

Utilité, utilisabilité, acceptabilité: interpréter les relations entre trois dimensions de l'évaluation des EIAH

André Tricot, Fabienne Plégat-Soutjis, Jean-François Camps, Alban Amiel,
Gladys Lutz, Agnès Morcillo

► To cite this version:

André Tricot, Fabienne Plégat-Soutjis, Jean-François Camps, Alban Amiel, Gladys Lutz, et al.. Utilité, utilisabilité, acceptabilité : interpréter les relations entre trois dimensions de l'évaluation des EIAH. Environnements Informatiques pour l'Apprentissage Humain 2003, Apr 2003, Strasbourg, France. pp.391-402. edutice-00000154

HAL Id: edutice-00000154

<https://edutice.archives-ouvertes.fr/edutice-00000154>

Submitted on 5 Nov 2003

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.

Utilité, utilisabilité, acceptabilité : interpréter les relations entre trois dimensions de l'évaluation des EIAH

André Tricot*[#], Fabienne Plégat-Soutjis*, Jean-François Camps*,
Alban Amiel*[#], Gladys Lutz*, Agnès Morcillo*

* CERFI, IUFM de Midi-Pyrénées

56 avenue de l'URSS

31 078 Toulouse cedex

[#]Laboratoire Travail & Cognition

CNRS et Université de Toulouse 2

andre.tricot@toulouse.iufm.fr

RÉSUMÉ. Ce papier méthodologique propose de définir puis de mettre en relation trois dimensions de l'évaluation des EIAH. Ces trois catégories de critères, si elles ne sont pas nouvelles, apparaissent dans des champs différents de recherches et de pratiques. L'objectif de cet article est donc de les définir, de montrer leurs différences, leur complémentarité et leurs liens. L'utilité concerne l'efficacité pédagogique. Cette catégorie de critères répond à la question : l'EIAH permet-il aux personnes visées d'apprendre ce qu'elles sont censées apprendre ? L'utilisabilité concerne la possibilité de manipuler l'EIAH. Cette catégorie de critères répond à la question : l'EIAH est-il aisé à prendre en main, à utiliser, à réutiliser, sans perdre de temps et sans faire d'erreur de manipulation ? L'acceptabilité concerne la décision d'utiliser l'EIAH. Cette catégorie de critères répond à la question : l'EIAH est-il compatible avec les valeurs, la culture, l'organisation dans lesquelles on veut l'insérer ?

MOTS-CLÉS : évaluation, utilisabilité, utilité, acceptabilité, méthodologie, EIAH

1. Introduction

L'évaluation de l'utilité ou de l'efficacité des outils ou des dispositifs éducatifs n'est ni un problème nouveau, ni un problème original. Cependant, depuis l'introduction de l'informatique et des technologies de l'information dans le monde de l'éducation et de la formation, une dimension jusque-là quasiment ignorée est apparue : l'utilisabilité. De très bons outils, très bien promus par leurs concepteurs, voire remarquablement étayés d'un point de vue didactique, se révèlent inutilisables [DILLON & GABBARD 98]. Par ailleurs, des outils pourtant faciles à utiliser ne parviennent pas à entrer dans les pratiques scolaires ou de formation professionnelle. Certains enseignants n'ont pas le temps, les compétences ou l'envie de les utiliser,

les ordinateurs ne sont ni assez nombreux, ni assez performants, ni correctement disposés [BARON & LEVY 01]. Un bon EIAH est donc non seulement utile à l'apprentissage visé, mais encore utilisable et acceptable. Chacune de ces qualités de l'EIAH a fait l'objet de nombreuses études et correspond à des champs de pratiques bien développés. Mais ces recherches et ces pratiques sont le plus souvent étanches les unes aux autres, alors qu'elles sont complémentaires en général [KEATES & CLARKSON 01] et dans le domaine particulier des EIAH [SQUIRES & PREECE 99].

Dans cet article, nous évoquons de façon très incomplète quelques techniques d'évaluation, en privilégiant certaines, sans que cela ne mesure leur supériorité. Nous admettons aussi deux positions, concernant le « comment » et le « quand » de l'évaluation des EIAH, alors que ces positions sont discutables.

Concernant le « comment » nous admettons avec [SENACH 90] la distinction entre l'évaluation par inspection et l'évaluation empirique. L'évaluation par inspection est réalisée par un « expert », qui applique de façon plus ou moins explicite des critères d'évaluation. Par exemple, un ergonomiste va réaliser une évaluation par inspection de l'utilisabilité d'un site Web présentant des cartes géographiques, un spécialiste de didactique des mathématiques va réaliser l'évaluation d'un Tuteur Intelligent pour l'enseignement des mathématiques au CP. L'évaluation empirique, quant à elle, consiste à interpréter les performances des usagers, à qui l'on prescrit une tâche, et plus généralement à interpréter leurs comportements, attitudes, opinions. Nous admettons que ces deux types d'évaluation sont strictement distincts et complémentaires. L'évaluation par inspection permet de repérer rapidement des erreurs grossières et de diagnostiquer « pourquoi » tel ou tel aspect de l'EIAH est défaillant. L'évaluation empirique permet de voir moins rapidement l'ensemble des erreurs mineures et majeures, et de diagnostiquer ce qui ne va pas dans l'EIAH, sans nécessairement en expliquer les raisons. Nous montrerons que les trois dimensions de l'évaluation (utilisabilité, utilité, acceptabilité) se prêtent à ces deux types d'approches.

Concernant le « quand », nous admettons la distinction entre l'évaluation en cours de conception, en fin de conception et *a posteriori*. De nombreux auteurs plaident pour l'intégration de l'évaluation dès les phases amont du processus de conception [NANARD & NANARD 98]. Une telle évaluation permet de corriger à moindres frais des problèmes qu'il sera très coûteux voire impossible de corriger en fin de conception et contribue, pour l'équipe de conception, au partage de représentations cognitives communes sur l'artefact en cours d'élaboration. L'évaluation en fin de conception (la validation) recouvre l'évaluation des maquettes et le *béta-testing*. Enfin, l'évaluation *a posteriori* est réalisée une fois que l'outil est non seulement conçu mais réalisé. Les trois dimensions de l'évaluation (utilisabilité, utilité, acceptabilité) se prêtent à l'évaluation en cours de conception, en fin de conception et *a posteriori*.

Le problème principal abordé ici est trivial, mais personne à notre connaissance ne l'a encore posé : comment interpréter les relations entre trois séries de tests mesurant l'utilité, l'acceptabilité et l'utilisabilité ? Imaginons le cas d'un EIAH qui,

suite à une série de tests, se révèle acceptable mais inutilisable et inutile. Comment savoir ce qu'il faut améliorer ? Doit-on croire que l'amélioration de l'utilisabilité va entraîner une amélioration de l'utilité ? Qu'il faut aussi améliorer l'utilité ? Ou que l'amélioration de l'utilité entraînera une amélioration de l'utilisabilité ? Ces questions « pratiques » d'interprétation des relations entre trois dimensions de l'évaluation soulèvent un problème théorique encore débattu aujourd'hui, et qui n'est rien moins que celui de la rationalité, c'est-à-dire le problème des relations entre moyens et buts [SEARLE 01]. Élaborer une rationalité c'est élaborer un modèle des relations entre moyens et buts. Mais élaborer un tel modèle ne garantit pas la valeur de ce modèle, en particulier quand il s'agit de l'exporter. On sait par exemple qu'en évaluation des interactions humains machines (IHM) un des problèmes réside dans l'exportation de critères rationnels d'un champ (*e.g.* les bases de données) vers un autre (*e.g.* les hypertextes). Un critère rationnel dans un champ peut ne pas l'être dans un autre [TRICOT et al. 99]. C'est pourquoi nous préférons une approche empirique de la rationalité. Est rationnelle une relation observée entre deux états : atteinte du but et mise en œuvre des moyens.

Nous considérons en outre que pour chaque objet finalisé, il existe une représentation mentale (individuelle ou collective) de cet objet (*e.g.* de ses attributs physiques) et de sa rationalité. Nous appelons acceptabilité la valeur de cette représentation.

Nous défendons la position suivante :

— quel que soit l'objet finalisé à évaluer, toute relation entre Utilisabilité (possibilité de mettre en œuvre les moyens), Utilité (possibilité d'atteindre le but) et Acceptabilité, est possible ;

— toute relation possible peut-être décrite de façon logique.

Nous allons donc définir ces trois dimensions de l'évaluation des EIAH, souligner leur pertinence et rappeler les variables qui permettent de mesurer par inspection et de façon empirique ces trois dimensions. Ensuite nous proposons un modèle interprétatif des différentes relations entre ces trois dimensions et montrons l'intérêt d'une telle interprétation.

2. Trois dimensions de l'évaluation

De façon résumée, le Tableau 1 présente des critères qu'il est possible d'utiliser en évaluation empirique et par inspection, pour l'utilité, l'utilisabilité et l'acceptabilité.

Type d'évaluation Dimensions	Empirique	Par inspection
Utilité	<p>Adéquation entre objectif défini et apprentissage effectif</p> <p>Adéquation entre dispositif et format de la connaissance à acquérir</p> <p>Différence entre niveau de connaissances initial et terminal</p> <p>Mesures par des tâches de :</p> <ul style="list-style-type: none"> - reconnaissance - rappel (contenu / structure) - résolution de problème - détection d'erreurs - production 	<p>Précision et présentation des objectifs</p> <p>Adéquation contenus /objectifs</p> <p>Précision du scénario didactique</p> <p>Adéquation scénario / objectifs / contenus</p> <p>Mise en œuvre des processus cognitifs et méta cognitifs</p> <p>Régulation</p> <p>Évaluation</p>
Utilisabilité	<p>Possibilité d'apprendre à utiliser le système</p> <p>Gestion et Prévention des erreurs</p> <p>Mémorisation du fonctionnement</p> <p>Efficience</p> <p>Sentiment de satisfaction</p> <p>Évaluation par :</p> <ul style="list-style-type: none"> - observations - entretiens - analyse des parcours 	<p>Guidage et Incitation</p> <p>Groupement / Distinction des items par localisation ou format</p> <p>Feed-back immédiat et nature du feed-back</p> <p>Charge de travail</p> <p>Contrôle explicite</p> <p>Adaptabilité</p> <p>Gestion des erreurs</p> <p>Qualité des messages</p> <p>Homogénéité et cohérence</p> <p>Signifiante des codes et dénominations</p>
Acceptabilité	<p>Motivation</p> <p>Affects</p> <p>Culture</p> <p>Valeurs</p> <p>Évaluation par :</p> <ul style="list-style-type: none"> - observations - entretiens - questionnaires 	<p>Acceptabilité en termes d'adéquation aux :</p> <ul style="list-style-type: none"> - besoins ou objectifs de l'institution - attentes des apprenants - caractéristiques des apprenants <p>Acceptabilité en termes de compatibilité avec :</p> <ul style="list-style-type: none"> - l'organisation du temps - l'organisation des lieux <p>Présence du matériel nécessaire</p> <p>Planification et suivis lisibles et cohérents</p> <p>Visibilité des résultats</p>

Tableau 1 . *Présentation des mesures et critères pour l'évaluation d'un EIAH*

2.1. Évaluation de l'utilité

L'évaluation de l'utilité relève du domaine général de la pédagogie, des didactiques, et plus généralement de l'évaluation telle qu'elle est habituellement conçue en enseignement et en formation. Il s'agit d'évaluer s'il y a bien adéquation entre l'objectif d'apprentissage défini par l'enseignant (ou le concepteur) et l'atteinte de cet objectif. Les EIAH ne sont pas toujours munis de scénarios didactiques. Le site Web contenant des cartes géographiques n'est pas en lui-même « didactique ». Ce n'est qu'au moment où un enseignant, dans le cadre de son enseignement, prescrit une tâche de recherche et compréhension de cartes que l'utilisation de ce site devient « didactique ». *A contrario*, le Tuteur Intelligent en mathématiques pour le CP contient un scénario didactique qui va entre autres présenter des tâches à l'élève (tâches dont la réalisation est censée produire un apprentissage). Ainsi, le contrôle des tâches prescrites et réalisées est central dans l'évaluation, que ces tâches soient définies à l'intérieur ou à l'extérieur de l'EIAH évalué.

L'évaluation de l'utilité recouvre en fait trois niveaux de questions emboîtées, celles relatives au « comment », au « quoi », et au « pourquoi » ? Ainsi, un déficit en utilité d'un EIAH relève-t-il d'un problème de contenu ou de scénario didactique ? Peut-il relever, à un niveau supérieur, de l'inadéquation entre le dispositif et le format de la connaissance à acquérir : un savoir, un savoir-faire, une attitude, etc. (pour une synthèse dans le domaine des EIAH, voir [DE VRIES 01]). Finalement, l'objectif d'apprentissage est-il en adéquation avec les besoins de l'apprenant ?

Nous voulons insister sur le fait que l'évaluation prend en compte la mise en œuvre des processus cognitifs (voire métacognitifs) impliqués dans l'activité d'apprentissage. Une littérature conséquente dans le domaine des EIAH a fourni la preuve que la facilitation de la mise en œuvre des processus qui permettent l'apprentissage favorise celui-ci : on pense par exemple à la compréhension [MAYER 01], à l'analogie [SANDER & RICHARD 97], à la production d'hypothèses [LIM 98], etc. Une même littérature existe pour les niveaux métacognitifs [GORDON 96].

Pour l'évaluation empirique, le principe général repose sur la comparaison de performances de groupes de sujets soumis à des tâches (*e.g.* résolution de problèmes) avant et après l'utilisation d'un EIAH avec ou sans utilisation, ou sur deux versions de l'EIAH [FASTREZ 02]. Il n'y a pas de consensus sur la question de l'évaluation initiale des connaissances des sujets [ROMISZOWSKI 86]. En effet, des phénomènes comme l'effet d'attente, l'anticipation, la reconnaissance de l'évaluation a posteriori peuvent biaiser les résultats. Rouet et Passerault [ROUET & PASSERAULT 99] ont présenté et discuté les principales variables importées de la psychologie cognitive vers l'analyse des apprentissages dans les environnements informatisés. Ils abordent en particulier le problème du grain de l'évaluation (*i.e.* le degré de précision optimal de la mesure en fonction du but d'évaluation recherché). Nous renvoyons le lecteur à cet article et au volume dont il est issu pour des informations plus complètes sur ce sujet. Pour une présentation et une discussion de

la pertinence des tâches utilisées en évaluation dans le domaine des EIAH voir [TRICOT 02].

Pour l'évaluation par inspection, nous proposons de retenir sept critères (Tableau 1). Rappelons qu'un EIAH étant intégré dans une situation d'enseignement ou de formation, la situation peut pendre en charge certains aspects. Ainsi, certains critères pourront être satisfaits par la situation et non par l'EIAH lui-même.

2.2. Évaluation de l'utilisabilité

L'utilisabilité désigne la possibilité d'utiliser l'EIAH : sa maniabilité. L'utilisabilité d'un EIAH se joue au niveau de son interface (sa cohérence, sa lisibilité, la façon dont elle représente les actions possibles, etc.), de sa navigation (la cohérence, la simplicité, l'exhaustivité des déplacements possibles, etc.) et de sa cohérence avec l'objectif et le scénario didactiques. Elle est aussi fonction de l'adéquation entre les objectifs du concepteur et ceux de l'utilisateur [SPOOL 99].

L'approche empirique de l'évaluation de l'utilisabilité la plus célèbre est sans doute celle de Nielsen [NIELSEN 93]. Cette méthode consiste à réaliser des tests utilisateurs (observation du comportement et mesures d'efficacité suite à une tâche prescrite) et des entretiens avec ces mêmes utilisateurs. Cette méthode d'évaluation et les autres dans le domaine des IHM, fondées sur l'application de critères, ne sont pas spécifiques au domaine des apprentissages, mais il existe des travaux sur l'importation de ces méthodes d'évaluation dans le domaine des EIAH [JEAN 00], importation dont nous avons déjà souligné qu'elle est délicate. L'évaluation empirique peut être complétée par l'analyse des parcours réalisés par chaque utilisateur lors des tests (quels éléments de l'EIAH ont été traités, dans quel ordre, avec quel temps passé sur chaque élément). Pour une synthèse sur l'analyse des parcours et leur complémentarité avec les autres mesures dans le domaine des EIAH voir [DE VRIES & TRICOT 98], [TRICOT et al. 99].

Les critères pour l'évaluation par inspection de Bastien et Scapin [BASTIEN & SCAPIN 96] sont souvent cités dans le domaine de l'évaluation des IHM. Nous les rapportons dans le tableau 1. Leur efficacité et leur précision ont été évaluées et comparées avec d'autres, comme la norme ISO [BASTIEN et al. 99].

2.3. Évaluation de l'acceptabilité

Nous définissons l'acceptabilité d'un EIAH comme la valeur de la représentation mentale (attitudes, opinions, etc. plus ou moins positives) à propos d'un EIAH, de son utilité et de son utilisabilité. Cette représentation mentale peut être individuelle ou collective. La valeur de cette représentation conditionnerait la décision d'utilisation de l'EIAH. L'acceptabilité peut être sensible à des facteurs très divers comme la culture et les valeurs des utilisateurs, leurs affects, leur motivation, l'organisation sociale et les pratiques dans lesquelles s'insère plus ou moins bien l'EIAH.

Les mesures empiriques de l'acceptabilité relèvent de champs très divers. L'acceptabilité ne concerne pas que la représentation de l'EIAH mais aussi celles de son utilité et de son acceptabilité. Dans une étude récente, Amiel et ses collègues [AMIEL 02] ont montré que les représentations de l'utilité et de l'utilisabilité d'un EIAH pouvaient être les variables d'acceptabilité les plus prédictives de la décision d'utiliser ou de prescrire l'utilisation de cet EIAH.

Les critères d'évaluation par inspection sont présentés dans le tableau 1.

3. Relations entre les trois dimensions de l'évaluation

Dans ce chapitre, nous examinons les relations entre utilité, utilisabilité et acceptabilité. Nous admettons que, selon le contexte, l'évaluateur choisit de ne retenir que quelques critères sur chacune des dimensions et que chaque test permet de diagnostiquer si telle dimension est satisfaisante ou non. Dans un premier temps, nous examinons quelques relations standard. Ensuite, nous présentons notre propre modèle.

3.1. Le modèle de Nielsen

Pour Nielsen [NIELSEN 93] l'acceptabilité comporte deux dimensions : pratique et sociale. L'utilisabilité et l'utilité « théorique » (c'est-à-dire le but que le système est censé permettre d'atteindre) sont deux dimensions de l'utilité « pratique » (c'est-à-dire le but que le système permet effectivement d'atteindre). L'utilité pratique est elle-même une sous-dimension de l'acceptabilité pratique. Les relations entre acceptabilité, utilité et utilisabilité se représentent comme dans la figure 1.

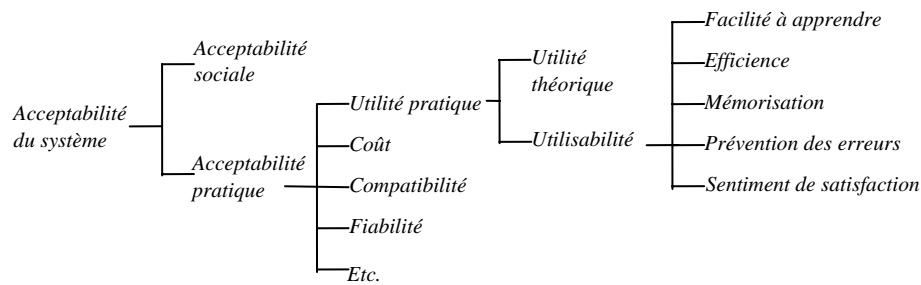


Figure 1. Le modèle de [NIELSEN 93], notre traduction

Il faut donc probablement lire le schéma de droite à gauche, chaque ensemble d'éléments de droite constituant une sorte de pré-condition pour l'élément placé à sa gauche. Si cette conception des relations entre les différentes dimensions de l'évaluation a valeur de référence aujourd'hui, quelques critiques peuvent être émises.

— Dans ce modèle, l'acceptabilité ne peut avoir d'effet sur l'utilisabilité ou l'utilité. Or [AMIÉL 02] analyse un exemple d'EIAH non accepté, qui, bien que n'ayant pas de défauts rédhibitoires en termes d'utilité et d'utilisabilité, se révèle inutilisé et inutile.

— Dillon et Morris [DILLON & MORRIS 96] notent qu'avec cette conception, l'utilisabilité, qui est pourtant au centre des préoccupations de Nielsen et de la plupart des évaluateurs, (1) apparaît comme une condition nécessaire mais non suffisante (2) ne permet de rien dire sur l'usage.

3.2. Le modèle de Dillon et Morris

Selon le modèle de Dillon et Morris, il faut introduire la notion d'attitude de l'utilisateur (perceptions, affects) pour pouvoir comprendre les relations entre utilité, utilisabilité et acceptabilité. En particulier, ces auteurs insistent sur la perception de l'utilité et de l'utilisabilité. Ces perceptions peuvent être très différentes entre elles et très différentes de l'utilisabilité effective ou de l'utilité pratique : je peux croire que cet EIAH sera facile à prendre en main mais peu utile pour moi, alors qu'il sera en réalité difficile à prendre en main mais utile au développement de mes compétences.

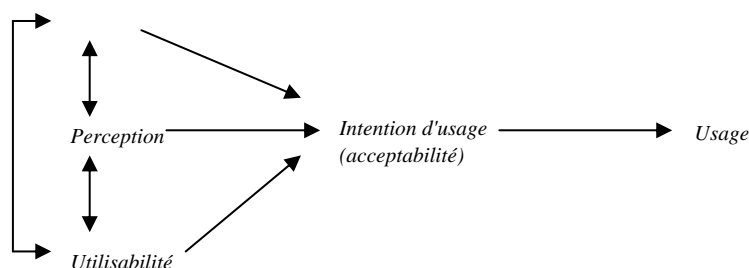


Figure 2. Le modèle de [DILLON & MORRIS 96], notre traduction

3.3. Le modèle Acceptabilité + Utilisabilité sont des pré-requis à l'utilité

De nombreux auteurs et praticiens, par exemple dans le domaine du marketing, semblent adopter un point de vue différent des deux premiers présentés. Pour eux, l'acceptabilité et l'utilisabilité sont des pré-requis à l'utilité [KETOLA & RÖYKKEE 01]. Une fois que l'EIAH est acceptable et utilisable, il est « prêt à être <éventuellement> utile ». Ainsi, l'utilité pratique peut devenir une dimension très floue, très peu définie, comme c'est le cas par exemple avec certains EIAH de type « ludo-éducatif », très bien vendus, dont on ne sait pas vraiment à quels apprentissages ils conduisent, mais qui sont indéniablement acceptés et utilisés. Parfois l'utilité n'est pas définie à l'avance mais créée par les usages. C'est le cas célèbre des *paggers* adoptés par les adolescents il y a quelques années pour une communication de proximité, ludique et relationnelle, sans qu'aucun concepteur de *paggers* ne l'ait jamais imaginé.

Une autre façon, assez proche, d'envisager la question, est de considérer l'acceptabilité comme un pré-requis à l'utilisation et l'utilisabilité comme pré-requis à l'utilité. Ce point de vue a été par exemple adopté dans le domaine de l'éducation où le manque de moyens matériels a été identifié comme la cause d'une non-motivation des enseignants à utiliser les nouvelles technologies, d'un manque d'expérience et d'aisance dans le domaine. La dotation d'ordinateurs à de nombreux établissements n'a pourtant pas réglé le problème de l'utilisation des TIC dans l'enseignement, peut-être parce que la relation des variables qui rendaient l'utilisation déficiente a été interprétée comme une relation causale. Cet exemple illustre d'une part que le lien entre les trois dimensions n'est pas aisé à évaluer et que, d'autre part, la spécificité de ce lien agit sur un ensemble qui doit être considéré comme tel du point de vue de l'évaluation. La question est de savoir quel modèle peut permettre de formaliser une évaluation capable de préciser quelle dimension (utilité, utilisabilité, acceptabilité) est en jeu dans la déficience d'un EIAH en spécifiant les probabilités de l'incidence de l'amélioration d'une dimension sur une autre et sur laquelle précisément. Par exemple, l'utilisabilité peut générer un problème d'acceptabilité au point de rendre non-opératoire le dispositif mais aussi, à l'inverse, avoir une incidence mineure sur les deux autres dimensions.

3.4. Le modèle proposé : toute relation formellement possible entre les trois dimensions peut exister

À partir de l'approche d'une relation non-définie à l'avance mais observée, nous proposons de formaliser par des évaluations (par inspection et empirique – préalable, en cours et *a posteriori* de la conception) un modèle des relations entre les trois dimensions en postulant que toute relation formellement possible entre les trois dimensions peut exister.

Ce que nous proposons d'interpréter est donc une série de résultats correspondant à un test à trois dimensions : utilisabilité, utilité, acceptabilité. Chaque dimension est donc une variable, sur laquelle on peut réaliser des mesures indépendantes des deux autres variables. Pour un même utilisateur-testeur, on peut noter ensemble les résultats obtenus selon les trois variables. Tout résultat à un test comportant les trois dimensions représente un point dans un espace à trois dimensions. Une série de résultats, obtenus avec n utilisateurs-testeurs, représente donc un ensemble de points. La forme de cet ensemble décrit la relation entre les trois dimensions.

Pour une première approche, imaginons le cas fictif d'une série de tests réalisés avec 15 utilisateurs-testeurs donnant comme résultat d'évaluation des valeurs binaires (-) et (+), (-) représentant un mauvais score et (+) un bon score. Il y aurait 8 triplets possibles {score utilité ; score utilisabilité ; score acceptabilité}. Une série de 15 tests est alors représentée dans un tableau par un ensemble de triplés et leurs effectifs respectifs (*cf.* Tableau 2).

Ainsi, on peut lire dans ce tableau que les scores positifs pour l'utilité sont 8/15, pour l'utilisabilité 4/15 et pour l'acceptabilité 3/15. Il conviendrait d'améliorer l'ensemble, mais peut-on s'attendre à ce qu'une amélioration de l'utilisabilité et/ou de

l'acceptabilité entraîne une amélioration de l'utilité ? Quand l'utilité est bonne, l'utilisabilité est autant mauvaise (n=4) que bonne (n=3+1). En revanche, quand l'utilité est mauvaise, l'utilisabilité est systématiquement mauvaise (n=7). L'analyse implicite permet de décrire les dépendances orientées entre de telles séries de variables [TRICOT & TRICOT 00]. Ainsi, l'analyse des scores d'utilisabilité et d'utilité peut être interprétée comme une relation d'implication Utilisabilité vers Utilité : l'utilisabilité n'est bonne que si l'utilité est bonne, mais une mauvaise utilisabilité est indépendante de l'utilité. L'analyse des autres résultats de ce cas fictif montre que Acceptabilité et Utilité sont dans une relation de tautologie tandis que Acceptabilité et Utilisabilité sont incompatibles. Conclusion : il faut améliorer l'acceptabilité et l'utilisabilité et les rendre compatibles. On peut raisonnablement espérer que l'amélioration de l'utilisabilité améliorera l'utilité.

		Acceptabilité			
		-		+	
		Utilisabilité		Utilisabilité	
		-	+	-	+
Utilité	-	5	0	2	0
	+	4	3	0	1

Tableau 2. *Table de contingence, récapitulatif d'une série fictive de 15 évaluations empiriques sur les trois dimensions : par exemple, la première cellule en haut à gauche indique que 5 sujets ont donné un score négatif aux 3 dimensions.*

De nombreuses autres techniques statistiques peuvent aussi permettre de décrire ces relations.

4. Conclusion

Dans ce papier nous avons défendu trois points de vue :

- l'utilité, l'utilisabilité et l'acceptabilité sont trois dimensions pertinentes et complémentaires dans l'évaluation des EIAH ;
- toute relation entre ces trois dimensions est *a priori* possible ;
- toute relation observée entre les trois dimensions est interprétable logiquement selon un degré de précision nécessaire et suffisant à une évaluation d'un EIAH.

Il nous semble qu'une approche itérative de l'évaluation au cours du processus de conception et qui s'inscrirait dans le cadre que nous proposons pourrait prévenir la grande majorité des défauts de l'EIAH conçu.

Notre approche mérite d'être affinée, notamment en ce qui concerne l'inférence statistique. Notre approche nécessite aussi d'être validée empiriquement, afin de prouver qu'elle permet d'identifier plus de défauts, de diagnostiquer plus d'erreurs que d'autres méthodes et de fournir des pistes d'amélioration plus pertinentes.

5. Bibliographie

- [AMIEL 02] AMIEL A., CAMPS J.-F., LUTZ G., PLEGAT-SOUTJIS F., TRICOT A., Acceptabilité de Form@lion : Évaluation et recommandations, rapport d'étude CERFI, octobre 2002, IUFM de Midi-Pyrénées, 25 p.
- [BASTIEN 99] BASTIEN J.M.C., SCAPIN D.L., LEULIER C., « The ergonomic criteria and the ISO/DIS 9241-10 dialogue principles : a pilot comparison in an evaluation task », *Interacting with Computers*, Vol. 11, n° 3, 1999, p. 299-322.
- [DE VRIES & TRICOT 98] DE VRIES E., TRICOT A., « Évaluer l'utilisation d'hypermédias : intérêts et limites des variables de performance », *Hypertextes et Hypermédias*, n° hors série, 1998, p. 175-190.
- [DE VRIES 01] DE VRIES E., « Les logiciels d'apprentissage, panoplie ou éventail ? », *Revue Française de Pédagogie*, n° 137, 2001, p. 105-116.
- [DILLON & GABBARD 98] DILLON A., GABBARD R., « Hypermedia as an educational technology : a review of the empirical literature on learner comprehension, control and style », *Review of Educational Research*, Vol. 68, n° 3, 1998, p. 322 - 334.
- [DILLON & MORRIS 96] DILLON A., MORRIS M., « User acceptance of information technology : theories and models », *Annual Review of Information Science and Technology*, 1996, p. 3-32.
- [FASTREZ 02] FASTREZ P., « La structuration d'un hypermédia éducatif influence-t-elle sur l'organisation des connaissances en découlant ? », *Colloque Compréhension et Hypermédia*, Albi, 10-11 octobre 2002.
- [GORDON 96] GORDON J., « Tracks for learning : Metacognition and learning technologies », *Australian Journal of Educational Technology*, Vol. 12, n° 1, 1996, p. 46-55.
- [JEAN 00] JEAN S., « Application de recommandations ergonomiques : spécificités des EIAO dédiés à l'évaluation », *Actes des Rencontres Jeunes Chercheurs en IHM*, mai 2000, pp. 39-42.
- [KEATES & CLARKSON 01] KEATES S., CLARKSON P.J., « Combining utility, usability and accessibility methods for universal access », *ACM CHI Workshop on Universal Design*, 2001, Washington, dec.
- [KETOLA & RÖYKKEE 01] KETOLA P., RÖYKKEE M., « The three facets of usability in mobile handsets Nokia », *CHI 2001*, Seattle, 31 March-5 April 2001.
- [LIM 98] LIM C.P., « The effect of a computer-based learning support package on the learning outcome of low-performance economics students », *The Virtual Edition*, Vol. 12, n°1, 1998.
- [MAYER 01] MAYER R.E., *Multimedia learning*, CUP, Cambridge, 2001.

- [NANARD & NANARD 98]. NANARD J., NANARD M., « La conception d'hypermédias », *Hypertextes et Hypermédias*, numéro hors série, 1998, p. 15-34.
- [NIELSEN 93] NIELSEN J., *Usability engineering*, Boston, Academic Press, 1993.
- [ROMISZOWSKI 86] ROMISZOWSKI A., *Developing auto-instructional materials*, Kogan Page, London, 1986.
- [ROUET & PASSERAULT 99] ROUET J.-F., PASSERAULT J.-M., « Analyzing learner-hypermedia interaction : An overview of online methods », *Instructional Science*, Vol. 27, n° 3/4, 1999, p. 201-219.
- [SANDER & RICHARD 97] SANDER E., RICHARD J.-F., « Analogical transfert as guided by an abstraction process : The case of learning by doing in text editing », *Journal of Experimental Psychology : Learning, Memory and Cognition*, Vol. 23, 1997, p. 1459-1483.
- [SCAPIN & BASTIEN 97] SCAPIN D.L., BASTIEN J.M.C., « Ergonomic criteria for evaluating the ergonomic quality of interactive systems », *Behavior & Information Technology*, Vol. 17, n° 4/5, 1997, p. 220-231.
- [SEARLE 01] SEARLE J., *Rationality in action*, Cambridge, MIT Press, 2001.
- [SENACH 90] SENACH B., Evaluation ergonomique des interfaces Homme/Machine : une revue de la littérature. Rapport INRIA n°1180, 1990.
- [SPOOL 99] SPOOL J., SCANLON T., SCHROEDER W., SNYDER C., DEANGELO T., *Web site usability : A designer's guide*, Morgan Kaufmann, San Francisco, 1999.
- [SQUIRES & PREECE 99] SQUIRES, D. PREECE, J., « Predicting quality in educational software : Evaluating for learning, usability and the synergy between them », *Interacting with Computers*, Vol. 12, 1999, p. 467-483.
- [TRICOT 99] TRICOT A., PUIGSERVER E., BERDUGO D., DIALLO M., « The validity of rational criteria for the interpretation of user-hypertext interaction », *Interacting with Computers*, Vol. 12, 1999, p. 23-36.
- [TRICOT & TRICOT 00] TRICOT A., TRICOT M., « Un cadre formel pour interpréter les liens entre utilisabilité et utilité des systèmes d'information (et généralisation à l'évaluation d'objets finalisés) », *Colloque Ergo-IHM 2000*, Biarritz, 3-6 octobre 2000.
- [TRICOT 02] TRICOT A., « IHM, cognition et environnements informatisés d'apprentissage », in G. Boy (Ed.), *L'ingénierie cognitive : IHM et cognition*, Hermès Science, Paris, 2002, p. 411-447.

Références sur le Web

- [BARON & LEVY 01] BARON G.-L., LEVY J.-F. Usages éducatifs des tic : quelles nouvelles compétences pour les enseignants ? publié en 2001 ; consulté en 2002. <http://www.inrp.fr/Tecne/Savoirplus/Rech40003/accueil.htm>