

Adaptive confidence-oriented self-assessment in an IMS LD Unit of Learning

This section presents an adaptive unit of learning, called "The Dangerous Knowledge Tour", which cumulates two novelties. On the one hand, the adaptive processes were modelled with the IMS-LD authoring tool Recourse at level B. On the other hand, it provides the first example of an adaptive sequencing based on a combination of cognitive and metacognitive formative tests. From a technical viewpoint (IMS-LD for adaptation) and from an instructional viewpoint (combination of learning and metalearning factors in assessment), the Dangerous Knowledge Tour (Verpoorten & Glahn, 2010) provides an up-and-running instantiation of a next generation assessment. However, the instructional design efforts highlighted shortages and drawbacks regarding IMS LD and QTI specifications.

Problem definition

Learning does not move someone from total ignorance to perfect knowledge. Often people will already have some knowledge or representation about what is taught, even if these representation or knowledge might be erroneous. So evaluation should not be limited to either knowledge or correct answer and ignorance or incorrect answer. "Partial information exists. To detect it is necessary and feasible" (De Finetti, 1965).

One way to detect partial knowledge and therefore to go beyond assessment practice based only on identifying correct and incorrect answers is to ask the student to make explicit the confidence he has in the accurateness of his answers. So doing, the cognitive evaluation (the answer) is coupled with a metacognitive evaluation (the confidence degree).

Practically, a confidence marking procedure implies that for each question, there should be an easy possibility to get two pieces of information from the student: the answer itself, a confidence degree expressed in percentages of chances of the answer being correct.

Different combinations of the cognitive and metacognitive parameters can be identified and represented in a "spectral distribution of knowledge" (Jans & Leclercq, 1999). On the left hand side are the incorrect answers. They are distributed by the learner chosen confidence, ranking from left to right from 100% down to 0%. On the right hand side are the correct answer, also distributed by confidence but ranking from left to right from 0% to 100%. The distribution allows to distinguish between four types of knowledge situations in which a person can be in relation to a piece of content : dangerous knowledge, unawareness, mid knowledge and usable knowledge. misinformed, uninformed, informed". In the middle, the grey area represents the not answered questions.

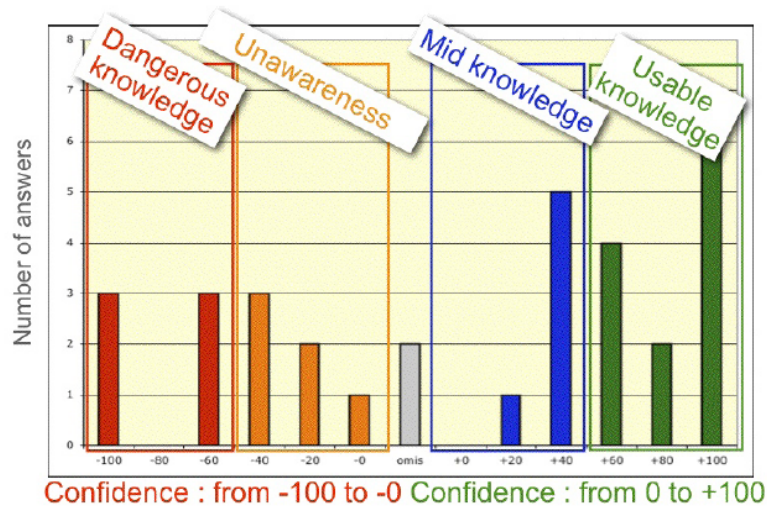


Fig. 1. The spectral distribution of knowledge

In the context of an assessment enriched with confidence marking, knowledge goes from the worst knowledge state (error with the highest confidence degree) to the best state (correct answer with the highest confidence degree). According to this typology, the dangerous knowledge is defined as incorrect knowledge associated with a high degree of confidence. The Dangerous Knowledge Tour postulates that this misinformed (dangerous) knowledge should be addressed in priority since¹:

- "the most useful piece of learning for the uses of life is to unlearn what is untrue" (Antisthenes).
- "To kill an error is as good a service as, and sometimes even better than, the establishing of a new truth or fact" (Darwin).
- "An education isn't how much you have committed to memory, or even how much you know. It's being able to differentiate between what you do know and what you don't". (Anatole France)

When students evaluate their confidence in their answer(s), they train their metalearning abilities. Without using confidence degrees, there are possibly crucial mental changes in the learner of which the system cannot have a trace. Confidence marking gives access to information about the learner that would remain totally invisible with the use of responses coded only correct or incorrect. In this respect, the implementation of this confidence marking procedure allowing for metacognition deepening in learning or in evaluation is an obvious real "plus" for instruction. Confidence marking allows a more subtle degree of differentiation in the learning progression (e.g. : frequently, responses do not change but the related confidence degree do). The analysis and the treatment of this data which remain on observable behaviour can make interesting learner's features appear (over- and under-estimation, detection of too/not enough complex material, level of internalization of electronic material, control degree of the workflow, etc). Whereas Confidence Degrees are not relevant for all situations, they are useful in a large range of applications without requesting any increase in the students' answering time or in the computing resource. They make it possible to measure much more subtle and diagnostic levels of assessment (research is available about the "stability" and inter tests variations of realism in the same person). They serve cognitive processes while offering an opportunity to train metacognitive skills.

¹ All quotations are given as cited by Miller, M. (2009). Overview of research methods. Western International University.

Current practice

On a general level, the practice of incorporating meta-learning activities in regular of online lessons is not largely spread. With specific regard to confidence degrees in distance education, neither eLearning platforms like Blackboard or Moodle or specialized assessment suites like QTI, Hot Potatoes or QuestionMark- Perception provide seamless way to incorporate evaluation of confidence. Furthermore, even if confidence degrees can be obtained by twisting existing test/quizzes facilities, this data should still be communicated to the IMS LD authoring tool so that it can be used in the adaptation process. Our attempts to have ReCourse and QTI working together in the Dangerous Knowledge Tour revealed a lack of integration between the tools. This is the reason why we decided not to implement the formative tests via QTI but straight with the facilities provided by Recourse, a workaround which turned out to be itself riddled with problems.

Proposed solution

The design of the "Dangerous Knowledge Tour" highlighted that a good implementation of the tests required to have tight interrelations between the LD and the QTI objects, what was not the case. Because the IMS LD- QTI integration is not fully supported by the IMS LD runtime and authoring environments, we opted for the global-element solution. These elements are directives for a runtime environment to inject special user interface elements that allow a participant in a UoL to change the underlying properties. It should be ascertained whether QTI 2 could solve the problem of interfacing so that, in the end, the QTI and the IMS LD global-element solution are equally tight integrated. But in our case, we decided to work on the conceptual level to reduce system dependencies (and sources for implementation related problems).

Validation

The key-idea of confidence degrees is that a genuinely valid assessment must pay attention both to students' answer correctness and the students' confidence in the correctness of the answer they gave.

Prototype

The "Dangerous Knowledge Tour" therefore addresses personal meta-cognitive support in addition to traditional assessment types. Within the domain of Web usability, students are supported in identifying when they build "dangerous knowledge". Dangerous knowledge are wrongly learned concepts connected to a learner's high confidence that these concepts are correct. The UoL primarily focuses on content personalisation, service integration, and assessment.

Assessment practice

Central to this Unit of Learning (UoL) is a series of frequent, quick and instantaneous formative tests. By "instantaneous", we mean tests displayed straight after the coverage of a specific piece of content. The test gives the learner immediate feedback on the quality of his study.

Critical to the assessment practice is double nature of the tests: cognitive and metacognitive. All tests are structured the same way: one MCQ or open question coupled with a self-reported estimation of the confidence in the accurateness of the answer.

Activity	Individual learner	teacher
Introduction		
Task 1	✓	
Formative test on principle 1	✓	
Task 2	✓	
Formative test on principle 2	✓	
test reading 2 open question		✓
Task 3	✓	
Formative test on principle 3	✓	
Task 4	✓	
Formative test on principle 4	✓	
Task 5	✓	
Formative test on principle 5	✓	
Conclusion	✓	

Fig. 2. The learning design organizes a systematic criss-crossing of study tasks and intermediate formative tests

Adaptation rules

Confidence marking overlays the learning dimension (answers to a test) with a metalearning one (confidence in the quality of the answers). Confidence marking allows having metacognitive elements taken into account in the feedback of an evaluation and, in the case of the Dangerous Knowledge Tour into the adaptive process. In this UoL, the adaptive sequencing of learning activities is defined by the combination of:

- the cognitive part of the assessment: the answer is right or wrong;
- the meta-cognitive part of the assessment: the student's level of confidence in his answer is 0%, 20%, 40%, 60%, 80% or 100%.

"Dangerous knowledge" is defined here as incorrect knowledge (wrong answer) associated with a high degree of confidence (>40% confidence). In this specific case, students are quite confident in the answer they gave, which is yet an erroneous information. When this double mistake is spotted, the student is brought back to the content whose appropriation is insufficient. After a new study sequence, he is invited to pass the formative test again. Only when he give the right answer with a higher confidence degree, is he allowed to move ahead in the UoL. When students give a wrong answer with a low confidence, they are also brought back to the previous content but in between they are requested to give an explanation for their low level of confidence.

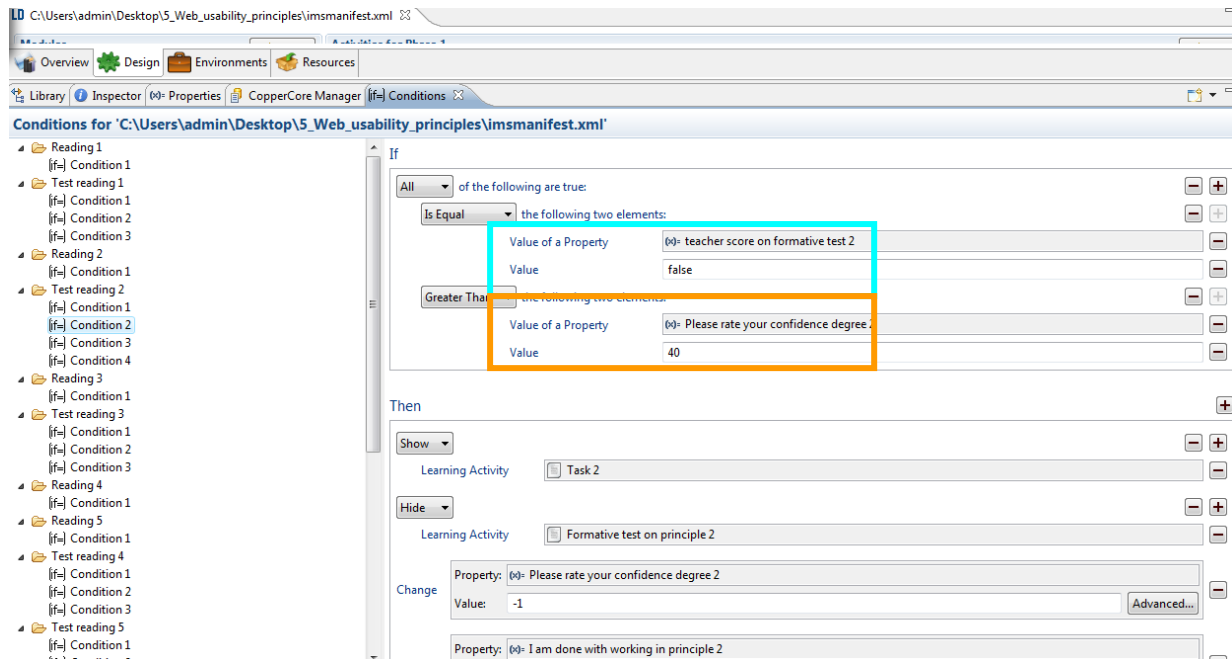


Fig. 3. A set of properties and conditions bearing on cognitive (blue rectangle) and metacognitive assessment (orange rectangle) determine the adaptive sequencing of activities

Validation results

The validation does not bear on students' performance or appreciation of the UoL. The UoL is here discussed against the ability of the IMS-LD Specification, IMS-LD editors/players and QTI specification to support course designers and learning supporters in the implementation of Adaptive confidence-oriented self-assessment seen as an example of a next generation evaluation.

In search of better interaction between IMS LD UoLs and Learning Resources

IMS LD Level B allows creating interactive UoLs by using the so-called “global elements” inside of learning resources. These elements are directives for a runtime environment to inject special user interface elements that allow a participant in a UoL to change the underlying properties. Alternatively, that the lexical naming of properties in resources and in a learning design should be used to exchange the underlying data. This has been discussed specifically with regard to the IMS LD and IMS QTIv2 integration (IMS Global Learning Consortium, 2005). Both approaches have the obvious drawback that in either way an interactive resource is tightly coupled to one learning design.

The first approach makes assumptions about the user interface presentation of the data stored in a property. These assumptions are specific to one IMS LD runtime environment and provide neither to the resource designer nor to an instructional designer to define guidelines for the presentation of the provided information. For appropriate presentation of the information, the resource designer has to reflect the specifics of the underlying IMS LD runtime environment. This limits the interoperability of a learning resource across different IMS LD runtime environments. Furthermore, the global properties limit the interoperability of learning resources because they rely on the internal identifiers of the properties. However, most of the recent IMS LD authoring environments hide this information from the instructional designer. This means that one interactive learning resource is difficult to be used across different UoLs.

The second approach relies on the lexical identity of property identifiers or variable names. Besides that this approach has the same drawback regarding the property identifiers as the global-element handling of the first approach, it also limits the interoperability of the resources in a similar way as the global-elements.

Both approaches have the drawback that a practitioner focus primarily on content that remains mostly static throughout a run of a UoL. Furthermore, in both cases a fundamental knowledge about the related interfaces is required during the modelling process. The concept of showing and hiding of CSS class names even assumes that all content and services in a UoL share a common style definition that is fully understood by the educational designer.

In the UoLs “Dangerous Knowledge Tour” the tight coupling between learning resources and the educational design became an immanent problem. This tight interrelation made it necessary to create several resources that all looked similar, but had references to different properties. At the same time the author had limited or no control over the layout and the design of the resources. The continuous switching between resource authoring related problems and educational design related problems caused major confusion, because the two types of authoring are very different activities.

Conclusion

By using the Recourse authoring tool, it was eventually possible to design an adaptive UoL that takes cognitive and metacognitive assessment as the determinants of the learning activities sequencing. However, this success has too much of a stopgap solution. Efforts highlighted that the QTI specification could not properly handle the pedagogical requirements of the UoL regarding assessment. The "global element" workaround also revealed limitations. Its management remains a hassle even for experimented instructional developers.

It clearly means that an average instructor cannot, in the current development stage of the specifications, without an intensive support, implement confidence degrees in a unit of learning. Despite the availability of complex tools and standards, the following basic requirement still cannot be met: any assessment tool of the next generation should, for each question, provide an easy possibility to get three answers from the student :

- the answer itself;
- a confidence degree expressed in percentages of chances,
- a comment or justification (a few sentences).

Ideally, such tool should also allow to use easily these inputs in order to compute a series of indices (for example the average confidence degree for all the correct responses, etc.).

References

De Finetti, B. (1965). Methods of discriminating levels of partial knowledge concerning a test item. *British Journal of Math. & Statist. Psychol.*, 18, 87-123.

IMS Global Learning Consortium (2005). *IMS Question and Test Interoperability Integration Guide*, Version 2.0 Final Specification.

http://www.imsglobal.org/question/qti_v2p0/imsqti_intgv2p0.html

Jans V., Leclercq D. Mesurer l'effet de l'apprentissage à l'aide de l'analyse spectrale des performances (Using spectral analysis to measure the effect of teaching). In: Depover, C., Noël, B. (eds). *L'évaluation des compétences et des processus cognitifs, Modèles, pratiques et contextes (The evaluation of competences and cognitive processes. Models, practical examples, and situations)*. Paris, Bruxelles: Coll. Pédagogies en développement, De Boeck Université; 1999. p. 304–17.

Verpoorten, D., & Glahn, C. (2010). *Dangerous Knowledge Tour on Web-usability*. IMS content package, Heerlen, The Netherlands. (handle URL pending approval).