

L'évaluation des SAMI (système d'apprentissage multimédia interactif) : de la théorie à la pratique

Max Giardina, Denis Harvey et Martine Mottet

Volume 24, numéro 2, 1998

URI : <https://id.erudit.org/iderudit/502015ar>

DOI : <https://doi.org/10.7202/502015ar>

[Aller au sommaire du numéro](#)

Éditeur(s)

Revue des sciences de l'éducation

ISSN

0318-479X (imprimé)

1705-0065 (numérique)

[Découvrir la revue](#)

Citer cet article

Giardina, M., Harvey, D. & Mottet, M. (1998). L'évaluation des SAMI (système d'apprentissage multimédia interactif) : de la théorie à la pratique. *Revue des sciences de l'éducation*, 24(2), 335–353. <https://doi.org/10.7202/502015ar>

Résumé de l'article

Cet article traite de la mise au point d'une stratégie de collecte de données visant à regrouper et à analyser les facteurs sous-jacents au développement d'une interactivité significative dans des situations d'apprentissage médiatisé. Les auteurs ont conçu un modèle adapté d'analyse axée principalement sur une approche qualitative; ils ont utilisé à cette fin des méthodes intégrant la technique de l'entrevue, l'observation directe et l'analyse de protocoles. Celles-ci ont permis de cerner les facteurs perceptifs, transactionnels, cognitifs, pédagogiques et évaluatifs qui peuvent diriger la prise de décision des concepteurs de systèmes d'apprentissage multimédia interactifs. Des mises à l'essai de ce modèle à l'aide de concepteurs issus des milieux de formation scolaire ont indiqué la pertinence et le raffinement des analyses possibles.

L'évaluation des SAMI (système d'apprentissage multimédia interactif): de la théorie à la pratique

Max Giardina
Professeur

Denis Harvey
Professeur

Martine Mottet
Étudiante

Université de Montréal

Résumé – Cet article traite de la mise au point d'une stratégie de collecte de données visant à regrouper et à analyser les facteurs sous-jacents au développement d'une interactivité significative dans des situations d'apprentissage médiatisé. Les auteurs ont conçu un modèle adapté d'analyse axée principalement sur une approche qualitative; ils ont utilisé à cette fin des méthodes intégrant la technique de l'entrevue, l'observation directe et l'analyse de protocoles. Celles-ci ont permis de cerner les facteurs perceptifs, transactionnels, cognitifs, pédagogiques et évaluatifs qui peuvent diriger la prise de décision des concepteurs de systèmes d'apprentissage multimédia interactifs. Des mises à l'essai de ce modèle à l'aide de concepteurs issus des milieux de formation scolaire ont indiqué la pertinence et le raffinement des analyses possibles.

Introduction

L'évaluation d'un projet pédagogique informatisé comprend au moins quatre phases distinctes, soit l'étude du contexte d'utilisation, le relevé des intrants, la description des stratégies utilisées et l'évaluation du système informatisé (Gerard et Roegiers, 1994). Cette dernière étape, qui retiendra principalement notre attention dans le présent article, doit permettre l'analyse interne du produit et l'évaluation externe de son efficacité.

Les évaluateurs se posent généralement des questions telles que «Le système permet-il d'atteindre les objectifs pédagogiques de la leçon? Est-il pertinent, bien fait, efficace et techniquement adéquat?» Les évaluateurs doivent aussi être à l'affût d'effets inattendus du système, dont des apprentissages non prévus (Zahner, Reiser, Dick et Gill, 1992).

Malheureusement, aucune théorie de l'enseignement n'est assez fiable ou assez précise pour répondre à ces questions sans contrôle pratique. Les innovations pédagogiques même les plus prometteuses doivent donc être soumises à une évaluation rigoureuse pendant les phases de planification, de conception, de production et d'implantation. De plus, comme les phénomènes reliés à l'apprentissage sont très complexes et, de ce fait, difficiles à mesurer, il est important d'avoir recours à différents types d'instruments d'observation pour obtenir des données valides à partir desquelles on peut inférer. L'évaluation d'un système d'apprentissage multimédia interactif (SAMI) nous oblige par ailleurs à prendre en compte le nouveau rôle de l'étudiant, le concept d'interactivité significative et la multimodalité des messages présentés (Giardina, 1992; Giardina et Meunier, 1993).

Ce travail requiert des outils efficaces et une méthodologie de travail éprouvée (Flagg, 1990; Reeves, 1990, 1993). Or, Reiser et Kegelman (1994) soulignent que méthodes et outils (sélection des utilisateurs évaluateurs, critères d'évaluation, etc.) actuellement employés varient beaucoup selon les auteurs.

Nous rapportons ici les résultats de nos efforts en vue d'élaborer un modèle et des outils d'évaluation de ces systèmes complexes que sont les SAMI. Pour cela, nous examinons d'abord l'ensemble de la démarche d'évaluation pour mieux situer notre approche sur le plan théorique. Nous présentons ensuite le modèle et la matrice d'analyse originaux que nous proposons de même que les premières observations réalisées avec ce modèle. Cette expérimentation préliminaire nous permettra de conclure en soulignant les perspectives nouvelles que ce modèle laisse entrevoir.

Méthodes et outils d'évaluation des systèmes d'apprentissage multimédia interactifs

Terminologie retenue

Dans le cadre d'une réflexion théorique sur l'évaluation des SAMI, il est essentiel de distinguer dès le départ les termes suivants: analyse, évaluation formative ou régulatrice (Depover, 1994) et évaluation *a posteriori*.

L'analyse d'un système vise à mettre au jour le discours du document. L'évaluation formative ou régulatrice d'un SAMI se définit, quant à elle, comme une collecte organisée d'informations ayant pour but d'aider le concepteur à améliorer le *design* du système d'enseignement de manière à produire un prototype optimisé qui permettra d'atteindre les objectifs pédagogiques (Reeves, 1993). Enfin, l'évaluation *a posteriori* doit permettre de vérifier l'efficacité pédagogique du système (Romizovski, 1986). Il importe de faire ces distinctions théoriques, car l'approche retenue pour la collecte et l'analyse des données est fonction des objectifs poursuivis.

Ces approches, qui diffèrent notamment par leurs critères de mesure et le genre de données recueillies, peuvent être de nature qualitative ou quantitative. Qualitative si elle s'intéresse aux relations entre une multitude d'éléments plus ou moins complexes et fournit des listes de difficultés ou de suggestions; quantitative si elle répond à des questions précises et permet de récolter des données sous forme de note, de pourcentage et de taux d'erreur (*Ibid.*). Comme chacune porte sur des aspects particuliers du problème, le choix de la méthodologie d'observation dépend des objectifs de l'évaluation (Salomon, 1991).

Si on vise l'optimisation du système, comme c'est généralement le cas dans l'évaluation formative ou régulative, on optera pour une approche plutôt prescriptive, systémique et, partant, qualitative. Si, par contre, on cherche à recueillir des informations descriptives et analytiques sur un sujet précis, comme c'est le cas dans l'évaluation *a posteriori*, on favorisera une approche empirique, systématique et donc quantitative (Salomon, 1991; Tripp et Bichelmeyer, 1990).

L'évaluation du système: quoi, quand et par qui?

Dans ses grandes lignes, l'évaluation d'un SAMI consiste à répondre à trois questions principales: «quoi évaluer? quand évaluer? et qui doit évaluer?» (Depover, 1994). Une fois cela déterminé, le «comment évaluer» va de soi.

Quoi évaluer? – Bien qu'il n'y ait pas encore de consensus au sujet des éléments clés de l'évaluation d'un système informatisé d'apprentissage, certains principes de base semblent se dégager de la documentation scientifique.

Ainsi, Barbara Flagg (1990) estime que l'évaluation complète d'un système informatisé d'apprentissage doit comprendre obligatoirement une critique de sa convivialité et de sa réception par les utilisateurs ainsi qu'une évaluation de son efficacité pédagogique.

En effet, un système performant à cet égard permettra à l'étudiant de se concentrer sur l'apprentissage du contenu plutôt que sur le fonctionnement du système pour une plus grande efficacité pédagogique. En accord avec ces affirmations, Reeves (1992) souligne que l'efficacité d'un SAMI est principalement limitée par deux facteurs clés: la qualité du *design* de l'interface, qui devrait assurer une interactivité maximale, et la motivation et le degré de compétence de l'apprenant par rapport au système. Tolhurst (1992) estime, pour sa part, qu'un bon dosage du caractère directif et la présence d'aides pertinentes à la navigation et de liens clairs entre les nœuds d'information sont des éléments fondamentaux d'un système bien structuré.

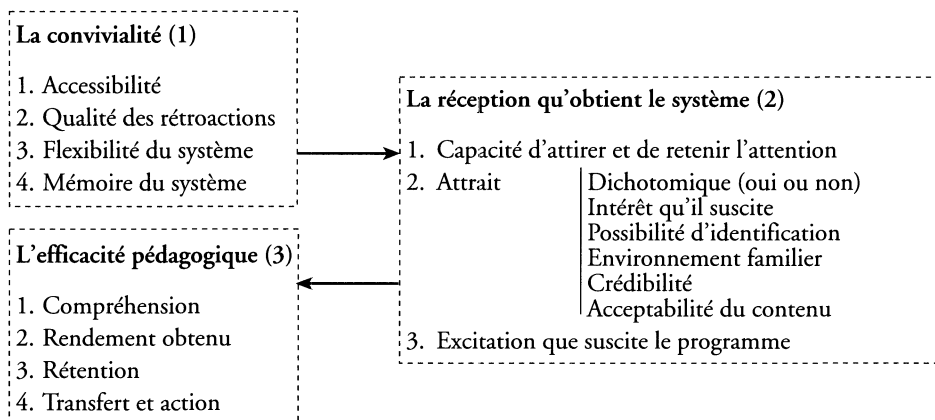


Figure 1 – Les éléments à évaluer dans un SAMI (selon Flagg, 1990)

À notre avis, bien d'autres facteurs doivent être pris en considération et nous y reviendrons lors de la présentation de notre modèle multicritérié.

Quand évaluer? – L'analyse ou l'évaluation *a posteriori* d'un système d'apprentissage informatisé à l'aide d'une méthodologie analytique de type expérimental ou quasi expérimental devrait être faite seulement une fois le système optimisé grâce à plusieurs séances d'évaluation régulative (Reeves, 1992).

Romizovski (1986), qui abonde dans le même sens, conseille d'évaluer le système en cours de production afin d'y apporter rapidement les correctifs nécessaires. Selon cet auteur, un prototype d'enseignement individualisé devrait être soumis à l'évaluation régulative dès qu'il y a suffisamment de matériel à présenter aux étudiants. Au début de la phase de production, on devrait vérifier la qualité de l'interaction (convivialité et réception du système) entre le prototype et l'utilisateur. L'efficacité du prototype sera testée plus tard à l'aide de prétests et de post-tests.

Qui doit évaluer? – Choisir de faire l'évaluation d'un système avant, pendant ou à l'issue du développement détermine non seulement ce qui est évalué mais aussi, dans une certaine mesure, qui procède à l'évaluation (Depover, 1994).

Plusieurs études ont montré que les professeurs, les technologues de l'éducation et les étudiants n'utilisent pas les mêmes critères pour évaluer les systèmes d'apprentissage informatisés. Les professeurs apprécient généralement les programmes fiables et conviviaux; les technologues de l'éducation mettent la priorité sur les fondements pédagogiques, l'utilisation adéquate des technologies disponibles et l'intégration du système dans le curriculum; les étudiants s'attardent surtout à la convivialité et à la réception du système (Flagg, 1990; Reiser et Kegelman, 1994; Zahner *et al.*, 1992).

L'évaluation d'un SAMI par des experts de divers domaines (contenu, *design*, graphisme, média, informatique, etc.) offre plusieurs avantages. En effet, les experts

peuvent intervenir très tôt dans la phase de production et même de planification, car ils n'ont pas besoin du produit fini pour faire des commentaires et des suggestions. Leur évaluation est généralement rapide et peu coûteuse. Par contre, ils sont souvent biaisés par leurs expériences personnelles et ont habituellement tendance à analyser un programme dans ses grandes lignes en prenant surtout en compte la qualité de son contenu et de son *design*. Pour sa part, l'étudiant, qui est généralement peu habile à identifier ces grandes lignes, s'attarde plutôt aux détails (Flagg, 1990). Experts et étudiants jouent donc des rôles complémentaires dans l'évaluation d'un SAMI et le produit fini gagnerait sans doute en qualité si, dès le début, les futurs utilisateurs d'un tel système étaient associés à son développement. (Depover, 1994).

Par ailleurs, pour inférer à partir des résultats de l'évaluation régulative d'un SAMI réalisée avec un nombre relativement restreint d'évaluateurs généralement choisis au hasard, il est essentiel que la population cible soit homogène. Moins elle l'est, plus le nombre d'évaluateurs devra être important. Il est aussi crucial de choisir les étudiants en fonction de la richesse potentielle de leur rétroaction (Boutin et Chinien, 1992) : il a été démontré que trois étudiants qui interagissent directement avec le responsable de l'évaluation fournissent autant d'information que huit à douze étudiants répondant individuellement à un questionnaire (Flagg, 1990).

Finalement, Romizovski (1986) suggère que l'évaluation régulative soit menée par le concepteur du prototype et non par des spécialistes de l'évaluation.

Méthodologies et outils

Nous avons indiqué plus haut que l'évaluation d'un SAMI peut procéder de trois objectifs différents mais complémentaires. Comme ils influencent grandement le choix de la méthodologie et des outils, il importe d'en discuter séparément.

— Objectif d'analyse

Nous avons souligné que l'analyse vise à mettre au jour le discours d'un document, à en saisir le sens plutôt qu'à porter un jugement de valeur comme on le fait avec une évaluation *a posteriori*. L'analyse est l'envers de la conception en ce qu'elle relève d'une démarche de déconstruction. Elle permet de constater ce qui est et non ce qui devrait être.

Malheureusement, les entreprises d'évaluation *a posteriori* font souvent l'économie de la phase d'analyse pour procéder directement à l'évaluation de l'efficacité du système. Pourtant, les critères d'évaluation devraient émerger de la connaissance des différents aspects d'un système. Autrement dit, les résultats de l'analyse peuvent fournir la matière première de l'évaluation *a posteriori* du système.

Nous constatons que, dans la plupart des méthodes d'analyse des SAMI, on demande à des évaluateurs externes de passer en revue un nombre très variable d'éléments du système selon la méthodologie retenue. Par ailleurs, la terminologie relative aux critères d'analyse change beaucoup d'un auteur à l'autre (Reiser et Kegelman, 1994). Comme nous le verrons plus loin, nous proposons d'analyser un grand nombre d'éléments avec notre outil.

— Objectif d'évaluation régulative du système

En accord avec Depover (1994), nous préférons parler d'évaluation régulative des systèmes et non d'évaluation formative pour éviter la confusion avec les méthodologies d'évaluation formative des connaissances.

D'après Romizovski (1986), pour faire une bonne évaluation régulative d'un système d'enseignement individualisé, il faut prévoir au moins quatre révisions consécutives des versions améliorées de l'objet en question par autant de groupes d'étudiants constitués d'au moins deux ou trois étudiants par groupe d'évaluateurs. Toutes les informations recueillies doivent être analysées en profondeur.

Au cours de cette phase de l'évaluation, on utilise souvent une approche qualitative comprenant généralement deux techniques de collecte de données, soit le rapport individuel d'appréciation et l'observation des utilisateurs en action. Le rapport individuel, qui est le plus utilisé, peut prendre plusieurs formes: réponse à un questionnaire écrit ou informatisé, participation à une entrevue ou le «commenter et penser à haute voix» pendant l'utilisation du système. L'observation des étudiants au travail exige beaucoup de temps et d'effort, mais elle permet de recueillir des informations que ne fournissent habituellement pas les experts. La technique du «commenter et penser à haute voix» est très efficace pour évaluer la convivialité d'un SAMI du point de vue de l'utilisateur. Le biais de l'observation subjective directe est alors contrebalancé par l'interprétation personnelle de l'étudiant. Par contre, cette technique, qui est peu naturelle, peut amener certains étudiants à dire des choses qu'ils ne pensent pas vraiment. Enfin, la plupart des évaluateurs auront tendance à interrompre leurs commentaires lorsque la tâche à effectuer requiert plus d'attention (Flagg, 1990). L'observation structurée ou non des utilisateurs est aussi très prisée. Non structurée, elle vise à évaluer le comportement général d'un étudiant pendant une certaine période de temps d'utilisation du SAMI et fournit donc des informations générales. Structurée, elle cherche à enregistrer un comportement précis en réponse à un *stimulus* déterminé. Pour pouvoir discuter des résultats obtenus à partir d'un petit nombre de répondants avec assez de confiance et sans l'aide d'analyses statistiques complexes, on doit mesurer de plusieurs manières chaque élément évalué.

— Objectif d'évaluation *a posteriori*

Étant donné les problèmes de fiabilité associés aux méthodes d'évaluation subjectives des systèmes informatisés de formation, plusieurs auteurs ont souligné

l'importance de toujours en vérifier l'efficacité pédagogique sur le terrain (Flagg, 1990; Reeves, 1992; Reiser et Dick, 1990; Romizovski, 1986).

En effet, on constate souvent une grande divergence selon la méthode employée; l'opinion des évaluateurs au sujet de la qualité ou de la pertinence d'un système d'apprentissage peut changer radicalement à la suite d'une évaluation sur le terrain (Owston et Wideman, 1987). En outre, l'évaluation subjective par le professeur ou par les étudiants est rarement un bon indicateur de l'efficacité pédagogique du système (Jolicœur et Berger, 1988).

Romizovski (1986) a proposé une approche quantitative détaillée pour l'évaluation *a posteriori* d'un SAMI dont voici les grandes lignes.

On doit d'abord soumettre les étudiants à un examen pour s'assurer qu'ils possèdent les préalables nécessaires. On leur fait ensuite subir un prétest sur la matière contenue dans le prototype afin d'évaluer leurs connaissances à l'entrée. Seuls les étudiants qui obtiennent près de 100 % à l'épreuve des préalables et une très faible note aux prétests doivent être sélectionnés. Après avoir utilisé le système pendant le temps requis pour atteindre les objectifs de la leçon, les étudiants passent un post-test équivalent au prétest au regard du contenu, du degré de difficulté et, bien sûr, des objectifs spécifiques... Plus de 90 % des étudiants doivent obtenir plus de 90 % (règle du double 90). Il peut être utile de diviser le post-test en deux parties dont la première se déroule immédiatement après l'explication d'un concept et la deuxième, à la fin de l'utilisation du prototype. Une bonne rétention initiale qui se dégrade en bout de ligne peut signifier que l'explication est sans doute adéquate, mais la pratique ou la révision, insuffisantes. Une faible rétention aux deux étapes dénote une explication déficiente (Romizovski, 1986).

Malgré sa précision, ce type de validation a rarement un impact sur le prototype de formation (Depover, 1994) puisqu'on l'applique généralement en toute fin de production, soit à un moment où les derniers tours en arrière sont malaisés. En outre, par leur caractère ouvert et la place qu'ils font aux initiatives de l'apprenant (simulation, exploration libre, etc.), les SAMI s'accommodent généralement mal d'une validation des seules compétences maîtrisées en fonction d'une liste de critères stricts prévus au départ: les apprentissages dépendent dans une large mesure des décisions et des préférences de l'apprenant lui-même.

Évaluation normative et évaluation critériée

Parce qu'ils permettent d'appréhender le réel dans toute sa complexité et limitent la part de subjectivité, seuls des outils précis, exhaustifs et validés devraient être retenus pour l'analyse en détail d'un système. En effet, pour des raisons évidentes, une approche uniquement qualitative peut se traduire par une grande variabilité dans l'évaluation et la pondération de chaque élément (Reiser et Kegelman, 1994).

Une autre dérive possible en l'absence de critères précis consiste à juger par référence à un autre système évalué antérieurement et donc à substituer l'appréciation normative à l'évaluation critériée (Depover, 1994).

Par ailleurs, comme les SAMI mettent souvent l'accent sur la résolution de problèmes ou la construction du savoir, il est difficile de cerner précisément la qualité du produit et les progrès réels accomplis par les étudiants avec les techniques traditionnelles d'évaluation de logiciels pédagogiques. On le voit, la mise au point de techniques et d'outils d'évaluation adaptés à ces systèmes complexes s'impose (Zahner *et al.*, 1992).

Vers un modèle multicritérié tridimensionnel

Le modèle tridimensionnel proposé vise précisément l'analyse méticuleuse et l'évaluation critériée des SAMI. Une fois complété et validé, ce modèle devrait nous fournir les instruments requis tant pour l'analyse des SAMI que pour leur évaluation régulatrice ou *a posteriori*.

Dans ce modèle composé de trois axes, soit le niveau d'intervention, les critères de qualité et les éléments structuraux (voir figure 2), nous nous efforçons de rendre compte des relations dynamiques à l'intérieur de la structure et du discours d'un SAMI.

Le modèle acquiert une signification contextuelle plus précise lorsque les axes sont envisagés en relation les uns avec les autres. Nous décrivons maintenant les différents critères généraux associés à chacun de ces trois axes.

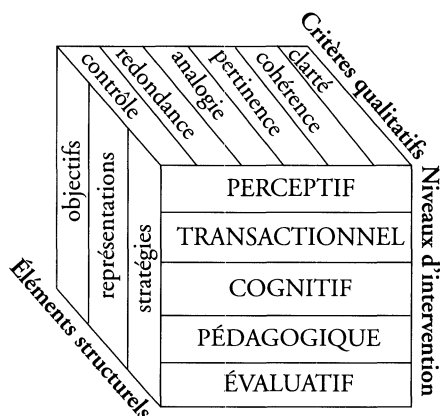


Figure 2 – Modèle tridimensionnel

Niveaux d'intervention

L'intervention peut se faire à cinq niveaux, c'est le premier axe de ce modèle (figure 2).

Au niveau perceptif, l'attention de l'évaluateur est retenue par les caractéristiques physiques des messages (image, son, texte écrit, narration, etc.) transmis par le système, caractéristiques qui déterminent quel canal sera sollicité. L'évaluateur se préoccupe de vérifier s'il y a adéquation entre l'organisation des informations, c'est-à-dire le type de message choisi, et le type d'interaction visé. En effet, cette adéquation est garante d'une meilleure perception des informations. L'intervention sur le plan perceptif, qui est constitué de l'ensemble des aspects sensoriels d'un système multimédia, vise donc l'atteinte d'une correspondance optimale entre l'interface et les objectifs d'interactivité en fonction des particularités des différents canaux de l'éventuel utilisateur.

Au niveau transactionnel, il s'agit de vérifier comment le système signifie à l'apprenant qu'une action est en cours, est activable ou a été activée, et quel canal devrait être privilégié pour instaurer le plus riche possible dialogue entre l'individu et le système.

Au niveau cognitif, l'évaluateur doit identifier quels outils cognitifs sont mis à la disposition de l'apprenant pour faciliter et optimiser son travail intellectuel. Il jugera le choix des stratégies, l'organisation du contenu, la représentation des connaissances et les procédures de modélisation et d'intervention.

Au niveau pédagogique, l'évaluateur doit identifier et évaluer les choix des concepteurs quant à la nature des objectifs d'apprentissage et aux modalités d'apprentissages offertes à l'utilisateur. On se préoccupe ici du discours pédagogique, de son orientation, de ses différentes manifestations et des réajustements possibles.

Au niveau évaluatif, l'attention porte sur la présence et la qualité des mécanismes diagnostiques, qui permettent d'analyser les interactions entre le système et l'apprenant, puis de suggérer de nouveaux échanges pour les améliorer.

Critères qualitatifs

Après une analyse approfondie (Norman, 1988; Shneiderman, 1992) des critères utilisés dans des modèles existants (Apple, Microsoft, Zenith), nous avons retenu les six critères qui nous paraissent les plus pertinents pour caractériser les choix de conception en ce qui a trait aux interventions et aux éléments structuraux qui s'y rattachent. Grâce à ces six critères qualitatifs (voir figure 2), on peut juger si ces choix de conception ont permis de raffiner, de diversifier et de rendre plus significatifs les échanges entre le système et l'apprenant.

La clarté, définie de façon abstraite comme la qualité de ce qui est facilement intelligible, est associée à un souci de netteté et de transparence.

La cohérence est définie par l'union étroite et l'harmonie entre les divers éléments d'un système.

La pertinence équivaut à la qualité de ce qui convient pour l'objet, c'est-à-dire ce qui est conforme à l'objet et au bon sens.

L'analogie se définit comme la ressemblance établie par l'imagination entre deux ou plusieurs objets de pensée différents.

La redondance se produit quand une information déjà donnée revient sous une autre forme. C'est l'abondance dans le discours, c'est-à-dire l'augmentation ou la diversification des signes sans accroissement corrélatif de la quantité d'information contenue dans le système.

Le contrôle sert à vérifier ou à surveiller. Ce critère intègre des aspects de conduite, de maîtrise, de direction et de commande.

Éléments structuraux

Les trois éléments qui composent ce dernier axe sont les objectifs, les représentations et les stratégies.

Les objectifs d'apprentissage décrivent de quelle manière l'apprenant sera transformé au cours de son interaction avec le système. Les objectifs du système doivent constituer un ensemble articulé et cohérent auquel l'évaluateur peut se référer en tout temps.

L'évaluateur doit aussi s'intéresser aux représentations; il doit déterminer comment la réalité est représentée et si on a recours aux stratégies appropriées (voir *infra*) pour amener l'apprenant à construire ses propres représentations. C'est en effet au moment de la conception qu'on détermine la nature des représentations qu'on souhaite voir se construire chez l'apprenant et, conséquemment, la nature des représentations dans le système qui, elles-mêmes, sont fonction des différents éléments de l'environnement multimédia interactif.

Quand il s'agit des stratégies, l'effort d'évaluation porte principalement sur l'identification et la critique des stratégies de communication, d'accès et de manipulation des connaissances organisées dans le système qui ont été privilégiées par le concepteur. Ces stratégies dites d'apprentissage, d'enseignement ou cognitives peuvent être mises en action par le système ou par l'apprenant (Boulet, Savoie-Zajc et Chevrier,

1995) lorsque ce dernier désire interagir en situation d'apprentissage. Il appartient au concepteur du système de prévoir des conditions favorables au déclenchement par l'apprenant de ces stratégies souvent latentes.

Relativement aux stratégies d'apprentissage, mentionnons la construction, la découverte, l'induction, l'expérimentation, l'analogie, le transfert, la systématisation, la mémorisation et l'implantation. Pour leur part, les stratégies d'enseignement portent davantage sur les modes de communication pédagogique d'un contenu et sa mise en forme générale; elles visent donc à présenter, à informer, à sensibiliser, à compléter, à questionner, etc. Enfin, les stratégies cognitives sont intimement liées aux stratégies d'apprentissage et d'enseignement, mais elles se réfèrent de manière plus précise aux processus internes, activés par l'apprenant pour se représenter, structurer et emmagasiner l'information qu'il reçoit. West, Farmer et Wolff (1991) les ont classées en stratégies organisationnelles, stratégies spatiales (schémas, plans), stratégies de raccord (métaphores, analogies) et stratégies multi-usages (mnémotechniques, imagerie, rappel). Ici, l'évaluateur doit aussi s'interroger non seulement sur la stratégie globale, mais aussi sur les stratégies partielles retrouvées dans le système.

Matrice

Le modèle multicritérié tridimensionnel est mis en œuvre par une matrice formée des trente croisements reliant les niveaux d'intervention et les critères qualitatifs. Nous avons donc six croisements pour le niveau perceptif : perceptif/clarté, perceptif/cohérence, perceptif/pertinence, perceptif/analogie, perceptif/redondance, perceptif/contrôle et ainsi de suite pour chacun des cinq niveaux d'intervention et des six critères qualitatifs. Les croisements permettent également d'examiner les éléments structuraux du SAMI, soit les stratégies, les représentations et les objectifs.

Chaque croisement est composé d'une série d'énoncés grâce auxquels l'évaluateur passe minutieusement en revue le SAMI et met au jour son discours. Dans sa présente version, la matrice compte 265 énoncés formulés en termes opérationnels. Les énoncés nombreux, précis et objectifs visent à minimiser la subjectivité inhérente à une évaluation trop générale ou globale et à rendre compte de la complexité des SAMI. Voici quelques exemples tirés de ces croisements: «Le pointeur de la souris se repère facilement sur les textes et les images» (perceptif/clarté); «Les outils de navigation permettent de comprendre par où on est déjà passé dans le SAMI» (transactionnel/clarté); «Les pouvoirs des apprenants sont les mêmes dans toutes les occurrences du même type de tableau» (cognitif/cohérence); «Les avantages et les limites des analogies et des simulations sont clairement précisés» (pédagogique/ analogie).

On trouvera dans les pages suivantes le court texte explicatif, destiné à l'évaluateur, qui précède la présentation de chacun des niveaux d'intervention de même que l'exemple d'un croisement pour chacun de ces niveaux.

En attribuant une note à chacun des énoncés, l'évaluateur dresse un portrait précis des caractéristiques du système et en fait l'évaluation critériée. L'échelle de notation comporte cinq échelons indiquant à quelle fréquence l'énoncé est vrai: jamais, parfois, souvent ou toujours. Le cas échéant, l'évaluateur peut aussi indiquer que l'énoncé est sans objet (S.O.) pour le système étudié.

Une brève définition et l'objectif de chaque croisement sont fournis afin de faciliter la compréhension de la matrice par l'évaluateur. Ainsi, la définition du croisement «Niveau pédagogique/ Redondance», se lit comme suit: «Le critère de redondance est atteint si les activités, les stratégies et les parcours d'apprentissage possibles sont bien diversifiés.»

On attire également l'attention de l'évaluateur sur les éléments du SAMI qui font l'objet du croisement (paramètres physiques, éléments des tableaux, propriétés sémantiques, stratégies pédagogiques, etc.). D'ailleurs, dans chacun des croisements, les énoncés sont regroupés en fonction des objets à évaluer. Ainsi, dans le croisement «niveau pédagogique/clarté», les énoncés sont classés en trois sous-catégories: objectifs, stratégies, contenu et activités.

À la rubrique «explication», on trouve des éléments théoriques susceptibles d'étayer l'évaluation du croisement (quelles activités d'apprentissage sont généralement associées aux types de connaissances à acquérir, etc.) ou encore des précisions sur la liste des caractéristiques à considérer (sexe, âge, culture, connaissance du domaine, culture informatique du public cible, etc.).

Enfin, les exemples et les contre-exemples aident l'évaluateur à saisir quels objets il doit examiner.

Comme nous le rapportons plus loin, nous avons pu constater, au cours de nos expérimentations, la pertinence de présenter la matrice sous cette forme.

Fondements théoriques et méthodologie d'élaboration

Habituellement, la méthodologie retenue pour faire l'évaluation des SAMI est plutôt empirique (Flagg, 1990). Par contre, comme on a pu le constater plus haut, notre matrice fait référence clairement aux sciences cognitives en général et au cognitivisme en particulier. En effet, le modèle que nous proposons utilise une approche résolument cognitive en ce sens que ses paramètres se réfèrent à un apprentissage qui favorise un processus de raffinement (West *et al.*, 1991) plutôt qu'une construction des référents. Elle tire également parti des principes directeurs généralement reconnus en ergonomie des interfaces humain-ordinateur.

Pour la réaliser, nous avons d'abord examiné un éventail de documents issus d'expériences ou d'études et liés de manière générale ou spécifique aux cinq niveaux d'intervention de notre modèle. Nous avons ainsi consulté des documents sur des modèles

d'évaluation des systèmes informatisés d'apprentissage, l'ergonomie des interfaces (Booth, 1989; Horn, 1989; Mayhew, 1992; Robert, 1988; Shneiderman, 1992), le multimédia et la télévision ainsi que sur la psychologie cognitive, notamment ses impacts en enseignement (Tardif, 1992), en *design* pédagogique (West *et al.*, 1991) et en évaluation des apprentissages (Scallan, 1988).

Pour chacun des niveaux d'intervention, les informations recueillies par critère qualitatif ont été classées. Nous avons également pris soin de les formuler de manière simple, claire, directe, précise et objective.

L'originalité de notre travail réside dans l'intégration des résultats de travaux effectués, d'une part, en pédagogie (niveaux d'intervention) et, d'autre part, en conception d'interfaces humain-ordinateur (critères qualitatifs). Il a fallu enrichir l'information recueillie sur les études et les principes directeurs, c'est-à-dire que l'information portant sur les interfaces humain-ordinateur ou sur la pédagogie a été revue en fonction du multimédia et des systèmes d'apprentissage. La documentation offre peu d'exemples d'un traitement particulier des éléments afférents au multimédia dans le cadre d'une démarche d'analyse et d'évaluation. Nous avons donc proposé d'autres énoncés en nous fondant sur notre réflexion théorique et sur notre expérience pratique issues de la conception, de l'analyse et de l'évaluation de plusieurs SAMI.

Enfin, dans le but de faciliter l'appropriation et la diffusion de la matrice, nous avons adopté la terminologie technique généralement acceptée en multimédia, télévision et ergonomie des interfaces (Vanderdonckt, 1995).

Expérimentations

Au cours des trois dernières années, nous avons mené dix expériences en entreprise et en milieu universitaire auprès d'étudiants du programme de maîtrise en technologie éducationnelle de l'Université de Montréal. Nous avons d'abord examiné avec les participants le modèle multicritérié tridimensionnel pour nous assurer de la bonne compréhension des objectifs visés et des principes théoriques sous-jacents. Les participants ont ensuite utilisé la matrice pour analyser puis évaluer les décisions de conception d'un SAMI. Par la suite, les étudiants à la maîtrise ont conçu un SAMI et mis au point un prototype. Au cours de la dernière expérience, nous avons aussi observé une douzaine d'étudiants pendant leurs séances d'analyse et d'évaluation, et avons recueilli leurs opinions sur la matrice au moyen d'un questionnaire écrit.

De manière générale, les participants ont réservé un bon accueil à la grille. En entreprise, on a particulièrement souligné combien elle venait combler un vide gênant. Si les évaluateurs relèvent souvent l'importante quantité de temps nécessaire pour passer en revue tous les énoncés (environ 10 heures), il ressort toutefois clairement que la matrice permet à la fois de dresser la liste des caractéristiques d'un SAMI et d'obtenir une bonne vue d'ensemble de la qualité des échanges potentiels entre l'apprenant et le système.

Par ailleurs, pour faciliter l'analyse des données recueillies, nous avons mis au point, pour notre dernière expérience, un outil très simple constitué d'un tableur permettant de générer automatiquement des graphiques (figures 3 à 6). Ce tableur permet d'attribuer une note aux énoncés des croisements, de ramener le résultat de chaque croisement à une note sur 100, puis d'établir la moyenne pour chacun des niveaux et chacun des critères. Les graphiques permettent de visualiser les résultats. Cet outil permet donc d'identifier rapidement et globalement les niveaux ou les critères qui présentent le plus de points forts ou le plus de points faibles: alors que les énoncés de la matrice sont regroupés par niveau, le tableur permet d'obtenir une vue par critère (croisée).

ASPECT PERCEPTIF											
Clarté		Cohérence		Pertinence		Analogie		Redondance		Contrôle	
Énoncé	Note	Énoncé	Note	Énoncé	Note	Énoncé	Note	Énoncé	Note	Énoncé	Note
n° 1	5	n° 1	4	n° 1	3	n° 1	1	n° 1	5	n° 1	4
n° 2	5	n° 2	4	n° 2	3	n° 2	1	n° 2	5	n° 2	4
n° 3	5	n° 3	4	n° 3	3	n° 3	1	n° 3		n° 3	4
n° 4	5	n° 4	4	n° 4	3	n° 4	1	n° 4		n° 4	4
n° 5	5	n° 5	4	n° 5	3	n° 5		n° 5		n° 5	4
n° 6	5	n° 6	4	n° 6	3	n° 6		n° 6		n° 6	4
n° 7	5	n° 7	4	n° 7	3	n° 7		n° 7		n° 7	
n° 8	5	n° 8	4	n° 8	3	n° 8		n° 8		n° 8	
n° 9	5	n° 9	4	n° 9	3	n° 9		n° 9		n° 9	
n° 10	5	n° 10		n° 10	3	n° 10		n° 10		n° 10	
n° 11	5	n° 11		n° 11		n° 11		n° 11		n° 11	
n° 12	5	n° 12		n° 12		n° 12		n° 12		n° 12	
n° 13		n° 13		n° 13		n° 13		n° 13		n° 13	
n° 14		n° 14		n° 14		n° 14		n° 14		n° 14	
n° 15		n° 15		n° 15		n° 15		n° 15		n° 15	
n° 16		n° 16		n° 16		n° 16		n° 16		n° 16	
n° 17		n° 17		n° 17		n° 17		n° 17		n° 17	
n° 18		n° 18		n° 18		n° 18		n° 18		n° 18	
n° 19		n° 19		n° 19		n° 19		n° 19		n° 19	
n° 20		n° 20		n° 20		n° 20		n° 20		n° 20	
n° 21		n° 21		n° 21		n° 21		n° 21		n° 21	
n° 22		n° 22		n° 22		n° 22		n° 22		n° 22	
Total	58		36		30		4		10		24
/100	97		80		60		20		100		80

Figure 3 – Exemple du tableur pour le niveau perceptif

RÉSUMÉ							
	Clarté	Cohérence	Pertinence	Analogie	Redondance	Contrôle	Moyenne
Perceptif	96,67	80,00	60,00	20,00	100,00	58,00	69,11
Transactionnel	89,47	71,43	80,00	75,00	58,00	58,00	71,98
Cognitif	75,45	70,00	75,00	50,00	85,45	71,43	71,22
Pédagogique	77,78	80,00	70,00	67,27	85,45	69,23	74,96
Évaluatif	62,50	80,00	66,67	67,27	80,00	69,23	70,95
Moyenne	80,37	76,29	70,33	55,91	81,78	65,18	71,64

Figure 4 – Exemple du tableur pour l'ensemble de la matrice

L'observation du plus récent groupe d'évaluateurs et leurs réponses au questionnaire écrit sur la matrice mettent en évidence combien les définitions et surtout les exemples et les contre-exemples orientent l'analyse du SAMI et facilitent la compréhension des énoncés. En effet, plusieurs évaluateurs ont tendance à passer directement aux énoncés sans lire les consignes, les définitions, les exemples, etc. Cependant, dès qu'ils éprouvent des difficultés à évaluer les énoncés, ils reviennent aux consignes pour y puiser de l'information ou une piste de travail.

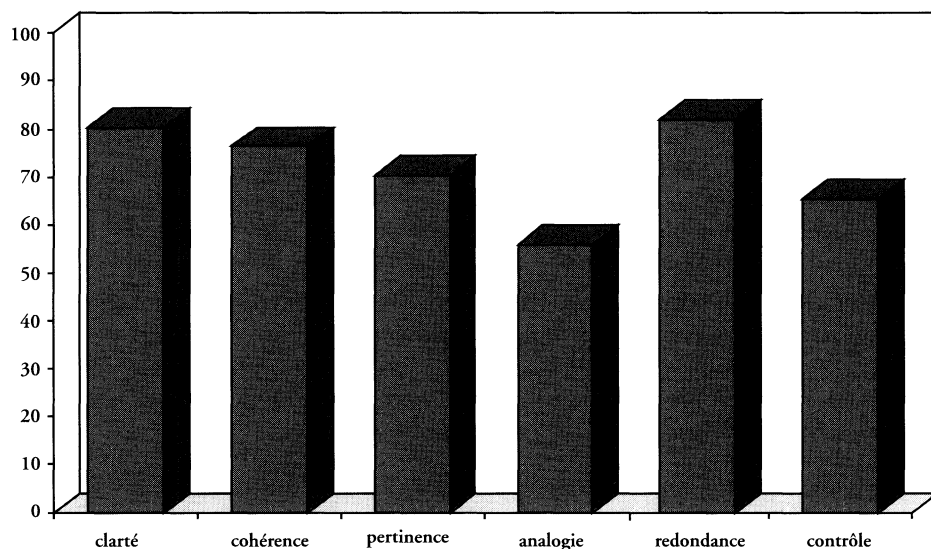


Figure 5 – Exemple de graphique pour la moyenne par critère

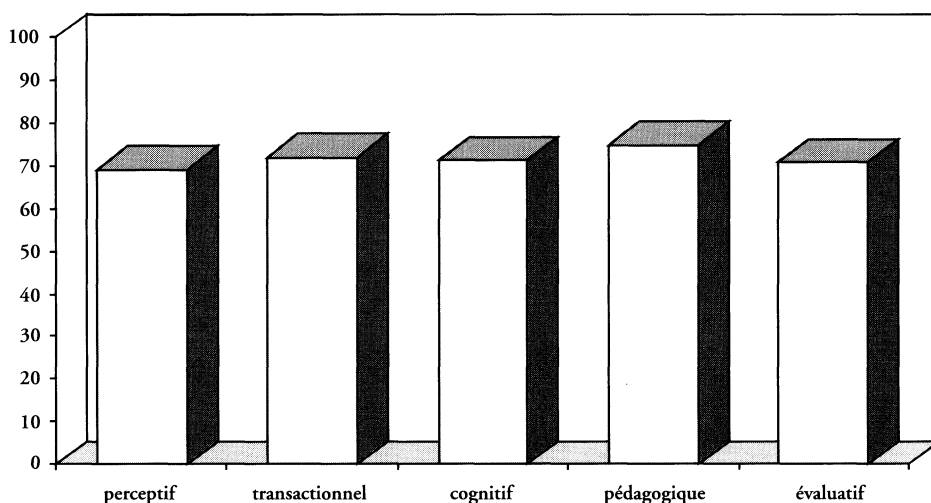


Figure 6 – Exemple de graphique pour la moyenne par niveau

La matrice favorise également la réflexion individuelle et les échanges entre évaluateurs d'un même SAMI, notamment en leur fournissant des critères et un langage communs. Le modèle, la matrice et le tableur constituent des outils non seulement d'analyse et d'évaluation, mais aussi d'apprentissage: le *design* et la conception des prototypes réalisée par des étudiants qui ont déjà eu recours à la matrice sont en effet supérieurs.

L'observation des étudiants a aussi permis de mettre en lumière les niveaux plus difficiles à analyser. Certains étudiants ont, par exemple, tendance à confondre niveau perceptif (aspect physique des messages) et niveau cognitif (sémantique des messages). L'observation des étudiants renseigne donc sur la technique même d'analyse des SAMI par le modèle multicritérié tridimensionnel.

Pistes de travail

Nous compléterons le modèle et la matrice par l'ajout obligatoire des dimensions affectives et métacognitives. Nous poursuivrons ensuite leur validation, notamment par la conception de SAMI à l'aide de ces outils et l'évaluation *a posteriori* de ces derniers avec la matrice. Nous devrions alors être davantage en mesure de déterminer les critères d'efficacité d'un SAMI.

Conclusion

Les deux principaux obstacles à l'implantation des logiciels éducatifs sont la complexité du processus, à laquelle bien des milieux de formation et d'éducation ne sont tout simplement pas préparés, et la pauvreté en matière de recherche empirique sur les critères précis d'efficacité d'un système d'apprentissage informatisé (Jolicœur et Berger, 1988).

La clé de l'évaluation efficace d'un système d'apprentissage informatisé demeure la fréquence et l'interactivité des évaluations tout au long de l'élaboration du projet. Avant tout, il faut bien sûr poser les bonnes questions! (Gerard et Roegiers, 1994). Mais pour y parvenir, il est essentiel de connaître le système grâce à une analyse détaillée de son discours et de sa structure.

Grâce à ses énoncés opérationnels, notre modèle facilite une évaluation qui minimise la subjectivité propre à toute tentative d'évaluation superficielle des SAMI. Il permet donc de rendre compte de la qualité du SAMI évalué dans toute sa complexité et sous tous ses angles intrinsèque et extrinsèque. En effet, le modèle tridimensionnel suggéré dans cet article vise l'analyse systématique des différents éléments d'un SAMI à l'aide d'un ensemble de critères. Par exemple, l'observation attentive du niveau perceptif nous autorise à porter un jugement éclairé sur l'ergonomie de l'interface du système.

En ce sens, tous les éléments de l'interface, de la lisibilité des caractères à la qualité de la mise en page sont pris en compte. Le niveau transactionnel et surtout les énoncés du niveau cognitif aident quant à eux à mieux cerner les points forts et les faiblesses cognitives du SAMI. Finalement, les neuf énoncés proposés sur le plan pédagogique de notre grille permettent de vérifier la qualité des éléments pédagogiques comme l'ancrage des connaissances, la participation active et interactive et l'approche de compagnonnage cognitif (Collins, Brown et Duguid, 1989).

Bien qu'il reste résolument qualitatif dans son approche, notre modèle permet une analyse critériée rigoureuse et reproductible des SAMI évalués. Ce modèle peut servir de base à la mise au point d'outils d'observation systématique qui permettront de dresser un portrait fidèle de l'environnement pédagogique étudié. Cette analyse complète d'un SAMI représente, à notre avis, la première étape d'une procédure rationnelle d'évaluation. Matière première qui alimente l'observation et le questionnement de l'apprenant, elle permet en définitive d'évaluer, de juger le système à l'étude.

Évaluer, c'est décomposer les objets du réel en catégories pour ensuite en établir la valeur. Une fois validée, la matrice multicritériée devrait devenir un «méta-outil» précieux pour l'analyse et l'évaluation régulative et *a posteriori* des SAMI et un aide-mémoire susceptible d'appuyer l'évaluateur dans la mise au point d'outils spécifiques comme des grilles d'évaluation ou des questionnaires. Si, par exemple, l'interface ou les interfaces du système doivent être évaluées à un moment précis de la mise au point du prototype, l'évaluateur pourrait se référer, entre autres, aux niveaux perceptif et transactionnel, et bâtir un questionnaire ou structurer des entrevues afin d'évaluer les différents critères de qualité tout en tenant compte des éléments structuraux du système.

Il faut toutefois garder à l'esprit que l'évaluation d'un SAMI est une opération fort complexe: souvent, les concepts sont souvent flous; les éléments, imprévisibles (Tripp et Bichelmeyer, 1990). La «matrice méta-outil» pourrait servir de point de référence et favoriser une observation plus pertinente et exhaustive des éléments cruciaux.

Finalement, il est important que les outils issus du modèle tridimensionnel (questionnaires, grilles d'évaluation ou entrevues) se réfèrent toujours, dans la mesure du possible, aux besoins et aux capacités du système cognitif et neurophysiologique de l'étudiant plutôt qu'aux possibilités de la technologie choisie ou disponible. À cet égard, l'analyse ou l'évaluation critériée est beaucoup plus pertinente que l'approche normative.

Abstract – This article describes the development of a data collection strategy whose aim is to classify and analyse the factors underlying interaction in mediated learning situations. The authors present a model for analysis that is centred principally on a qualitative approach which integrates an interview technique, direct observation, and protocol analysis. These methods were used to describe perceptual, transactional, cognitive, pedagogical, and evaluative factors

that could influence decisions made by those who develop interactive multimedia learning systems. Results of trials using this model by developers who work in training and in school contexts show the possible value and the detail that these analyses offer.

Resumen – Este artículo discute la preparación de una estrategia de recolección de datos que busca reagrupar y analizar los factores subyacentes al desarrollo de una interactividad significativa en situaciones de aprendizaje mediatizado. Los autores han creado un modelo adaptado de análisis basado principalmente en un enfoque cualitativo, utilizando para esto métodos que integran la técnica de entrevistas, la observación directa y el análisis de protocolos. Estos métodos han permitido distinguir los factores perceptuales, transaccionales, cognoscitivos, pedagógicos y evaluativos que pudieran dirigir la toma de decisiones de los diseñadores de sistemas de aprendizaje interactivo multimedia. Ensayos de este modelo con diseñadores en ámbitos de formación y escolar han indicado la pertinencia y el refinamiento del análisis posible.

Zusammenfassung – Dieser Artikel befasst sich mit der Aufstellung einer Datenerfassungsstrategie, die bezweckt, die Voraussetzungen für die Entwicklung einer signifikanten Interaktion in mediatisierten Lernsituationen zusammenzustellen und zu analysieren. Die Verfasser haben ein hauptsächlich auf qualitatives Vorgehen hin konzipiertes Untersuchungsmodell entworfen, wobei sie Methoden benützten, die das Interview, die direkte Beobachtung und die Protokollanalyse integrieren. Mit diesen Methoden konnten die Perzeptions-, Transaktions- und Kognitionsfaktoren, sowie die pädagogischen und die Beurteilungsfaktoren erfasst werden, die für die Autoren neuer interaktiver Multimedia-Lernsysteme wichtig sind. Das Modell wurde mit Hilfe von von der Bildung oder Schule her kommenden Autoren erprobt; die Relevanz der Untersuchungen wurde dabei bestätigt.

RÉFÉRENCES

- Booth, P. (1989). *An introduction to human-computer interaction*. East Sussex, Grande-Bretagne: Lawrence Erlbaum Associates.
- Boulet, A., Savoie-Zajc, L. et Chevrier, J. (1995). L'abc de la réussite universitaire. Les stratégies d'apprentissage. *Réseau*, novembre, 14-19.
- Boutin, F. et Chinien, C. A. (1992). Synthesis of research on student selection criteria in formative evaluation. *Educational Technology*, 32, 28-31.
- Collins, A., Brown, J. S. et Duguid, P. (1989). Cognitive apprenticeship: Teaching the craft of reading, and mathematics. In L. Resnick (dir.), *Knowing, learning and instruction: Essays in honor of Robert Glaser* (p. 453-494). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Depover, C. (1994). Problématique et spécificité de l'évaluation des dispositifs de formation multimédias. *Éducatotechniques*, 1, 29-46.
- Flagg, B. N. (1990). *Formative evaluation for educational technologies*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Gerard, F. M. et Roegiers, X. (1994). Évaluer un projet d'informatique pédagogique: une question de questions. *Recherche en éducation-théorie et pratique*, 16, 35-43.
- Giardina, M. (1989). *Évolution du concept d'interactivité à dimension cognitive*. Thèse de doctorat, Faculté des sciences de l'éducation, Université Laval, Sainte-Foy.
- Giardina, M. (1992). L'interactivité dans un environnement d'apprentissage multimédiatisé. *Revue des sciences de l'éducation*, XVIII(1), 43-66.

- Giardina, M. et Meunier, C. (1993). *La problématique de l'évaluation des systèmes multimédia interactifs*. Communication présentée au Conseil interinstitutionnel pour le progrès de la technologie éducative, 18 mai 1993, Université du Québec à Montréal, Québec.
- Horn, R. (1989). *Mapping Hypertext*. Lexington, Grande-Bretagne: The Lexington Institute.
- Jolicœur, K. et Berger, E. (1988). Implementing educational software and evaluating its academic effectiveness: Part II. *Educational Technology*, 28, 13-19.
- Mayhew, D. J. (1992) *Principles and guidelines in software user interface design*. Englewood Cliffs, NJ: Prentice-Hall.
- Norman, D. A. (1988) *The psychology of everyday things*. New York, NY: Basic Books.
- Owston, R. D. et Wideman, H. H. (1987). The value of supplemental panel software reviews with field observations. *Canadian Journal of Educational Communication*, 16, 295-308.
- Reeves, T. C. (1990). Redirecting evaluation of interactive video: The case for complexity. *Studies in Educational Evaluation*, 16, 115-131.
- Reeves, T. C. (1992). Evaluating interactive multimedia. *Educational Technology*, 32, 47-53.
- Reeves, T. C. (1993). Evaluating interactive multimedia. In D. M. Gayeski (dir.), *Multimedia for learning: Development, application, evaluation* (p. 97-112). Englewood Cliffs, NJ: Educational Technology Publications.
- Reiser, R. A. et Dick, W. (1990). Evaluating instructional software. *Educational Technology Research and Development*, 38, 43-50.
- Reiser, R. et Kegelmann, H. W. (1994). Evaluating instructional software: A review and critique of current methods. *Educational Technology Research and Development*, 42, 63-69.
- Robert, J.-M. (1988). *Ergonomie des interfaces humains-ordinateurs*. Guide du participant. Cours préparé pour Info J.E.D. inc.
- Romizovski, A. J. (1986). *Developing auto-instructional materials: From programmed texts to CAL and interactive video*. New York, NY: Nichols Publishing Company.
- Salomon, G. (1991). Transcending the qualitative-quantitative debate: The analytic and systemic approaches to educational research. *Educational Researcher*, 20, 10-18.
- Scallon, G. (1988). *L'évaluation formative des apprentissages*. Québec: Les Presses de l'Université Laval.
- Shneiderman, B. (1992). *Designing the user interface. Strategies for effective human-computer interaction* (2^e éd.). Reading, MA: Addison-Wesley Publishing.
- Tardif, J. (1992). *Pour un enseignement stratégique. L'apport de la psychologie cognitive*. Montréal: Les Éditions Logiques.
- Tolhurst, D. (1992). A checklist for evaluating content-based Hypertext computer software. *Educational Technology*, 32, 17-21.
- Tripp, S. D. et Bichelmeyer, B. (1990). Rapid prototyping: An alternative instructional design strategy. *Educational Technology Research and Development*, 38, 31-44.
- Vanderdonckt, J. (1995). *Guide ergonomique des interfaces homme-machine*. Namur: Presses universitaires de Namur.
- West, K. W., Farmer, A. J. et Wolff, P. M. (1991). *Instructional design: Implications from cognitive science*. Needham Heights, MA: Allyn and Bacon.
- Zahner, J. E., Reiser, R. A., Dick, W. et Gill, B. (1992). Evaluating instructional software: A simplified model. *Educational Technology Research and Development*, 40, 55-62.