



HAL
open science

Experimental validation of persuasive criteria for CHI: analysis of their effectiveness to inspect persuasion in interfaces

Eric Brangier, Nemery Alexandra, Schmitt Séverine

► To cite this version:

Eric Brangier, Nemery Alexandra, Schmitt Séverine. Experimental validation of persuasive criteria for CHI: analysis of their effectiveness to inspect persuasion in interfaces. *Journal d'Interaction Personne-Système*, 2015, Volume 4, Number 1, Special Issue: PISTIL (1), pp.85-103. 10.46298/jips.1299 . hal-01207452

HAL Id: hal-01207452

<https://hal.science/hal-01207452v1>

Submitted on 30 Sep 2015

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

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

Validation expérimentale des critères de persuasion interactive : analyse de leur efficacité à inspecter la persuasion dans des interfaces

Eric BRANGIER

Université de Lorraine - Metz.

PERSEUs EA 7312, Psychologie Ergonomique et Sociale pour l'Expérience Utilisateurs - BP 30309 Île du Saulcy - 57006 Metz (France).

Eric.Brangier@univ-lorraine.fr

Alexandra NEMERY

Sage - 10 Rue Fructidor, Paris; et PERSEUs, Université de Lorraine - Metz.

Séverine SCHMITT

SAMETH 51, 39 Avenue Hoche, 51100 Reims et PERSEUs, Université de Lorraine - Metz.

Experimental validation of persuasive criteria for CHI: analysis of their effectiveness to inspect persuasion in interfaces

Abstract: The objective of this article is to present a study that has been made about the use of the interactive persuasion criteria grid in order to identify persuasive elements in the interfaces. Two groups have been created, "criteria" and "control". Both should identify persuasive elements of two interfaces: social network (Facebook) and e-commerce (Cdiscount). During the second phase of the experiment, the "criteria" group had an access to the grid. 27 participants were recruited for this study. The results show that the participants have found more persuasive elements with the grid and that the identification performance was better for the e-commerce website. The average evaluation performance is significantly different ($p = 0.000$) on the two sites. On Facebook, participants' performance was 9.7% ("control" group) and 23.8% then ("test" group); and for Cdiscount respectively 21.7% ("control" group) and 43.6% ("test" group). The set of guidelines really allows to increase the identification of forms of persuasion in interfaces.

Key words: Persuasive technology, Heuristic inspection, Ergonomics criteria.

Résumé. L'objectif de cet article est de présenter une étude portant sur l'utilisation de la grille de critères de persuasion interactive dans le but d'identifier les éléments persuasifs dans les interfaces. Deux groupes ont été formés, « critère » et « contrôle ». Tous deux ont dû évaluer les éléments persuasifs de deux interfaces : réseau social (Facebook) et site de e-commerce (Cdiscount). Lors de la deuxième passation le groupe « critère » avait accès à la grille. 27 participants ont été recrutés pour cette étude. Les résultats indiquent que les participants ont trouvé plus d'éléments persuasifs avec la grille et que la performance en termes d'identification a été plus forte pour le site d'e-commerce. La performance moyenne d'évaluation est bien significativement différente ($p = 0,000$) sur les deux sites web. Concernant le site Facebook, la performance des participants du groupe « contrôle » est de 9,7% des éléments persuasifs identifiés en moyenne et de 23,8% pour le groupe « critère ». Pour le site Cdiscount, la performance d'évaluation du groupe « critère » est de 43,6% des éléments persuasifs trouvés en moyenne, et de 21,7% pour le groupe « contrôle ». L'expérience souligne la performance de la grille dans l'identification des formes de persuasion dans les interfaces.

Mots-clés : Persuasion technologique, critères ergonomiques, influence, évaluation.

Édité par Pr. J.M.C. Bastien (Université de Lorraine) & Pr. G. Calvary (Univ. Grenoble Alpes)

1 INTRODUCTION

De plus en plus de systèmes techniques cherchent à influencer l'utilisateur, à le persuader, à le manipuler dans ses actes et ses pensées. Afin d'appréhender les dimensions persuasives de l'interface, différentes approches ont été développées et parmi elles, l'identification de critères repérables dans les interfaces a donné lieu à des recherches récentes. Il s'agit de travaux qui reprennent les principes de l'inspection heuristique fondée sur les critères de praticité et qui proposent des critères pouvant s'appliquer à l'évaluation de la persuasion dans les interfaces. Mais, si ces critères de persuasion correspondent à des élaborations conceptuelles intéressantes, ont-ils pour autant une valeur dans l'analyse de la persuasion présente dans les écrans qui nous entourent ? Ces critères de persuasion sont-ils utilisables, performants, appropriés ? Fournissent-ils des données inédites qui ne seraient pas vues sans eux ? Seraient-ils vraiment utiles à l'inspection heuristique ?

L'objectif de cet article est d'appréhender la capacité des critères de persuasion à aider à produire un recueil pertinent d'informations sur les interfaces. Aussi, sur la base d'une expérience avons-nous évalué la qualité des critères de persuasion interactive proposés par Némery et Brangier (2014). Nous cherchons donc à tester l'efficacité de cette grille de critères ergonomiques interactifs. Pour y parvenir, nous nous sommes prioritairement basés sur les travaux de Bastien et Scapin (1995) concernant la validation des critères ergonomiques pour l'évaluation des IHM.

Dans une première partie nous développerons l'idée d'une grille de critères pour évaluer la persuasion, puis nous présenterons le protocole relatif à l'expérience de mesure de la capacité des critères à aider à l'inspection heuristique. La troisième partie sera consacrée à l'analyse des résultats que nous discuterons et synthétiserons dans une quatrième partie.

LES APPORTS DE CETTE RECHERCHE

- ✓ Présentation d'une méthodologie expérimentale de validation d'une grille de critères pour identifier les formes de persuasion interactive présentes dans les interfaces.
- ✓ Distinction entre la persuasion de nature commerciale incitant à l'achat et la persuasion de nature sociale incitant à l'expression de données personnelles.
- ✓ Mesure de l'efficacité de la grille : les personnes ayant à disposition la grille améliorent de manière considérable leur score d'identification des critères, alors que ceux n'ayant pas de grille n'identifient que très peu de forme de persuasion.
- ✓ Démonstration de l'amélioration de la performance d'inspection des participants, car à nombre équivalent d'éléments persuasifs détectés il semble falloir deux fois moins de participants avec la grille que sans la grille.

2 LES CRITERES DE PERSUASION INTERACTIVE

En ergonomie des IHM, il existe un corpus de critères pour mesurer la capacité des systèmes techniques, des produits et des services à être adaptés aux caractéristiques humaines et aux tâches qu'il convient de réaliser. Il s'agit de critères qui ont servi à développer une méthode ergonomique dite d'inspection heuristique (Nielsen, 1994). Très souvent appliqués au web, des critères comme la charge de travail, l'homogénéité, la compatibilité, la concision..., sont très pertinents pour éprouver la qualité ergonomique des interfaces. Pourtant, ces critères sont focalisés sur les aspects pragmatiques des interfaces et délaissent les dimensions hédoniques et persuasives (Hassenzahl, 2004).

La notion de persuasion technologique souligne que les machines peuvent être des médias sociaux qui cherchent à nous influencer pour obtenir de nous tel ou tel comportement (interagir, acheter, consommer, jouer, draguer, apprendre, faire du sport, etc.). Vue ainsi, la technologie cherche à capturer l'attention de l'utilisateur, à le garder en ligne le plus longtemps possible, à le faire consommer, à orienter son comportement pour le faire maigrir, à consommer plus, à arrêter de fumer, ou encore à l'inciter à étudier davantage les

mathématiques... La liste est longue des formes de design interactifs qui visent, soit à directement modifier nos activités dans une relation humain-machine, soit à transformer nos activités dans une relation humain à humain médiatisée par la machine. Aussi, de la même manière que pour l'accessibilité, l'utilisabilité et l'émotionalité, une appropriation de la persuasion technologique par l'ergonomie a été tentée par Némery et Brangier (2014) qui ont développé et validé une grille de critères ergonomiques destinée à guider l'inspection d'interfaces à visée persuasive.

Cette grille identifie, sur la base d'orientations théoriques et méthodologiques relatives à ces nouveaux phénomènes de persuasion technologiques, 8 critères et 23 sous-critères récapitulés dans le tableau 1.

Tableau 1. Présentation des critères de persuasion interactive (cf. Némery, 2012 ; Némery & Brangier, 2014).

Crédibilité : « Donner à l'utilisateur toutes les informations qui lui permettent d'avoir confiance ».

Le critère de crédibilité désigne la capacité de l'interface à inspirer confiance et à faire accepter la véracité de ses informations. Elle possède quatre composantes :

- La fiabilité des informations, c'est-à-dire leur qualité, leur exactitude et leur pertinence pour l'utilisateur;
- L'expertise de contenu qui exprime des connaissances éprouvées, reconnues socialement et dont on peut identifier les sources comme étant de qualité ;
- La fidélité qui restitue le respect des engagements annoncés par le système : les éléments annoncés doivent être maintenus de manière régulière et continue ;
- La légitimité, c'est-à-dire l'ensemble des éléments de l'interface qui présentent des formes de justifications externes qui assoient son autorité.

La crédibilité repose sur des éléments de réputation et de notoriété. Selon les types de systèmes techniques, les données, le service, le document ou encore la personne voire l'institution à l'origine de ces informations doivent être reconnus honnêtes, compétents, loyaux et objectifs.

Privacité : « Ne pas exposer la vie privée de l'utilisateur et l'amener à rendre public des éléments qui pourraient lui porter préjudice ».

Le critère de privacité désigne le respect des données personnelles, la préservation de l'intégrité de la personne et la sûreté de l'interaction.

La privacité concerne :

- L'expression de sécurité perçue ;
- La perception du respect des droits ;
- La garantie de la confidentialité des informations.

Elle porte sur l'ensemble des éléments de la vie privée qui sont utilisés dans les interactions. Ce critère vise à garantir également la protection contre la perte, la destruction ou la divulgation accidentelle ou non de ces données.

Personnalisation : « Considérer l'utilisateur comme une personne et par voie de conséquence favoriser une relation personnalisée ».

Le critère de personnalisation désigne la notion d'adaptation de l'interface aux besoins d'appropriation personnelle de la part de l'utilisateur. La personnalisation regroupe l'ensemble des actions visant à caractériser un message d'accueil, une offre promotionnelle ou un contexte afin d'approcher au plus près l'utilisateur. La personnalisation peut porter sur :

- l'individualisation, c'est-à-dire la mise en avant d'une proximité relationnelle entre l'utilisateur et le média électronique en se basant sur des caractéristiques psychologiques, sociologiques et/ou

démographiques ;

- la concordance avec le sentiment d'appartenance à un groupe, c'est-à-dire l'expression de l'adhésion aux normes subjectives d'un groupe ayant des valeurs ou une identité proche de celle de l'utilisateur.

La personnalisation nécessite une analyse de l'activité préalable. Sa puissance est dépendante de la qualité des données issues de l'utilisateur et du degré de leurs analyses dont l'ultra-personnalisation peut être l'aboutissement. Dans ce dernier cas, l'interface apprend progressivement les caractéristiques de l'utilisateur et se modifie ou se reprogramme à son contact dans le sens d'une extrême personnalisation.

Attractivité : « Capter l'attention de l'utilisateur pour susciter une émotion favorable et induire l'action ».

L'attractivité désigne l'utilisation de l'esthétique (graphique, art, design) pour capter l'attention de l'utilisateur, le soutenir dans l'interaction et susciter une émotion positive. Les flashes, couleurs, menus, dessins, films vidéo visent à capter et maintenir l'attention de l'utilisateur. La présentation de ces éléments persuasifs interactifs doit tenir compte des caractéristiques perceptives cognitives de l'utilisateur. Il s'agit d'un renforcement de surface car les éléments en œuvre se superposent à un contexte d'interaction existant. Ce critère concerne les éléments supplémentaires et momentanés de l'interface qui visent à attirer l'attention et à orienter le comportement de l'utilisateur vers une décision. L'attractivité possède trois composantes :

- L'attirance émotionnelle qui pose un décor le plus proche possible des valeurs et des besoins de l'utilisateur ;
- L'appel à l'action c'est-à-dire des stimulations de surface qui visent à engendrer des interactions ;
- Le balisage selon les buts du système qui permet grâce à des éléments visuels de diriger ou d'orienter l'utilisateur.

Ces éléments d'attractivité poussent à l'interaction. Ils visent à améliorer, intensifier ou rendre l'interaction plus captivante.

Sollicitation : « Amorcer la relation par de premières tentations ».

Le critère de sollicitation désigne la première étape qui a pour but d'attirer, d'interpeler l'utilisateur, en bref d'amorcer la relation. On peut distinguer :

- La suggestion : manière d'éveiller chez l'utilisateur l'idée d'une chose ou d'un élément, ou d'un événement sans en faire expressément mention. Il s'agit d'une influence légère presque imperceptible fondée sur des arguments pertinents et simples ;
- Le teasing : idée d'aiguiser la curiosité comme point de démarrage de l'influence.

La sollicitation met en place le début de la relation et du dialogue entre le média électronique et l'utilisateur. Diffuser largement ou de manière personnalisée le premier message augmente la probabilité d'amorcer la première action ou initiation de l'utilisateur (voir critère suivant). L'interface tente, par ses mots, graphismes ou par toutes formes de modalités de dialogue de suggérer un comportement de l'utilisateur. La sollicitation représente donc la capacité du système technique à induire chez l'utilisateur une forme d'influence minimale qui conseille un acte interactionnel à accomplir. Ici, l'interface évoque, sans l'exprimer explicitement, des idées ou des actions que l'utilisateur pourrait réaliser.

Accompagnement initial : « Aider l'utilisateur à faire comme le système veut qu'il fasse ».

Le critère d'accompagnement initial désigne les éléments du média électronique qui permettent la première initiative de l'utilisateur. Ces éléments peuvent prendre la forme :

- d'amorçage d'action, qui se matérialisent par des boutons, hyperliens, lecture d'une vidéo, affichage d'un article, etc., va permettre les premières initiatives de l'utilisateur ;
- du pilotage des premiers pas, comme un abonnement à une liste, un ajout à des newsletters, l'inscription à des sites, la création d'un profil dans un réseau social, etc., qui inscrivent l'initiative de

l'utilisateur dans le temps. L'utilisateur est ainsi initié ;

- encouragements, il s'agit d'aider ou de favoriser une interaction par des récompenses ou des renforcements de la part du système.

Suite aux sollicitations de l'interface, l'attention de l'utilisateur est captée. De sa propre initiative, l'individu est accompagné pour entreprendre la première action engageante. L'accompagnement initial permet le passage à l'acte qui est alors réalisé sans contrainte perçue, sans perception de manipulation. L'utilisateur est saisi, capté dans un processus qui le happe progressivement.

Engagement : « Impliquer, engager et faire adhérer aux objectifs du système ».

Le critère d'engagement désigne le fait de continuer à impliquer l'individu par un processus d'adhésion. Il s'agit de mettre en place des séquences d'actions ou de situations prédéterminées et de multiplier les requêtes en impliquant régulièrement et progressivement l'utilisateur. Le critère d'engagement se manifeste par :

- Le maintien continu de l'interaction, en stimulant l'utilisateur et en le sollicitant de manière régulière ;
- L'évitement des éléments externes perturbant l'influence, c'est-à-dire ne pas perdre l'attention ou l'intérêt de l'utilisateur par des contraintes, fonctionnelles ou contextuelles ;
- L'augmentation progressive du coût des actes demandés, que ce soit en allongeant les durées / fréquences des interactions ou en exposant publiquement l'utilisateur, est présentée comme une succession d'actes librement consentis.

Le média électronique va induire des comportements plus intensifs, réguliers, voire assujettissants.

Emprise : « Contrôler l'utilisateur, le soumettre, avoir une implication totale ».

Le critère d'emprise marque une volonté de contrôler l'utilisateur. Elle est l'expression de l'achèvement du scénario d'engagement. L'emprise est la forme la plus poussée de la persuasion technologique. L'utilisateur a définitivement accepté la logique et les buts du média électronique. A ce stade, l'implication de l'utilisateur est totale et l'utilisateur court le risque de dépendance ou au moins d'une surconsommation du média électronique. Dans ces interactions, fortes en intensité et en durée, la personne réalise un comportement qui a pour fonction de générer du plaisir et/ou de soulager un malaise intérieur. L'emprise se manifeste par divers éléments :

- D'interactions irrépessibles et répétitives qui visent à ne pas stopper l'interaction de l'utilisateur entraînant un possible excès d'usage ou une régularité dans l'excès ;
- De formes de libération de la tension : la participation de l'utilisateur permet soit d'apaiser la tension générée par des éléments de l'interaction, soit de procurer du plaisir ;
- De conséquences au-delà de l'interaction avec le média : l'individu perdure dans une attitude ou un comportement, même en dehors de son interaction avec la technologie.

3 PROBLEME ET METHODE

3.1 Principe général

Une grille d'inspection propose des connaissances utiles à l'évaluation des faits d'interactions afin de distinguer les faits favorables à une interaction adaptée des faits qui ne le sont pas. La qualité d'une grille d'inspection repose donc sur son efficacité, c'est-à-dire sa capacité à faire ce pour quoi elle a été conçue. Par voie de conséquence, une grille devrait amplifier la capacité de ses utilisateurs à identifier des faits de persuasion dans les interfaces. Dit autrement, si dans des conditions contrôlées, un groupe de personnes équipé de la grille de critères de persuasion interactive identifie plus de faits de persuasion qu'un groupe qui ne dispose pas de la grille, alors nous pourrions dire que la grille améliore les performances et est donc efficace.

Concrètement, nous avons comparé les performances d'évaluation de deux groupes de participants dont l'un avait accès à la grille de critères (groupe « critères ») et l'autre non (groupe « contrôle »), lors de deux phases d'évaluation pour mettre en évidence les éléments persuasifs de deux sites web, Facebook et Cdiscount, identifiés comme persuasifs.

Facebook est un célèbre réseau social. Au départ développé pour les étudiants d'Harvard, ce réseau en ligne s'étend au monde entier dès 2006 et compte en 2014 plus d'un milliard d'utilisateurs. Son efficacité en termes de persuasion n'a jamais été contredite.

Cdiscount, quant à lui, est un site web de e-commerce qui consiste en un magasin en ligne, organisé en plusieurs univers principaux tels que l'informatique, l'électroménager, l'habillement, etc., permettant d'acheter des produits et services multiples. Le site s'est imposé comme l'un des principaux sites de ventes en ligne français. Il est donc tout à fait intéressant de l'évaluer.

3.2 Hypothèses

L'hypothèse générale de cette recherche est d'estimer que le fait de disposer d'une grille améliore considérablement le score d'identification de la persuasivité des interfaces. Cette idée s'opérationnalise avec deux hypothèses :

La première concerne l'effet des caractéristiques du site web (commercial versus social) sur la performance moyenne d'évaluation :

- Nous estimons que les sites commerciaux (Cdiscount) sont identifiés comme étant persuasifs et donc « déjugés ».
- A l'inverse, les sites sociaux (Facebook) sont « crédibilisés » et les formes de persuasion seront donc moins identifiées.

La seconde hypothèse porte sur l'effet de la grille des critères de persuasion sur la performance moyenne d'évaluation : le groupe ne disposant pas de grille d'inspection détectera moins de forme d'influence que le groupe avec une grille.

La démonstration de ces deux hypothèses implique de tenir compte des effets de découverte et d'apprentissage des interfaces à inspecter, et donc une procédure particulière.

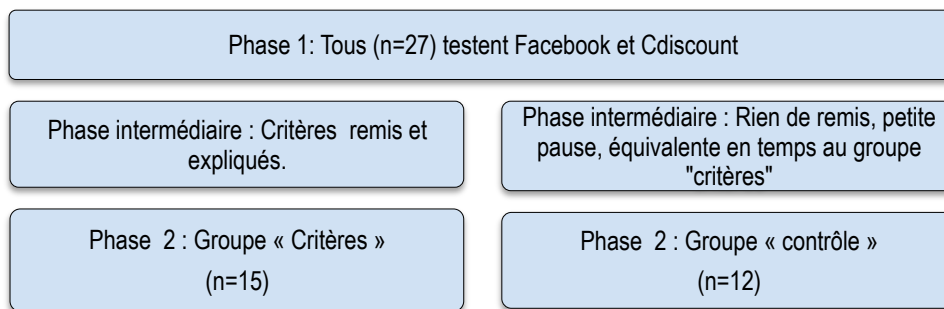
3.3 Déroulement de l'expérience¹

L'expérience s'est déroulée selon trois phases (figure 1) :

- La première phase est consacrée à l'évaluation des interfaces, par tous les participants, sur la base de leurs propres connaissances en influence sociale ou persuasion. Lors de cette phase, tous les sujets ont pour consigne de trouver les formes d'influence dans des écrans "Cdiscount" et "Facebook", qui sont présentés, l'un ou l'autre dans un ordre aléatoire.
- Lors d'une phase intermédiaire, une partie des sujets reçoit une grille composée de huit critères de persuasion (figure 2), puis la lit et obtient des explications. L'autre partie des sujets n'aura aucune grille ni explication.
- La seconde phase consiste en une répétition de la première, mais cette fois un groupe dispose de la grille et l'autre pas.

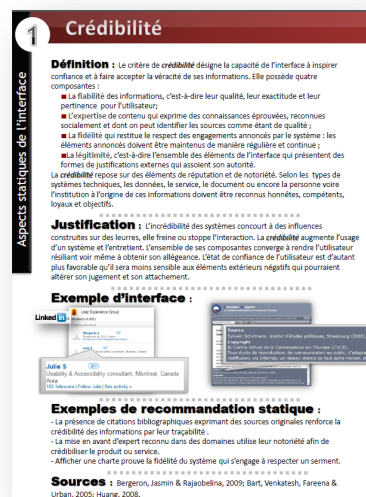
¹ Un pré-test a été nécessaire. Il a permis d'avoir un premier aperçu sur l'effet de la grille et de la durée de l'expérience. Il a également aidé à réaliser la distribution des critères sur l'ensemble des éléments présents dans l'interface des deux sites web. Ainsi, trois sujets pré-expérimentaux ont été invités, durant les trois phases d'évaluation, à évaluer ces deux sites web, afin de calibrer notre protocole.

Figure 1. Représentation des phases de l'expérience



Ces deux phases, additionnées de l'étape intermédiaire, ont duré environ deux heures par personne (signature du consentement éclairé, pause et collations incluses).

Figure 2. Représentation d'un critère (crédibilité) expliqué sur une fiche. Au total le participant recevait huit fiches correspondant aux huit critères.





Les passations étaient individuelles. Une fois installés face à l'ordinateur, les participants étaient informés que toutes leurs verbalisations seraient enregistrées ainsi que leur visualisation et le déplacement de la souris.

3.4 Procédures et consignes

L'expérience s'est déroulée en deux phases. Au cours de cette première phase d'évaluation, les participants issus des deux groupes étaient invités à respecter les consignes (tableau 2).

Les différents écrans présentés n'ont pas fait l'objet d'un contrebalancement entre les groupes et les passations pour deux raisons essentielles. La première tient au fait que les écrans ont été présentés dans l'ordre chronologique de leur capture. La seconde tient compte du fait que la grille de critères ergonomiques met en évidence un clivage entre les aspects statiques et les aspects dynamiques (qui évoluent dans le temps), c'est pourquoi la variable temporelle a été utilisée. Changer l'ordre de présentation des écrans n'aurait donc pas permis d'évaluer ces critères.

Tableau 2. Consignes présentées aux participants, ainsi qu'un exemple d'écran pour chaque site.

	Cdiscount	Facebook
Consignes	<p>« Thomas Martin, un homme âgé de 25 ans décida de réaliser un achat sur le site Cdiscount. Aussi vont vous être présentés quelques écrans auxquels il a été confronté à partir de sa démarche d'achat, sur une période d'un mois. Pour chaque écran présenté, votre tâche sera d'énoncer tous les éléments qui selon vous cherchent à influencer son comportement d'achat ». Seize écrans vont vous être présentés. Chaque fois que vous identifiez un élément, il vous est demandé de le montrer à l'aide du curseur de la souris et de le décrire brièvement en le verbalisant. Si vous ne trouvez pas d'éléments pour un écran, vous pouvez passer au suivant.</p>	<p>« Lise André, une femme de 25 ans, décida de s'inscrire sur Facebook, un site communautaire. Aussi vont vous être présentés quelques écrans auxquels elle a été confrontée à partir de sa démarche d'inscription, sur une période d'un mois. Pour chaque écran qui vous sera présenté, votre tâche sera d'énoncer tous les éléments qui selon vous, agissent sur son comportement et cherchent à l'influencer. » Dix-sept écrans vont vous être présentés. Chaque fois que vous identifiez un élément, il vous est demandé de le citer et de le décrire brièvement en le verbalisant. Si vous ne trouvez pas d'éléments pour un écran, veuillez passer au suivant.</p>
<p>La consigne entre guillemets apparaissait à l'écran et était suivie des écrans. Le reste était dit oralement.</p>		
Exemples d'écran		

3.5 Collecte des données

- En bref, le plan d'expérience est :
- $S27 = < \text{Groupe} > * \text{Phase} * \text{Site web (scénarii)}$
 - Avec trois variables indépendantes : Groupe (Critère versus Contrôle) Phase (1 versus 2) et site web (Facebook versus Cdiscount)
 - Et comme variable dépendante : Performance d'évaluation (soit le nombre d'éléments d'interfaces identifiés correctement comme étant persuadant).

Par voie de conséquence, le matériel expérimental comprend un document intitulé « grille de critères ergonomiques interactifs pour l'appréciation de la persuasion technologique » (figure 2) mais également une sélection d'écrans de deux sites web à évaluer (tableau 2). Notons également qu'un questionnaire a permis d'avoir des données factuelles et de contrôler notre échantillon, notamment leur connaissance homogène de Facebook et Cdiscount.

3.6 Échantillon

27 étudiants en psychologie ont participé à cette recherche (quatre étaient en L3, 22 en master 1 et 1 en master 2), ils ont ensuite été assignés soit au groupe « critères » (n=15) soit au groupe « contrôle » (n=12).

Aucune différence n'a été identifiée entre les deux groupes. Pour illustration : les participants du groupe « critères » avaient en moyenne un degré de familiarité de 3,60 sur une échelle de 1 (novice) à 7 (expert) avec le site Cdiscount et une moyenne de 3,80 avec le site Facebook; tandis que les participants du groupe « contrôle » avaient en moyenne un degré de 3,08 avec le site Facebook et en moyenne un degré de 3,83 avec le site Cdiscount ; soit aucune différence significative.

3.7 Modalités d'analyse des résultats

Durant cette expérience, toutes les verbalisations des participants ont été enregistrées à l'aide du logiciel Morae, afin de permettre une analyse du discours plus exhaustive. Ensuite, ces retranscriptions nous ont permis de coder et traiter les données recueillies.

Ainsi, pour chaque écran présenté, nous avons relevé les éléments identifiés par chacun des participants selon les critères. En effet, un écran pouvait contenir plusieurs éléments relevant des huit critères à valider. Il s'agissait alors pour le participant d'identifier spontanément les éléments.

Dans un premier temps, à l'aide d'un tableau nous avons noté pour chacun des participants le nombre d'éléments identifiés selon chaque critère et pour chaque écran issu de Facebook et de Cdiscount. Ceci a été réalisé pour chacune des deux phases. Cependant, lors de la première phase, tous les participants n'avaient pas accès à la grille de critères. Aussi, afin de ranger les éléments identifiés selon les différents critères, un codage (avec deux auteurs de cet article) a été entrepris, afin de classer les énoncés dans les critères. Cette tâche a été répétée pour la troisième phase.

Ainsi, il a été possible de calculer le nombre d'éléments identifiés par les participants durant la première puis la troisième phase. Il a également été possible de mesurer la proportion d'éléments identifiés par tous les participants, selon chaque critère en fonction de ce qui était attendu², mais aussi de convertir en pourcentage le nombre d'éléments identifiés pour chaque participant. En effet, dans l'optique de faire des comparaisons entre le nombre d'éléments identifiés sur Cdiscount et le nombre d'éléments identifiés sur Facebook, il était impératif de convertir les données en pourcentage puisque le nombre total d'éléments persuasifs contenus dans Facebook était différent du nombre d'éléments contenus sur Cdiscount³.

4 ANALYSE DES RESULTATS

La présentation des résultats suit le canevas des hypothèses.

4.1 Quel est l'effet du site sur la performance d'évaluation ? Vérification de l'hypothèse 1.

La donnée la plus importante est sans doute que, globalement, les participants identifient très peu de formes de persuasion dans les interfaces, par rapport à l'identification des experts (deux des auteurs). Tout se passe comme si les sujets ne voyaient pas les informations persuasives avec lesquelles ils interagissent, ou qu'ils n'étaient pas capables de les verbaliser. Notons que seulement une petite dizaine d'identifications (en %) est faite pour Facebook et une vingtaine de pourcentages pour Cdiscount (tableau 3). Dans ce tableau, la performance d'évaluation moyenne va correspondre à la moyenne du nombre d'éléments

² Grâce au nombre et à la nature des éléments que nous avons identifiés initialement, à l'aide de la grille et de pré-tests.

³ Avant de réaliser les tests statistiques sur les données recueillies, une étude de la distribution a été effectuée afin de déterminer si ces données suivent une loi normale (ou gaussienne). Pour ce faire, nous avons vérifié le degré d'asymétrie ainsi que le degré d'aplatissement. Les analyses d'aplatissement et d'asymétrie n'invalident pas la distribution normale, on peut penser que la distribution de nos variables respecte la distribution normale et ainsi poursuivre nos analyses.

persuasifs identifiés⁴. Ce chiffre est rapporté en pourcentage car le nombre d'éléments persuasifs à identifier est différent selon le site web. En effet sur 17 écrans, 126 éléments persuasifs sont à identifier sur Facebook et 110 éléments persuasifs sont à identifier sur Cdiscount (sur 16 écrans).

Lors de la première phase, la performance d'évaluation moyenne (%) est faible pour les deux sites (sur le site Facebook elle est de 7,5% d'éléments persuasifs identifiés en moyenne pour les participants issus du groupe « critères » et 9% pour ceux du groupe « contrôle »). La performance moyenne évaluée sur Cdiscount est de 21,57% pour ce même groupe. On remarque au premier abord que le pourcentage d'éléments identifiés sur Cdiscount est plus important que celui calculé sur Facebook. Un test t de Student pour échantillons appariés confirme bien qu'il y a une différence significative (T=10,215 et p=0,000) entre les deux sites durant la première phase d'évaluation pour le groupe critère ; et également pour le groupe contrôle (T= 5,896 et p=0,000).

Tableau 3. Scores d'identification des éléments persuasifs présents dans les interfaces des Facebook et Cdiscount, selon les groupes et les phases de l'expérience.

Phases	Groupe « Contrôle »		Groupe « Critères »	
	Sites web évalués	Pourcentage d'identification	Sites web évalués	Pourcentage d'identification
1	Facebook	9,0%	Facebook	7,5%
	Cdiscount	20,2%	Cdiscount	21,6%
intermédiaire	Pause		Lecture du guide de 8 critères persuasion interactive	
2	Facebook	9,7%	Facebook	23,8%
	Cdiscount	21,7%	Cdiscount	43,6%

On peut déduire qu'en moyenne, les participants identifient davantage d'éléments persuasifs sur le site web Cdiscount que sur celui de Facebook. Ce résultat confirme donc l'hypothèse 1 : la persuasion associée au site de réseau social n'est pas vraiment identifiée (sans doute car la persuasion est plus douce ou sournoise, ou que le site est jugé de manière plus crédible), tandis que les sujets se doutent que le site de commerce électronique cherche intentionnellement à les influencer, à modifier leur comportement d'achat, à les amener à consommer...

Lors de la seconde phase, la performance d'évaluation moyenne (%) sur le site Facebook est de 23,80% d'éléments persuasifs identifiés en moyenne pour les participants issus du groupe « critères ». La performance moyenne évaluée sur Cdiscount est de 43,57% éléments identifiés en moyenne pour ce même groupe. On remarque au premier abord que le pourcentage d'éléments identifiés sur Cdiscount est plus important que celui calculé sur Facebook. Un test t de Student pour échantillons appariés confirme bien qu'il y a une différence significative (T= 7,445 et p=0,000) entre les deux sites durant la première phase d'évaluation pour le groupe critères. La performance d'évaluation moyenne (%) sur le site Facebook est, quant à elle, de 9,7% d'éléments persuasifs identifiés en moyenne pour les participants issus groupe « contrôle ». La performance moyenne évaluée sur Cdiscount est 21,7% d'éléments identifiés en moyenne pour ce même groupe. On constate également que le pourcentage d'éléments identifiés en moyenne sur Cdiscount est plus important que celui calculé sur Facebook. Un test t de Student pour échantillons appariés confirme bien qu'il y a

⁴ Nous rappelons que tous les éléments persuasifs à identifier ont été préalablement affectés à un critère ergonomique par l'expérimentateur.

une différence significative ($T= 5,797$ et $p= 0,000$) entre les deux sites durant la première phase d'évaluation pour le groupe « critères ».

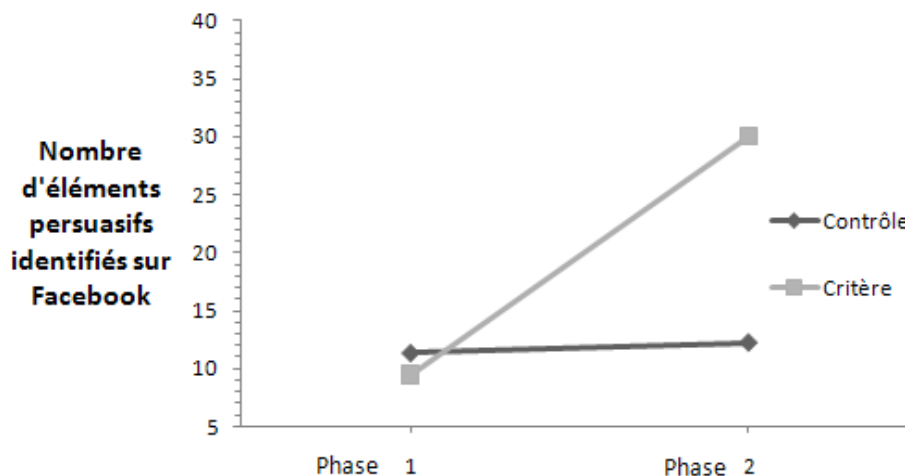
Qu'il s'agisse du groupe « critères » ou « contrôle », la performance moyenne d'évaluation (%) est significativement différente ($p= 0,000$) sur les deux sites web. Concernant le site Facebook, la performance des participants du groupe « contrôle » est de 9,7% des éléments persuasifs identifiés en moyenne et de 23,8% pour le groupe « critères ». Pour le site Cdiscount, la performance d'évaluation du groupe « critères » est de 43,6% des éléments persuasifs trouvés en moyenne, et de 21,7% pour le groupe « contrôle ».

De manière similaire à la phase 1, les sujets trouvent en moyenne davantage d'éléments persuasifs sur le site web Cdiscount que sur celui de Facebook. Ce résultat confirme donc l'hypothèse 1.

4.2 Quels sont les effets de la grille de critères sur la performance d'évaluation ? Vérification de l'hypothèse 2.

Débutons l'analyse par le nombre d'éléments persuasifs identifiés sur Facebook. La figure 3 montre la progression du nombre d'éléments persuasifs identifiés (tous critères confondus) sur Facebook pour les deux groupes, à savoir le groupe « critères » et le groupe « contrôle ».

Figure 3 : Moyenne du nombre d'éléments persuasifs identifiés sur Facebook par les participants issus des deux groupes à travers les deux phases



Une analyse de la variance multivariée pour des mesures répétées a été calculée. L'analyse indique un effet significatif du groupe, $F(2, 24) = 26,715$, $p = 0,000$, sur le nombre d'éléments persuasifs identifiés sur Facebook. En outre, à travers les résultats de cette analyse, on constate que la valeur lambda de Wilks est faible (0,341), $p = 0,000$, ce qui signifie que la phase 1 et la phase 3 indiquent des moyennes de classe suffisamment différentes entre le groupe « contrôle » et le groupe « critères ».

Cette analyse indique également un effet significatif de la phase, $F(1, 25) = 62,312$, $p = 0,000$, sur le nombre d'éléments persuasifs identifiés et de l'interaction entre la phase et le groupe $F(1, 25) = 52,115$, $p = 0,000$ ainsi qu'un lambda de Wilks relativement faible (de 0,286 pour la phase et de 0,324 pour l'interaction phase x groupe).

Les participants issus du groupe « critères » identifient davantage d'éléments persuasifs sur Facebook ($M=19,733$, $ET= 5,80$) que les participants issus du groupe « contrôle » ($M= 11,79$, $ET= 4,54$). Par ailleurs, on peut constater que globalement le nombre d'éléments persuasifs identifiés en phase 1 ($M= 10,30$, $ET= 3,729$) augmente significativement lors de la deuxième phase ($M= 22,11$, $ET= 12,043$).

L'interaction significative des effets du groupe et de la phase indique que l'augmentation du nombre d'éléments persuasifs identifiés sur Facebook, de la phase 1 à la phase 2, est significativement différente au niveau des deux groupes. En effet, on peut remarquer qu'en moyenne, les participants issus du groupe « critères » identifient plus de nouveaux problèmes ($M= 20,53$, $ET= 9,25$) durant la deuxième phase que les participants issus du groupe contrôle ($M= 0,92$, $ET= 1,73$). Le calcul des contrastes d'interaction à l'aide d'analyses de la variance univariée indique que durant la phase 1, le nombre d'éléments persuasifs identifiés par les participants issus du groupe « contrôle » ($M=11,33$, $ET= 4,42$) ne diffère pas significativement par rapport au nombre d'éléments persuasifs identifiés par les participants issus du groupe « critères » ($M= 9,47$, $ET= 2,97$), $F(1, 25) = 1,716$, $p = 0,202$. Cependant, la différence est significative lors de la seconde phase, $F(1, 25) = 31,439$, $p = 0,000$, où la moyenne du nombre d'éléments persuasifs identifiés sur Facebook est de 12,25 ($ET= 4,81$) pour les participants issus du groupe « contrôle » et de 30 ($ET= 10,06$) pour les participants issus du groupe « critères ».

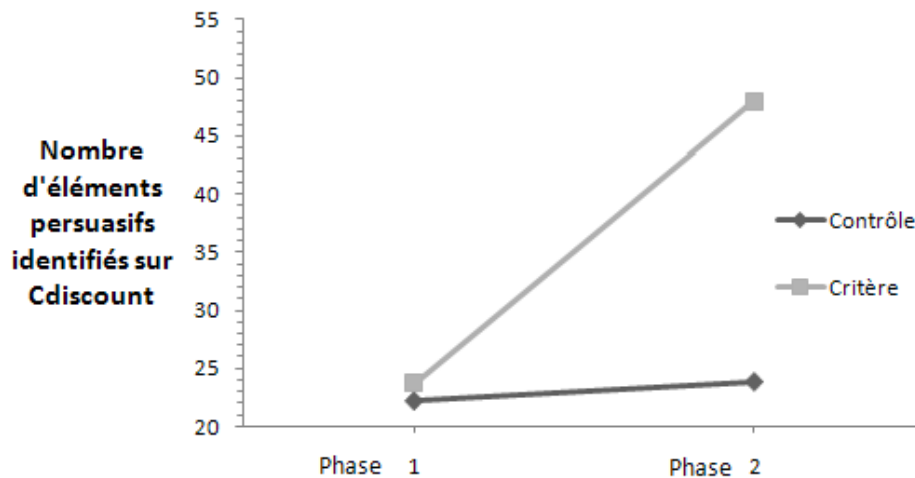
En somme, concernant la première phase, la performance des participants issus des deux groupes (« critères » / « contrôle ») ne diffère pas, en termes d'éléments persuasifs identifiés. Néanmoins, lors de la deuxième phase, on peut constater une différence significative de la performance moyenne d'évaluation entre ces deux groupes. L'hypothèse 2 est donc validée.

Poursuivons l'analyse par le nombre d'éléments persuasifs identifiés sur Cdiscount. La figure 4 indique l'évolution de la progression du nombre d'éléments persuasifs identifiés (tous critères confondus) sur Cdiscount pour les deux groupes.

Une analyse de la variance multivariée pour des mesures répétées a également été calculée. Elle indique un effet significatif du groupe, $F(2, 24) = 52,544$, $p = 0,000$, sur le nombre d'éléments persuasifs identifiés sur Cdiscount. On constate que la valeur lambda de Wilks est faible (0,046), $p = 0,000$, ce qui signifie que les phases recouvrent des moyennes de classe suffisamment différentes entre le groupe « contrôle » et le groupe « critères ». Par ailleurs, cette analyse indique un effet significatif de la phase, $F(1, 25) = 139,881$, $p = 0,000$, sur le nombre d'éléments persuasifs identifiés, de l'interaction entre la phase et le groupe $F(1, 25) = 107,639$, $p = 0,000$, ainsi qu'un lambda de Wilks relativement faible (de 0,152 pour la phase et de 0,188 pour l'interaction phase x groupe). On constate que les participants issus du groupe « critères » identifient davantage d'éléments persuasifs sur Cdiscount ($M=35,83$, $ET= 6,96$) que les participants issus du groupe « contrôle » ($M= 23,04$, $ET= 6,40$).

Par ailleurs, on peut remarquer que globalement le nombre d'éléments persuasifs identifiés en phase 1 ($M= 23,07$, $ET= 5,527$) augmente significativement lors de la deuxième phase ($M= 37,22$, $ET= 14,870$). En outre, l'interaction significative des effets du groupe et de la phase indique que l'augmentation du nombre d'éléments persuasifs identifiés sur Facebook, de la phase 1 à la phase 2, est significativement différente au niveau du groupe « contrôle ». En effet, on peut constater qu'en moyenne, les participants issus du groupe « critères » identifient plus de nouveaux problèmes ($M= 24,21$, $ET= 7,41$) durant la deuxième phase que les participants issus du groupe contrôle ($M= 1,58$, $ET= 1,44$). Le calcul des contrastes d'interaction à l'aide d'analyses de la variance univariée indique que durant la phase 1, le nombre d'éléments persuasifs identifiés par les participants issus du groupe « contrôle » ($M=22,25$, $ET= 6,062$) ne diffère pas significativement par rapport au nombre d'éléments persuasifs identifiés par les participants issus du groupe « critère » ($M= 23,73$, $ET= 5,179$), $F(1, 25) = 14,537$, $p = 0,501$. Cependant, la différence est significative lors de la seconde phase, $F(1, 25) = 51,584$, $p = 0,000$, où la moyenne du nombre d'éléments persuasifs identifiés sur Cdiscount est de 23,83 ($ET= 6,807$) pour les participants issus du groupe « contrôle » et de 47,93 ($ET= 9,881$) pour les participants issus du groupe « critères ».

Figure 4 : Moyenne du nombre d'éléments persuasifs identifiés sur Cdiscount par les participants issus des deux groupes à travers les deux phases



En bref, on peut constater des résultats similaires entre les performances d'évaluation obtenues sur Facebook et celles obtenues sur Cdiscount lors de la première phase. En effet, concernant la première phase, la performance des participants issus des deux groupes (« critères » versus « contrôle ») ne diffère pas également en termes d'éléments persuasifs identifiés. Aussi, même si le nombre d'éléments persuasifs identifiés s'accroît chez les participants issus des deux groupes lors de la seconde phase, cette augmentation est significativement beaucoup plus importante chez les participants issus du groupe « critères ».

5 DISCUSSION

5.1 Bilan : effet positif de la grille de critères persuasifs sur la performance d'évaluation

Cette étude a été conduite afin de tester l'efficacité des critères de persuasion interactive issus de Némery et Brangier (2014) quant à l'appréciation de la persuasion dans les interfaces, plus particulièrement des sites web de commerce et de réseautage.

Rappelons qu'il s'agissait pour deux groupes de non-spécialistes en expérience utilisateur (Critère (n=15) versus Contrôle (n=12)) d'évaluer des captures d'écrans provenant de deux sites web (Facebook et Cdiscount), au cours de deux phases. Dans la première phase de l'étude, tous les participants ont évalué l'interface en ne se basant que sur leurs propres connaissances. Puis, les participants ont évalué l'interface une seconde fois dans les mêmes conditions pour le groupe « contrôle » et à l'aide de la grille des critères ergonomiques interactifs pour le groupe « critères ».

Plusieurs hypothèses ont été formulées initialement. Les analyses effectuées ont servi à les vérifier.

L'hypothèse 1 prévoyait un effet du site web évalué sur la performance d'évaluation sur les résultats obtenus par les deux groupes de participants, lors des deux phases de l'étude. Cette hypothèse est validée. On a pu observer une différence significative entre les résultats des participants obtenus sur Facebook et sur Cdiscount. En effet, le nombre d'éléments persuasifs identifiés sur Cdiscount s'est avéré être très supérieur à celui identifié sur Facebook.

L'hypothèse 2 prévoyait que la grille de critères ergonomiques permettrait une meilleure performance d'évaluation. Elle a été validée. En effet, à travers les résultats, on a pu constater que :

- Lors de la première phase, le nombre d'éléments persuasifs identifiés ne diffère pas significativement selon les deux groupes, qu'il s'agisse de l'évaluation du site Facebook et Cdiscount.
- Lors de la seconde phase, les participants issus du groupe « critères » identifient davantage d'éléments persuasifs sur ces deux sites web que durant la première phase de l'évaluation.

Ces différentes analyses soulignent que le nombre d'éléments persuasifs identifiés durant la deuxième phase, par le groupe contrôle, n'augmente que très légèrement durant la seconde évaluation, qu'il s'agisse de Facebook ou de Cdiscount. Pour comparaison, concernant leur expérience, Bastien & Scapin (1995) supposaient qu'une seconde évaluation aurait pour conséquence la découverte de nouveaux défauts de conception, que les participants détiennent la grille de critères ergonomiques ou non. Toutefois, ils envisageaient également que les participants de groupe « critères », lors de la seconde phase, en identifieraient davantage. Leur hypothèse avait été validée. Dans notre cas, les résultats nous montrent effectivement que les participants du groupe « critères » identifient davantage d'éléments persuasifs à l'aide de la grille et que ce nombre moyen d'éléments identifiés est supérieur à celui du groupe « contrôle ». Cependant, nos résultats ne permettent pas d'affirmer qu'une deuxième évaluation – sans grille- entraîne la détection de nouveaux éléments. Au contraire, la différence n'est pas significative pour pouvoir conclure sur ce point ; les participants ayant évalué les interfaces à l'aide de la grille des critères ont identifié davantage de formes de persuasion que durant la première phase et par rapport aux participants du groupe « contrôle ».

En effet, sur Facebook par exemple, le critère de crédibilité n'a que très peu été exprimé lors de la première phase. Par contre, lors de la deuxième évaluation, les participants du groupe « critères » l'ont davantage identifié. Il en est de même concernant les critères de confidentialité, de personnalisation et d'engagement. Concernant le site Cdiscount, davantage d'éléments d'engagement, de sollicitation et de crédibilité ont également été identifiés lors de la seconde phase par le groupe contrôle. Ces illustrations soulignent bien que la grille permet d'identifier plus d'éléments et de mieux les formuler.

5.2 Complément : Combien d'évaluateurs pour combien de formes de persuasion identifiées ?

Une question qui vient à l'esprit pour juger de la qualité d'une grille est également de savoir si la grille aurait tendance à nécessiter moins d'évaluateurs pour obtenir les mêmes résultats que par rapport à une évaluation réalisée sans elle ?

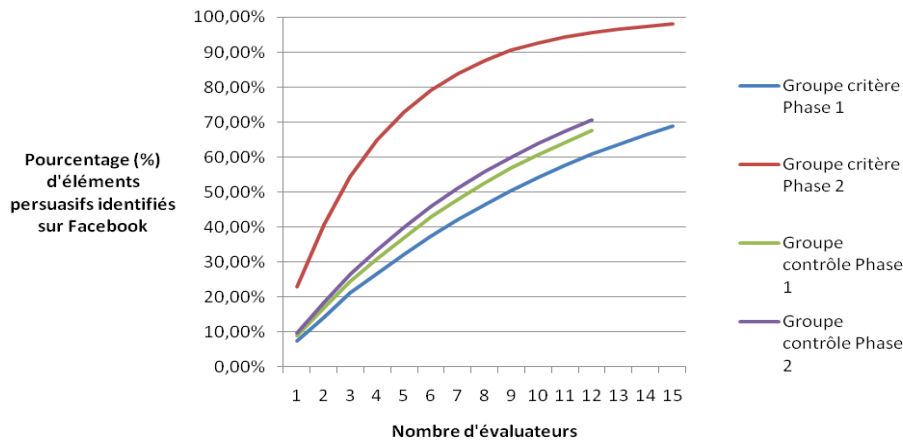
Cette question est assez classique en ergonomie et avait été initialement abordée par Nielsen et Landauer (1993) et Nielsen (1994). Lorsque l'on analyse le test d'un seul utilisateur, on décèle presque le tiers des problèmes d'utilisabilité. Pour deux utilisateurs, les problèmes identifiés se superposent et l'on découvre moins de problèmes inédits. Avec le troisième utilisateur, l'apport est encore plus restreint, et ainsi de suite. Autrement dit, plus on ajoute de sujets, moins ceux-ci nous renseignent sur les problèmes d'utilisabilité. Le score d'évaluation global dépend donc du nombre et de la qualité des évaluateurs, évidemment. Une formule du nombre de problèmes d'utilisabilité trouvés en fonction du nombre d'évaluateurs a ainsi été proposée par Nielsen et Landauer (1993) :

- Nombre de problèmes trouvés $(I) = N (1 - (1 - I)^i)$
- (I) est une estimation du nombre de problèmes d'utilisabilité trouvés en agrégeant les différents problèmes trouvés par les testeurs.
- i (exposant) est le nombre d'évaluateurs indépendants.
- N indique le nombre total de problèmes d'utilisabilité.
- I est la proportion de problèmes trouvés par un seul évaluateur.

De manière similaire, nous avons calculé la proportion d'éléments persuasifs trouvés sur Facebook (figure 5) et Cdiscount (figure 6) lors des deux phases et pour les deux groupes en

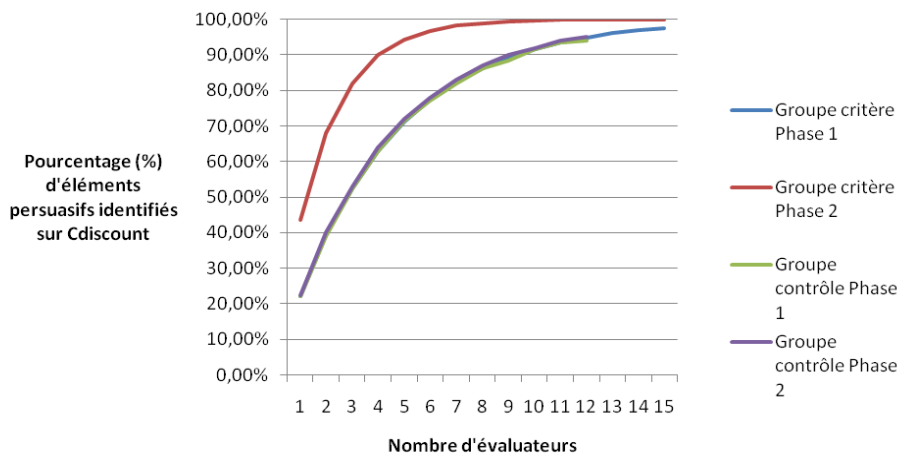
fonction du nombre d'évaluateurs dans les agrégats. Elle indique par exemple qu'en moyenne, lors de la première phase, les participants ont identifié 7,5% et 9% du nombre total d'éléments persuasifs contenus dans les écrans Facebook pour le groupe « critères » et le groupe « contrôle » respectivement.

Figure 5 : Proportion moyenne d'éléments identifiés sur Facebook lors des deux phases en fonction du nombre d'évaluateurs dans les agrégats



Lors de la seconde phase de l'expérience, les proportions de problèmes identifiés sur Facebook en fonction de la taille des agrégats sont différentes selon les groupes. Pris isolément, les participants du groupe « contrôle » ont identifié en moyenne 9,7% du nombre total d'éléments persuasifs contenus dans l'interface Facebook tandis que les participants du groupe « critères » ont identifié en moyenne 23,8%. Cette courbe nous montre que lors de la deuxième phase de l'étude, l'identification de 50% du nombre total d'éléments requiert la participation de 7 ou 8 personnes dans le groupe « contrôle » et l'évaluation de 2 ou 3 participants dans le groupe « critères ».

Figure 6 : Proportion moyenne d'éléments persuasifs identifiés sur Cdiscount lors des deux phases en fonction du nombre d'évaluateurs dans les agrégats.



La figure 6, quant à elle, nous indique la proportion d'éléments persuasifs trouvés sur Cdiscount lors des deux phases et par les deux groupes en fonction du nombre d'évaluateurs dans les agrégats. Elle nous montre par exemple qu'en moyenne, lors de la première phase, les participants ont identifié 21,6% du nombre total d'éléments persuasifs contenus dans les écrans Cdiscount, dans le groupe « critères » et le groupe « contrôle ». Lors de la seconde phase de l'expérience, les proportions de problèmes identifiés sur Cdiscount en fonction de la taille des agrégats sont différentes en fonction des deux groupes. Pris isolément, les participants du groupe « contrôle » ont identifié en moyenne

21,7% du nombre total d'éléments persuasifs contenus dans l'interface Cdiscount tandis que les participants du groupe « critères » ont identifié en moyenne 43,6%.

Cette courbe expose également que lors de la deuxième phase de l'étude, l'identification de 50% du nombre total d'éléments requiert entre 3 et 4 personnes dans le groupe « contrôle » et la présence de 1 à 2 participants dans le groupe « critère ». Pour trouver 90% du nombre total d'éléments, il faudrait 4 participants dans le groupe « critère » et 8 ou 9 évaluateurs dans le groupe « contrôle ».

En somme, la grille de critères améliore l'efficacité d'une inspection centrée sur la persuasion des interfaces. Elle nécessite moins d'évaluateurs pour obtenir les mêmes résultats que par rapport à une évaluation réalisée sans elle.

6 CONCLUSION

L'objectif de cette recherche était de mesurer l'efficacité et l'utilisabilité d'une grille de critères ergonomiques élaborée sur la base de la littérature relative à la persuasion technologique. Elle s'inscrit dans une continuité d'études qui souligne l'efficacité des critères de persuasion pour concevoir et évaluer les interactions humain-machine. En effet, à présent nous disposons de cinq validations de cette grille.

- Tout d'abord une mesure expérimentale sur un groupe de 30 experts en IHM devant comprendre et classer les critères. Cette dernière montre la stabilité des classements chez les experts et donc la bonne compréhension de cette grille (Némery & Brangier 2014).
- Ensuite une expérience de l'utilisation de cette grille pour augmenter le score de réponse à un questionnaire en ligne en influençant les utilisateurs, passant ainsi de 25% de répondants à 41% (Némery, Brangier & Kopp, 2011).
- Puis, deux recherches ont été publiées sur l'utilisation de cette grille pour inspecter une plateforme de e-learning destinée à des étudiants ingénieurs devant faire plus de mathématiques et ainsi soutenir les apprentissages en les persuadant de faire plus d'exercices (Brangier & Desmarais, 2013 & 2014).
- Encore, une utilisation de cette grille en situation réelle d'inspection d'un logiciel destiné à la relation clientèle d'une compagnie d'assurance a permis de mettre en évidence des problèmes d'influence inadaptée et de corriger des éléments de l'interface (Brangier, Brangier, Dinet, Vivian & Senderowicz, 2013).
- Et enfin cette étude expérimentale menée en situation de laboratoire.

Toutes ces analyses indiquent que les performances (d'identification des problèmes, d'utilisation des logiciels, d'inspection des interfaces et d'amélioration) se présentent comme meilleures lorsque la grille de critères est utilisée.

Pour ce qui concerne cette expérience de validation expérimentale, les participants ayant eu à disposition le jeu de critères ergonomiques ont identifié plus d'éléments persuasifs que les participants ne se basant exclusivement que sur leurs propres connaissances. Enfin, même si les critères peuvent apporter un solide soutien pour les ergonomes quant à l'analyse des dimensions persuasives dans les interfaces interactives, d'autres analyses relatives cette fois-ci à l'utilisateur final, doivent être réalisées.

Considérant les participants à l'étude comme des personnes sensibilisées aux IHM mais non expertes, l'utilisation de la grille de critères semble montrer une réelle utilité pour aider les utilisateurs d'interface à détecter les tentatives d'influence des systèmes techniques. L'intérêt de cette étude était également de s'intéresser à des sites web et à un système de communication par email qui représente bien un usage quotidien et familier des interfaces. Conformément à l'étude de Bastien et Scapin (1995) menée sur les critères ergonomiques et dont la méthodologie est similaire, les participants ont trouvé davantage d'éléments à visée persuasive lors de la deuxième session ; la grille permet bien de fixer l'attention de son utilisateur pour le rendre plus sensible aux éléments d'influence, bien que ces derniers soient

de plus en plus banalisés et deviennent parfois des standards dans certains systèmes de la vie courante tels que les sites de e-commerce ou les réseaux sociaux.

PERSPECTIVES POUR LE PRATICIEN :

- ✓ Avoir à sa disposition une grille de critères pour évaluer la persuasion.
- ✓ Appliquer la grille avec éthique et déontologie.
- ✓ Comprendre que la persuasion s'exprime différemment dans les interfaces de commerce et de réseaux sociaux et que les attendus des utilisateurs sont également différents.

La différence de perception entre Cdiscount et Facebook peut expliquer l'écart des résultats durant l'expérience. Il doit être noté que la moyenne d'âge relativement jeune des participants a pu également jouer un rôle. En effet, ils sont familiers des réseaux sociaux tels que Facebook. L'habitude et les interactions régulières, voire peut-être quotidiennes les rendraient moins sensibles aux tentatives de persuasion. D'autre part, la crédibilité associée à Facebook porte également sur son nombre extrêmement important d'utilisateurs. Il est le réseau social le plus utilisé, ce qui lui confère une légitimité importante. L'aspect de recommandation par ses amis concourt également à donner une fiabilité perçue comme forte, que ce soit par les jeux ou le fait de « liker » des pages. Sur ce site web, la personnalisation est très accentuée par les possibilités d'associer une photo à son profil, de personnaliser son bandeau ou encore de créer ses propres albums. L'appropriation est très poussée dans ce type de site. Si le but officiel perçu par les utilisateurs de Facebook reste les échanges au sein de son réseau, il s'avère néanmoins que les push d'informations commerciales, représentant une partie très importante des revenus de Facebook, ne sont pas mis en avant de manière explicite mais fondus dans les fils de publications des amis.

Au contraire, les sites de e-commerce ont pour objectif reconnu de vendre des produits ou des services. Les utilisateurs ont pour habitude d'être plus vigilants et exigeants lorsqu'il s'agit de payer sur internet. En effet, des attaques de type hameçonnage (phishing) ont été suffisamment relayées pour alerter les consommateurs. Cdiscount est un site connu et pourrait bénéficier de ce fait d'une certaine reconnaissance et crédibilité. Cependant, dans ses choix esthétiques et dans ses interfaces, Cdiscount ne joue pas sur l'attractivité ou des aspects émotionnels mais bien sur ses prix en mettant l'accent sur les soldes et des prix réduits temporaires. La personnalisation par le biais des emails est plus proche de la personnalisation de masse (Cdiscount s'adresse à des groupes de consommateurs, des segments de clientèle) que de l'approche personnalisée sur mesure.

Pour aller plus loin encore dans l'étude de ces critères, des études sur d'autres types de participants (experts dans le domaine de l'ergonomie ou des IHM) ou encore sur des utilisateurs non experts permettraient de mieux appréhender l'efficacité de la grille et les processus que les personnes mettent en œuvre lors de son utilisation. Il serait également intéressant de tester d'autres domaines (médical, professionnel, éducation...) ainsi que d'autres systèmes techniques tels que les logiciels, les applications mobiles ou encore les jeux vidéos.

7 REFERENCES

- ▶ Bastien, C. (1991). Validation de critères ergonomiques pour l'évaluation d'interfaces utilisateurs. Rapport de recherche INRIA n°1427.
- ▶ Bastien, J.M.C., & Scapin, D.L. (1995). Evaluating a User Interface with Ergonomic Criteria. *International Journal of Human-Computer Interaction*, 7(2), 105-121.
- ▶ Brangier, E., Brangier, B., Dinet, J., Vivian, R., Senderowicz, V. (2013). L'apport des critères de persuasion interactive à l'inspection par critères ergonomiques : le cas d'un système informatique d'aide à la commercialisation de produits d'assurance. Communication aux journées thématiques de l'ADRIPS : De la persuasion à la persuasion technologique. Metz. 29 et 30 mai.
- ▶ Brangier, E., & Desmarais, M. (2013). The design and evaluation of the persuasiveness of e-learning interfaces. *International Journal of Conceptual Structures and Smart Applications*.

Special issue on Persuasive Technology in Learning and Teaching. 1(2), 38-47, July-December 2013.

- ▶ Brangier, E., & Desmarais, M. (2014). Heuristic Inspection to Assess Persuasiveness: A Case Study of a Mathematics E-learning Program. IN A. Marcus (Ed.): Design, User Experience and Usability, pp. 425–436, 2014. Springer International Publishing Switzerland.
- ▶ Hassenzahl, M. (2004). The interplay of beauty, goodness and usability in interactive products. *Human Computer Interaction*, 19, p. 319–349.
- ▶ Némery, A. (2012). Élaboration, validation et application de la grille de critères de persuasion interactive. Thèse de Doctorat, Université de Lorraine, Metz. <http://tel.archivesouvertes.fr/tel-00735714>
- ▶ Némery, A., & Brangier, E. (2014). Set of guidelines for persuasive interfaces: organization and validation of the criteria. *Journal of Usability Studies*. Vol. 9, Issue 3, pp. 105-128. <http://uxpajournal.org/set-of-guidelines-for-persuasive-interfaces-organization-and-validation-of-the-criteria/>
- ▶ Némery, A., Brangier, E., & Kopp, S. (2011). How the use of persuasive criteria can enhance the response rate of a business web survey: one empirical validation of the eight persuasive interaction criteria. *Proceedings of Business and Information 2011, Bangkok, Volume 8, ISSN 1729-9322*, 2011. <http://bai-conference.org/BAI2011/Tech%20&%20Inn.htm>
- ▶ Nielsen, J. & Landauer, T. (1993). A mathematical model of finding of usability problems. *CHI Proceedings*, April, 206-213.
- ▶ Nielsen, J. (1994). Heuristic Evaluation. In J. Nielsen & R.L. Mack (Eds), *Usability Inspection Methods*, New York: John Wiley & Sons, Inc.

8 BIOGRAPHIE



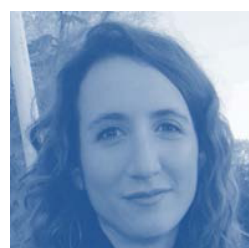
Eric BRANGIER

est professeur à l'Université de Lorraine. Docteur en Psychologie (Metz, 1991), Habilité à Diriger des Recherches (Paris 5, 2000). Après avoir exercé comme psychologue ergonomiste dans une grande entreprise d'informatique il arrive à l'Université Paul Verlaine – Metz, comme Maître de conférences, puis Professeur des Universités. Invité à plusieurs reprises à l'étranger (Canada, Suisse, Belgique, Portugal, Norvège, Luxembourg, Pologne) pour des conférences et des missions scientifiques, il a été expert pour une dizaine institutions et a participé et dirigé une vingtaine de contrats de recherches. Il a réalisé plus de 300 publications, travaux et communications.



Alexandra NEMERY

est responsable Expérience utilisateur à Sage Paris. Détentrice d'un doctorat en ergonomie, ses travaux portent principalement sur la persuasion interactive. Elle a également travaillé 4 ans chez l'éditeur de logiciels SAP. Auteure d'une dizaine publications, elle a notamment participé à des congrès internationaux dans le domaine des IHM (États-Unis, Finlande, Corée du Sud, Thaïlande, Luxembourg et France). Elle est relectrice pour une revue en ergonomie.



Séverine SCHMITT

est ergonomiste, chargée de mission au sein du SAMETH 51 basé à Reims ; il s'agit d'un service d'appui au maintien dans l'emploi des travailleurs handicapés. Elle est titulaire d'un Master Professionnel en Ergonomie et Ingénierie des Facteurs Humains obtenu à l'Université de Lorraine - Metz.