
Vers une scénarisation de l'évaluation en EIAH

L'évaluation comme activité scénarisable dans un dispositif de scénarisation pédagogique

Guillaume Durand

*Laboratoire SysCom(Université de Savoie)
Bâtiment Mont-Blanc
F-73370 Le Bourget du lac
Guillaume.Durand@etu.univ-savoie.fr*

RÉSUMÉ. Que ce soit dans la manière de décrire, d'indexer les contenus pédagogiques mais aussi de scénariser les activités pédagogiques, les Environnements Informatiques pour l'Apprentissage Humain (EIAH) connaissent actuellement un important effort de standardisation. Dans cette évolution, l'évaluation est le parent pauvre, car très souvent absente ou obsolète dans les formalismes proposés. Or l'évaluation, qui est une part importante de l'activité pédagogique, mérite de disposer elle aussi des méthodes, des techniques et des outils qui font d'une manière générale évoluer le contexte des EIAH. Après avoir discuté de l'état de l'art de l'évaluation dans les EIAH, nous présentons dans cet article les choix opérés pour le traitement d'un exemple typique d'évaluation formative et les réflexions que ce traitement nous inspire pour nos futurs travaux.

ABSTRACT. That it is in the manner of describing, of indexing, managing learning contents, Technology Enhancing Learning (TEL) are making a standardisation effort. But, in this evolution assessment is rarely considered. However, assessment which is a significant share of the learning activity should also have it own methods, techniques, tools which nowadays make evolve TEL. After having discussed the state of the art of assessment in TEL, we present in this article the choices which we have operated for the treatment of a typical example of formative assessment and the reflections that this treatment inspires to us for our future work.

MOTS-CLÉS : évaluation, scénario pédagogique, scénario d'évaluation.

KEYWORDS: assessment, learning script, assessment script.

1. Introduction

Autrefois développés de manière empirique sur des modèles implicites, sans ambition d'interopérabilité avec des systèmes extérieurs, les EIAH connaissent ces dernières années d'importants efforts de formalisation et de normalisation. Que ce soit dans la manière de décrire, d'indexer les contenus pédagogiques mais aussi de scénariser les activités pédagogiques, des formalismes s'imposent. Dans cette évolution, l'évaluation est le parent pauvre. Il n'existe pas de formalisme particulier pour spécifier l'évaluation des apprenants. L'évaluation, dimension cruciale de l'activité pédagogique, dans sa fonction certificatrice, formative¹, normative² ou sommative³, se retrouve largement traitée en EIAH. Cependant, les approches et solutions proposées sont bien souvent ad hoc et spécialisées. Il n'y a pas d'approche de formalisation de l'évaluation comme il peut en exister par exemple avec les propositions de formalisation des activités pédagogiques. Après avoir discuté de l'état de l'art de l'évaluation dans les EIAH et introduit brièvement les critiques relatives à l'absence de la prise en compte de l'évaluation dans les normes et standards, nous présentons dans cet article les choix opérés pour le traitement d'un exemple typique d'évaluation formative et les réflexions que ce traitement nous inspire pour nos futurs travaux.

2. L'évaluation dans les EIAH

2.1. Rôle et portée de l'évaluation en EIAH

Que ce soit en classe ou au sein des EIAH, l'évaluation est un dispositif clé. En effet, cette évaluation dans le contexte d'un EIAH doit (Juwah, 2003) :

- être motivante pour l'apprenant,
- encourager une activité d'apprentissage soutenue,
- contribuer à la progression de l'apprenant,
- être faible en coût humain et facilement maintenable.

Cette évaluation assistée par ordinateur améliore la détection des progrès et lacunes des apprenants suivant la rapidité et la qualité des interactions qu'ils ont avec et dans l'EIAH (Bull, 1999). De même que l'évaluation joue différents rôles en EIAH, elle porte sur des objets différents. Lorsqu'elle porte sur les productions des apprenants, on parle alors d'évaluation de production. En revanche, lorsque l'évaluation porte sur la démarche mise en oeuvre pour réaliser des productions, on parle alors d'évaluation de la démarche⁴.

1. Evaluation formative : évaluation dont l'ambition est de contribuer à la formation.

2. Evaluation normative : évaluation qui situe les individus les uns par rapport aux autres.

3. Evaluation sommative : évaluation par laquelle on fait un inventaire des compétences acquises.

4. Synthèse sur l'évaluation : <http://www.grenoble.iufm.fr/departement/shs/campeval>

2.2. Quelles sont ces évaluations ?

L'évaluation est tributaire des activités proposées par l'EIAH. En effet, l'évaluation menée dans le cadre d'une activité collective⁵ n'est pas toujours valide dans une activité individuelle⁶. Il est possible de distinguer l'évaluation dans le contexte des activités individuelles, de celle menée dans le contexte des activités collectives.

2.2.1. L'évaluation dans des activités individuelles

Evaluation de production : le bilan de compétences.

Le bilan de compétence consiste à prendre une "photographie" du niveau de compétence d'un apprenant en évaluant ses productions. Ce type d'évaluation sommative est mené dans des environnements d'apprentissage basés sur des banques de questionnaires à choix multiples (QCM) organisés en parcours et sanctionnant l'apprenant par une note finale (Green *et al.*, 1984). Le test du TOEFL⁷ est certainement le plus connu. Cette évaluation de compétence a été poussée à son paroxysme avec des évaluations diagnostiques, qui ne se contentent pas de corriger les productions des apprenants mais qui proposent des diagnostics. Chaque réponse de l'apprenant est analysée pour essayer de l'associer à une erreur envisagée par les concepteurs de l'exercice et à son origine potentielle. Le logiciel Pépite (Delozanne *et al.*, 2004) et ses dérivés illustrent cette évaluation diagnostique.

Evaluation de production : l'auto-évaluation.

Dans certains cas, il peut être utile de proposer à l'apprenant de s'auto-évaluer afin de favoriser l'auto-régulation de ses apprentissages en s'insérant dans le cadre d'une évaluation formative. Le logiciel GenEval (David, 2003) propose une liste de questions, dans laquelle l'apprenant peut naviguer. Pour chaque question, l'apprenant peut bénéficier ou non d'indications pour répondre. Une fois la réponse donnée et corrigée, l'apprenant évalue sa maîtrise des compétences mises en jeu dans les questions en se donnant une note.

Evaluation de la démarche : l'assistance à l'évaluation.

S'il est possible d'évaluer relativement finement une production, il n'en est pas de même pour une démarche. Aussi, que ce soit dans des activités individuelles ou collectives, des solutions d'assistance à l'évaluation sont proposées. Elles consistent bien souvent à fournir à l'enseignant des vues sur les activités par le biais d'indicateurs, de mesures (Merceron *et al.*, 2004). Ces mesures permettent à l'enseignant d'évaluer le déroulement de l'activité.

2.2.2. L'évaluation dans des activités collaboratives

Evaluation de production : l'auto-évaluation collaborative.

Une des premières pratiques d'évaluation des activités collaboratives en EIAH est l'auto-évaluation par pair. Cette évaluation est menée par un groupe d'apprenants

5. Activité réalisée en groupe avec des apprenants qui interagissent les uns avec les autres.

6. Activité réalisée par un apprenant seul.

7. TOEFL : <http://www.ets.org>

et porte sur une production réalisée par un autre groupe. La difficulté consiste entre autres, pour les membres du groupe qui évaluent à se mettre d'accord sur les réponses attendues (Juwah, 2003).

Evaluation de la démarche : l'évaluation de la participation.

Bien que commune à toute activité collective, l'évaluation de la participation se retrouve quasi essentiellement dans le cadre d'activités usant d'un forum. L'évaluation porte alors sur la qualité des discussions basées sur des techniques de "text mining" et s'intéressant au rapport entre mots clés (ceux attendus par l'enseignant) et nombre d'échanges. Tel est le cas du logiciel DIAS⁸ dont le but est d'assister un enseignant pour gérer l'activité d'apprenants dans un forum (Bratitsis *et al.*, 2005) en évaluant la qualité de la participation des intervenants.

L'évaluation de la démarche : l'assistance à l'évaluation.

Alors que l'assistance à l'évaluation trouvait du sens dans une activité individuelle, elle devient nécessaire dans le cadre de l'activité collective qui démultiplie la complexité pour l'enseignant d'en évaluer le déroulement. Dans ce contexte, plus que des outils de mesure (Mazza *et al.*, 2005), il existe des systèmes qui proposent de guider l'activité. Certains de ces EIAH procèdent par notifications qui prennent la forme de conseils destinés aux apprenants (George, 2001). Par exemple, si le groupe d'apprenant est un retard sur son planning, le système lui conseillera de se dépêcher.

3. Les efforts de standardisation

3.1. La nécessité d'une standardisation

Après avoir introduit les différentes pratiques d'évaluation en EIAH, il est possible de constater une certaine diversité des pratiques rencontrées. Les EIAH qui portent ces usages sont aujourd'hui nombreux. Cependant, chaque EIAH utilise des formalismes et des méthodes qui lui sont propres pour décrire, référencer, organiser, et jouer les contenus et les activités qu'il propose. Dans un souci de lisibilité, d'interopérabilité et de réutilisabilité en EIAH, il est nécessaire de spécifier des formalismes en vue de les faire adopter comme standards (souvent de fait). C'est la démarche qui est entreprise en EIAH depuis quelques années.

3.2. Les exemples importants

En EIAH, la majeure partie des contenus interactifs s'apparente à des questionnaires. IMS-QTI⁹ propose un format XML pour construire ces contenus. Ce format permet d'échanger des contenus entre EIAH compatibles et de diminuer les coûts de production grâce aux outils auteurs et moteurs d'exécution qui apparaissent (Respon-

8. DIAS : Data Recording and Usage Interaction Analysis in Asynchronous Discussions

9. IMS Question and Test Interoperability : <http://www.imsglobal.org/question/>

us, Canvas, etc¹⁰). IMS-QTI jouit d'un grand succès, et se positionne aujourd'hui comme un standard de fait.

Tout comme les contenus eux-mêmes, leur organisation et interaction avec l'(les) apprenant(s) au sein d'une activité pédagogique nécessite d'être décrite. Des formalismes (IMS-LD¹¹, LDL¹²(Ferraris *et al.*, 2005)), ont tout d'abord été proposés afin d'exprimer des activités pédagogiques avant que ne soient imaginées et réalisées des infrastructures pour les exécuter dans des systèmes de gestion de l'apprentissage (SGA) (Reload¹³, LDI¹⁴(Ferraris *et al.*, 2005)). Grâce à ces formalismes, il est possible de créer dans un langage structuré des scénarios qui décrivent un modèle d'activité pédagogique. Ces scénarios sont ensuite instanciés avec les objets d'un SGA cible (participants, outils) et joués par une infrastructure d'exécution de scénario pédagogique (LDI, Reload, etc..).

3.3. *L'évaluation est-elle intégrée à cet effort de standardisation ?*

Que ce soit dans l'expression des contenus ou la scénario des activités pédagogiques, l'évaluation devrait être largement présente. Or, IMS-QTI qui initialement devait être un format d'évaluation a manqué son objectif. IMS-QTI décrit finement les contenus mais ne propose que des préconisations¹⁵ pour leur évaluation. En outre, le paradigme d'évaluation proposé est pauvre (évaluation individuelle de compétence) au regard des situations rencontrées en EIAH. Certes, il n'est pas du ressort d'un formalisme d'expression de contenu pédagogique de décrire son usage. C'est plutôt celui des formalismes de scénarisation. Or, dans les dispositifs de scénarisation, l'évaluation n'est que très peu abordée. Elle n'est pas considérée comme partie intégrante de l'activité pédagogique. Elle sert au mieux à ordonnancer le parcours pédagogique, grâce aux résultats obtenus dans les QCM (IMS-QTI).

En conclusion, l'évaluation pratiquée en EIAH est quasi absente de l'effort de formalisation et de standardisation actuel.

4. Scénariser l'évaluation

4.1. *L'évaluation est-elle une activité ?*

Qu'elle porte sur des productions ou des démarches, qu'elle soit normative, formative ou sommative, l'évaluation est une activité composée d'interactions entre un (des) évaluateur(s) et un (des) évalué(s), portant sur des productions ou sur la dé-

10. Liste des outils IMS-QTI : <http://support.imsglobal.org/question/qtiSupportIndex.html>

11. IMS-Learning Design : <http://www.imsglobal.org/learningdesign/>

12. LDL : Learning Design Language

13. Reload : <http://reload.ac.uk>

14. LDI : Learning Design Infrastructure

15. http://www.imsglobal.org/question/quiv2p1pd/imsqti_infov2p1pd.html

marche choisie pour la réalisation de ces productions. L'évaluation peut être considérée comme une activité, et devrait pouvoir à ce titre être scénarisée tout comme une activité pédagogique. Il doit être possible d'exprimer l'évaluation dans un formalisme de scénarisation. Ce qui permettrait de formaliser l'évaluation en utilisant les formalismes de scénarisation existants (tels que LDL) et ainsi de pallier son absence dans le mouvement actuel de standardisation.

4.2. *Peut-on parler de scénario d'évaluation ?*

De cette hypothèse découle deux propositions. Ou bien la description de l'évaluation fait partie du scénario pédagogique, ou bien elle nécessite un scénario qui lui est propre : c'est à dire un **scénario d'évaluation**.

L'évaluation est une activité dynamique qui doit pouvoir être adaptée au cours de l'activité pédagogique, tout comme l'enseignant adapte son barème au fur et à mesure de ses corrections. Dans le cas où, l'évaluation est définie avec l'activité pédagogique dans un même scénario, adapter l'évaluation ne peut se faire sans influencer sur l'activité pédagogique puisque ces activités sont alors confondues. Adapter l'évaluation, nécessite dans ce cas l'arrêt de l'activité pédagogique, ce qui n'est pas chose souhaitable. En outre, il doit être possible de réutiliser certaines évaluations d'une activité pédagogique à une autre (Williamson *et al.*, 2003). Or, il est plus simple de s'échanger des scénarios d'évaluation que d'isoler dans la description de l'activité pédagogique les éléments relatifs à l'évaluation et de les réimplanter dans un scénario pédagogique cible. C'est pourquoi nous avons orienté nos recherches vers une scénarisation de l'évaluation indépendante ou presque des activités évaluées. Nous proposons l'usage de scénarios d'évaluation dont voici un exemple.

4.3. *Exemple de scénario d'évaluation*

Le scénario d'évaluation choisi, décrit une évaluation individuelle diagnostique et formative. L'exemple est relativement pauvre, il porte sur des contenus simples à traiter tels que des questionnaires, mais illustre la faisabilité de la proposition. Cette activité d'évaluation se déroule simultanément à une activité de réponse (l'activité pédagogique) dans laquelle un apprenant répond à une série de questions ayant toutes un même objectif pédagogique. Chaque question possède une et une seule solution, ainsi qu'une liste d'erreurs envisagées avec l'origine possible de chaque erreur. Le scénario d'évaluation décrit la correction de chaque réponse de l'apprenant (grâce à la réponse attendue) mais aussi en cas d'erreur, la détermination de l'origine possible de l'erreur, et la proposition d'une activité de remédiation (cf Figure 1). Cette activité de remédiation est aussi scénarisée, elle propose à l'apprenant un cours expliquant la notion non comprise. Une fois l'activité de remédiation terminée, l'apprenant continue l'activité de réponse. Si l'apprenant ne commet pas d'erreur, il passe à la question suivante prévue par le scénario de réponse.

Afin de démontrer la faisabilité de notre proposition, nous implémentons actuellement

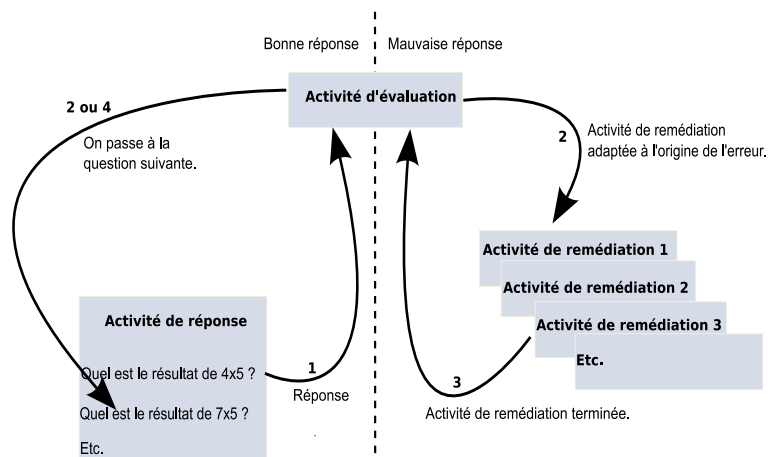


Figure 1. Illustration du déroulement de l'activité d'évaluation exemple.

différents scénarios d'évaluation (dont celui-ci) représentatifs des pratiques en EIAH. Nous utilisons le formalisme de scénarisation LDL et l'infrastructure LDI.

5. Les problèmes rencontrés

Au fur et à mesure de cette implémentation apparaissent un certain nombre de difficultés auxquelles nous devons répondre. En effet, un scénario d'évaluation décrit l'évaluation de production(s), de démarche(s) décrite(s) par un scénario pédagogique. L'activité décrite par le scénario pédagogique produit des observables qui sont des vues du déroulement de l'activité pédagogique. Ce sont ces vues qui constituent les observables évalués par une activité d'évaluation. Se pose donc le problème de la communication des observables entre deux activités scénarisées : comment l'activité d'évaluation obtient les observables de l'activité pédagogique ? Mais encore, comment connaître et définir les observables disponibles dans l'activité pédagogique ? En outre, l'évaluation ne se limite pas au recueil de mesures issues d'observables (cas de l'assistance à l'évaluation). L'évaluation demande aussi des critères d'évaluation. Dans l'exemple choisi, ce sont la solution de la question mais aussi les erreurs possibles et leurs origines. Comment et où définir ces critères (dans les métadonnées des contenus, dans les contenus eux-mêmes, au niveau du scénario pédagogique) ?

6. Conclusion

Avec l'effort de normalisation des EIAH et le fort degré de liberté que proposent les dispositifs de scénarisation quant aux activités pédagogiques qu'ils permettent,

l'évaluation ne peut rester aussi spécialisée et figée qu'elle ne l'est aujourd'hui. La solution choisie est de considérer l'évaluation comme une activité scénarisable pouvant être décrite par un scénario d'évaluation dans un formalisme de scénarisation. Cette proposition n'est pas sans poser quelques problèmes. Pour démontrer la faisabilité de cette proposition, nous voulons implémenter différents scénarios d'évaluation significatifs des pratiques en EIAH. Cette implémentation est actuellement consolidée par la formalisation des objets qui lui sont nécessaires tels que les observables. Enfin, ce travail va donner lieu à une expérimentation dans le cadre d'un projet européen afin d'en vérifier la pertinence.

Remerciements

Je souhaite particulièrement remercier mon directeur de thèse Christian Martel, les sociétés Pentila, Odile Jacob Education et le laboratoire SysCom.

7. Bibliographie

- Bratitsis T., Dimitracopoulou A., « Data Recording and usage Interaction Analysis in Asynchronous Discussions : The D.I.A.S. System », *AIED Workshops (AIED'05)*, juillet, 2005.
- Bull J., « Computer-Assisted Assessment : Impact on Higher Education Institutions », *Journal of Educational Technology and Society*, vol. 2(3), p. 123-126, 1999.
- David J.-P., « Modélisation et production d'objets pédagogiques », *Sciences et Techniques Educatives*, avril, 2003.
- Delozanne E., Grugeon B., « Pépites et lingots : des logiciels pour faciliter la régulation par les enseignants des apprentissages en algèbre », *Cahiers Éducation et Devenir*, vol. Hors série, Les TIC à l'école : miracle ou mirage ?, p. 82-92, septembre, 2004.
- Ferraris C., Lejeune A., Vignollet L., David J., « Modélisation de scénarios d'apprentissage collaboratif pour la classe », *actes de la conférence EIAH (EIAH'05)*, juin, 2005.
- George S., SPLACH : un environnement informatique support d'une pédagogie par projet, PhD thesis, Université du Maine, Apprentissage collectif à distance, 2001.
- Green B., Bock R., Humphreys L., Linn R., Reckase M., « Technical Guidelines for Assessing Computerized Adaptive Tests », *Journal of Educational Measurement*, vol. 21 session 4, p. 347-360, 1984.
- Juwah C., « Using Peer Assessment to Develop Skills and Capabilities », *United States Distance Learning Association*, vol. janvier 2003, p. 39-50, 2003.
- Mazza R., Milani C., « Exploring Usage Analysis in Learning Systems : Gaining Insights From Visualisations », *AIED Workshops (AIED'05)*, juillet, 2005.
- Merceron A., Acef K., « Train, store, analyse for more adaptive teaching », *Technologies de l'Information et de la Connaissance dans l'Enseignement Supérieur et l'Industrie*, vol. octobre 2004, p. 52-58, 2004.
- Williamson D., Bauer M., Mislevy R., Behrens J., « An ECD Approach to Designing for Reusability in Innovative Assessment », *American Educational Research Association (AERA'03)*, avril, 2003.

ANNEXE POUR LE SERVICE FABRICATION
A FOURNIR PAR LES AUTEURS AVEC UN EXEMPLAIRE PAPIER
DE LEUR ARTICLE ET LE COPYRIGHT SIGNE PAR COURRIER
LE FICHER PDF CORRESPONDANT SERA ENVOYE PAR E-MAIL

1. ARTICLE POUR LES ACTES :
RJC-EIAH'2006
2. AUTEURS :
Guillaume Durand
3. TITRE DE L'ARTICLE :
Vers une scénarisation de l'évaluation en EIAH
4. TITRE ABRÉGÉ POUR LE HAUT DE PAGE MOINS DE 40 SIGNES :
Scénariser l'évaluation
5. DATE DE CETTE VERSION :
13 avril 2006
6. COORDONNÉES DES AUTEURS :
 - adresse postale :
Laboratoire SysCom(Université de Savoie)
Bâtiment Mont-Blanc
F-73370 Le Bourget du lac
Guillaume.Durand@etu.univ-savoie.fr
 - téléphone : 06 60 86 64 52
 - télécopie : Guillaume.Durand@etu.univ-savoie.fr
 - e-mail :
7. LOGICIEL UTILISÉ POUR LA PRÉPARATION DE CET ARTICLE :
L^AT_EX, avec le fichier de style `article-hermes.cls`,
version 1.2 du 03/03/2005.
8. FORMULAIRE DE COPYRIGHT :
Retourner le formulaire de copyright signé par les auteurs, téléchargé sur :
<http://www.revuesonline.com>

SERVICE ÉDITORIAL – HERMES-LAVOISIER
14 rue de Provigny, F-94236 Cachan cedex
Tél : 01-47-40-67-67
E-mail : revues@lavoisier.fr
Serveur web : <http://www.revuesonline.com>