

# IHM et évaluation ergonomique : une approche pragmatique des statistiques

Marion Wolff

Univ. Paris Descartes/LATI  
UFR Biomédicale  
45 rue des Saints-Pères, 75 006 Paris  
(+33) 142 862 074

[marion.wolff@parisdescartes.fr](mailto:marion.wolff@parisdescartes.fr)

## RESUME

L'originalité de ce cours est d'organiser la présentation autour d'études de cas concernant des évaluations IHM *WIMP* (*Windows, Icons, Menus, Pointing device*) et *POST-WIMP*, et de prendre en compte les différents types de données recueillies sur le terrain, plutôt qu'autour des concepts.

## Mots clés

Étude de cas, protocoles expérimentaux, analyse de données

## ABSTRACT

This unconventional course is organized around the presentation of *WIMP* (*Windows, Icons, Menus, Pointing device*) and *Post-WIMP*, and takes into account the different types of field data collected, rather than around the concepts.

## Categories and Subject Descriptors

J.4 [Computer application] Social and Behavioral Sciences – Psychology.

## General Terms

Measurement, Experimentation, Human Factors, Verification.

## Keywords

Case Study, experimental Protocol, Data Analysis.

## 1. INTRODUCTION

Les types de données à analyser varient en permanence selon les procédures mises en place par le chercheur qui se pose un certain nombre de questions auxquelles il veut tenter de répondre. Ces procédures peuvent par conséquent utiliser différents protocoles : observations, expérimentations, débriefings, questionnaires, qui

peuvent être analysés séparément ou conjointement.

Les données issues de ces protocoles sont, d'un point de vue statistique, de différentes natures : données univariées numériques et qualitatives, données bivariées numériques et qualitatives, ou mixtes, à mesures répétées, données multivariées ...

Ainsi pour chaque type de données qui sera résumé sous l'appellation de *variable statistique* (qualitative/quantitative ou numérique) et pour chaque type de question posée sur ces données, sont présentées l'ensemble des procédures nécessaires à l'analyse complète de ces données [1]: des procédures descriptives (statistiques et graphiques) et exploratoires de type multivarié [2] [3] aux procédures inférentielles (les tests). Les concepts et procédures sont ainsi mis en contexte.

En organisant le cours autour des types de données, il devient plus aisé de confronter les indices, les méthodes (quel est l'ensemble des méthodes disponibles pour analyser tel type de données ?) et de mettre en avant les aspects méthodologiques (pourquoi utiliser telle méthode plutôt que telle autre ?). Ces aspects méthodologiques étant peu pris en compte dans la plupart des « cours de stats » qui sont actuellement plutôt centrés sur le calcul de différentes statistiques. Il est à noter que les analyses des verbalisations peuvent également faire l'objet d'analyses statistiques assez sophistiquées ; les procédures sont les mêmes que celles utilisées pour les analyses des protocoles numériques multivariés quels qu'ils soient [4] [5] [6]

Pour ce cours nous proposons trois études de cas, une concernant des évaluations IHM *WIMP* (*Windows, Icons, Menus, Pointing device ou fenêtres, icônes, souris, pointeur*) et deux *post-WIMP*. Les IHM *post-WIMP* mettent en œuvre des techniques d'interaction innovantes qui vont au-delà des interfaces classiques *WIMP*. Ces techniques innovantes incluent entre autres l'interaction à deux mains (surfaces multi-touch) ou l'interaction à plusieurs utilisateurs simultanément.

## 2. Évaluation IHM WIMP : identification d'icônes pour un équipement militaire

Cette évaluation a fait l'objet d'une communication [7]. Son objectif est de déterminer parmi un ensemble de 72 icônes celles qui sont spontanément identifiées correctement et celles qui engendrent des erreurs, même après un court apprentissage. Les sujets de cette évaluation sont 78 militaires (61 grenadiers-voltigeurs et 17 gradés : officiers et sous-officiers), volontaires pour passer cette expérience et préalablement informés des objectifs et des conditions de passation. Les icônes testées se répartissent en 12 familles de 6 icônes, chacune correspondant à un certain contexte de transaction (par exemple : préparation de mission, suivi de mission, carte, tir, armes, etc.). Ce contexte est indiqué lors la présentation de chacune des icônes.

L'expérience, dont la passation est entièrement automatique, se déroule en 3 phases : identification spontanée (association d'une signification à une icône présentée), phase d'apprentissage (répétition), puis identification inverse (c'est-à-dire association d'une icône à une signification présentée). Pour chaque phase, l'ordre des présentations des stimuli est aléatoire. Le programme enregistre toutes les réponses des sujets (bonnes et mauvaises). On désire analyser d'une part les performances des trois groupes de sujets lors des trois phases, et d'autre part distinguer dans ce lot d'icônes celles qui sont spontanément reconnues de celles qui résistent à l'apprentissage.

Plusieurs analyses vont ainsi être mises en œuvre dans un souci également de généralisation puisque ce dispositif doit être utilisé par l'ensemble des militaires français.

*De quelle nature sont ces variables ? Quelles statistiques mettre en œuvre pour répondre aux questions posées ? Comment généraliser ces résultats à l'ensemble de la population concernée à partir de cet échantillon ? Quelles précautions doit-on prendre ? Comment identifier les lots d'icônes non pertinentes ?*

C'est à partir de ces questions et de l'identification des variables que l'on pourra mettre en œuvre les analyses pertinentes.

### 3. Évaluation IHM Post-WIMP et étude des multisollicitations

Le protocole de test utilisateur a été mis en œuvre par la PEPSS (Plateforme d'Évaluation, de Prototypage et de teSt d'usageS)<sup>1</sup>. Cette étude simule les tâches qu'un opérateur aurait à effectuer en parallèle (multisollicitations) [8] [9] : situation d'accueil de clientèle/visiteurs et/ou de surveillance d'événements (situation-type du guichet d'accueil où l'opérateur doit fournir des renseignements multiples tout en effectuant des tâches de surveillance).

Dans ce protocole, la tâche principale du sujet est d'effectuer une surveillance d'événements informatisés (dérives de curseurs, contrôle de voyants) via un test attentionnel. Le test utilisé pour la PEPSS est inspiré de la *MultiAttribute Task Battery* (MATB) [10], fréquemment employée dans les recherches en aéronautique pour mesurer la charge mentale des opérateurs. Durant ce test, en plus de cette tâche attentionnelle, un certain nombre de distracteurs-stresseurs interviennent régulièrement, venant perturber l'opérateur dans sa tâche. Le protocole mis en œuvre pour ce test de multisollicitations se déroule en trois phases, chacune d'une durée de 6 minutes. Durant chacune de ces phases, trois types de distracteurs vont intervenir pendant l'exécution de la tâche principale : sollicitations de personnes venant demander la localisation d'un individu, alarmes sonores et visuelles à stopper, demande de codes d'accès via un talkie-walkie. Chaque phase débute de la même manière : la première minute le sujet n'effectue que la tâche principale, puis lors de chaque minute suivante les distracteurs interviennent.

Ces trois phases sont de difficulté croissante : lors de la phase 1, les sollicitations interviennent l'une après l'autre, elles sont ensuite couplées en phase 2, pour apparaître enfin en simultané et en continu lors de la phase 3. Après chaque phase sont données aux sujets : des échelles d'évaluation quant au niveau de leur performance, de confiance dans cette évaluation, de difficulté de

la tâche, de l'effort fourni, ainsi que des échelles d'auto-évaluation des émotions « *Self-Assessment Manikin scale (SAM)* » [11] se déclinant en trois échelles illustrées par des pictogrammes que le sujet doit cocher pour indiquer son état émotionnel : plaisir, éveil, et dominance de la situation.

Les sujets évalués sont en position debout (situation d'accueil du public appliquée de plus en plus fréquemment en secteur public, tertiaire ou encore de surveillance dans l'industrie [12]), face à un pupitre sur lequel est posé un ordinateur. Ils sont équipés avec différents capteurs permettant les relevés physiologiques (réponse électrodermale/fréquence cardiaque), et leurs performances (en termes d'omissions de réponses aux tâches de surveillance ou de temps de réponse) sont directement enregistrés via le programme informatique. À la fin du test, un débriefing est proposé au sujet lors duquel les verbalisations sont enregistrées.

Pour les sujets testés on a ainsi relevé leurs :

- Sc : Skin conductance (conductance électrodermale) mesurée en microsiemens ( $\mu S$ )
- Fc : Fréquence cardiaque
- Omissions (échelle MAT/Alarmes)
- Temps de réponse (échelle MAT/Alarmes)
- Réponses aux échelles analogiques (notes de 0 à 10) pour les évaluations subjectives (Performance, Confiance, Difficulté, Effort) et les évaluations SAM.
- Verbalisations

*Questions initiales : 1) l'expérimentation est-elle bien adaptée aux besoins de l'étude IHM (vérification que les phases sont bien de difficulté croissante et induisent des réactions différentes aux sujets) ?*

*2) L'analyse de l'ensemble des variables permet-elle de dégager des profils de comportements, de stratégies ?*

### 4. Évaluation IHM Post-Wimp : apports des surfaces interactives

Cette étude exploratoire est actuellement en cours. Elle est issue des travaux menés dans le domaine des surfaces interactives et est menée à l'aide de la table *Tangisense* utilisant des objets tangibles [13] [14] et [15].

Nous proposons une application permettant la reconnaissance et l'apprentissage des couleurs aux enfants (2 à 5 ans environ selon le scénario). Les scénarii sont basés sur le programme d'enseignement dans les écoles maternelles. Nous avons demandé à un professeur des écoles d'imaginer un ou plusieurs scénarii en utilisant une table interactive et un ensemble d'objets sans donner de limite ou de contrainte. Il en est ressorti une application simple où les enfants doivent replacer un ensemble d'objets ayant "perdu leur couleur" dans le cadre de la couleur approprié (i.e. un citron en "noir et blanc" à venir placer dans le cadre de couleur jaune). Pour cela, l'enfant dispose d'un ensemble d'objets (petits cubes portant chacun une image en noir et blanc). L'enfant doit ensuite déterminer ce que représente l'image, y associer la couleur appropriée et venir placer l'objet dans le cadre coloré sur la table interactive. Une fois les objets placés, l'enfant est invité à vérifier ses choix à l'aide d'un personnage représentant un magicien. Le magicien lance alors la procédure de vérification, et annonce à l'enfant ses éventuelles erreurs ou bonnes réponses. Toutes les interactions enfant-enfant, enfant-maîtresse, enfant-objets sont filmées et des grilles d'observations ont été élaborées dans l'objectif d'analyser ces interactions.

<sup>1</sup> PEPSS : Estia, Technopôle Izarbel, 64210 Bidart & Univ. Paris Descartes/LATI axe Ergonomie, 45 rue des Saints-Pères, 75006 Paris.

*Quelles questions peut-on se poser ? De quel(s) type(s) sont les variables à analyser ? Quelles analyses peut-on mettre en œuvre ?*

Pour ces trois études de cas, ce cours s'organise systématiquement autour des types de données et des méthodes à utiliser en fonction des questions initiales que le chercheur se pose.

## 5. REFERENCES

- [1] Corroyer, D., & Wolff, M. 2003. L'analyse statistique des données pour la Psychologie : concepts et méthodes de base. Paris, Armand Colin.
- [2] Wolff, M. 2003. Apports de l'analyse géométrique des données pour l'analyse de l'activité. In J.-C. Sperandio & M. Wolff (Eds.), *Formalismes de modélisation pour l'analyse du travail et l'ergonomie* (pp. 195-227). Paris, PUF.
- [3] Wolff, M. (2010). Basics of geometric data analysis and PCA. *World Engineering Anthropometry Resource Association (WEAR)*. New-Zeland, Auckland, February.
- [4] Wolff, M., Burkhardt, J.M., & De la Garza C. 2005. Analyse exploratoire de « points de vue » : une contribution pour outiller les processus de conception. *Le Travail Humain*, 68, 3, 253-284.
- [5] Lu, J.M., Wang, M.J., & Mollard, R. 2010. The effect of arm posture on the scan-derived measurements.- *Applied Ergonomics*, 41(2), 236-241.
- [6] Cappe, E., Wolff, M., Bobet, R., & Adrien, J.-L. Sous presse. Quality of life: a key variable to consider in the evaluation of adjustment in parents of children with autism spectrum disorders and in the development of relevant support and assistance programs. *Quality of Life Research*.
- [7] Sperandio, J.-C., Wolff, M., & Todeschni, L. 2003. Evaluation d'icônes utilisées comme base d'une communication médiatisée en milieu militaire. In J.M.C. Bastien (Ed.), *Deuxièmes journées d'étude en psychologie ergonomique : Epique' 2003* (pp. 15-28). Rocquencourt, France : INRIA.
- [8] Couture, N., Wolff, M., Mollard, R., & Todeschni, L. 2010. Etude du comportement et des émotions dans un contexte de multisollicitations : approche exploratoire avec la Plateforme d'Evaluation, de Prototypage et de tests d'usageS (PEPSS). *ACM Ergo'IA 2010, October 13-15, 2010* (pp. 67-74). Biarritz, France. ACM : ISBN 978-1-4503-0273-9
- [9] Mollard, R., Wolff, M., Couture, N., & Clay, A. (à paraître). Développement d'une plateforme d'évaluation personnalisable et adaptable pour l'étude du comportement émotionnel en situation de multisollicitations. *Le Travail Humain*.
- [10] Comstock, J. L., & Arnegard, R. J. 1992. *The Multiattribute Task Battery for human operator workload and strategic behavior research*. Technical Report 104174. Hampton, VA:NASA Langley Research Center.
- [11] Gil, S. 2009. Comment étudier les émotions. *Revue électronique de psychologie sociale*, 4.
- [12] Mollard, R., Alban, G., & Wolff, M. 2010. Transformation du travail et prévention des risques : aménagement des espaces d'activités et conception des postes de travail dans les bureaux de poste. *ACM Ergo'IA 2010, October 13-15, 2010* (pp. 193-199). Biarritz, France. ACM : ISBN 978-1-4503-0273-9
- [13] Kubicki, S., Lepreux, S., Kolski, C., and Caelen, J. 2010. Towards New Human-Machine Systems in contexts involving interactive table and tangible objects. *11th IFAC/IFIP/IFORS/IEA Symposium on Analysis, Design, and Evaluation of Human-Machine Systems*. Valenciennes, France.
- [14] Kubicki, S., Lepreux, S., Kolski, C., Perrot, C., & Caelen, J. 2009. TangiSense: présentation d'une table interactive avec technologie RFID permettant la manipulation d'objets tangibles et traçables. *IHM '09: Proceedings of the 21st International Conference on Association Francophone d'Interaction Homme-Machine*, 351-354, ACM Press.
- [15] Kubicki, S., Lepreux, S., Lebrun, Y., Dos Santos, P., Kolski, C., & Caelen, J. 2009. New Human-Computer Interactions using tangible objects: application on a digital tabletop with RFID technology. In J. A. Jacko (Ed.), *13th Int. Conf., HCI International, vol. 5612, LNCS*. San Diego, CA.