexes et exigeants de Commandement et Contrôle (C2) (e.g., Durlach, Kring, & Bowens, 2009). En fait, les opérateurs peuvent même regarder directement un changement essentiel, mais sans le «voir» réellement (e.g., Vachon, Valières, Jones, & Tremblay, 2012). Un objet est plus facilement détecté s'il est proche du centre de l'attention visuospatiale (Most, Simons, Scholl, & Chabris, 2000), mais la proximité n'est pas une condition suffisante pour la détection et il peut toujours être manqué (Newby & Rock, 1998 ; Simons & Chabris, 1999).

Dans des tâches requérant une attention importante, les opérateurs peuvent expérimenter un rétrécissement du champ attentionnel (Chan & Courtney, 1993) et se fixer sur un aspect particulier de leur tâche à l'exclusion d'autres aspects pouvant être du même niveau voire plus important, dans le déroulement de l'ensemble de leur activité (Dehais, Tessier, Christophe, & Reuzeau, 2010). La tunnellation attentionnelle désigne un mécanisme de focalisation excessive de l'attention. Dans des activités de supervision de systèmes critiques, une focalisation excessive sur un des sous-ensembles des paramètres supervisés peut aboutir à des accidents graves comme dans le cas du vol Eastern 401 qui s'est crashé dans les Everglades en 1972 parce que les trois membres d'équipages à bord s'étaient focalisés sur les indicateurs de sortie du train d'atterrissage (l'une des ampoules était grillée). Des contres mesures cognitives (Dehais, Causse, Tremblay 2011 ; Dehais Tessier et Chaudron 2003) se sont révélées efficaces dans ce type de situation.

Un second problème concernant les systèmes de notification est que l'apparition d'une nouvelle information dans le champ visuel peut être faite de façon trop saillante et détourner l'attention au détriment de l'activité en cours si celle-ci est en fait plus prioritaire. Les connaissances sur la vision montrent que certains aspects de la visualisation peuvent susciter une réaction automatique, par exemple la couleur (Bauer, Jolicœur, & Cowan, 1996 ; D'Zmura, 1991), la luminance (Turatto & Galfano, 2000), le mouvement (Faraday & Sutcliffe, 1997), ou l'apparition d'un nouvel objet (Remington, Johnston, & Yantis, 1992 ; Yantis & Jonides, 1996; Franconeri, Hollingworth, & Simons, 2005), ce qui peut diminuer la performance dans la réalisation de la tâche principale puisque moins de ressources attentionnelles sont disponibles pour la tâche en cours. Les interruptions sont un des éléments ayant un effet négatif sur la tâche principale en terme d'erreurs humaines (Reason, 1990), d'anxiété (Bailey et al., 2001), d'augmentation du temps de décision (Hodgetts, Vachon, & Tremblay, 2014), et de diminution de la conscience de la situation (St. John, Smallman, & Manes, 2005). L'impact peut être particulièrement fort si le temps manque pour consolider des aspects importants de la tâche principale avant de basculer sur le traitement de la nouvelle information (Hodgetts & Jones, 2006 ; Trafton, Altmann, Brock, & Mintz, 2003). Ainsi, bien qu'il soit souhaitable que la notification capte l'attention, il est important que ce ne soit pas fait au détriment de toutes les autres tâches. Concevoir et évaluer un système de notification nécessite d'établir un compromis entre le type de design et sa facilité d'intégration et d'assimilation dans l'ensemble des tâches d'un opérateur (Maglio & Campbell, 2000 ; McCrickard & Chewar, 2003).

Dans le cas des notifications visuelles sur écrans multiples, la configuration de l'espace de travail, le nombre et la taille des écrans, peut entraîner la représentation d'informations critiques dans des zones de vision périphé-

riques. Dans ce cas le design utilisé pour leur représentation est crucial pour la réalisation de la tâche. Des études de laboratoire ont montré que des attracteurs visuels en vision périphérique peuvent capturer l'attention lors de processus dirigés par les stimuli (bottom-up), mais la soudaine apparition d'un nouvel objet dans la scène ne va pas nécessairement attirer l'attention si celle-ci est focalisée sur une tâche différente ou un autre endroit (Jonides, 1981 ; Theeuwes, 1991 ; Yantis & Jonides, 1990). De plus, le pouvoir attentionnel d'un nouveau stimulus peut être considérablement atténué dans un contexte de tâche dynamique si l'on ne considère pas la totalité de l'interface et de ses paramètres comme la similarité des couleurs entre la cible et le fond ou le déplacement d'autres d'objets en arrière-plan (Nikolic, Orr, & Sarter, 2004). Les déplacements sont mieux détectés en vision périphérique que la couleur ou des formes (Bartram, Ware, & Calvert, 2001). Des études ont montré l'impact de différents paramètres concernant le mouvement comme la vitesse, l'amplitude, la régularité ou le type de mouvement (e.g., linéaire, zoom, clignotement, disparition) sur la vitesse et la précision de la détection d'une notification (Bartram, Ware, & Calvert, 2003 ; McCrickard, Catrambone, Chewar, & Stasko, 2003 ; McCrickard, Catrambone, & Stasko, 2001 ; Ware, Bonner, Cater, & Knight, 1992). Les alertes visuelles ont été étudiées spécifiquement dans des environnements de Commandement et Contrôle (C2) ; par exemple, dans l'utilisation d'une tâche de caractérisation tactique, l'apparence de la bordure de l'écran était une alerte plus efficace qu'une barre d'icônes et cela qu'elle soit statique ou clignotante, aucune différence sur la vitesse de détection n'a été démontrée (Crebolder, 2012). En revanche, dans le cas de l'utilisation du clignotement au lieu d'une alerte statique, un impact positif sur la mémoire prospective et la détection des conflits dans une tâche ATC a été démontré (Loft, Smith, & Bhaskara, 2011). Dans le contrôle aérien, la couleur ainsi que les animations de couleurs (clignotements) sur des champs textes spécifiques sont le plus souvent utilisées pour représenter les notifications visuelles. Par exemple, lorsqu'un des systèmes servant à alimenter l'image radar tombe en panne, une information de panne avec une couleur spécifique apparaît dans un coin de l'écran. Lorsque deux avions ont des trajectoires aboutissant à une séparation inférieure aux normes de sécurité (5 miles nautiques en horizontal et 1000 ft en vertical), un champ texte avec l'inscription STCA ou ALRT s'affiche en clignotant dans l'étiquette radar de chaque avion. Des guides de conception ont été proposés pour la création de signaux visuels pour les images radars (Athènes et al., 2000), mais il est clairement nécessaire de faire de nouvelles études concernant le pouvoir attentionnel, le coût cognitif associé et les éventuelles non-détections des notifications dans un contexte de visualisations contenant un grand nombre d'informations et sur des écrans de grandes dimensions.

2.3.2. Notifications sonores

Dans le contrôle aérien français, des notifications sonores sont utilisées sur des positions de travail destinées au contrôle d'approche. Le contrôle d'approche est très différent du contrôle en-route ; le trafic aérien y est très évolutif, l'environnement (relief, conditions météo, zones prohibées) est crucial pour la sécurité des aéronefs ; des systèmes d'alerte spécifiques y sont installés. MSAW (détection de collisions avec le relief) et APW (détection d'intrusion dans une zone prohibée) émettent un son lorsque la trajectoire

d'un aéronef concorde avec les zones de protections décrites dans le système. Ces sons sont diffusés dans l'ensemble de la salle d'approche. Le contrôleur concerné par l'alerte c'est-à-dire celui qui est responsable du volume dans lequel se trouve l'avion problématique est sensé associer ce son avec les informations présentées sur son image radar. Il est donc difficile d'associer directement le son avec les vols qu'ils gèrent, la notification visuelle associée sur la supervision radar est donc cruciale pour la perception correcte du problème et limite l'intérêt de la notification sonore.

Dans les cockpits, des signaux sonores peuvent être émis par un nombre important de sous-systèmes. En aviation générale, l'avertisseur de décrochage (perte de portance pouvant aller à la chute d'un avion) ou l'oubli de la sortie du train d'atterrissage en dessous d'une certaine vitesse se traduisent par l'émission d'un son strident. Dans les cockpits modernes, différents systèmes utilisent le son, parfois en complément de notifications visuelles ou haptiques pour prévenir les équipages d'alertes critiques pour le déroulement du vol. Le système TCAS (Traffic Collision Avoidance System) a pour but d'éviter les collisions entre aéronefs et notifie les actions à prendre en cas de collision imminente. Grâce à des messages sonores (descend descend) donnant la manœuvre à suivre et une information visuelle fournie sur le PFD (Primary Flight Display) donnant les valeurs à respecter pour résoudre le conflit. Le GPWS (Ground Proximity Warning System) a pour but d'éviter qu'un avion ne percute le sol, un message audio est émis lorsque celui-ci s'en rapproche trop (WHOOOP, WHOOOP - TERRAIN, TERRAIN - PULL UP, PULL UP). Ces divers systèmes n'ont pas été conçus de façon intégrée pour un appareil. Si une notification sonore doit être prioritaire ou mieux mise en avant en raison de sa criticité, les pilotes devront faire le tri et analyser la situation pour réagir. La problématique du tri des notifications par niveau d'intérêt dans le domaine des notifications sonores se pose de la même façon que dans celui des notifications visuelles.

Des notifications sonores sont utilisées dans un grand nombre de systèmes de supervision, elles sont en général associées à des informations critiques du type alarmes ou avertissements, mais elles peuvent aussi être utilisées pour donner des informations sur une tâche en cours d'exécution (feedback). Les notifications sonores ont de bonnes propriétés puisqu'elles sont perçues plus rapidement que les notifications visuelles (Wheale, 1981). Elles sont par ailleurs perçues sans nécessiter de focaliser son regard sur une partie des interfaces de supervision ou d'effectuer un mouvement de tête dans la direction de la source sonore, mais dans certaines conditions, elles peuvent diminuer l'efficacité des opérateurs en devenant distrayantes ou stressantes (Edworthy, 1991). À l'instar des variables visuelles (Dallet & Veitl 2001), plusieurs variables permettent de concevoir et caractériser les notifications sonores comme l'enveloppe d'amplitude, le timbre ; le son peut être caractérisé par sa représentation spectrale (analyse de Fourier) qui permet de voir la contribution de chaque onde sonore sinusoïdale par son amplitude au signal sonore global. Une étude de 2004 (Arrabito, Mondor & Kent) dans le domaine du cockpit a montré que lorsqu'elles sont conçues dans le but de montrer un niveau d'urgence différent en utilisant certaines propriétés acoustiques du signal sonore, celui-ci était correctement perçu par les pilotes. Dans le domaine du contrôle aérien, plusieurs auteurs décrivent l'intégration des notifications sonores, les problèmes de design associés, les problèmes rencontrés lorsque plusieurs types différents d'alarmes sonores

sont diffusées et la difficulté liée à leur interprétation et discrimination (Cabrera, Ferguson, et Laing, 2006 ; Ahlstrom, V., Panjwani 2003). Selon la théorie de la charge perceptuelle de Lavie (1995), il a été montré que les tâches nécessitant un niveau élevé d'attention laissent peu de capacité de traitement pour les distracteurs que sont les alarmes visuelles ou sonores (Lavie 2010, Hughes, Hurlstone, Marsh, Vachon, & Jones, 2012). De nombreux cas d'incidents dans le monde aéronautique montrent que des signaux sonores dans des situations de forte charge de travail ne sont pas perçus par les opérateurs (Bliss, 2003). L'absence de réaction ou la non-perception de notifications sonores fait écho au mécanisme de cécité attentionnelle décrit précédemment, et est qualifiée de surdité attentionnelle (Koreimann, Strau, & Vitouch, 2009 ; McDonnald & Lavie, 2011) ; ce mécanisme a été mis en évidence lors d'une étude sur la perception des notifications sonores dans le cockpit (Dehais et al., 2013). Lorsque la tâche visuelle est complexe, le nombre d'évènements mettant en jeux la surdité attentionnelle est plus élevé (Macdonald et Lavie, 2011).

2.4. Résumé

Les visualisations de type supervision sont des interfaces homme-machine particulières ; elles sont associées à des tâches de suivi d'information et de prise de décision dans des environnements complexes et dynamiques. Lors de la conception d'un système de visualisation, le designer pourra manipuler plusieurs variables visuelles ayant des propriétés et un pouvoir de sélection différent.

Différentes théories visent à guider le designer pour faciliter l'acquisition d'informations par l'opérateur. Certaines propriétés visuelles rendent l'information représentée plus aisément interprétée par le cerveau, limitant la charge cognitive induite par leur acquisition. On parle alors de perception préattentive. De la même façon, il est possible de limiter la charge cognitive en concevant des objets en utilisant la théorie de la Gestalt et aider le cerveau à établir des surformes sur un ensemble d'items graphiques.

Les théories de Gibson sur la perception écologique de l'information ont été mises en œuvre dans des méthodes de design dites 'écologiques' visant à limiter la charge cognitive pour la perception d'information en s'appuyant sur des représentations ou interactions que l'homme trouve habituellement dans son environnement.

D'une façon plus globale, en s'intéressant aux informations nécessaires à l'établissement d'une bonne conscience de la situation des opérateurs, Endsley propose une méthode de design orientée conscience de la situation basée sur plusieurs règles de conception et une analyse de la tâche des opérateurs.

Les systèmes de supervision peuvent délivrer des informations dynamiques et les notifier aux opérateurs qui devront les prendre en compte dans leur tâche. Les informations notifiées peuvent avoir des niveaux de criticités variables et les designs employés, qu'ils soient visuels sonores ou haptiques, ont des pouvoirs attentionnels différents. Une adéquation entre pouvoir attentionnel et niveau de criticité doit être trouvée. Une notification est une interruption, elle a un coût cognitif dans le déroulement de la tâche. Par ailleurs, différentes études montrent l'impact négatif de la charge de travail sur la perception des notifications. D'autres études mettent en évidence la charge visuoperceptive et les interactions entre les différentes modalités de notification.

En résumé, la conception d'une visualisation de supervision doit viser à faciliter l'acquisition d'information avec un coût cognitif le plus limité possible en utilisant les différentes théories sur la perception pour réaliser son design pour que l'opérateur soit à même de traiter les informations rares que sont les notifications. Les types de notifications utilisées doivent être en rapport avec l'importance de l'information notifiée dans le déroulement de la tâche pour limiter leur impact. Les quelques exemples d'accidents graves dans le domaine aéronautique cités précédemment mettent en évidence plusieurs phénomènes et problématiques :

- La transmission d'informations visuelles à l'humain doit être optimisée pour adapter leur impact au déroulement de la tâche de supervision,
- la multiplication des notifications sonores et visuelles rend leur hiérarchisation et interprétation problématique,

- lorsque les sons de différentes notifications sont proches, des confusions sont à craindre lors de situations dégradées,
- l'humain a des capacités cognitives limitées, le traitement des notifications visuelles et sonores est disruptif de la tâche en cours posant le problème de la reprise sur interruption,
- lors de situations dégradées et de niveau de charge cognitive importante, des phénomènes de surdit  attentionnelle peuvent rendre les notifications sonores inop rantes.

La repr sentation d'information et le design des diff rentes modalit s de notification doivent prendre en compte les capacit s et les limites du syst me perceptif humain. Il est n cessaire d'avoir une conception globale des diff rents types de notifications susceptibles d' tre  mises pour  viter la surcharge cognitive associ e   leur traitement. Un design bas  sur la perception pr attentive et  cologique de l'information diminue la charge de travail li e   son traitement et am liore la conscience de la situation. Les m thodes de design orient  SA permettent de faire  merger les informations n cessaires   la r alisation des objectifs de la t che, la repr sentation de ces informations. Ces travaux visent   am liorer la conscience de la situation des contr leurs a riens au travers de la transmission d'informations par la visualisation.

L'augmentation du trafic a rien et par voie de cons quence des informations   afficher va  tre limit  par la capacit  de perception et de traitement des contr leurs. Notre approche durant ces travaux est de tenter d'adapter la visualisation en filtrant les informations pertinentes gr ce au contexte. Ce type de filtrage est une forme d'automatisation et d'adaptation de la visualisation. Le chapitre suivant a pour but de pr senter l'automatisation et son impact sur l'humain ainsi que les perspectives concernant les syst mes adaptatifs.

3. AUTOMATISATION ET ADAPTATION

3.1. Introduction

Les objectifs d'augmentation de capacité des secteurs contrôlés ne pourront être atteints qu'à l'aide d'une automatisation accrue. Selon Sheridan et Parasuraman (2006) l'automatisation se définit comme l'ensemble des processus artificiels allant de perception des éléments de l'environnement jusqu'à l'action sur celui-ci. La complexité et la variabilité du trafic aérien rendent son automatisation totale inatteignable aujourd'hui et très certainement à long terme. L'automatisation partielle, qui est la solution envisagée pour l'évolution du système de gestion du trafic aérien, présente un certain nombre de problèmes, dont le couplage entre l'humain et l'automate. Si elle est considérée comme la seule solution pour augmenter la capacité des systèmes voire de limiter l'erreur humaine ; l'automatisation apporte aussi de nombreux défis, dont celui du rôle de l'humain et de son interaction avec le système automatisé à fortiori lorsque ce dernier est défaillant. Plusieurs études d'analyse d'accident dans le domaine aéronautique montrent la problématique des interactions homme-système et l'impact de l'automatisation (Parasuraman et Riley, 1997 ; Byrne 2003). Les systèmes automatisés n'apportent pas forcément un bénéfice sur la sécurité ou la charge de travail. Leur introduction peut aboutir à une évolution du rôle d'opérateur vers celui de superviseur et diminuer sa capacité à détecter et traiter les problèmes critiques.

Plusieurs approches visent à améliorer le couplage humain-automate en agissant sur leur relation ou le niveau d'automatisation. L'automatisation adaptative est la nouvelle voie d'intégration de l'automatisation dans le couple homme-système. Son but est d'offrir une flexibilité qui permettrait de prendre en compte l'humain dans le fonctionnement du système et agir sur le ou les automates pour améliorer ce couplage.

3.2. Niveaux d'automatisation

Les travaux sur l'automatisation ont introduit la notion de «level of automation» ou niveau d'automatisation. Plusieurs niveaux peuvent être atteints ; différentes classifications existent dans la littérature, dont l'échelle LOA (Level Of Automation) de Parasuraman, Sheridan et Wickens (2000) définissant quatre étapes dans le traitement de l'information :

- l'acquisition d'information, N1
- l'analyse d'information, N2
- la prise de décision, N3
- l'action, N4.

Les systèmes automatisés peuvent être mis en œuvre à chaque étape pour aider l'humain. Au niveau le plus élémentaire, c'est le cas du filtrage utilisé

dans les visualisations de supervision radar. Les avions sont identifiés par des codes à quatre chiffres, certains filtrages permettent de supprimer l'affichage des avions de tourisme depuis les années 1980 (code transpondeur 7000). D'autres types de filtrage vont permettre de ségréguer visuellement les flux de trafic. Sur la plateforme de Roissy, la variable visuelle couleur est utilisée pour différencier les flux de départ ou d'arrivée ou les avions se posant sur la piste nord ou la piste sud (Roissy dispose de deux pistes de décollage et deux pistes d'atterrissage). Le STCA peut être considéré comme un automate de N2 puisqu'il détecte en analysant les trajectoires des avions ceux qui risquent de se trouver dans le futur sous les normes de séparation horizontales et verticales. Un automate de N3 proposera des solutions de résolutions de problèmes. Dans le cas du contrôle aérien, cela pourrait être des propositions de consignes en niveau ou en cap afin de gérer un conflit. Enfin un automate de N4 effectuera lui-même les actions correctrices des problèmes détectés. Dans l'aviation commerciale, l'Airbus A320 a été le premier appareil à disposer de ce type d'automate. Lorsque l'avion a une incidence et une vitesse l'amenant à décrocher, celle-ci est corrigée automatiquement et la puissance envoyée dans les moteurs pour remettre l'appareil dans une configuration viable.

Dans le domaine du trafic aérien français, tous les niveaux d'automatisation ont été testés, des niveaux les plus basiques (tri d'information, filtrages sur la visualisation radar) aux plus élevés c'est-à-dire « l'action » (N4) comme avec le projet Erasmus (www.atm-erasmus.com) et plus tard le projet SESAR WP4.7.2 dans lequel s'insèrent ces travaux. L'automate présent dans ce projet a pour but de dé-conflictualiser le trafic aérien en envoyant aux avions (en simulation) de légères régulations de vitesse en amont des conflits identifiés. L'automate interagit donc directement avec les avions sans intervention de l'humain. L'algorithme n'étant pas capable de résoudre tous les conflits, l'opérateur humain devait gérer les conflits résiduels à l'aide d'un agenda de problèmes à résoudre. Dans ce même projet, des outils d'aide à la décision ont été testés pour montrer au contrôleur l'impact d'un ordre donné à un avion sur la situation globale (fonction What if). D'autres visaient à améliorer la projection de l'état futur du système à l'aide d'un outil d'extrapolation.

3.3. Impact de l'automatisation sur l'humain

La problématique de l'automatisation et le changement de rôle du superviseur qui va toucher le contrôleur aérien dans le futur se pose dans le domaine aéronautique depuis plusieurs décennies. À ce titre, l'expérience acquise dans l'aviation qui est proche du contrôle aérien peut aider à anticiper les problèmes qui se poseront dans le futur. L'automatisation partielle du cockpit est en œuvre depuis l'introduction du pilote automatique en 1914 jusqu'à l'introduction du système automatisé de gestion du vol (FMS), utilisé aujourd'hui dans tous les avions de transport de passager. Les analyses d'incidents dans ce domaine montrent que si l'automatisation n'est pas mauvaise en soi, le changement du rôle des opérateurs dans le système provoque l'apparition de nouvelles erreurs, tout en corrigeant d'autres et que les opérateurs se déqualifient. À titre d'exemple, un incident impliquant l'automatisation et des actions inappropriées du pilote liées à son incompréhension de l'état du système s'est produit en 1994 sur le vol Taron 381. Cet Airbus A310 était en finale à Orly, le pilote avait désactivé

le pilote automatique, mais laissé les automanettes enclenchées (régime moteur automatique sur ralenti en approche) ; l'avion arrivait un peu trop vite en phase de capture de l'ILS (Instrument Landing System) et le pilote a décidé de sortir les traînées à la limite des vitesses maximales autorisées ; le passage des volets à 20° a été effectué au-dessus de cette vitesse ce qui a entraîné le basculement automatique des automanettes dans le mode CLIMB, dans le même temps certainement en raison d'un mauvais réflexe, l'un des pilotes a mis l'avion à cabrer avec le TRIM (actionneur sur le manche) à 13° d'assiette. L'avion s'est alors trouvé moteur à fond assiette à cabrer sans que les pilotes en comprennent la raison. Plusieurs actions ont par la suite été faites par l'équipage qui montre leur incompréhension totale de l'état de l'appareil et des automates sensés les aider (BEA rapport Taron 381 940924, 1994). Toutes les notifications visuelles et sonores émises par l'appareil n'ont visiblement pas été prises en compte et commentées par l'équipage durant l'incident.

Les problèmes concernant l'interaction entre l'homme et le ou les automates ont été identifiés et étudiés depuis les années 80 (Bailey, 1983) et plus récemment dans des études concernant le monde aéronautique (Billings, 1996 ; Parasuraman et Riley, 1997 ; Parasuraman et Byrne, 2003). D. Norman a enquêté sur l'adéquation de la rétroaction et de l'interaction avec les humains qui doivent contrôler la conduite globale d'une tâche. Il a montré les difficultés auxquelles les contrôleurs humains peuvent faire face en cas de défaillance ou de conditions de fonctionnement particulières conduisant à des situations qui dépassent les capacités du système automatisé. En ce domaine, de nombreux problèmes ont été étudiés dans la littérature spécialisée en facteurs humains concernant à la fois les problèmes liés aux outils (interaction, design) ou à l'humain lui-même utilisant ces outils (désengagement de la tâche, hypo vigilance, charge trop élevée, incapacité à reprendre la main lorsque l'automate est en défaut, confiance dans le système, déqualification).

L'étude de l'impact de l'automatisation sur l'humain a été étudiée du point de vue la charge de travail, de la mémorisation et de la détection de conflit dans le trafic aérien par Metzger et Parasuraman (2001 & 2005). Ils ont montré qu'un niveau élevé d'automatisation diminuait la performance dans la détection des conflits et mis en évidence l'importance de la fiabilité des outils d'aide.

Des accidents dramatiques récents ont montré comment, dans un contexte d'automatisation, les êtres humains peuvent rapidement passer d'une charge de travail faible à trop élevée, avec une trop grande quantité d'informations à traiter et analyser (Martins, Soares 2012). La relation entre la charge mentale et la performance des pilotes est souvent discutée dans la littérature sur les facteurs humains. Proctor et Zandt (2011) supposent une fonction en forme de U inversé entre la charge mentale et la performance (Figure 15 : performance en fonction de la charge de travail). Selon le niveau de la charge mentale, la performance varie. Si la charge mentale est trop basse ou trop élevée, la performance n'est pas acceptable (Proctor, 11).

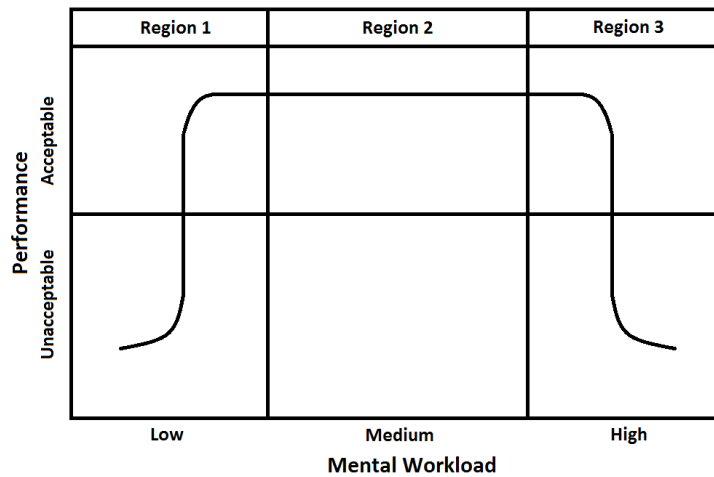


Figure 15 : performance en fonction de la charge de travail

La charge de travail, si l'on est dans la région 3 (Figure 15 : performance en fonction de la charge de travail), le stress ou la fatigue peuvent significativement réduire la conscience de la situation en réduisant la capacité de l'opérateur à traiter les informations. En condition de stress, les opérateurs de supervision portent moins attention aux informations disponibles en vision périphérique, sont plus sujets à la tunnellation attentionnelle, également moins organisés dans le balayage des informations et sont plus enclins à prendre une décision sans prendre en compte tous les paramètres. La fatigue liée par exemple à des cycles de travail inadaptés (contrôleurs et pilotes sont soumis à des horaires les amenant à travailler la nuit), peut entraîner des temps de réponse élevés aux notifications et des difficultés à la mémorisation des informations à prendre en compte pour établir un bon niveau de conscience de la situation.

Dans un contexte d'automatisation poussée, la charge de travail peut devenir suffisamment faible pour provoquer un désengagement de la tâche. Le terme 'out-of-the-loop' ou hors de la boucle a été défini par Wickens et Holland (2000) et traduit le désengagement de la tâche d'un opérateur dont la vigilance a diminué. L'automatisation peut aboutir à diminuer la charge cognitive, si cela peut être judicieux en période de forte charge de travail, en cas de faible charge la conscience de la situation peut s'en trouver affectée. Par ailleurs, si l'automate fait défaut, l'opérateur peut ne plus être en mesure de reprendre la situation en main parce ce qu'il n'est plus capable de détecter et d'interpréter le problème ou d'agir correctement et rapidement pour reprendre la main.

3.4. Confiance dans le système

Les différentes études portant sur les niveaux d'automatisation élevés (par exemple sur l'intégration d'informations, les aides à la décision) montrent que si un haut niveau de fiabilité ne peut être garanti, surtout dans les systèmes critiques, les inconvénients pour les opérateurs sont plus importants que les bénéfices observés. Le syndrome de 'Cry-Wolf' par analogie au conte, décrit la non-réaction d'un opérateur à des alarmes trop fréquentes ou trop souvent erronées. Dans ce type de situation, l'opérateur finit par filtrer inconsciemment ces alertes rendant l'automate les ayant produites

inutile. Dans tout système d'automatisation produisant des informations d'aide, la fiabilité et la qualité des éléments remontés par l'automate sont cruciales si l'on veut que l'opérateur en tire parti (Lee & See, 2004 ; Parasuraman, Riley, 1997). Dans le domaine du contrôle aérien, l'automate le plus généralisé dans le monde est le filet de sauvegarde (STCA), en 2009 Wickens, Rice, Keller, Hutchins Hughes et Clayton ont réalisé une étude basée sur l'analyse de 495 déclenchements de STCA aux USA dont 45% étaient fausses après analyse, l'effet Cry-Wolf n'a pas été démontré dans cette étude. Il n'en reste pas moins que ce taux de fausses alarmes a été identifié dans un rapport du NTSB (National Transportation Safety Board) par les contrôleurs comme une source importante de nuisances et que la méfiance dans le système d'alarme a un impact sur la charge de travail (Bliss, Dunn 2000). Les notifications d'alertes graduées sont tout de même plus efficaces, y compris dans le cas où le taux de fausses alertes est important, ce qui diminue la confiance dans le système (Bliss & Dunn, 2000).

3.5. Automatisation et conscience de la situation

Dans les environnements complexes et dynamiques, la conscience de la situation est au centre du processus de décision. La performance d'un opérateur de supervision dans la réalisation de sa tâche est intimement liée à sa capacité à avoir une image mentale claire de la situation courante et de son état futur en disposant de ressources cognitives libres pour gérer les situations inattendues. Des décisions inappropriées peuvent être prises si l'opérateur a une mauvaise représentation de l'état de l'automate l'assistant dans sa tâche.

Le principal résultat attendu de l'automatisation est d'augmenter la capacité de traitement des opérateurs en diminuant leur charge de travail. Peu d'études s'intéressent à l'impact de l'automatisation sur la conscience de la situation des opérateurs. Dans l'une d'entre elles, concernant les accidents et incidents dans le domaine aéronautique (contrôle aérien et aviation), Endsley et Jones ont trouvé que 35% des cas de défaillance de la conscience de la situation concernaient l'incapacité à suivre tous les paramètres. Un nombre trop important de données à traiter peut amener à une dégradation importante de la conscience de la situation. D'autres études portant sur l'analyse de la performance ou de la conscience de la situation dans une tâche de contrôle aérien ont été réalisées par Kaber et Endsley (2004). L'étude de Kaber et al. portait sur une simulation de contrôle aérien se déroulant dans le contexte des quatre niveaux d'automatisation (acquisition d'information, analyse d'information, prise de décision et action automatisée). L'analyse de la conscience de la situation a été réalisée avec la méthode SAGAT (Situation Awareness Global Assessment Technique). Cette étude a permis de démontrer l'intérêt de l'automatisation pour l'amélioration de la conscience de la situation et de la performance. Elle a aussi montré que le bénéfice de l'automatisation s'appliquait essentiellement à l'acquisition d'informations ou à l'action, l'automatisation des fonctions cognitives (analyse d'information) quant à elle augmentant la charge de travail. Ces résultats doivent néanmoins être nuancés dans la mesure où le design de la représentation de l'information, l'interaction pour réaliser les actions, ainsi que les outils mis à disposition du contrôleur pour l'analyse et la prise de décision, sont cruciaux. Des outils mal conçus aboutissent fatalement à de mauvais résultats.

3.6. Interrompre à quel prix ?

Dans de nombreuses situations, un évènement extérieur vient interrompre un utilisateur en train de réaliser une tâche l'obligeant à gérer une interruption rapidement. Dans certains cas un avertissement pourra prévenir d'une interruption future (ex. sonnerie de téléphone). Le temps entre cet avertissement et la réalisation de l'interruption est qualifié d'interruption lag, celui pour repasser à la tâche principale à l'issue de l'interruption resumption lag. Ces deux mesures ont été utilisées dans la littérature concernant l'interruption et plus particulièrement le resumption lag qui est le plus fréquemment utilisé comme indicateur de performance.

Des études récentes ont montré que plus la charge de travail est élevée plus il est difficile de récupérer d'une interruption. Iqbal et al. (2005) ont validé cette hypothèse en étudiant trois tâches principales interrompues périodiquement et prouvé que le temps pour rebasculer sur la tâche principale était plus important lorsque les utilisateurs étaient interrompus pendant une des sous-tâches plutôt qu'entre elles et qu'un index basé sur la charge de travail permettrait de déterminer le moment judicieux pour interrompre. Cette étude s'est appuyée sur l'analyse des variations de diamètre pupillaire pour estimer la charge de travail. Toujours dans le domaine du desktop, Hurter & al. (2012) proposent d'utiliser les moments pendant lesquels normalement une progress bar est affichée pour proposer des tâches secondaires à l'utilisateur, les résultats montrent que les utilisateurs sont satisfaits d'utiliser le système pendant le temps mort de la tâche principale.

Un niveau moyen d'automatisation pour l'acquisition d'informations consiste à les organiser en fonction de certains critères tout en maintenant l'affichage de toutes les données brutes, un niveau plus élevé consiste à en filtrer l'affichage. Dans ces deux types de LOA, des erreurs sur la mise en évidence d'informations, en rendant par exemple une information trop saillante par rapport à sa pertinence opérationnelle ou de mauvais filtrages, peuvent dégrader la performance des opérateurs (Metzger et Parasuraman 2005). Endsley, Bolte et Jones (2003) recommandent de rendre ce type de fonctionnalités déclenchables par l'utilisateur plutôt que par un système automatique. De la saillance des informations présentées aux opérateurs de supervision va dépendre leur interprétation, la célérité avec laquelle elles vont être prises en compte et la capacité de l'opérateur à revenir à sa tâche précédente.

La concurrence des différents systèmes de notifications, visuelles, sonores et haptiques peut considérablement dégrader la conscience de la situation et provoquer un accident. Dans le contrôle aérien en-route en France, seules les notifications visuelles sont utilisées et la saillance des informations notifiée est remise en question dans les rapports d'analyse d'incident et les retours d'expérience. Dans le cockpit où ces trois types de notifications peuvent être mises en œuvre, des analyses d'accidents montrent la difficulté pour l'opérateur de reprendre en main une situation critique lorsqu'il est interrompu par des systèmes de notifications concurrents. Dans le cas du vol Aeroperù 603, un défaut de mesure de pression atmosphérique (qui sert à mesurer la vitesse et l'altitude) a provoqué un défaut du système de management de vol et des différents systèmes d'alerte de l'appareil. Le FMS a remonté des alertes de survitesse et de décrochage provoquant des alertes sonores

et des vibrations dans le manche, l'avion se rapprochant du sol, le GPWS (Ground Proximity Warning System) c'est lui aussi mis à envoyer des notifications sonores. Les pilotes surchargés de notifications diverses accordant une confiance excessive à ces différents automates n'ont pas su interpréter l'état de leur avion qui s'est crashé en mer.

Interrompre a un coût, celui-ci peut devenir rédhibitoire lors de situations critiques et empêcher l'opérateur de se faire une idée globale de l'état du système qu'il supervise.

3.7. Modalités d'initiative et automatisation adaptative

Trois types de modalités d'initiatives sont susceptibles de définir la relation humain-automate : l'autonomie adaptative, l'autonomie ajustable et l'initiative mixte (Hardin et Goodrich, 2009). Le mode d'initiative concerne la dynamique de l'autonomie c'est dire l'aptitude qu'ont l'humain ou l'automate de modifier le niveau d'autonomie de l'automate. Quel que soit ce mode d'autonomie, l'opérateur humain doit avoir une connaissance des capacités de l'automate qui est sensé l'aider dans sa tâche et doit être capable de déterminer son état futur ; sa conscience de la situation doit donc s'élargir au fonctionnement des automates devant l'aider dans sa tâche. De la même façon, si l'automate est autonome, il doit avoir un modèle des capacités de l'opérateur humain et de son état.

« L'autonomie adaptative » donne à l'automate uniquement la capacité de changer son niveau d'autonomie. Ce type de modalité peut être mise en place grâce par exemple à l'utilisation de données physiologiques pour déclencher le changement du niveau d'automatisation d'un agent intelligent (Byrne & Parasuraman, 1996 ; Parasuraman et al., 1992). Kaber & Endsley (2004) ont montré l'intérêt de l'autonomie adaptative en étudiant son effet sur la charge de travail et la conscience de la situation.

« L'autonomie ajustable » (Bradshaw et al., 2003) donne uniquement à l'humain la capacité de changer le niveau d'autonomie de l'automate. Dans le cas de l'autonomie adaptative un système expert prend la décision du changement de niveau d'automatisation. Ce concept laisse le contrôle à l'humain du changement. Le principal défaut de cette méthode est qu'elle repose sur l'humain et sa conscience de la situation qui peut être mise en défaut. Son avantage est de maintenir l'humain dans la boucle de décision.

« L'initiative mixte » est caractérisée par l'aptitude de l'humain et de l'automate à modifier le niveau d'autonomie de l'automate (Horvitz, 1999 ; Hardin et al., 2009 ; Sellner et al., 2006), l'autorité est partagée entre l'humain et l'automate. L'initiative mixte a pour but de capitaliser sur les avantages de deux solutions précédentes. Elle a été mise en œuvre dans différents domaines comme le contrôle de robot ou l'adaptation de Microsoft Office par Allen et Horvitz (1999).

Une problématique cruciale est de déterminer les critères pour modifier l'interaction homme-automate. Plusieurs stratégies d'adaptation ont été proposées comme celle de Scerbo, Freeman et Mikulka (2003) qui consiste à pouvoir changer dynamiquement le niveau d'automatisation grâce à des événements durant l'activité comme le déclenchement d'une situation d'urgence ou une tâche spécifique à réaliser ou bien grâce à un système d'aide

intelligent analysant la situation et les objectifs de la tâche pour adapter le système (Miller et Hannen, 1999). Une autre stratégie consisterait à répartir les tâches entre opérateur et système pour limiter les interactions homme-système. La dernière approche, qui se développe ces dernières années avec l'apparition de capteurs physiologiques performants et économiques, prendrait en compte l'état de l'opérateur pour adapter le niveau d'automatisation.

Dans le contexte de l'automatisation croissante de la gestion du trafic aérien et des fonctions avancées sur les postes de supervision, il serait extrêmement intéressant d'utiliser des informations sur l'état du contrôleur aérien pour améliorer la transmission d'informations. L'état de vigilance, le niveau de charge cognitive, l'éventuel état cognitif dans lequel il peut être engagé comme la tunnelisation attentionnelle seraient pertinents pour adapter la représentation des informations et ainsi l'aider à faire le tri pour éviter des incidents concernant le niveau 1 de la conscience de la situation SA1 (perception des éléments de l'environnement dans le temps et l'espace).

Ainsi une approche alternative à Scerbo, Freeman et Mikulka est l'utilisation de moyens de mesure psychophysiologiques pour déclencher de l'automatisation adaptative. Parasuraman évoque dans (Parasuraman, 1996) le rôle de ces mesures, et présente quelques résultats expérimentaux dans un environnement de laboratoire. Ces résultats montrent comment cette approche peut être particulièrement utile dans la prévention de la détérioration de la performance dans des conditions de sous-charge qui peuvent accompagner l'automatisation. De nombreuses études ont eu pour but d'évaluer cet état cognitif en utilisant des moyens de mesures le moins intrusifs possibles. Différents moyens de mesure physiologiques ont été utilisés lors de simulation ATC comme la conductivité de la peau (Boucsein, 1992), la fréquence cardiaque (Oehme, 2002), la pression artérielle, la respiration, et le suivi du regard (Fitts, 50). D'autres moyens permettent d'avoir une estimation du déroulement de la tâche pour l'opérateur comme des mesures de performance associés à l'instrumentation de la visualisation de supervision ou une analyse du trafic aérien géré par le contrôleur suivant des critères de complexité et de performance.

Lors de travaux dans le domaine de l'aviation militaire Wilson et al. (2000) ont utilisé des méthodes neurométriques, c'est-à-dire l'analyse en temps réel de l'activité du cerveau. Ils ont surveillé les paramètres neurophysiologiques d'un opérateur dans un cadre expérimental pour mettre en œuvre un système adaptatif d'aide à l'opérateur (Wilson 2000, 2003, 2007). La tâche a été adaptée en réduisant le nombre de sous-tâches lorsque des niveaux élevés de charges de travail ont été détectés. Ils ont également mis en évidence la manière dont le niveau de charge de travail et l'activité cognitive peuvent être classifiés lorsque des mesures neurophysiologiques sont employées. Des résultats comparables ont été obtenus récemment lors d'une étude pour l'aviation militaire italienne qui a également porté sur les interactions entre les pilotes (Astolfi 2011). Dans une étude récente, Parasuraman (2008) résume les perspectives de recherche et montre comment les développements dans le domaine de la neuro-imagerie pourraient aider à évaluer la charge de travail dans les tâches complexes et déclencher l'automatisation adaptative.

L'automatisation adaptative est aussi évaluée dans le domaine de l'ATC au travers d'un projet SESAR conduit par la NLR (National Aerospace Labora-

tory of the Netherlands) Aviator II, pour évaluer l'utilisation d'un système de gestion de la séquence d'arrivée dans un aéroport qui serait capable de s'adapter à la charge de travail perçue (et non mesurée) par l'opérateur. Les premiers résultats montrent un bénéfice significatif dans l'activité des contrôleurs (Uebbing-Rumke, 2012).

3.8. Résumé

L'utilisation ou pas d'automatisation dans les systèmes de supervision n'est plus un choix, mais une obligation ; la complexité de ces systèmes, la quantité d'information à mémoriser et à prendre en compte, impose désormais l'utilisation de systèmes d'aide.

Dans le chapitre 1, nous avons pu voir que les analyses d'incidents dans le contrôle aérien et leur corrélation aux différents niveaux de conscience de la situation, montrent que la perception et la mémorisation des informations constituent une part importante des erreurs humaines. Dans le domaine de l'automatisation, c'est dans les premiers niveaux (N1, N2) d'automatisation, c'est-à-dire des automates dédiés à l'acquisition et l'analyse des informations que les bénéfices sont les plus importants et le couplage avec l'humain le plus aisé à réaliser.

Trois types de relation humain-automate peuvent être définis « L'autonomie adaptative » (l'automate uniquement peut changer son niveau d'autonomie), « L'autonomie ajustable » (donne uniquement à l'humain la capacité de changer le niveau d'autonomie de l'automate) et « L'initiative mixte » (l'humain et l'automate peuvent modifier le niveau d'autonomie de l'automate). L'étude de l'impact de l'automatisation sur l'humain dans le contrôle aérien a montré qu'un niveau élevé d'automatisation diminuait la performance dans la détection des conflits et mis en évidence l'importance de la fiabilité des outils d'aide. L'automatisation peut aboutir à une réduction de la charge cognitive ce qui est positif en période de forte charge de travail, mais peut dans le cas contraire provoquer un désengagement de la tâche (out-of-the-loop syndrom). Par ailleurs, dans le cas des niveaux d'automatisation élevés (N3, N4), si la fiabilité de l'automate n'est pas garantie, les inconvénients pour les opérateurs sont plus importants que les bénéfices observés (par exemple avec l'effet cry-wolf).

Les automates doivent remonter des informations aux superviseurs, leur pouvoir attentionnel va conditionner leur prise en compte. La concurrence des différents systèmes de notifications, visuelles, sonores et haptiques présents dans les systèmes de supervision, peut considérablement dégrader la conscience de la situation, et l'interruption d'un opérateur dans sa tâche a un coût pour y revenir (le 'resumption lag'). Pour optimiser le couplage humain-automate, l'automate et l'éventuel système de notification afférent doivent donc prendre en compte ce coût et le contrebalancer par le bénéfice attendu sur la globalité de l'activité du superviseur.

Pour tenter d'adresser les problèmes liés à l'automatisation partielle décrits précédemment et améliorer le couplage entre l'humain et le système supervisé, se développe depuis quelques années la notion d'interfaces adaptatives qui visent à intégrer des données sur l'humain dans le système pour optimiser la transmission des informations. D'autres approches visent à adapter l'automatisation à l'aide d'évènements issus de l'environnement ou d'informations issues d'agents intelligents analysant la tâche, sa réalisation et les objectifs à atteindre.

Pour conclure, l'automatisation partielle pose un nombre important de problèmes qu'il faut adresser pour optimiser le couplage humain-automate :

- maintenir l'humain au centre du système et engagé dans sa tâche,

- tenir compte des capacités limitées de l'humain et adapter le système à ses capacités,
- permettre à l'opérateur de reprendre la main quand le système automatisé est défaillant ou sans solutions,
- le système automatisé doit interrompre (notifications) l'opérateur à bon escient.

Dans le prochain chapitre, nous aborderons le processus d'évaluation des interfaces et les différentes approches méthodologiques possibles. Par ailleurs, nous traiterons en détail les méthodes d'évaluation de la performance de l'humain dont l'utilisation de capteurs physiologiques qui ont été évoqués ici dans le cadre de la réalisation d'interfaces adaptatives basées sur l'état de l'opérateur.

4. STRATÉGIE MÉTHODOLOGIQUE

La complexité des interfaces de supervision de systèmes critiques est élevée ; concevoir un nouveau design dans le but d'améliorer la conscience de la situation rend son évaluation complexe. Dans ce chapitre nous aborderons l'évaluation des interfaces dans différents contextes allant des études de laboratoire aux observations en situation ainsi que les différentes méthodes qualifiées d'objectives ou subjectives d'évaluation de la performance de l'humain.

4.1. De l'étude de laboratoire à l'opérationnel

Le cadre de l'ingénierie cognitive (Rasmussen et al., 1994) propose quatre étapes clés dans la recherche cognitive appliquée qui ont pour but d'atteindre différents niveaux entre la validité écologique et le contrôle expérimental :

- Test des processus cognitifs (tests de laboratoire),
- Tests de fonctions (tests de laboratoire appliqués au contexte),
- Tests de fonctions dans un environnement de simulation complexe (étude ou haute-fidélité),
- Observations comportementales dans un environnement opérationnel.

4.1.1. Etudes de laboratoire

La première étape est la plus appropriée quand des concepts nouveaux doivent être évalués de manière à établir les processus cognitifs mis en œuvre à un niveau fondamental et définir l'espace des problèmes. Dans ce type d'évaluation, des tests répétitifs et des stimuli déclenchés sont mis en œuvre dans la tâche expérimentale. La performance est mesurée par instrumentation du code en enregistrant par exemple le temps mis pour réaliser une action ou les erreurs et l'utilisation éventuelle de capteurs physiologiques. Le but des études de laboratoire est d'évaluer dans un contexte expérimental contrôlé des mécanismes de bas niveau sur la réalisation d'une tâche.

Dans le domaine de l'interface homme-machine, des études de laboratoire vont permettre par exemple de tester de nouveaux types de menus, de nouveaux designs pour représenter des informations et de nouveaux moyens d'interaction. Le principal avantage des études de laboratoire est de pouvoir très facilement réaliser un grand nombre de mesures comportementales sur des tâches élémentaires. Pour illustrer ce type d'études dans le domaine de l'IHM, les deux exemples suivants sont des études de laboratoires réalisées au sein de l'équipe PII laboratoire spécialisé en IHM pour l'ATC, dont le but est de tester plusieurs moyens de désignation (Wacom et souris) grâce à une tâche développée spécifiquement. La première étude mesurait la performance comparative des moyens d'interaction (souris ou stylet Wacom en désignation directe ou indirection) (Alonso et al., 2009). La seconde testait différentes surfaces de zone d'indirection sur la Wacom pour optimiser le

compromis précision rapidité de la désignation des cibles (Alonso et al., 2011).

La tâche, qui a été conçue pour réaliser ces études de laboratoire (Figure 16: Wacom training, interaction directe au stylet) consiste à faire un glisser/déposer de ronds de couleurs dans les zones de réception carrée se trouvant sur les côtés de l'interface à l'aide de différents moyens d'interaction (souris, stylet Wacom en interaction sur un écran ou en indirection). Régulièrement, les cibles carrées sont déplacées sur un autre côté de l'interface, les ronds apparaissent de façon aléatoire.

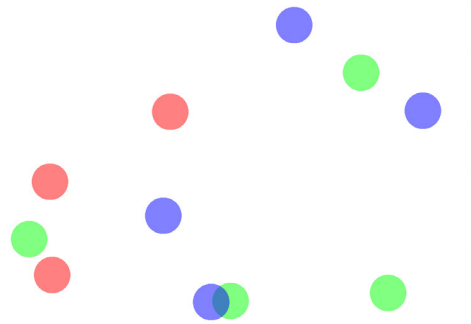


Figure 16 : Wacom training, interaction directe au stylet

La figure suivante (Figure 17 : Wacom Training en mode indirection) montre la même interface déportée sur un écran 30", l'interaction de déplacement des cercles se fait toujours sur la Wacom dans la zone noire, mais en indirection. Cette zone (en noir sur la Figure 17) représente en absolu la position du curseur au contraire des touchpads que l'on trouve habituellement sur les portables qui travaillent en position relative. Un pointage en haut à gauche de la zone d'indirection déplace le curseur au même endroit sur l'écran 30". Cette tâche a aussi été réalisée à l'aide d'une souris. La performance de l'opérateur peut être mesurée et comparée avec les différentes solutions de pointage.

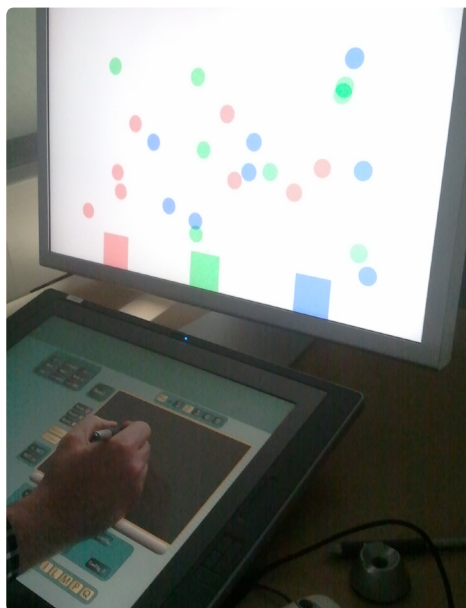


Figure 17 : Wacom Training en mode indirection

Les deux versions de la tâche ont été utilisées dans l'expérimentation décrite au chapitre 8 de cette thèse pour former les contrôleurs à l'utilisation de la Wacom que nous utilisons comme moyen de saisie et désignation en indirection dans notre simulateur d'étude pour le contrôle aérien mis en œuvre dans le cadre du projet SESAR WP4.7.2.

4.1.2. Tests de laboratoire appliqués au contexte : les micro-mondes

Les micro-mondes offrent un environnement de simulation dynamique en fournissant un haut degré de réalisme par rapport aux études de laboratoire tout en conservant un niveau de contrôle sur la manipulation et la mesure des variables de la tâche qui n'est pas possible dans le contexte opérationnel (Brehmer & Dörner, 1993 ; Gonzalez, Vanyukov & Martin, 2005). Du point de vue de l'ingénierie cognitive (Rasmussen, Pejtersen, et Goodstein, 1994 ; Woods & Roth, 1988), les micro-mondes se placent à un point important dans le continuum entre le travail de laboratoire initial et les études menées dans des contextes opérationnels, ils permettent d'atteindre un optimum entre la validité écologique et le contrôle expérimental.

Les micro-mondes sont des environnements de travail informatisés interactifs qui permettent l'étude du comportement humain pour des individus ou des équipes dans un scénario contrôlé (par exemple, Jobidon, Tremblay, Lafond, et Breton, 2006 ; Tremblay, Vachon, Lafond, & Kramer, 2012). Ils sont en général cognitivement exigeants et mettent en œuvre diverses fonctions cognitives telles que l'évaluation de la situation, la prise de décision, de suivi, la résolution de problèmes complexes, l'apprentissage de la causalité et de la planification (Gonzalez, Vanyukov et Martin, 2005).

Les micro-mondes couvrent un large éventail de domaines, tels que le domaine militaire (Smallman & St. John, 2003 ; St. John, Smallman, et Manes, 2005; Vachon, Vallières, Jones, & Tremblay, 2012; Hodgetts, Vachon, & Tremblay, en choix presse), économique / investissement (DiFonzo & Bordia, 1997 ;. Hantula & Crowell, 1994), lutte contre l'incendie (Tremblay et al, 2012), la défense de la cybersécurité (Rajivan, Venkatanarayanan, et Cooke, 2011), et les interventions d'urgence (Gagnon, Couderc, Rivest, & Tremblay, 2011).

Contrairement à de nombreuses études qui gèrent simplement une tâche et affichent des informations statiques, les micro-mondes sont interactifs et exigent la prise dynamique de décision (DDM – Dynamic Decision Making). Les environnements DDM sont caractérisés par de multiples variables interdépendantes, avec des facteurs de stress supplémentaires tels que l'incertitude, la pression temporelle et des ressources limitées (Gonzalez, Vanyukov, & Martin, 2005; Granlund et Johansson, 2004).

Bien que les micro-mondes conservent certaines caractéristiques du monde réel, ils ne sont pas simplement une version réduite de la vie réelle. Ils sont construits pour être conceptuellement pertinents et généralisables en termes de concepts théoriques (Mook, 1983). Il est donc important que la simulation recrée les relations fonctionnelles de la tâche d'origine, mais pas nécessairement toutes ses caractéristiques. Les micro-mondes peuvent avoir diverses applications comprenant les simulateurs de formation, les jeux vidéo, ou pour la recherche théorique dans les processus de décision.

Les micro-mondes partagent certaines caractéristiques clés avec leurs homologues du monde réel en ce sens qu'ils sont *dynamiques*, *complexes* et *opaques* (Brehmer, 2004). Un micro-monde *dynamique* évolue de manière autonome en temps réel afin que son état dépende d'une combinaison de ces deux facteurs endogènes (actions faites par l'individu) et des facteurs exogènes qui échappent à sa commande (Edwards, 1962). La *complexité* est influencée par le nombre de composants dans le système ainsi que le nombre et le type de relations entre ces composants (Gonzalez et al., 2005), pour lequel il peut y avoir plusieurs options différentes et des objectifs contradictoires (Brehmer, 1992). Diverses questions peuvent avoir besoin d'être traitées en parallèle, ce qui nécessite la programmation des actions futures et le suivi de l'information et des changements dans l'environnement. Un micro-monde est *opaque* dans le sens où il ne peut être compris directement, la relation entre les différentes variables n'est pas immédiatement apparente et doit être déduite. Les hypothèses concernant l'impact des différentes variables sont formulées et testées, en utilisant l'expertise des opérateurs dont le micro-monde s'inspire afin d'obtenir une meilleure compréhension de la façon dont le système fonctionne (Brehmer & Dörner, 1993).

Bien que le plus souvent menée dans un environnement de laboratoire, la capacité des micro-mondes à simuler des situations complexes et dynamiques les rend plus pertinents que les tâches en laboratoire classiques pour la compréhension des mécanismes cognitifs. Tandis que les tâches cognitives de base nous permettent d'étudier des variables spécifiques et isolées, généraliser ces conclusions à la nature complexe et interactive du monde réel peut parfois sembler assez difficile. En outre, cette complexité et le dynamisme de la tâche crée un monde qui semble plus intéressant et réaliste pour le participant. Avec des essais répétitifs et des stimuli artificiels, on peut se demander si les participants sont vraiment engagés dans les tâches traditionnelles de laboratoire et donc de savoir si le comportement peut être généralisé à partir d'un tel cadre naturel. Toutefois, les caractéristiques d'un micro-monde créent une expérience immersive : ils sont dynamiques, axés sur les buts, sémantiquement riches, et comportent également une dimension de temps (délais, retards) qui peut évoquer l'urgence et le défi. Leur nature interactive donne aux participants un sentiment de contrôle à partir duquel ils sont capables de voir les conséquences de leurs actions. Toutes ces fonctionnalités permettent de provoquer l'intérêt et la pression temporelle, ce qui augmente la probabilité que les participants s'engageront dans la tâche à un point tel que leur comportement deviendra complètement naturel (Dörner & Schaub, 1994 ; Gray 2002). Par ailleurs, certaines études suggèrent que la performance des tâches réalisées avec des micro-mondes peut prédire le succès dans leur équivalent de la vie réelle (Rigas, Carling, et Brehmer, 2002).

Un deuxième avantage des micro-mondes est son niveau élevé de *malléabilité* (Gray, 2002), contrairement aux évaluations conduites sur le terrain, le chercheur peut facilement contrôler l'environnement. Comme les micro-mondes impliquent une interaction homme-machine, le biais de l'expérimentateur est réduit (par exemple, le ton de la voix, les expressions faciales, les signes non verbaux). Les variables externes sont minimales puisque que la tâche est exactement la même pour chaque participant et cela permet de s'assurer que seules les variables pertinentes sont manipulées. C'est cette capacité à isoler les variables spécifiques qui permet d'aller au-delà de simples associa-

tions et d'établir des relations de cause à effet qui peuvent être généralisées à des environnements plus complexes. Les études de terrain sont en fait moins généralisables à des situations du monde réel. Le contrôle offert par les micro-mondes les rend par ailleurs idéaux à des fins de formation, des scénarios difficiles peuvent être mis en œuvre, et cela dans un contexte d'utilisation simplifié. Un outil de formation de ce type aussi appelé « *serious game* » dans la littérature, a le potentiel de fournir une rétroaction en temps réel et de manière uniforme afin d'améliorer l'expérience d'apprentissage.

Les micro-mondes offrent d'autres avantages par rapport à la recherche ou à la formation dans le domaine en termes de contraintes pratiques ou logistiques. L'utilisation de scénarios générés par ordinateur évite les considérations éthiques qui peuvent se poser. Il y a aussi des avantages financiers considérables puisque leur mise en œuvre nécessite moins de moyens aussi bien au niveau humain que matériels, les équipements spécialisés de leur équivalent opérationnel ne sont pas nécessaires et différents essais et conditions peuvent être exécutés sans coût supplémentaire ou mise en œuvre différente. Ils sont aussi utiles en termes de plateformes de tests pour la validation de choix de conception dès le début du cycle de développement de la technologie. Ils permettent d'établir la faisabilité et d'évaluer de nouveaux designs ou solutions d'interaction avant de les développer et les mettre en œuvre dans des simulateurs plus complexes. Ils favorisent et facilitent la collecte précise et efficace de gros volumes de données, qui sont enregistrées et codées rapidement sans erreurs. Cette mesure précise de variables réduit les erreurs et réduit l'écart entre les études de laboratoire et sur le terrain. La gamme de mesures dépendantes qui peuvent être collectées en parallèle facilite l'utilisation d'une approche holistique (Lafond, Vachon, Rousseau, et Tremblay, 2010), permettant aux chercheurs d'atteindre une compréhension globale de la performance dans son ensemble, et d'effets interactifs entre les variables (par exemple Hodgetts et al, sous presse ; Vachon, Lafond, Vallières, Rousseau, et Tremblay, 2011). Un autre avantage pratique est que les micro-mondes peuvent être utilisés avec des experts ou des novices. D'assez haute-fidélité pour les opérateurs experts, ces tâches simplifiées ne nécessitent pas nécessairement des connaissances ou des compétences spécifiques au domaine, il est donc possible pour les chercheurs de recruter un grand nombre de participants pour des tests d'hypothèses plutôt que d'avoir à compter sur un nombre limité de participants qui possèdent l'expertise technique spécifique pour faire fonctionner le système.

Bien sûr les micro-mondes ont leurs limites et ne représentent qu'une seule étape dans le cadre de l'ingénierie cognitive des systèmes (Rasmussen et al., 1994). Ils sont utiles pour comprendre les fonctions sous-jacentes impliquées dans une tâche, et cette connaissance peut à son tour aider à proposer des moyens pour mieux la prendre en compte. Les micro-mondes sont également utiles pour l'élaboration et l'évaluation de ces solutions technologiques dans un scénario simplifié, ces choix devront être testés dans un environnement plus complexe et réaliste afin de s'assurer que les performances auparavant observées sont bien reproduites. En termes de formation, les micro-mondes peuvent aider à l'acquisition de compétences initiales et l'apprentissage des règles techniques spécifiques, mais des simulations dans des environnements plus réalistes sont tout de même nécessaires. Dans les paragraphes suivants, nous présenterons plusieurs micro-mondes utilisés dans le domaine aéronautique.

4.1.2.1. *Un micro monde aéronautique en contexte multitâche : MATB*

Dans le domaine aéronautique, MATB (Multi-Attribute Task Battery) qui est un micro-monde open source de la NASA (Comstock & Amegard, 1992) reprenant des éléments de la tâche des pilotes a été fréquemment utilisé pour réaliser des études visant à évaluer la performance humaine dans des tâches mettant en œuvre l'utilisation d'automates et les aptitudes au multi-tâche (Figure 18 : MATB (Multi-Attribute Task Battery)). Depuis les années 90, il a été utilisé dans plus de 40 études. Suivant les objectifs de l'étude, il permet en suivant un script défini par l'expérimentateur de programmer l'apparition d'évènements pouvant porter sur cinq types de tâches : monitoring système, maintien de cible (tracking), communications, management de ressources et agenda pour anticiper la charge de travail. Dans le contrôle aérien, cette tâche a été utilisée lors d'expérimentations mettant en situation contrôleurs et pilotes lors d'évènement TCAS pour maintenir un niveau élevé de charge de travail pour les pilotes (Rome Cabon et al., 2006). Ce type de micro-monde permet de s'intéresser aux mécanismes sous-jacents au travail en contexte multitâche. Par ailleurs il ne nécessite pas d'expertise particulière, ce qui rend sa mise en œuvre possible avec tout type de sujets.

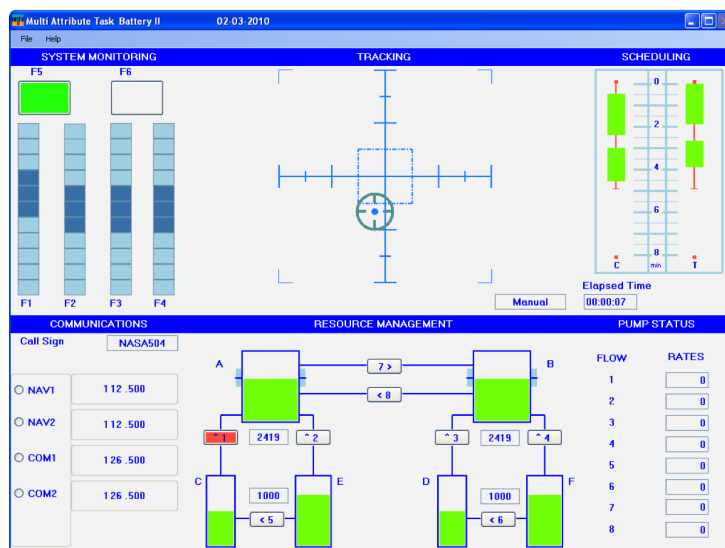


Figure 18 : MATB (Multi-Attribute Task Battery)

4.1.2.2. *Un micro monde ATC pour l'évaluation du travail collaboratif : CTEAM*

Un micro-monde ATC (Bailey, 1999) appelé CTEAM (Controller TeamWork Evaluation and Assessment Methodology) a été développé par le laboratoire de recherche des ressources humaines de l'Institut de médecine aéronautique civile (CAMI) de la FAA (Federal Aviation Administration USA). Il fournit une IHM radar très simplifiée (Figure 19 : le micro-monde ATC CTEAM) et peut être exécuté pour plusieurs participants dans un contexte mono ou multisecteur. Aucune formation spécifique de contrôle de la circulation aérienne n'est nécessaire pour utiliser CTEAM et des expérimentations ont été réalisées avec des élèves. Chaque participant ne gère que son secteur et doit réaliser des actions afin de maintenir la séparation entre aéronefs et les guider en fonction de leurs itinéraires prescrits (ce qui peut être un autre secteur ou un aéroport dans le secteur). La charge de travail peut être

ajustée par le nombre d'aéronefs ou la route qui leur est assignée. Diverses données peuvent être recueillies lors de l'expérimentation afin de mesurer la performance des participants ou de l'équipe globalement dans les scénarios.

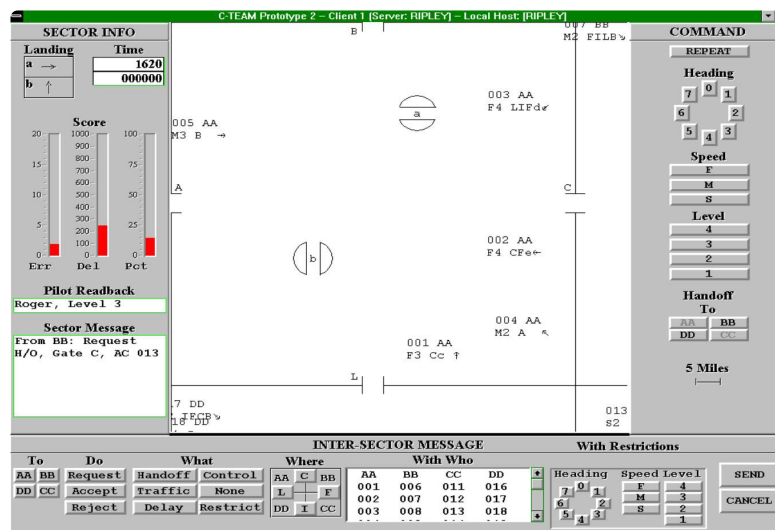


Figure 19 : le micro-monde ATC CTEAM

4.1.2.3. Un micro-monde pour l'ATC : Laby

Laby est un projet développé en interne de l'équipe PII (Performance et Innovation des IHM) du centre de R&D Français dans le domaine ATC (Parise R., Imbert JP et al. 2012, Imbert, Hodgetts H. et al. 2014). Il a été développé à l'origine pour prototyper et évaluer dans un contexte expérimental contrôlé de nouveaux designs ou de nouvelles interactions pour la visualisation des contrôleurs aériens opérationnelle. À ce titre et dans un souci de réalisme, il reprend un certain nombre de caractéristiques de cette visualisation opérationnelle.

Ce micro-monde est conçu autour d'une tâche principale consistant à guider un ou plusieurs avions sur une route déterminée matérialisée par un chemin de couleur verte (Figure 20 : le micro-monde ATC Laby). Les participants doivent saisir des valeurs numériques telles que le cap, le niveau de vol, la vitesse à l'aide d'interaction sur l'étiquette radar faisant apparaître des listes de valeurs afin de maintenir les avions en vol le long de la trajectoire et renseigner le système avec des consignes qui apparaissent au fur et à mesure du trajet de l'avion dans le labyrinthe. Par ailleurs, il est possible de déclencher l'apparition d'obstacles ou d'avions interférents gérés par le système pour complexifier la tâche du contrôleur aérien. Par exemple, des avions gérés automatiquement par le système peuvent être utilisés pour créer Cette tâche centrale est très flexible et de nombreux scénarios différents et de configurations peut être générés en fonction des besoins de l'utilisateur, que ce soit pour le développement du système et de l'évaluation, ou à des fins de formation.

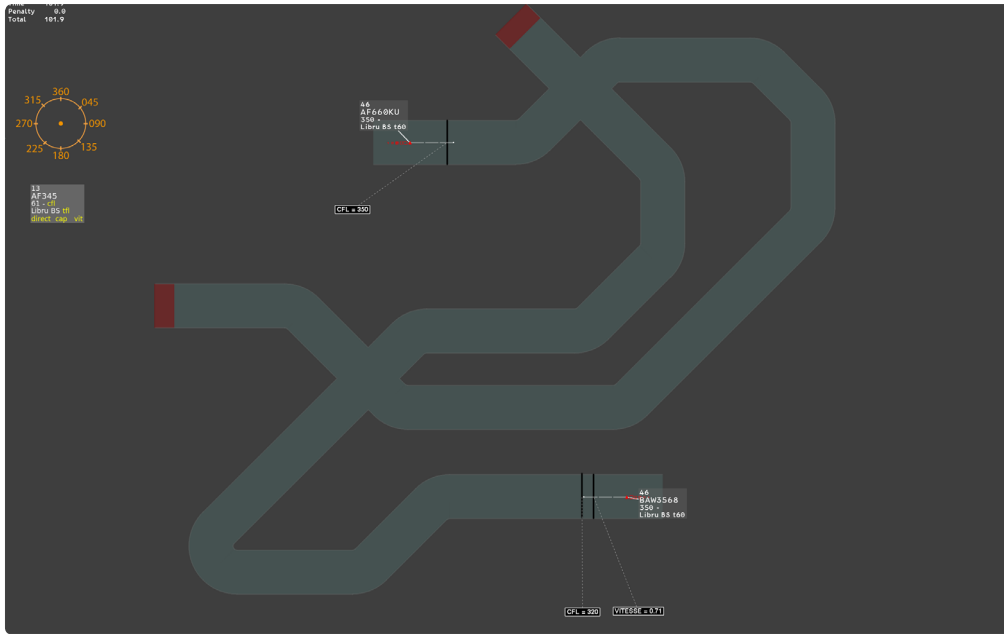


Figure 20 : le micro-monde ATC Laby

Pour concevoir un labyrinthe, un outil appelé LabyEdit permet de dessiner le ou les trajets à faire suivre par les avions guidés et positionner les consignes ainsi que les valeurs à saisir le long de ces trajets. La figure 21 présente l'édition d'un Labyrinthe minimal à l'aide de LabyEdit (voir annexes pour plus d'explications sur cet outil) avec un tronçon droit, une zone de départ et d'arrivée ainsi qu'une zone de consigne. Cette dernière possède un côté début de consigne (côté rectangle noir). C'est par là que l'avion en entrant va déclencher l'apparition de la consigne associée. Un côté fin de consigne matérialise le seuil qui va provoquer l'effacement de la consigne associée si celle-ci n'a pas été réalisée (c'est-à-dire correctement renseignée). L'avion guidé commence son trajet dans la zone de départ, son passage dans la zone d'arrivée termine le labyrinthe et la tâche pour le sujet.

Bouton pour inverser le sens de l'objet (ici, la poignée passerait à gauche)

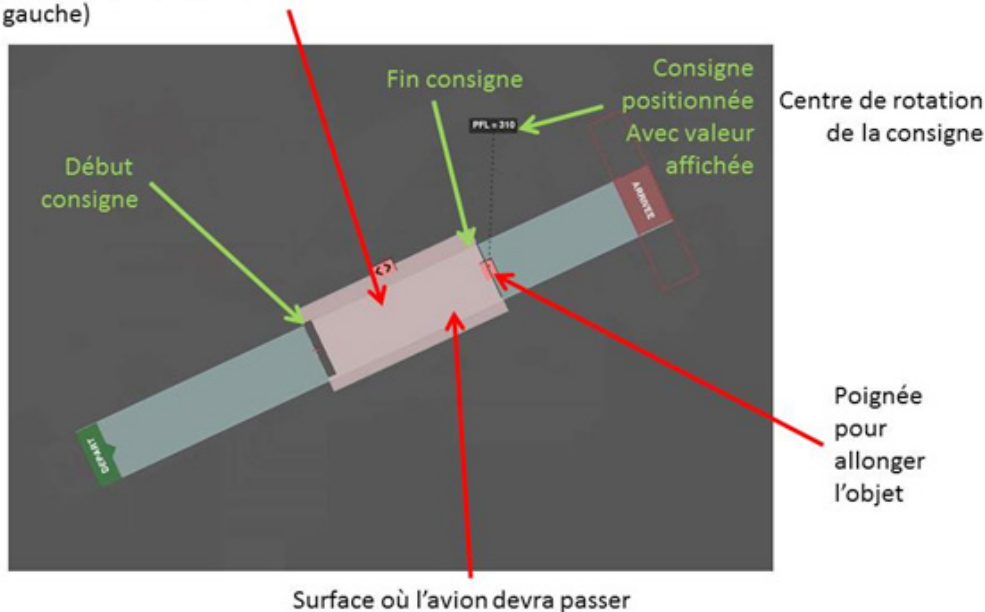


Figure 21 : un Labyrinthe minimaliste édité

Il est possible de collecter de grandes quantités de données afin d'obtenir une vision globale de la façon dont la performance peut être affectée par des changements dans l'interface ou dans la façon dont le parcours a été conçu (nombres de consignes, d'avions, vitesse des avions etc.). Un score est affiché au participant qui est calculé dynamiquement en fonction des erreurs (par exemple, des pénalités sont encourues pour dévier de la route, entrer dans une zone interdite, ne pas maintenir la distance minimale de séparation entre deux avions, etc.), celui-ci ne préjuge pas de la complexité de la tâche. Ce score donne une mesure de la performance qui couvre les principales caractéristiques et paramètres de la tâche Laby pour la gestion du trafic et des conflits potentiels (pour des mesures de complexité dans l'ATC voir Giannazza & Guittet, 2006), et il inclut aussi des indicateurs basés sur la capacité du contrôleur à renseigner le système avec les informations demandées pendant le parcours. Pendant la passation, ce score peut également être utilisé pour fournir un retour aux participants sur leur performance et de cette façon les aider à rester concentrés sur la tâche principale, même lorsque les tâches secondaires doivent être traitées simultanément. En plus du score de performance, les temps précis de toutes les actions se déroulant pendant une passation sont enregistrés par Laby, de sorte que la performance peut être évaluée en termes de temps mis pour saisir une consigne, son exactitude, etc.

Laby est dynamique dans le sens où la situation évolue de manière autonome (avions interférents commandés par le système, consignes à effectuer par l'opérateur) et comme un résultat direct des actions effectuées par le participant (avions guidés). Il s'agit d'un environnement multitâche complexe, ce qui peut impliquer le contrôle de plusieurs aéronefs et l'engagement dans d'autres tâches concurrentes en fonction de la configuration particulière du scénario comme répondre aux notifications de consignes du système. Des scénarios *opaques* (Brehmer, 2004) peuvent être créés dans le sens où le participant a besoin de construire une compréhension de l'impact du changement de différentes variables sur l'appareil qu'ils guident, par exemple, la vitesse ou la correction de cap nécessaire pour éviter une zone interdite ou un conflit potentiel avec un autre avion.

Pour illustrer le continuum de l'ingénierie cognitive (Rasmussen et al, 1994) dans une perspective ATC, il est tout à fait approprié d'utiliser un paradigme d'évaluation avec une tâche de laboratoire et des répétitions sur un échantillon important de participants novices afin d'aborder une question de recherche appliquée (par exemple l'utilisation d'une souris ou d'un stylet pour désigner des avions), si les concepts à l'étude sont assez nouveaux et si les processus impliqués dans la tâche (mémorisation de séquences verbales/spatiales de stimuli) sont également comparables à la performance sur la tâche dans le monde réel (Tremblay, Parmentier, Hodgetts, Hughes, & Jones, 2012). Inversement, si le but de l'expérience est d'appliquer des théories cognitives bien établies dans un cadre plus réaliste (par exemple, l'impact des communications sur l'exécution des tâches de vol), il pourrait être plus approprié à ce stade d'utiliser un simulateur de vol réaliste, mais sur un échantillon plus restreint de pilotes qualifiés (Hodgetts, Farmer, Joose, Parmentier, & Jones, 2005). Laby se situe entre ces deux niveaux, fournissant un environnement multitâche, mais réaliste en fonction des tâches requises par les participants et n'impliquerait pas forcément une expertise spécifique au domaine.

La nature complexe et dynamique de Laby, ainsi que la facilité avec laquelle les variables spécifiques peuvent être manipulées et mesurées, le rend utile pour les chercheurs dans un large éventail de domaines concernant les facteurs humains et la psychologie appliquée. Par exemple, les interruptions (Hodgetts et al, sous presse; Trafton, Altmann, Brock, et Mintz, 2004), le bruit de fond (Tremblay et al, 2012.), la surcharge d'information (Cummings & Mitchell, 2008), la fatigue (Hockey et al, 1998 ; Sauer et al, 2003), et la pression temporelle (Rieskamp & Hoffrage, 2007 ; Pachur & Hertwig, 2006) sont tous des facteurs de stress qui peuvent affecter les performances dans une mission de surveillance. Utiliser Laby afin de parvenir à une meilleure compréhension des processus cognitifs mis en œuvre lors de son utilisation pourrait permettre ultérieurement d'élaborer des dispositifs de soutien dans ce type de conditions. La flexibilité de Laby pour différentes configurations permet de faire usage de différentes modalités d'entrée d'information et de représentation d'informations (signaux auditifs, indices visuels), ce qui pourrait contribuer au développement théorique sur le concept d'intégration multimodale et des interactions entre les différents moyens (Prabhakaran, Narayanan, Zhao, et Gabrielli, 2000 ; Tremblay et al., 2012). La nature de la tâche de surveillance est particulièrement indiquée pour le domaine de la recherche sur la cécité au changement (Simons, 2000 ; Vachon et al, 2012) ; il serait par exemple intéressant de vérifier si les participants réagissent à un changement inattendu de vitesse ou de trajectoire d'un aéronef. Par ailleurs les alarmes sonores pourraient servir de base pour une étude sur le phénomène équivalent dans le domaine auditif : la surdit  attentionnelle (Eramudugolla, Irvine, McAnally, Martin, et Mattingley, 2005 ; Vitevitch, 2003). La quantité d'informations affichées sur l'écran signifie que la tâche pourrait se prêter facilement à des études sur la recherche visuelle (Boot, Be-cic, & Kramer, 2007 ; Neider, et Zelinsky, 2007), la charge cognitive (Lavie, 2005), la vigilance (Molloy et Parasuraman, 1996), la capture attentionnelle à l'aide d'attracteurs auditifs/visuels (Hughes, Vachon, & Jones, 2007 ; Yantis & Jonides, 1996) et la conscience de la situation (Endsley, 1994 ; Wickens, 2002). À l'instar de MATB, Laby offre aussi des possibilités pour l'étude des différences individuelles sur les performances multitâches (la capacité de travail, la mémoire). La tâche de guider un avion le long d'un parcours utilise les caractéristiques générales d'une activité de commandement et de contrôle (comme la surveillance, l'attention, l'anticipation, décision dynamique, l'exécution d'une séquence d'actions) qui sont pertinentes pour un certain nombre d'autres domaines tels que la conduite, la sécurité, la surveillance des cyberattaques, la gestion de crise et l'intervention d'urgence.

Laby fournit une plateforme de recherche afin de mieux caractériser les processus cognitifs et les limites de la connaissance humaine en situations multitâches en s'appuyant sur une série de mesures de performance et des indicateurs de processus, comme les délais de décision/temps de réaction, le taux d'erreur, des questionnaires sur la conscience de la situation ou la charge de travail.

Avec la quantité croissante d'informations à afficher, souvent sur plusieurs écrans, les exigences cognitives de la tâche de supervision des contrôleurs peuvent parfois être trop importantes. Laby représente une première étape importante dans le continuum de l'ingénierie des systèmes permettant d'évaluer des prototypes dans un environnement simulé à faible coût et sans risque. Alors que les plateformes expérimentales plus fondamentales ne

fournissent pas un niveau assez élevé de réalisme - par exemple, les contrôleurs ne peuvent montrer certaines activités mentales dans des situations complexes dynamiques et sous pression temporelle - Laby permet de réaliser des expériences qui sont considérées comme suffisamment engageantes permettant une première évaluation ou des choix de designs avant une utilisation réelle. Il offre un environnement contrôlé pour le processus itératif avec la possibilité de changer la dynamique, de modifier les caractéristiques de conception, et d'examiner des solutions de rechange.

Laby s'appuie sur l'utilisation d'une approche holistique, en permettant aux chercheurs d'examiner une série de variables à la fois isolément ou en combinaison, afin d'acquérir une compréhension globale de l'impact d'une variable particulière sur la performance dans son ensemble. Par exemple, un gain sur une dimension (par exemple la précision nécessaire pour ouvrir un menu) peut être accompagné par une perte dans une autre dimension (comme le temps de réaction réduit) et il peut être important qu'un outil d'aide ne compromette pas les performances dans un autre domaine (Vachon et al, 2011 ; Vallières et al, soumis).

Laby n'est pas exempt de limitations, la première étant qu'il ne permet pas de reproduire les interactions verbales qui existent dans un environnement de travail réel entre les contrôleurs et les pilotes à travers la radio VHF. La communication verbale est une partie importante de l'ATC, il n'est donc pas possible de généraliser complètement les résultats des scénarios menés dans Laby avec ceux obtenus dans un simulateur plus complexe qui incluent cette tâche supplémentaire. L'ajout d'une autre modalité pour transmettre et recevoir des informations augmentera certainement la charge cognitive. Par ailleurs, même le bruit de fond qui n'est pas pertinent pour la tâche est connu pour affecter certains aspects de l'exécution des tâches en vol (Hodgetts et al, 2005 ; Tremblay et al, 2012) il est donc possible que la saisie d'informations dans Laby soit impactée par la communication verbale. Il y aurait plusieurs possibilités pour intégrer la communication verbale dans Laby, par exemple en intégrant un pseudo-pilote qui réaliserait les consignes saisies par le contrôleur sur Laby ou en utilisant de la reconnaissance vocale et de la synthèse vocale pour répondre au contrôleur. De cette façon, les avions à l'écran ne modifieraient pas leurs paramètres par une conséquence directe de commandes saisies par le contrôleur, mais par l'action du pseudo-pilote suite à un échange verbal. Ce type de développement rendrait Laby plus réaliste tout en rendant sa mise en œuvre plus complexe puisqu'il nécessiterait deux personnes pour réaliser un scénario. Laby peut intégrer de nombreux aspects de l'exécution des tâches, mais il doit être pris comme une simple étape sur le continuum décrit par Rasmussen (Rasmussen, 1994). Cet environnement ATC simplifiée peut aider à comprendre les fonctions et les interactions tâche sous-jacente, mais les résultats doivent ensuite être validés dans les simulateurs plus réalistes. De même, Laby peut aider avec les premières étapes de la formation des contrôleurs, mais l'expérience avec le système actuel est toutefois nécessaire afin d'apprendre les aspects techniques spécifiques à l'IHM opérationnelle.

Laby satisfait la plupart des exigences décrites dans le guide défini par Buchner (1995) et par Frensch et Funke (1995), mais contrairement au micro-monde CTEAM (Controller TeamWork Evaluation and Assessment Methodology), il ne peut pas être utilisé pour évaluer la performance des contrôleurs en équipe et ne fonctionne qu'en mono secteur. Néanmoins, il

offre plus de possibilités que CTEAM pour se conformer à une situation réelle ATC et son IHM est plus proche de l'opérationnel.

4.1.3. Environnements de simulations complexes

Dans le domaine de la simulation de systèmes complexes, on peut définir deux niveaux de simulateurs, ceux qualifiés de 'Part task' aussi appelés simulateurs d'études et ceux qualifiés de 'Full Scale' aussi appelés simulateurs 'haute-fidélité'. Ces deux types de simulateurs répondent à des besoins très différents, le premier par sa possible évolutivité permet de tester de nouvelles fonctions dans un environnement réaliste de l'activité opérationnelle, le second est une version simulée de la version opérationnelle du système qui est en général utilisé à des fins de formation ou d'entraînement. Un simulateur d'étude est moins cher et plus facilement modifié pour intégrer de nouveaux dispositifs ou fonctions logicielles, il est aussi généralement pensé pour intégrer les mécanismes permettant de réaliser l'évaluation expérimentale. La mise en œuvre de simulateurs haute-fidélité est coûteuse et compliquée puisque ce type de simulateur est généralement utilisé à des fins de formation. Ces deux types de simulateurs offrent des niveaux de contrôle différents sur le processus expérimental ; augmenter le réalisme diminuant ce contrôle (Folleso & Volden, 1993). Dans le domaine aéronautique est considéré comme 'full scale' un simulateur reproduisant précisément le cockpit ainsi que le modèle de vol de l'avion simulé et monté sur vérin pour reproduire les accélérations et assiettes lors des phases de transition. Un simulateur d'étude pourra être fixe, moins réaliste au niveau des modèles ou des systèmes bord, mais permettra d'intégrer facilement des prototypes de nouveaux équipements de bord, des capteurs physiologiques pour réaliser l'évaluation de l'opérateur humain. Il peut être modifié pour sortir des paramètres internes au simulateur pour de la mesure de performance.

Des simulateurs d'études et haute-fidélité dans le domaine du contrôle aérien sont utilisés en France depuis une vingtaine d'années dans les services qui développent les futurs systèmes de contrôle comme la DTI. Les simulateurs d'études utilisés dans le cadre des projets SESAR ont en général un niveau de réalisme assez poussé pour réaliser des simulations multisecteurs tout en offrant la possibilité d'intégrer des composants logiciels (automates, nouvelles visualisations). Dans le cadre du projet WP4.7.2, une nouvelle interface homme-machine a été intégrée au simulateur.

4.2. Évaluation de l'utilisation d'une interface homme-machine

Les méthodes d'évaluations fonctionnelles en Facteurs Humains peuvent être catégorisées en deux types, les évaluations dites subjectives et celles qualifiées d'objectives ; la dernière catégorie incluant les mesures physiologiques et comportementales. La plupart des outils ou techniques les plus utilisées dans le domaine des facteurs humains reposent sur de la mesure subjective fondée sur l'auto-évaluation (par exemple évaluation périodique de la charge cognitive perçue par l'opérateur pendant l'activité), des questionnaires le plus souvent remplis à la fin du protocole expérimental, de l'observation ou de la verbalisation durant l'activité. Ce que pensent les par-

ticipants lors d'expérimentations n'est pas forcément cohérent par rapport à ce qu'ils ont fait en réalité. C'est particulièrement vrai lors de l'utilisation de systèmes interactifs complexes où ils ne sont pas forcément capables de verbaliser leur mode opératoire. Par ailleurs, les reports périodiques pendant l'activité, dans des cas de charges de travail élevées peuvent rentrer en conflit avec le déroulement de la tâche. De la même façon, la verbalisation durant l'activité peut interférer avec l'activité de compréhension ou le déroulement de la tâche si celle-ci nécessite des échanges vocaux.

Les mesures objectives reposent sur l'analyse de la performance et dépendent du type d'étude réalisée.

4.2.1. Utilisation de méthodes d'évaluation subjectives

Il existe de nombreux tests d'évaluation subjective, certains s'appliquant à l'évaluation de la charge de travail, d'autres à la conscience de la situation ou au travail en équipe. Des tests d'évaluation ad hoc peuvent aussi être réalisés pour répondre à des questions spécifiques liées à l'activité. Nous n'aborderons dans ce paragraphe que les méthodes d'évaluation de la charge mentale ou de la conscience de la situation.

Pour l'évaluation des interfaces homme-machine, et à fortiori dans les supervisions de systèmes critiques, la charge de travail est un critère essentiel ; pour assurer la sécurité et le confort des opérateurs, il est important d'assurer un niveau correct de charge cognitive de façon à ce qu'ils ne soient jamais sous ou surchargés (Figure 15 : performance en fonction de la charge de travail).

Plusieurs tests ou questionnaires sont couramment utilisés comme le NASA Task Load Index (NASA TLX – Hart & Staveland, 1988), le Subjective Workload Assessment Technique (SWAT – Reid & Nygren, 1988) ou le Workload Profile (WP Tsang & Velazquez, 1996). Une étude (Rubio & Martin, 2004) a montré une forte corrélation entre les résultats obtenus avec ces trois tests lors d'une expérimentation. D'autres tests plus simples peuvent être utilisés durant l'activité comme ISA (Instantaneous Self-Assessment of workload technique – Jordan, 1992) ; il s'agit pour le sujet, sur une base périodique, en général toutes les deux minutes, de renseigner sur une échelle de 1 (faible) à 5 (élevée) la charge de travail ressentie. Ce test a été utilisé essentiellement dans le domaine du contrôle aérien comme en 1997 par Kirwan et plus récemment lors du projet SESAR WP4.7.2 pour mesurer l'impact de l'utilisation de nouveaux outils sur la charge de travail des contrôleurs.

L'évaluation de la conscience de la situation (voir chapitre 3), selon la définition d'Endsley (Endsley, 1995) a pour but de déterminer la perception des éléments dans l'environnement dans un volume de temps et d'espace, la compréhension de leur signification et la projection de leur état dans le futur proche. Les tests les plus couramment utilisés en ce domaine sont SAGAT (Situation Awareness Global Assessment Technique – Endsley, 1995) et SART (Situation Awareness Rating Technique – Taylor, 1990) pour les tâches en situation réelle. SAGAT ne fonctionne qu'en environnements simulés et la simulation doit être arrêtée à certains moments et les écrans éteints. Les sujets doivent alors répondre à un questionnaire sur la situation courante. Ces réponses sont ultérieurement comparées à l'analyse

d'experts (SME – Subject Matter Expert) disposant des images à l'arrêt de la simulation. SART consiste en une échelle de 10 points auxquels les sujets répondent en fin de passation, dont les questions peuvent être catégorisées en 3 groupes concernant les ressources attentionnelles (Demands), l'exigence d'attention (Supply) et la compréhension de la situation (Understanding). Si SAGAT est critiquée pour sa difficulté de mise en œuvre, SART l'est aussi en raison de son utilisation uniquement à la fin de la tâche. Elle ne donne qu'une impression globale de la conscience de la situation pendant une passation. D'autres méthodes temps réel mais plus intrusives existent comme SPAM (Situation Present Assessment Method – Durso, 04). Des questions sont posées périodiquement sur les trois niveaux de SA en utilisant le temps de réponse comme indicateur supplémentaire.

Pour un contrôleur aérien, maintenir une bonne conscience de la situation signifie avoir une image mentale correcte de la situation aérienne courante et de son évolution. Dans ce domaine, Eurocontrol a développé une méthode appelée SASHA (Eurocontrol, 2003) (<http://www.skybrary.aero/index.php/SHAPE>) dans le but d'évaluer l'impact sur la SA de différents niveaux d'automatisation qui a deux déclinaisons SASHA_(L|Q) dont l'une est pendant la passation et l'autre à la fin de la passation). SASHA_L est basée sur la méthode SPAM (Durso, 2004), comme SART, SASHA_Q est en questionnaire en 10 points visant à évaluer la SA globalement pendant la tâche. Dans le cas de l'évaluation de fonctions particulières d'une IHM de contrôle aérien, un questionnaire ad hoc permet d'avoir des retours plus fins de la part des SME (Subject Matter Experts) que les questionnaires plus généraux sur la conscience de la situation.

4.2.2. Mesures comportementales

Les mesures comportementales permettent d'obtenir une vision objective de la performance de l'humain pendant la réalisation de sa tâche. Elles consistent à définir des métriques associées à l'utilisation du système ou bien à la performance au regard des objectifs de plus haut niveau de son activité et à les intégrer dans la tâche de fa.

Par exemple, dans le cadre d'une étude de laboratoire ou de micro-monde, la performance est mesurée à l'aide de métriques propres aux tâches comme le temps de validation d'une notification, le taux d'erreur (Brangier & Barcenilla, 2003), le nombre d'actions par secondes. Le plus souvent, des tâches multiples sont mises en œuvre de façon à saturer la capacité de travail de l'opérateur (Leplat & Spérandio, 1967 ; Baddeley & Della Sala, 1996). C'est le cas de MATB présenté au paragraphe 4.1.2.1 dont les différentes tâches (jusqu'à 6) peuvent être déclenchées à loisir pendant une passation. Dans l'exemple cité au paragraphe 4.1.1, la performance du moyen de désignation (souris ou stylet) a été mesurée de deux façons. La première basée sur un score de haut niveau lié au nombre de cibles déposées dans les zones rectangulaires, la seconde basée sur le temps de réaction après l'apparition d'un cercle jusqu'à l'action de touché pour réaliser le glisser/déposer. Dans le cas du micro-monde Laby, la tâche principale consiste à guider un avion dans un labyrinthe tout en exécutant des consignes présentées durant la tâche. La performance est mesurée à l'aide d'un score global basé sur la trajectoire suivie et la complétion des consignes. Elle sera mesurée de façon plus pré-

cise avec le taux d'erreur des saisies de consignes ou le temps moyen pour les entrer dans le système.

Dans le cas de CTEAM, trois types de mesures peuvent être réalisées : la complexité de la tâche, la performance et l'efficacité des contrôleurs. La performance est évaluée par le nombre et la durée des erreurs de séparation. L'efficacité est mesurée par le nombre de commandes émises pour diriger les avions tout en maintenant la séparation minimale. La complexité de la tâche dépend du nombre d'avions, la durée de vol dans le secteur ainsi qu'une mesure de l'activité du système. Une mesure axée sur le comportement de l'expertise, l'indice de Cochran- Weiss - Shanteau (CWS) (Weiss, Shanteau, 2003) a été utilisé et validé comme un bon indicateur de la performance au cours des expérimentations utilisant CTEAM et de simulations haute-fidélité de l'ATC. De grandes valeurs de CWS signifient une meilleure performance ; cet indice permet également d'observer des changements dans la complexité de la tâche ou de l'efficacité du contrôleur.

Dans le cadre d'études basées sur des simulateurs haute-fidélité, de par la nature complexe des moyens mis en œuvre, la mesure de performance est en général établie par rapport à des critères liés à l'activité. Dans le domaine du contrôle aérien (Gianazza et al. 2006), des indicateurs de la complexité de la tâche ont été établis et permettent d'obtenir des métriques par rapport aux scénarios simulés. Des indicateurs plus simples peuvent permettre d'évaluer le déroulement de la tâche du contrôleur ; le trafic est enregistré pendant la simulation, son analyse après coup va permettre d'estimer à quel point le contrôleur a bien géré la situation et correctement séparé les avions en mesurant les distances minimales de croisement pour les avions en conflit ou bien en mesurant la durée des trajectoires pour voir si les séparations créées par le contrôleur ont trop rallongé ces trajectoires. L'importance d'un critère par rapport à un autre dépend de l'objectif de l'analyse. S'il s'agit de sécurité ou de consommation de carburant, les indicateurs seraient bien différents.

4.2.3. Mesures physiologiques

Une approche originale pour améliorer la compréhension de l'interaction homme-système s'est développée ces dix dernières années. Elle consiste à utiliser des capteurs physiologiques et mesurer lors de l'activité l'impact de l'utilisation du système pendant le déroulement de la tâche, et ce à des fins d'études ou plus récemment à des fins d'adaptation du système à l'état de l'humain. Différents types de capteurs sont utilisés dont certains concernent l'activité du cerveau, d'autres la réponse du corps à des stimuli ou bien l'étude du suivi du regard qui permet d'analyser l'utilisation de l'interface.

Dans l'approche objective, les capteurs physiologiques peuvent aussi être utilisés pour mesurer des niveaux de charge mentale, des réponses physiologiques à des stimuli ou des états cognitifs de type tunnellation attentionnelle, surcharge cognitive, divagation attentionnelle, anxiété, stress, etc. Les sujets sont appareillés pendant les expérimentations, les mesures issues des capteurs sont corrélées avec les données de performance provenant des tâches expérimentales lors de processus de post-traitement qui peuvent s'avérer complexes et fastidieux.

Plusieurs types de capteurs physiologiques sont couramment utilisés lors d'études en laboratoire ou en situation pour la mesure de la performance des opérateurs. Les cinq types principaux étant :

- Mesure de la résistance électrodermale GSR (Galvanic Skin Response) (Clarion, 2009),
- mesure du rythme cardiaque ECG (ElectroCardioGram) (Michaut & Pin, 1965 ; Kalseek & Ettema, 1965 ; Dinand & Defayolle, 1969 ; Delignières, Legros & Famose, 1991),
- mesure de l'activité du cerveau EEG (ElectroEncephaloGram), FNIR (Functional Near-Infrared spectroscopy) (Wilson & O'Donnell, 1988 ; Gevins, Smith, Leong, McEvoy, Raabe, Rutschmann, Schrauf & Greenlee, 2005 ; Ayaz, 2012),
- mesure de l'activité oculaire (Beatty, 1982 ; Janisse, 1977 ; Di Stasi et al., 2010 ; De Waard, 1996 ; Golberg & Kotval, 1998),
- mesure de l'activité du système respiratoire (Gamberale, 1972 ; Seip, Snaed, Pierce, Stein & Weltman, 1991).

L'utilisation de moyens physiologiques dans des études concernant le trafic aérien date de la fin des années 90. L'expérimentation Stress (Rome Cabon et al., 2006) conduite au Centre d'Etudes de la Navigation Aérienne mesurait le rythme cardiaque de contrôleurs aériens et de pilotes lors de simulation les mettant dans des situations de TCAS (Traffic Collision Avoidance System). La résistance électrodermale a été utilisée pour mesurer l'effet de la charge de trafic sur les contrôleurs aériens (Averty, 1998). Un capteur FNIR (Ayaz, 2012) a été mis en œuvre pour réaliser une évaluation de la charge cognitive grâce à l'évaluation de l'oxygénation des zones préfrontales du cerveau sur des contrôleurs aériens. Des analyses de variation du diamètre pupillaire ont aussi été réalisées dans ce même but (Martin, 2012).

4.2.3.1. Mesures oculométriques

« Seuls les objets qui émettent ou réfléchissent de la lumière nous sont donnés à voir : une des fonctions de l'œil est de sélectionner et réarranger une partie de l'énergie des radiations visibles, de telle sorte que le système nerveux qui lui est associé, aidé des résultats acquis lors d'expérience antérieure et des informations fournies simultanément par les autres organes des sens puissent analyser les éléments pertinents du 'contact visuel', tels que la luminance, la couleur, la taille, la position, la forme, le mouvement, la profondeur et la signification »

M. Imbert, Physiologie T2 (Ch. Kayser Ed.) Masson, Paris, 1976

La vision n'est donc pas limitée à un capteur qui serait l'œil, mais il s'agit d'un système intégré à composantes multiples traitant l'information en fonction du contexte et de l'environnement.

L'oculométrie est une méthode non invasive permettant de collecter des indices du fonctionnement cognitif, qui est bien adapté à des situations dynamiques. Elle peut permettre de comprendre le traitement de l'information en temps réel (par exemple, Pearson et Sahraie, 2003 ; Zelinsky, 2008), en analysant où l'attention est dirigée sur un affichage et combien de ressources sont appliquées à un objet particulier. Il existe de nombreuses études sur le suivi du regard comme un moyen d'évaluer l'utilisabilité des interfaces (Baccino, 2002 ; Jacob & Karn, 2003). L'oculométrie a également

été utilisée dans des domaines tels que la lecture (par exemple, Rayner, 2009), la prise de décision (par exemple Morrison, Marshall, Kelly, et Moore, 1997 ; Vallières, Chamberland, Vachon, et Tremblay, 2013) et la mémoire (par exemple, Theeuwes, Belopolsky, et Olivers, 2009 ; Tremblay, Saint-Aubin, et Jalbert, 2006).

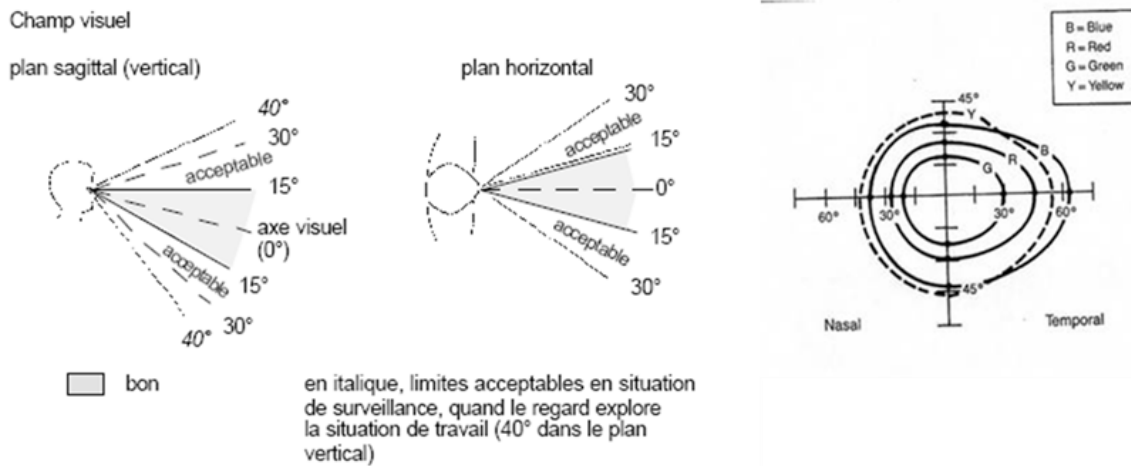
L'œil a nombre de limitations tout comme les appareils photographiques. La vision précise ne représente que 1-5° d'angle visuel (vision fovéale) du champ visuel. La zone de confort visuel telle que définie dans la norme AFNOR NF X 36-101-2 couvre +/- 15° du champ visuel, dans cette zone les couleurs sont correctement perçues. En dehors de la plage de +/- 30°, l'angle visuel est considéré comme non acceptable pour l'affichage information. Le temps de vision est surtout composé de fixations, le mouvement rapide de la fovéa quand l'attention est dirigée vers un autre endroit est appelé saccade, même pendant les fixations. L'œil effectue de très petits mouvements appelés microsaccades (Hubel, 1988). Les saccades vont de 10 ms à 100 ms, les fixations sont considérées comme significatives de l'attention portée à un objet à un seuil de 150 ms (Serenio et Rayner, 2003), 90 % du temps de vision est dédié aux fixations. Pour simplifier la modélisation de l'œil, les microsaccades pendant les fixations sont souvent considérées comme du bruit, celui-ci étant inférieur à 5° d'angle visuel (Carpenter, 1977, p.105). Ce seuil peut être utilisé à des fins de filtrage des coordonnées brutes du regard. Le système visuel humain peut également poursuivre des cibles en mouvement (Reimer et Sodhi, 2006).

Chaque type de mesure physiologique ayant des limites, des études sont réalisées en utilisant une approche multisenseurs permettant de compenser les limites de chaque capteur pris indépendamment. Dans la suite de ce document, nous détaillerons plus avant l'utilisation des systèmes oculométriques et des électroencéphalogrammes (EEG) que nous avons utilisés dans ce travail de thèse.

La vision est généralement considérée comme *fovéale* ou *périphérique*. La vision *périphérique* a des caractéristiques spécifiques ; le mouvement y est correctement perçu au contraire de la couleur dont la perception est directement corrélée à l'angle visuel (Figure 22 : Champ visuel et vision des couleurs en fonction de l'écart à la zone fovéale). Le bleu et le jaune sont perçus dans un angle visuel plus large que le rouge et le vert. La sensibilité au mouvement dans la vision périphérique est corrélée à la vitesse de déplacement ; une vitesse plus élevée est préférable dans la région périphérique (Boff et Lincoln 88). Pour corréler un angle visuel avec un objet sur un écran ou des coordonnées pixels, il est nécessaire de connaître avec précision la distance entre l'œil et l'écran et la taille en cm de cet objet sur l'écran, il peut être calculé avec la formule suivante :

$$AngleVisuel = 2 \times \tan^{-1}\left(\frac{size}{2 \times Distance}\right)$$

Par exemple, l'angle de confort visuel défini dans la norme AFNOR NF X 36-101-2 (+/- 15 °) sur un écran situé à 70 cm de la tête d'un sujet représente un cercle de 18,75 cm de rayon.



Limites d'utilisation du champ visuel, d'après NF X 35-101-2 (1997).

Figure 22 : Champ visuel et vision des couleurs en fonction de l'écart à la zone fovéale

Divers systèmes ont été conçus et utilisés pour suivre le mouvement des yeux (voir Duchowsky, 2007 pour plus de détails). Ces oculomètres peuvent être montés sur la tête (portés) comme le système Pertech (<http://www.pertech.fr/>) ou posés sur une table (Tobii – www.Tobii.com, Facelab - www.seeing-machines.com, SMI – www.smivision.com) c'est-à-dire déportés (Figure 23 : oculomètres portés ou posés). Ils utilisent des caméras et des logiciels de traitement pour calculer la position du regard à partir de la réflexion d'une source émissive infrarouge sur la cornée. Pour augmenter la précision des dispositifs posés, il est possible d'utiliser une mentonnière fixée sur la table, afin de limiter les mouvements de tête.

Avec les systèmes portés, la position de la tête du sujet dans la scène peut être calculée, mais cela nécessite un calcul complexe basé sur l'analyse de la vidéo prise par une caméra frontale. En général, des indices (vignettes avec une trame) facilitant la calibration doivent être disposés sur la scène pour que le logiciel de traitement puisse faire ce calcul. Dans le cas de l'oculomètre déporté, lorsqu'il n'est pas intégré à un écran, il est nécessaire de décrire avec précision la configuration de l'appareil pour l'expérimentation. Une scène 3D donnant la position et la distance du ou des écrans avec leur taille et leur résolution doit être créée.

Un processus d'étalonnage est également obligatoire pour assurer la précision du système. Habituellement, le processus de l'étalonnage consiste à afficher plusieurs points consécutifs dans toutes les parties de la scène de visualisation que le sujet doit regarder au fur et à mesure de leur apparition, le logiciel de suivi de l'œil va calculer une transformation qui fait correspondre la position de la tête par rapport à l'environnement et estimer la position du regard sur le ou les écrans définis dans la scène (Goldberg & Wichansky, 2003).

Les oculomètres déportés sont généralement binoculaire et permettent de calculer la vergence des yeux et calculer les coordonnées brutes du point d'intersection du regard (GIP : Gaze Intersection Point) en coordonnées x,y pixels appliqués à un écran en temps réel . Cette fonction permet entre autres d'intégrer les coordonnées pixel du regard dans l'application à des fins de monitoring ou comme moyen d'interaction.

Des zones statiques d'intérêt (AOI : Area Of Interest) peuvent être définies et des informations spécifiques seront envoyées lorsque le regard rencontre une AOI. Cette fonction peut par exemple être intéressante dans l'analyse de l'utilisation d'une page web. Il est possible de définir une AOI recouvrant un bandeau de publicité, à chaque fois que le regard se portera sur cette zone, l'oculomètre donnera des coordonnées brutes et l'éventuelle AOI sur laquelle s'est porté le regard.

Le suivi des zones d'intérêt dynamique (dAOI : dynamic Area Of Interest) est beaucoup plus complexe puisqu'il faut notifier à l'oculomètre en temps réel la position de zones d'intérêt pouvant se déplacer. A notre connaissance, aucun dispositif oculométrique ne permet de la faire, mais en 2012, Jambon et Luengo (2012) ont proposé une architecture logicielle qui permet la détection d'une fixation de zones d'intérêt dynamiquement basée sur l'utilisation du bus logiciel Ivy (Buisson & al, 2002). Les données issues de l'oculomètre sont traitées avec celles de positionnement des dAOI dans une brique logicielle spécifique qui détermine les fixations sur ces zones. Une autre étude propose un outil pour la détection dynamique des AOI sur une vidéo ou une animation (Papenmeier, Frank & Huff, Markus, 2010).

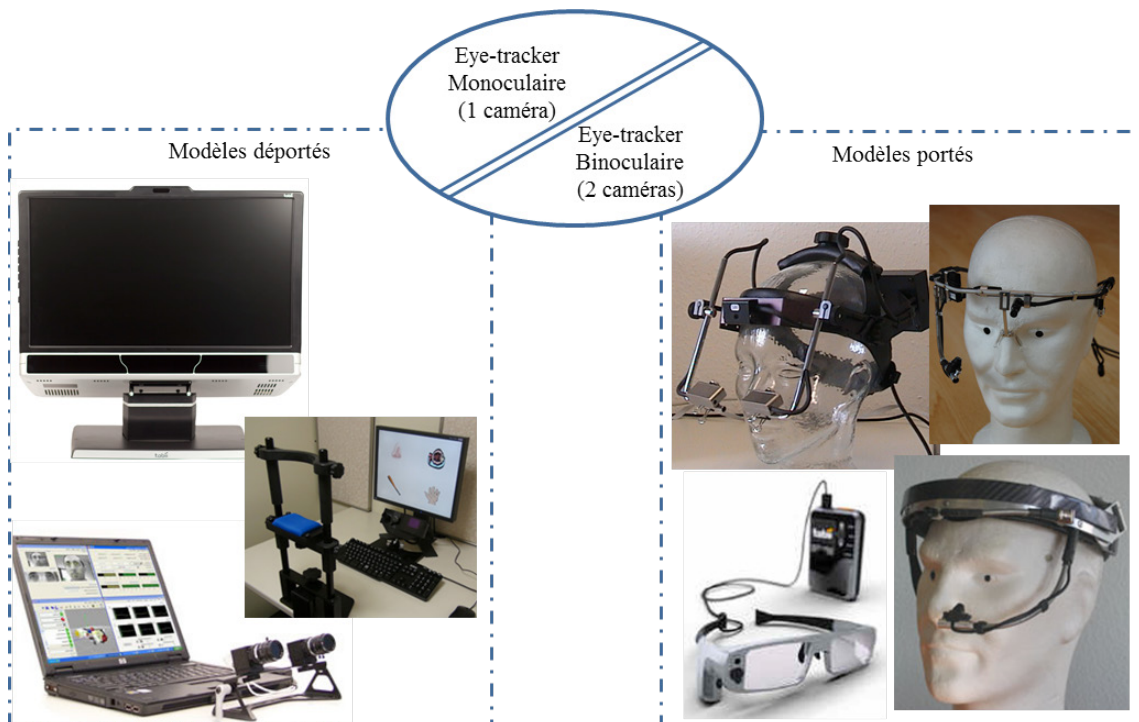


Figure 23 : oculomètres portés ou posés

Il existe deux sortes de méthodes pour recueillir des données de suivi du regard. La première et certainement la plus courante est d'utiliser le logiciel fourni par le fabricant de l'appareil. Ces outils peuvent enregistrer ou visualiser les données de suivi du regard. La seconde est d'utiliser un SDK (System Developer Kit) fourni avec l'oculomètre et de développer le module de logiciel spécifique à l'aide de l'API (Application Programming Interface) de l'oculomètre qui sera relié à l'appareil pour recueillir des données.

Divers paramètres auront un impact sur la précision des données brutes délivrées par le système de suivi du regard. Parmi eux, la fréquence de rafraîchissement (en Hertz) et la résolution de la caméra sont critiques pour le logiciel de suivi. Les systèmes existants utilisent une fréquence de capture qui va de 30 Hz à 1000 Hz. Une fréquence élevée permet d'améliorer la

précision des données en donnant la possibilité de faire des filtrages complexes, mais provoquera également l'augmentation considérable de la taille des données enregistrées ainsi que leur temps de traitement.

Les données physiologiques de l'œil peuvent être traitées en ligne à des fins d'analyse, afin d'adapter dynamiquement l'IHM ou bien d'utiliser le regard comme un dispositif de pointage (Sibert & Jacob, 2000). Les fixations, les saccades et la poursuite d'objets lents (Reimer & Sodhi, 2006) peuvent être détectées à partir des coordonnées brutes du regard sur l'écran. Pour associer ces informations avec l'IHM, des données doivent être collectées à la fois en provenance de l'oculomètre (coordonnées brutes, diamètre pupillaire, saccades etc.) et de l'IHM. Par exemple, les coordonnées de tous les objets fixes ou mobiles pour lesquels on souhaite savoir s'il y a eu fixation ainsi que des événements applicatifs (actions de haut niveau) et des événements détectés par rapport à l'activité de l'utilisateur peuvent être nécessaires. Les fixations et la poursuite lente sont considérées comme le désir de fixer son attention sur un objet spécifique (Goldberg & Kotval, 1998 ; Recarte & Nunes, 2003). Au contraire, les saccades sont la manifestation de la volonté de changer le centre de l'attention. Le nombre de fixations sur un objet spécifique est considéré comme un indicateur de l'importance attachée à celui-ci (Poole, Ball & Phillips, 2004). La durée de fixation peut être reliée aux ressources attentionnelles associées à la conception de l'objet (Just & Carpenter, 1976) ou au niveau d'engagement associé à l'objet durant la tâche. D'autres données peuvent être collectées à partir des systèmes d'oculométrie comme les clignements d'yeux ou le diamètre de la pupille. La durée du clignement d'œil peut aller de 100 ms à 400 ms. Elle peut être utilisée pour évaluer la charge de travail (Brookings, Wilson, & Swain, 1996 ; Bruneau, Sasse & McCarthy, 2002). La variation du diamètre de la pupille peut également être utilisée comme une indication de la charge de travail (Marshall, 2000 ; Pomplun & Sunkara, 2003), mais les sources de lumière (environnement, informations affichées sur l'écran) du dispositif expérimental doivent être strictement contrôlées puisque le réflexe pupillaire lié au changement de lumière est plus prononcé que l'impact de la cognition sur la taille de la pupille. En outre, même la luminance de la zone de fixation (alors que la luminance de l'écran de l'ordinateur ne change pas) a un impact sur la taille de la pupille. La dilatation de la pupille est en corrélation avec la demande cognitive perçue de la tâche (Matthews, Middleton, Gilmartin & Bullimore, 1991), l'effet de cette demande sur la pupille a été défini par Beatty (Beatty, 1982) « TEPR » (Task Evoked Pupillary Response). La trajectoire du regard peut également être analysée pour évaluer l'utilisation de l'IHM et la confronter à la méthode de travail préconisée ; par exemple, dans le cockpit un circuit visuel est enseigné qui balaye les différents indicateurs dans un certain ordre, l'oculomètre permet de valider cet enseignement.

Après la collecte de données et la phase de tri et de remise en forme des données, celles-ci peuvent être analysées pour en déduire des liens de cause à effet, les statistiques et l'analyse du comportement. De nombreux travaux sont disponibles dans ce domaine, citons deux exemples récents avec le cube de visualisation de l'espace-temps (Li, Çoltekin, Kraak, 2010) et les usages de l'algorithme de regroupement de trajectoires (bundling) (Hurter, Ersoy, Fabrikant, Klein, Telea, 2013).

4.2.3.2. Mesures neurophysiologiques

Les mesures neurophysiologiques regroupent plusieurs techniques dont l'électro-encéphalographie (EEG) qui consiste à mesurer l'activité cérébrale à l'aide d'un casque comportant des électrodes, l'imagerie par résonance magnétique fonctionnelle (IRMf) qui permet d'observer la variation des propriétés du flux sanguin lorsque des zones du cerveau sont stimulées, la fNIR (functional near-infrared spectroscopy) qui consiste à mesurer l'oxygénation d'une zone du cerveau pour en déduire son activité.

Dans le domaine de l'EEG qui sera mis en œuvre dans ces travaux de recherche, les électrodes mesurent l'activité électrique générée par l'activation de zones du cerveau, on parle alors d'activité neuroélectrique. Deux types de mesures peuvent être effectuées, l'analyse continue qui consiste à étudier les ondes cérébrales Alpha, Beta, Delta, Gamme et Thêta (Figure 24: ondes cérébrales et EEG), et la méthode dite des potentiels évoqués ERP (Event Related Potential) qui consiste à étudier la réponse du cerveau à un stimulus.

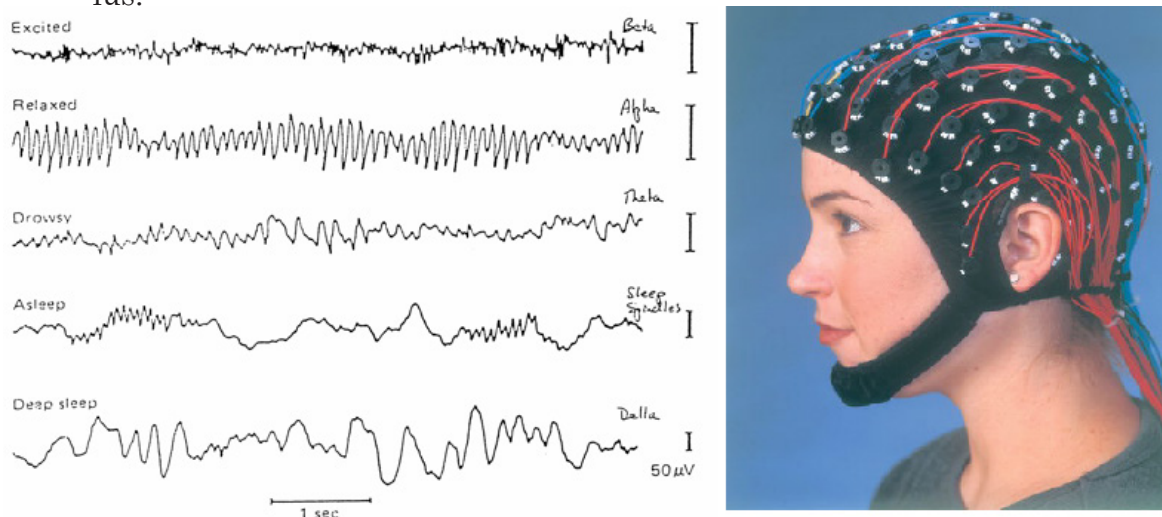


Figure 24: ondes cérébrales et EEG

4.2.3.2.1. Méthode continue

L'EEG présente l'activité neuroélectrique collectée sur plusieurs électrodes résultant de la sommation des potentiels d'action post-synaptiques d'un grand nombre de neurones. L'analyse de cette activité et leur classification par gamme de fréquences ont mis en évidence cinq types de signaux :

- • Alpha (8-12 Hz), apparaissant pendant la relaxation ou lorsque les yeux sont fermés,
- • Beta (12-30 Hz), retrouvée pendant l'alerte, l'activité intellectuelle, la concentration ou l'anxiété,
- • Delta (jusqu'à 4Hz), retrouvées au cours du sommeil profond,
- • Gamma (30-100+ Hz), apparaissant au cours d'une intense activité mentale et dans certaines fonctions motrices,
- • Thêta (4-7 Hz), retrouvées dans la somnolence.

Dans ces signaux, l'évolution des ondes Alpha et Thêta est particulièrement utilisée dans les études sur le changement d'exigence de la tâche. La baisse

globale de l'activité Alpha est reliée à des niveaux de vigilance augmentée, alors que la diminution de l'activité Thêta est liée au traitement de l'entrée visuelle (Barry, Clarke, Johnstone, Magee, Rushby, 2007). Les signaux EEG ont été étudiés dans une multitude d'exécution de tâches cognitives. L'activité Thêta semble être liée à des processus de mémoire, l'encodage et la récupération ainsi que le transfert d'informations entre la mémoire à court et à long terme (Sauseng et al., 2006), (Ward, 2003). L'activité Thêta augmente lors de la récupération dans la mémoire (Klimesch, Doppelmayr, Schimke, Ripper, 1997), alors que la diminution de l'activation de l'Alpha semble refléter l'activation de la mémoire à long terme (Sauseng et al., 2006), ou des changements dans l'attention requise pour une tâche (Ward, 2003). Les changements liés à une tâche dans la bande Thêta sont généralement plus forts dans les régions frontales (Figure 25 : les zones du cerveau), tandis que les changements de l'activité Alpha sont plus prononcés dans les zones postérieures (Klimesch, Doppelmayr, Schimke, Ripper, 1997). L'activité EEG Thêta dans les zones frontales semble augmenter lorsque la demande d'une tâche dans les études sur la mémoire de travail (Smith, McEvoy, Gevins, 2002), les tâches de recherche visuelle (Yamada, 1998), les tâches multiples (Fournier, Wilson, Swain, 1999). En plus des sites frontaux, l'augmentation de l'activité Thêta a également été signalée dans les zones pariétales en réponse à l'augmentation des exigences des tâches (Fairclough, Venables, 2006). L'activité Thêta frontale est également reliée à une demande accrue de l'attention portée (Doppelmayr, Finkenzeller, Sauseng, 2008). Inversement, l'activité Alpha sur la zone pariétale semble diminuer avec l'augmentation de la charge de la mémoire de travail (Fairclough, Venables, 2006), (Ryu, Myung, 2005).

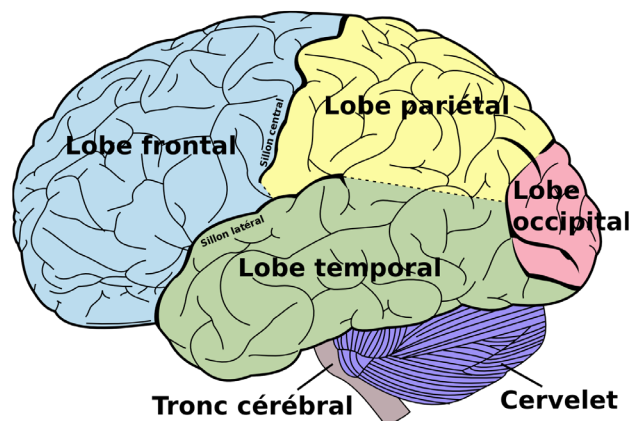


Figure 25 : les zones du cerveau

Le niveau d'activité Thêta permet de différencier une situation de simple tâche ou de baseline (condition de calibration) d'une situation multitâche (Fairclough, Venables, 2006). La pression temporelle dans une tâche visuo-motrice augmente l'activité Thêta des zones frontales et centrales, et diminue le signal Alpha dans la zone pariétale (Sloubounov, Fukada, Simon, Rearick, Ray, 2000).

Les différentes études dans la littérature scientifique semblent montrer une qu'une augmentation de l'activité Thêta dans la zone frontale associée à une diminution de l'activité Alpha dans les zones pariétales est corrélée à une augmentation de la charge de travail.

4.2.3.2.2. Méthode ERP (Evoked Related Potential)

Le Potentiel Evoqué (PE ou ERP en anglais) représente la réponse du cerveau à un stimulus sensoriel, moteur ou lié à un événement cognitif. L'ERP est calculé par moyennage des données brutes provenant de l'EEG durant la tâche. Les applications de l'ERP dans le domaine des facteurs humains portent sur l'évaluation de la charge de travail, de la vigilance, de la surveillance, de la fatigue bien que la plupart des études portent sur l'évaluation de la charge de travail. Voir Fu et Parasuraman (2007) pour une liste exhaustive de ces travaux et les types d'ERP étudiés.

Différents ERP sont étudiés dans la littérature (N100, P300, N400, P600, VCN, error-related negativity) ; nous nous intéresserons plus particulièrement à la N100 et la P300 qui permettent de caractériser la détection d'un stimulus, sa prise en compte et d'évaluer la charge de travail.

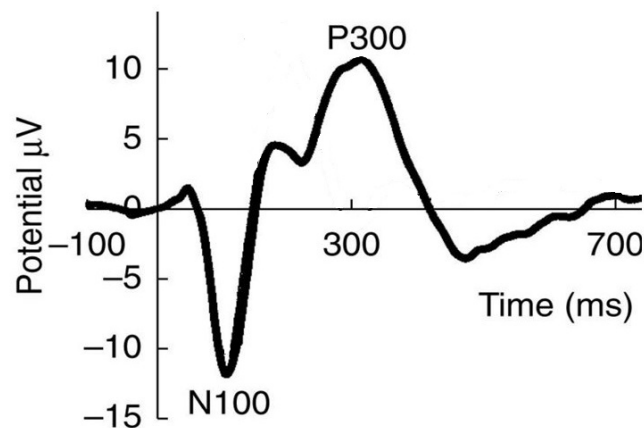


Figure 26 : ERP N100 et P300

La N100, ou N1, est un potentiel évoqué négatif (N) apparaissant entre 75 et 180 ms après la présentation d'un stimulus (Figure 26 : ERP N100 et P300). La N100 est répartie principalement dans les aires frontales et centrales du cerveau. Ce PE apparaît après un stimulus imprévisible ; qu'il soit auditif (Spreng, 1980 ; Keidel & Spreng, 1965) ou visuel (Warnke, Hennighausen, Renschmidt, 1994). La N100 est dite préattentive, c'est-à-dire qu'elle se produit de façon précoce, avant la prise de conscience et le traitement cognitif du stimulus. Dans le cas des stimuli auditifs, son amplitude peut varier notamment selon l'amplitude des sons, leur fréquence, l'intervalle de temps qui sépare 2 stimuli et la variation de fréquence avec le son.

La P300, ou P3, est un potentiel évoqué positif (P) apparaissant vers 300 ms après le stimulus et mis en évidence par Sutton, Braren & Zubin en 1965. Elle est principalement étudiée dans les analyses de charge de travail (Humphrey et Kramer, 1994). La P300 apparaît dans la région pariétale et est associée au traitement de stimuli associés à la tâche et rares comme une alarme. Son amplitude est d'autant plus importante que son occurrence est faible ; plus il est difficile de discriminer le stimulus plus l'apparition de la P300 est retardée (Kutas, Mc Carthy, Donchin, 1977). Il a été aussi montré que l'amplitude de la P300 est corrélée avec la quantité de ressources attentionnelles allouées à la tâche (Johnson, 1984), en clair en contexte multitâche, la P300 se trouvera diminuée si la charge de travail est élevée.

La procédure la plus couramment employée pour provoquer la P300 dans le cas des stimuli visuels ou sonores est le paradigme Oddball (Prinzel,

Freeman, Scerbo, 1994 ; Kramer, Trejo, Humphrey, 1995) ; on présente aux sujets deux types de stimuli, l'un fréquent, l'autre rare qui implique une réaction du sujet. L'amplitude de la P300 est significativement plus élevée sur les stimuli rares que sur les fréquents. L'amplitude de la P300 va diminuer sur les stimuli rares si la difficulté de la tâche principale et donc la charge de travail augmente. Ce phénomène a été vérifié entre autres dans le domaine du jeu vidéo ; en agissant sur le niveau de difficulté, un effet significatif sur l'amplitude de la P300 sur la perception de sons rares a été mis en évidence (Alisson, Polish, 2008). Dans le cockpit, ce paradigme a été utilisé sur une tâche de pilotage en approche finale (Fowler, 1994), la P300 était obtenue grâce à des tâches secondaires visuelles et sonores (flashes ou sons). Dans cette expérience la latence de la P300 a été plus significativement corrélée à la charge de travail que son amplitude.

4.3. Résumé

La complexité des interfaces de supervision rend problématique leur évaluation. Le cadre de l'ingénierie cognitive (Rasmussen et al., 1994) propose quatre étapes clés dans la recherche cognitive appliquée qui ont pour but d'atteindre différents niveaux entre la validité écologique et le contrôle expérimental : les tests de laboratoire, les tests de laboratoire appliqués au contexte, les tests dans un environnement de simulation complexe, les observations dans un environnement opérationnel.

Les deux premiers niveaux permettent en conservant un contrôle suffisant de réaliser facilement des expérimentations, les tests en environnement de simulation nécessitent des développements coûteux et complexes pour obtenir des indications de performance dans le déroulement de la tâche.

Différentes méthodes permettent d'évaluer la performance de l'humain lors de l'utilisation d'une interface homme-machine, les méthodes dites subjectives et les méthodes qualifiées d'objectives. Les méthodes subjectives reposent essentiellement sur l'auto-évaluation (par exemple évaluation périodique de la charge cognitive perçue par l'opérateur pendant l'activité), des questionnaires le plus souvent remplis à la fin du protocole expérimental, de l'observation ou de la verbalisation durant l'activité. Les méthodes objectives reposent sur la mesure de la performance durant l'activité au moyen de métriques basées sur l'utilisation des interfaces ou à l'aide de capteurs physiologiques permettant d'avoir une indication de l'impact de leur utilisation sur l'humain.

Cinq familles de capteurs physiologiques sont couramment utilisées lors d'études en laboratoire ou en situation pour la mesure de la performance des opérateurs. Dans ces travaux nous nous intéresserons plus particulièrement aux dispositifs oculométriques et neurophysiologiques. Les systèmes oculométriques permettent en analysant les fixations sur les objets, d'obtenir des éléments concernant l'utilisation de l'interface et d'avoir une estimation de la charge de travail grâce à l'analyse des trajets oculaires ou du diamètre pupillaire. Les dispositifs électroencéphalographiques peuvent être utilisés dans deux types d'analyses : continue, ou potentiel évoqués. Dans la mesure continue, les signaux électriques en provenance des différentes électrodes placées sur le cerveau sont analysés du point de vue fréquentiel et découpés en cinq gammes d'ondes (Alpha, Béta, Delta, Gamma et Thêta). Leur analyse permet d'obtenir des éléments sur le fonctionnement du cerveau durant l'activité ; à titre d'exemple, une augmentation de l'activité Thêta dans la zone frontale associée à une diminution de l'activité Alpha dans la zone pariétale est corrélée à une augmentation de la charge de travail. La méthode d'analyse des potentiels évoqués consiste à observer la réponse du cerveau à un stimulus sensoriel, moteur ou lié à un événement cognitif. L'analyse de l'amplitude de la P300 (une onde apparaissant environ 300ms après un stimulus) permet de déduire la complexité d'une tâche pour le sujet et d'estimer la charge de travail.

Un système performant pourra un jour s'adapter dynamiquement à l'état cognitif de l'opérateur grâce à des capteurs physiologiques pour par exemple modifier en temps réel la représentation de l'information, notifier de l'état dégradé d'une position de travail (couple humain-automate) ou bien utiliser des contres mesures cognitives en cas de défaillance de l'opérateur humain.

La suite de ce manuscrit présente la problématique de la thèse, les choix méthodologiques pour adresser cette problématique et les expérimentations réalisées dans le cadre de ces travaux.

Problématique de la thèse

*Améliorer la conscience de la situation
des contrôleurs aériens*

Ce travail de recherche a pour domaine d'application le contrôle aérien. Celui-ci est dans une phase de profondes mutations liées à l'intégration des services nationaux de prestation de service de contrôle aérien au sein de l'Europe ainsi qu'aux évolutions technologiques amenant des changements radicaux des outils permettant d'assurer ce service. Le trafic aérien dans le monde est en augmentation depuis plusieurs décennies. L'enjeu principal de cette augmentation est de faire évoluer les outils et méthodes pour permettre aux contrôleurs d'assurer le service dans de bonnes conditions de sécurité tout en ayant la responsabilité d'un nombre plus important d'avions que par le passé.

Le projet SESAR dans lequel s'insèrent ces travaux a pour objectif de moderniser les systèmes existants en Europe dans le but de tripler la capacité de l'espace aérien, améliorer la sécurité, limiter l'impact du trafic accru sur l'environnement tout en diminuant le coût associé au service. Pour réaliser ces objectifs, différentes pistes d'amélioration du système permettant de réaliser le service sont envisagées, dont l'automatisation partielle du contrôle mis en œuvre pour les positions de travail 'En-Route' dans le cadre du work-package 4.7.2 du projet SESAR dans lequel se sont insérés ces travaux. Ce projet comporte deux volets, le premier concerne le développement d'un automate visant à supprimer les conflits entre avions en agissant sur leur vitesse. Le second porte sur le développement d'une nouvelle position de contrôle intégrant cet automate et des outils permettant de gérer un nombre très important d'avions.

Pour réaliser leur tâche de supervision, les contrôleurs doivent surveiller les trajectoires de tous les avions présents dans leur secteur de contrôle et s'assurer du bon suivi des consignes qu'ils leur donnent ainsi que du déroulement conforme de leur plan de vol. L'augmentation du nombre d'avions gérés par le contrôleur, qui est un des objectifs du projet SESAR implique mécaniquement un nombre beaucoup plus important qu'auparavant d'informations à traiter et mémoriser. Or une part importante des causes d'incidents dans le domaine du contrôle aérien est déjà due à des problèmes de perception et de mémorisation. Ces problèmes de perception sont particulièrement critiques dans le cas de la perception des alarmes et des avertissements donnés par le système sur la visualisation radar. En cas de défaut de détection de conflits entre aéronefs, un outil d'aide, le filet de sauvegarde, est là pour mettre en évidence sur la supervision radar les avions dont les trajectoires vont aboutir à un passage sous la norme de séparation. Percevoir ces alarmes le plus rapidement possible, c'est-à-dire éventuellement interrompre le contrôleur pendant le déroulement de sa tâche est donc nécessaire pour ce type d'information. D'autres informations, moins critiques, peuvent être notifiées au contrôleur, dans ce cas l'interruption n'est pas forcément nécessaire, mais leur design doit permettre de les mettre en évidence. Le design actuel de ce type d'alertes en France qui n'utilise que la couleur fait aujourd'hui débat et il a été recommandé dans un bulletin de sécurité aérienne de le faire évoluer. Compte tenu de la taille des écrans de supervision radar, il n'est pas rare pour un contrôleur de manquer l'apparition d'une alerte parce que son attention est maintenue par la gestion d'une situation dans une autre partie de l'écran ; le travail en binôme permet entre autres de limiter les conséquences de ce type de problèmes. Néanmoins, les problèmes de perception et la mémorisation des informations courantes des avions ainsi que la

détection tardive des notifications ne feront que s'aggraver avec l'augmentation du nombre d'avions à gérer par les contrôleurs.

La perception des informations nécessaires à l'établissement d'une bonne conscience de la situation aérienne des contrôleurs est au centre de cette thèse. L'objet principal de ce travail est d'améliorer la conscience de la situation des contrôleurs en s'assurant que les éléments pertinents à leur disposition dans le déroulement de leur tâche sont perçus dans des délais conformes à leur importance et que les actions nécessaires qui en découlent sont bien réalisées.

Nous reformulons cette problématique de conscience de la situation et de la perception des informations nécessaires à la réalisation de la tâche de contrôle sous la forme de plusieurs questions :

- Quelles sont les informations nécessaires au contrôleur pour maintenir une image mentale correcte de la situation aérienne ?
- Comment détecter que ces informations n'ont pas été perçues et aider le contrôleur à les percevoir ?
- Dans le cas des alertes propres aux systèmes de supervision, comment s'assurer que leur design a un pouvoir attentionnel suffisant et conforme à l'importance de l'information dans le déroulement de leur tâche ?

Afin de répondre à ces questions de recherche, notre travail s'est porté sur deux axes. Le premier consiste à étudier la tâche des contrôleurs de façon à analyser les informations nécessaires à la réalisation de certains objectifs et proposer un agent de suivi de la tâche qui pourrait épauler le contrôleur et améliorer sa conscience de la situation. Le second volet porte sur la perception des notifications et plus particulièrement en vision périphérique et la conception de design propres à améliorer leur perception ainsi que leur impact sur la réalisation de la tâche.

La deuxième partie de cette thèse présente les développements et expérimentations réalisées pour concevoir et évaluer une position de contrôle dont le design a été orienté pour améliorer la conscience de la situation des contrôleurs. Le prochain chapitre décrit la conception et le fonctionnement de l'agent d'aide au suivi des informations nécessaires à la réalisation de la tâche de contrôle ainsi que les différents choix méthodologiques que nous avons faits pour évaluer les designs mis en œuvre pour la représentation des notifications issues de cet agent. Les deux chapitres suivants décrivent les expérimentations conduites ainsi que les résultats associés pour évaluer plusieurs types de designs de notification visuelles et leur impact sur la tâche de contrôle. L'avant-dernier chapitre présente une expérimentation mettant en œuvre l'agent de suivi de la tâche lors de simulations de contrôle et les résultats préliminaires obtenus sur son impact dans la réalisation de la tâche. Enfin, le dernier chapitre présente les conclusions de ce travail de recherche ainsi que les perspectives de recherche que nous envisageons.

PARTIE II : DÉVELOPPEMENTS ET VALIDATION

Avion pout

Avion notifié

37	AF660KU	208	238
	LIBU B5 t60	220	210
	DINGUE...	200	
		190	
		180	
		170	

Avion statique

Notifications

- TAP172 State: Flight not shoeted
- BMM777 State: Flight not shoeted
- Root: Flight not shoeted
- 17: Previous position did not shoot the Flight
- 400: Previous position did not shoot the Flight
- Warn: Previous position did not shoot the Flight
- 44: Previous position did not shoot the Flight
- AAL524 State: Flight not shoeted
- IBE925 State: Flight not shoeted
- DLH419 State: Previous position did not shoot the Flight
- BA15972 State: Previous position did not shoot the Flight
- DLH059 State: Flight not shoeted
- AUA359 State: Previous position did not shoot the Flight
- NJE211 State: Previous position did not shoot the Flight
- RAE17: Previous position did not shoot the Flight

44 BMM773
347 \ 290
Juven M2 t290

50 BMM773
360
Juven M2 t360
k430 -> Juven

45 AZA806
360 -
Vulca N3 t360

48 DLH824
380 -
Mozao G2 t380

45 SAS248
400 -
Sevet G2 t400

Niveau 1: warning couleur

Niveau 2: animation couleur

Niveau 3: animation cadre étiquette

5. CONCEPTION D'UN DESIGN ORIENTÉ CONSCIENCE DE LA SITUATION

5.1. Introduction

Ce travail de recherche a pour principal objectif d'améliorer la conscience de la situation des contrôleurs aériens.

Nous avons mis en évidence plusieurs interrogations liées à cette problématique :

- *Quelles sont les informations nécessaires au contrôleur pour maintenir une image mentale correcte de la situation aérienne ?*
- *Comment détecter que ces informations n'ont pas été perçues et aider le contrôleur à les percevoir ?*
- *Dans le cas des alertes ou notifications propres aux systèmes de supervision comment s'assurer que leur design a un pouvoir attentionnel suffisant et conforme à l'importance de l'information dans le déroulement de leur tâche ?*

Pour répondre à ces questions, les travaux conduits lors de cette thèse comportent deux volets qui se sont traduits par trois expérimentations. Le premier volet porte sur la conception d'un système d'aide dont l'objectif est de mettre en valeur des informations qui doivent normalement être détectées par le contrôleur dans son processus de supervision (questions 1 et 2). Le second volet porte sur la présentation des informations issues des systèmes d'aide en général et le compromis à établir entre leur mode de représentation et leur importance dans la tâche (question 3).

Séquentiellement nous avons commencé par étudier l'aspect visualisation et la manière de notifier les alertes visuelles de façon à établir une hiérarchie en termes de pouvoir attentionnel et d'impact sur la tâche. Pour cela nous avons utilisé le micro-monde ATC Laby présenté au chapitre 4 nous permettant de réaliser une évaluation dans un contexte assez réaliste tout en conservant un bon contrôle expérimental. Par la suite, nous avons étudié d'une façon plus globale la tâche du contrôleur avec des contrôleurs experts pour identifier les informations dynamiques nécessaires à l'établissement d'une bonne conscience de la situation. Cette étude a été réalisée dans le but de développer un agent destiné à les aider dans leur tâche de supervision. Cet agent détectant des informations dynamiques pertinentes pour la conscience de la situation doit les transmettre à la visualisation pour les notifier. Ce nouveau couple visualisation-automate a été évalué dans le contexte des simulations du projet WP4.7.2 décrit au chapitre 1.

Dans les sections suivantes, nous présentons les développements qui ont été réalisés pour évaluer différents types de design de notifications visuelles puis l'agent de suivi de la tâche des contrôleurs destinés à une nouvelle visualisation dont le design est orienté conscience de la situation. L'approche méthodologique mise en œuvre lors de l'évaluation des designs de notifica-

tions ainsi que leur emploi dans une visualisation mettant en œuvre l'agent de suivi de la tâche, y sont par ailleurs décrits.

5.2. Évaluation du design de notifications visuelles

5.2.1. Évolutions de Laby vers un environnement de test multitâche

Les notifications visuelles utilisées dans les positions de travail opérationnelles pour les alertes ou les avertissements ont été décrites au chapitre 1.1.2. Les retours d'expérience des contrôleurs attestent que ceux-ci peuvent parfois ne pas les percevoir lorsque leur attention est attirée dans d'autres parties de l'écran. Dans le chapitre 2, nous avons abordé les différents types de notifications ainsi que les différentes variables visuelles sur lesquelles il est possible d'agir pour modifier leurs propriétés. Enfin, dans le chapitre 4 nous avons décrit plusieurs approches méthodologiques visant à des compromis différents au niveau contrôle expérimental en opposition au réalisme opérationnel, ainsi que les méthodes d'évaluation de l'utilisation d'une interface homme-machine.

Parmi les différentes approches méthodologiques décrites, l'utilisation d'un micro-monde nous a semblé la plus adaptée pour évaluer et mesurer la performance de signaux visuels dans un environnement suffisamment écologique par rapport à la tâche de supervision des contrôleurs aériens. Pour réaliser les deux études sur les notifications visuelles et leur confrontation à des notifications sonores, nous nous sommes appuyés sur Laby décrit au chapitre 4.1.2.3 (Parise, Imbert, Marais, 2012 ; Imbert, Hodgetts, Parise, Vachon & Tremblay, 2014). En apport des mesures comportementales offertes par Laby sur la performance dans la réalisation de la tâche de guidage des avions, nous avons intégré dans notre architecture l'utilisation de deux types de capteurs physiologiques (EyeTracker et EEG) décrits au chapitre 4. Le premier dans le but d'avoir des informations sur l'utilisation de Laby grâce aux trajets visuels, le second pour mesurer par des mesures neurophysiologiques l'impact de plusieurs designs sur une tâche de détection de signaux sonores dans les mêmes conditions que l'expérimentation précédente.

Pour répondre à la troisième interrogation de notre problématique, nous avons modifié Laby afin d'intégrer le paradigme de tâches multiples comme permet de le faire NASA MATB. Ce paradigme a été fréquemment utilisé dans des études sur la cognition. Il permet d'évaluer l'allocation de ressources attentionnelles entre différentes tâches ainsi que leurs impacts respectifs dans la réalisation des tâches. Ces évolutions ont permis de réaliser plusieurs expérimentations dont deux sont présentées dans ces travaux (expérimentation Saillance et expérimentation Bip).

Nous souhaitons dans un premier temps, utiliser cette plate-forme expérimentale pour tester le pouvoir attentionnel de plusieurs types de notifications en vision périphérique ainsi que l'impact de ces designs dans la réalisation de la tâche principale de Laby (guidage dans le labyrinthe et saisie des consignes). Pour cela, des modifications ont été apportées à Laby afin

de pouvoir ajouter une nouvelle tâche consistant à valider des notifications présentées en vision périphérique sur des avions non guidés et statiques (ne se déplaçant pas sur l'image radar).

Pour déterminer à quelle distance montrer les notifications par rapport à l'avion guidé dans le labyrinthe, nous nous sommes appuyés sur la norme AFNOR NF X 36-101-2 qui définit une zone de confort visuel dans laquelle les couleurs sont correctement perçues par un cercle de $\pm 15^\circ$ autour du point de fixation. Dans les centres de contrôle opérationnels, la position de contrôle utilise des écrans radars de 64 cm (30") avec les contrôleurs placés à une distance d'environ 70 cm (Figure 27 : position du contrôleur). Cette zone de confort visuel représente sur l'écran radar un cercle de rayon 18,7 cm.

La nouvelle tâche de validation de notification consistait à demander à l'opérateur de cliquer sur un avion notifié lors de l'apparition d'une notification sur un avion statique à plus de 18,7 cm de l'avion guidé. Le choix de l'avion notifié est réalisé par un tirage aléatoire sur la liste d'avions hors zone de confort visuel. Pour déclencher l'apparition d'une alarme sur l'un de ces avions, un nouveau type de consigne (alarme) a été créé dans l'éditeur LabyEdit. La pénétration de l'avion guidé dans cette zone d'alarme provoque le tirage aléatoire d'un avion satisfaisant le critère de distance minimale par rapport à l'avion guidé.

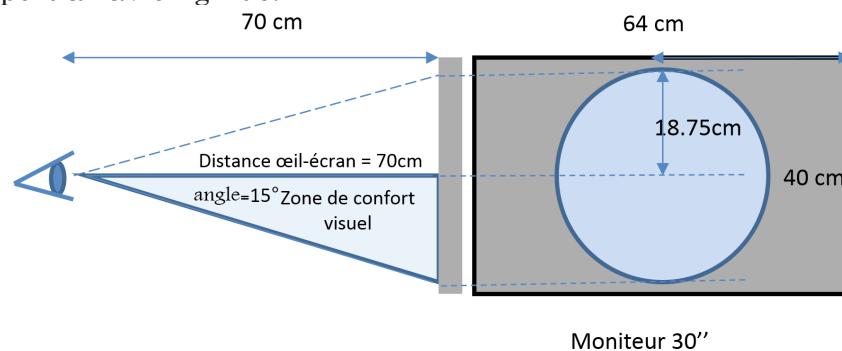


Figure 27 : position du contrôleur

Laby a été modifié de façon à pouvoir implémenter et évaluer plusieurs types de design. Des mesures comportementales ont été ajoutées à Laby de façon à pouvoir évaluer le temps mis par le sujet à valider une notification, le type de notification associé ainsi que la distance entre l'avion guidé et l'avion statique notifié. Pour faire disparaître une notification, le sujet doit cliquer sur l'étiquette de l'avion notifié. Lorsqu'il le fait, un événement de validation de notification est enregistré dans le log de Laby. Le type de design est tiré aléatoirement, à chaque passation l'ordre des designs présentés ainsi que l'avion sur lequel ils s'appliquent et par voie de conséquence, la distance à laquelle ils apparaissent, est différente.

5.2.2. Conception du labyrinthe de l'expérimentation Saillance

L'objet de l'expérimentation Saillance est de mesurer le pouvoir attentionnel de cinq designs de notifications. Parmi eux trois types de designs sont des prototypes conçus pour cette expérimentation qui ont pour but d'être efficaces en vision périphérique. L'un d'entre eux a été créé dans le cadre de la nouvelle visualisation radar du projet SESAR WP4.7.2 ; les deux autres, de par leurs propriétés, ont pour objectifs de capturer l'attention en cas de cécité inattentionnelle (voir chapitre 2.3.1). Deux autres designs sont utilisés

aujourd'hui dans la visualisation radar opérationnelle en France. Il s'agit des designs de warnings et d'alertes décrits au chapitre 1.1.2. Des analyses d'incidents ont montré que dans certains cas elles n'avaient pas été vues par les contrôleurs. Nous souhaitons donc les évaluer avec des mesures de performance quantitative et les comparer aux designs prototypes que nous avons maquetés.

Le labyrinthe saillance (Figure 29 : le labyrinthe Saillance) a été conçu pour permettre d'afficher 20 avions disposés de façon à ne pas se superposer au tracé de l'avion guidé. Le sujet n'a pas à piloter ces avions ; ils restent immobiles durant la passation. La figure suivante (Figure 28 : édition du labyrinthe Saillance) montre le labyrinthe qui a été conçu pour l'expérimentation Saillance avec LabyEdit. Il s'agit d'un parcours unique avec un seul avion guidé par les sujets de façon à focaliser leur attention sur un endroit de l'écran. 40 notifications étaient présentées pendant une passation (8 par design, 5 designs étaient évalués). 83 consignes de niveau et 6 consignes de virages étaient également présentées durant une passation. Lorsque l'avion guidé arrive à un virage, une consigne demandant la réalisation d'une directe sur un point apparaît, son exécution par le sujet amène l'avion dans la direction du segment suivant.

Une zone de consigne ou de notification est l'une des boîtes rectangulaires présentées dans la figure suivante (Figure 28 : édition du labyrinthe Saillance). Le temps d'affichage d'une consigne correspond au temps passé par l'avion guidé dans cette boîte (il est donc fonction de sa vitesse et de la trajectoire suivie). Le sujet dispose de ce temps pour entrer correctement la consigne qui ne disparaît que si la saisie est correcte.

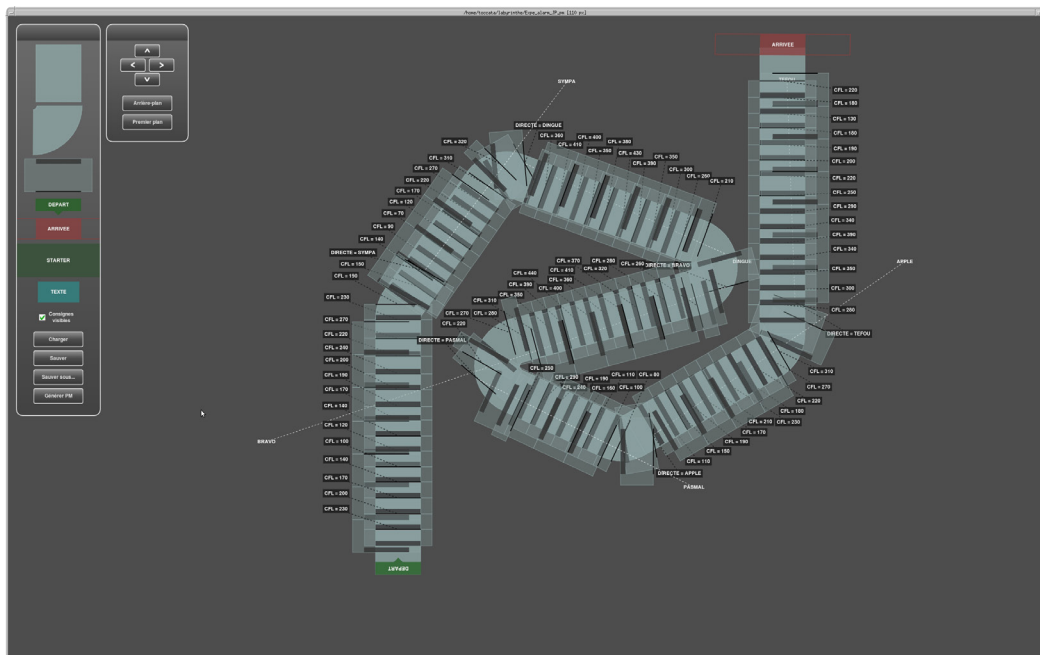


Figure 28 : édition du labyrinthe Saillance

Si la valeur saisie est incorrecte ou si elle n'est pas réalisée à temps, ces informations sont enregistrées et le score (affiché en haut à gauche) est impacté par cette mauvaise réponse. Lorsqu'il est nécessaire d'entrer une consigne, un message texte avec la consigne, ainsi qu'un trait noir matérialisant le délai maximum pour l'entrer apparaissent. Pour pouvoir entrer cette consigne il est nécessaire de cliquer sur le champ de l'étiquette radar

incriminé pour faire apparaître un menu et entrer la valeur correcte (Figure 30 : menus de saisie).

Le labyrinthe a été conçu de telle façon qu'il y avait systématiquement 2 consignes à renseigner pendant la durée théorique d'une notification. La durée moyenne des zones de notifications est de 19.62s ; la durée moyenne d'une zone de consigne est de 9.14s. Une passation durait environ une vingtaine de minutes.

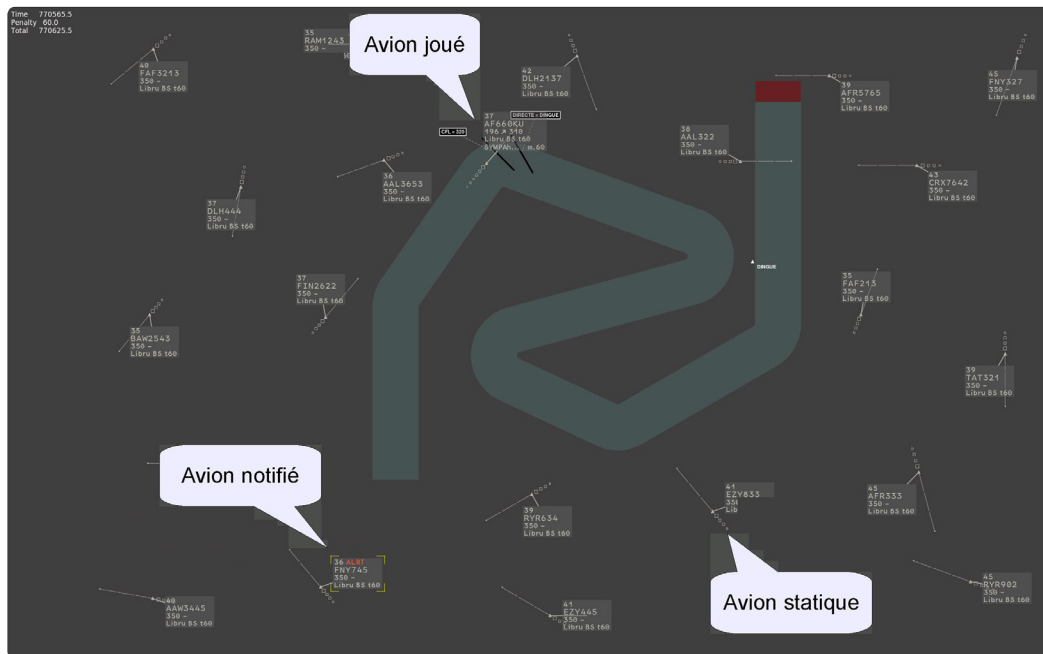


Figure 29 : le labyrinthe Saillance

Le design des étiquettes radars, des menus, les couleurs utilisées reprennent celui utilisé dans la visualisation radar opérationnelle.

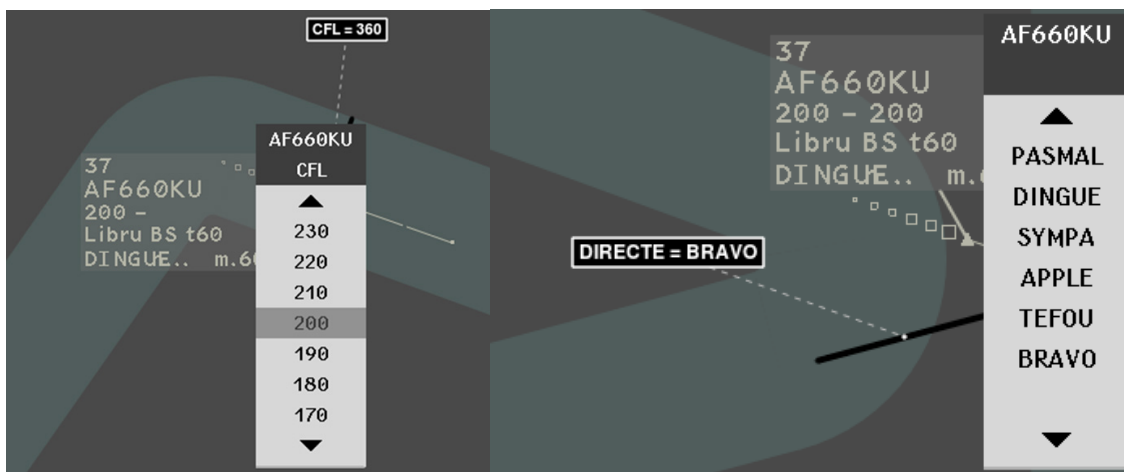


Figure 30 : menus de saisie niveau et directe

Nous aurions pu créer des consignes plus complexes qui auraient mis en œuvre des compétences spécifiques aux contrôleurs, par exemple en ne donnant pas d'indication pour maintenir l'avion guidé dans le labyrinthe ou en provoquant l'apparition d'avions évolutifs ou des obstacles à faire éviter. Ce choix n'a pas été fait, car l'objectif de l'expérimentation était d'évaluer des designs de notification en vision périphérique, ça n'est pas une compétence

spécifique de contrôle et en guidant les sujets sur ces zones de virage nous pouvions nous permettre de prendre des sujets non experts.

5.2.3. Types de notifications évaluées

Cinq designs ont été évalués lors de l'expérimentation Saillance :

1) **Color** est une notification statique de couleur (aucune animation) qui affiche le texte «FNIV» (mauvais niveau) dans une couleur rouge-orangé en ligne 0 de l'étiquette radar. Elle est utilisée dans les visualisations opérationnelles et correspond à un avertissement sur l'altitude de l'avion qui se trouve à un niveau de vol incorrect compte tenu de son niveau d'équipement (Figure 31 a).

2) **ColorBlink** est une notification de couleur animée, le texte «ALRT» est affiché en ligne 0 de l'étiquette radar et clignote en passant de blanc à rouge orangé. Le blanc est affiché 200 ms, le rouge-orangé 800 ms. Il est actuellement utilisé dans les visualisations radar opérationnelles pour les alertes filets de sauvegarde (Figure 31 b).



Figure 31 : notifications: (a) Color (b) ColorBlink (c) BoxAnimation (d) ShadowMask

3) **BoxAnimation** est une notification utilisant le mouvement (Figure 31 c). Le texte ALRT en rouge orangé est affiché en ligne 0 de l'étiquette radar ; quatre chevrons jaunes placés autour de l'étiquette de l'avion notifié se déplacent vers l'extérieur pour revenir vers le cadre de l'étiquette avec un cycle de l'animation de 1 Hz suivant un profil d'animation slow in/slow out. La vitesse d'animation change au cours du temps, en début et en fin de cycle celle-ci diminue ; l'effet obtenu est une sorte de pulsation. Ce design a été prototypé pour une nouvelle image radar utilisée dans le cadre du work package 4.7.2 du projet SESAR. L'animation s'arrête lorsque le sujet la valide en cliquant sur l'avion notifié.

Le design rationale derrière BoxAnimation était d'attirer l'attention en produisant un mouvement autour de l'étiquette, les chevrons en jaune (couleur la mieux perçue en périphérie) matérialisent une étiquette qui changerait de taille.

4) **ShadowMask** est une notification utilisant le mouvement et la couleur. Le texte ALRT en rouge orangé est affiché en ligne 0 de l'étiquette radar. Deux animations consécutives sont réalisées. La première utilise l'opacité, toute l'image sauf l'avion notifié est opacifiée en utilisant le canal Alpha (transparence), l'animation dure 300ms. L'avion notifié se retrouve beaucoup plus lumineux que tout le reste de l'image. À la fin de l'animation de transparence, l'avion notifié vibre pour attirer l'attention du sujet en vision périphérique, la vibration dure 2.5. La totalité de l'avion bouge (étiquette et comète radar) en se décalant de 60 pixels de la position d'origine (figure 31 d). L'écran radar reste sombre pendant 20s ou jusqu'à ce que le participant valide la notification en cliquant sur l'avion notifié.

La conception de ShadowMask est inspirée par le concept de contre-mesures cognitives mis en œuvre par (Dehais, Causse & Tremblay, 2011) où des informations étaient temporairement retirées afin d'attirer l'attention sur un aspect critique de la tâche et éviter la persévérance sur des éléments moins importants. Avec ce design, nous avons choisi de dégrader la visualisation de l'information non pertinente au lieu de l'enlever complètement, les vibrations ont été ajoutées de façon à capturer l'attention en vision périphérique.

5) **Halo** est une notification utilisant le mouvement et la couleur. Le texte ALRT en rouge orangé est affiché en ligne 0 de l'étiquette radar. Un cercle jaune centré sur l'avion notifié et dont le rayon correspond à la position de l'avion guidé dans le labyrinthe est créé, une animation fait diminuer le rayon de ce cercle pour l'amener sur l'avion notifié, à la fin de ce cycle le cercle est recrée en partant de la position de l'avion guidé.

La conception de Halo s'inspire de celui de Baudisch (2003) au départ conçu pour notifier une information sur des écrans de très petite taille, en montrant un arc de cercle dans la direction d'une information notifiée. Un cercle, par l'ouverture de son arc, donne une indication de la distance à l'information notifiée (plus il est ouvert plus celle-ci est éloignée) sa direction indique où chercher. Il est conçu pour attirer l'attention des participants à l'endroit où ils réalisent la tâche principale vers l'avion notifié (Figure 32 : notification Halo).

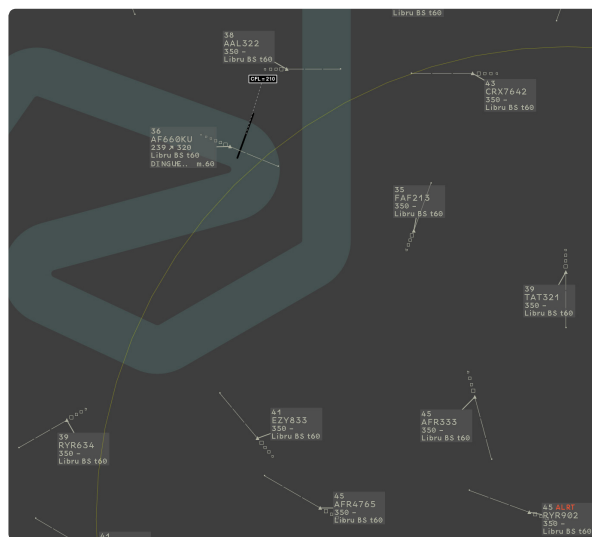


Figure 32 : notification Halo

5.2.4. Architecture de la plateforme Laby, intégration d'un oculomètre

Laby a été développé en langage Perl/Tk et utilise plusieurs bibliothèques développées en interne au département R&D dont le widget Zinc qui est un canvas évolué permettant d'afficher des items graphiques dédiés à la création d'images radar. Parmi ces bibliothèques, on trouve une couche de communication Ivy (Buisson et al. 2002), un bus logiciel permettant très simplement d'échanger des informations entre des processus à l'aide de messages textes et d'abonnements basés sur des expressions régulières. Le bus Ivy est utilisé dans un grand nombre de plateformes des plus simples aux plus complexes interconnectant très facilement plusieurs dizaines d'applications. Des li-

brairies Ivy existent pour les langages de programmation les plus utilisés, elles sont disponibles sur plusieurs systèmes d'exploitation (<http://www.eei.cena.fr/products/ivy>).

Conformément à notre approche méthodologique holistique, nous avons souhaité intégrer un oculomètre pour compléter les mesures comportementales par des mesures physiologiques. Dans le chapitre 4.2.3.1 nous avons vu les deux méthodes permettant de traiter les données oculométriques ; hors ligne par l'enregistrement de données brutes à l'aide du logiciel fourni par l'oculomètre ou bien en temps réel en se connectant à l'aide d'un SDK permettant de se connecter sur le driver de l'oculomètre.

Laby étant en mesure d'utiliser le bus Ivy pour échanger des informations avec d'autres applications, nous avons fait le choix de développer une passerelle (gateway sur la figure 33) entre le bus Ivy et l'oculomètre et échanger ces informations en temps réel. Ce choix technique a plusieurs avantages :

- Il devient possible d'effectuer dynamiquement la corrélation entre le point de fixation et les objets graphiques mobiles (avions) présents dans le Laby ce qui évite des post-traitements fastidieux pour corrélérer les coordonnées brutes du regard avec les objets affichés
- L'acquisition des données oculométriques dans Laby permet d'avoir un log unifié avec les informations liées à l'exécution des différentes tâches et les fixations sur les avions
- La mise à disposition en temps réel des informations provenant de l'oculomètre rend possible le développement futur d'applications les utilisant dynamiquement

L'architecture présentée dans la figure 33 a été utilisée lors de l'expérimentation Saillance. On y voit le gateway assurant la communication entre l'oculomètre et le Laby au travers du bus Ivy, la collecte des données est centralisée au niveau du Laby et enregistrée sous forme de fichier csv pour faciliter leur traitement sous Excel et Statistica. Les questions de synchronisation entre les deux machines (une pour le Facelab et son logiciel propriétaire sous Windows 7 et une pour le Laby sous Linux debian) ont été résolues en utilisant le protocole (NTP : Network Time Protocole) en se connectant à un serveur de temps universel. Le Laby était présenté sur un écran 30 pouces identique à ceux utilisés dans les salles de contrôle opérationnelles.

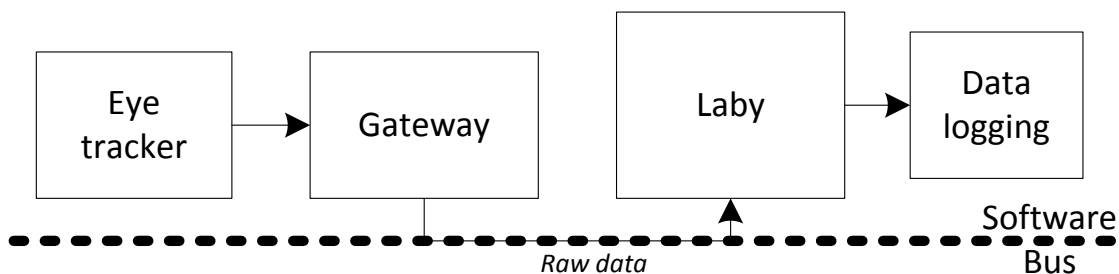


Figure 33 : plateforme expérimentation Saillance

Pour réaliser les mesures oculométriques, nous disposons d'un oculomètre Facelab 5 (Figure 34 : oculomètre Facelab 5). Un SDK était fourni avec ce dispositif, il est possible au niveau du pilote de l'oculomètre de choisir d'enregistrer les données localement ou bien de les envoyer sur une socket TCP puis de les traiter dans une application dédiée à l'aide d'une structure de

données décrite en C++. Le gateway que nous avons développé est relativement simple fonctionnellement, il a pour but de récupérer ces données lorsqu'elles arrivent sur cette socket et les transformer en message à envoyer sur le bus Ivy. Les objets émis par le driver Facelab permettent de récupérer la position du regard en coordonnées pixel, la position de la tête des pupilles et leur diamètre. Si des zones d'intérêt statiques ont été déclarées dans la scène et que le regard s'établit dessus, le champ `objectName` est renseigné avec le nom de l'AOI (Area Of Interest).



Figure 34 : oculomètre Facelab 5

Le message suivant est émis sur Ivy par la passerelle Facelab (`Facelab_netivy`) en fonction de la fréquence de mise à jour demandée sur le PC Facelab (15Hz ou 30Hz). Les valeurs `(.*)` sont remplacées par les valeurs mesurées par l'oculomètre. Un nombre important d'informations sont fournies par `Facelab_netivy` comme les coordonnées pixel du regard, la position de la tête, les clignements de paupières ou le diamètre et la position de la pupille :

```
EyePixelCoordinate xPercent=(.*) yPercent=(.*) xPixel=(.*) yPixel=(.*)
gazeCalibrated=(.*) gazeSaccade=(.*) gazeQualityLevelLeft=(.*)
gazeQualityLevelRight=(.*) headPositionX=(.*) headPositionY=(.*)
headPositionZ=(.*) headRotationX=(.*) headRotationY=(.*) headRotationZ=(.*)
headTrackingState=(.*) headPositionConfidence=(.*) eyeClosureRight=(.*)
eyeClosureRightConfidence=(.*) eyeClosureLeft=(.*) eyeClosureLeftConfidence=(.*)
pupilPositionRightX=(.*) pupilPositionRightY=(.*) pupilPositionRightZ=(.*)
pupilPositionLeftX=(.*) pupilPositionLeftY=(.*) pupilPositionLeftZ=(.*)
pupilDiameterRight=(.*) pupilDiameterLeft=(.*) frameNum=(.*) objectName=(.*)
```

Pour que Facelab transmette les coordonnées pixels du regard sur un écran, il est nécessaire de décrire la scène à l'aide du modèleur 3d simplifié du logiciel livré avec le Facelab. L'écran, sa taille, sa position, le nombre de pixels de la dalle, la position du Facelab en coordonnées `x,y,z` et la distance prévu du sujet par rapport au Facelab sont renseignés. Dans notre cas, il s'agit d'un écran 30 pouces de résolution 2560x1600 situé à 70 cm de la tête des sujets (Figure 35 : le dispositif expérimental). Nous n'avons pas défini d'AOI spécifique puisque Laby a été modifié pour analyser en temps réel les coordonnées du regard et les associer aux objets avions présentés.



Figure 35 : le dispositif expérimental Saillance

Pour que les données soient transmises sur le réseau, une phase de calibration préalable doit être exécutée par le sujet l'amenant à regarder 9 points présentés à différents endroits de l'écran (Figure 36 : phase de calibration dans Facelab). À l'issue de cette phase, les données oculométriques peuvent être exportées vers le réseau.

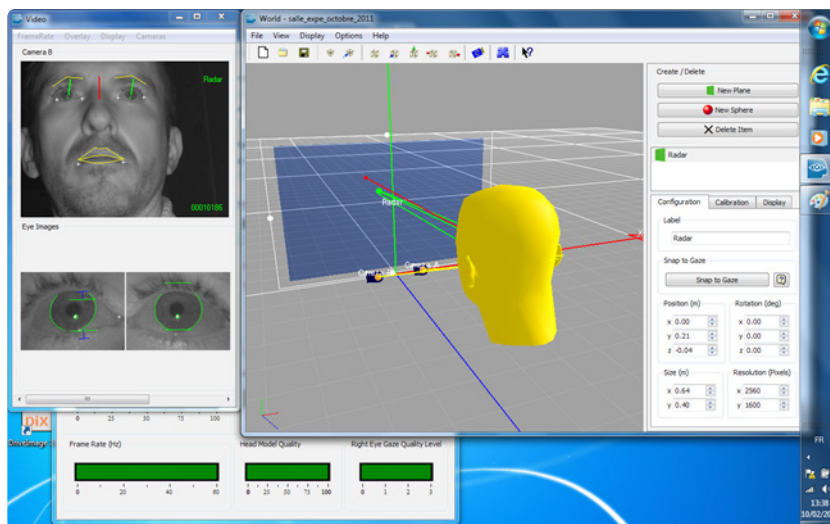


Figure 36 : phase de calibration dans Facelab

Différents objets graphiques sont présentés sur le Laby comme le tracé du parcours (en vert clair sur la figure 35), les aides pour les saisies ou le score. Seules les fixations sur les étiquettes radar sont utilisées pour l'évaluation de la performance de l'opérateur pendant son utilisation, puisque nous souhaitons connaître à quel moment une notification était perçue par l'opérateur.

Les données oculométriques envoyées par Facelab sont beaucoup trop bruitées pour être utilisées directement dans une application évaluant les fixations sur des objets. Pour pouvoir les utiliser, nous avons dû implémenter un algorithme visant à filtrer ce bruit. L'analyse des fixations peut être réalisée avec de nombreux algorithmes (Salvucci et Goldberg 2000). Compte tenu des possibilités offertes par la toolkit TkZinc et que nous souhaitions détecter des fixations sur des objets graphiques connus de TkZinc en temps réel, plutôt que de développer deux modules séparés de détection de fixation puis de détection d'AOI comme cela est généralement le cas dans la littérature nous avons développé un algorithme fusionnant ces deux étapes.

Sur réception d'un message en provenance de l'oculomètre, tous les avions (étiquettes radar) présents dans une boîte carrée de 140 pixels de côté, centrée sur les coordonnées brutes en pixel du regard sur l'écran sont détectés (Figure 37 : algorithme de calcul de fixation). La taille de cette boîte de détection a été réglée après plusieurs essais pour minimiser le temps de calcul tout en ayant un volume d'incertitude suffisant pour gommer le bruit lié à la mesure. La distance en pixels entre le centre de chaque étiquette et la position du regard est calculée pour chacun de ces avions même si les coordonnées du regard ne sont sur aucune des étiquettes. L'avion avec la distance minimale est retenu. Sur l'exemple de la figure suivante, deux avions ont leur étiquette en recouvrement avec la boîte de 140 pixels. Le calcul de distance donne une valeur plus faible pour FNY745 ($d_1 < d_2$), la fixation est donc comptée pour celui-ci. Si plusieurs mesures consécutives sont effectuées sur le même avion (il y a 30 mesures par secondes) un compteur de fixations est incrémenté. Lorsque la boîte de détection ne porte plus sur l'avion précédemment détecté, une fixation sera sortie uniquement si l'avion a été observé au moins 50 ms. Les fixations sont normalement considérées à partir de 150ms, ce seuil de 50ms a été fixé pour permettre l'éventuelle fusion ultérieure de micro fixations en post-traitement des logs d'expérimentation (deux fixations consécutives pour le même avion séparées par un temps faible peuvent être fusionnées a posteriori).

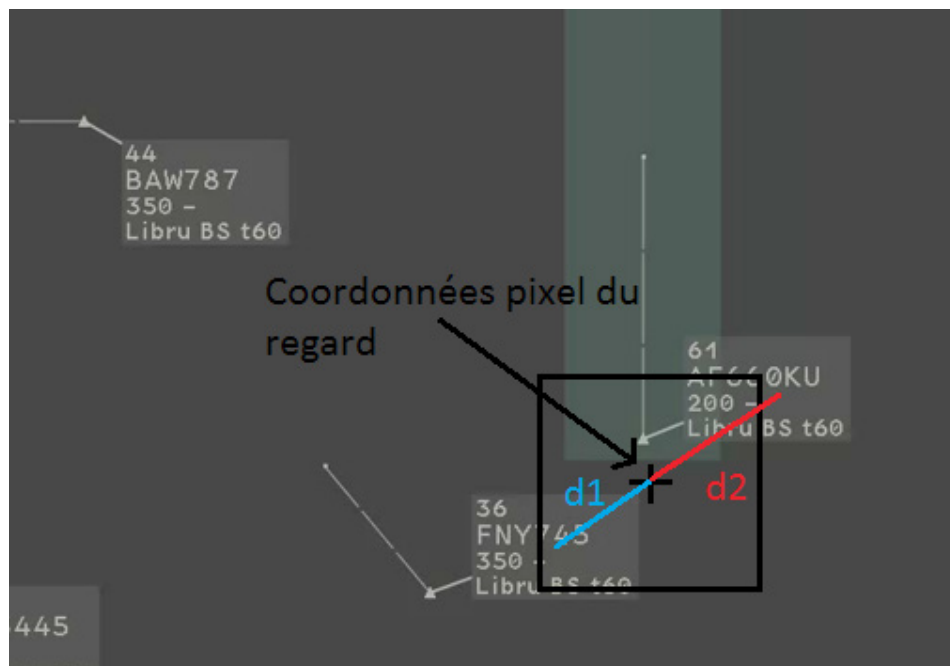


Figure 37 : algorithme de calcul de fixation

5.3. Confrontation de notifications visuelles à des notifications sonores

Au chapitre 4.2.3.2, nous avons décrit une méthode d'analyse ERP basée sur des mesures EEG permettant à l'aide du paradigme Oddball qui consiste à présenter des sons rares et des sons fréquents en contexte de double tâche. Lorsque la charge de travail augmente, un effet significatif sur l'amplitude de la P300 en cas de sons rares peut être mis en évidence.

La théorie de la charge perceptuelle de Lavie (1995) montre que les tâches nécessitant un niveau élevé d'attention laissent peu de capacité de traitement pour les distracteurs visuels ou sonores. Le phénomène de surdité inattentionnelle aux notifications sonores a été mis en évidence dans le cockpit en cas de période de forte charge de travail (Bliss, 2003), mais pas dans le contrôle aérien où ce type de notification est aussi utilisé. Dans l'hypothèse où les ressources attentionnelles sont réparties entre la vision et l'audition (Brand-D'Abrescia & Lavie, 2008), nous avons voulu en réutilisant l'expérimentation Saillance, confronter des designs de notifications visuelles à des notifications sonores dans différentes conditions et vérifier si le type de design de notification a un effet sur la perception des sons rares par l'analyse de l'amplitude de la P300.

Parmi les conditions évaluées, la première est la vitesse d'exécution de la tâche de guidage de l'avion dans le labyrinthe qui entraîne mécaniquement une augmentation de la charge de travail. La seconde est le nombre d'avions statiques affichés et susceptibles d'être notifiés en vision périphérique, l'objet étant d'augmenter la charge visuelle. Les résultats de l'expérimentation Saillance ont montré que seuls deux groupes de notifications présentaient des différences significatives avec ou sans mouvement. Nous avons donc réalisé cette expérimentation en n'utilisant qu'un design de chaque groupe, ColorBlink et BoxAnimation. ColorBlink, car il s'agit du design utilisé pour les alertes filets de sauvegarde dans l'image opérationnelle, BoxAnimation, car c'est celui offrant le meilleur compromis pouvoir attentionnel/impact sur la tâche de guidage.



Figure 38 : lanceur de la tâche laby expé Bip

Laby a été modifié pour offrir le choix sur deux types de vitesse, en vitesse lente la position de l'avion guidé change toutes les quatre secondes, en vitesse rapide toutes les deux secondes (Figure 38 : lanceur de la tâche laby expé Bip). Une nouvelle option au lancement du Laby permet aussi d'agir

sur le critère de charge visuelle (Figure 39 : laby en condition low ou high visual load) en modifiant le nombre d'avions statiques (donc notifiables). Si la charge visuelle est faible, 5 avions statiques sont affichés. Si elle est élevée il y a 22 avions qui peuvent recevoir une notification. Dans les deux cas et comme dans l'expérimentation Saillance, l'avion notifié est choisi aléatoirement en dehors de la zone de vision de confort centrée sur l'avion guidé dans le labyrinthe.

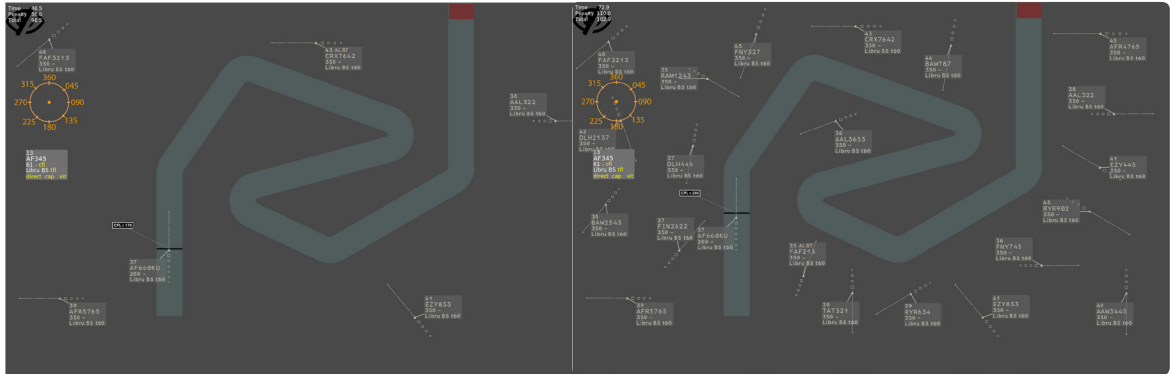


Figure 39 : laby en condition low ou high visual

La tâche auditive ajoutée à la tâche de guidage et de validation de notification utilisée dans Saillance consiste à appuyer sur un bouton poussoir lorsqu'un son d'alarme était émis (paradigme Oddball cf. chapitre 4.2.3.2.2) et à ne rien faire quand il s'agissait d'un son standard.

Cette tâche comportait 25 sons d'alarme pour 100 sons standards présentés de façon aléatoire (tirage sans remise) avec une période aléatoire dont la valeur moyenne est centrée sur 4.26s pour le facteur vitesse lente et 2.6s pour la vitesse rapide. En cas de sons d'alarme, les sujets devaient les valider en appuyant sur un bouton poussoir. Cette tâche a été ajoutée aux options de lancement de Laby.

Pour estimer l'impact du design des notifications visuelles sur la charge de travail dans les différentes conditions ; des mesures de type potentiel évoqué à l'aide d'un EEG ont été réalisées ; l'amplitude de la P300 est un indicateur de charge de travail (voir chapitre 4.2.3.2.2). Par ailleurs, des mesures de performance des différentes tâches ont été réalisées par Laby portant sur les mêmes critères que l'expérimentation Saillance.

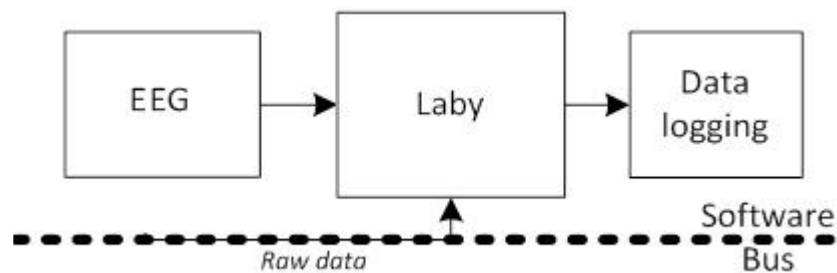


Figure 40: architecture expérimentation Bip

Pour réaliser la synchronisation de Laby avec l'EEG, un pulse était envoyé vers l'EEG au moyen d'une liaison port parallèle au travers d'une interface USB. Lorsqu'un son était émis par Laby pour la tâche auditive, un front montant était envoyé sur l'une des broches du port.

Les données EEG ne sont pas intégrées directement dans les logs de Laby. L'architecture de la plateforme (Figure 40: architecture expérimentation Bip) est donc plus simple que lors de l'expérimentation Saillance puisqu'aucun autre agent que le lanceur de la plateforme n'est présent sur Ivy en dehors de Laby.

5.4. Réalisation d'un agent de suivi de la tâche de supervision des contrôleurs aériens : SAmonitoring

Le contrôleur doit, en effectuant un circuit visuel permanent sur sa visualisation radar, établir une image mentale du trafic aérien sous sa responsabilité. L'analyse de ces informations lui permet de projeter son état futur pour détecter les éventuels conflits et pertes de séparation dans son secteur de contrôle. Il s'agit donc de suivre tous les éléments de chaque avion disponibles sur l'étiquette radar décrite au chapitre 1.1.2 (cap, altitude, vitesse, route, plan de vol déposé). Par exemple, le contrôleur va s'assurer que s'il a donné un ordre à un avion pour effectuer une montée vers un niveau, celui-ci va bien la réaliser et se stabiliser au niveau demandé ou qu'en donnant une directe sur un point de sa route il va bien prendre le cap adéquat.

Les positions de contrôle aérien aujourd'hui en service utilisent des strips papiers. Lorsque les contrôleurs donnent des ordres aux avions, ils notent sur ces strips les valeurs des consignes. Ces informations ne sont pas connues du système de traitement des données qui alimentent les positions de supervision. L'évolution des supervisions vers le tout électronique (plus de strips papier) oblige le contrôleur à renseigner sa supervision lorsque :

- il donne des ordres aux avions
- il prend un avion en fréquence ou le transfère à un autre secteur

Pour le deuxième volet de ces travaux dont le but est de concevoir et mettre en œuvre une nouvelle visualisation dont le design vise à améliorer la conscience de la situation, nous nous sommes appuyés sur les données disponibles dans ces nouvelles supervisions.

Grâce à ces données, il est désormais possible de s'assurer pendant l'activité du contrôleur que ce qui est en train de se passer est cohérent avec :

- ce que font les avions (leur trajectoire), basé sur les données radar
- ce que le contrôleur a renseigné dans sa supervision radar (les ordres donnés aux avions), les informations connues de l'IHM
- ce que le contrôleur doit faire conformément au fonctionnement de son secteur de contrôle (par exemple prendre les avions en fréquence, les délivrer au secteur suivant, les amener au niveau de vol de sortie du secteur), les règles associées au secteur de contrôle et les informations issues des plans de vols des avions

Endlsey & Jones (2012) dans leur ouvrage 'Designing for situation awareness : A user centered design', ont proposé une méthode de design orientée conscience de la situation. Cette méthode de conception est basée sur la mise en évidence des informations dynamiques nécessaires à la réalisation de la tâche à l'aide d'une analyse GDTA (voir chapitre 2.2.4) et le suivi de plusieurs règles de design pour concevoir une supervision qui améliorerait sa conscience de la situation. Nous nous sommes inspirés de cette méthode, pour réaliser un agent de monitoring de la tâche du contrôleur.

Pour réaliser une analyse préliminaire de la tâche des contrôleurs, nous avons fait appel à deux SME (Subject Matter Experts), des experts contrôleurs (plus de 15 ans d'activité opérationnelle) affectés dans les équipes de recherche qui ont été impliqués dès le début de ce projet. Une analyse GDTA a déjà été réalisée pour les contrôleurs en-route aux États-Unis (Endsley,

Rodgers, 1994). Elle fait apparaître un nombre très important d'objectifs et de sous objectifs associés à la tâche des contrôleurs. Compte tenu de cette complexité, il est apparu avec les SME qu'un agent relativement simple au niveau fonctionnel et limité quant aux nombres d'objectifs qu'il serait capable de suivre en parallèle du contrôleur rendrait déjà un nombre important de services dont aucun système ne dispose à l'heure actuelle. Pour cette raison, nous n'avons donc pas réalisé une nouvelle analyse GDTA complète de l'activité du contrôleur aérien.

Les SME ont permis de dégager un certain nombre d'objectifs de leur tâche pour lesquels ils doivent actuellement se forger une bonne conscience de la situation en surveillant dynamiquement les informations présentées sur l'image radar. De ces objectifs, nous avons déduit les informations dynamiques les plus intéressantes que les contrôleurs doivent détecter pendant leur tâche afin de réaliser un agent de suivi de ces informations. Celles-ci présentent des caractéristiques de criticité différentes pour le contrôleur. Nous avons éliminé de cette analyse la détection des informations concernant les interactions entre avions pouvant aboutir à des conflits qui sont déjà traités dans d'autres systèmes d'aide (STCA et MTCD voir 1.1.2) pour nous focaliser sur la surveillance des paramètres individuels associés à chaque avion. Tous les autres sous-objectifs de la tâche individuels à un avion, comme l'optimisation de la trajectoire pour limiter le carburant ou la prise en compte des demandes compagnie de niveau ou de vitesse de croisière, n'ont pas été pris en compte dans cet agent, d'une part pour le simplifier et d'autre part parce qu'il s'agit de sous-objectifs secondaires dans la réalisation de la tâche.

L'objectif de cet agent est de contrôler divers paramètres concernant la tâche du contrôleur aérien et de notifier la supervision radar des éventuels écarts par rapport à des seuils réglés avec les SME. Grâce à un niveau de criticité associé à l'information détectée déterminée par les SME, il sera possible de choisir le design adapté à la représentation de l'information sur la visualisation pour optimiser le niveau de saillance en regard des objectifs de la tâche.

Les SME ont défini que trois niveaux de notification étaient suffisants par rapport à l'activité : mineur, majeur, critique. À titre d'exemple, un filet de sauvegarde a un niveau critique dans l'activité. Le niveau mineur correspondrait à la notification d'une information pertinente, mais non urgente pour le contrôleur, à lui de gérer le moment où il va prendre en compte cette notification en fonction des tâches en cours. Le niveau majeur est une notification importante par rapport à l'activité entraînant une réponse obligatoire de la part du contrôleur, mais sans contrainte temporelle forte. Le niveau critique est une notification avec réponse immédiate obligatoire de la part du contrôleur, si dans les deux premiers niveaux le contrôleur ne doit pas nécessairement être interrompu pendant sa tâche, dans le cas du niveau critique, tout doit être fait au niveau de la supervision pour que cette notification soit perçue y compris si l'attention est focalisée sur une zone éloignée de celle-ci.

De l'analyse de tâche basée sur la conscience de la situation, c'est-à-dire basée sur le suivi des éléments dynamiques de la supervision radar, a émergé l'intérêt de surveiller avec l'agent de monitoring les éléments suivants :

- le suivi des clairances (consignes) de niveau de vol basé sur les données radar et IHM

- le suivi des clairances de vitesse basé sur les données radar et IHM
- le suivi des clairances de cap ou de virage basé sur les données radar et IHM
- le suivi des clairances de directe sur un point de la route basé sur les données radar, plan de vol et IHM
- la conformité du niveau de vol courant avec le niveau de vol attendu en entrée du secteur (niveau d'entrée) basé sur les données plan de vol et IHM
- la conformité du niveau courant de l'avion avec le niveau de vol de transfert en sortie du secteur (niveau de sortie) basé sur les données plan de vol et IHM

le suivi des états de coordination et des états fréquences (voir l'explication détaillée paragraphe 5.4.5) basé sur les coordonnées du secteur de contrôle et les données IHM. Comme dans le cas de Laby, cet agent a été développé de façon à s'intégrer à une plateforme existante basée sur le bus logiciel Ivy. Sur ce bus l'agent va récupérer les informations en provenance du simulateur de trafic (les données radar et plans de vol) ainsi que toutes les informations nécessaires en provenance de la supervision radar (consignes renseignées par le contrôleur).

L'architecture de cet agent est simple. Elle consiste à acquérir les informations pertinentes sur le bus Ivy et les comparer à des valeurs attendues ou des seuils définis avec les SME. Tous les paramètres sont surveillés en parallèle et des messages sont délivrés en sortie de l'agent SAmonitoring vers le bus Ivy (Figure 41 : architecture plateforme). En sus des éléments d'information décrits précédemment, une source d'information sur l'opérateur basée sur des données physiologiques est prévue pour réaliser de l'automatisation adaptative. Dans le cadre de ces travaux, elle n'a été utilisée que pour réaliser des mesures destinées à l'évaluation de cette nouvelle position de contrôle aérien.

Pour émettre les messages de notification en sortie de SAmonitoring sur Ivy, le format suivant a été choisi :

```
MonitoringEvent Flight=(\d) Level=(Warning|Alert|Critical)
Type=(State|Heading|Speed|Cfl|Efl|Tfl) Meaning=(\s)
```

Sur détection d'une anomalie, un message identifiant le numéro de vol concerné est émis avec le niveau de criticité associé (warning, alerte ou alerte critique), le type d'anomalie détecté (état de coordination, niveau de vol ou autres) ainsi que la raison de déclenchement de la notification (par exemple vitesse verticale trop faible). Dans les paragraphes suivants, nous présentons en détail chacun de ces éléments et le type d'information fourni par l'agent de monitoring.

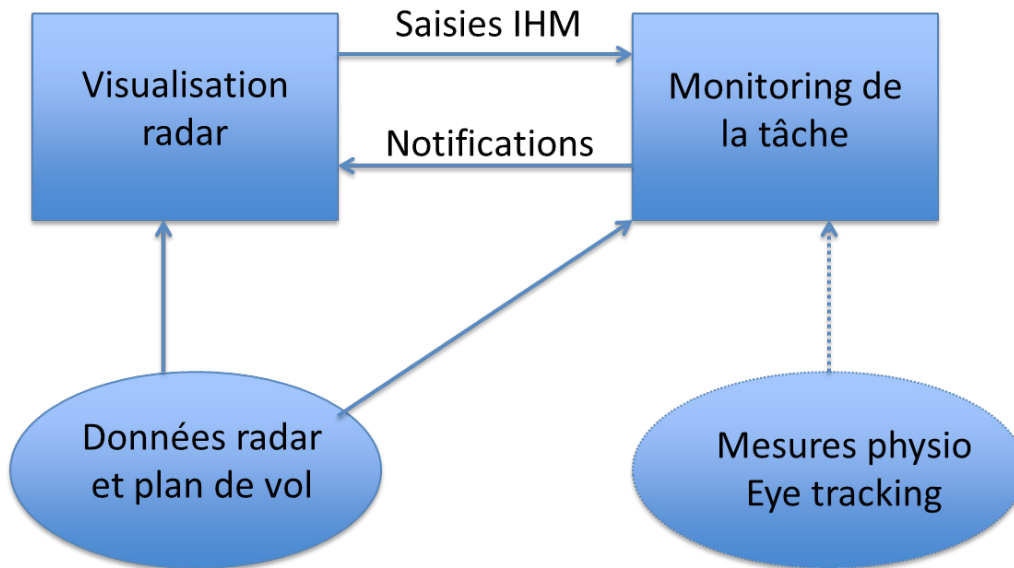


Figure 41 : architecture plateforme

5.4.1. Suivi des clairances de niveau de vol

À aucun moment, dans un vol contrôlé, les pilotes ne décident du niveau courant de l'avion, car celui-ci est forcément donné par un contrôleur. Dans le système opérationnel, après avoir donné une consigne à un avion, les contrôleurs se contentent d'écrire le niveau de vol assigné sur un strip papier, dans la version tout électronique en cours de déploiement, celui-ci doit renseigner sur la supervision radar le niveau de vol qu'il assigne à un avion, cette information est propagée dans le système et le contrôleur du secteur suivant en dispose également. Lorsqu'un avion a un niveau de vol assigné, il n'est pas censé en changer sans autorisation, lorsqu'on lui donne l'ordre de changer de niveau, il est supposé le faire dans un délai raisonnable et s'arrêter au niveau autorisé. À l'aide des SME nous avons déterminé 7 cas pour lesquels l'avion ne fait pas ce qui est attendu du contrôleur. (Figure 42 : synoptique du monitoring de niveau de vol CFL) Pour chacun de ces cas, un niveau de criticité associé ou un seuil a été déterminé. Il est apparu que seuls deux niveaux de criticité étaient nécessaires pour le suivi du niveau de vol. En effet l'écart entre deux niveaux de vol et la norme de séparation verticale ne sont que de 300 m, des problèmes de suivi des clearances de niveau peuvent donc avoir très rapidement des conséquences graves. Le suivi de l'altitude d'un avion est donc critique dans l'activité du contrôleur, pour cette raison nous allons détailler spécifiquement cette partie de l'agent SAmonitoring.

Les 7 cas classés par niveaux de criticité déterminés avec les contrôleurs sont les suivants :

Niveau mineur :

- *vitesse verticale trop faible ou trop forte (dans ce cas l'avion n'évolue pas de façon nominale pour le contrôleur, il n'est pas dans le volume attendu d'évolution du vol.*

- *avion toujours stable après un seuil temporel de détection (une clairance a été donnée par le contrôleur, mais le pilote ne l'effectue pas)*

Niveau critique :

- *stabilisation prématurée, l'avion arrête son évolution avant le niveau clearé*
- *dépassement du niveau, l'avion poursuit son évolution après avoir croisé le niveau clearé*
- *l'avion part dans le sens inverse de la manœuvre demandée*
- *changement de niveau sans clearance*

Le système de monitoring délivre une notification associée au CFL (niveau de vol clearé-assigné) avec un niveau de criticité et une raison de déclenchement. Le niveau actuel de l'avion est nommé AFL, le niveau de l'avion au moment de la clairance AFLO. Les seuils ont été définis avec les SME.

Si l'on définit les données suivantes :

T1 : temps maximum pour déclencher la manœuvre

T2 : temps maximum pour finir la manœuvre, à calculer sur la base d'une vitesse verticale 'réaliste' compte tenu de l'altitude de l'avion et de l'écart en niveau

T3 : premier seuil de détection de non-accomplissement de la clairance

T4 : second seuil de détection de non-accomplissement de la clairance

DtSteady : écart entre l'heure courante et l'heure à laquelle l'avion est passé stable

Dt: temps entre le moment de notification de la manœuvre et l'heure courante

Voici les conditions de déclenchement des notifications et les raisons associées :

Si $Dt > T1$ et $AFL == AFLO$

Type de notification : mineure

Raison : avion toujours stable après timeout

Si $AFL != AFLO$ et $AFL - CFL$ varie dans le mauvais sens

Type de notification : critique

Raison : l'avion part dans le sens inverse de la manœuvre

Si $AFL != CFL$ et $DT > T2$

Type de notification : mineure

Raison : avion évolue très lentement en montée ou en descente

Si $\text{abs}(\text{AFL}-\text{CFL})$ augmente et CFL déjà croisé

Type de notification : critique

Raison : overshoot du niveau

Si $\text{abs}(\text{AFL}-\text{CFL})$ stable et $\text{AFL} \neq \text{CFL}$ et $\text{DtSteady} > T3$ et $< T4$

Type de notification : mineure

Raison : avion stabilisé avant la clairance

Si $\text{abs}(\text{AFL}-\text{CFL})$ stable et $\text{AFL} \neq \text{CFL}$ et $\text{DtSteady} > T4$

Type de notification : critique

Raison : avion stabilisé avant la clairance

Si $\text{AFL} \neq \text{CFL}$ et avion était stable au CFL

Type de notification : critique

Raison : avion change de niveau sans clairance

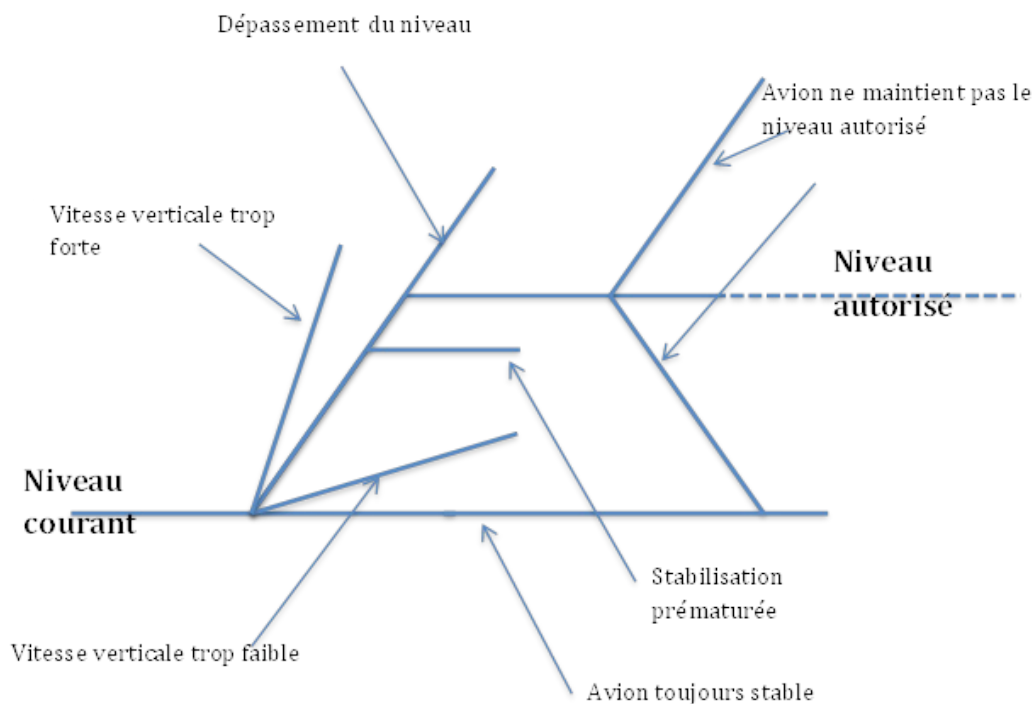


Figure 42 : synoptique du monitoring de niveau de vol (CFL)

5.4.2. Suivi des clairances de cap

Une clairance de cap est un ordre de cap absolu (entre 0 et 360°) ou un ordre de virage par rapport au cap courant de l'avion. En contrôle en route, ce type d'ordre est le plus souvent donné pour éviter une zone ou un conflit avec un autre avion. Les mêmes variables sont utilisées que dans le cas du CFL, le cap initial est nommé CAPO , le cap donné par le contrôleur CAPSELECTE .

T2 : temps maximum pour finir la manœuvre calculée sur la base d'un taux de virage 'réaliste' (environ 3° par secondes)

Si $DT > T1$ et $CAP == CAP0$

Type de notification : mineure

Raison : avion toujours pas en virage après timeout

Si $CAP != CAP0$ et $CAP - CAP0$ varie dans le mauvais sens

Type de notification : critique

Raison : l'avion tourne dans le sens inverse de la manoeuvre

Si $CAP != CAPSELECTE$ et $Dt > T2$

Type de notification : mineure

Raison : l'avion tourne très lentement

Si $abs(CAP - CAPSELECTE)$ augmente et $CAPSELECTE$ déjà passé

Type de notification : majeure

Raison : overshoot du cap sélectionné

5.4.3. Suivi des consignes de vitesse

Au moment de la clairance la vitesse de l'avion est égale à $VEL0$. La vitesse donnée par le contrôleur $VELSELECTE$, VEL la vitesse courante

Si $DT > T1$ et $VEL == VEL0$

Type de notification : mineure

Raison : avion toujours stable en vitesse après timeout

Si $VEL != VEL0$ et $VEL - VEL0$ augmente dans le mauvais sens

Type de notification : critique

Raison : l'avion ralenti au lieu d'accélérer ou réciproquement

Si $VEL != VELSELECTE$ et $dt > T2$

Type de notification : mineure

Raison : avion accélère/décélère très lentement

Si $abs(VEL - VELSELECTE)$ augmente et $VELSELECTE$ déjà croisée

Type de notification : majeure

Raison : vitesse sélectionnée non stabilisée (intégrer une louche de quelques kts)

5.4.4. Suivi des consignes de directe

Les directes sont une consigne demandant à un avion de se diriger vers un point nommé de l'espace aérien (Waypoint). Le plus souvent il s'agit de directes sur un point faisant partie de la route d'un avion. Au moment de la clairance le cap de l'avion est égal à CAP0

T2 : Temps max pour finir la manœuvre, à calculer sur la base d'un taux de virage 'réaliste'

CAPDIRECTE : le cap calculé pour rejoindre la balise donnée en directe.

ECART : écart toléré en degré par rapport à la directe calculé à l'heure de CAP0

Si $DT > T1$ et $CAP == CAP0$

Type de notification : mineure

Raison : avion toujours stable après timeout

Si $DT > T1$ et $CAP == CAP0$

Type de notification : mineure

Raison : avion toujours pas en virage après timeout

Si $CAP \neq CAP0$ et $CAP - CAP0$ varie dans le mauvais sens

Type de notification : critique

Raison : l'avion tourne dans le sens inverse de la manoeuvre

Si $CAP \neq CAPDIRECTE \pm ECART$ et $dt > T2$

Type de notification : mineure

Raison : avion tourne très lentement

Si $abs(CAP - CAPSELECTE)$ augmente et $CAPDIRECTE \pm ECART$ déjà passé

Type de notification : majeure (rang à vérifier, peut être critique)

Raison : dépassement du cap pour la directe selectée

5.4.5. Suivi des états de coordination et des états fréquences

Les états de coordination correspondent à des états système pour la gestion de la vie du vol. Ils sont mis à jour par le système de traitement de plan de vol. Lorsqu'un vol se dirige vers l'entrée d'un secteur, une dizaine de minutes avant l'entrée 'géographique' dans celui-ci, un strip va être imprimé sur la position de contrôle et récupéré par le contrôleur organique pour une analyse des caractéristiques du vol comme l'heure de passage sur les waypoints du secteur de façon à détecter les éventuels conflits pouvant se

produire et gérer d'éventuelles coordinations pour éviter un conflit avec le secteur amont.

MAE (mouvement accepté en entrée) : le vol passe dans cet état 10 min avant l'entrée dans le secteur. Sur les anciennes positions le passage MAE signifiait l'impression du strip papier sur la position, sur les nouvelles positions il déclenche le changement de format et de couleur de l'étiquette radar sur la supervision.

MAS (mouvement accepté en sortie) : le vol est annoncé pour le secteur suivant. Cet état qui se déclenche en général 10 minutes avant la sortie du secteur correspond à l'édition du strip sur la position suivante : le vol devient MAE pour la position suivante.

Les états fréquences correspondent à la prise en charge à la radio du vol par le contrôleur radariste. Lorsqu'un vol va sortir du secteur géographique en charge d'un contrôleur, celui-ci le transfère au secteur suivant en lui donnant la fréquence sur laquelle il doit contacter le prochain contrôleur.

VSEF : vol sans état fréquence : vol en fréquence dans le secteur amont

VATF : vol attendu en fréquence : vol shooté par le secteur amont. Ce vol va donc bientôt appeler la position

VASS : vol assumé en fréquence. Le vol est en contact radio avec la position, cet état est initié par une action d'assume du contrôleur sur la visualisation

VSHO : vol shooté ou transféré vers le prochain secteur (action du contrôleur sur la visualisation), mais il n'a pas encore été assumé par celui-ci (encore sous la responsabilité du secteur donnant)

VTRF : vol transféré : vol en contact avec le secteur suivant et assumé par le contrôleur.

6 cas d'anomalies à propos des états de coordination et de fréquence ont été identifiés par les SME, 3 majeurs et 3 mineurs, aucun cas critique :

Si MAE et VSEF et heure=TES

Type de notification : majeure

Raison : avion dans le secteur géographiquement, mais pas encore shooté par le secteur précédent

Si MAE et VATF et heure=TES

Type de notification : mineure

Raison : avion dans le secteur géographiquement, mais pas encore assumé par le contrôleur

Si MAE et VATF et heure>T1

Type de notification : majeure

Raison : avion non assumé dans le secteur et temps depuis l'entrée secteur supérieure au seuil T1

Si le vol n'est pas MAS et VSHO

Type de notification : majeure

Raison : vol shooté bien que pas annoncé en sortie (shooté beaucoup trop tôt par encore MAE dans le secteur suivant)

Si le vol n'est pas TSS et VASS

Type de notification : mineure

Raison : vol toujours assumé alors qu'il vient de franchir la frontière géographique du secteur (aurait du être shooté)

Si le vol est VSHO et pas encore VTRF après seuil T2

Type de notification : majeure

Raison : vol shooté et toujours pas assumé par le secteur suivant après seuil T2 (le vol a été ignoré par le secteur suivant, action possible à réaliser de la part du contrôleur organique)

5.4.6. Suivi du niveau d'entrée dans le secteur

Le niveau d'entrée (EFL) correspond au niveau de vol auquel le contrôleur du secteur précédent doit délivrer l'avion avant son entrée dans le secteur. Ce niveau de transfert est défini dans les règles associées au secteur de contrôle. Lorsqu'un avion rentre dans le secteur, il peut déjà être à un niveau courant conforme à l'EFL, à minima le contrôleur du secteur amont est présumé l'avoir clairé à l'EFL avant de transférer le vol. S'il ne l'a pas fait, il est censé avoir coordonné avec le secteur suivant l'entrée à un niveau de vol différent et mis à jour le niveau de transfert du vol.

Soit TES : heure entrée secteur

Un seul cas de type majeur a été identifié par les SME :

Si $CFL \neq EFL$ et VATF

Type de notification : majeure

Raison : vol shooté par le secteur précédent sans qu'il ait été clairé au bon niveau d'entrée pour le secteur

5.4.7. Suivi des niveaux de transfert en sortie

Le niveau de transfert (TFL) correspond au niveau de vol auquel le contrôleur est censé délivrer l'avion en sortie de son secteur. Lorsqu'un avion rentre dans le secteur, il peut déjà être à un niveau courant conforme au TFL, si ça n'est pas le cas, nous calculons un Top Of Descent (ou Climb) optimal pour que l'avion soit au niveau de transfert à la limite géographique du secteur. Si le contrôleur donne l'ordre à l'avion de rejoindre le TFL à l'heure de ce TOD il sera précisément au TFL au moment du passage à la frontière géographique du secteur. Il est acceptable pour les contrôleurs de délivrer un avion à un niveau différent du TFL si celui-ci a été clairé au TFL avant

d'être transféré au secteur aval ; pour cette raison la notification émise au moment du TOD est considérée comme mineure par les SME.

TOD : heure de descente/montée. Calculée par rapport à l'heure de sortie secteur, l'écart de niveau et la vitesse verticale moyenne à l'altitude de l'avion.

TSS : heure sortie secteur

3 cas ont été déterminés par les SME, un de chaque niveau de criticité. Pour le dernier qui est de type critique et concerne un avion qui aurait été transféré au secteur aval sans avoir eu de clairance de niveau l'amenant au TFL, il y a eu débat parmi les contrôleurs pour savoir si il devait être classé majeur ou critique. Dans le cadre des expérimentations, il a été décidé de mettre un statut critique.

Si $CFL \neq TFL$ et $heure = TOD$

Type de notification : mineure

Raison : avion pas encore clairé au TFL au seuil TOD

Si $CFL \neq TFL$ et $heure = TSS$ et VASS

Type de notification : majeure

Raison : avion pas encore clairé au TFL alors qu'il passe la frontière secteur

Si $CFL \neq TFL$ et VSHO

Type de notification : critique

Raison : avion shooté alors que pas encore clairé au TFL

5.5. Intégration de SAmonitoring dans un simulateur d'étude

Le SAmonitoring a été intégré à la plateforme d'un simulateur d'étude Droïd mettant en œuvre des interfaces homme-machine innovantes. Ce simulateur a été utilisé dans le cadre du projet SESAR WP4.7.2 (Figure 43 : position de travail Droïd équipée d'oculomètres Tobii). Le but de Droïd est de mettre en œuvre et de tester de nouvelles fonctions pour le contrôle aérien ainsi que de nouveaux moyens d'interaction ou de saisie d'information. Ce prototype de position de travail pour les contrôleurs a été réalisé au sein de l'équipe PII (Prospection et Innovation pour les Interfaces) du centre R&D de la Navigation Aérienne.

Cette position de supervision est composée de deux écrans ; un écran de 30" vertical utilisé pour afficher la supervision radar et un écran horizontal de 21" de type Wacom sur lequel sont affichés des interacteurs permettant de manipuler l'image radar (zoom, excentrement, curseur à l'aide du stylet Wacom).



Figure 43 : position de travail Droid équipée d'oculomètres Tobii

La supervision radar de Droid (Figure 44 : partie haute : image radar) est centrée sur le secteur courant de la position de contrôle (matérialisé en vert clair). En fonction de leur état de coordination et de fréquence, les vols sont représentés avec un format et une couleur d'étiquette différents. En vert et en périphérie du secteur on peut voir les vols MAE avant que les contrôleurs ne les assument, en blanc les vols assumés et en fréquence avec le secteur contrôlé. Les balises (waypoints) et les routes aériennes les reliant sont matérialisées sur le fond de carte. Des listes de vols présentent les avions listés en fonction du secteur amont dont ils proviennent. Une liste de vols assumés présente tous les avions actuellement en fréquence dans le secteur.

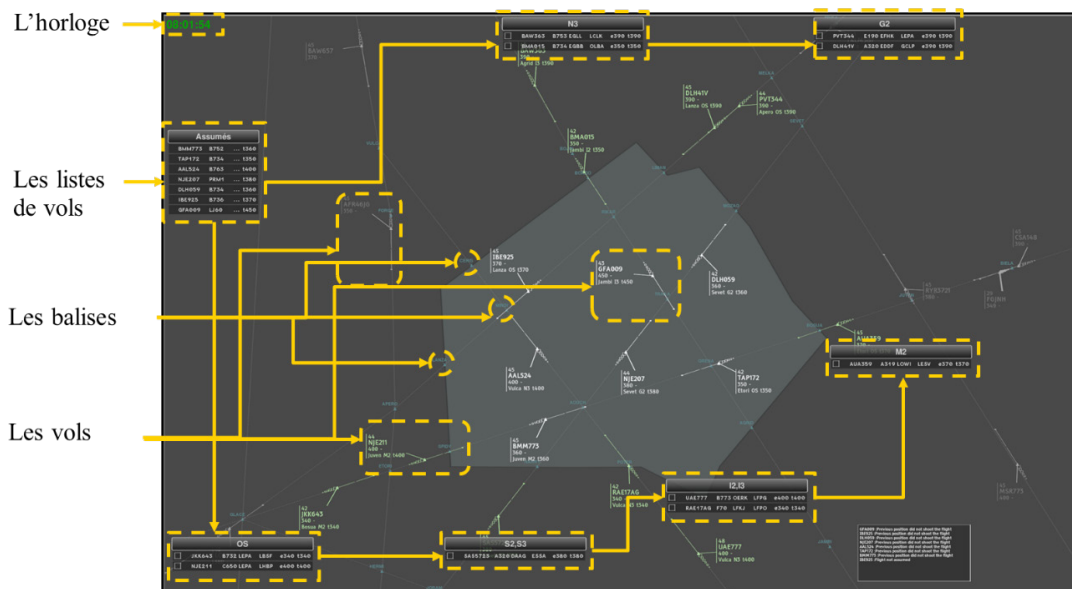


Figure 44 : partie haute : image radar

L'écran Wacom de 21'' en partie basse n'est pas tactile, mais on peut réaliser toutes les opérations normalement faites à la souris sur un écran conventionnel à l'aide d'un stylet (Figure 45 : partie basse écran Wacom). Toutes les interactions normalement réalisées à la souris sur l'écran de supervision radar sont faites à l'aide de l'écran Wacom et son stylet en utilisant une zone sensible réservée de l'écran en indirection (rectangle noir sur l'écran 21''). Un pointage sur un coin de cette zone positionne le curseur de l'image radar au même endroit sur le 30''. On peut comparer cette zone de désignation à un touchpad qui fonctionnerait en absolu au contraire des touchpads d'ordinateurs portables qui fonctionne en relatif (il faut réaliser plusieurs gestes sur leur touchpad pour traverser l'écran). Les menus de saisies associés aux étiquettes radar sont également accessibles de cette façon ; la saisie des clairances est réalisée par reconnaissance d'écriture en ouvrant l'un de ses menus puis en écrivant la valeur de la clairance sur l'écran Wacom. Sur cet écran est affiché un strip électronique (dypinfo) qui montre de façon synthétique les informations de plan de vol pour l'avion sélectionné sur l'image radar. Les autres boutons permettent d'interagir avec la supervision radar en changeant le zoom, l'excentrement et l'affichage des balises.

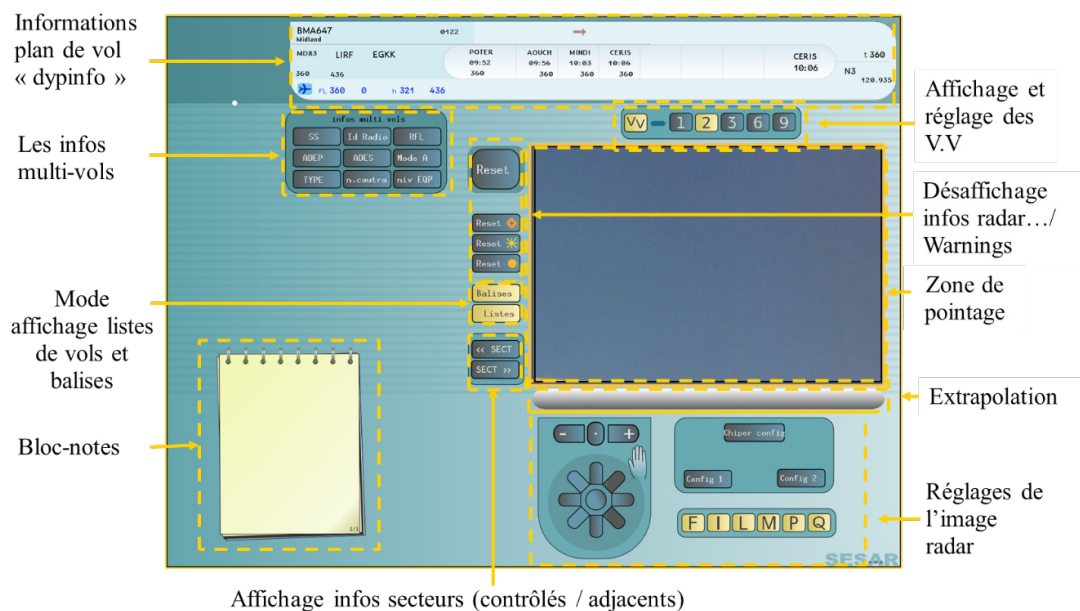


Figure 45 : partie basse écran Wacom

Lors des simulations, le binôme de contrôleurs doit interagir avec un pseudo-pilote qui joue le rôle des avions en fréquence dans le secteur (il est en communication avec le contrôleur radariste uniquement) et avec les contrôleurs des secteurs amont et aval (contrôleur organique uniquement). Ces deux positions de travail (Figure 46 : poste de pseudo-pilote et secteurs adjacents) sont placées dans une salle séparée de la salle de contrôle, les communications sont réalisées à l'aide d'une radio et d'un téléphone VOIP (Voice Over IP). Pour jouer une simulation, quatre personnes sont donc nécessaires en sus des personnels qui supervisent ou observent le déroulement des scénarios.

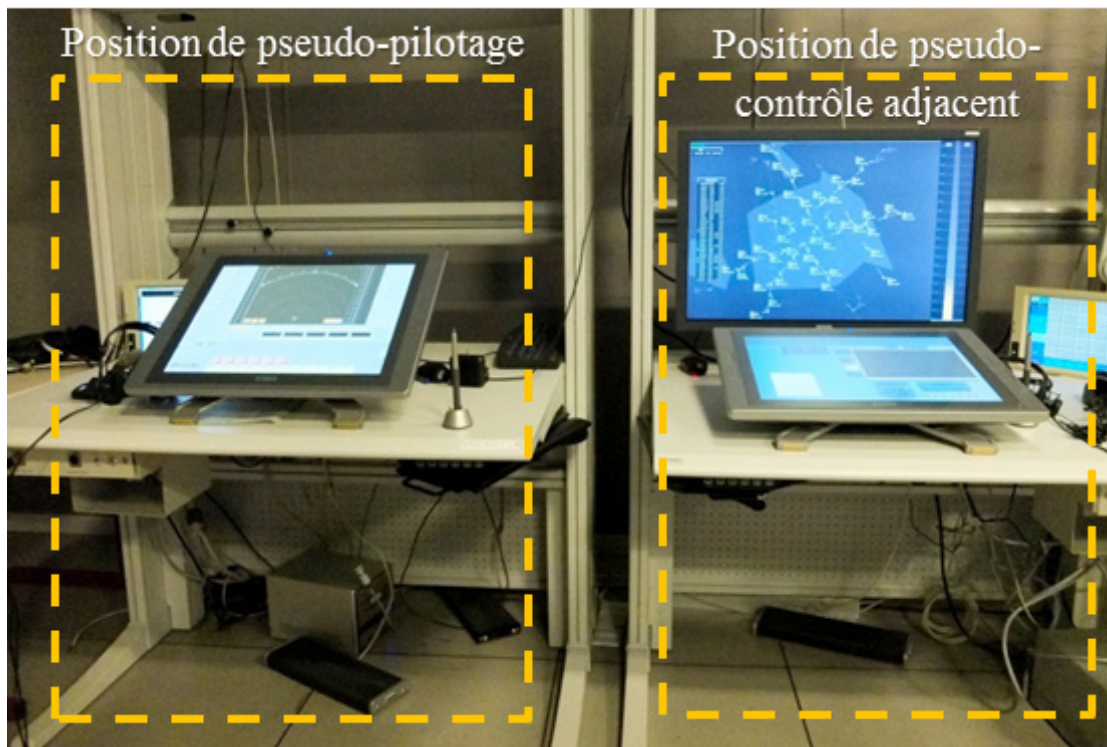


Figure 46 : poste de pseudo-pilote et secteurs adjacents

5.5.1. Architecture Droïd et intégration d'un dispositif oculométrique

Un oculomètre posé Facelab du même type que celui utilisé lors de l'expérimentation Saillance a été mis en oeuvre pendant nos expérimentations. Contrairement à cette dernière, l'expérimentation Monitoring est beaucoup plus complexe en termes d'architecture, elle fait appel à un grand nombre d'agents interconnectés dont un simulateur de trafic aérien. Nous souhaitons utiliser un oculomètre pour détecter lorsque les notifications étaient perçues et contrairement à l'expérimentation Saillance dans laquelle l'un des objectifs était de valider les notifications dès que perçues, dans l'expérimentation Monitoring le contrôleur gère ses objectifs et priorités en fonction de la situation aérienne. Pour savoir si le contrôleur a perçu les notifications il était donc nécessaire d'intégrer un oculomètre à la plate-forme Droïd.

L'expérience acquise sur l'expérimentation Saillance et d'autres projets mettant en oeuvre deux modèles d'oculomètres posés (Tobii et Facelab) nous a permis d'aller plus loin dans le processus d'intégration et de modularisation du traitement des données oculométriques.

Deux défis techniques devaient être adressés dans cette expérimentation. Le premier est l'intégration de l'oculomètre dans cette plateforme de simulation complexe. Le second est la complexité de l'image radar et la difficulté liée à l'intégration des données de suivi du regard à des logiciels complexes pour la détection de zones d'intérêt dynamique.

Pour le premier défi technique, une architecture modulaire beaucoup plus flexible que dans l'expérimentation Saillance a été testée. Dans la mesure où l'algorithme de fixation intégré dans le Laby n'a pas besoin des coordonnées d'objets de l'IHM (ceux-ci sont uniquement nécessaires pour détecter si une fixation correspond à un objet), elle a été séparée dans un module spécifique.

Pour sortir une fixation, un module de ce type a besoin des coordonnées brutes issues du gateway connecté à l'oculomètre, si le point de fixation du regard reste dans une zone géographique (dont la taille est paramétrable et liée à la précision de l'oculomètre) pendant une durée supérieure à un seuil paramétrable (normalement 150ms), alors il délivre une fixation couvrant la zone géographique. Selon la fréquence d'échantillonnage de l'oculomètre, il est possible d'avoir un module de fixation avec un algorithme différent sans que cela impacte les autres modules oculométriques ou agents Ivy présents dans la plateforme. Dans ce type d'architecture, chaque fonction est un agent logiciel dédié afin de permettre une meilleure flexibilité future de l'expérimentation. S'il est nécessaire d'obtenir des fixations sur des objets dynamiques, un autre module peut se charger de collecter les fixations et les coordonnées des zones contenant les objets pour délivrer des fixations cette fois-ci sur des objets.

D'autres modules optionnels pourraient être créés et intégrés dans cette architecture :

- *un module de détection de fixations sur des zones d'intérêt dynamiques,*
- *un module TEPR (Task Evoked Pupillary Response) analyse les données brutes venant du gateway de l'oculomètre et peut produire une sortie pour un classifieur de l'état mental,*
- *un classifieur d'état mental alimenté en données en exemple par le module TEPR, un module d'analyse de la balistique de l'œil ou des clignements de paupières (indicateur de la fatigue) ou de données d'un autre capteur physiologique (c.-à-d. EEG).*

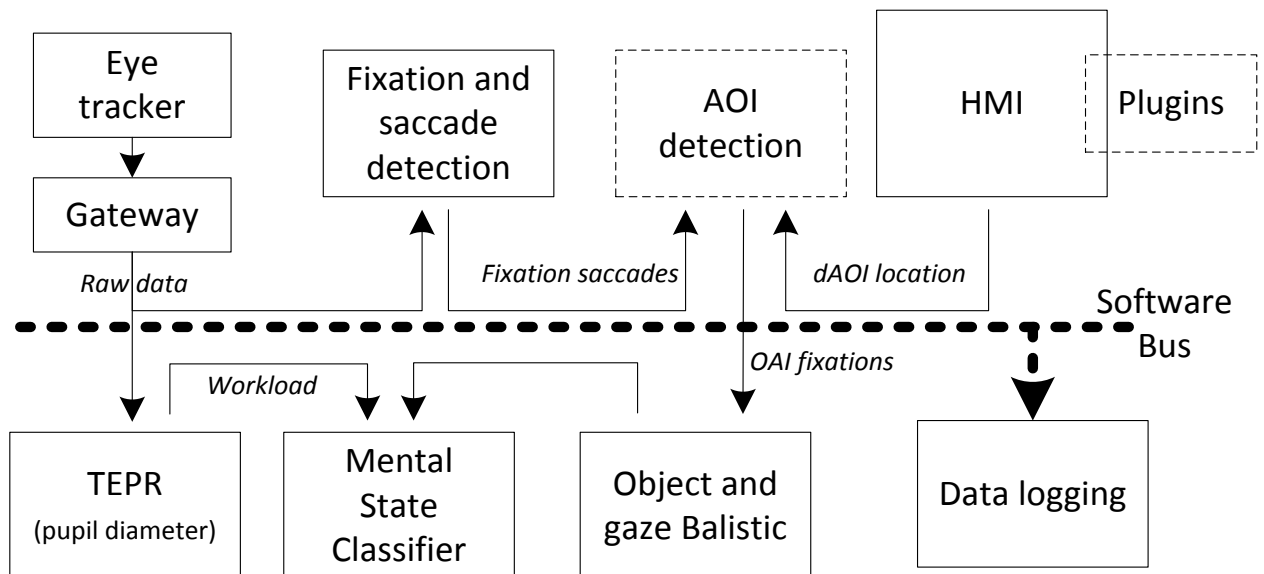


Figure 47: une architecture flexible pour l'intégration de l'oculométrie

Avec cette architecture (Figure 47), il est possible d'ajouter ou de supprimer un module selon les objectifs de l'expérimentation. Par exemple pour enregistrer des données pertinentes afin de réaliser une analyse a posteriori. Nous proposons par exemple un module (non réalisé) qui intégrerait les données de diamètres pupillaires et celle d'un module analysant la balistique de l'œil pour réaliser un classifieur d'états mentaux qui pourrait servir d'input pour l'IHM. Cette architecture simplifie la phase de post-traitement et la

phase d'analyse pour les spécialistes des facteurs humains. Elle permet aussi l'adaptation de l'IHM avec les données issues de modules effectuant des traitements de plus ou moins haut niveaux.

Pour le deuxième défi technique, c'est-à-dire la corrélation entre les objets de la visualisation et les données oculométriques, deux choix étaient possibles. Le premier était d'instrumenter le code de l'image radar de façon à ce que les coordonnées des objets significatifs pour notre expérimentation soient sortis sur le bus Ivy et de les associer dans un module d'AOI (zone en pointillé, Figure 47). La seconde était d'intégrer les fixations dans l'image radar de les corréler aux objets à l'intérieur de l'application (zone pointillée plugins sur la Figure 47). Nous avons choisi la seconde solution parce que notre supervision radar utilise un modèle de conception du type Subject-Observer (Wolfgang, 1994) et une architecture utilisant les plugins. Dans ce type de design, un mécanisme d'attachement et de notification de message permet la communication entre objets graphiques et les données. En utilisant l'architecture de plugin (identique à celle utilisée par exemple dans Chrome ou Firefox) à l'intérieur de l'image radar, il est possible de créer un module dédié à un usage que nous ne voulons pas intégrer complètement dans le logiciel. Cette architecture utilisant les plugins a été introduite dans l'image radar afin de permettre le développement agile de nouvelles fonctionnalités sans modifier le cœur de l'application.

Divers plugins ont déjà été créés pour prendre en charge l'utilisation d'un oculomètre dynamiquement comme un curseur piloté par le regard, un plugin de validation d'alarmes à partir de longues fixations sur un objet notifié ainsi qu'un plugin qui permet de détecter les fixations sur des avions et sortir un message sur le bus Ivy. Des informations pertinentes comme les fixations sur des objets/données graphiques sont envoyées sur le bus logiciel et collectées par un agent dédié. Si l'objectif de l'expérimentation n'est pas d'adapter l'IHM, mais seulement de corréler et enregistrer les données provenant de l'oculomètre et les événements IHM ou les coordonnées des objets, une architecture sans plugins et modules supplémentaires est plus pertinente pour la réutilisation dans d'autres architectures. Pour l'expérimentation SAmonitoring, nous avons utilisé le module de fixation et un plugin inclus dans l'image radar pour la détection des fixations de dAOI et la corrélation avec les avions. L'enregistrement des données a été faite par un module dédié connecté sur Ivy.

Un plugin nommé FacelabIR a donc été développé pour la supervision radar. Il est chargé au lancement de la supervision de s'abonner aux messages en provenance de l'oculomètre et faire la corrélation entre les coordonnées du regard et les objets graphiques présentés sur la supervision radar. L'algorithme utilisé est le même que celui intégré dans Laby, lorsqu'une fixation est détectée sur un avion ; un message est envoyé vers le module d'enregistrement des données d'expérimentation présent sur le bus Ivy.

5.5.2. Intégration des données de SAmonitoring dans la supervision radar

L'expérience saillance a montré que dans les cinq designs testés nous n'avons pu trouver que deux groupes de designs avec des différences significatives en termes de saillance. L'analyse effectuée avec les SME sur l'agent

SAmonitoring a fait émerger trois niveaux de criticité, warning, alerte et alerte critique. Pour les deux premiers il n'y a pas d'exigence de performance concernant le temps mis pour les percevoir, il a donc été décidé en accord avec les experts d'utiliser les deux designs opérationnels Color et ColorBlink pour représenter les niveaux mineurs et majeurs et d'utiliser BoxAnimation qui offrait le meilleur compromis saillance/reprise sur interruption pour les alertes de type critique.

L'étiquette radar est composée de plusieurs champs présentés sur cinq lignes (Figure 48 : détail d'une étiquette). En ligne 0 est montrée la vitesse sol, en ligne 1 l'indicatif de l'avion (callsign), en ligne 2 l'altitude courante de l'avion (AFL) sa tendance verticale (montée/stable/descente) et le niveau de vol clairé (CFL), en ligne 4 la balise de sortie le niveau de transfert (TFL) et le nom du secteur suivant. Sur la dernière ligne sont présentées les éventuelles clairances renseignées pour l'avion (cap/vitesse/directe sur une balise).

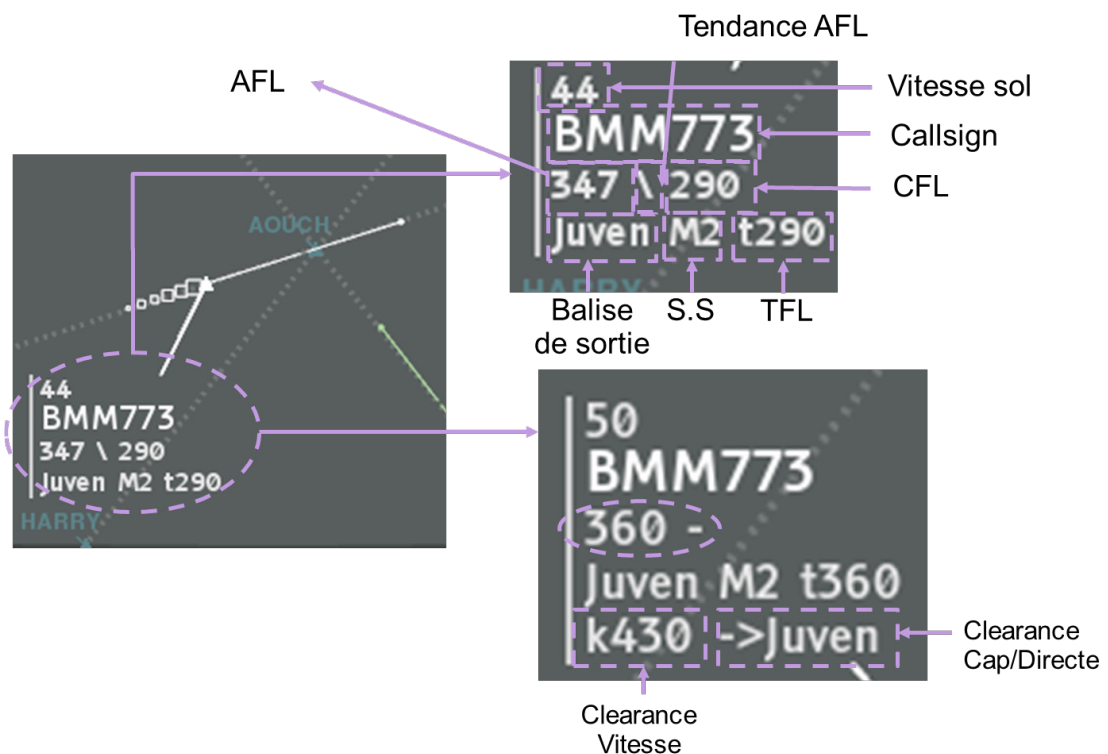


Figure 48 : détail d'une étiquette

Pour afficher les notifications nous avons fait le choix d'appliquer les designs Color, ColorBlink et BoxAnimation sur le type d'information détectée par SAmonitoring, c'est-à-dire les zones CFL, EFL/TFL/ Cap/Directe et Callsign pour les états de vols. Par exemple un problème de CFL va changer la couleur du champ CFL de l'étiquette, si l'alerte est critique le champ passe en rouge et les chevrons de BoxAnimation sont affichés jusqu'à validation de l'alerte. Les problèmes détectés sur les états de coordination ou de fréquence sont affichés sur l'identifiant du vol ligne 1.

Pour valider une alerte (ne plus l'afficher), il suffit de passer le curseur sur l'étiquette radar plus de 800ms. Les contrôleurs souhaitaient pouvoir les faire disparaître sans interaction complexes. La figure suivante (Figure 49 : les 3 niveaux de notification) montre les trois niveaux de notifications appliqués sur l'identifiant de vol, le TFL et l'AFL.



Figure 49 : les 3 niveaux de notification

Lorsqu'une alerte est supprimée de l'affichage par le contrôleur, celle-ci reste tout de même affichée jusqu'à ce que le problème détecté soit corrigé dans une liste de notifications (Figure 50 : la liste des notifications en cours) qui donne l'indicatif de l'avion, le type de problème concerné (état de vol, CFL, EFL, cap, vitesse, directe) ainsi que la raison qui a déclenché la notification. Si le statut de l'alerte évolue au cours du temps, par exemple une alerte Tfl de type mineur qui a été affichée puis supprimée de l'affichage sans correction par le contrôleur passe au type majeur, la notification sur l'étiquette sera réaffichée. Ce mécanisme est aussi appliqué si le niveau de criticité ne change pas, mais que la raison (par exemple on passe de vitesse verticale trop faible à vitesse verticale trop élevée) a changé.

La liste de notification est un élément très important de l'interface vers SA-monitoring puisqu'elle permet de comprendre la raison de déclenchement d'une notification. Tout problème détecté ne disparaîtra pas de la liste avant sa résolution même si le contrôleur a supprimé le feedback sur l'étiquette. Parfois certaines alertes critiques en particulier celles concernant le niveau de vol pourront être déclenchées parce que le contrôleur a omis de renseigner son système après avoir autorisé un avion à changer de niveau. Dans ce cas il est essentiel pour eux de rapidement pouvoir supprimer la notification sur l'étiquette, à eux de gérer leur temps et le moment où ils mettront à jour le CFL dans le système, la liste de notification est là pour maintenir l'action à réaliser sur le vol. En fonctionnement nominal cette liste doit être vide.

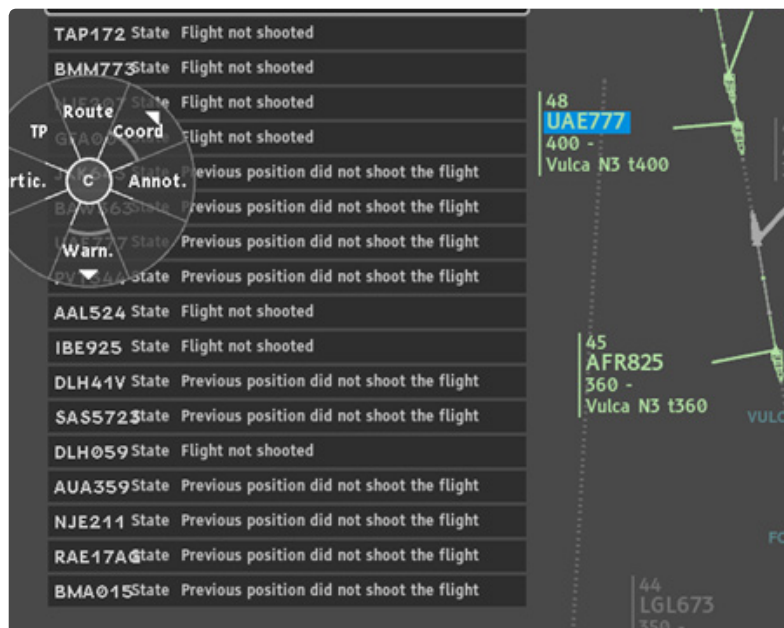


Figure 50 : la liste des notifications en cours

Pour afficher ces notifications, les supprimer par survol et afficher la liste de notifications, un plugin a été développé et est chargé au lancement de la supervision radar. Il se connecte au bus Ivy et s'abonne aux messages issus de SAmonitoring. Il délivre sur Ivy un message lorsqu'une notification est enlevée par l'utilisateur pour le module de log de l'expérimentation.

5.5.3. Retour d'expérience de l'intégration d'oculomètre dans un simulateur d'étude

Des différentes expériences où nous avons mis en œuvre de l'oculométrie, nous avons identifié un certain nombre de questions à se poser pour faire les choix d'architecture :

→ Quelles données oculométriques sont nécessaires au processus d'évaluation ?

données brutes, diamètres pupillaires, données couplées à l'IHM etc..

→ Quelles sont les contraintes environnementales ?

expérimentation en labo, en salle ou il est possible de contrôler la luminosité, en extérieur pour faire le choix du type d'oculomètre utilisable posé ou porté

→ Quelles sont les exigences en matière d'enregistrement de données ?

La plateforme d'évaluation utilise t'elle plusieurs ordinateurs/agents ? est-il nécessaire de synchroniser temporellement pour pouvoir comparer les données? La collecte des données se fait-elle de façon centralisée ou répartie ?

→ Quelles sont les exigences en matière d'analyse de données ?

Doit-on réaliser des corrélations entre différentes sources de données (c.-à-d. données IHM/applicatives/oculométrique), que faire pour faciliter le post traitement ?

De ces expériences nous formulons les recommandations suivantes pour l'intégration d'un oculomètre dans un simulateur d'étude :

- Utiliser un bus logiciel (middleware) pour permettre une architecture simple et modulaire.
- Utiliser le protocole NTP pour résoudre les problèmes de synchronisation temporelle.
- Même dans le cas d'évaluation en laboratoire (étapes 1 et 2 du cadre CST) où les ingénieurs de recherche ont un contrôle total sur le code source, l'utilisation de l'architecture la plus flexible (cas d'utilisation 2) est la plus pertinente puisqu'elle conduira à la création de composants réutilisables pour des expérimentations futures.
- Il est préférable de séparer les fonctions identifiées dans un composant logiciel unique et de les connecter via un bus logiciel. Par exemple, un algorithme de fixation de coordonnées brutes ou la détection de TEPR peut être considéré comme un module optionnel et réutilisable lors de futures expérimentations.
- La passerelle entre l'oculomètre et le bus logiciel ne doit être qu'une courroie de transmission et délivrer uniquement des données brutes pro-

venant de l'oculomètre pour être traitées par un autre composant logiciel, aucun traitement ne doit être fait. Cela permet de remplacer un modèle d'oculomètre par un autre simplement en changeant la passerelle.

- Les coordonnées brutes provenant de la passerelle vers l'oculomètre doivent être stockées dans le fichier de log même si les données de suivi oculaire de niveau élevé sont déjà stockées, car ils peuvent être utiles plus tard à des fins d'analyses complémentaires ou être traités avec par exemple un nouvel algorithme de fixation ou un nouveau module.
- Les calculs en temps réels, par exemple la corrélation avec les événements IHM, doivent être effectués lorsque c'est possible afin de stocker des données de niveau élevés dans les logs (par exemple les fixations sur les objets) afin de simplifier la phase de post-traitement et d'analyse.
- Utiliser si possible l'architecture de plugin pour l'IHM. Elle permet une intégration plus facile des données de haut niveau pour adapté l'IHM en utilisant des données oculométriques et permet une corrélation directe entre les fixations et les objets graphiques.

6. ÉVALUATION DE DESIGNS DE NOTIFICATIONS VISUELLES : L'EXPÉRIMENTATION SAILLANCE

6.1. Présentation

L'objectif de l'expérimentation Saillance est d'évaluer des designs de notifications dans un environnement expérimental contrôlé, mais réaliste du point de vue opérationnel en termes de design et de position de travail en s'aidant de capteurs oculométriques pour la mesure de performance. Dans le contexte du projet SESAR, plusieurs IHM innovantes ont été développées pour intégrer de nouveaux moyens d'interaction et de saisie ainsi que de nouvelles informations provenant d'automates ayant pour but d'améliorer la capacité des secteurs de contrôle aérien (Martin, 2012). Dans le cadre de ce projet, il était intéressant de comparer les performances de plusieurs designs de notifications et de les confronter à des designs utilisés dans l'opérationnel aujourd'hui de façon à intégrer les plus adaptés d'entre eux dans ces nouvelles IHM.

6.2. Protocole

6.2.1. Participants

Trente spécialistes ATC travaillant pour le centre de recherche de l'aviation civile française (principalement des ingénieurs et cinq contrôleurs) se sont portés volontaires pour participer à l'expérience. Ils étaient tous bien informés sur les interfaces opérationnelles et sur l'activité de contrôleur aérien. Les contrôleurs avaient tous plus de 20 ans d'expérience dans des centres français, ils sont actuellement employés comme experts (SME) en support aux projets réalisés dans le centre de R&D.

6.2.2. Procédure de passation

À leur arrivée les participants étaient positionnés à 70 cm de l'écran affichant Laby et la phase de calibration de l'oculomètre Facelab était réalisée. Ensuite ils devaient réaliser une phase de formation avec un exercice d'entraînement différent du scénario joué pendant l'expérience jusqu'à ce qu'ils soient parfaitement à l'aise dans l'exécution de la tâche de guidage de Laby et la tâche de validation des notifications. Les sujets devaient alors réaliser la tâche expérimentale qui durait environ 20 minutes. À la fin de la passation, il y avait une séance de discussion d'une quinzaine de minutes dans laquelle les participants avaient l'occasion de fournir leurs opinions subjectives sur les cinq types de notifications. Une passation complète incluant la formation, la calibration de l'oculomètre, la tâche expérimentale et le débriefing final durait environ 50 minutes.

6.2.3. Hypothèses de travail

Au regard des points théoriques développés, nous proposons les hypothèses suivantes :

H1 : En fonction des designs, des différences significatives dans les temps de perception devraient apparaître.

H2 : Le temps mis pour valider les notifications devrait dépendre du pouvoir intrusif de la notification sur la tâche principale.

H3 : Les designs prototypes devraient être plus efficaces que les designs opérationnels en raison de l'utilisation du mouvement.

H4 : Les designs très saillants devraient avoir un impact sur la tâche principale.

6.2.4. Variables

Conformément à l'approche multifactorielle décrite dans le chapitre évaluation, les différentes mesures enregistrées lors d'une passation portent sur la tâche principale (faire voler l'avion dans le labyrinthe en exécutant correctement les consignes demandées), la tâche secondaire de validation des notifications dès leur apparition ainsi que les mesures oculométriques et les reports subjectifs.

Sur la tâche de guidage et de saisie sur l'avion central nous collectons différentes données nous permettant d'estimer le temps de validation consigne (TVC), son type (niveau ou directe) et si elle a été renseignée correctement ou pas. Sur la tâche de validation des notifications, nous enregistrons le temps de validation de la notification (TV) si elle a été validée ou pas et son type (TN) (par exemple Color, ColorBlink) ainsi que la distance à laquelle elle est apparue de l'avion guidé (D). La distance en cm va nous permettre de contrôler a posteriori la distance réelle entre l'avion et l'alarme. Cette variable de contrôle nous permet de vérifier qu'il n'y a pas une trop grande disparité de distance entre les conditions. Pour les données oculométriques sont stockés la durée de fixation (TF), le callsign de l'avion regardé (E), le temps de perception avant la validation d'une notification (TPAV) et le temps de validation après la perception d'une notification (TVAP).

6.3. Résultats

Des ANOVAs unidirectionnelles à mesures répétées ont été réalisées pour évaluer l'effet du type de notification sur chaque variable dépendante. La procédure de Greenhouse-Geisser (1959) a été appliquée sur chaque effet pour lequel l'hypothèse de sphéricité de la matrice de variance/covariance n'a pas été validée. Pour les effets significatifs, les tests de comparaisons multiples ont été effectués en utilisant la correction de Bonferroni (Dunn 1961) avec un seuil de significativité pour chaque comparaison dépendant de leur nombre global. Les barres d'erreur représentent les intervalles de confiance inter participants à 95% obtenus par la méthode de Masson et Loftus (2003).

6.3.1. Réalisation de la tâche de guidage

Nous avons évalué la performance sur la tâche principale (la tâche de guidage et de saisie des consignes pour l'avion dans le labyrinthe) en regardant le taux d'erreur de saisie des consignes. La performance de saisie a été comparée pour les cinq designs de notification (Figure 51 : % de saisie correcte tâche principale par type de notification).

La consigne passée aux sujets était de valider les notifications dès qu'ils les percevaient même si une saisie était en cours et un menu ouvert. Valider une notification ne fermait pas les menus, il était donc aisé de finir la saisie après la validation. Cette consigne explique les taux d'erreurs montrés dans la figure suivante (entre 75 et 90%) qui peuvent sembler assez faibles, mais sont cohérents compte tenu de la consigne donnée. Un niveau de performance de 85% semble toutefois acceptable dans la mesure où il s'agit d'une double tâche sous pression temporelle (durée moyenne des notifications de 19s, durée des consignes 9s).

Une ANOVA unidirectionnelle a mis en évidence un effet du type de notification sur la performance de la tâche principale, $F(4,116) = 11.91$, $p < 0.001$. Les tests de comparaisons multiples montrent que ShadowMask dégradait beaucoup plus la tâche principale que Color ColorBlink et BoxAnimation ($ps < 0.006$). De la même façon, la performance était sensiblement dégradée lorsque des notifications de type Halo étaient présentées en comparaison de ColorBlink et BoxAnimation ($ps < 0.002$).

Halo et ShadowMask semblent donc dégrader la tâche principale bien plus que les autres designs. BoxAnimation et ColorBlink sont ceux qui ont permis aux sujets d'obtenir les meilleurs résultats sur la tâche de saisie. L'hypothèse **H4** est validée, l'analyse statistique montre deux groupes, Color, ColorBlink et BoxAnimation d'une part et Halo, ShadowMask d'autre part.

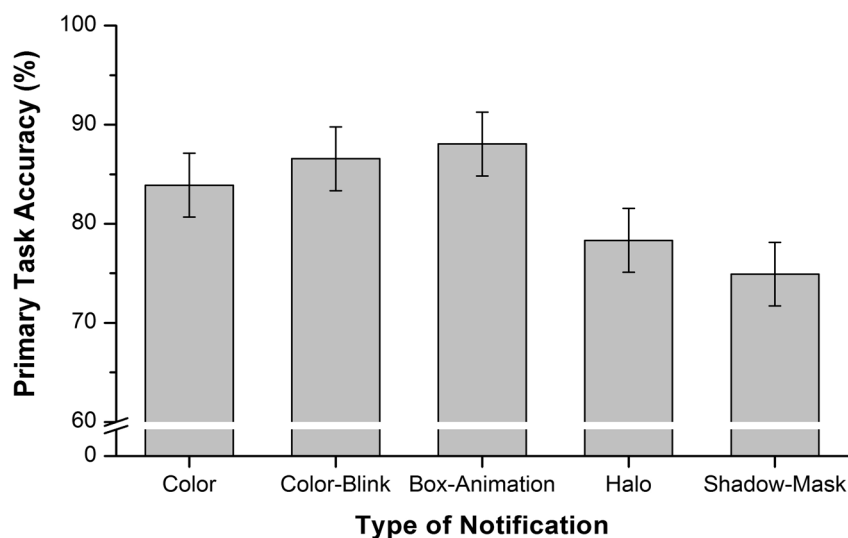


Figure 51 : % de saisie correcte tâche principale par type de notification

6.3.2. Réalisation de la tâche de validation des notifications

La tâche de validation des notifications consistait à cliquer sur l'étiquette du vol notifié pour la faire disparaître dès que les sujets percevaient le stimulus, il ne pouvait y avoir qu'une seule notification à la fois. La durée moyenne des notifications était de 19.62 s. Le délai entre l'apparition de notification et sa validation a été comparé entre les conditions de façon à déterminer la saillance de designs évalués.

Une transformation logarithmique a été appliquée aux données avant la réalisation de l'ANOVA afin de normaliser la distribution des temps de validation pour chaque condition dont l'asymétrie était positive.

L'analyse menée sur les données transformées a montré une différence significative entre les cinq types de notifications, $F(4,116) = 119.75$, $p < 0.001$ (Figure 52 : Délai de validation (s) pour chaque type de notification

Les tests post-hoc ont révélé que les participants étaient significativement plus lents à valider les notifications Color et ColorBlink que les notifications BoxAnimation, Halo ou ShadowMask ($p_s < 0.001$ pour tous les tests). En outre, les temps de validation transformés étaient plus faibles pour Halo et à ShadowMask que pour BoxAnimation ($p_s < 0.04$).

Les notifications apparaissaient hors de la zone de confort ($0-15^\circ$) centrée sur l'avion objet de la tâche principale. Un tirage aléatoire sur les avions hors de cette zone était effectué pour sélectionner l'avion-cible de la notification. La médiane de la distribution des distances tirée aléatoirement se situait à environ 30cm de l'avion principal soit 23° d'angle visuel. Nous avons voulu vérifier s'il y avait des différences de temps de validation entre les différents designs lié à la distance ; les mesures ont donc été séparées en fonction de ce seuil de 23° et des ANOVA à mesures répétées ont été réalisées sur les données transformées.

Elles ont montré un effet principal de la distance, $F(1,29) = 6.76$, $p = 0.015$, avec des temps de validation transformés plus importants pour les grandes distances. Ce résultat semble assez logique dans la mesure où le geste à réaliser est plus important, mais sans interaction, $F(4,116) < 1$.

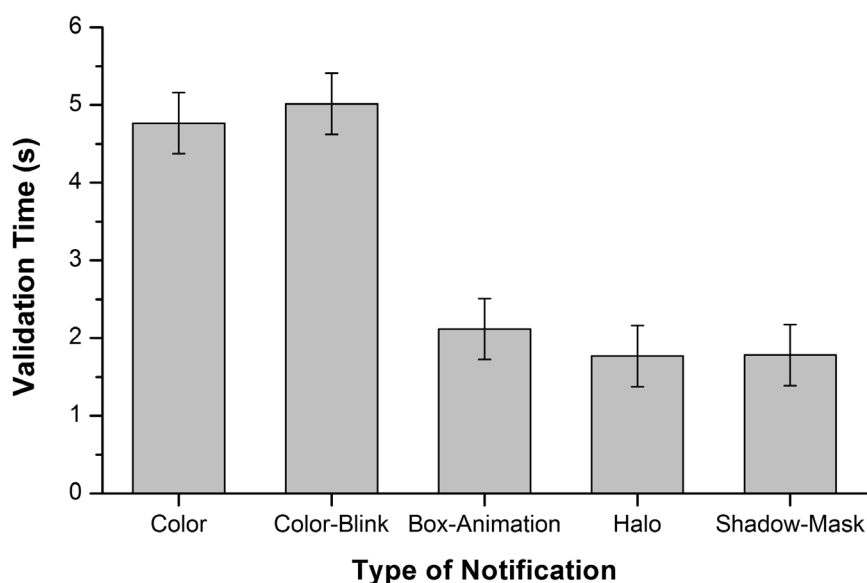


Figure 52 : Délai de validation (s) pour chaque type de notification

Des notifications n'ont pas été perçues ou validées à temps (avant 20s) pendant les passations. Sur les 1200 notifications présentées, 20 notifications de type Color et 21 notifications de type ColorBlink n'ont pas été validées (variable E). Il n'y a eu aucune notification manquée pour les trois autres designs. Le nombre de notifications manquées aurait sans doute été plus important si celles-ci avaient été présentées moins longtemps ; mais un délai plus long pour valider n'aurait sans doute pas amélioré les choses puisque 95% des notifications ont été traitées dans un délai de 6.4 s. Le nombre d'erreurs ne semble pas lié à la distance, 13 erreurs ont été faites pour les notifications qui sont apparues à moins de 30 cm / 23°, et 10 autres concernant des notifications à une distance supérieure à 30 cm. La distance moyenne d'une notification manquée était de 29.93cm (24.50 variance, asymétrie 0.15).

H2 et **H3** sont donc validées par ces résultats. Deux groupes apparaissent encore une fois dans ces analyses, Color, ColorBlink d'une part avec des résultats en performance de validation inférieurs à l'autre groupe de designs (prototypes) qui utilisent du mouvement : BoxAnimation, Halo et ShadowMask.

6.3.3. Résultats oculométriques

Trois mesures ont été effectuées à l'aide du dispositif oculométrique FaceLab : la durée des fixations (TF), le temps mis pour percevoir une notification (TPAV) avant sa validation et le temps mis pour valider une notification après sa perception (TVAP).

La première analyse porte sur le TPAV (Figure 53 : Temps en (s) mis pour percevoir une notification (TPAV)), il s'agit du délai entre l'instant d'apparition de la notification et l'heure de début de la première fixation sur l'avion notifié. L'ANOVA a montré un effet lié au type de design, $F(4,116) = 34.54$, $p < 0.001$, avec une tendance similaire à celle des temps de validation : les participants étaient plus rapides à détecter et à fixer les trois designs prototypes utilisant du mouvement BoxAnimation, Halo et ShadowMask que Color ou ColorBlink ($ps < 0.001$ pour tous les tests multiples).

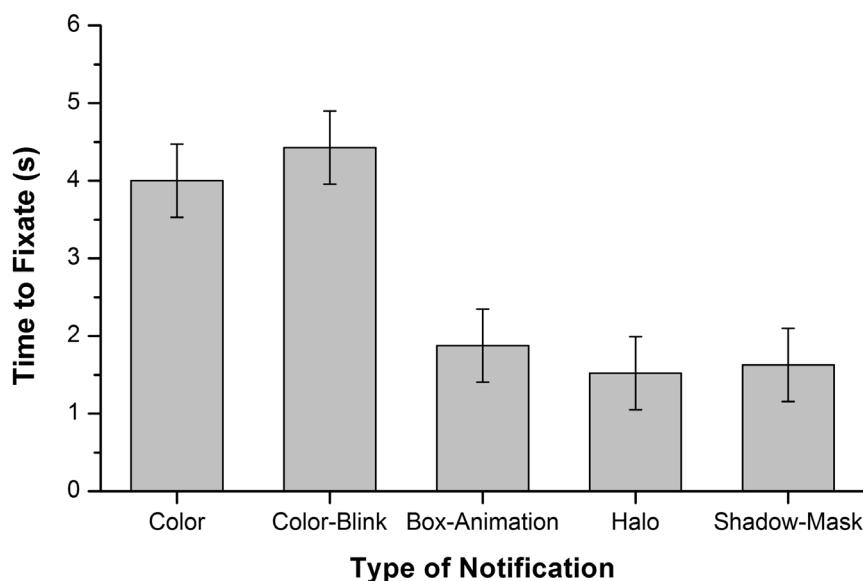


Figure 53 : Temps en (s) mis pour percevoir une notification (TPAV)

La deuxième analyse porte sur le TVAP, c'est-à-dire le temps mis pour valider une notification en cliquant sur l'avion notifié après sa perception (Figure 54 : Temps mis-en (s) pour valider une notification après perception (TVAP)). Le but de cette analyse était de vérifier si après interruption et perception d'une notification, un effet du design pouvait être observé sur le temps mis pour le valider et donc quantifier l'impact intrusif du design.

On observe un effet significatif entre les designs $F(4,76) = 3.987$, $p < 0.05$ et la même tendance que celle observée sur les mesures précédentes, avec des temps de réponse plus lent pour Color et ColorBlink.

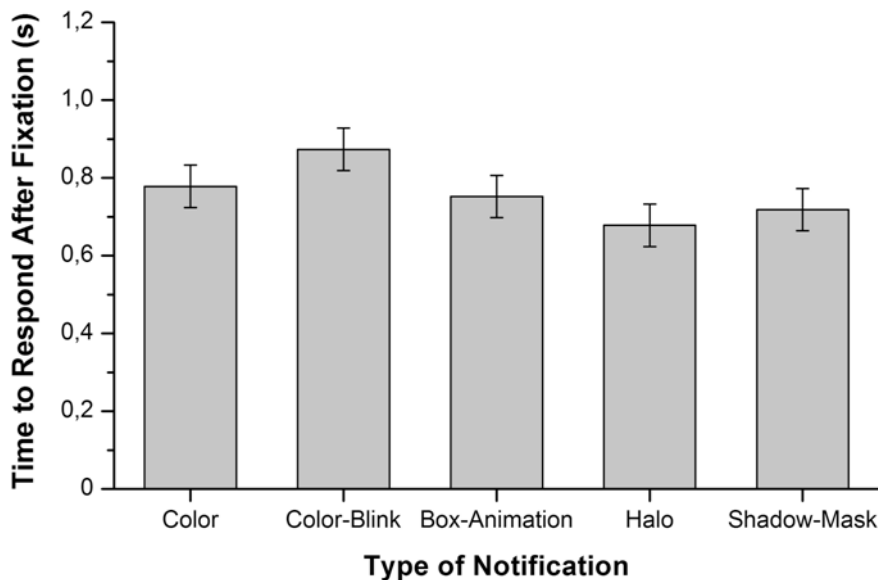


Figure 54 : Temps mis-en (s) pour valider une notification après perception (TVAP)

La dernière analyse porte sur la durée de fixation (TF) avec un seuil minimum de 50 ms après la perception (Figure 55 : durée de fixation), qui représente le temps nécessaire pour les participants de comprendre la notification afin de donner une réponse. Dès que perçus, nous émettons l'hypothèse que l'ensemble des designs sont bien lisibles et ne posent pas de problèmes spécifiques d'acquiescement en fonction du type de notification. La significativité n'est d'ailleurs pas confirmée, il n'y a pas de différence significative entre les différents designs, $F(4,116)=2.22$, $p>.05$).

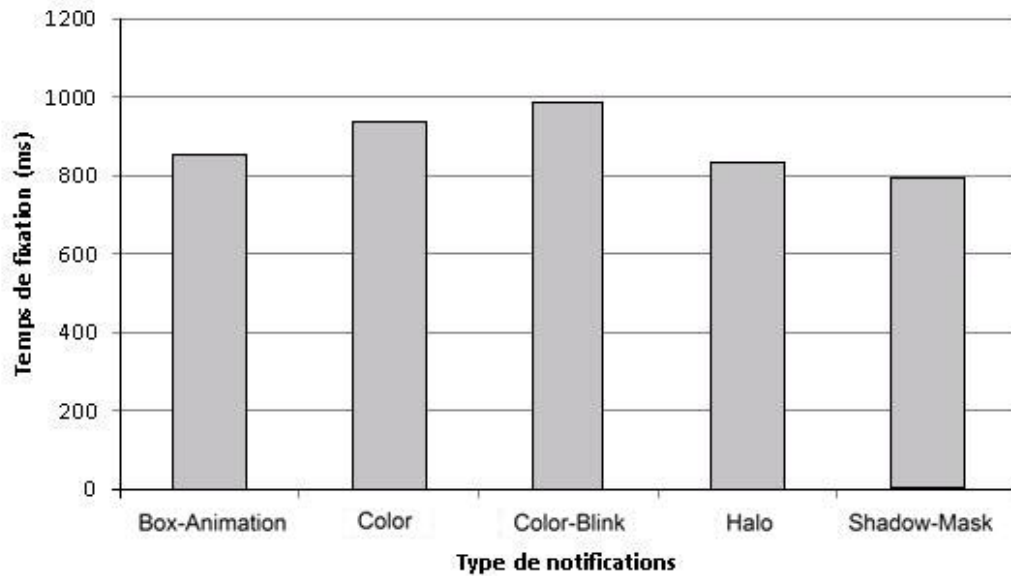


Figure 55 : durée de fixation

Toutefois, les tendances au niveau des différences entre les designs, même si elles ne sont pas ici statistiquement validées, laissent supposer que le type de conception peut parfois ralentir l'efficacité des actions. Par exemple, Halo et ShadowMask se détachent avec des temps de fixation les plus courts. BoxAnimation se place en intermédiaire alors que les temps les plus longs sont sur le Color et ColorBlink.

6.3.4. Résultats qualitatifs

Les commentaires des participants de la session de débriefing après l'expérience sont résumés dans le tableau 1. Ces rapports subjectifs correspondent relativement bien aux mesures objectives en suggérant que les designs Color et ColorBlink sont mieux perçus dans la tâche principale, mais peuvent être manqués, tandis que Halo et ShadowMask bien que perçus instantanément sont peut-être trop intrusifs. BoxAnimation apparaît clairement comme le meilleur compromis pour la conscience de la situation dans ces reports subjectifs.

Parmi les participants, et tout particulièrement les contrôleurs, certains ont jugé ShadowMask comme trop gênant pour pouvoir être un jour utilisé dans une position de contrôle opérationnelle. Les contrôleurs ne voient à priori pas de raison valable qui justifierait de faire disparaître même partiellement tout ou partie de leur visualisation radar.

	Commentaires positifs	Commentaires négatifs
Color	N'interromps pas la tâche principale La couleur rouge attire l'attention lors du scanning visuel de l'image radar	Utile pour les alertes non urgentes Pas perceptible si l'attention est portée sur une autre partie de l'écran Seulement perçue pendant le scanning visuel de l'image radar
ColorBlink	N'interromps pas la tâche principale Rouge clignotant est très efficace (synonyme de danger) Approprié pour les alertes non urgentes Mieux perçue que Color (mais encore insuffisante)	La notification est manquée ou ralentie si l'attention est focalisée trop loin Uniquement perçue lors des phases de scanning visuel de l'image radar
BoxAnimation	Pas trop intrusive, attire l'attention juste ce qu'il faut Perceptible immédiatement sans surcharger l'image Meilleur compromis entre la perception de l'alerte et le maintien de l'attention sur la tâche principale Mieux perçue que ColorBlink parce qu'elle occupe plus d'espace En utilisation opérationnelle ça serait le meilleur compromis à court terme	Il est difficile de regarder ailleurs lorsque les chevrons pulsent Perception plus lente des alertes BoxAnimation quand l'avion guidé dans le Laby sur lequel se porte l'attention se trouve loin de l'alerte
Halo	Alerte perçue quasiment instantanément Alerte qui ne peut être manquée Attire l'attention sans être trop intrusive et sans trop interrompre	Charge trop l'écran alors qu'il contient déjà un grand nombre d'informations Attire trop l'attention, et interrompt trop la tâche principale
ShadowMask	Alerte qui ne peut être manquée Idée intéressante, mais le masque (voile) devrait être moins opaque.	Extrêmement intrusive Oblige à interrompre une tâche qui est possiblement plus importante que l'alerte Il est plus long et plus difficile de reprendre la tâche principale après cette alerte. Comment peut on gérer des alertes simultanées ? Le seul des cinq designs qui semble inutilisable en condition réelle dans l'ATC

Table 1 : Synthèse des débriefings

6.4. Conclusion

6.4.1. Analyse des résultats

Nous formulions plusieurs hypothèses avant de réaliser cette expérimentation, les résultats obtenus par rapport à celles-ci sont les suivants :

H1 : Partiellement validée, des différences significatives dans les temps de perception ne sont apparues qu'entre deux groupes de design ceux avec ou sans mouvement.

H2 : Validée, le temps mis pour valider les notifications dépend du pouvoir intrusif de la notification sur la tâche principale.

H3 : Validée, les designs prototypes sont plus efficaces que les designs opérationnels en raison de l'utilisation du mouvement.

H4 : Validée, les designs très saillants ont un impact sur la tâche principale en la dégradant.

Cette expérience avait pour objet d'évaluer l'intérêt et la performance de cinq types de designs de notifications visuelles dans un contexte ATC (tâche de surveillance sur grand écran avec des notifications en vision périphérique) en tenant évaluant les critères suivants :

- la vitesse à laquelle les designs sont perçus,
- le coût attentionnel (perturbation liée à l'interruption de design trop saillant),
- les problèmes de perception (notifications manquées lorsque les designs ne sont pas assez saillants).

Les trois designs utilisant du mouvement (BoxAnimation, Halo, ShadowMask) ont été perçus rapidement et sans erreur. Color et ColorBlink sont moins efficaces avec des temps de validation plus lents et des notifications en vision périphérique qui sont manquées. Cependant ces résultats sont à mitiger puisque le pouvoir attentionnel de Halo et ShadowMask s'est révélé être un handicap lorsque l'on prend en compte la performance des sujets dans la réalisation de la tâche principale. Cela illustre la nécessité d'adopter une approche holistique qui tient compte de multiples aspects de la tâche de l'opérateur lors de l'introduction de nouvelles solutions technologiques.

La principale caractéristique des attracteurs visuels est de pouvoir capter l'attention lors d'une tâche en cours pour avertir l'opérateur d'une information à traiter dans l'environnement de travail. Toutes les alertes utilisant le mouvement ont été perçues et nous avons pu mettre en évidence le phénomène de cécité inattentionnelle pour les deux designs utilisant uniquement la couleur. Le mouvement est connu pour être bien plus efficace en terme de saillance attentionnelle (Bartram et al., 2003; McCrickard et al., 2001). La recherche dans le domaine de la vision indique que les caractéristiques de couleur d'un stimulus ont la capacité d'attirer l'attention (par exemple, Turratto & Galfano, 2000), mais ce pouvoir attentionnel peut être réduit dans les environnements de type C2 complexes avec des affichages de données très riches (Nikolic et al, 2004). Nous avons pu observer moins de notifications manquées avec l'animation de couleur de ColorBlink en comparaison de Color, mais étant donné que ces deux designs sont actuellement utili-

sés de façon opérationnelle dans l'ATC, le fait que certaines notifications soient passées inaperçues est relativement inquiétant dans un environnement aussi critique pour la sécurité. Lorsqu'un opérateur de supervision de système complexe est profondément engagé dans une tâche exigeante, son champ visuel tend à se réduire ce qui rend difficile la perception et l'utilisation des informations périphériques (Chan & Courtney, 1993). La cécité inattentionnelle est plus susceptible de se produire dans des conditions de forte charge perceptive (Cartwright, Finch & Lavie, 2007) ou dans le cas d'une augmentation de la charge de la mémoire visuospatiale (Todd et al., 2005), deux conditions susceptibles de se produire dans la tâche de supervision des contrôleurs aériens. Les temps de perception de Color et ColorBlink varient en fonction de l'endroit où la notification est apparue : une réponse immédiate est possible pour celles affichées à proximité du point de focalisation visuelle de l'opérateur, mais lorsque le stimulus apparaît dans la vision périphérique la réponse peut être retardée ou le stimulus manqué.

Des expérimentations complémentaires permettraient de déterminer si le choix de contraste de Color et le cycle d'animation pour ColorBlink pourraient être rendus plus efficace, mais nous formulons l'hypothèse que dans le cas où un opérateur se trouve dans une phase de forte charge de travail, il se retrouverait dans la même situation de tunnellation attentionnelle et pourrait ne pas remarquer une alerte plus lumineuse ou qui clignote avec un cycle différent dans la zone de vision périphérique.

Le deuxième critère à prendre en compte dans la conception de notifications visuelles est que, même si elles doivent attirer l'attention, cela doit se faire en conformité avec leur importance relative dans le déroulement de la tâche. Dans le cas de Halo et ShadowMask, la perception a été presque immédiate, mais leur caractère intrusif était également un inconvénient en termes de performance sur la réalisation de la tâche en cours. Comme la littérature le suggère, l'œil est attiré par le mouvement (Bartram et al., 2003), et, dans le cas de Halo qui démarre sur l'objet de la tâche en cours, l'attention se porte naturellement sur les cercles concentriques et la zone désignée par la direction du cercle. L'inconvénient est qu'il tend à occuper l'écran en raison de sa conception et par conséquent charge visuellement une image qui contient déjà un grand nombre d'informations. Lorsque dans des situations particulières la réponse à une notification doit être retardée, car l'opérateur est en train de gérer un problème plus important, la notification déjà déclenchée se poursuivra et pourra devenir une source de distraction qui compromet une tâche potentiellement de plus grande importance. Dans le domaine des notifications sonores, des notifications de bas niveau opérationnel peuvent être préjudiciables à des tâches hautement prioritaires (Banbury, Fricker, Tremblay, & Emery, 2003), ce qui peut entraîner des conséquences plus graves que celles provoquées par l'alerte d'origine. Pour ShadowMask, l'attention est automatiquement dirigée vers l'avertissement visuel, car toutes les autres informations sur l'écran sont dégradées par l'application d'un masque opaque. Ce design exige donc une attention immédiate et oblige l'opérateur à interrompre la tâche en cours. Il a été démontré que les interruptions peuvent porter atteinte à la conscience de la situation dans des contextes de type C2 (Hodgetts et al., en cours de publication), et peuvent être particulièrement perturbatrices s'il n'y a pas la possibilité de consolider des aspects de la tâche en cours avant de basculer à une autre tâche (Hodgetts et Jones, 2006). Ainsi conformément avec notre approche holistique,

nous constatons que les designs aboutissants à une amélioration des temps de perception sont sujets à une diminution de la performance sur la tâche principale. De plus, nos mesures subjectives sont conformes avec les résultats qualitatifs, les alertes de Halo et ShadowMask étaient perçues comme impossibles à manquer, mais très déstabilisantes dans la tâche de guidage. Ces deux designs soulèveraient également des problèmes dans le cas de notifications multiples.

Utilisant l'approche de l'ingénierie cognitive et d'un système cognitif global, la technologie devrait aider l'opérateur à optimiser la performance du système homme-machine sans encourir de coût dans d'autres domaines. Nous constatons donc que la conception de BoxAnimation atteint un meilleur compromis entre le coût attentionnel et l'efficacité en termes de temps de détection. Il était très efficace même en cas de forte charge, mais en impactant faiblement la performance dans la réalisation de la tâche principale. Contrairement à ColorBlink, l'animation a lieu en dehors de l'étiquette augmentant la taille perçue de l'objet, l'effet s'additionnant à une meilleure détection du mouvement en vision périphérique ce qui explique pourquoi cette notification est perçue rapidement. Il a également l'avantage sur Halo et ShadowMask de ne pas surcharger la visualisation ou forcer l'opérateur à se détourner de la tâche principale, et semble donc montrer les notifications de façon optimale.

Plutôt que d'abandonner complètement les designs Halo et ShadowMask, nous pourrions envisager leur utilisation dans des situations d'urgence extrême, où il peut être nécessaire d'utiliser des méthodes radicales pour alerter afin que des mesures soient prises immédiatement. Dans le cas de l'effet de tunnellation attentionnelle (Wickens et Alexander, 2009), les opérateurs se concentrent parfois trop sur un élément particulier d'une tâche au détriment d'autres paramètres, et ne prennent pas en compte (ou ne perçoivent pas) les notifications (Beringer & Harris, 1999 ; Dehais et al., 2010). ShadowMask qui dégrade significativement la visualisation pour attirer l'attention sur l'objet notifié est impossible à manquer, ce type de design est comparable aux techniques de contre-mesures cognitives dans lesquelles l'information est supprimée afin d'aider au désengagement de la tâche en cours dans le cas de tunnellation attentionnelle (voir Dehais, Causse, & Tremblay, 2011). Surtout, le type d'alerte doit être adapté à la gravité de la situation qu'il rapporte.

6.4.2. Discussion et applications

Lors de la conception des notifications, il est nécessaire d'atteindre un équilibre entre le pouvoir attentionnel du design et son importance relative dans le contexte de la tâche. Cette étude a démontré que les designs utilisés dans l'opérationnel aujourd'hui devraient peut-être être modifiés. Si Color qui est utilisé pour des avertissements a des temps de perception moyens satisfaisants compte tenu de leur importance dans la tâche, le design ColorBlink qui correspond à une alerte critique dans le déroulement de la tâche de contrôle a des temps de détection trop longs. Il n'est pas acceptable que certaines alertes montrées en vision périphérique aient pu être manquées. En outre, ces deux designs opérationnels ne diffèrent pas significativement en saillance entraînant des temps de détection comparables alors que le besoin opérationnel était de les différencier. Ces deux designs pourraient

donc être employés pour des notifications non urgentes qui ne nécessitent pas l'interruption de l'opérateur. Les designs prototypés pour l'expérimentation étaient beaucoup plus efficaces grâce à l'utilisation de mouvement, bien qu'il n'y ait pas de différence significative de saillance établie entre eux, c'est leur impact sur la tâche principale qui nous a permis de distinguer l'un des designs. Les données quantitatives et qualitatives obtenues lors de cette expérimentation ont démontré que l'alerte BoxAnimation permettrait d'obtenir des temps de réaction très courts tout en laissant le contrôle à l'opérateur (il n'est pas 'obligé' de réagir), par ailleurs c'est le design produisant l'impact le plus faible sur la réalisation de la tâche principale. Ce type de design devrait donc être mis en place pour des alertes de type Filet de Sauvegarde (STCA) puisqu'il garantit la perception des alertes même dans des zones éloignées du point de focalisation de l'attention du contrôleur aérien.

En cas d'urgence absolue, il peut y avoir une place pour une alerte plus intrusive comme ShadowMask qui 'force' l'opérateur à changer d'activité, mais cela doit être examiné avec soin avec les opérateurs et utilisé dans des situations exceptionnelles. Une alerte de ce type ne doit être utilisée qu'en dernier recours, lorsque la perturbation de la tâche en cours est jugée impérative par un système intelligent de surveillance de la tâche. Dans le domaine du contrôle aérien, dans une situation de collision imminente qui est le genre d'incident qui viendrait à l'esprit pour une contre mesure cognitive de ce type, le système embarqué TCAS donne des instructions aux pilotes pour éviter la collision et le contrôleur aérien n'a plus ni le temps ni l'autorité pour faire éviter l'accident (les pilotes sont sensés suivre l'avis du TCAS quel que soit l'ordre donné par le contrôle aérien). Par ailleurs, on pourrait considérer que cela devrait être à l'opérateur et non à un système automatisé de choisir et de gérer la priorité de ses tâches et qu'il ne devrait pas être obligé de traiter une alerte très intrusive immédiatement et sans préavis, au détriment de tout le reste. Dans un contexte complexe et à haut risque comme celui du contrôle aérien, l'opérateur peut difficilement se permettre de perdre l'accès à l'écran radar pour traiter une alerte. En tout état de cause, des conditions de déclenchement extrêmement sévères devraient être mises en place afin d'éviter les fausses alarmes qui pourraient compromettre la sécurité dans une situation qui n'est pas une urgence vitale.

Dans le cas de Halo, une application possible de ce design a été trouvée grâce au contexte particulier de centre de contrôle en-route de Reims. En raison d'une organisation du trafic aérien particulière liée à ses interconnexions avec des centres étrangers et pour délivrer les avions avec un cadencement conforme aux lettres d'accord inter-centres, les contrôleurs doivent anticiper la prise en compte des aéronefs très en amont de leur propre secteur. Pour cela ils doivent dézoomer et excentrer considérablement leur image radar (normalement centrée sur leur secteur) pour voir une zone normalement affectée au secteur amont. Cet excentrement peut aboutir à ne plus voir une partie de leur secteur, si une alerte se produisait dans une zone non visible, aucun des designs évalués lors de cette expérimentation ne fonctionnerait hormis Halo qui montrerait dans quelle direction se produit l'alerte.

Sur le plan méthodologique, nous avons démontré que le micro-monde Laby est une plateforme expérimentale idéale pour l'évaluation des innovations dans le domaine de l'interface homme-machine pour le contrôle aérien, car il offre un équilibre entre la validité écologique et le contrôle expérimental.

Les micro-mondes représentent une étape importante dans le continuum du CSE de la recherche appliquée (Rasmussen, Pejtersen, et Goodstein, 1994) ; au contraire des études de laboratoire classiques qui utilisent des stimuli statiques pour évaluer des variables isolées, la nature complexe et dynamique des micro-mondes ainsi que leur réalisme signifie qu'ils sont les mieux placés pour comprendre les mécanismes cognitifs sous-jacents. Laby fournit une plateforme d'évaluation d'image radar pour le contrôle aérien qui offre un réalisme suffisant pour le participant, mais qui nécessite moins de ressources humaines lors de la mise en œuvre que des simulations plus complexes. Sa simplicité et sa flexibilité ont permis de développer rapidement l'extension permettant de réaliser des notifications en vision périphérique, de connecter un oculomètre et d'optimiser son fonctionnement pour le besoin expérimental avec des cycles itératifs courts. Le principe de scénario et l'éditeur de Laby permettent de concevoir différents types d'expérimentations en ajustant le nombre de consignes à réaliser pour éventuellement réguler un niveau de charge de travail. Pour généraliser les résultats obtenus avec Laby aux interfaces homme-machine opérationnelles, il sera nécessaire de valider les effets obtenus dans une simulation plus complexe.

Cette étude illustre l'importance d'une approche globale dans la conduite de l'évaluation concernant les nouveaux systèmes ou les outils de support à l'activité. Plutôt que de se concentrer sur une seule variable directement associée à l'objectif principal de l'expérience (par exemple, la notification du temps de validation), cette approche utilisant plusieurs mesures simultanées permet d'avoir une vue plus large sur l'effet que chaque design pourrait avoir sur la performance globale. Un effet positif sur une dimension peut être accompagné par un effet négatif sur une autre, ce que nous avons mis en évidence en analysant les temps de perception des notifications et en contrebalançant le résultat avec l'impact sur la tâche principale. Une approche holistique devrait donc être utilisée pour s'assurer que la mise en œuvre d'un nouveau système est réellement bénéfique dans tous les aspects de la réalisation de la tâche (Lafond et al., 2010).

7. ÉVALUATION DE DESIGNS VISUELS CONFRONTÉS À DES NOTIFICATIONS SONORES : L'EXPÉRIMENTATION BIP

7.1. Présentation

Cette expérimentation s'est déroulée à l'Université Laval, Québec dans le cadre d'une collaboration avec l'équipe Neuroergonomie et facteurs humains de l'ISAE par l'intermédiaire du docteur M. Causse, de L. Giraudet (doctorante de M. Causse) et du professeur S. Tremblay directeur du laboratoire CoDot de l'Université Laval.

L'expérimentation Saillance a permis de comparer plusieurs designs de notifications visuelles. En France, les positions de travail en approche utilisent à la fois des notifications visuelles et sonores. La théorie de la charge perceptuelle de Lavie (1995) évoque un couplage entre le visuel et l'auditif, et l'effet de l'augmentation de la charge de travail sur les capacités de traitement des distracteurs visuels et auditifs. Avec cette expérimentation, et en réutilisant les résultats de l'expérimentation Saillance, nous avons voulu évaluer les interactions entre les notifications visuelles et sonores dans le contexte ATC.

Dans le cadre de la collaboration avec L.Giraudet, nous avons poursuivi plusieurs objectifs :

- évaluer deux des designs les plus significatifs issus de l'expérience Saillance en les confrontant à des notifications sonores
- mesurer l'impact des designs de notifications visuelles sur la charge de travail à l'aide de mesures EEG ERP P300
- tenter de provoquer les conditions d'apparition de la surdit  inattentionnelle (sujet de thèse de L. Giraudet) avec Laby en jouant sur la difficulté de la tâche. La difficulté était modulée grâce à la vitesse de l'avion guidé dans le labyrinthe et la charge visuelle en jouant sur le nombre d'avions susceptibles de recevoir une notification.

7.2. Protocole

7.2.1. Participants

42 sujets, étudiants de l'Université Laval, Québec ont participé à l'expérience dont 16 femmes et 26 hommes. L'âge des participants variait entre 19 et 46 ans (moyenne de 24,4 ans avec un écart type de 6 ans). Ils étaient tous naïfs sur le but de l'expérience et ne connaissaient pas le phénomène de la surdit  inattentionnelle.

N'ont pas été sélectionnés les sujets prenant des substances affectant le système nerveux central ; avec antécédents neurologiques ou psychiatriques ou des problèmes de vue et d'audition.

Les sujets sélectionnés ont dû effectuer le test de Pichot (mesure de la fatigue), obtenir un score inférieur à 22 et réussir la tâche de contrôle auditive avec un score supérieur à 85%.

7.2.2. Procédure de passation

Les sujets devaient réaliser la tâche Laby avec les différentes conditions de design, vitesse ou charge visuelle accompagnés de la tâche auditive, la vitesse de la tâche auditive était conditionnée à la condition de vitesse du Labyrinthe. Ensuite ils devaient réaliser une tâche de contrôle où ne leur était présentée que la tâche auditive, 25 sons d'alarme pour 100 sons standards présentés avec une période aléatoire centrée sur 4.26s pour la tâche de contrôle à vitesse lente et 2.6s pour celle à vitesse rapide ; en cas de sons d'alarme ils devaient les valider en appuyant sur un bouton poussoir.

L'expérience se déroulait sur un ordinateur portable dans une salle insonorisée. Le participant était installé face à un écran de 30', distant de 60cm, une souris dans la main droite pour la tâche de guidage et la tâche de validation des notifications et un bouton-poussoir Powermate de Griffin Technology dans la main gauche pour la tâche d'acquiescement des notifications sonores (Figure 56 : dispositif expérimental expérimentation Bip). Des enceintes de part et d'autre de l'écran généraient les sons du paradigme oddball et le bruit blanc. Les puissances sonores étaient : 42dB pour le bruit blanc, 52.5dB pour les sons graves et aigus du oddball (sonomètre : Simpson 886-2 Sound Level Meter). Fréquences : sons standards à 1 000 Hz, alarmes sonores à 2 000 Hz, tous deux présentés pendant 500ms. La moyenne entre chaque son était de 4.2s pour les scénarios lents et 2.6s pour les rapides.

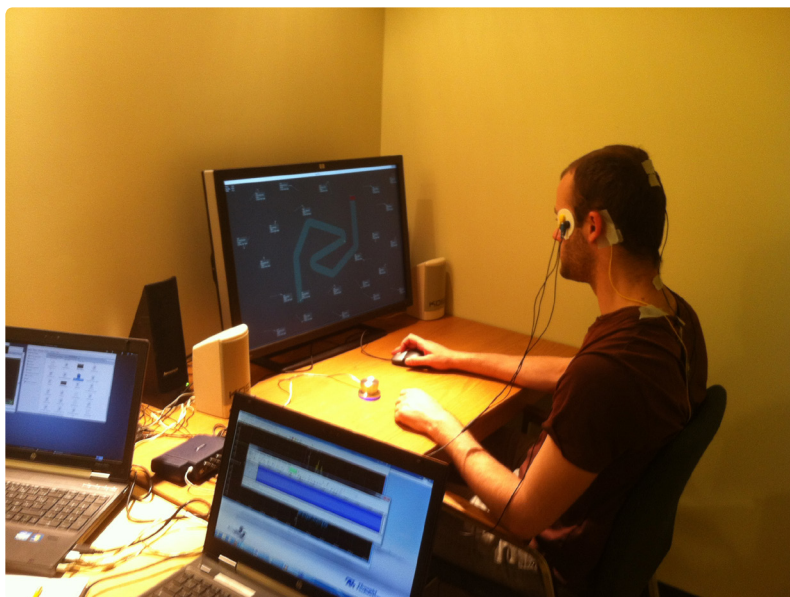


Figure 56 : dispositif expérimental expérimentation Bip

Les enregistrements EEG étaient effectués sur un portable dédié à l'aide d'un appareil Infinity FlexComp (Thought Technology Ltd). Trois électrodes ont été placées pour les mesures bipolaires : l'électrode positive sur le site

Pz (lobe pariétal), l'électrode de référence sur le côté gauche du front et l'électrode de masse avec un clip sur le lobe de l'oreille droite. Le signal EEG a été enregistré à une fréquence d'échantillonnage de 256Hz avec un filtre passe-haut de 0,5 Hz et un filtre passe-bas à 49Hz. L'analyse des données EEG a été effectuée à l'aide EEGLAB 11.0.3.1b (Delorme & Makeig, 2004) fonctionnant sous MATLAB 7.1 (The Mathworks). Pour synchroniser les logs de Laby et ces enregistrements, une voie de l'EEG a été utilisée pour recevoir des pulses émis par Laby au moment de l'émission des sons sur les haut-parleurs au travers d'une interface usb/port parallèle (un front montant était envoyé sur un bit du port parallèle). Pour cette raison, il n'a pas été nécessaire de connecter l'EEG à Ivy.

Une passation durait 2h30. À leur arrivée les participants devaient remplir la charte de consentement éthique et les questionnaires d'auto-évaluation de la fatigue (test de Pichot & Brun, 1984) et d'anxiété (STAI formulaire YB, Spielberger, 2010). Ensuite ils étaient formés à l'utilisation du Laby jusqu'à ce qu'ils se jugent suffisamment à l'aise avec, puis équipés avec les électrodes de l'EEG.

Les participants ont été séparés en deux groupes, selon le type d'alarmes visuelles utilisées (BoxAnimation ou ColorBlink). Chacun d'entre eux devait jouer les 4 scénarios Laby correspondants aux configurations suivantes :

- A : vitesse lente, peu d'avions périphériques
- B : vitesse lente, beaucoup d'avions périphériques
- C : vitesse rapide, peu d'avions périphériques
- D : vitesse rapide, beaucoup d'avions périphériques.

À l'issue de chaque scénario, un questionnaire d'auto-évaluation de la charge de travail (NASA-TLX) devait être rempli.

La tâche de contrôle était ensuite déroulée dans les deux conditions de vitesse (lente et rapide). Pendant la tâche de contrôle, une croix était affichée à l'écran que les sujets devaient regarder pendant son déroulement.

À l'issue des passations, les participants devaient remplir un questionnaire ad hoc sur l'expérience et la stratégie globale adoptée (notamment priorités données aux différentes tâches, ressenti général sur l'expérience) ainsi que sur les notifications visuelles et auditives.

7.2.3. Hypothèses de travail

Au regard des points théoriques développés, nous proposons les hypothèses suivantes :

H1 : En fonction des designs de notifications visuelles, des différences significatives de performance dans la tâche de guidage ou la tâche auditive devraient apparaître.

H2 : On devrait observer une meilleure détection des sons rares dans la tâche auditive lors de la tâche de contrôle par rapport aux conditions avec le Laby.

H3 : La vitesse aura un effet sur la surdité inattentionnelle et sur la réalisation de la tâche de guidage et la tâche de validation des notifications.

H4 : L'encombrement visuel aura un effet sur la surdité inattentionnelle et sur la réalisation de la tâche de guidage et la tâche de validation des notifications.

H5 : L'amplitude de la P300 (effet de la workload) devrait être plus élevée avec la tâche de contrôle qu'avec Laby.

H6 : L'amplitude de la P300 (effet de la workload) pendant la tâche Laby devrait être plus élevée avec la notification BoxAnimation qu'avec ColorBlink.

7.2.4. Variables

Cette expérimentation comporte deux variables indépendantes supplémentaires par rapport à l'expérimentation Saillance : **(V) la vitesse** de déroulement du scénario (lente ou rapide), et **(C) la charge visuelle** (faible ou élevé). Les participants sont répartis en deux groupes en fonction de **(T) le type de notification** avec 2 modalités de design ColorBlink et BoxAnimation. Durant un scénario (ils en jouaient 4), un sujet va voir 34 notifications visuelles et 25 sons rares pour 100 sons standards.

Pour chacune des tâches nous nous sommes intéressés à deux variables dépendantes sur les mouvements de pointages le temps de validation (TVC : temps de validation consigne, TVN ; temps de validation notifications, TVS : temps de validation sons) et les loupés d'acquiescement (EC : erreur consigne, EN : erreur notification et ES : erreur son) des notifications ou des consignes par le sujet. Le temps de validation correspond au temps entre le moment où l'alarme ou la consigne va se déclencher et celui où le sujet va la valider ou rentrer la consigne correcte. Dans la suite des analyses, les loupés d'acquiescement seront exprimés en ratio sur le nombre total de consignes ou de notifications à valider.

Conformément à l'approche multifactorielle décrite dans le chapitre évaluation, les différentes mesures enregistrées lors d'une passation portent sur la tâche de guidage de l'avion dans le labyrinthe, les deux tâches de validation des notifications (visuelles et sonores), des mesures EEG de type potentiel évoqués et les reports subjectifs.

7.3. Résultats

Les résultats provenant de 6 participants sur les 42 qui ont passé l'expérimentation ont été supprimés en raison de données non exploitables (s'expliquant principalement pour des raisons matérielles ou d'acquisition). Nous avons deux groupes de 18 participants, dont les caractéristiques sont données ci-dessous :

Groupe ColorBlink : moyenne d'âge : 24 ans, écart-type : 5.4 ans ; 28% de femmes ;

Groupe BoxAnimation : moyenne d'âge : 25 ans ; écart-type : 6.3 ans ; 61% de femmes.

7.3.1. Réalisation de la tâche de guidage

Les sujets devaient saisir des consignes de niveau et de virage pour réaliser la tâche de guidage. Nous nous sommes intéressés à deux indicateurs de performance dans la réalisation de cette tâche, le pourcentage de consignes bien renseignées et le temps mis pour réaliser la saisie.

Sur le pourcentage de consignes bien renseignées, l'analyse ne montre pas d'effet du design des notifications en vision périphérique ni de la charge visuelle (nombre d'avions en périphérie affichés). Seul un effet principal significatif de la vitesse est observé $F(1,34) = 44.70, p < 0.001, \eta^2p = 0.57$.

Lorsque le tempo est plus rapide, le pourcentage de consignes bien renseignées diminue, passant de 96% à 85% (figure 57).

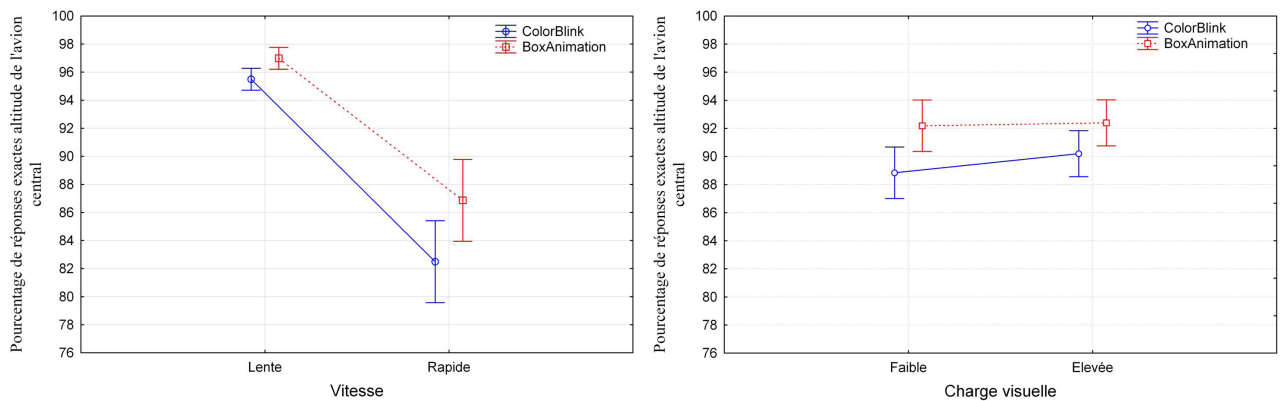


Figure 57 : effet de la vitesse et de la charge visuelle sur le % de consignes exactes saisies

Sur le temps de saisie des consignes, l'analyse (figure 58) encore une fois ne montre pas d'effet du design des notifications, mais un effet principal significatif de la vitesse observé $F(1,34) = 25.47, p < 0.001, \eta^2p = 0.43$ ainsi que de la charge visuelle $F(1,34) = 8.495, p=0.006, \eta^2p = 0.2$.

La pression temporelle augmentant avec la vitesse on peut observer que les sujets réalisent leurs saisies plus rapidement (gain de 11% soit 3.3s).

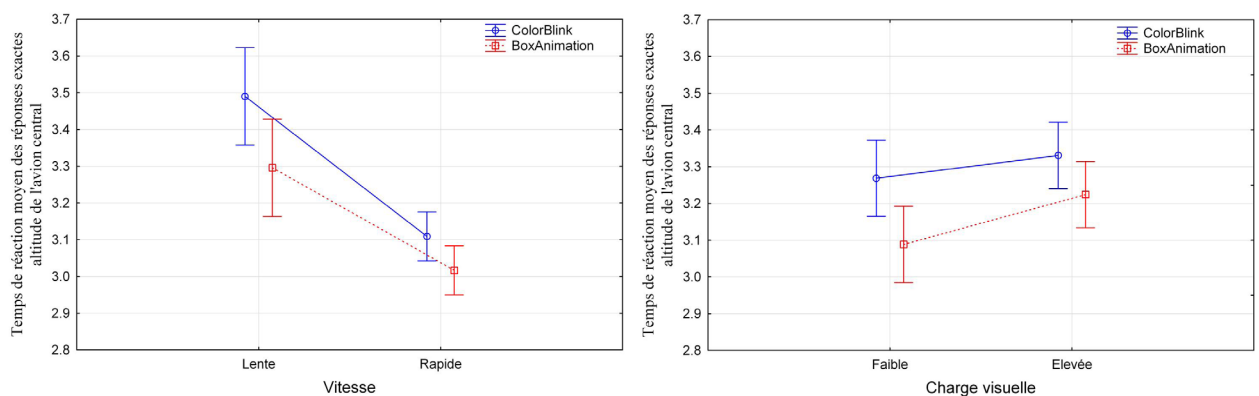


Figure 58 : effet de la vitesse et de la charge visuelle sur le temps de réaction pour saisir les consignes d'altitude

La charge visuelle liée à la présence d'un nombre plus important d'avions susceptibles de recevoir une notification en vision périphérique ralentit logiquement, mais assez légèrement les temps de réaction (0.1s).

7.3.2. Réalisation de la tâche de validation des notifications

Comme pour la tâche de guidage de l'avion central, nous nous intéressons aux indicateurs de pourcentage de notifications validées dans les temps et le temps de réaction pour réaliser la validation de la notification.

L'analyse statistique (figure 59) montre un effet principal significatif pour le type de design de notification (ColorBlink-CB ou BoxAnimation-BA) $F(1,34) = 20.14$, $p < 0.001$, $\eta^2p = 0.37$ ainsi que pour la vitesse $F(1,34) = 14.78$, $p < 0.001$, $\eta^2p = 0.3$ et la charge visuelle $F(1,34) = 7.74$, $p = 0.009$, $\eta^2p = 0.18$.

Lorsque les notifications sont du type ColorBlink 95.7% sont correctement validées, ce chiffre passe à 99.8% lorsqu'elles sont du type BoxAnimation. Dans l'expérimentation Saillance nous avons observé un taux de 100% pour BoxAnimation, nous formulons l'hypothèse que les quelques notifications BoxAnimation manquées sont dues à la pression temporelle (le temps pour valider est réduit dans le tempo rapide) et à l'augmentation de la complexité de la tâche avec l'introduction de la tâche auditive. Un capteur oculométrique aurait permis de vérifier si les notifications non validées ont bien été perçues en utilisant les fixations sur les objets.

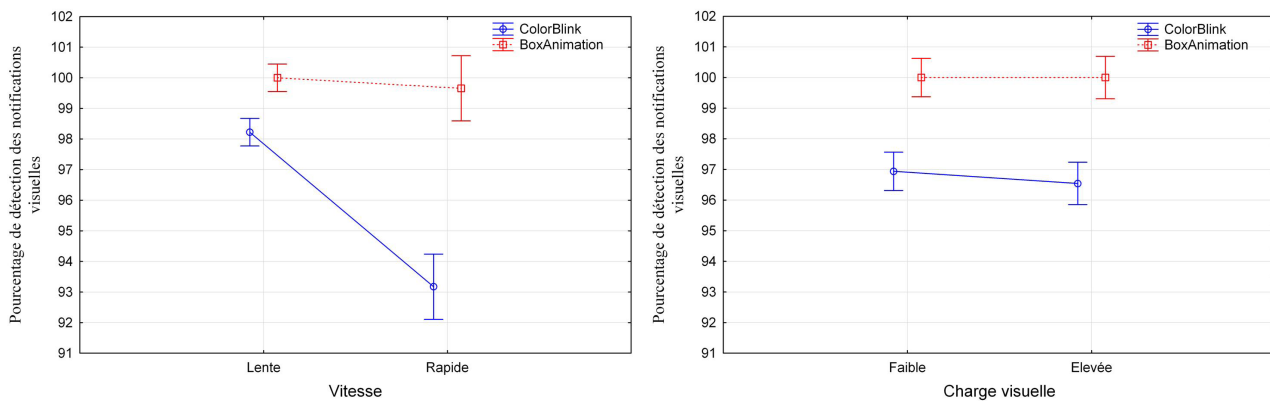


Figure 59 : effet du type de design sur le taux de notifications validées

Lorsque la vitesse augmente, le taux d'alarmes visuelles signalées diminue significativement, passant de 99.1% à 96.5%, ce qui correspond à une perte de 2.6%. Résultat que nous attendions dans l'hypothèse H3.

La charge visuelle (nombre d'avions susceptibles d'être notifiés) a un impact significatif sur le pourcentage de notifications validées (-1.5%). La charge visuelle avait pour objectif de complexifier le scanning visuel de l'image, ce résultat valide l'hypothèse H4.

L'analyse de comparaisons multiples (test HSD de Tukey) montre un effet significatif de la vitesse sur le pourcentage de notifications validées sur le design ColorBlink ($p < 0.001$), mais pas sur le design BoxAnimation ($p = 0.95$). La performance de validation du BoxAnimation ne se dégrade donc pas avec l'augmentation de la charge de travail. Le pourcentage de notifications ColorBlink validées se dégrade très nettement lorsque la charge de travail augmente avec la vitesse ce qui valide l'hypothèse H3.

Sur la charge visuelle, le test post-hoc montre un effet significatif sur le pourcentage de notifications validées avec le design ColorBlink ($p = 0.007$), mais pas sur le design BoxAnimation ($p = 0.96$). La performance de validation du BoxAnimation ne se dégrade donc pas avec l'augmentation de la charge

visuelle. Le pourcentage de notifications ColorBlink validées se dégrade très nettement lorsque la charge visuelle augmente ce qui valide l'hypothèse H4. La qualité de l'attracteur visuel BoxAnimation semble supprimer l'effet lié à la complexité de la tâche de scanning visuel nécessaire dans le cas du ColorBlink.

L'analyse des temps de réaction (figure 60) pour valider les notifications montre un effet principal significatif du type de design $F(1,34) = 102.218$, $p < 0.001$, $\eta^2p = 0.75$ et de la vitesse $F(1,34) = 11.12$, $p=0.002$, $\eta^2p = 0.24$. On observe aussi effet marginal de la charge visuelle $F(1,34) = 3.35$, $p=0.07$, $\eta^2p = 0.09$ (augmentation faiblement significative des temps de réaction lorsque la charge visuelle est élevée). Nous avons déjà observé le même effet du type de design dans l'expérimentation Saillance, ColorBlink était moins performant que BoxAnimation.

L'analyse post-hoc (test HSD de Tukey) montre un effet de la vitesse très intéressant puisque l'on peut constater que le design BoxAnimation provoque des temps de réaction comparables qu'elle soit lente ou rapide au contraire des temps avec le design ColorBlink qui s'améliorent significativement en fonction de la vitesse. Le résultat était attendu pour BoxAnimation (hypothèse H1,H3), dans la mesure où la charge de travail sur la tâche de guidage augmente avec la vitesse, le temps disponible pour le scanning visuel des notifications s'en trouve amoindri ; le design BoxAnimation étant un bien meilleur attracteur visuel, il permet au sujet de se reposer sur cette spécificité et limiter le scanning visuel. Dans le cas de ColorBlink, on peut émettre l'hypothèse que le sujet est plus engagé dans la tâche ce qui diminuerait les temps de validation, effet qu'il faut néanmoins nuancer puisque cette amélioration des temps de réaction s'accompagne d'une diminution importante du nombre de notifications détectées.

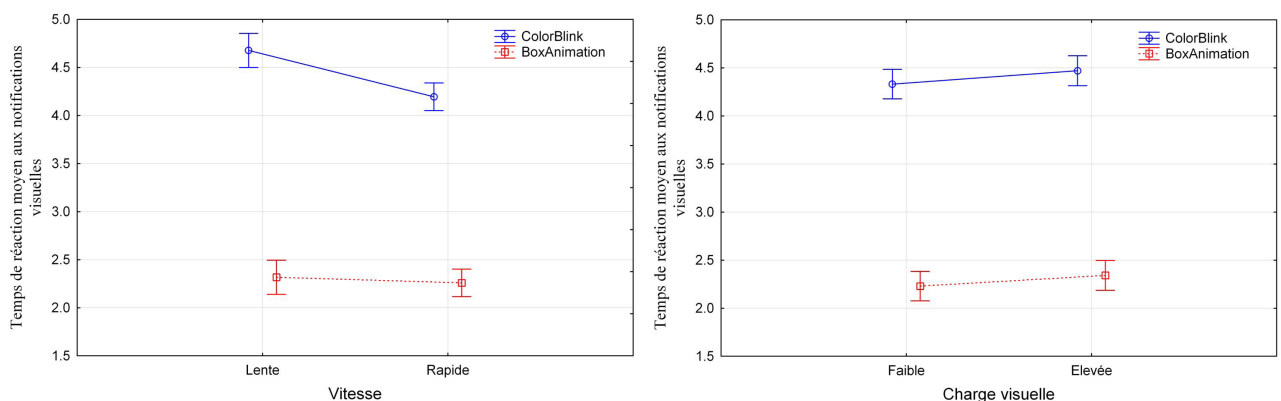


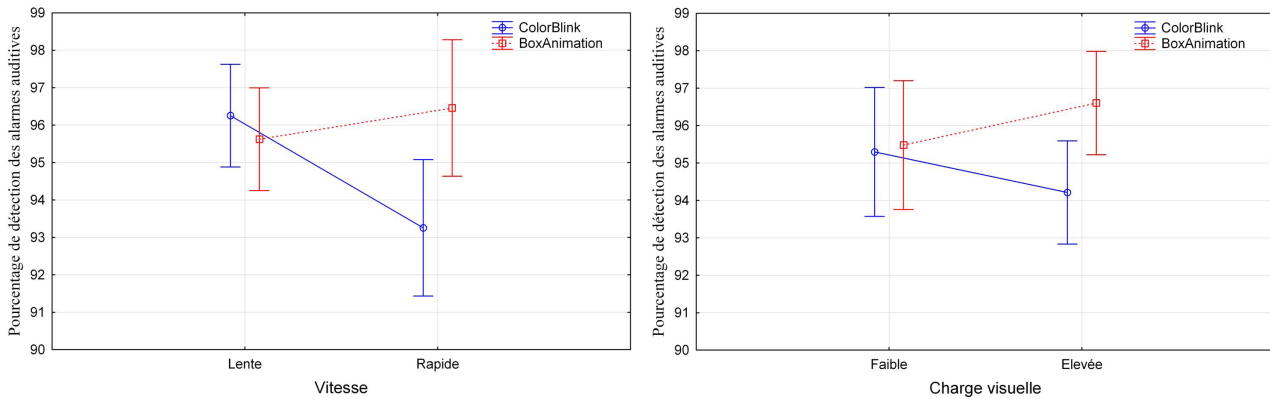
Figure 60 : effet de la vitesse et de la charge visuelle sur le temps de réaction pour valider les notifications

7.3.3. Réalisation de la tâche auditive

Comme pour les deux premières tâches nous nous intéressons plus particulièrement au pourcentage de notifications sonores correctement validées ainsi qu'au temps de réaction pour le faire.

La vitesse qui pouvait varier sur la tâche de contrôle ne provoquait pas d'effet significatif sur le taux de détection des sons rares ($p=0.291$), les deux conditions ont donc été fusionnées dans une seule. La comparaison a en-

suite été faite entre la tâche de contrôle fusionnée et Laby tous scénarios confondus. Dans la figure 61 est représenté le pourcentage de notifications sonores non validées en tâche de contrôle (entraînement à la tâche auditive) par rapport à la situation de triple tâche (Laby), l'analyse montre un effet significatif entre la tâche de contrôle et Laby $F(1, 35) = 17.16, p < 0.001, \eta^2p = 0.33$. 3.5% d'alarmes sonores ne sont pas validées dans la condition Laby. Les tests post-hoc ne montrent pas d'effet significatif de la charge visuelle ($p=0.984$) et de la vitesse ($p=0.351$) sur le pourcentage de notifications sonores détectées avec la tâche Laby.



.Figure 61 : effet de la vitesse et de la charge visuelle sur le pourcentage de détection des alarmes sonores

Les hypothèses H3 et H4 d'augmentation de la surdité inattentionnelle avec la vitesse (augmentation de la charge de travail) ou la charge visuelle ne sont donc pas validées pour la tâche Laby intra. L'hypothèse H2 est validée sur la comparaison tâche de contrôle/Laby.

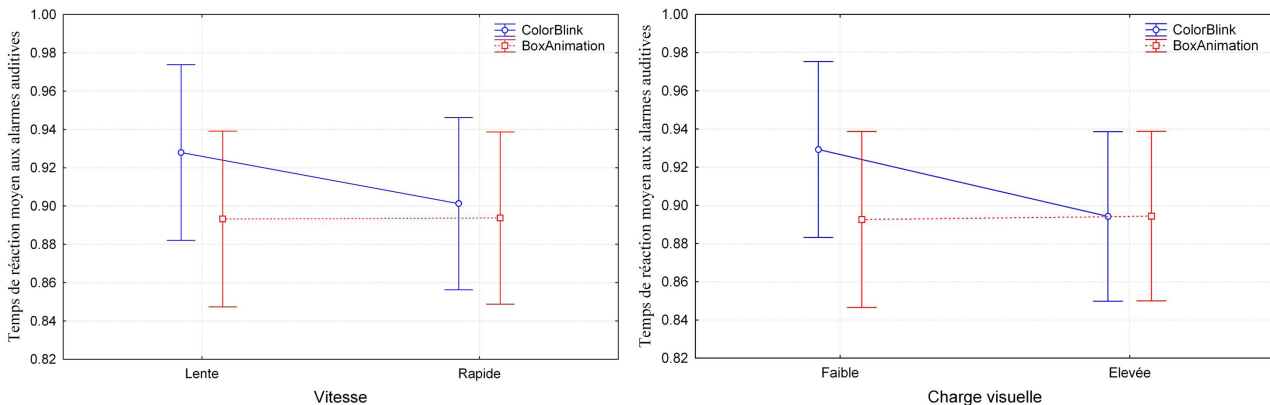


Figure 62 : effet de la vitesse et de la charge visuelle sur le temps de réaction moyen aux alarmes sonores

Sur les temps de réaction à valider les notifications sonores (figure 62), elles sont signalées en 0.90 s en vitesse lente et en 0.91 s en vitesse rapide dans la condition Laby. On n'observe aucun effet significatif du design, de la vitesse ou de la charge visuelle. Dans le cas de la tâche de contrôle qui avait aussi été jouée avec deux vitesses, une diminution du temps de réponse de 10% était observée dans la condition vitesse rapide. Nous formulons l'hypothèse que ce résultat s'explique en raison d'un engagement plus important des sujets dans la réalisation de la tâche lié à la pression temporelle.

7.3.4. Résultats physiologiques, analyse de la P300

L'objectif de ces mesures était d'observer l'effet de la charge de travail sur l'amplitude de la P300 dans les différentes conditions expérimentales. L'effet attendu est une diminution de l'amplitude de la P300 en condition de charge plus élevée (hypothèses H5,H6).

Les signaux EEG ont été segmentés en plage de temps autour du stimulus auditif (à partir de 200 ms avant le début du stimulus jusqu'à 1200 ms après). L'analyse de la P300 a été effectuée sur une fenêtre de 100ms centrée sur le pic moyenné de la tâche de contrôle à 414 ms. La vitesse qui pouvait varier sur la tâche de contrôle ne provoquait pas d'effet significatif sur l'amplitude de la P300 ($p=0.364$); les deux conditions ont donc été fusionnées dans une seule.

Nous avons ensuite comparé la tâche de contrôle fusionnée à la tâche Laby tous scénarios confondus. Un effet significatif sur l'amplitude de la P300 avec la tâche Laby a été observé $F(1, 34) = 14.29, p < .001, \eta^2p = .30$ (figure 63).

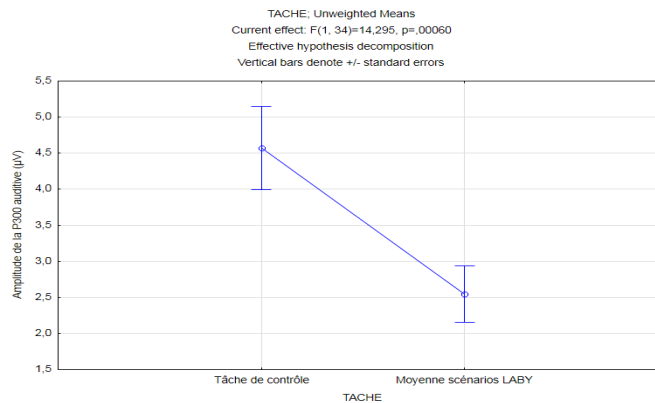


Figure 63 : amplitude moyenne de la P300 tous scénarios confondus tâche de contrôle/Laby

Un test post-hoc a ensuite été effectué pour chaque condition de la tâche Laby, on observe une interaction significative entre le type de sons et le design. Il n'y a aucun effet significatif de la vitesse ($p=0.274$) et de la charge visuelle ($p=0,789$) sur l'amplitude de la P300. Le test montre que BoxAnimation élicite une meilleure P300 par rapport à ColorBlink. Ce résultat suggère une libération des ressources attentionnelles provoquée par un meilleur design visuel (théorie de la charge attentionnelle étendue de Lavie).

La figure suivante (Figure 64 : amplitude de la P300 (μV) pour les deux types de design) montre l'impact du design sur l'amplitude de la P300, on constate une augmentation significative de la P300 dans le cas BoxAnimation, $F(1,34)=4.26 p=0.046$.

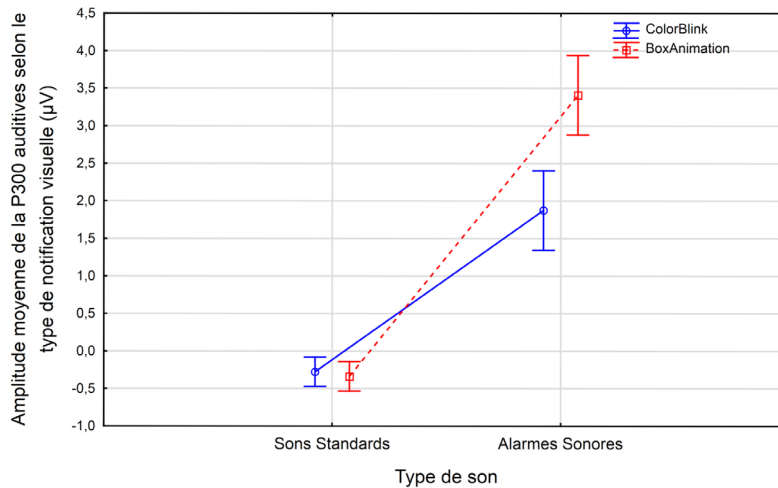


Figure 64 : amplitude de la P300 (µV) pour les deux types de design

7.3.5. Résultats qualitatifs

Nous avons analysé les scores NASA-TLX pour les dimensions suivantes : exigence mentale, exigence temporelle, performance et effort. Un effet significatif de la vitesse a été observé sur les quatre dimensions ($p < 0.001$). Un effet significatif de la charge visuelle n'a été observé que sur la dimension exigence mentale ($p=0.04$). Sur l'exigence mentale, Il n'y a pas d'effet principal du groupe ($p=0.18$), et aucune interaction avec la vitesse ($p=0.13$) et la charge visuelle ($p=0.51$).

L'exigence mentale et temporelle, la performance et l'effort ont été perçus plus élevés dans la condition vitesse rapide confirmant les résultats quantitatifs et le choix de l'utilisation de la vitesse dans le protocole expérimental pour augmenter la charge de travail.

La figure suivante (Figure 65) montre l'effet de la vitesse et de la charge visuelle sur la perception de la charge de travail pour les deux types de design.

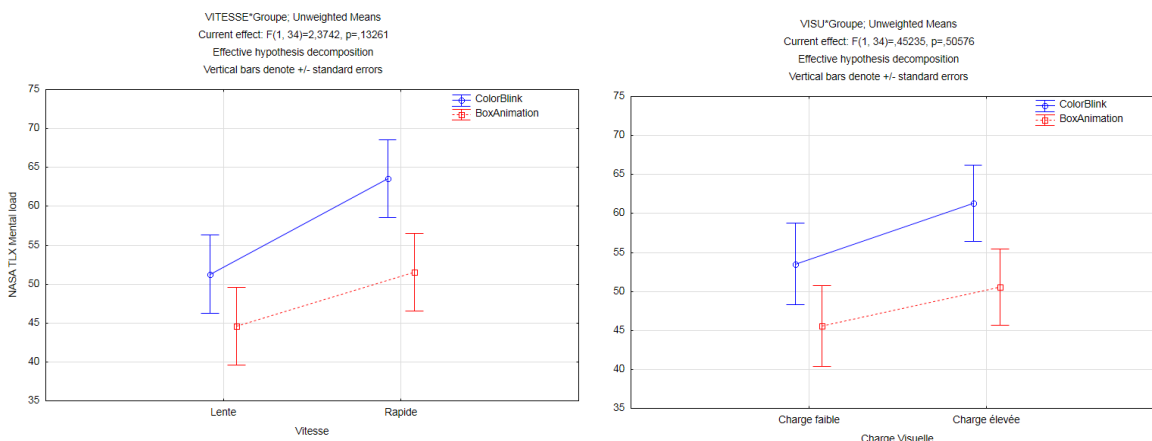


Figure 65 : index « Mental Load » du NASA TLX en fonction de la vitesse et de la charge visuelle.

Un questionnaire spécifique (voir annexes) sur les notifications visuelles a été soumis aux participants. L'analyse des réponses (Figure 66: questionnaire notifications visuelles

) a fait émerger des différences significatives entre les deux designs que sur deux questions : « Je remarquais immédiatement les alertes dès leur apparition » et « J'étais obligé de faire attention aux avions périphériques pour voir ces alertes ». Pour cela, nous avons attribué aux différentes possibilités de réponses (« Presque jamais », « Parfois », « Souvent », « Presque Toujours ») respectivement les poids 1, 2, 3 et 4, et comparer la moyenne de ces réponses entre les deux groupes. 2 participants du groupe BoxAnimation ont dû être retirés des analyses pour absence de réponse à cette question. Nous observons un effet significatif du groupe dans les deux cas ($F(1,32) = 11.92$, $p = 0.002$, $\eta^2p = 0.27$ et $F(1,32) = 17.53$, $p < 0.001$, $\eta^2p = 0.35$). Les résultats confirment les données comportementales, les sujets ont tendance à voir plus tôt les notifications dans le groupe BoxAnimation, et estiment avoir moins besoin d'y prêter attention pour les voir.

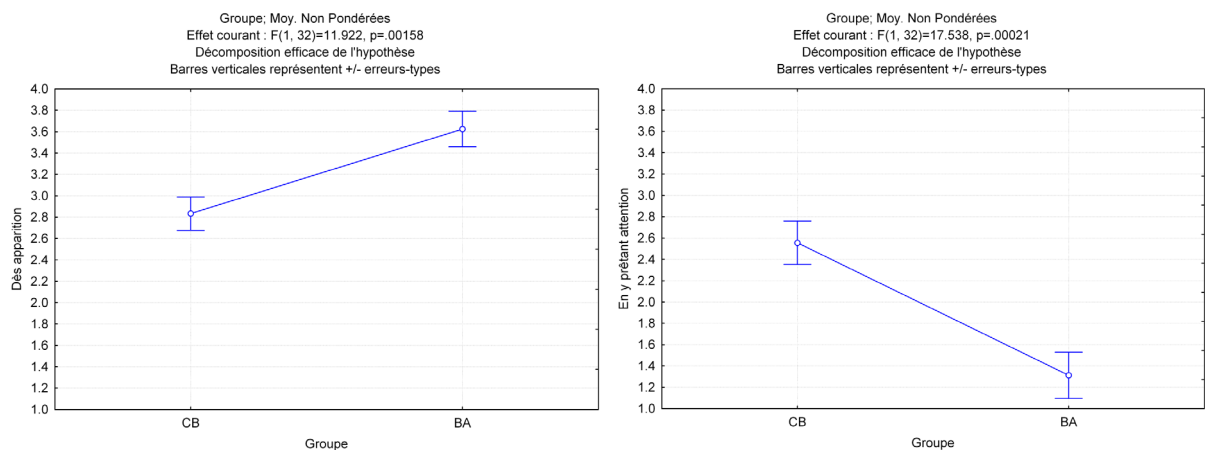


Figure 66: questionnaire notifications visuelles

7.4. Conclusion

7.4.1. Analyse des résultats

Tout d'abord, les résultats comportementaux et physiologiques ont confirmé que les participants qui utilisaient Laby avec le design BoxAnimation ont mobilisé moins de ressources pour valider les notifications. Par ailleurs, ils ont mis moins de temps à les valider et le taux de validation est plus élevé que dans le cas du groupe ColorBlink. BoxAnimation s'est en outre révélé plus efficace lorsque la difficulté de la tâche ATC (vitesse et charge visuelle) augmentait. Le pourcentage de validation des notifications n'a pas été affecté par l'augmentation de la difficulté avec BoxAnimation alors qu'il s'est dégradé dans le cas du groupe ColorBlink. Du point de vue subjectif, les résultats viennent confirmer l'objectif des expériences : un sentiment de débordement lorsque l'on augmente les charges visuo-perceptives et de travail. Dans les faits, tel que le protocole était réalisé, nous n'avons malheureusement pas mis en évidence l'effet de la charge visuelle sur les trois tâches évaluées dans les mesures de performances.

Cette expérience a mis en évidence un lien entre la charge de travail et la surdité inattentionnelle, lorsque la charge augmente significativement

plus de notifications sonores peuvent être manquées. Aucun lien n'a pu être établi entre la charge visuelle et la surdité inattentionnelle.

En gras les hypothèses de travail validées :

H1 : En fonction des designs de notifications visuelles, des différences significatives de performance dans la tâche de guidage ou la tâche auditive ont été mises en évidence.

H2 : Nous avons observé une meilleure détection des sons rares dans la tâche auditive lors de la tâche de contrôle par rapport aux conditions avec le Laby.

H3 : La vitesse a eu un effet sur la surdité inattentionnelle et sur la réalisation des tâches de guidage et de validation des notifications.

H4 : L'encombrement visuel n'a eu d'effet sur la surdité inattentionnelle et sur les tâches de guidage et de validation des notifications.

H5 : L'amplitude de la P300 (qui doit diminuer lorsque la tâche se complexifie) est plus élevée avec la tâche de contrôle qu'avec Laby.

H6 : L'amplitude de la P300 (effet de la workload) pendant la tâche Laby s'est révélée significativement plus élevée avec la notification BoxAnimation qu'avec ColorBlink.

Le principal résultat quantitatif obtenu grâce à cette expérience concerne l'analyse de l'amplitude de la P300 et l'impact du type de design des notifications sur celle-ci. On observe un effet significatif sur l'amplitude de la P300 qui est un indicateur de charge de travail. En d'autres mots un design efficace en termes de saillance va libérer des ressources attentionnelles pour réaliser d'autres tâches, ce résultat est confirmé par les mesures comportementales, physiologiques et par les réponses aux questionnaires des participants.

7.4.2. Discussion et applications

À notre connaissance, très peu d'études ont été réalisées dans le domaine de l'IHM pour faire des choix de conception à l'aide de mesures comportementales et physiologiques basées sur des analyses ERP. Nos résultats corroborent d'autres travaux dans la littérature sur l'utilisation des mesures ERP et l'analyse de l'amplitude de la P300 comme indicateur de la charge de travail (Ullsperger et al., 2001 ; Allison et al., 2008). Ils confirment la nécessité de valider des choix de designs des notifications, aussi bien visuelles que sonores à l'aide de mesures comportementales et physiologiques pour évaluer de façon complète leur impact sur l'humain.

Dans le cockpit, la problématique des alarmes sonores manquées est déjà étudiée depuis plusieurs années (Dehais et al., 2013; Bliss, 2003). Celles-ci sont peu utilisées aujourd'hui en France (uniquement en approche dans les grands aéroports), mais cela n'est pas le cas dans d'autres pays et ces résultats sont tout à fait intéressants pour la communauté du contrôle aérien. Le développement de l'emploi des notifications sonores (Cabrera, Ferguson, & Laing, 2005, 2006) posera à terme le même type de problèmes que dans le cockpit. Leur emploi peut provoquer du stress, un effet cry-wolf (Edworthy, Loxley, & Dennis, 1991) ou bien ne pas être perçues, phénomène décrit par Koreiman et al. (2009) comme de la surdité inattentionnelle. Leur mise

en œuvre au sein d'une supervision, couplée à des notifications visuelles doit donc être évalué dans un cadre expérimental contrôlé. L'utilisation d'un micro-monde ATC suffisamment réaliste pour être considéré comme écologique, permet de capitaliser sur les résultats obtenus pour aider à la conception des futurs systèmes de supervision.

Cette étude montre que le système perceptif humain doit être considéré dans sa globalité, le couplage entre le visuel et l'auditif est très fort. L'approche multifactorielle montre son intérêt comme dans le cas de l'expérimentation Saillance et la mesure de l'activité cérébrale permet d'obtenir des informations pertinentes pour l'évaluation de la charge de travail. Elle peut aider à établir un compromis adapté à la tâche dans le design des notifications visuelles ou sonores. Ce compromis entre le pouvoir attentionnel, la vitesse de détection de la notification et l'impact sur la tâche principale, s'enrichit de l'évaluation de l'impact du design sur les ressources attentionnelles mises en œuvre. L'utilisation d'un oculomètre associé à un EEG permet en plus des mesures de comportementales d'atteindre le meilleur compromis pour la notification d'informations qu'elles soient sonores ou visuelles vers les superviseurs.

8. ÉVALUATION D'UNE VISUALISATION ORIENTÉE CONSCIENCE DE LA SITUATION : LES EXPÉRIMENTATIONS MONITORING & WP4.7.2

8.1. Présentation

Le deuxième volet de ces travaux de thèse a pour but de concevoir et mettre en œuvre une nouvelle position de contrôle aérien dont le design vise à améliorer la conscience de la situation. Pour atteindre cet objectif, nous avons déduit les informations dynamiques les plus intéressantes que les contrôleurs doivent détecter pendant leur activité et réaliser un agent de suivi de la tâche : SAmonitoring. Cet agent, lorsqu'il trouve des informations pertinentes pour le contrôleur, émet un message vers la supervision radar décrivant le problème et son niveau de criticité. La supervision radar a été modifiée pour afficher ces notifications en choisissant un design dont la saillance dépend de la criticité de l'information. Ces designs ont été évalués lors des expérimentations Saillance et Bip. La nouvelle position de contrôle aérien intègre SAmonitoring et le système de notification.

Ce chapitre a pour but de présenter les résultats préliminaires obtenus lors de l'évaluation de cette position de contrôle aérien lors de simulations. L'objectif de ces expérimentations est de s'assurer de son utilité pour améliorer la conscience de la situation des contrôleurs en particulier lors des situations critiques pour la sécurité.

Une simulation réaliste nécessite la mise en œuvre d'une position de contrôle comportant un contrôleur radariste et un contrôleur organique, un contrôleur gérant les secteurs adjacents au secteur contrôlé, un pseudo-pilote pour interagir avec le contrôleur radariste ainsi que le personnel technique gérant la plateforme. Par ailleurs, dans le cadre du processus d'évaluation à l'aide de questionnaires ou d'interviews ou bien de la formation des contrôleurs, des psychologues et des contrôleurs experts sont en général impliqués durant les simulations. Ce type d'expérimentation est très coûteux en moyens humains et matériels.

Le type de problèmes détectés par SAmonitoring sont de différents niveaux (warning, alerte, alerte critique). Nous souhaitons nous assurer à l'aide de cette expérimentation de son apport sur les alertes critiques qui sont les plus importantes pour la sécurité aérienne. Ce genre de problème est évidemment exceptionnel et ne peut pas être rencontré de façon systématique lors de simulations mêmes de plusieurs heures sans actions délibérées des personnes impliquées dans le déroulement de la simulation (pseudo-pilote(s) ou contrôleur de secteur adjacent). Nous avons donc réalisé deux expérimentations. Dans la première, l'expérimentation **Monitoring**, nous provoquons en temps réel lors des simulations des situations dans lesquelles le SAmonitoring détecterait des cas d'alertes critiques. Dans la seconde, l'expérimentation **WP4.7.2** nous nous sommes placés dans un cadre écologique par rapport à la tâche et avons réalisé des simulations sans provoquer volontairement des problèmes aboutissant à des alertes critiques.

L'expérimentation Monitoring s'est insérée dans le planning d'une expérimentation plus importante se déroulant en plusieurs sessions de deux jours qui concernait le work package WP4.7.2 du projet SESAR. Ce projet a pour but de réduire le nombre de conflits à gérer par les contrôleurs à l'aide d'un automate agissant sur la vitesse des avions en amont des conflits. Une nouvelle position de travail pour les contrôleurs devait être créée pour ce projet de façon à offrir des outils évolués permettant l'interaction avec le solveur de conflits. C'est dans ce cadre que la position Droïd a été créée afin d'apporter de nouveaux services et outils aux contrôleurs. La position Droïd est composée d'un écran vertical de 30" affichant la supervision radar et d'un écran horizontal Wacom de 21" servant à afficher des informations plan de vol et à interagir avec la supervision radar affichant (voir section 5.5.2 pour plus de détails). L'équipe projet SESAR WP4.7.2, ayant décidé d'intégrer SAmonitoring et le plugin de notification de la supervision radar lors de leurs expérimentations de validation, nous avons pu bénéficier de ces simulations pour évaluer leur impact dans un contexte écologique.

Dans le cas de l'expérimentation Monitoring, nous devons provoquer le déclenchement de cas d'alertes critiques. Le déroulement d'une simulation n'étant pas déterministe, chaque contrôleur ayant la possibilité d'agir différemment lors des situations conflictuelles, il n'était pas possible de préparer à l'avance des situations dont nous aurions la certitude qu'elles se produiraient durant les simulations. Nous avons donc déterminé à l'aide de contrôleurs experts (SME) plusieurs types de scénarios (une vingtaine) pouvant déclencher des notifications critiques, certaines portant sur des avions impliqués dans des conflits pendant la traversée du secteur, d'autres portant sur des avions n'ayant aucune interaction avec les autres avions présents dans le secteur ; tous étant des cas parfaitement réalistes pour des contrôleurs expérimentés. Lors d'une simulation, un contrôleur expert choisissait en temps réel de déclencher les scénarios en fonction des actions du contrôleur radariste de façon à répartir les dix alertes critiques sur la durée de la simulation. L'expérimentation Monitoring a permis de réaliser 12 simulations de 40 minutes pendant lesquelles nous provoquons 10 alertes critiques. Un oculomètre posé a été mis en œuvre pour compléter les mesures comportementales réalisées sur la supervision radar (voir section 5.5.1).

Lors de l'expérimentation WP4.7.2, nous avons pu collecter 39h de données comportementales, 18h30 avec le système de notification, 10h30 sans. L'oculomètre ne pouvait être mis en œuvre pendant ces simulations. Les cas d'alertes critiques détectés n'ont pas été provoqués intentionnellement. Ils ont été détectés par SAmonitoring soit en raison d'erreurs de saisie du contrôleur soit en cas de mauvaise exécution des ordres du contrôleur par le pseudo-pilote. Ces situations sont donc tout à fait réalistes pour le contrôleur.

Dans la suite de ce chapitre nous présenterons les résultats d'analyse de ces deux campagnes d'enregistrement (expérimentation **Monitoring** et expérimentation **WP4.7.2**).

8.2. Protocole

8.2.1. Participants

12 contrôleurs en route provenant de plusieurs centres de contrôle aérien en Europe ont participé à ces expérimentations dont 8 Français, 2 venant de Skyguide en Suisse et 2 venant du centre de contrôle Européen de Maastricht (MUAC). Ils étaient tous qualifiés premier contrôleur dans leurs centres de contrôle respectifs. L'âge moyen des participants était de 35 ans (std : 8.5 ans) pour une expérience moyenne de 8.5 années et demi (std 6 ans). Tous les participants ont été formés à l'utilisation de Droïd et participé à 4 demi-journées de simulations d'entraînement durant lesquelles ils réalisaient des scénarios de 2h. Les contrôleurs ont aussi été formés à la gestion du secteur générique qui a été créé pour les expérimentations WP4.7.2 ; il dispose d'un réseau de route spécifique et des règles de transfert pour les secteurs adjacents (Figure 67 : secteur générique). Dans ce secteur toutes les routes le traversant (en rouge les routes Nord/Sud, en vert les routes Sud/Nord) sont rectilignes ; plusieurs points de conflits sont présents (croisement de routes). Les exercices joués dans les simulations ont été calibrés pour présenter un nombre de conflits important sur ces points de croisement.

Tous les contrôleurs ont été formés au fonctionnement du SAmonitoring ainsi qu'à l'utilisation du plugin Notification présentant les alertes et la liste de cas détectés. SAmonitoring et son plugin graphique ont été utilisés sur une partie des sessions d'expérimentation du WP4.7.2 par tous les participants avant de réaliser l'expérimentation Monitoring. Compte tenu des objectifs de SESAR d'augmentation de capacité, les scénarios joués pendant les simulations sont exceptionnellement chargés et réclamaient un engagement important dans la tâche.

Plusieurs conditions ont été mises en œuvres durant les expérimentations du WP4.7.2, l'oculométrie a été utilisée pour évaluer l'impact de ces conditions sur la charge de travail à l'aide de mesures de diamètres pupillaires (Martin C. et Imbert JP 2012). Nous avons pu nous insérer dans le planning de ces expérimentations en obtenant un créneau d'une heure avec chaque binôme de contrôleurs pour réaliser l'expérimentation monitoring. Lors des simulations WP4.7.2, nous ne pouvions pas connecter notre oculomètre à la visualisation radar. Nous ne disposons donc pas des données oculométriques de fixation sur les notifications corrélées avec les données de la visualisation radar et du SAmonitoring.



Figure 67 : secteur générique

8.2.2. Hypothèses de travail

H1 : SAmonitoring et son système de notification améliorent la détection des situations critiques que l'avion

H2 : lorsque les avions sont en conflit, ils sont mieux surveillés par le contrôleur

H3 : SAmonitoring et son système de notification diminuent le temps mis par le contrôleur pour agir sur le vol

H4 : SAmonitoring et son système de notification permettent de diminuer la durée des situations critiques (temps entre la détection du problème par l'agent et l'action correctrice du contrôleur le faisant disparaître).

8.2.3. Variables

Les contrôleurs aériens, lorsqu'ils ont détecté un conflit potentiel, surveillent en théorie plus attentivement l'évolution des avions. Nous avons décidé d'évaluer l'apport du monitoring dans les situations critiques impliquant ou pas une situation de conflit avec un autre aéronef.

Quatre conditions expérimentales ont été évaluées . Le SAmonitoring fonctionnait dans tous les cas, les notifications pouvaient être ou pas affichées sur la visualisation radar (groupes ON et OFF), les avions sur lesquels portait la notification pouvaient être impliqués ou pas dans un conflit. Facteur Conflit (0 | 1).

Nous avons enregistré et analysé trois variables :

Durée avant fixation : lorsque l'oculomètre est présent, cela correspond à l'écart de temps entre la détection du problème par SAmonitoring et la première fixation sur le vol incriminé.

Durée avant entrée label : cela correspond à l'écart de temps entre la détection du problème et la première action sur le vol incriminé (curseur

souris dans l'étiquette radar, consigne saisie). La visualisation radar a été modifiée pour sortir ces informations dans les logs.

Durée du problème détecté : il s'agit de la durée du problème du point de vue du SAmonitoring c'est-à-dire le temps mis par le contrôleur pour détecter et corriger le problème soit en renseignant sa supervision radar, soit en donnant un ordre au pseudo-pilote, soit lorsque l'avion quitte le secteur contrôlé. Lorsque le problème est ignoré pendant toute la durée de traversée du secteur, cette durée correspond au temps entre le début du problème et l'heure de sortie de secteur.

8.2.4. Procédure de passation expérimentation Monitoring

Les 12 contrôleurs ont été répartis en deux groupes de 6 : le groupe « ON » disposait de nos notifications sur la supervision radar (ON) ; le groupe « OFF » en était dépourvu. Dans tous les cas, l'agent SAmonitoring fonctionnait et permettait d'enregistrer les cas détectés. Pour chaque groupe nous avons essayé de provoquer 10 cas d'alertes critiques dont 5 portaient sur des avions en conflit avec un autre appareil et 5 sur des avions n'ayant aucun conflit avec un autre appareil.

Après la calibration du Facelab, un briefing contrôleur sur la simulation était effectué par un SME. Les contrôleurs n'étaient pas renseignés sur l'objectif exact de l'expérimentation, mais à l'issue du briefing le SME évoquait le fait que la simulation serait très difficile, qu'ils ne seraient pas jugés sur leur performance, mais que ne nous nous intéressions à des aspects IHM très fins ce qui justifiait la présence de l'oculomètre.

Dans le cas du groupe ON, l'objet du système de notification a été rappelé en insistant sur l'importance d'agir pour corriger rapidement les problèmes remontés par le système. Nous leur avons demandé a minima de faire disparaître immédiatement les alertes critiques (en mettant le curseur sur l'étiquette) et dans la mesure du possible entreprendre rapidement l'action correctrice pour faire disparaître le problème.

Pour ceux ne disposant pas de l'aide, il leur était demandé de s'assurer que tous les états de vols et coordination ainsi que toutes les clairances données soient bien entrées dans le système et bien réalisées par le pseudo-pilote pour assurer un fonctionnement correct de la simulation et des systèmes d'aides.

À l'issue du briefing, les contrôleurs devaient réaliser un scénario de quarante minutes pendant lesquelles nous provoquions des incidents concernant des avions à partir d'une liste de situations prévues par le SME pour des avions en conflits ou pas. Nous avons provoqué les alertes critiques avec deux types de situation, des avions ne conservant pas leur niveau de vol ou dépassant leur clairance de niveau, et des déviations de route à l'opposé de la direction donnée par le contrôleur.

Parmi les situations prévues, certaines étaient faisables systématiquement. Par exemple, un avion non impliqué dans un conflit demandait une déviation de 10° à gauche pour éviter un orage et partait 10° droite après l'autorisation du contrôleur. Dans un autre cas, un avion se mettait à descendre très lentement sans autorisation préalable du contrôleur, si celui-ci s'en rendait compte et contactait le pilote il lui était répondu que l'avion avait un pro-

blème de pilote automatique. Dans un dernier cas, un avion se mettait en descente très rapide sans autorisation préalable, puis finissait par contacter le contrôleur pour déclarer un problème grave de dépressurisation cabine. Dans l'essentiel des cas provoqués, lorsque les contrôleurs ordonnaient des directes sur des points de leur route, le pilote partait en légère déviation (environ 5°) à l'opposé du cap nécessaire pour réaliser la directe réclamée.

Le contrôleur expert décidait dans quel cas le pseudo-pilote devait suivre précisément l'ordre du contrôleur ou faire le contraire pour provoquer une notification de façon à repartir les alertes critiques tout au long de la passation (40 mn). Il disposait d'une liste précisant si les avions étaient en conflit ou pas ce qui lui permettait d'équilibrer les cas provoqués dans chaque catégorie (5 en conflit avec un autre avion, 5 pas impliqués dans un conflit).

À l'issue de la simulation, un questionnaire ad hoc était rempli par les contrôleurs. Deux questionnaires avec ou sans monitoring ont été conçus à l'aide de contrôleurs experts pour obtenir leur auto-évaluation sur des aspects contrôle et conscience de la situation. Suivait un débriefing expliquant l'objectif de l'expérimentation et une interview sur leur ressenti par rapport à un système de ce type. Au total une passation durait 1h10.

8.2.5. Procédure de passation expérimentation WP4.7.2

Le SAmonitoring a été utilisé dans plusieurs sessions de simulations du projet WP4.7.2. Dans ces simulations, aucune situation n'était provoquée de façon intentionnelle. Les cas détectés sont donc issus de problèmes liés soit au non-renseignement de l'interface de supervision avec une consigne que le contrôleur a donné à la voix au pseudo-pilote, soit à une erreur de saisie ou une mauvaise compréhension de la consigne par le pseudo-pilote.

12 contrôleurs sont passés sur ces expérimentations. Ils ont réalisé 19 sessions de 1h30 avec le SAmonitoring et les notifications (groupe ON) et 7 sessions de 1h30 avec le SAmonitoring sans les notifications (groupe OFF). Pour des problèmes de temps de mise en œuvre, l'oculomètre n'a malheureusement pas été utilisé dans ces sessions. Lors de ces expérimentations, nous n'avons enregistré que le temps mis pour réaliser la première interaction sur l'étiquette radar de l'avion notifié et la durée du problème critique du point de vue du SAmonitoring.

Aucun briefing spécifique au système notification n'a été réalisé lors de ces expérimentations. Les contrôleurs ont simplement été formés à son utilisation durant les phases d'entraînement.

8.3. Résultats

8.3.1. Résultats quantitatifs

8.3.1.1. Expérimentation Monitoring

Pendant les expérimentations, nous avons provoqué 10 alertes critiques sur la durée du scénario pour chaque contrôleur. Dans deux cas, nos actions en temps réel durant l'expérimentation n'ont pas provoqué la détection d'une alerte critique pour le groupe OFF. Les alertes et warning ont été détectés par l'agent SAmonitoring sans que nous en soyons à l'origine. Le nombre important de warning n'est pas significatif d'un mauvais déroulement de la tâche. Le contrôleur adjacent devait en parallèle de sa tâche aider le pseudo-pilote à provoquer les notifications critiques. Il n'était donc pas en mesure de gérer correctement les états fréquences en entrée et sortie du secteur contrôlé ce qui provoquait l'apparition de warnings sur la position contrôlée.

Nos analyses ne portent que sur les alertes critiques puisque nous n'avons pas essayé de provoquer intentionnellement les deux autres types d'alerte. Le tableau suivant présente les comptages pour les 3 types d'alertes :

Condition\Statut	Warning	Alerte	Alerte critique
OFF	59	17	58
ON	80	27	60

Les données oculométriques se sont révélées peu fiables, sur les 58 alertes critiques du groupe OFF seules 32 fixations valides ont été observées avant une action délibérée sur l'étiquette radar. Pour le groupe ON, 33 fixations valides ont été mesurées.

La durée moyenne avant fixation est de 7.22 s (std 8.13) pour le groupe ON contre 22.55s (std 22.66) pour le groupe OFF.

Dans les cas de notifications critiques, nous avons identifié les avions en conflit ou non et formulé l'hypothèse (**H2**) que les avions en conflits seraient surveillés plus fréquemment limitant dans ce cas le bénéfice du SAmonitoring. Cette hypothèse n'a pas été validée, il n'y a pas d'effet du facteur conflit sur le temps mis à percevoir un problème critique $F(1,62)=1.53$ $p=0.22$.

Par contre, l'analyse (Figure 68 : durée avant la première fixation sur l'avion notifié) montre un effet très significatif du groupe sur le temps mis à percevoir un problème critique $F(1,62)=14.03$, $p<0.001$ validant l'hypothèse **H1**. Les contrôleurs disposant des notifications issues du SAmonitoring les ont perçus nettement plus rapidement que ceux qui devaient réaliser leur tâche de supervision.

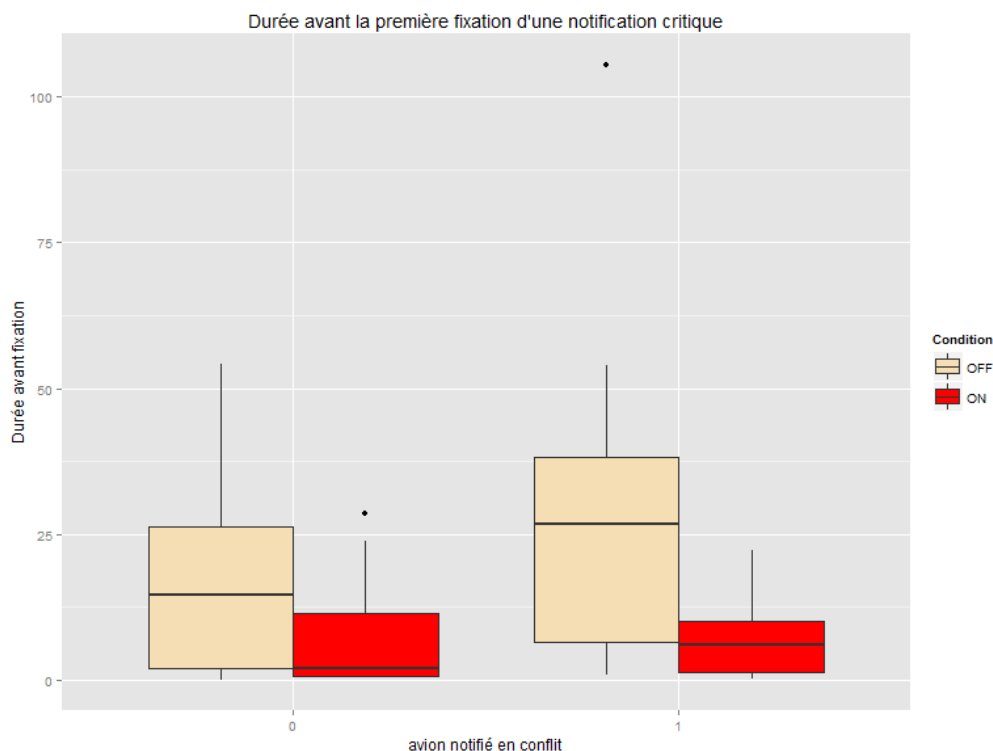


Figure 68 : durée avant la première fixation sur l’avion notifié

Pour le groupe OFF, le contrôleur interagit avec l’étiquette de l’avion cible d’un événement du SAmonitoring dans 45 cas sur 58. Pour le groupe ON dans 56 cas sur 60. L’analyse statistique démontre que le groupe ON a significativement plus agit sur les étiquettes liés à des alertes critiques que le groupe OFF ($p=.01$).

La durée moyenne avant une action du contrôleur sur l’étiquette radar des avions notifiés est de 19s (std 24.5) pour le groupe ON et 97.8s (std 115.6s) . L’analyse de cette durée (Figure 69 : durée avant la première interaction sur l’avion notifié) donne des résultats comparables à celle des temps de fixation, la variable conflit n’est pas significative $F(1,98)=0.204$ $p=0.653$; la condition (ON/OFF) est très significative $F(1,98)=24.8$ $p<0.001$. Les sujets du groupe disposant du SA monitoring ont donc significativement détecté et agi sur l’étiquette des avions sur lesquels un problème était provoqué bien plus rapidement que le groupe OFF ce qui valide l’hypothèse **H3**.

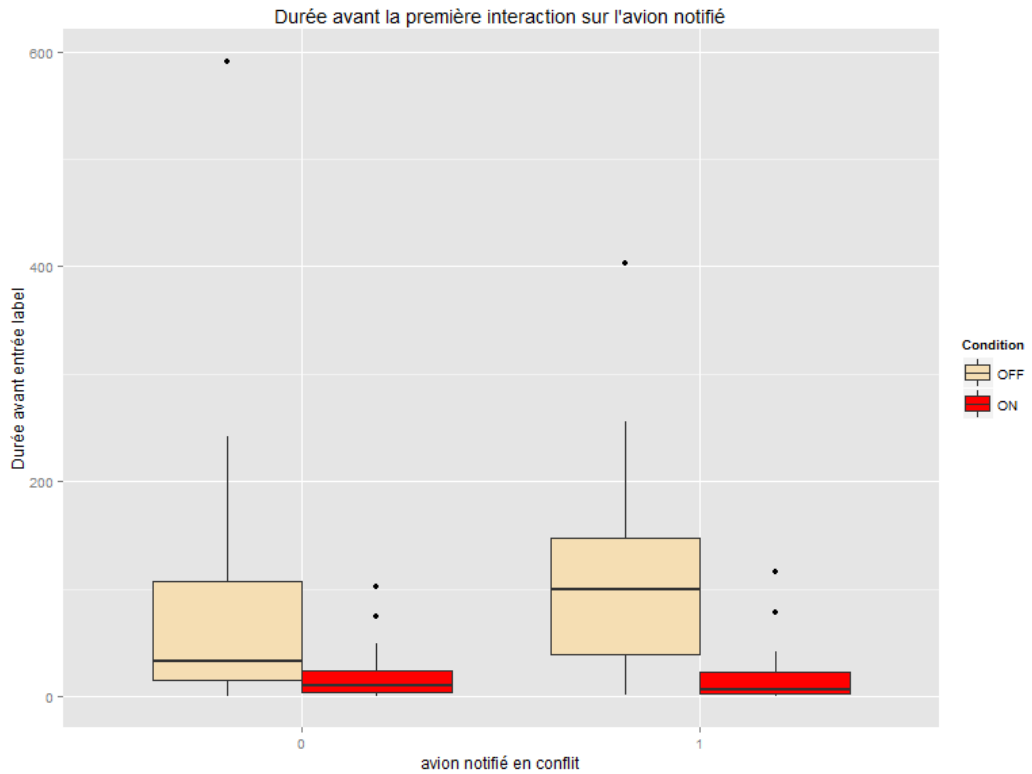


Figure 69 : durée avant la première interaction sur l'avion notifié

Nous avons réalisé un test de corrélation de Pearson entre les durées avec la première fixation et les durées avant l'entrée dans le label des avions en alerte critique (Figure 70 : durée avant fixation =f(durée avant entrée label)), les deux variables sont très corrélées ($r(58)=.60$, $p<.001$). Ce qui montre qu'une fois l'alerte perçue, les contrôleurs ont bien interagi avec l'étiquette radar.

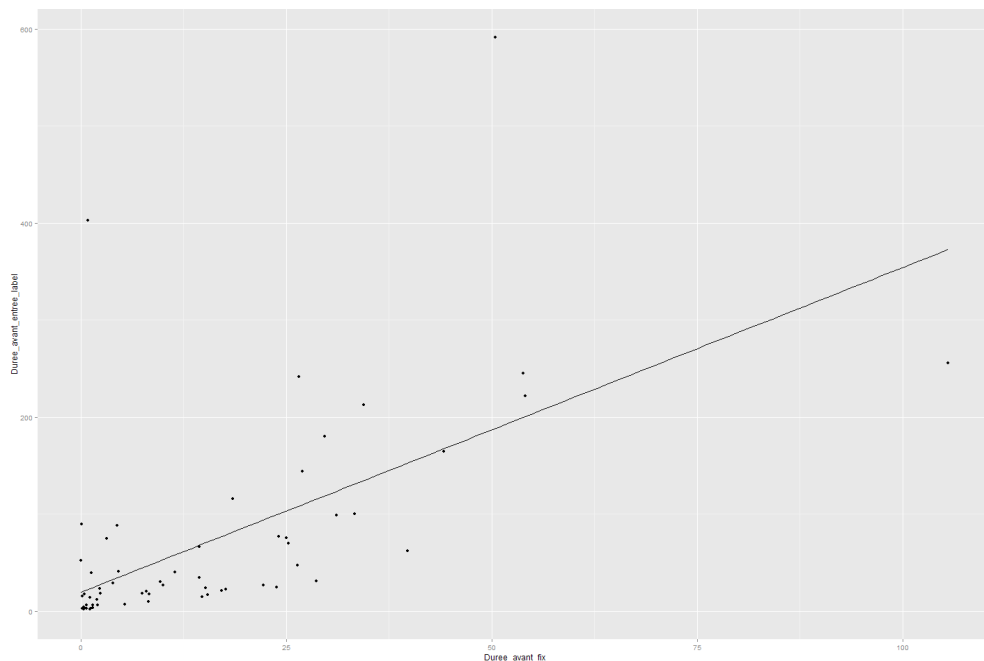


Figure 70 : durée avant fixation =f(durée avant entrée label)

La dernière analyse porte sur la durée des problèmes du point de vue de SAmonitoring (Figure 71 : durée du problème détecté par l'agent de monitoring), c'est-à-dire le temps mis par le contrôleur pour agir sur le vol soit en donnant un ordre au pseudo-pilote (ce qui va faire disparaître le problème) soit en modifiant les informations connues sur le vol en agissant sur le label. La durée moyenne d'une alerte est de 129s (std 164) pour le groupe avec les notifications et 241s (std 239) pour le groupe sans.

L'analyse de la durée du monitoring ne montre elle aussi aucun effet de la variable conflit $F(1,115)=0.257$ $p=0.613$. Le groupe (On/Off) en revanche est très significatif $F(1,115)=8.762$ $p=0.0037$. Les problèmes détectés par le SAmonitoring ont donc été résolus significativement plus vite par les contrôleurs du groupe ON validant ainsi l'hypothèse **H4**.

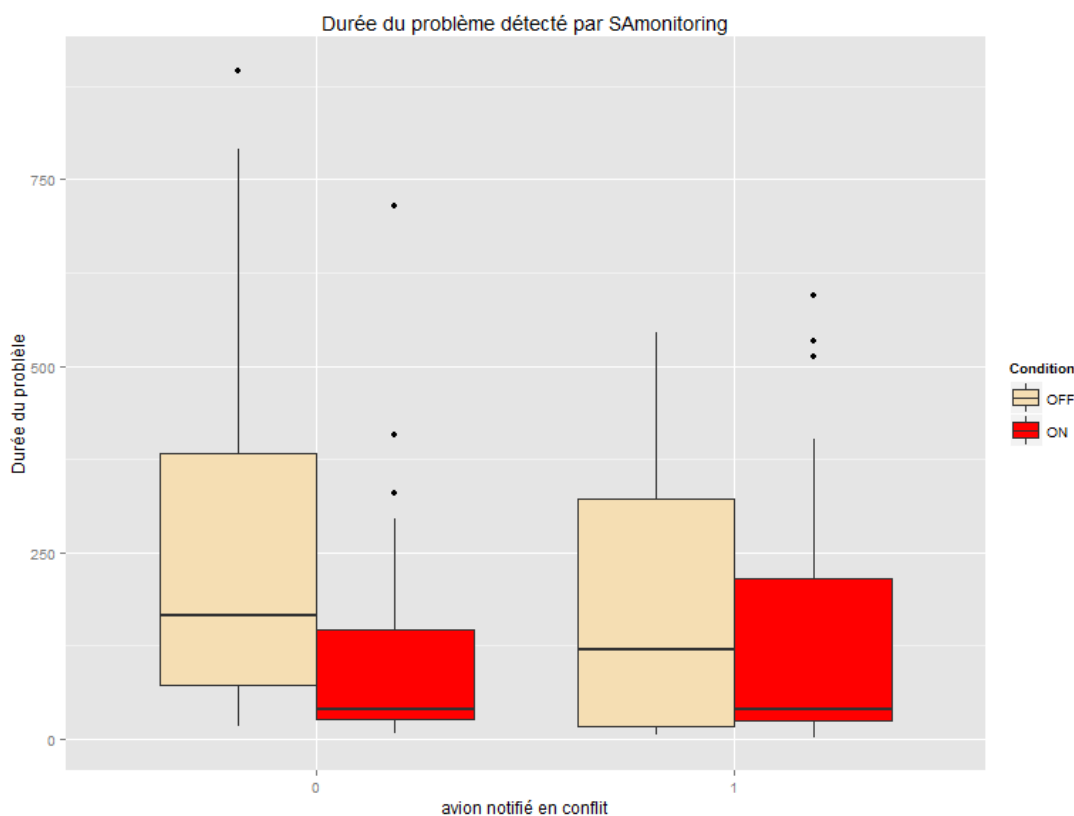


Figure 71 : durée du problème détecté par l'agent de monitoring

8.3.1.2. Expérimentation WP4.7.2 sans oculométrie

Pendant les 28h30 de simulation réalisées avec SAmonitoring et le système de notification actif (groupe ON), 31 notifications critiques ont été présentées aux contrôleurs. Ces cas sont en grande partie dus au renseignement tardif par les contrôleurs des clairances qu'ils ont donné aux avions, les autres à des incompréhensions ou des erreurs de saisie du pseudo-pilote. Par ailleurs et contrairement à l'expérimentation Monitoring, il n'était pas demandé explicitement aux contrôleurs de traiter rapidement les problèmes détectés par le système.

Pendant les 10h30 de simulation réalisées avec SAmonitoring sans le système de notification (groupe OFF), 9 cas critiques ont été détectés par SAmonitoring. Un chiffre comparable au taux d'alerte du groupe ON. La moyenne des durées avant l'entrée dans un label pour le groupe ON est de

10.48s (std 12.16) pour le groupe ON contre 61.22s (std 125.4) pour le groupe OFF.

Contrairement à l'expérimentation SAmonitoring, il n'y a pas de cas de problème non détecté au-delà des 60s pour le groupe OFF, ce qui était attendu puisque nous n'avons pas provoqué de cas difficiles à détecter pour le contrôleur pendant ces simulations.

Compte tenu des faibles effectifs du groupe OFF et de l'écart d'effectif entre les deux groupes, nous avons comparé les deux conditions à l'aide du test de Mann-Whitney-Wilcoxon de la somme des rangs (test non paramétrique ne requérant pas de condition d'application). Le résultat de ce test $p=0.043$ est significatif.

Il est toutefois difficile de conclure compte tenu de la disproportion des effectifs des deux groupes. On peut néanmoins constater un nombre très important de cas détectés précocement dans le groupe ON ce qui a un impact très important sur la valeur moyenne. Les outliers sont très comparables pour les deux groupes ON et OFF, alors que le nombre de cas non détectés sur de grandes durées était bien plus important sur l'expérimentation Monitoring. Cette différence est liée au fait que nous avons essayé lors de Monitoring de provoquer des cas subtils et difficiles à détecter pour le contrôleur en l'absence d'un système d'aide en condition de charge de travail élevée.

8.3.2. Résultats qualitatifs

Les questionnaires fournis en annexe ont été remplis par les 12 sujets répartis en deux groupes, avec ou sans monitoring. Le questionnaire 'avec' monitoring reprend toutes les questions du questionnaire 'sans' en ajoutant des questions spécifiques à l'utilisation et la représentation des informations issues du monitoring sur la supervision radar.

Compte tenu des faibles effectifs ($n_1=6$ et $n_2=6$) nous avons comparé les deux conditions à l'aide du test de Mann-Whitney-Wilcoxon de la somme des rangs (test non paramétrique ne requérant pas de condition d'application).

Évaluation subjective	Statistique W	p-value
Satisfaction	31.5	.03*
Réaliste	19.5	.86
Trafic subjectif	21	.59
Conscience de la situation	26	.22
Détection de problèmes	28.5	.10
Sécurité	33	.02*
Performance	27	.14

* : significatif à $p<.05$

Seules deux questions font apparaître des différences significatives entre les deux groupes. Au contraire du groupe sans, le groupe avec le monitoring affiché sur la supervision radar a pensé réaliser la simulation de façon satisfaisante avec un niveau de sécurité élevé. Les réponses sur la conscience de la situation sont trop dispersées et l'échantillon trop faible pour en tirer

des conclusions (Figure 72 : résultats questionnaires expérimentation monitoring).

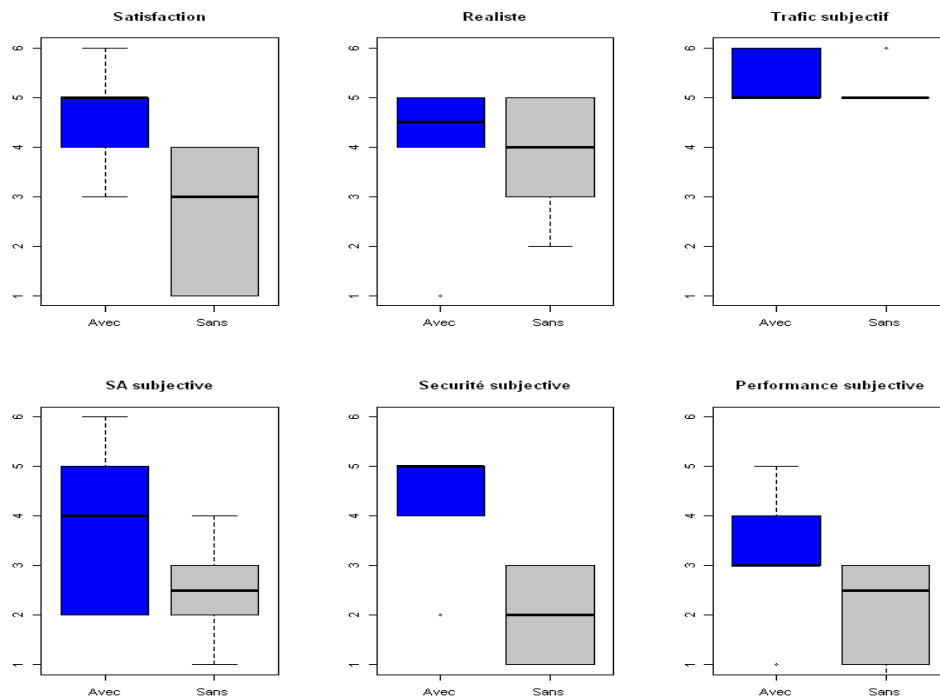


Figure 72 : résultats questionnaires expérimentation monitoring

Sur les questions spécifiques au groupe avec monitoring (Figure 73 : résultats questionnaire avec monitoring), les contrôleurs ont globalement trouvé les informations représentées suffisantes et s'intégrant bien à la méthode de travail sans trop perturber. Sur la conscience de la situation là encore, les réponses sont très dispersées pour en tirer des conclusions.

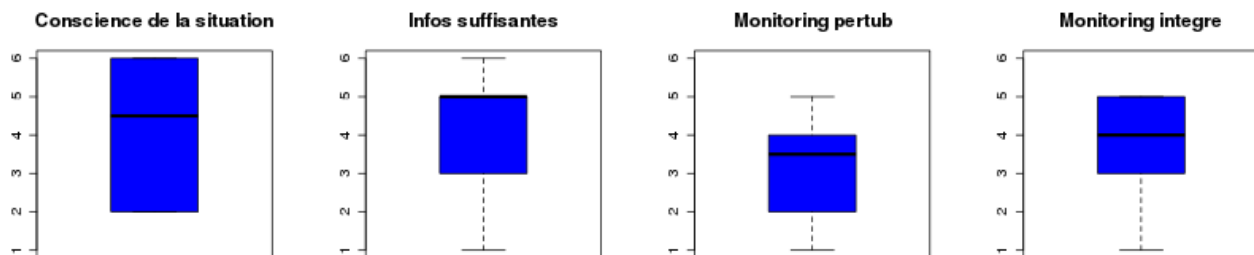


Figure 73 : résultats questionnaire avec monitoring

8.4. Conclusion

Nous avons formulé 4 hypothèses préalablement à ces expérimentations, 3 d'entre elles ont été validées (en gras) :

H1 : SAmonitoring et son système de notification améliorent la détection des situations critiques

H2 : lorsque les avions sont en conflit, ils ne sont pas mieux surveillés par le contrôleur

H3 : SAmonitoring et son système de notification diminue le temps mis par le contrôleur pour agir sur le vol

H4 : SAmonitoring et son système de notification permettent de diminuer la durée des situations critiques

Les résultats préliminaires de cette expérimentation sont conformes à ce que nous attendions sur l'utilité du SAmonitoring et de son système de notifications. Ce nouvel outil d'aide et les notifications que nous avons évalués et mis en place dans ce prototype permettent d'améliorer très significativement la conscience de la situation des contrôleurs dans le suivi des avions sous leur responsabilité et tout particulièrement dans le cas des avions sur lesquels les problèmes à détecter sont subtils. Nous avons formulé l'hypothèse que ce système serait peut-être moins intéressant dans le cas des avions impliqués dans un conflit puisqu'ils sont normalement plus surveillés par les contrôleurs, cela n'a pas été le cas ; il semble que ceux-ci surveillent tous les avions de la même façon.

Nous limiterons la portée de nos résultats en raison des choix expérimentaux que nous avons dû faire pour nous insérer dans le planning des expérimentations du WP4.7.2. Nous n'avons qu'un créneau pour une simulation d'une durée maximum de 40mn et 12 contrôleurs. Pendant ces passations, qui étaient parfaitement réalistes du point de vue de l'activité de contrôle, nous avons provoqué une dizaine de situations qui ne se produisent que de façon rare dans leur activité. Ce biais a limité le réalisme de nos simulations, mais les résultats obtenus montrent un gain très important sur la détection des alertes critiques par les contrôleurs lorsque le système est actif.

Les résultats obtenus sur les questionnaires subjectifs ne font pas apparaître de différences significatives pour le groupe ON que sur leur perception du déroulement et du niveau de sécurité de la simulation. Les questions portant sur la conscience de la situation n'ont pas permis de distinguer les deux groupes, un questionnaire basé sur la SA aurait sans doute été plus adapté, mais nous ne disposions pas du temps nécessaire pour les réaliser pendant le créneau disponible.

En ce qui concerne les passations sans oculométrie, il est difficile de conclure sur les résultats sans une analyse détaillée des outliers présents dans nos données. Par ailleurs le déséquilibre entre les deux groupes ON/OFF limite la puissance statistique de nos analyses. Des simulations complémentaires du système en condition réaliste (sans provoquer d'alerte critique) devront donc être faites pour confirmer les résultats obtenus dans l'expérimentation SAmonitoring. Néanmoins, les contrôleurs du groupe ON ont réagi très rapidement en agissant sur l'étiquette radar des avions notifiés, en moyenne en 10s alors qu'ils ont mis en moyenne 19s lors de l'expérimenta-

tion Monitoring. Cette différence est très certainement liée à la difficulté de la simulation qu'ils devaient réaliser lors de Monitoring et la rareté de ces alertes dans l'expérimentation 4.7.2. Une alerte était présentée en moyenne toutes les 8min dans Monitoring contre une en moyenne toutes les 55 min de simulation dans l'expérimentation 4.7.2.

Une étude complémentaire permettant de mesurer la charge de travail pendant les simulations permettrait de conclure au lien possible entre celle-ci et le temps mis par le contrôleur pour percevoir et agir sur les avions-cibles d'une alerte critique.

9. CONCLUSION

9.1. Bilan et synthèse des résultats

Les outils et méthodes des contrôleurs aériens vont évoluer dans le futur pour faire face à l'augmentation prévue du trafic aérien. Pour permettre aux contrôleurs de continuer à assurer leur tâche dans de bonnes conditions en contrôlant un nombre plus important d'avions que par le passé, le système de gestion du contrôle aérien va évoluer en intégrant des automates qui aideront le contrôleur dans sa tâche. Le métier de contrôleur aérien va évoluer du rôle d'opérateur vers celui de superviseur. Il devra maintenir une bonne conscience de la situation aérienne tout en prenant en compte les notifications issues des systèmes l'aidant dans sa tâche.

Ces travaux de thèse nous ont amené à nous intéresser à la conscience de la situation des contrôleurs aériens. Celle-ci s'établit grâce au suivi et à la mémorisation des informations présentées sur la visualisation radar. Parmi ces informations, la perception des notifications ou alertes issues des systèmes d'aide est d'ores et déjà critique pour le contrôleur aérien.

Pour tenter d'améliorer la conscience de la situation des contrôleurs, notre travail de recherche s'est porté sur deux axes. Nous avons étudié la tâche des contrôleurs aériens et réalisé un agent s'assurant de la réalisation de plusieurs objectifs la composant. Nous avons proposé une méthode d'évaluation de designs de notification s'appuyant sur des mesures comportementales, physiologiques et subjectives et un micro-monde ATC. Cette méthode a été employée lors de deux expérimentations visant à évaluer les caractéristiques de plusieurs designs dont deux sont utilisés dans les visualisations opérationnelles. Enfin, nous avons évalué l'impact sur la réalisation de la tâche du contrôleur aérien, d'une nouvelle position de contrôle mettant en œuvre l'agent de suivi de la tâche ainsi qu'un système de notification dont les designs sont adaptés à la criticité de l'information.

La première étude, l'expérimentation Saillance avait pour objet d'évaluer l'intérêt et la performance de cinq types de designs de notifications visuelles dans un contexte ATC (tâche de surveillance sur grand écran avec des notifications en vision périphérique) à l'aide de mesures comportementales et physiologiques basées sur l'utilisation d'un oculomètre. Parmi ces designs, deux étaient basés sur des designs de notifications utilisés sur des visualisations radar opérationnelles, les trois autres étaient des designs prototypes utilisant des animations. Les sujets devaient réaliser en parallèle une tâche de guidage d'un avion dans un labyrinthe en saisissant des consignes et une tâche de validation des notifications. Les critères d'évaluations de la tâche de guidage portaient sur le pourcentage de consignes correctes saisies ainsi que le temps mis pour réaliser les saisies. La tâche de validation des notifications était évaluée par le temps mis valider une notification et du nombre de notifications non validées.

L'expérimentation Saillance a permis de mettre en évidence les résultats suivants :

- Les trois designs de notification utilisant du mouvement ont été perçus rapidement et sans erreur
- Les deux designs utilisés actuellement dans les visualisations opérationnelles sont les moins efficaces en termes de temps de perception
- Parmi les trois designs utilisant le mouvement, deux ont dégradé significativement la tâche principale

Cette étude montre l'intérêt d'adopter une approche holistique qui tient compte de multiples aspects de la tâche de l'opérateur lors de l'introduction et l'évaluation de nouvelles solutions technologiques. L'utilisation de l'oculométrie permet d'enrichir les résultats comportementaux obtenus lors du déroulement d'une tâche et permet de faciliter l'analyse de l'utilisation des visualisations. Les designs évalués au cours de cette étude ont des propriétés différentes. Leur évaluation dans ce contexte expérimental contrôlé a permis de les mettre en évidence et d'établir leur coût attentionnel c'est-à-dire le rapport entre le niveau de saillance et l'impact de l'interruption de l'attracteur visuel sur la réalisation d'une autre tâche.

La seconde étude, l'expérimentation Bip avait pour objet en utilisant la tâche réalisée avec Saillance de provoquer le phénomène de surdité inattentionnelle et d'évaluer deux designs de notifications à l'aide de deux nouvelles variables : la charge visuelle et la charge de travail dans un contexte de triple tâche. Aux tâches de guidage et de validation des notifications de l'expérimentation précédente, une tâche de validation d'alarmes sonores a été ajoutée. À l'aide des résultats obtenus lors de l'expérimentation Saillance, deux groupes de sujets utilisant deux designs avec ou sans mouvement ont été établis. La charge visuelle a été manipulée en jouant sur le nombre d'avions pouvant faire l'objet d'une notification. La charge de travail quant à elle a été ajustée à l'aide de la vitesse de l'avion objet de la tâche de guidage. Lors de la tâche de validation sonore, deux types de sons : rares ou standards étaient présentés ; en cas de sons rares les sujets devaient appuyer sur un bouton. Là encore des mesures physiologiques ont été mises en œuvres à l'aide d'un EEG et ont permis par l'analyse de l'amplitude de la P300 (potentiel évoqué positif 300ms après un stimulus) d'évaluer l'impact des différentes variables sur la charge de travail.

L'expérimentation Bip a permis de mettre en évidence les résultats suivants :

- Le design des notifications visuelles a eu un impact significatif sur la réalisation de la tâche de validation sonore,
- En augmentant la charge de travail (vitesse), l'effet de la qualité du design mis en évidence dans l'expérimentation Saillance s'accroît,
- L'amplitude de la P300 augmente en présence de sons rares lorsque le design de notification visuelle est plus performant,
- la charge visuelle n'a pas d'impact sur la performance des sujets

Cette étude a apporté des résultats complémentaires très intéressants à l'expérimentation Saillance. Elle a permis de mettre en évidence, de façon plus fine et grâce à l'utilisation d'un EEG le coût attentionnel d'un design de notification visuelle et de voir l'impact de la qualité d'un design sur la charge de

travail globale d'un opérateur réalisant des tâches multiples. Par ailleurs, nos résultats vont dans le sens d'une répartition des ressources attentionnelles entre la vision et l'audition. Nous avons pu constater une dégradation significative de la tâche sonore lorsque la tâche visuelle devient plus difficile. Elle met en évidence l'intérêt d'obtenir une indication en temps réel de la charge de travail pour adapter les interfaces et ajuster la quantité d'information à traiter qu'elles soient visuelles ou sonores.

Dans la troisième étude, l'expérimentation Monitoring, nous avons évalué une nouvelle position de contrôle dont le design est orienté conscience de la situation. Cette nouvelle position intègre l'agent SAmonitoring et un système de notification développés dans le cadre de ces travaux. Les designs utilisés par le système de notification ont été choisis en fonction de leur saillance grâce aux résultats de l'expérimentation Saillance de façon à avoir un pouvoir attentionnel conforme à la criticité de l'information délivrée. Lors de cette expérimentation, des incidents difficiles à détecter étaient provoqués par un contrôleur complice qui aboutissaient à la détection par SAmonitoring d'alertes critiques.

Deux groupes de sujets ont dû réaliser ces simulations, l'un des deux ne disposait pas du système de notification sur sa visualisation radar, l'autre en disposait et était donc prévenu dès l'apparition des incidents.

Cette expérimentation a permis de mettre en évidence les résultats suivants :

- L'agent SAmonitoring et son système de notification permettent de détecter significativement plus rapidement les problèmes de type critiques lorsque les notifications sont présentées sur la visualisation radar
- L'agent SAmonitoring et son système de notification sont utiles pour améliorer la conscience de la situation des contrôleurs
- L'agent SAmonitoring et son système de notification réduisent la durée des situations critiques

SAmonitoring et son système de notification ont aidé le contrôleur à maintenir une bonne conscience de la situation en particulier dans une situation de très forte charge de travail. Par ailleurs, dans le cas de problèmes difficiles à détecter, SAmonitoring permet une détection et un traitement précoce, en son absence certaines situations n'ont tout simplement pas été vues par les contrôleurs. Cette dernière expérimentation vient clôturer ces travaux de thèse dont l'objet était d'améliorer la conscience de la situation des contrôleurs en détectant les informations dynamiques pertinentes pour la réalisation de leur tâche et en les montrant à l'aide d'un design adapté à leur importance. Un travail important reste à faire sur la représentation des informations de façon à limiter la charge cognitive liée à leur perception en s'inspirant des techniques utilisées dans le design écologique. Ces travaux sont aussi une première étape vers des IHM adaptatives grâce à l'intégration en temps réel d'informations en provenance de l'analyse du contexte et de capteurs physiologiques. L'état de l'opérateur permettrait d'agir sur le niveau d'automatisation des systèmes l'aidant dans sa tâche et la représentation des informations qui lui sont présentées pour optimiser leur transmission à l'humain.

9.2. Perspectives de recherche

9.2.1 Impact du moyen de saisie sur la mémoire à court terme

La conception de l'interface homme-machine mise en œuvre dans la position de travail Droid a été réalisée en postulant que l'écriture est plus efficace d'un point de vue conscience de la situation et mémorisation que la sélection d'un item dans une liste de valeurs comme cela est le cas dans les supervisions radar tout électronique utilisées aujourd'hui dans le monde. L'écriture est un mécanisme lent, associé à une gestuelle ; la sélection d'un item dans un menu déroulant est très rapide, mais peut être source d'erreurs. Il serait très intéressant de justifier ou infirmer nos choix en étudiant l'impact sur la mémoire du moyen d'interaction et de saisie. Laby permettant de changer de moyen de saisie très simplement, nous envisageons de réaliser une expérimentation où nous évaluerons l'impact sur la mémoire à court ou moyen terme du moyen de saisie. Nous pourrions aussi nous intéresser à la durée de l'interaction, à son taux d'erreur et à son impact sur les diverses tâches que nous pouvons mettre en œuvre grâce à Laby.

9.2.1. Utilisation dynamique de la détection de l'état cognitif : le projet NINA

Le projet NINA financé par le projet SESAR work package WPE a démarré fin 2013. Il a pour but de développer une boîte à outils pour surveiller l'état cognitif des contrôleurs aériens grâce à une combinaison de mesures physiologiques. L'un des objectifs de ces travaux est de développer une nouvelle approche afin de déterminer le degré approprié d'automatisation en utilisant l'état cognitif de l'opérateur. Les mesures physiologiques concernent les mouvements des yeux, la fréquence et la durée de clignotement, le rythme cardiaque (ECG), la résistance électrodermale (GSR) et de la mesure continue EEG. La combinaison de l'EEG avec d'autres mesures neurophysiologiques (EOG, ECG, GSR) peut améliorer la discrimination de l'état mental de l'opérateur (Qian, 2009).

Le projet vise à développer un ensemble d'indicateurs temps réels pour suivre et mesurer la charge de travail, la fatigue. Plusieurs phases dans la réalisation de ce projet ont été définies. La première qui a déjà fait l'objet de publication (voir section 8.3.1) a été réalisée à l'aide de Laby couplé à un EEG et un oculomètre dans le but de déterminer le niveau d'apprentissage des sujets (Borghini et al., 2014) ainsi qu'un indicateur de la charge de travail basé sur de l'analyse continue des ondes cérébrales. Les étapes suivantes prévues pour 2015, mettront ces données à profit pour intégrer à la plateforme Droid mise en place dans le cadre de ces travaux de thèse. Un classifieur temps réel d'états mentaux sera utilisé par l'agent de monitoring et la visualisation radar de façon à changer leur comportement en fonction de la charge de travail.

Nous évaluerons à l'aide d'expérimentations réalisées avec des contrôleurs si l'adaptation en temps réel de nos IHMs permet de réduire la charge de travail lorsque celle-ci est déjà à des niveaux élevés.

9.2.2. Extension aux notifications haptiques

Les notifications haptiques sont généralisées aujourd'hui pour notifier des informations aux utilisateurs de téléphone mobile, des applications dans les transports apparaissent dans des véhicules haut de gamme pour éviter sorties de route et collisions avec d'autres véhicules. Dans le domaine de l'IHM, différentes techniques utilisant des feedbacks haptiques ont été mises en œuvre dans le cadre de l'utilisation d'écrans tactiles pour permettre aux utilisateurs de 'ressentir' l'interface (Poupyrev, Okabe & Maruyama, 2004) ou notifier des avertissements ou alertes. Un système développé par General Motors a pour but d'améliorer la conscience de la situation des conducteurs d'automobiles en produisant un retour haptique sur le siège du conducteur. Deux vibreurs indépendants sont installés sous le siège et entrent en fonction en cas de détection d'un risque de collision avec un autre objet, leur déclenchement individuel est lié à la position de l'obstacle par rapport au véhicule. Ce 'Safety Alert Seat' fait partie d'un ensemble de systèmes de sécurité nommé 'Driver Awareness'. Qu'elles soient sonores ou haptiques, les notifications sont plus efficaces sur le conducteur si elles sont graduées (Lee, Hayes, 2004). Dans cette étude, les notifications haptiques ont été reçues plus positivement que les notifications sonores sans graduation de niveau.

Dans le domaine de l'aviation, des retours haptiques ont été ajoutés aux appareils modernes équipés de commandes de vol électrique pour simuler le comportement des avions et de leurs commandes de vol de l'époque hydraulique et mécanique. Le phénomène du 'buffeting' qui se traduit par une vibration sur le manche lors du début d'un décrochage dans les avions à commandes de vol hydrauliques traduisant la vibration des voilures mobiles a été reproduit dans les nouveaux appareils à commandes de vols électriques ou la commande est découplée physiquement de la voilure. À ce jour et à notre connaissance, aucun système haptique n'a été mis en service dans les systèmes de visualisation pour les contrôleurs aériens. Un prototype de feedbacks haptiques a néanmoins été développé au Centre d'Études de la Navigation Aérienne (Imbert, Bustico, 2004) pour rendre tangibles les interactions d'un outil de manipulation de strips électroniques dans une interface tactile (Figure 74 : notifications haptiques avec Digistrrips 4). Ces objets électroniques sont une représentation des strips papiers utilisés encore aujourd'hui par les contrôleurs aériens en France. Ils peuvent être déplacés dans des colonnes grâce à de la manipulation directe tactile ; lorsqu'une colonne est pleine et que l'opérateur veut relâcher un strip dedans, l'application ne le permet pas et produit une vibration sur l'écran pour notifier cette impossibilité à l'opérateur. L'appui sur un des boutons de l'application produit aussi une vibration avec une forme d'onde différente. Toujours dans le domaine ATC, une étude basée sur l'utilisation d'un simulateur ATC basse fidélité (Ngo, Pierce, Spence, 2012) a comparé l'utilisation de plusieurs types de modalités de notifications (visuelles sonores et haptiques) a mis en évidence des temps de réaction significativement plus faibles aux stimuli lors de l'utilisation d'un feedback audio couplé à un feedback haptique. Ce type de résultat montre l'intérêt d'une approche multi sensorielle dans les systèmes de notifications.

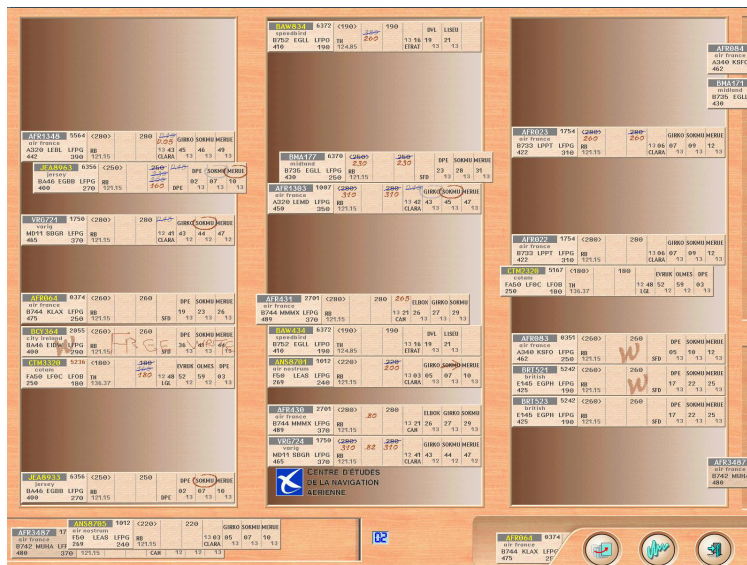


Figure 74 : notifications haptiques avec Digistrrips 4

Il serait intéressant d'étendre l'étude Bip aux notifications haptiques pour essayer d'évaluer et comparer l'impact et la performance des divers moyens de notification de façon à les utiliser à bon escient.

9.2.2.1. Évaluation de nouveaux designs de notifications

Le projet 4Flight en cours de développement à la DTI va bénéficier dans le futur d'un agent comparable à SAmonitoring. Des notifications sur la visualisation radar issues des informations provenant de cet agent seront montrées aux contrôleurs. Plusieurs designs ont été prototypés pour cette nouvelle visualisation radar. De façon à obtenir des données comportementales et physiologiques sur ces nouveaux designs et faire des choix adaptés, nous intégrerons ceux-ci à Laby avec le même protocole que celui utilisé dans l'expérimentation Saillance ce qui nous permettra de les comparer avec ceux déjà évalués.

9.3. Contributions

9.3.1. Contributions scientifiques

- Imbert J-P., Hodgetts H.M., Parise R., Vachon F., Dehais F. & Tremblay S. Attentional costs and failures in air traffic control notifications, Ergonomics (in-press)
- Imbert J-P., Hodgetts H.M., Parise R., Vachon F., & Tremblay S. The LABY microworld: À Platform for Research, Training and System Engineering, HFES 2014
- Parise R., Imbert J-P, Marais F., Alonso R., Laby Un support d'aide à l'évaluation de choix de conception d'IHM pour le contrôle aérien, Ergo'IHM 2012, Biarritz, France.
- Martin C., Imbert J-P Introduction of a more automated environment in En-Route Air Traffic Control Sesar Innovation Days 2012

- Borghini G., Aricò P., Babiloni F., Granger G., Imbert J-P, Benhacene R.; Napoletano L., Pozzi S., NINA: Neurometrics Indicators for ATM, SESAR Innovation Days 2013
- Giraudet G., Bérenger M., Imbert J-P, Tremblay S. and Causse M. Inattentional Deafness in Simulated Air Traffic Control Tasks: A Behavioral and P300 Analysis Proceedings of the 5th International Conference on Applied Human Factors and Ergonomics AHFE 2014, Kraków, Poland 19-23 July 2014
- Gianluca Borghini, Pietro Aricò, Federico Ferri, Ilenia Graziani, Simone Pozzi, Linda Napoletano, Jean-Paul Imbert, Geraud Granger, Railane Benhacene, Fabio Babiloni A neurophysiological training evaluation metric for Air Traffic Management Conference: 36th Annual International IEEE EMBS Conference (EMBC'14)

En cours de publication

- Imbert J-P, Peysakhovich V., Blatter C., Camachon C., Dehais F. Design requirements to integrate eye trackers in experimental environments
- Giraudet L., Imbert J-P, Causse M., Tremblay S. Evaluation of visual alerts in Air Traffic Control using behavioral and auditory P300 measures (journal)

9.3.2. Réalisations

Différents logiciels ont été réalisés ou modifiés dans le cadre de cette thèse :

- Gateway Facelab/Ivy
- Évolutions de Laby pour permettre l'évaluation de notifications visuelles et sonores en contexte multitâche
- Intégration des données oculométriques dans Laby pour la détection des fixations sur des zones d'intérêt dynamiques
- Développement d'un mécanisme de synchronisation entre Laby et un EEG corrélé avec l'émission de notifications sonores
- Développement de l'agent de suivi de la tâche SAmonitoring
- Développement d'un plugin pour l'image radar utilisant les informations de SAmonitoring pour notifier des informations à l'aide de designs adaptés
- Développement d'un plugin pour l'image radar utilisant en temps réel les données brutes de l'oculomètre

Références Bibliographiques

- Ahlstrom, V., & Panjwani, G. (2003). Auditory alarms in airways facilities environment (DOT/FAA/CT-TN04/04). Atlantic City Airport, NJ: Federal Aviation Administration, William Hughes Technical Center.
- Allen, J. E., Guinn, C. I., & Horvitz, E. (1999). Mixed-initiative interaction. *Intelligent Systems and their Applications, IEEE*, 14(5), 14-23.
- Allison, B. Z., & Polich, J. (2008). Workload assessment of computer gaming using a single-stimulus event-related potential paradigm. *Biological psychology*, 77(3), 277-283.
- Alonso, R., Causse, M., Vachon, F., Parise, R., Dehais, F., & Terrier, P. (2013). Evaluation of head-free eye tracking as an input device for air traffic control. *Ergonomics*, 56, 246-255.
- Alonso, R., P. Terrier, R. Parise, & Cellier, J.M. (2009). Effet d'une de corrélation visuomotrice sur le mouvement de pointage. *Le Travail Humain*, 72, 245-66.
- Ancman, E. (1991). Peripherally located CRTs: Color perception limitations. In *Proceedings of the IEEE National Aerospace and Electronics Conference* (pp. 960 – 965). New York: Institute of Electrical and Electronics Engineers.
- Arrabito G.R, Mondor T.A., Kent K. J. (2004) Judging the urgency of non-verbal auditory alarms: a case study *Ergonomics* Vol. 47, Iss. 8
- Astolfi L., Toppi J., De Vico Fallani F., Vecchiato G., Cincotti F., Wilke C. T., Yuan H., Mattia D., Salinari S. He B., Babiloni F., *Imaging the Social Brain by Simultaneous Hyperscanning During Subject Interaction, IEEE Intelligent Systems*, Vol. 26, No. 5, October 2011, pp. 38-45.
- Athènes, S., Chatty, S. & Bustico (2000). Human factors in ATC alarms and notifications design: an experimental evaluation - Proc. ATM'2000 R&D seminar, 2000 - atmseminar.org
- Averty, P. (1998). Les effets de la charge de trafic sur le niveau d'activation psychophysiologique du contrôleur aérien. Thèse de Doctorat en Psychologie Cognitive, Université Lumière Lyon 2.
- Ayaz, H., Shewokis, P. A., Bunce, S., Izzetoglu, K., Willems, B., & Onaral, B. (2012). Optical brain monitoring for operator training and mental workload assessment. *Neuroimage*, 59(1), 36-47.
- Baccino, T. (2002). Oculométrie Cognitive. In G.Tiberghien (Ed.), *Dictionnaire des Sciences Cognitives* (pp. 202-203). Paris: Armand Colin.
- Baddeley, A.D. & Della Sala, S. (1996). Working memory and executive control. *Proceedings of the Royal Society London*, B,351, 1397-1404.
- Bailey, L. L., Broach, D. M., Thompson, R. C., & Enos, R. J. (1999). Controller teamwork evaluation and assessment methodology: (CTEAM): A scenario

- calibration study.(DOT/FAA/AAM-99/24). Washington, DC: Federal Aviation Administration Office of Aviation. Medicine. Available from: National Technical Information Service, Springfield, VA 22161.
- Bailey B. P., Konstan J. A., & Carlis J. V. (2001). The effects of interruptions on task performance, annoyance, and anxiety in the user interface. In: M. Hirose (Ed.) *Human-Computer Interaction - INTERACT 2001 Conference Proceedings*. Amsterdam: IOS Press, 593-601.
- Banbury, S., Fricker, L., Tremblay, S., & Emery, L. (2003). Using auditory streaming to reduce disruption to serial memory by extraneous auditory warnings. *Journal of Experimental Psychology: Applied*, 9, 12–22.
- Barry R.J., Clarke A.R., Johnstone S.J., Magee C.A., Rushby J.A. (2007), EEG differences between eyes-closed and eyes-open resting conditions. *Clinical Neurophysiology*, Vol. 118, No. 12, 2007, pp 2765-2773.
- Bartram, L., Ware, C., & Calvert, T. (2001). Moving icons: detection and distraction. In: *Proceedings of the IFIP TC.13 International Conference on Human-Computer Interaction (INTERACT 2001)*. Tokyo, Japan.
- Bartram, L., Ware, C., & Calvert, T. (2003). Moticons: detection, distraction and task. *International Journal of Human-Computer Studies*, 58, 515-545.
- Baudisch, P. Halo: supporting spatial cognition on small screens.(2003) In *UIST'03 posters and demos companion (demo paper)*, Vancouver Canada
- Bauer, B., Jolicœur, P., & Cowan, W.B. (1996). Visual search for colour targets that are or are not linearly separable from distractors. *Vision Research*, 36, 1439–1466.
- BEA rapport Tarom 381 940924 (1994) <http://www.bea.aero/docspa/1994/yr-a940924/pdf/yr-a940924.pdf>
- BEA rapport final d'enquête AF447 2013 <http://www.bea.aero/fr/enquetes/vol.af.447/rapport.final.fr.php>
- Beatty, J. (1982). Task-evoked pupillary responses, processing load, and the structure of processing resources. *Psychological Bulletin*, 91, 276-292.
- Benhacène, R., Marion, A., & Rousselle, M. P. (2005). As Rapid as PaperStrips? Evaluation of VertiDigi, a new control tool for Terminal Sectors. *ATM2005*.
- Beringer, D. B., & Harris, H. C., Jr. (1999). Automation in general aviation: Two studies of pilot responses to autopilot malfunctions. *International Journal of Aviation Psychology*, 9, 155–174.
- Bertin Jacques, *Sémiologie graphique*, Paris, Mouton/Gauthier-Villars, 1967.
- Bertin, J. 1983. *Semiology of Graphics*. University of Wisconsin Press.
- Bliss, J., & Dunn, M. (2000). The behavioral implications of alarm mistrust as a function of task workload. *Ergonomics*, 43, 1283–1300
- Bliss, J. P. (2003) "Investigation of alarm-related accidents and incidents in aviation"- *The International Journal of Aviation Psychology*, 5(3), 257–276.

- Boff, K. R., & Lincoln, J. E. (Eds.). (1988). *Engineering Data Compendium: Human Perception and Performance*. Wright-Patterson AFB, OH: USAF Harry G. Armstrong Aerospace Medical Research Laboratory (AAMRL)
- Boot, W.R., Becic, E., & Kramer, A.F. (2007). Temporal limitations in multiple target detection in a dynamic monitoring task. *Human Factors*, 49, 897-906.
- Boucsein, W.: *Electrodermal Activity*. Springer, Berlin, 1992 (2nd edition. 2010)
- Bradshaw, J. M., Sierhuis, M., Acquisti, A., Feltovich, P., Hoffman, R., Jeffers, R., ... & Van Hoof, R. (2003). Adjustable autonomy and human-agent teamwork in practice: An interim report on space applications. In *Agent Autonomy* (pp. 243-280). Springer US.
- Brand-D'Abrescia, M., & Lavie, N. (2008). « Task Coordination between and within sensory modalities: Effects on distraction ». *Perception & Psychophysics*, 70(3), 508–515.
- Brangier, E. & Barcenilla, J. (2003). *Concevoir un produit facile à utiliser*, Editions d'Organisation.
- Brehmer, B. (1992). Dynamic decision making: Human control of complex systems. *Acta Psychologica*, 81, 211-241.
- Brehmer, B., & Dörner, D. (1993). Experiments with computer-simulated microworlds: Escaping both the narrow straits of the laboratory and the deep blue sea of the field study. *Computers in Human Behavior*, 9(2-3), 171–184. doi:10.1016/0747-5632(93)90005-D
- Brehmer, B., Schifflett, S. G., Elliott, L. R., Salas, E., & Covert, M. D. (2004). Some reflections on microworld research. *Scaled worlds: Development, validation and applications*, 22-36.
- Brennen, S.D. (1992). *An Experimental Report on Rating Scale Descriptor Sets for the Instantaneous Self Assessment (ISA) Recorder, DRA/TM/CAD5/92017*. Defence Research Agency. Portsmouth.
- Brookings, J., Wilson, G., & Swain, C. (1996). Psychophysiological responses to changes in workload during simulated air traffic control. *Biological Psychology*, 42(3), 361-377.
- Bruneau, D., Sasse, M. A. & McCarthy, J.D. (2002). *The Eyes Never Lie: The use of eyetracking data in HCI research*. Proceedings of the CHI2002 Workshop on Physiological Computing, Minneapolis.
- Buisson & al, 2002. Ivy: un bus logiciel au service du développement de prototypes de systèmes interactifs. In *Proceedings of the 14th French-speaking conference on Human-computer interaction (IHM '02)*.
- Buchner, A. (1995). Basic topics and approaches to the study of complex problem solving. *Complex problem solving: The European perspective*, 27-63.
- Bulling, A., Dachsel, D., Duchowski, A., Jacob, R., Stellmach, S., Sundstedt, V. 2012. Gaze interaction in the post-WIMP world. In *CHI '12 Extended Abstracts*. ACM, New York, NY, USA, 1221-1224.

- Burns, C. M. & Hajdukiewicz, J. R. (2004). *Ecological Interface Design*. Boca Raton, FL: CRC Press. ISBN 0-415-28374-4
- Byrne, E. A., & Parasuraman, R. (1996). Psychophysiology and adaptive automation. *Biological psychology*, 42(3), 249-268.
- Cabrera, D., Ferguson, S., & Laing, G. (2005). Development of auditory alerts for air traffic control consoles. In *Audio Engineering Society Convention 119*. Audio Engineering Society.
- Cabrera, D., Ferguson, S., & Laing, G. (2006). Considerations Arising From The Development Of Auditory Alerts For Air Traffic Control Consoles. *group*, 14(5), 2.
- Card, S. K., Mackinlay, J. D., & Shneiderman, B. (Eds.). (1999). *Readings in information visualization: using vision to think*. Morgan Kaufmann.
- Carpenter, R.H.S. (1977). *Movements of the Eyes*. London: Pion
- Cartwright-Finch, U., & Lavie, N. (2007). The role of perceptual load in inattentive blindness. *Cognition*, 102, 321–340.
- Chan, H. S., & Courtney, A. J. (1993). Effects of cognitive foveal load on a peripheral single-target detection task. *Perceptual and Motor Skills*, 77, 515–533.
- Chapman, R.M. & Bragdon, H.R. (1964). Evoked responses to numerical and non-numerical visual stimuli while problem solving. *Nature*, 203, 1155-1157.
- Clarion, A. (2009). Recherche d'indicateurs électrodermiques pour l'analyse de la charge mentale en conduite automobile. Thèse de doctorat, Mention : Neurosciences du comportement. Université de Lyon.
- Comstock, J.R., & Arnegard, R.J. (1992). The multi-attribute task battery for human operator workload and strategic behavior research. (Technical Memorandum No.104174). Hampton, VA: National Aeronautics and Space Administration.
- Crebolder, J. M. (2012). Investigating visual alerting in complex command and control environments. *Journal of Human Performance in Extreme Environments*, 10, DOI: <http://dx.doi.org/10.7771/2327-2937.1000>.
- Cummings, M. L., & Mitchell, P. J. (2008). Predicting controller capacity in supervisory control of multiple UAVs. *IEEE Transactions on Systems, Man and Cybernetics, Part A: Systems and Humans*, 38(2), 451-460.
- Cushing, S. (1994) *Fatal Words – Communication clashes and aircraft crashes*. The University Press of Chicago: London.
- Cushing, S. (1995) *Pilot-Air Traffic Communication- It's not (only) what you say, it's how you say it*. *Flight Deck*, Winter 1995/6.
- Dallet S., Veitl A. *Du sonore au musical: Cinquante années de recherches concrètes (1948-1998)* Editions L'Harmattan
- Dehais, F., Causse, M., Vachon, F., & Tremblay, S. (2012). Cognitive conflict in human-automation interactions: A psychophysiological study. *Applied*

- Ergonomics, 43, 588-595.
- Dehais, F., Tessier, C., Christophe, L., & Reuzeau, F. (2010). The perseveration syndrome in the pilot's activity: Guidelines and cognitive countermeasures. *Human Error, Safety and Systems Development. Lecture Notes in Computer Science*, 5962, 68-80.
- Dehais, F. (2004). *Modélisation des conflits dans l'activité de pilotage*. École doctorale : Systèmes, avril 24.
- Dehais, F., Causse, M., & Tremblay, S. (2011). Mitigation of Conflicts with Automation use of cognitive countermeasures. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 53(5), 448-460.
- Dehais, F., Mickaël Causse, François Vachon, Nicolas Régis, Eric Menant, Sébastien Tremblay. Failure to Detect Critical Auditory Alerts in the Cockpit: Evidence for Inattentive Deafness Human Factors: *The Journal of the Human Factors and Ergonomics Society* 0018720813510735, first published on November 11, 2013 doi:10.1177/0018720813510735
- Dehais, F., S. Mercier, et C. Tessier. 2009. « Conflicts in Human Operator–Unmanned Vehicles Interactions ». *Engineering Psychology and Cognitive Ergonomics*: 498-507.
- Dehais, F., C. Tessier, et L. Chaudron. 2003. GHOST: experimenting conflicts countermeasures in the pilot's activity. *International Joint Conference on Artificial Intelligence*, 18:163-168.
- Delignières, D., Legros, P., & Famose, J.P. (1991). Perception de l'effort et difficulté de la tâche. *Science et Motricité*, 13, 14-18.
- De Waard, D. (1996). *The measurement of drivers' mental workload*. PhD thesis, University of Groningen. Haren, The Netherlands: University of Groningen, Traffic Research Centre.
- DiFonzo, N., & Bordia, P. (1997). Rumor and prediction: Making sense (but losing dollars) in the stock market. *Organizational Behavior and Human Decision Processes*, 71(3), 329-353.
- Dinand, J. P. & Defayolle, M. (1969). Utilisation des potentiels évoqués moyennés pour l'estimation de la charge mentale. *Agressologie*, 1969, 10, 525-533.
- Di Stasi, L. L., Renner, R., Staehr, P., Helmert, J. R., Velichkovsky, B. M., Cañas, J. J., ... & Pannasch, S. (2010). Saccadic peak velocity sensitivity to variations in mental workload. *Aviation, space, and environmental medicine*, 81(4), 413-417.
- Doppelmayr M., Finkenzeller T., Sauseng P. (2008) Frontal midline theta in the pre-shot phase of rifle shooting: differences between experts and novices. *Neuropsychologia*, Vol. 46, No. 5, pp. 1463-1467.
- Dörner, D., & Schaub, H. (1994). Errors in Planning and Decision-making and the Nature of Human Information Processing. *Applied Psychology*, 43(4), 433-453.
- Drew, T., Vö, M. L.-H., & Wolfe, J. M. (2013). The invisible gorilla strikes again: Sustained inattentive blindness in expert observers. *Psychological*

Science, 24, 1848-1853.

- Duchowski, A. T. (2007). *Eye Tracking Methodology: Theory and Practice* (2nd Ed.). Springer, London.
- Dunn, O. J. (1961). "Multiple Comparisons Among Means". *Journal of the American Statistical Association* 56 (293): 52–64. doi:10.1080/01621459.1961.10482090
- Durlach, P. J., Kring, J. P., & Bowens, L. D. (2009). Effects of action video game experience on change detection. *Military Psychology*, 21, 24-39.
- Durso, F. T., Dattel, A. R., Banbury, S., & Tremblay, S. (2004). SPAM: The real-time assessment of SA. A cognitive approach to situation awareness: Theory and application, 1.
- Dyson B.J., Alain C., He Y. (2005), Effects of visual attentional load on low-level auditory scene analysis *Cognitive, Affective & Behavioral Neuroscience*, 5 pp. 319–338
- D’Zmura, M. (1991). Color in visual search. *Vision Research*, 31, 951-966.
- Edwards, W. (1962). Dynamic decision theory and probabilistic information processing. *Human Factors*, 4, 59–73.
- Edworthy, J., Loxley, S., & Dennis, I. (1991). 'Improving auditory warning design: Relationship between warning sound parameters and perceived urgency'. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 33(2), 205-231.
- Endsley, M. R., & Rodgers, M. D. (1994, October). Situation awareness information requirements analysis for en route air traffic control. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* (Vol. 38, No. 1, pp. 71-75). SAGE Publications.
- Endsley, M. R. (1995). Toward a theory of situation awareness, *Human Factors*, 37 (1), 32-64.
- Endsley, M. R. (1998). A comparative analysis of SAGAT and SART for evaluations of situation awareness. In *Proceedings of the Human Factors and Ergonomics Society 42nd Annual Meeting* (pp. 82-86). Santa Monica, CA: The Human Factors and Ergonomics Society.
- Endsley, M. R., Bolte, B., & Jones, D. G. (2003). *Designing for situation awareness: An approach to user centered design*. New York: Taylor & Francis.
- Endsley, M. R. (2012). *Designing for situation awareness: An approach to user-centered design*. CRC Press.
- Eramudugolla, 2005 Eramudugolla, R., Irvine, D. R. F., McAnally, K. I., Martin, R. L., & Mattingley, J. B. (2005). Directed attention eliminates 'change deafness' in complex auditory scenes. *Current Biology*, 15, 1108–1113.
- Edworthy, J., Loxley, S., & Dennis, I. (1991). Improving auditory warning design: Relationship between warning sound parameters and perceived urgency. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 33(2), 205–231.

- Eurocontrol (2003) The Development of Situational Awareness Measures in ATM Systems. (HRS/HSP-005-REP-01). Edition 1.0. Brussels, Belgium: Eurocontrol.
- Fairclough S. H., Venables L. (2006) Prediction of subjective states from psychophysiology: A multivariate approach. *Biological Psychology*, Vol. 71, No. 1, pp. 100 – 110
- Faraday, P., & Sutcliffe, A. (1997). Designing effective multimedia presentations. In *Proceedings of ACM CHI '97*. ACM, New York, pp. 272–279.
- Federal Aviation Administration (1990). Profile of operational errors in the national airspace system: Calendar Year 1988. Washington, DC.
- Fitts P.M., Jones R.E., Milton J.L., Eye movements of aircraft pilots during instrument-landing approaches. *Aeronautical Engineering Review* Vol. 9, No. 2, 1950, p. 24-29
- Folleso K. & Volden F.S. (1993) Lessons learned on test and evaluation methods from test and evaluation activities performed at the OECD Halden reactor project, Institutt for Energiteknikk, Halden, Norway.
- Fougnie, D., & Marois, R. (2007). Executive working memory load induces inattentive blindness. *Psychonomic Bulletin & Review*, 14, 142-147.
- Fournier L. R., Wilson G. F., Swain C.R., Electrophysiological, behavioral, and subjective indexes of workload when performing multiple tasks: manipulations of task difficulty and training. *International Journal of Psychophysiology*, 31, 1999, pp. 129-145.
- Fowler, B. (1994). P300 as a measure of workload during a simulated aircraft landing task. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 36(4), 670-683.
- Franconeri, S. L., Hollingworth, A., & Simons, D. J. (2005). Do new objects capture attention? *Psychological Science*, 16(4), 275-281.
- Frensch P.A., & J. Funke, (Eds.), *Complex problem solving* (pp. 27-63). Hillsdale, NJ: Lawrence Erlbaum.
- Friel, B. M., Thomas, R. P., Raacke, J., & Shanteau, J. (2001). Utilizing CWS to track the longitudinal development of expertise. *Proceedings from the 45th Annual Meeting of the Human Factors and Ergonomics Society*. Minneapolis/St. Paul, MN.
- Friel, B. M., Thomas, R. P., Shanteau, J., & Raacke, J. (2002). CWS applied to an air traffic control simulation task (CTEAM). *Proceedings from the International Symposium on Aviation Psychology*. Columbus, OH.
- Fu S., Parasuraman R. *Event-Related Potentials (ERPs) in Neuroergonomics (2007) - Neuroergonomics, the brain at work*, Oxford series in human-technology interaction ISBN-13 978-0-19-5368665-9
- Gagnon, J. F., Couderc, F., Rivest, M., & Tremblay, S. (2011, April). Capturing the Task Model of Experts in Emergency Response using SYnRGY. In *Proceedings of the 9th International ISCRAM Conference Vancouver* (Vol. 2).

- Gamberale, F. (1972). Perceived exertion, Heart Rate, Oxygen Uptake and Blood Lactate in Different Work Operations. *Ergonomics*, 15, 5, 545-554.
- Gevins, A., Smith, M. E., Leong, H., McEvoy, L., Whitfield, S. & Du, R., (1998). Monitoring working memory load during computer-based tasks with EEG pattern recognition methods. *Human Factors*, 40, 79–91.
- Gianazza, D., Guittet, K. (2006) Selection and evaluation of air traffic complexity metrics 25th DASC
- Gibson James J. (1979), *The Ecological Approach to Visual Perception*, ISBN 978-0-89859-959-6
- Goldberg, J.H. & Kotval, X.P. (1998). A brief review of eye-movement based interface evaluation followed by a discussion of temporally and spatially-based measures.
- Goldberg, J.H. & Kotval, X.P. (1998). Eye movement-based evaluation of the computer interface. In: Kumar, S. (ed.) *Advances in occupational ergonomics and safety*, pp. 529-532. ISO press, Amsterdam.
- Goldberg, H. J., & Wichansky, A. M. (2003). Eye tracking in usability evaluation: A practitioner's guide. In J. Hyöna, R. Radach, & H. Deubel (Eds.), *The mind's eye : Cognitive and applied aspects of eye movement research*, pp. 573-605). Amsterdam Elsevier.
- Gonzalez, C. (2005). Decision support for real-time, dynamic decision-making tasks. *Organizational Behavior and Human Decision Processes*, 96, 142–154.
- Gonzalez, C. (2005). Task workload and cognitive abilities in dynamic decision making. *Human Factors*, 47, 92-101.
- Gonzalez, C., Vanyukov, P., & Martin, M. K. (2005). The use of microworlds to study dynamic decision making. *Computers in Human Behavior*, 21, 273–286. doi:10.1016/j.chb.2004.02.014
- Granlund, R., & Johansson, B. (2004). Monitoring distributed collaboration in the C3 Fire microworld. Dans S. G. Schifflet, L. R. Elliott, E. Salas, & M. D. Coovert (Eds.), *Scaled Worlds: Development, Validation and Application* (pp.37-48). Cornwall, GB: Ashgate.
- Gray, W. D. (2002). Simulated task environments: The role of high-fidelity simulations, scaled worlds, synthetic environments, and laboratory tasks in basic and applied cognitive research. *Cognitive Science Quarterly*, 2, 205–227
- Greenhouse, S. W., & Geisser, S. (1959). "On methods in the analysis of profile data." *Psychometrika*, 24, 95-112.
- Ham, D. H., Youn, W. C., (2001), The effect of presenting functionality abstracted information in fault diagnosis tasks. *Reliability Engineering and System Safety*, 73, 103-119
- Hardin, B., & Goodrich, M. A. (2009, March). On using mixed-initiative control: a perspective for managing large-scale robotic teams. In *Proceedings of the 4th ACM/IEEE international conference on Human robot interaction* (pp. 165-172). ACM.

- Hantula, D. A., & Crowell, C. R. (1994). Behavioral contrast in a two-option analogue task of financial decision making. *Journal of applied behavior analysis*, 27(4), 607-617.
- Hantula, D. A., & Crowell, C. R. (1994). Intermittent reinforcement and escalation processes in sequential decision making: A replication and theoretical analysis. *Journal of Organizational Behavior Management*, 14(2), 7-36.
- Hart, S. G., & Staveland, L. E. (1988). Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. *Advances in psychology*, 52, 139-183.
- Healey, C.G., *Visualization Of Multivariate Data Using Preattentive Processing*, Master's Thesis, 1992, Department of Computer Science, University of British Columbia.
- Hennighausen, K., Remschmidt, H., & Warnke, A. (1994). Visually evoked potentials in boys with developmental dyslexia. *European Child & Adolescent Psychiatry*, 3(2), 72-81.
- Ho, D., & Burns, C. M. (2003, October). Ecological interface design in aviation domains: Work domain analysis of automated collision detection and avoidance. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* (Vol. 47, No. 1, pp. 119-123). SAGE Publications.
- Hockey, G. R. J., Wastell, D. G., & Sauer, J. (1998). Effects of sleep deprivation and user interface on complex performance: a multilevel analysis of compensatory control. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 40(2), 233-253.
- Hodgetts, H., Farmer, E., Joose, M., Parmentier, F., Schaefer, D., Hoogeboom, P., ... & Jones, D. (2005). The effects of party line communication on flight task performance. *Human factors in design, safety, and management*, 1-12.
- Hodgetts, H. M. & Jones, D. M. (2006). Contextual cues aid recovery from interruption: The role of associative activation. *Journal of Experimental Psychology: Learning, Memory & Cognition*, 35, 1120-1132.
- Hodgetts, H. M., Vachon, F., & Tremblay, S. (in press). Background sound impairs interruption recovery in dynamic task situations: Procedural conflict? *Applied Cognitive Psychology*.
- Hollnagel, E. (1993). *Human reliability analysis: context and control* (Vol. 145). London: Academic Press.
- Horvitz, E. (1999, May). Principles of mixed-initiative user interfaces. In *Proceedings of the SIGCHI conference on Human Factors in Computing Systems* (pp. 159-166). ACM.
- Hubel, D. H. (1988). *Eye, Brain, and Vision* New York: Scientific American Library.
- Hughes, R. W., Hurlstone, M. J., Marsh, J. E., Vachon, F., & Jones, D. M. (2013). Cognitive control of auditory distraction: Impact of task difficulty, foreknowledge, and working memory capacity supports duplex-mechanism account. *Journal of Experimental Psychology: Human Perception and Performance*, 39(2), 539.

- Humphrey D.G., Kramer A.F. (1994), Toward a psychophysiological assessment of dynamic changes in mental workload *Human Factors*, 36 pp. 3–26
- Hurter, C., Conversy, S., Jestin, Y., Vinot, J-L. Représentations écologiques de données temporelles: exemples et apports. (IHM'09) In Proceedings of the 21st International Conference on Association Francophone d'Interaction Homme-Machine. ACM, New York, NY, USA, 33-42. doi=10.1145/1629826.1629832
- Hurter, C., Cowan, B. R., Girouard, A., & Riche, N. H. (2012, September). Active Progress Bar: Aiding the switch to temporary activities. In Proceedings of the 26th Annual BCS Interaction Specialist Group Conference on People and Computers (pp. 99-108). British Computer Society.
- Hurter, C.; Ersoy, O.; Fabrikant, S.; Klein, T.; Telea, A., Bundled Visualization of Dynamic Graph and Trail Data, *IEEE Transactions on Visualization and Computer Graphics*, vol.PP, no.99, pp.1,1
- Imbert, J. P., Hodgetts, H. M., Parise, R., Vachon, F., Dehais, F., & Tremblay, S. (2014). Attentional costs and failures in air traffic control notifications. *Ergonomics*, (ahead-of-print), 1-16.
- Imbert, J.-P, Hodgetts, H. M., Parise, R., Vachon, F., & Tremblay, S. (2014). The LABY microworld: A platform for research, system engineering, and training in air traffic control. HFES'14 (Human Factors and Ergonomic Society annual conference)
- Imbert. J-P, Hurter C., Conversy S., (in preparation) The use of characterizing dimensions to design efficient visualizations for Air Traffic Controllers
- Imbert J-P, Bustico A. Implémentation de feedbacks haptiques dans digistrips4 (Note technique Centre d'Etudes de la Navigation Aérienne)
- Iqbal, S. T., Adamczyk, P. D., Zheng, X. S., & Bailey, B. P. (2005, April). Towards an index of opportunity: understanding changes in mental workload during task execution. In Proceedings of the SIGCHI conference on Human factors in computing systems (pp. 311-320). ACM.
- Jacob, R. J. K., & Karn, K. S. (2003). Eye tracking in Human-Computer Interaction and usability research: Ready to deliver the promises. In J. Hyönä, R. Radach, and H. Deubel (Eds.), *The mind's eye: Cognitive and applied aspects of eye movement research* (pp. 573-605). Amsterdam: Elsevier.
- Jambon, F., Vanda ; L., Analyse oculométrique'on-line'avec zones d'intérêt dynamiques: application aux environnements d'apprentissage sur simulateur. In : Ergo'IHM 2012.
- Janisse, M-P. (1977). *Pupillometry : The psychology of the pupillary response*. Hemisphere Pub. Corp.
- Jobidon, M., Tremblay, S., Lafond, D., & Breton, R. (2006). The role of cognition in team functioning: A matter of information sharing and coordination among team members. *Proceedings of Cognitio 2006*, 22-32.
- Jonides, J. (1981). Voluntary versus automatic control over the mind's eye's movement. *Attention and performance IX*, 9, 187-203.

- Johnson R. P300: A model of the variables controlling its amplitude. In: Karrer R, Tueting P, Callaway E, eds. Brain and Information: Ann NY Acad Sci., 425, 1984. p.223-30.
- Johnson, N., Wiegmann, D., Wickens, C. (2006). Effects of Advanced Cockpit Displays on Aviation Pilot's Decision to Continue Visual Flight Rules Flight into Instrument Meteorological Conditions. Proceedings of the Humans Factors and Ergonomics Society 50 th annual Meeting.
- Jordan.C.S. (1992). Experimental Study of the Effect of An Instantaneous Self Assessment Workload Recorder on Task Performance, DRA/TM/CAD5/92011.Defence Research Agency. Portsmouth.
- Just, M. A. & Carpenter, P. A. (1976). Eye Fixations and Cognitive Processes. *Cognitive Psychology*, 8, 441-480.
- Kaber, D. B., & Endsley, M. R. (2004). The effects of level of automation and adaptive automation on human performance, situation awareness and workload in a dynamic control task. *Theoretical Issues in Ergonomics Science*, 5(2), 113-153.
- Kalsbeek, J. W. H. & Ettema, J. H. (1965). L'arythmie sinusale comme mesure de la charge mentale. Communication au IIIe congrès d'Ergonomie de Langue Française, Paris.
- Keidel, W. D.; Spreng, M. (1965). "Neurophysiological Evidence for the Stevens Power Function in Man". *The Journal of the Acoustical Society of America* 38: 191-195.
- Keidel, W. D., & Spreng, M. (1965). Computed audio-encephalograms in man (a technique of "objective" Audiometry). *International Journal of Audiology*, 4(1), 56-60.
- Kemeny, J-G., 1979, 'Final report of the President's commission on the accident at Three Mile Island' Washington DC, USA
- Kinney, G.C., Spahn, M.J. and Amato, R.A. (1977). The human element in air traffic control: Observations and analyses of the performance of controllers and supervisors in providing ATC separation services. METRIEK Division of the MITRE Corporation: MTR-7655:.
- Kirwan, B., Evans, A., Donohoe, L., Kilner, A., Lamoureux, Atkinson, T., & MacKendrick, H. (1997) Human Factors in the ATM System Design Life Cycle. FAA/Eurocontrol ATM R&D Seminar, Paris, France. Internet source.
- Klimesch W., Doppelmayr M., Schimke H., Ripper B. (1997) Theta synchronization and alpha desynchronization in a memory task. *Psychophysiology*, Vol. 34, No. 2, pp. 169-176.
- Koffka, K., Principles of Gestalt psychology. Routledge 1935.
- Koreimann S, Strauß S., Vitouch O. (2009). "Inattentive deafness under dynamic musical conditions". URN:NBN:fi:jyu-2009411269.
- Kramer, A. F.; Trejo L. J.; Humphrey D. (1995) - "Assessment of mental workload with task-irrelevant auditory probes" *Biological Psychology*, Elsevier.

- Koreimann, S., Strauß, S., & Vitouch, O. (2009). Inattentive deafness under dynamic musical conditions.
- Kutas M, McCarthy G, Donchin E. Augmenting mental chronometry: the P300 as a measure of stimulus evaluation time. *Science* 1977; 197: 792-5.
- Lafond, D., Vachon, F., Rousseau, R., & Tremblay, S. (2010). A cognitive and holistic approach to developing metrics for decision support in command and control. In D. B. Kaber & G. Boy (Eds.), *Advances in Cognitive Ergonomics* (pp. 65–73). Danvers, MA: CRC Press.
- Lavie, N. (1995). “Perceptual load as a necessary condition for selective attention”. *Journal of Experimental Psychology: Human Perception and Performance*, Vol 21(3), 451-468.
- Lavie, N. (2005). Distracted and confused?: Selective attention under load. *Trends in Cognitive Sciences*, 9, 75-82.
- Lavie, N. (2010). Attention, distraction, and cognitive control under load. *Current Directions in Psychological Science*, 19(3), 143-148.
- Lee, J. D., & See, K. A. (2004). Trust in automation: Designing for appropriate reliance. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 46(1), 50-80.
- Leplat J., Sperandio J.-C. (1967). La mesure de la charge de travail par la technique de la tâche ajoutée. In: *L'année psychologique*, 1967. Vol. 67, n°1. pp. 255-277.
- Li, X., Çöltekin, A., Kraak, MJ. 2010. Visual exploration of eye movement data using the space-time-cube. In *Proceedings of GIScience'10*, Springer-Verlag, Berlin, Heidelberg, 295-309.
- Loft, S., Smith, R. E., & Bhaskara, A. (2011). Prospective memory in an air traffic control simulation: External aids that signal when to act. *Journal of Experimental Psychology: Applied*, 17, 60-70.
- Macdonald J. S. P.; Lavie N. (2011). “Visual perceptual load induces inattentive deafness” *Attention, Perception, & Psychophysics*, Springer.
- Mack, A., & Rock, I. (1998). “Inattentive blindness: Perception without attention”. In R. Wright (Ed.), *Perception and Attention*, New York, Oxford. U
- Macdonald, J. & Lavie, N. (2011). Visual perceptual load induces inattentive deafness. *Attention Perception & Psychophysics*, 73(6), 1780-1789.
- Maglio, P. P. & Campbell, C.S. (2000). Tradeoffs in displaying peripheral information. In *Proceedings of the SIGCHI conference on Human factors in computing systems (CHI '00)*. ACM, New York, NY, USA, 241-248.
- Martin, C. & Imbert, J-P (2012). Introduction of a more automated environment in En-Route Air Traffic Control. In D. Schaefer (ed) *Proceedings of the SESAR Innovation Days EUROCONTROL*. ISBN 978-2-87497-024-5.
- Martins, E., & Soares, M. (2012). Automation under suspicion—case flight AF-447 Air France. *Work: A Journal of Prevention, Assessment and Rehabilitation*, 41, 222-224.

- Marshall, S. P. (2000). Cognitive workload and pupil dilation: A useful and usable indicator of cognitive activity. In E. Granholm & S. Steinhauer (chairs), *In the mind's eye: A view through the pupil at cognition in psychopathology*. Symposium presented at the 40th Annual Meeting of the Society for Psychophysiological Research, San Diego.
- Matthews, G., Middleton, W., Gilmartin, B. & Bullimore, M. A. (1991). Pupillary diameter and cognitive load. *J Psychophysiology*, 5, 265-271.
- Masson, M. E., & Loftus, G. R. (2003). Using confidence intervals for graphically based data interpretation. *Canadian Journal of Experimental Psychology/Revue canadienne de psychologie expérimentale*, 57(3), 203.
- McCrickard, D. S., Catrambone, R., Stasko, J. T. (2001). Evaluating animation in the periphery as a mechanism for maintaining awareness. *Proceedings of the IFIP TC.13 International Conference on Human-Computer Interaction (INTERACT 2001)* (pp. 148-156). Tokyo, Japan.
- McCrickard D. S., Catrambone R., Chewar C. M., & Stasko J. T. (2003). Establishing tradeoffs that leverage attention for utility: Empirically evaluating information display in notification systems. *International Journal of Human-Computer Studies*, 58, 547-582.
- McCrickard, D. S., & Chewar, C. M. (2003). Attuning notification design to user goals and attention costs. *Communications of the ACM*, 46(3), 67-72.
- Metzger, U., & Parasuraman, R. (2001). The role of the air traffic controller in future air traffic management: An empirical study of active control versus passive monitoring. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 43(4), 519-528.
- Metzger, U., & Parasuraman, R. (2005). Automation in future air traffic management: Effects of decision aid reliability on controller performance and mental workload. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 47(1), 35-49.
- Michaut, G. & Pin, M.C. (1965). Effets sur quelques variables psychophysiologiques de la conduite automobile urbaine, communication au IIIème Congrès d'Ergonomie de langue française, Paris.
- Miller, C. A., & Hannen, M. D. (1999). The Rotorcraft Pilot's Associate: design and evaluation of an intelligent user interface for cockpit information management. *Knowledge-Based Systems*, 12(8), 443-456.
- Mook, D. G. (1983). In defense of external invalidity. *American psychologist*, 38(4), 379.
- Morrison, J. G., Marshall, S. P., Kelly, R. T., & Moore, R. A. (1997). Eye Tracking in Tactical Decision Making Environments: Implications for Decision Support Evaluation. *Third International Command and Control Research and Technology Symposium*, National Defense University, June 17-20.
- Most, S. B., Scholl, B. J., Clifford, E., & Simons, D. J. (2005). What you see is what you set: Sustained inattentive blindness and the capture of awareness. *Psychological Review*, 112, 217-242.

- Most, S. B., Simons, D. J., Scholl, B. J., & Chabris, C. F. (2000). Sustained inattentive blindness: the role of location in the detection of unexpected dynamic events. *Psyche*, 6,
- Neider, M. B., Brotzen, S., & Zelinsky, G. J. (2007). Cutting through the clutter: Searching for targets in evolving realistic scenes. *Journal of Vision*, 7(9), 1056a.
- Newby, E. A., & Rock, I. (1998). Inattentive blindness as a function of proximity to the focus of attention. *Perception*, 27, 1025-1040.
- Ngo, M. K., Pierce, R. S., & Spence, C. (2012). Using multisensory cues to facilitate air traffic management. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 54(6), 1093-1103.
- Nikolic, M. I., Orr, J. M., & Sarter, N. B. (2004). Why Pilots Miss the Green Box: How Display Context Undermines Attention Capture. *The International Journal of Aviation Psychology*, 14,
- Nikolic, M. I., & Sarter, N. B. (2001). Peripheral visual feedback: A powerful means of supporting attention allocation and human-automation coordination in highly dynamic data-rich environments. *Human Factors*, 43, 30–38.
- Norman Donald, *The Design of Everyday Things*, (ISBN 978-0-465-06710-7)
- Oehme O., Wiedenmaier S., Schmidt L., Luczak H., Comparison between the Strain Indicator HRV of Head Based Virtual Retinal Display and LC-Head Mounted Displays for Augmented Reality. In *Proceedings of the Conference WWDU 2002 World Wide Work*, Berchtesgaden, 2002.
- Pachur, T., & Hertwig, R. (2006). On the psychology of the recognition heuristic: retrieval primacy as a key determinant of its use. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 32(5), 983
- Papenmeier, Frank et Huff, Markus. DynAOI: A tool for matching eye-movement data with dynamic areas of interest in animations and movies. *Behavior research methods*, 2010, vol. 42, no 1, p. 179-187.
- Parasuraman, R., Bahri, T., Deaton, J. E., Morrison, J. G., & Barnes, M. (1992). Theory and design of adaptive automation in aviation systems. Catholic University of America Washington dc. Cognitive Science Lab.
- Parasuraman, R., Mouloua, M., & Molloy, R. (1996). Effects of adaptive task allocation on monitoring of automated systems. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 38(4), 665-679.
- Parasuraman, R., & Riley, V. (1997). Humans and automation: Use, misuse, disuse, abuse. *Human Factors*, 39, 230–253.
- Parasuraman, R., Sheridan, T. B., & Wickens, C. D. (2000). A model of types and levels of human interaction with automation. *IEEE Transactions on Systems, Man, and Cybernetics—Part A: Systems and Humans*, 30, 286–297.
- Parasuraman, R., & Byrne, E. A. (2003). Automation and human performance in aviation. *Principles and practice of aviation psychology*, 311-356.

- Parasuraman, R. (2003) “ Neuroergonomics: Research and practice” Theoretical Issues in Ergonomics Science 4 (1-2), 5-20.
- Parasuraman, R., et C. D Wickens. (2008). « Humans: Still vital after all these years of automation ». Human Factors: The Journal of the Human Factors and Ergonomics Society 50 (3): 511.
- Parise, Imbert, Morais, & Alonso (2012). LABY : Un support d’aide à l’évaluation de choix de conception d’IHM pour le contrôle aérien Proceedings of ErgoIHM 2012
- Pawlak, W. S., & Vicente, K. J. (1996). Inducing effective operator control through ecological interface design. International Journal of Human-Computer Studies, 44(5), 653-688.
- Pearson, D., & Sahraie, A. (2003). Oculomotor control and the maintenance of spatially and temporally distributed events in visuo-spatial working memory. The Quarterly Journal of Experimental Psychology Section A: Human Experimental Psychology, 56, 1089-1111.
- Pomplun, M. Sunkara, S. (2003). Pupil dilation as an indicator of cognitive workload in human-computer interaction. In Human-Centred Computing: Cognitive, Social, and Ergonomic Aspects. HCII 2003, Crete, Greece, 542-546.
- Poole, A., Ball, L. J. & Phillips, P. (2004). In search of salience : A response time and eye movement analysis of bookmark recognition. In : People and Computer XVIII, Proceedings of HCI 2004. Springer, London, 363-378.
- Poupyrev I., Okabe M., Maruyama S.. 2004. Haptic feedback for pen computing: directions and strategies. In CHI ‘04 Extended Abstracts on Human Factors in Computing Systems (CHI EA ‘04).
- Prabhakaran, V., Narayanan, K., Zhao, Z., & Gabrieli, J. D. E. (2000). Integration of diverse information in working memory within the frontal lobe. Nature neuroscience,3(1), 85-90.
- Prinzel, L. J., Freeman, F. G., Scerbo, M. W., Mikulka, P. J., & Pope, A. T. (2003). Effects of a psychophysiological system for adaptive automation on performance, workload, and the event-related potential P300 component. Human Factors: The Journal of the Human Factors and Ergonomics Society, 45(4), 601-614.
- Proctor, R. W., & Van Zandt, T. (2011). Human factors in simple and complex systems. CRC press.
- Raabe, M., Rutschmann, R.M., Schrauf, M. & Greenlee, M.W. (2005). Neural Correlates of Simulated Driving: Auditory Oddball Responses Dependent on Workload. In Schmorrow, D. D. (Ed.), Foundations of Augmented Cognition (pp.1067-1076). Mahwah, NJ: Lawrence Erlbaum Associates, Inc. Publishers.
- Rajivan, Venkatanarayanan, & Cooke, 2011
- Rasmussen, J. (1983). Skills, Rules, and Knowledge; Signals, Signs, and Symbols, and Other Distinctions in Human Performance Models. IEEE

Transactions on Systems, Man, and Cybernetics, 13, 257-266.

- Rasmussen, J., Pejtersen, A. M., & Goodstein, L. P. (1994). Cognitive systems engineering. Wiley.
- Rayner, K. (2009). The 35th Sir Frederick Bartlett Lecture: Eye movements and attention in reading, scene perception, and visual search. *Quarterly Journal of Experimental Psychology*, 62, 1457–1506.
- Reason, J. T. (1990). Human error. New York: Cambridge University Press.
- Recarte, M. A., & Nunes, L. M. (2003). Mental workload while driving: effects on visual search, discrimination, and decision making. *Journal of experimental psychology: Applied*, 9(2), 119.
- Reimer, B., & Sodhi, M. (2006). Detecting eye movements in dynamic environments. *Behavior Research Methods*, 38, 667-682.
- Reid, G.B., & Nygren, T.E. (1988). The subjective workload assessment technique: A scaling procedure for measuring mental workload. In P.A. Hancock & N. Meshkati (Eds.), *Human mental workload*, 185–218.
- Remington, R. W., Johnston, J. C. & Yantis, S. (1992). Involuntary attentional capture by abrupt onsets. *Percept. Psychophys.* 51, 279–290.
- Rome F., Cabon P., Favresse A. & Mollard R., Figarol S., Hasquenoph B. (2006), *Human Factors Issues of TCAS: a Simulation Study, HCI-Aero'06*
- Rigas, G., Carling, E., & Brehmer, B. (2002). Reliability and validity of performance measures in microworlds. *Intelligence*, 30(5), 463-480.
- Rieskamp, J., & Hoffrage, U. (2008). Inferences under time pressure: How opportunity costs affect strategy selection. *Acta psychologica*, 127(2), 258-276.
- Rubio, S., Diaz, E., Martin, J., & Puente, J. M. (2004). Evaluation of subjective mental workload: A comparison of SWAT, NASA-TLX, and Workload Profile, *Applied Psychology: An International Review*, 53 (1), 61-86.
- Ryu K., Myung R., Evaluation of mental workload with a combined measure based on physiological indices during a dual task of tracking and mental arithmetic. *International Journal of Industrial Ergonomics*, No. 35, 2005, pp. 991-1009.
- Salvucci, D.D. et Goldberg, J.H. Identifying fixations and saccades in eye-tracking protocols. *Proceedings of the Eye Tracking Research and Applications Symposium (Palm Beach Gardens, FL, USA), 2000*, pp. 71-78.
- Sarter, N. B., & Woods, D. D. (1994). Pilot interaction with cockpit automation II: An experimental study of pilots' model and awareness of the Flight Management System (FMS). *International Journal of Aviation Psychology*, 4, 1-28.
- Sarter, N., & Sarter, M. (2003). Neuroergonomics: opportunities and challenges of merging cognitive neuroscience with cognitive ergonomics. *Theoretical issues in ergonomics science*, 4(1), 142-150.
- Sauer, J., Wastell, D. G., Robert, G., Hockey, J., & Earle, F. (2003). Performance in a complex

- multiple-task environment during a laboratory-based simulation of occasional night work. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 45(4), 657-670.
- Sauseng P., Klimesch W., Freunberger R., Pecherstorfer T., Hanslmayr S., Doppelmayr M. (2006) Relevance of EEG alpha and theta oscillations during task switching. *Experimental Brain Research*, Vol. 170, No. 3, pp 295-301.
- Scerbo, M. (2006). Adaptive automation. *Neuroergonomics: The brain at work*, 239-252.
- Scerbo, M. W., Freeman, F. G., & Mikulka, P. J. (2003). A brain-based system for adaptive automation. *Theoretical Issues in Ergonomics Science*, 4(1-2), 200-219.
- Seip, R.L., Snaed, D., Pierce, E.F., Stein, P., & Weltman, A. (1991). Perceptual responses and blood lactate concentration: effect of training state. *Medecine and Science in Sports and Exercise*, 23, (1), 80-87.
- Sellner, B., Heger, F. W., Hiatt, L. M., Simmons, R., & Singh, S. (2006). Coordinated multiagent teams and sliding autonomy for large-scale assembly. *Proceedings of the IEEE*, 94(7), 1425-1444.
- Sereno, S.; Rayner, K. (2003). Measuring word recognition in reading: eye movements and event-related potentials. *Trends in Cognitive Science* 7 (11): 489 - 493.
- SESAR JU. The ATM Target Concept - D3, Rep. DLM-0612-001-02-00, September 2007.
- Sheridan, T., & Parasuraman, R. (2006). Human-automation interaction. In R. S. Nickerson (Ed.), *Reviews of human factors and ergonomics* (Vol. 1, pp. 89–129) Santa Monica, CA: Human Factors and Ergonomics Society.
- Sibert L. E. & Jacob. R. J. K. 2000. Evaluation of eye gaze interaction. In *Proceedings of the SIGCHI (CHI '00)*.
- Simons, D. J. (2000). Current approaches to change blindness. *Visual Cognition*, 7, 1–15.
- Simons, D. J. & Chabris, C. F. (1999). Gorillas in our midst: sustained inattention blindness for dynamic events. *Perception* 28, 1059–1074.
- Sloubounov S.M., Fukada K., Simon R., Rearick M., Ray W. (2000) neurophysiologic and behavioral indices of time pressure effects on visuomotor task performance. *Brain Research. Cognitive Brain Research*, Vol.9, No. 3, pp 287-298.
- Smallman, H. S., & St. John, M. (2003). CHEX (Change History EXplicit): New HCI concepts for change awareness. *Proceedings of the Human Factors and Ergonomics Society 47th Annual Meeting* (pp. 528–532). Santa Monica, CA: Human Factors and Ergonomics Society.
- Smith M. E., McEvoy L.K., Gevins A. (2002) The impact of moderate sleep loss on neurophysiologic signals during working-memory task performance. *Sleep*, Vol. 25, No. 7, pp. 784-794.

- Spreng, M. (1980) Objective neuro-electrophysiological evaluation of noise effects. In: Noise as a Public Health Problem, Proc. Third Int. Congr., Tobias, J.V., Jansen, G. & Ward, W.D. eds., ASHA-Report, Rockville, Maryland, pp.254-260
- Spreng, M. (1980). "Influence of impulsive and fluctuating noise upon physiological excitations and short-time readaptation". *Scandinavian audiology. Supplementum (Suppl 12): 299–306.*
- Stanton, N. A. (Ed.). (1994). *Human factors in alarm design.* CRC Press.
- St. John, M., Smallman, H. S., & Manes, D. I. (2005). Recovery from Interruptions to a Dynamic Monitoring Task: The Beguiling Utility of Instant Replay. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting September 2005 vol. 49 no. 3 473-477*
- Sutton S., Braren M., Zubin J., John E. R. (1965) Evoked potential correlates of stimulus uncertainty. *Science; 150 : 1187-8.*
- Taylor, R. M. (1990). Situational Awareness Rating Technique(SART): The development of a tool for aircrew systems design. *AGARD, Situational Awareness in Aerospace Operations 17 p(SEE N 90-28972 23-53).*
- Theeuwes, J. (1991). Exogenous and endogenous control of attention: The effects of visual onsets and offsets. *Perception and Psychophysics, 49, 83–90.*
- Theeuwes, J., Belopolsky, A., & Olivers, C. N. L. (2009). Interactions between working memory, attention and eye movements. *Acta Psychologica, 132, 106-114.*
- Thomas, L.C., et C.D. Wickens. (2004). "Eye-tracking and individual differences in off-normal event detection when flying with a synthetic vision system display". *Human Factors and Ergonomics Society Annual Meeting Proceedings, 48:223–227.*
- Thomas, R. P., Willem, B., Shanteau, J., Raacke, J., & Friel, B. M., (2001). Measuring performance of expert controllers: An application to air traffic control *Proceedings of the 45th Annual Meeting of the Human Factors and Ergonomics Society. Minneapolis/St. Paul, MN.*
- Thomas, R. P., Willem, B., Shanteau, J., Raacke, J., & Friel, B. M. (2002). CWS applied to controllers in a high-fidelity simulation of ATC. *Proceedings from the International Symposium on Aviation Psychology. Columbus, OH.*
- Trafton, J. G., Altmann, E. M., Brock, D. P., & Mintz, F. E. (2003). Preparing to resume an interrupted task: Effects of prospective goal encoding and retrospective rehearsal. *International Journal of Human Computer Studies, 58, 582-602.*
- Treisman, A. (1985). Preattentive processing in vision. *Computer vision, graphics, and image processing, 31(2), 156-177.*
- Tremblay, S., Granlund, R., Berggren, P., Turner, P., Holmberg, M., & Jobidon, M-E. (2012). A multiteam international simulation of joint operations in crisis response. *ISCRAM2012, 9th International Conference on Information*

- Systems for Crisis Response and Management. Vancouver, Canada, April 22-25, 2012.
- Tremblay, S., Vachon, F., Lafond, D., & Kramer, C. (2012). Dealing With Task Interruptions in Complex Dynamic Environments Are Two Heads Better Than One?. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 54(1), 70-83.
- Tremblay, S., Parmentier, F. B. R., Hodgetts, H. M., Hughes, R. W., & Jones, D. M. (2012). Disruption of verbal-spatial serial memory by extraneous air-traffic speech. *Journal of Applied Research in Memory and Cognition*, 1, 73-79.
- Todd, J. J., Fougny, D., & Marois, R. (2005). Visual Short-Term Memory Load Suppresses Temporo-Parietal Junction Activity and Induces Inattentive Blindness. *Psychological Science*, 16, 965-972.
- Tremblay, S., Saint-Aubin, J., & Jalbert, A. (2006) Rehearsal in serial memory for visual-spatial information: Evidence from eye movements. *Psychonomic Bulletin & Review*, 13, 452-457.
- Tsang, P.S., & Velazquez, V.L. (1996). Diagnosticity and multidimensional subjective workload ratings. *Ergonomics*, 39(3), 358 – 381.
- Turatto, M., & Galfano, G. (2000). Color, form and luminance capture attention in visual search. *Vision Research*, 40, 1639-1643.
- Uebbing-Rumke M., Gürlük H., Schulze-Kissing D., Adaptive Automation Support for Time-Based Operations in ATC. Third International Air Transport & Operations Symposium ATOS, 2012, Delft, Nederland.
- Ullsperger, P., Freude, G., & Erdmann, U. (2001). Auditory probe sensitivity to mental workload changes—an event-related potential study. *International Journal of Psychophysiology*, 40(3), 201-209.
- Vachon, F., Lafond, D., Vallières, B. R., Rousseau, R., & Tremblay, S. (2011, February). Supporting situation awareness: A tradeoff between benefits and overhead. In *Cognitive Methods in Situation Awareness and Decision Support (CogSIMA)*, 2011 IEEE First International Multi-Disciplinary Conference on (pp. 284-291). IEEE.
- Vachon, F., Vallieres, B. R., Jones, D. M., & Tremblay, S. (2012). Nonexplicit change detection in complex dynamic settings: What eye movements reveal. *Human Factors*, 54, 996-1007.
- Vallieres, B. R., Vachon, F., & Tremblay, S. (submitted). When setting the table for dynamic change detection is not enough.
- Vallières, B. R., Chamberland, C., Vachon, F., & Tremblay, S. (2013). Insights from eye movement into dynamic decision-making research and usability testing. In C. Stephanidis (Ed.), *HCI International 2013 – Posters' Extended Abstracts, Part I. Communications in Computer and Information Science 373* (pp. 169-174). Berlin: Springer-Verlag.
- Vicente, K.J.; Rasmussen, J., "Ecological interface design: theoretical foundations," *Systems, Man and Cybernetics, IEEE Transactions on*, vol.22, no.4,

pp.589,606, Jul/Aug 1992

- Vitevitch, M. S. (2003). Change deafness: The inability to detect changes between two voices. *Journal of Experimental Psychology: Human Perception & Performance*, 29, 333-342.
- Waard, K. Brookhuis, R. van Egmond, and T. Boersema (Eds.), *Human Factors in Design, Safety, and Management*, (pp. 327–338). Maastricht, Netherlands: Shaker publishing.
- Ware, C., Bonner, J., Cater, R., & Knight, W., (1992). Simple animation as a human interrupt. *International Journal of Human-Computer Interaction*, 4, 341–348.
- Ward L.M. (2003) Synchronous neural oscillations and cognitive processes. *Trends in cognitive sciences*, Vol.7, No. 12, 2003, pp. 553-559.
- Warnke, A.; Remschmidt, H.; Hennighausen, K. (1994). “Verbal information processing in dyslexia--data from a follow-up experiment of neuropsychological aspects and EEG”. *Acta paedopsychiatrica* 56 (3): 203–208.
- Waard, K. Brookhuis, R. van Egmond, and T. Boersema (Eds.), *Human Factors in Design, Safety, and Management*, (pp. 327–338). Maastricht, Netherlands: Shaker publishing.
- Wheale, J. L. (1981). “The speed of response to synthesized voice messages”. *British Journal of Audiology*, 15(3), 205-212.
- Weiss, D. J., & Shanteau, J. (2003). Empirical assessment of expertise. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 45(1), 104-116.
- Wickens, C. D. & Alexander, A. L. (2009). Attentional tunneling and task management in synthetic vision displays. *The International Journal of Aviation Psychology*, 19, 182-199.
- Wickens, C. D., Hollands, J., *Engineering psychology and human performance* (3rd ed.), Prentice Hall, NY, 2000.
- Wickens, C. D., Rice, S., Keller, D., Hutchins, S., Hughes, J., & Clayton, K. (2009). False alerts in air traffic control conflict alerting system: Is there a “cry wolf” effect?. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 51(4), 446-462.
- Wilson, G.F. & O'Donnell, R.D. (1988). Measurement of operator workload with the neuropsychological workload test battery. *Human Mental Workload*. (pp.63-115.)
- Wickens, C. D. (2002). Multiple resources and performance prediction. *Theoretical Issues in Ergonomics Science*, 3, 159-177.
- Wilkinson, L., *The grammar of Graphics*. New York: Springer Verlag, 1999.
- Wilson, G. F., Lambert, J. D., & Russell, C. A. (2000, July). Performance enhancement with real-time physiologically controlled adaptive aiding. In *Proceedings of the human factors and ergonomics society annual meeting* (Vol. 44, No. 13, pp. 61-64). SAGE Publications.

- Wilson, G. F., & Russell, C. A. (2007). Performance enhancement in an uninhabited air vehicle task using psychophysiological determined adaptive aiding. *Human factors: the journal of the human factors and ergonomics society*, 49(6), 1005-1018.
- Wolfgang, P. (1994). *Design patterns for object-oriented software development*. Addison-Wesley.
- Woods, D. D., & Roth, E. M. (1988). Cognitive engineering: Human problem solving with tools. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 30(4), 415-430.
- Yamada F. (1998) Frontal midline theta rhythm and eyeblinking activity during a VDT task and a video game: useful tools for psychophysiology in ergonomics. *Ergonomics*, Vol. 41, No. 5,98, pp. 678-688.
- Yantis, S. & Jonides, J. (1990). Abrupt visual onsets and selective attention: voluntary versus automatic allocation. *Journal of Experimental Psychology: Human Perception and Performance*, 16, 121–134.
- Yeh, M., Wickens, C. D., & Seagull, F. J. (1999). Target cuing in visual search: the effects of conformality and display location on the allocation of visual attention. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 41(4), 524-542.
- Zelinsky, G. J. (2008). A theory of eye movements during target acquisition. *Psychological Review*, 115, 787-835.

Table des illustrations

Figure 1: du contrôle en-route au contrôle d'aérodrome.....	20
Figure 2: visualisation radar opérationnelle des années 50 à nos jours	21
Figure 3: contrôleurs et pilotes ; deux métiers étroitement liés	22
Figure 4 : l'étiquette radar	23
Figure 5 : alerte filet de sauvegarde.....	24
Figure 6 : évolution du trafic en France entre 1988 et 2012 ; radial : Paris/province, transversal : province/province (source DGAC/DTA/SDE).....	26
Figure 7 : EEE ERATO Environnement Electronique.....	26
Figure 8 : causes et facteurs contributifs aux incidents 2011- Rapport Sécurité Circulation Aérienne n°52.....	29
Figure 9 : variables visuelles de Bertin.....	35
Figure 10: pouvoir sélectif des variables visuelles	35
Figure 11 : vecteurs de transmission d'informations	36
Figure 12 : exemple de la gelstat sur un scatterplot.....	37
Figure 13 : la comète radar des visualisations radar	39
Figure 14 : design écologique pour la représentation de la vitesse verticale.....	40
Figure 15 : performance en fonction de la charge de travail	52
Figure 16 : Wacom training, interaction directe au stylet.....	62
Figure 17 : Wacom Training en mode indirection.....	62
Figure 18 : MATB (Multi-Attribute Task Battery).....	66
Figure 19 : le micro-monde ATC CTEAM.....	67
Figure 20 : le micro-monde ATC Laby	68
Figure 21 : un Labyrinthe minimaliste édité.....	68
Figure 22 : Champ visuel et vision des couleurs en fonction de l'écart à la zone fovéale.....	78
Figure 23 : oculomètres portés ou posés	79
Figure 24: ondes cérébrales et EEG	81
Figure 25 : les zones du cerveau	82
Figure 26 : ERP N100 et P300.....	83
Figure 27 : position du contrôleur	95
Figure 28 : édition du labyrinthe Saillance	96

TABLE DES FIGURES

Figure 29 : le labyrinthe Saillance.....	97
Figure 30 : menus de saisie niveau et directe.....	97
Figure 31 : notifications: (a) Color (b) ColorBlink (c) BoxAnimation (d) ShadowMask.....	98
Figure 32 : notification Halo.....	99
Figure 33 : plateforme expérimentation Saillance.....	100
Figure 34 : oculomètre Facelab 5.....	101
Figure 35 : le dispositif expérimental Saillance.....	102
Figure 36 : phase de calibration dans Facelab.....	102
Figure 37 : algorithme de calcul de fixation.....	103
Figure 38 : lanceur de la tâche laby expé Bip.....	104
Figure 39 : laby en condition low ou high visual.....	105
Figure 40: architecture expérimentation Bip.....	105
Figure 41 : architecture plateforme.....	110
Figure 42 : synoptique du monitoring de niveau de vol (CFL).....	112
Figure 43 : position de travail Droid équipée d'oculomètres Tobii.....	118
Figure 44 : partie haute : image radar.....	118
Figure 45 : partie basse écran Wacom.....	119
Figure 46 : poste de pseudo-pilote et secteurs adjacents.....	120
Figure 47: une architecture flexible pour l'intégration de l'oculométrie.....	121
Figure 48 : détail d'une étiquette.....	123
Figure 49 : les 3 niveaux de notification.....	124
Figure 50 : la liste des notifications en cours.....	124
Figure 51 : % de saisie correcte tâche principale par type de notification.....	129
Figure 52 : Délai de validation (s) pour chaque type de notification.....	130
Figure 53 : Temps en (s) mis pour percevoir une notification (TPAV).....	131
Figure 55 : durée de fixation.....	133
Table 1 : Synthèse des débriefings.....	134
Figure 56 : dispositif expérimental expérimentation Bip.....	142
Figure 57 : effet de la vitesse et de la charge visuelle sur le % de consignes exactes saisies.....	145
Figure 58 : effet de la vitesse et de la charge visuelle sur le temps de réaction pour saisir les consignes d'altitude.....	145
Figure 59 : effet du type de design sur le taux de notifications validées.....	146

Figure 60 : effet de la vitesse et de la charge visuelle sur le temps de réaction pour valider les notifications	147
.Figure 61 : effet de la vitesse et de la charge visuelle sur le pourcentage de détection des alarmes sonores.....	148
Figure 62 : effet de la vitesse et de la charge visuelle sur le temps de réaction moyen aux alarmes sonores.....	148
Figure 63 : amplitude moyenne de la P300 tous scénarios confondus tâche de contrôle/Laby.....	149
Figure 64 : amplitude de la P300 (μ V) pour les deux types de design	150
Figure 65 : index « Mental Load » du NASA TLX en fonction de la vitesse et de la charge visuelle.	150
Figure 66: questionnaire notifications visuelles	151
Figure 67 : secteur générique.....	158
Figure 68 : durée avant la première fixation sur l'avion notifié.....	162
Figure 69 : durée avant la première interaction sur l'avion notifié.....	163
Figure 70 : durée avant fixation =f(durée avant entrée label).....	163
Figure 71 : durée du problème détecté par l'agent de monitoring	164
Figure 72 : résultats questionnaires expérimentation monitoring.....	166
Figure 73 : résultats questionnaire avec monitoring.....	166
Figure 74 : notifications haptiques avec Digistris 4	174
Figure 75 L'éditeur de parcours LABYedit.....	203
Figure 76 : Palette des objets	204

Annexes

CONCEPTION D'UNE TÂCHE EXPÉRIMENTALE AVEC LABY

La construction de la tâche expérimentale et la manipulation d'objets est effectuée et contrôlée en utilisant l'éditeur LABYedit. Cet éditeur fournit à l'expérimentateur des éléments graphiques pour construire une route, par exemple, des segments de tracé (lignes droites ou courbes), les zones de contrainte (où l'opérateur est tenu de délivrer un commandement, comme un changement de vitesse), des zones de déclenchement de vol interférent, ou d'autres obstacles (voir Figure 1).

La charge de travail peut être modifiée en fonction du nombre d'avions que le participant doit contrôler, le nombre et le type de consignes à saisir pendant une passation, l'activation de tâches complémentaires comme la validation d'un son ou d'une notification visuelle, ainsi que le nombre de vols ou d'obstacles interférents, tandis que la pression du temps peut être augmentée en augmentant la vitesse de l'avion. La charge cognitive peut être établie à un niveau continu en utilisant des demandes de saisies périodiques ou discontinues avec des événements imprévus comme les croisements ou les zones à éviter.

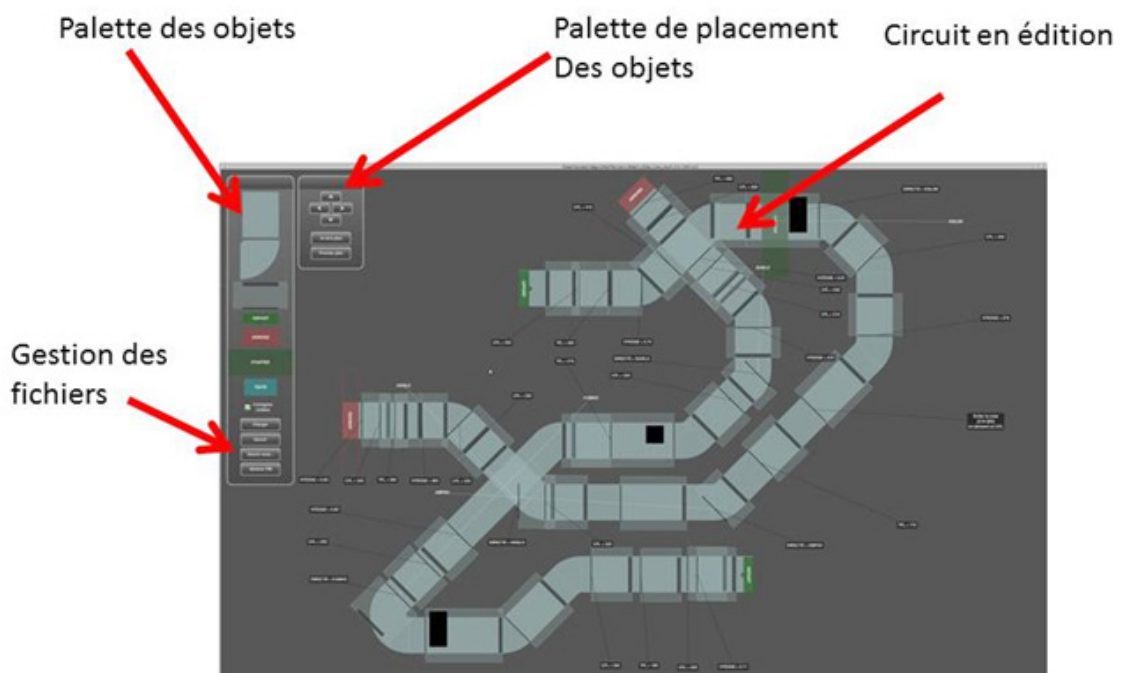


Figure 75 L'éditeur de parcours LABYedit

Les autres caractéristiques du système peuvent également être modifiées à des fins de recherche ou de formation, telles que la taille de l'écran cible, les modalités d'entrées d'information (menus déroulants, pie menu, reconnaissance d'écriture), le moyen de saisie d'information (souris, stylet, doigt) et la présence de notifications visuelles et/ou auditives qui permettent de réaliser des scénarios en contexte de tâches multiples.

La liste actuelle des saisies n'est pas limitée, et d'autres types d'instructions ou de contraintes pourraient être mis en œuvre. L'idée est que l'opérateur se concentre sur la tâche principale qui consiste à guider les avions (lecture des instructions, suivi des avions, anticipation des trajectoires) tout en réalisant des tâches annexes qui peuvent être des distracteurs (obstacles, vols interférents) ou des tâches secondaires (validation de notification visuelles ou sonores), la performance des opérateurs se mesurant sur la réalisation de toutes ces tâches.

En conséquence, la performance sur la tâche de guidage dépendra de facteurs tels que l'efficacité des moyens d'interaction pour la saisie des consignes (périphérique d'entrée, type de menus), ou la conception des notifications (designs, types de sons), selon les différentes configurations de la tâche.

Tous les événements apparaissant lors d'une passation sont déclenchés par la pénétration des avions guidés par l'opérateur dans des zones de déclenchement de consigne. Ces zones de taille variable conditionnent l'apparition ou la disparition d'une consigne si elle n'a pas été renseignée correctement avant que l'avion guidé ne sorte de la zone. La palette des objets permet de concevoir un parcours (Figure 76 : Palette des objets).

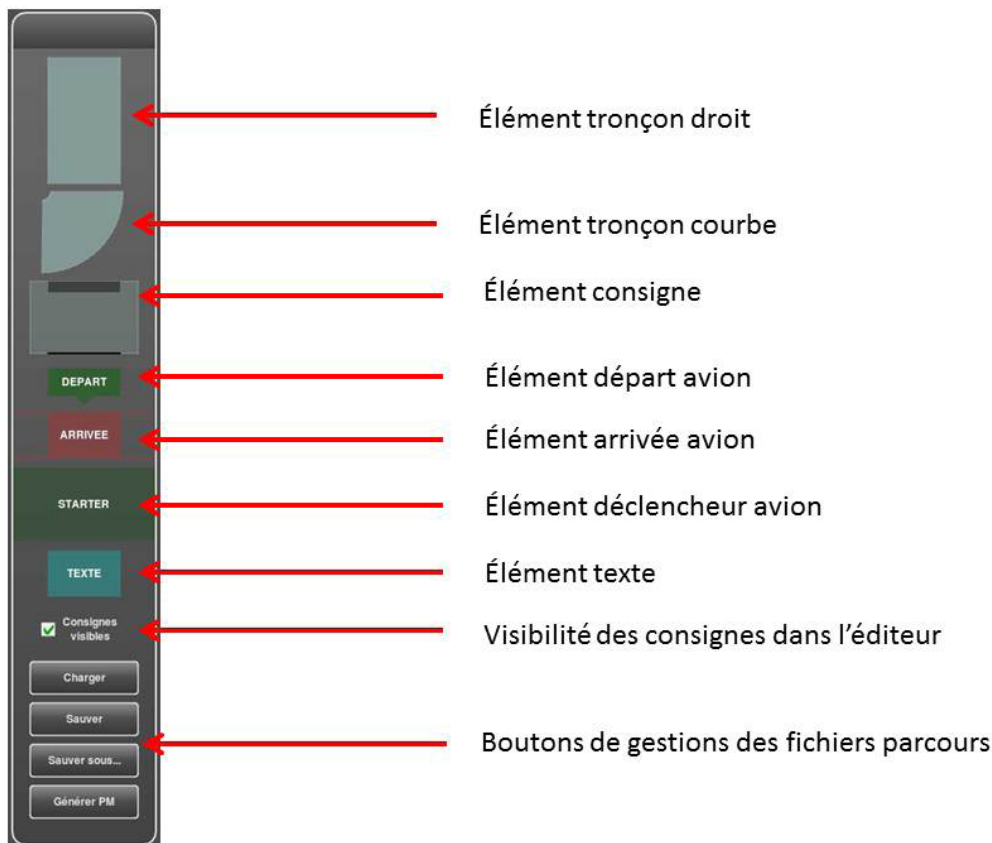


Figure 76 : Palette des objets

L'élément tronçon droit permet de construire des tronçons rectilignes dans un parcours. On peut poser autant de tronçons que nécessaire et changer leur direction (angle). L'élément tronçon courbe permet de construire des tronçons de virages, leur amplitude angulaire peut être modifiée. L'élément consigne permet de créer une zone consigne sur le parcours, sa taille définit la durée pendant laquelle une consigne est présentée à l'opérateur, lorsque celle-ci a été correctement exécutée, la consigne disparaît même si l'avion

n'est pas encore sorti de la zone. L'élément de départ défini le lieu où les avions guidés apparaissent, la zone arrivée provoque leur disparition. Lorsque tous les avions guidés sont passés sur la zone arrivée, le Laby s'arrête. L'élément déclencheur va provoquer l'apparition d'un nouvel avion guidé si plusieurs avions sont prévus pendant une passation. L'élément texte permet d'afficher un texte permanent sur le circuit. Une fois la construction de la route terminée, LabyEdit génère un fichier pour l'application finale (Laby).

QUESTIONNAIRE MONITORING CONDITION : SANS MONITORING

Pour chacune des questions suivantes, veuillez entourer la réponse de votre choix.

Lors de la simulation, j'ai pu réaliser de façon satisfaisante mon activité de contrôle.

En désaccord total	1	2	3	4	5	6	En accord total
--------------------	---	---	---	---	---	---	-----------------

COMMENTAIRES:

Les situations présentées m'ont paru réalistes

En désaccord total	1	2	3	4	5	6	En accord total
--------------------	---	---	---	---	---	---	-----------------

COMMENTAIRES:

La séquence de trafic m'a paru

Très peu chargée	1	2	3	4	5	6	Très chargée
------------------	---	---	---	---	---	---	--------------

COMMENTAIRES :

Durant la simulation, j'ai maintenu un bon niveau de conscience de la situation

En désaccord total	1	2	3	4	5	6	En accord total
--------------------	---	---	---	---	---	---	-----------------

COMMENTAIRES:

Au cours de la simulation, j'ai détecté rapidement les situations problématiques

En désaccord total	1	2	3	4	5	6	En accord total
--------------------	---	---	---	---	---	---	-----------------

COMMENTAIRES:

Au cours de la simulation, j'ai pu maintenir un niveau de sécurité élevé

En désaccord total	1	2	3	4	5	6	En accord total
--------------------	---	---	---	---	---	---	-----------------

COMMENTAIRES:

Sur une échelle de 0 à 5, le niveau de performance que j'ai atteint est

COMMENTAIRES:

QUESTIONNAIRE MONITORING CONDITION : AVEC MONITORING

Pour chacune des questions suivantes, veuillez entourer la réponse de votre choix.

Lors de la simulation, j'ai pu réaliser de façon satisfaisante mon activité de contrôle

En désaccord total	1	2	3	4	5	6	En accord total
--------------------	---	---	---	---	---	---	-----------------

COMMENTAIRES:

Les situations présentées m'ont paru réalistes

En désaccord total	1	2	3	4	5	6	En accord total
--------------------	---	---	---	---	---	---	-----------------

COMMENTAIRES:

La séquence de trafic m'a paru

Très peu chargée	1	2	3	4	5	6	Très chargée
------------------	---	---	---	---	---	---	--------------

COMMENTAIRES :

Durant la simulation, j'ai maintenu un bon niveau de conscience de la situation

En désaccord total	1	2	3	4	5	6	En accord total
--------------------	---	---	---	---	---	---	-----------------

COMMENTAIRES:

Le monitoring m'a permis d'améliorer ma conscience de la situation

En désaccord total	1	2	3	4	5	6	En accord total
--------------------	---	---	---	---	---	---	-----------------

COMMENTAIRES:

Au cours de la simulation, j'ai détecté rapidement les situations problématiques

En désaccord total	1	2	3	4	5	6	En accord total
--------------------	---	---	---	---	---	---	-----------------

COMMENTAIRES:

Le monitoring m'aide à maintenir un niveau de sécurité élevé

En désaccord total	1	2	3	4	5	6	En accord total
--------------------	---	---	---	---	---	---	-----------------

COMMENTAIRES:

Les informations fournies par le monitoring sont suffisantes pour identifier le contenu de la situation

En désaccord total	1	2	3	4	5	6	En accord total
--------------------	---	---	---	---	---	---	-----------------

COMMENTAIRES:

Le monitoring m'a perturbé durant la simulation (interruption de tâche nuisible par exemple)

En désaccord total	1	2	3	4	5	6	En accord total
--------------------	---	---	---	---	---	---	-----------------

COMMENTAIRES:

Le monitoring s'intègre à ma méthode de travail

En désaccord total	1	2	3	4	5	6	En accord total
--------------------	---	---	---	---	---	---	-----------------

COMMENTAIRES:

Sur une échelle de 0 à 5, le niveau de performance que j'ai atteint est

COMMENTAIRES:

QUESTIONNAIRE ALERTES VISUELLES

Questionnaire sur les alertes visuelles

Identifiant du participant : _____

1. A votre avis, combien y avait-il d'alertes (sur les avions périphériques) dans chaque scénario ? (Dans quel(s) scénario vous a-t-il semblé qu'il y en avait le plus ? le moins ?)

Numéro du scénario	Nombre d'alarmes
1	
2	
3	
4	

2. Il est possible que vous ayez manqué de signaler des alertes visuelles. Remplissez le tableau suivant de manière à décrire au mieux les situations que vous avez rencontrées, avec soit :

J : Presque jamais

P : Parfois

S : Souvent

T : Presque toujours

ainsi que le(s) numéro des scénario concernés.

	J, P, S ou T ?	N° scénario concerné(s)
Je n'ai pas traité l'alerte parce que je n'ai pas eu temps		
Je n'ai pas traité l'alerte parce que j'ai oublié		
Je n'ai pas traité l'alerte parce que je ne l'ai pas vue		
J'ai traité l'alerte		
Autre : (Spécifiez)		

3. Concernant les alertes visuelles, complétez le tableau suivant selon ce qui décrit le mieux votre situation :

J : Presque jamais
P : Parfois
S : Souvent
T : Presque toujours

	J, P, S ou T ?
Je remarquais immédiatement les alertes dès leur apparition	
Je les voyais au bout de quelques secondes	
J'étais obligé de faire attention aux avions périphériques pour voir ces alertes	
Autre : (Précisez)	

4. Avez-vous des remarques concernant les signaux visuels, leur détection ou leur traitement ? (Difficultés, rapidités, ou au contraire facilités...)

MERCI DE VOTRE PARTICIPATION

Résumé

Le contrôle aérien rentre dans une phase de profondes mutations liées à l'augmentation du trafic et à l'évolution des outils permettant d'assurer le service. L'augmentation du nombre d'avions gérés par le contrôleur implique un nombre beaucoup plus important qu'auparavant d'informations à traiter et mémoriser ; or une part importante des causes d'incidents est déjà due à des problèmes de perception et de mémorisation des informations amplifiés par la taille des écrans de contrôle. Ceux-ci sont particulièrement critiques dans le cas de la perception des alarmes et des avertissements donnés par le système sur la visualisation radar. Le design actuel de ce type d'alertes en France qui n'utilise que la couleur, fait aujourd'hui débat et il a été recommandé dans un bulletin de sécurité aérienne de le faire évoluer. La perception des informations nécessaires à l'établissement d'une bonne conscience de la situation aérienne des contrôleurs est au centre de cette thèse. L'objet principal de ce travail est d'améliorer la conscience de la situation des contrôleurs en s'assurant que les éléments pertinents à leur disposition dont les alarmes sont perçus dans des délais conformes à leur importance et que les actions nécessaires qui en découlent sont bien réalisées. Pour répondre à cette problématique, notre travail s'est porté sur deux axes. Le premier consiste à étudier la tâche des contrôleurs de façon à analyser les informations nécessaires à la réalisation de certains objectifs et proposer un agent de suivi de la tâche qui pourrait les épauler. Le second porte sur la perception des notifications, plus particulièrement en vision périphérique, et la conception de designs propres à améliorer leur perception ainsi que l'étude de leur impact sur la réalisation de la tâche. Grâce à une approche holistique basée sur l'utilisation d'un micro-monde ATC (Laby) et l'utilisation de capteurs physiologiques, nous avons pu évaluer plusieurs designs de notifications. Deux expérimentations ont été conduites, la première utilisant de l'oculométrie, visait à mesurer le pouvoir attentionnel de cinq designs et leur impact sur la réalisation de la tâche. La seconde, en utilisant des données neurophysiologiques, visait à mesurer l'impact de deux designs sur la charge de travail. Les designs évalués ont servi à concevoir un système de notification intégré dans une nouvelle supervision radar qui a été couplé à l'agent de suivi de la tâche. La dernière expérimentation conduite durant ces travaux avait pour objectif d'évaluer l'impact de cette nouvelle position de contrôle dont le design est orienté conscience de la situation sur la détection de problèmes critiques. Les résultats obtenus montrent l'intérêt de ces nouveaux outils et leur impact positif sur la réalisation de la tâche des contrôleurs ainsi que la nécessité d'évaluer dans un contexte expérimental contrôlé les caractéristiques des designs de notifications pour les visualisations de supervision.

Mots Clés :

contrôle aérien, saillance, notifications, adaptabilité, oculométrie, neuroergonomie

Abstract

Air traffic control is undergoing a great change due to the increase of traffic and the evolution of the control tools. The greater number of aircrafts managed by the controller implies a much greater load of information to deal with and memorize than before. A significant part of accidents' causes is already due to the problem of information perception and memorization which is worsened by the size of the control screens which are particularly critical concerning the perception of alarms and warnings displayed on the radar image. The current design in France of those types of alerts which rely on color is controversial and an Aviation Safety bulletin recommended it should be improved. The perception of information contributing to a satisfactory situation awareness by the air traffic controllers is central to this thesis which aims at making sure alarms and relevant information are detected early enough and according to priorities and that the corresponding actions are actually performed. To answer the issue at stake our first work focus dealt with the study of the controllers' task so as to analyze the required information in order to achieve certain objectives and provide a task monitoring agent that could support them. The second work focus dealt with the notifications perception, more specifically in the peripheral vision and the conception of designs such as to improve their perception and the study of their impact on the performance of the task. Thanks to a holistic approach based on the use of an ATC microworld (Laby) and the use of physiological sensors we managed to assess several notification designs. Several experiments have been conducted, one using eye-tracking aimed at measuring attention capacity of five designs and their impact on the performance of the task. The second used neurophysiologic data and aimed at measuring the impact of two designs on the work load. The assessed designs enabled to conceive a notification system integrated in a new radar supervision together with the task monitoring agent. The last experiment conducted in the course of this work aimed at measuring the impact of this new position of control whose design is situation awareness oriented, on the detection of critical problems. The obtained results show the relevance of these new tools and their positive impact on the performance of the task by controllers as well as the need to assess, in a controlled experimental context, the characteristics of notification designs for supervision visualization.

Keywords :

air traffic control, saliency, notifications, adaptability, eye tracking, neuroergonomics