
Interface humanisée : quels effets sur l'implication dans la tâche et la qualité de l'interaction ?

Michel Dubois,

ERIHST, Laboratoire de Psychologie Sociale, Université Pierre Mendès France, Grenoble.

Michel.Dubois@upmf-grenoble.fr

Federico Tajariol,

ARCADE, Laboratoire Clips (Imag), Université Joseph Fourier, Grenoble.

Federico.Tajariol@imag.fr

RESUME

Les agents logiciels humanisés suscitent de plus en plus d'intérêts auprès des concepteurs d'interfaces. Mais quels sont les effets de ces interfaces sur l'implication dans la tâche et la qualité de l'interaction Homme-Machine ? Nos résultats expérimentaux montrent que plus les interfaces sont humanisées, plus les sujets sont impliqués dans la tâche, mais que la qualité de l'interaction dépend à la fois de l'humanisation de l'interface et du contexte de réalisation de la tâche.

MOTS CLES

Agents logiciels, interfaces humanisées, implication dans la tâche, qualité de l'interaction.

Au cours de la dernière décennie, de nombreux concepteurs informatiques ont créé, pour améliorer le dialogue Homme-Machine, des assistances (« agents logiciels »). Ces agents reposent sur des programmes informatiques qui s'exécutent indépendamment des actions de l'utilisateur, et pour interagir avec l'utilisateur, leurs interfaces sont de plus en plus sous forme de personnages humains représentés de manière caricaturale ou réaliste. Quels sont alors les effets de ces « interfaces humanisées » sur l'implication dans la tâche et le déroulement de l'interaction ?

1. INTERFACES HUMANISEES ET INTERACTION HOMME-MACHINE

1.1. L'interface humanisée : développements actuels et postulats de base

De nombreux développements informatiques d'interfaces humanisées ont été réalisés (Don & al., 1992 ; Lee & al., 1995 ; Sheth, 1994 ; Maes, 1995 ; Bell, 1996 ; Bates, 1994 ; Elliott, 1994, etc.). Ces développements s'appuient sur l'idée empirique que l'utilisateur attribue à ce type d'interface de la personnalité, de la motivation et des émotions. L'interaction avec la machine devient alors plus conviviale et plus facile à gérer car l'utilisateur se sert de ses propres connaissances interpersonnelles quotidiennes. Ainsi, pour Oren & al. (1990) l'intérêt de l'humanisation des guides logiciels est, d'une part, de réduire la charge cognitive lors de l'interaction Homme-Machine et, d'autre part, d'augmenter l'engagement de l'utilisateur. Pour cela il suffit de reproduire les traits caractéristiques des émotions utilisées lors des interactions interpersonnelles courantes (Laurel, 1990). L'interface humanisée est donc initialement posée comme permettant une gestion plus efficace de l'interaction Homme-Machine.

1.2. L'interface humanisée : une métaphore pertinente des modalités d'interaction humaine ?

Un grand nombre d'enquêtes de satisfaction après l'utilisation de programmes utilisant des interfaces humanisées ont montré que l'interaction est jugée satisfaisante car plus "naturelle" et plus "émotionnelle" (Thorisson, 1996 ; Kozierok, 1993 ; Nass & al., 1994 ; Lashkari & al., 1994 ; Blumberg, 1994 ; Maes, 1995 ; Schneiderman, 1995 ; Thorisson, 1996 ; Bell & al., 1996). Cependant, pour Lanier (1995), ce type d'interface peut rendre l'utilisateur plus paresseux, sans prise d'initiative car se conformant encore plus dans une relation sociale d'assistantat. L'ensemble de ces évaluations empiriques montre cependant que l'humanisation des interfaces aide à gérer l'interaction Homme-Machine. Mais ces évaluations se rapportent spécifiquement aux applications développées (de nature très hétérogène) sans chercher à comparer différents dispositifs de conception d'interfaces à partir d'une application identique.

1.3. Quels intérêts réels à utiliser une interface humanisée ? Etudes comparatives

Takeuchi et al. (1995) ont comparé un agent logiciel agent représenté soit par une flèche, soit par un visage synthétique. La perception d'engagement par l'utilisateur est supérieure avec une interface humanisée. La flèche est jugée comme un outil sérieux et utile ; l'interface faciale comme plus amusante et plus divertissante. King et al. (1996) ont eux aussi testé l'effet des caractéristiques des interfaces (géométriques vs faciales soit caricaturée, soit réaliste) sous forme statique ou dynamique. Les attributions d'intelligence sont plus importantes pour les interfaces faciales réalistes, puis les caricaturées, puis les géométriques. Walker et al. (1994) ont montré qu'un visage plus émotionnel (visage triste vs visage neutre) engage plus l'utilisateur dans l'interaction. Sproull et al. (1996) ont comparé, lors d'une enquête par ordinateur, une interface textuelle avec une interface humanisée présentant un visage à l'expression soit plaisante, soit triste et montré qu'avec les interfaces faciales, le niveau d'attention, d'élaboration et de temps de réponses est plus élevé (effets de facilitation sociale). Les traits de personnalité sociale (serviabilité, intelligence, etc.) sont plus attribués aux interfaces faciales. Les questions au contenu peu désirable socialement sont moins traitées devant les interfaces faciales (biais de désirabilité sociale et tendance à l'auto-présentation positive). L'ensemble de ces recherches a le mérite de comparer différentes interfaces à partir d'une même application. Globalement, une interface humanisée est perçue par l'utilisateur comme plus engageante et comme améliorant la qualité de l'interaction. Cependant ces recherches, surtout de nature évaluative, n'abordent jamais la variation du contexte d'interaction. Or ce contexte n'est jamais neutre socialement et exerce un ensemble d'effets sur la réalisation de la tâche et l'auto-estimation des sujets (cf. théorie de l'efficacité personnelle, Bandura, 1986).

Notre recherche se propose d'identifier plus précisément en faisant varier le contexte de passation : (a) les effets réels de l'interface humanisée sur le niveau d'implication dans la tâche, (b) les représentations qu'elles suscitent auprès de l'utilisateur notamment en terme de qualité de l'interaction.

2. EXPERIENCE

2.1. Méthode

Dans cette étude, les participants ont été invités à réaliser une tâche dans sa phase finale impossible à résoudre pour perturber émotionnellement les sujets, puis à se former une impression et à porter un jugement sur la qualité de l'interaction Homme-Machine.

Matériel :

Il a été réalisé avec le logiciel Authorware de Macromedia et Flash 4.0. L'écran de l'ordinateur est divisé en trois fenêtres distinctes (schéma 1). Pour la séquence d'apprentissage, deux fenêtres

informatives, situées dans la moitié gauche de l'écran, placées l'une au-dessus de l'autre, correspondent à la leçon (solfège rythmique). Elles présentent, en cinq écrans, les bases théoriques d'une leçon musicale avec du texte (partie haute) et des illustrations graphiques et sonores (partie basse). Différents thèmes sont abordés (note de référence, pulsation, succession de notes, division du temps, etc.). Les outils de navigation sont très simples. Le sujet peut revenir sur l'ensemble des écrans sans limite de temps. Lorsque le participant estime avoir suffisamment consulté la leçon, il clique sur une croix qui ferme la partie leçon et active une nouvelle configuration qui correspond à la partie « exercices ». Le participant est alors sollicité pour reconnaître des extraits de rythmes à partir de différentes portées musicales. Cette partie débute avec une page sous forme textuelle qui présente la démarche à suivre pour réaliser les exercices pré-testés par 33 sujets pour obtenir un niveau de succès respectif d'environ 100% (exercice 1), 75% (exercice 2) et 50% (exercice 3). L'exercice 4 est très difficile (10% de succès moyen) et l'exercice 5 impossible à résoudre. Lors de la réalisation des exercices, l'ordinateur dialogue avec l'utilisateur en fonction de la qualité des réponses (correcte ou fausse). Selon l'exactitude de la réponse, l'interface affiche soit des commentaires à valence positive pour l'encourager, soit des commentaires à valence négative pour le mettre en garde et susciter sa vigilance. La valence de ces dialogues de renforcement a été pré-testée par 27 sujets sur une échelle en 11 points (de « 1 » très négatif à « 11 » très positif). Nous avons retenu les propositions les plus représentatives des valences positives (n = 7 ; Moyenne > 8,5) et négatives (n = 14, M < 3) avec un écart type faible. Lorsque le sujet estime qu'il n'arrive plus à réaliser les exercices, il clique sur un bouton « abandon » qui ferme l'application.

Participants

120 étudiants de DEUG (âge moyen, 21,43 ans) ont participé à cette étude (80% de femmes). Tous les sujets sont novices en musique (pré-test).

Modalités d'interaction ordinateur-participant et contexte de passation

La présentation de la leçon est toujours identique.

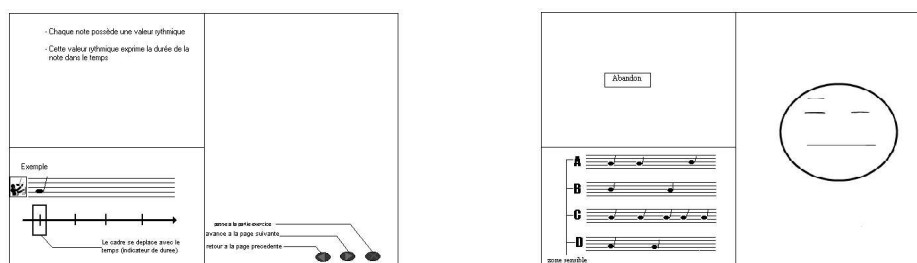


Schéma 1 : Fenêtres de l'application lors de la leçon (à gauche) et des exercices (à droite).

La partie « exercices » présente des modalités d'interaction différentes. Quatre conditions expérimentales ont été réalisées pour permettre à l'ordinateur de dialoguer sous forme de commentaires : soit textuels, soit sonores (voix humaine enregistrée avec des tonalités correspondant aux valences des messages), soit sonores avec simultanément la présentation statique d'un visage iconique à émotion négative ou positive correspondant à la valence des commentaires (cf. schéma 2), soit sonores avec une présentation dynamique de ce même type de visage. Le visage neutre constitue toujours le point de retour après la traduction d'une émotion faciale.



Schéma 2 : Les représentations iconiques des visages (de gauche à droite : expression neutre, négative et positive)

Parallèlement à ces différentes modalités d'interaction, les sujets sont, par effet de consigne, soit répartis dans une condition avec enjeu personnel (leur propre prestation fait l'objet d'une évaluation personnelle : processus d'auto-évaluation), soit dans une condition plus neutre en terme d'image de soi (leur rôle est de réaliser les tâches pour évaluer la qualité ergonomique et pédagogique de l'application dans son ensemble, leçon et exercices : processus d'évaluation externe de la part du sujet).

Mesures

Deux catégories de mesure ont été réalisées. Une première série de variables (mesures de temps) a été enregistrée automatiquement, en cours de passation, par l'ordinateur pour quantifier le degré d'implication du sujet dans les tâches. Une seconde série de mesures permet à l'utilisateur, après la passation de l'épreuve, d'évaluer la qualité de l'interface à l'aide d'un questionnaire (échelles en 5 points en majorité inversées pour éviter tous biais de désirabilité sociale). Cinq mesures d'évaluation ont été réalisées¹. Deux concernent la perception globale de la relation avec l'ordinateur : sentiment de dérangement pendant la passation de l'exercice (de 0 « Peu dérangement » à 4 « Très dérangement ») et de motivation à poursuivre (de 0 « Très motivante » à 4 « Peu motivante »). Les trois autres portent sur l'évaluation de la nature de l'interaction avec l'ordinateur durant la réalisation des exercices difficiles (de 0 « Peu perturbante » à 4 « Très perturbante » ; de 0 « Très sympathique » à 4 « Très antipathique » ; de 0 « Très encourageante » à 4 « Très décourageante »).

2.2. Principaux résultats

Le plan 4 (degré d'humanisation : texte vs voix vs voix + visage iconique statique vs voix + visage iconique dynamique) X 2 (contexte de passation : « avec enjeu personnel » vs « sans enjeu personnel ») a été traité par analyse de variance sur chaque facteur dépendant.

Implication dans la tâche

Le temps passé sur la partie leçon (temps d'apprentissage) est identique pour toutes les conditions (N.S.). Par contre, le temps passé par les utilisateurs dans la partie application suite à la première erreur (première difficulté) montre des différences significatives entre les conditions [$F(3,120) = 4,509$; $p < .005$].

Types d'interface	Moyenne	Ecart type
Sans visage :	322,35	(255,04)
- Textuelle	294,18	(257,15)
- Audio	350,47	(254,12)
Avec visage :	480	(261,05)
- Audio+Visage Iconique Statique	525,79	(262,29)
- Audio+Visage iconique animé	434,21	(255,95)

Tableau 1 : Temps d'implication moyen (en secondes) selon les conditions

Plus le degré d'humanisation augmente (« Avec visage »), plus les sujets persévèrent dans la réalisation des exercices [$F(1,120)=5,515$; $p < .03$]. Les sujets qui ont utilisé les interfaces avec visage [$M = 480,0$; $SD = 261,05$] ont passé plus de temps après la première erreur que les sujets qui ont utilisé les interfaces sans visage [$M = 322,3$; $SD = 255,04$]. Aucun effet du contexte de passation n'est trouvé, ni aucun effet d'interaction.

¹ Les items sélectionnés pour cette évaluation ont été pré-testés au préalable parmi un questionnaire de 21 items qui permettait à des sujets (population homogène) d'identifier, sur une échelle en 5 points, les termes les plus pertinents pour évaluer les caractéristiques d'une interaction avec une interface d'un cours multimédia. Nous avons retenu les 5 items ayant les plus fortes moyennes ($M > 4,27$).

Evaluation de la relation avec l'interface

Les sujets qui utilisent une interface avec visage ont plus tendance que les sujets de la condition « Sans visage » à percevoir les réactions de l'ordinateur comme dérangeantes [$F(1,119) = 2,94$; $p < .09$; $M = 3,53$ vs $3,20$] et peu motivantes [$F(1,119) = 5,29$; $p < .03$; $M = 3,42$ vs $2,98$].

Type d'interface	Evaluation du dérangement	Evaluation de la motivation
Sans visage	3,20 (1,36)	2,98 (1,07)
Avec visage	3,53 (1,05)	3,42 (0,98)

Tableau 2 : Perception globale de la relation avec l'ordinateur selon les conditions « sans visage vs avec visage »

Si on s'intéresse maintenant plus précisément aux effets d'interaction, lors de la réalisation des exercices difficiles, des facteurs « types d'interfaces » et « contexte de passation », on note un effet significatif pour l'évaluation du sentiment de perturbation [$F(3,120) = 3,93$; $p < .01$] et d'encouragement [$F(3,120) = 3,02$; $p < .03$] et de sympathie [$F(3,120) = 4,97$; $p < .003$]. Avec les conditions humanisées, les sujets « sans enjeu personnel » perçoivent la relation avec l'ordinateur comme plus perturbatrice, plus antipathique et plus décourageante que les personnes de la condition « avec enjeu personnel ». Au contraire, avec les interfaces non humanisées (textuelle et sonore), les personnes de la condition « sans enjeu personnel » jugent moins perturbantes, plus encourageantes et plus sympathiques les interfaces que les sujets de la condition « avec enjeu personnel ».

Type d'interface	Evaluation de la perturbation (de 0 Peu à 4 Très)		Evaluation de la sympathie (de 0 Très à 4 Peu)		Evaluation de l'encouragement (de 0 Très à 4 Peu)	
	Contexte « sans enjeu »	Contexte « avec enjeu »	Contexte « sans enjeu »	Contexte « avec enjeu »	Contexte « sans enjeu »	Contexte « avec enjeu »
Textuelle	2.53 (1.06)	2.93 (1.03)	2.46 (1.24)	3.33 (1.23)	2.86 (1.18)	3.53 (0.91)
Audio	2.28 (1.38)	3.06 (1.27)	3.21 (1.18)	4.06 (0.80)	3.0 (1.03)	3.80 (0.77)
Audio + Visage iconique statique	3.13 (1.18)	2.33 (1.11)	3.47 (1.06)	2.93 (1.27)	3.73 (0.60)	3.33 (1.11)
Audio + Visage iconique animé	3.20 (0.77)	2.40 (1.05)	3.60 (1.29)	2.60 (1.18)	3.53 (1.18)	3.26 (0.80)

Tableau 3 : Evaluation lors des exercices difficiles de l'interaction selon les conditions

3. DISCUSSION

Cette étude met en évidence les réactions différentes des utilisateurs selon le degré d'humanisation de l'interface. Pour les mesures d'implication du sujet dans la tâche, nos résultats corroborent ceux relevés dans la littérature (Takeuchi, 1995 ; Sproull & Walker, 1996, etc.). Plus l'interface est humanisée, plus les sujets persistent à résoudre les exercices qui sont difficiles et/ou impossibles à terminer. On retrouve ici l'effet de facilitation sociale certainement imputable à la volonté de vouloir se présenter de manière positive. Par contre, en terme d'évaluation relationnelle, nous ne trouvons pas de valorisations positives en faveur des interfaces humanisées. Elles sont globalement perçues comme peu motivantes et plutôt perturbatrices. Cette évaluation négative est en contradiction avec les résultats de la plupart des recherches. Elle découle des différences de contexte de passation utilisées dans notre expérience. En effet si l'on se place dans un contexte évaluatif pour le sujet (condition « avec enjeu personnel »), l'interface humanisée apparaît comme moins perturbante, plus encourageante et plus

sympathique (on retrouve les résultats de soutien social conformes à ceux évoqués dans la littérature). Inversement, dans le contexte « sans enjeu personnel », l'interface humanisée est perçue comme plutôt perturbante, antipathique et peu encourageante. Différentes explications, à ces résultats contradictoires, sont possibles :

- Les sujets ne s'attendent pas à une situation d'échec (statut d'évaluateur), d'où un sentiment de frustration (réaction négative) lorsqu'ils n'arrivent pas à terminer l'exercice. L'interface humanisée est alors dévalorisée car elle nuit plus à l'auto présentation positive du sujet que les interfaces plus « matérielles ».
- Les sujets sont embarrassés de devoir dialoguer avec une interface humanisée car elle revêt intrinsèquement des caractéristiques anthropomorphiques de nature plus « évaluative » alors que la tâche ne s'y prête pas (sans enjeu pour leur propre image personnelle). Elle devient donc plus déstabilisante pour l'estime de soi du sujet et ils lui préfèrent alors une interface « plus traditionnelle ».
- Enfin, les sujets mis dans une situation d'évaluateur ont plus de réticences pour émettre des jugements à l'encontre d'un dispositif aux caractéristiques plus « humaines », d'où un sentiment de malaise qui dégrade la perception de l'interface humanisée.

On retiendra que les interfaces humanisées affectent bien l'interaction Homme-Machine au niveau du comportement et de l'attitude des sujets (effet d'anthropomorphisme). Mais ces modifications ne sont pas toujours similaires, elles varient selon le contexte de réalisation de l'activité (« avec enjeu personnel » vs « sans enjeu personnel »). Des approfondissements dans cette perspective pourraient apporter de nouveaux éléments de réflexion utiles au développement des interfaces humanisées.

Bibliographie

- Bandura, A. (1986). *Social foundations of thought and action*. Englewood cliffs: Prentice-Hall.
- Bates, J., (1994). The role of emotion in believable agents. *Communications of the ACM*. 37, 7, 122-125.
- Bell, G., Ling, D., Kurlander, D., Miller, J., Pugh, D., Skelly, T., Stankosky, A., Thiel, D., Van Dantzych, M., Wax, T., (1996). Lifelike computer characters: the persona project at Microsoft Research. In J. Bradshaw (ed.). *Software Agents*. MA: MIT Press.
- Blumberg, B., (1994). Action selection in Amsterdam: lessons from ethology. In Cliff, D., Husbands, P., Meyer, J.-A. and Wilson, S. (Eds.), *From Animals to Animats 3: Proceedings of the third International Conference on Simulation of Adaptive Behavior*, pp.108-117. Cambridge, MA: MIT Press
- Don, A., Brennan, S., Laurel, B., Shneiderman, B. (1992). Anthropomorphism: from Eliza to Terminator 2. In *Proceedings of the CHI'92 Conference of Human Factors in Computing Design*, 67-69. Reading, MA: Addison-Wesley.
- Elliot, C., (1994). Multi-Media communication with emotion-driven "Believable Agents". In *Working Notes for the AAAI spring symposium on believable agents*. American Association for Artificial Intelligence. 16-20.
- King, W.J., & Ohya, J. (1996). The representation of agents : anthropomorphism, agency, and intelligence. In *CHI '96 Conference Companion of Human Factors in Computing System*, 289-290. MA : Addison – Wesley.
- Kozierok, R., (1993). A learning approach to knowledge acquisition for Intelligent Interface Agents. *SM Thesis*, Department of Electrical Engineering and Computer Science. MIT: Boston.
- Lanier, J., (1995). Agents of Alimentation. *Interactions*, July, 66-72. ACM Press.
- Lashkari, Y., Metral, M. & Maes, P., (1994). Collaborative Interface Agents. In *Proceedings of the national Conference of Artificial Intelligence*. Cambridge, MA: MIT Press.
- Laurel, B., (1990). Interface Agents: Metaphors with Character. In B. Laurel (ed.) *The Art of Human-Computer Interface Design*. Reading, MA: MIT Press.
- Lee, Y. & al. (1995). Realistic modeling for facial animation. In *SIGGRAPH'95 Conference Proceedings on Computer Graphics Annual Conference Series*. 35-62.
- Maes, P., (1995). Intelligent software. *Scientific American*. September, 84-86.
- Nass, C., Moon, Y., Fogg, B., Reeves, B., Dryer, D.C. (1995). Can Computer Personalities Be Human Personalities ? In *Proceedings of CHI'95 Conference of Human Factors in Computing Systems*. Reading, Ma : Addison-Wesley, 228-229.

- Oren, T. & al. (1990). Guides: Characterizing the Interface. In B. Laurel (ed.). *The Art of Human-computer Interface Design*. Reading, MA : Addison-Wesley, 367-381.
- Schneiderman, B., (1995). Looking for the Bright Side of User Interface Agents. *Interaction*, January, 13-15.
- Sheth, B., (1994). *A learning approach to personal information filtering*. Master's Thesis. MIT Media Lab.
- Sproull, L., Subramani, M., Kiesler, S., Walker, J.H. & Waters, K. (1996). When the interface is a face. *Human-Computer Interaction*, 11, 125-156.
- Takeuchi, A. & al. (1995). Situated facial Displays: Toward Social interaction. In *Proceedings CHI'94, Conference of Human Factors in Computing Systems*. Reading, MA: Addison-Wesley. 450-454.
- Thorisson, K. (1996). *Communicative humanoids: a computational model of psychosocial dialogue skills*. Dottoral Dissertation, Program in Media Arts and Sciences, MIT. July.
- Walker, J., & al. (1994). Using a Human Face in an Interface. In *Proceedings CHI'94, Conference of Human Factors in Computing Systems*. MA: Addison-Wesley, 85-91.