Giesbers, B., van Bruggen, J., Hermans, H., Joosten-ten Brinke, D., Burgers, J., Koper, R., & Latour, I. (2007). Towards a methodology for educational modelling: a case in educational assessment. *Educational Technology & Society*, 10 (1), 237-247.

Towards a methodology for educational modelling: a case in educational assessment

Bas Giesbers, Jan van Bruggen, Henry Hermans, Desirée Joosten-ten Brinke, Jan Burgers, Rob Koper
Educational Technology Expertise Centre, Open University of the Netherlands, Heerlen, The Netherlands
Tel: +31 45 576 2209 // Fax: +31 45 576 2907
bas.giesbers@ou.nl

Ignace Latour
Citogroep, Nieuwe Oeverstraat 50, P.O. Box 1034, 6801 MG Arnhem, The Netherlands
Tel: +31 26 352 1111 // Fax +31 26 352 1356
ignace.latour@citogroep.nl

ABSTRACT

Educational modelling is the modelling of educational [sub-] systems. Such a model is a framework containing important concepts, processes and relations. Several models have been published but their development, which we call educational modelling, still is a tedious process. We lack clear guidelines or a methodology. In this article we present a case study in which we take first steps towards the development of a methodology for educational modelling. We do so by analysing our current practice that we typify as expert-driven, model-centred and consensus-based. We explicate the assumptions under this approach and investigate whether these assumptions are explicit and confirmed by our case study. The results give rise to a number of guidelines that can be used by future projects and that provide a first step towards a more systematic approach to educational modelling.

Keywords

Educational modelling, interoperability of educational [sub-] systems, educational ontologies, case study

Introduction

Educational modelling refers to the modelling of educational systems or sub-systems, such as instructional design or assessment. Such a model is a framework that contains important concepts, processes and relations. Instructional design, for example, is modelled in Educational Modelling Language (EML) (see Koper, 2001) and IMS Learning Design (IMS Global Learning Consortium Inc., 2003). Tattersall et al. (2003) present a curriculum model used to help students navigate a curriculum. Hermans et al. (2005) developed a model for educational assessment, described and compared to other models in Joosten – ten Brinke et al. (in press).

Educational modelling can be seen as the building of an ontology – an interrelated collection of entities and their relationships. Although educational modelling is a highly specialized field within educational technology, its products may have a wide-reaching impact through consortia such as IMS and IEEE that foster the development of interoperability specifications and standards in education. Interoperability of an educational system like instructional design or assessment requires a model of the system in question. There still is a lot of work to be done in this field of which the results (i.e., open standards to increase interoperability) have a significant impact.

Several languages have been developed to support the development of ontologies. For example, the OWL Web Ontology Language (W3C, 2004) formally describes the meaning of terminology used in web documents. It was designed to meet the need for a language that facilitates the machine readability of (semantic) web content. A second example is the Unified Modelling Language (UML) (see Booch, Jacobson & Rumbaugh, 1998), which is frequently used as a notation language for a domain or conceptual model in a broad spectrum of software development for which it was originally designed.

Although we now have gained considerable experience in creating educational models, their formulation tends to be a tedious and time-consuming process (Hermans, van den Berg, Vogten, Brouns & Verhooren, 2002). There is a need for guidelines, or a methodology, that helps to build these models in a more efficient way.

The type of methodology that we are looking for combines elements of software and knowledge engineering. Methodologies that cover only one of these domains will not suffice as it may lead to an deficient ontology. Consider, for example, the Unified Process (Jacobson, Booch & Rumbaugh 1999). This is a software development methodology that uses the UML notations that are being used in educational modelling. Thus, UP seems a likely candidate to consider. Its strength is in modelling information and process flows (Abdullah, Benest, Evans & Kimble, 2002) and that is what we need to develop further instrumentation. However, our prime interest lies in the development of a domain model (or ontology), meaning a depiction of the concepts of a domain and the interrelations between them. Here, however, UP only offers some weak heuristics. Knowledge engineering methodologies, on the other hand, offer more support to stipulate domain knowledge. They tend to emphasize the procedural aspects of knowledge, rather than the declarative knowledge represented in a domain model. Thus, using only one type of development methodology may cause to overlook certain types of knowledge categories or the relationship(s) between concepts. This would lead to a deficiency in our ontology for we both need concepts and the interrelations between them.

Further, use-case modelling may be adopted in educational modelling. As Arlow & Neustadt (2002) point out, a use-case is a "*description of a sequence of actions that a system performs to yield value to a user*" (p.15). Modelling and collecting use-cases has proven useful in other instances of educational modelling (for example, see Manderveld, van den Berg, Hermans, Koper & Brouns, 2001; Gorissen & Tattersall, 2005). The work reported here, however, addresses a preceding stage in which a domain model is construed independently from services provided by a system.

The current practice in educational modelling combines knowledge elicitation techniques with UML modelling and is performed more or less ad lib depending upon the context in which the modelling takes place. This does not automatically lead to efficient modelling. A set of guidelines that structure the modelling process will greatly enhance the modelling. Therefore a case study was prepared as part of a joint project of the Open University of the Netherlands and the Central Institute for Test Development (CITO). This project aimed to define a model of educational assessment to '*provide insight into gaps between the different specifications to support assessment exchange initiatives*' (Joosten- ten Brinke et al. (submitted)). The major gap identified refers to the QTI interoperability specification (IMS, 2006) that is effectively limited to "classical" multiple-choice assessments and variations thereof. To create more advanced learning environments one needs more advanced assessment types. Other specifications do not fulfil this need because they are either specifically directed at the (technical aspects of) portfolio assessment (IMS ePortfolio (2005); IMS Rubric (2005), IMS Learner Information Package (2005)) or make use of QTI, such as IMS LD (2003). More information on the model can be found in this article under the heading "Modelling in UML" and in Joosten- ten Brinke et al. (in press).

In the case study data were collected to improve our understanding of current practice in educational modelling. The case study is limited to the initial phase of the project, during which a domain model for educational assessment was construed in a series of expert meetings.

Case characteristics and assumptions

As a first step towards a methodology, we concentrate on a number of general characteristics of our current practice in modelling, and in particular on the assumptions underlying these characteristics. If these assumptions are not met, any methodology based on them is likely to fail in educational modelling.

The first characteristic is that of the *expert-driven* nature of the sessions, in terms of content as well as the organization of the meetings. Experts acted as the source for the domain objects and relations to be used in the model. Furthermore, they defined the goals and procedures for each session. As an alternative, one could have chosen to limit the role of the experts to providing knowledge for the construction of the model and, eventually, evaluating the model.

To ensure that the domain of educational assessment would be covered as complete as possible a deliberate selection of multiple experts was made to cover various levels of education and a wide variety of assessment types. First experts from CITO were invited to the project. CITO is the leading testing and assessment institute in the Netherlands. CITO's main focus is on primary and secondary education but its services are also used in vocational education, higher education, adult education, government and trade & industry. Second assessment experts from Fontys University of Applied Sciences were invited. Fontys offers several higher professional education programmes and is working with innovative assessment strategies. In conclusion, assessment experts from the Educational Technology Expertise Center (OTEC) were chosen because of their experience with various assessment types in higher distance education.

The experts who contributed to the project also covered various expertise areas within assessment, such as measurement theory , item and test construction and validations, deployment of various assessment types, et cetera. In using these experts we expected that coverage and quality of a wide gamut of information could be guaranteed (McGraw et al., 1989). Moreover multiple experts may help to identify and solve differences in, for example, terminology and importance of concepts by discussion. During each meeting, topics to be included in the model were identified and prioritised by the experts. This way an agenda for topics to be discussed was determined. The agenda was guarded by two moderators to prevent topics to be left unspoken.

Second, the sessions were *model-centred*, i.e., all activities were directed towards further development of the educational model. During the sessions the objects and relations that the experts identified were cast 'on the fly' in UML diagrams. Experts were to validate these initial diagrams. An alternative approach would, for example, separate knowledge elicitation and modelling.

Finally, the procedures during the sessions were *consensus-based*. Participants, in particular the experts, had to agree about the procedures as well as about the content of the model. Here an important alternative to consider is whether all work has to be done in *one* team of experts.

In this case study we concentrate on the assumptions associated with the characteristics mentioned and test whether these assumptions were met in the case. For example, the expert-driven approach assumes that experts can define and maintain the overall goals of the project, as well as the goals of the individual sessions. It also assumes that experts can and will express their knowledge. The model-centred approach assumes that experts can translate their knowledge to the UML notations and test the models presented to them against their knowledge. The consensus-based approach assumes that experts can and will collaborate and reach consensus about procedures as well as about the content of the model. Before we present the data collected to test these hypotheses, we describe the case in somewhat more detail.

Case description

Team composition

The team for the project consisted of ten participants. Six assessment experts participated in eight expert sessions. Their task was to identify the building blocks of assessment and their relations. The experts came from three different institutions. They were selected to strike a balance between theory and practice. All experts had more than five years experience in assessment and they all were still active in the field, thus they would allow to capture state of the art knowledge.

Two moderators facilitated the sessions. They were experienced in moderating groups and both were knowledgeable in the educational assessment, although not at an expert level. A modeller, who had experience in modelling of ICT systems, took the lead in all technical UML modelling activities. Finally, a scribe was added to the team to record session notes and to collect evaluative data. Pictures were taken from the result of concept mapping exercises and from writings on the blackboard and flip-over.

Structure and work formats of the expert sessions

During the first session, the dates for seven subsequent experts sessions were planned over a period of two and a half months. Later on, it was decided to use the eighth session as a debriefing session.

Since no firmly established methods for educational modelling are available, the choice of appropriate methods was based on available literature on knowledge elicitation. Cooke (1994) provides an extensive overview of knowledge elicitation techniques discusses their strengths, weaknesses and application across different fields ranging from psychology to business management. She divides the methods into three families called (1) “observations and interviews”, (2) “process tracing” and (3) “conceptual techniques”. The latter contain knowledge elicitation techniques that result in representations of domain concepts, their structure and interrelations. Following McGraw and Harbison-Briggs (1989), Cooke recommends using a combination of techniques to minimize errors and to maximize the scope of domain coverage. In order to elicit declarative knowledge, she recommends concept brainstorming, free association, concept listing and (hierarchical) concept clustering. These then were techniques suggested to the experts. In the first session these techniques were used to formulate a first domain demarcation. Hierarchical concept clustering was then used to create and order ten topics that were scheduled for discussion and modelling in the next sessions. In this way, the experts defined the core agenda for the sessions (this agenda was dropped in the second session however). Note that only concepts and relations brought forward by the experts were included in the model. No background literature or models were incorporated by reference alone, emphasizing the expert-driven nature of the work methods used.

The techniques of brainstorming, free association, concept listing and hierarchical clustering were used throughout the sessions. It was assumed that using these techniques the experts would manage to reach consensus. If, however, no consensus would be reached within the time allotted, the moderators and the experts would make a joint decision on whether or not to stick to the agenda.

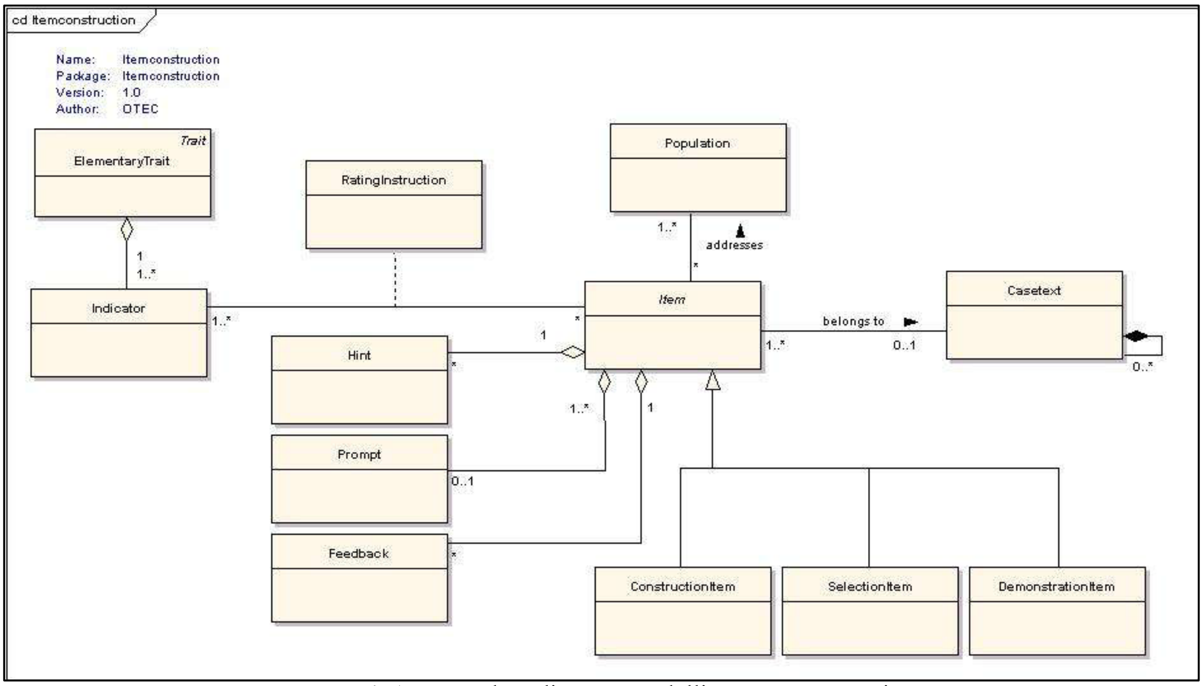

Figure 1. A UML class diagram modelling Itemconstruction

Modelling in UML

UML class diagrams (see Arlow & Neustadt, 2002 for an overview of UML diagrams) were used to model the different classes of objects and their relations in the assessment domain that the experts identified. The model consists of sub models, each depicting a different stage in the assessment process. The first sub-model is *Assessment design*, which describes the initial stage of assessment by stating the reasons for using assessment as well as different assessment types, a description of the assessment policy, an assessment plan, and so on . This is followed by *Item construction* which is built upon the definition of an elementary trait to be assessed and indicators to give evidence for this trait. Third, *Assessment construction* may take place. The central element here is a unit of assessment (UoA) that contains a unit of assessment definition, an assessment type and different items.

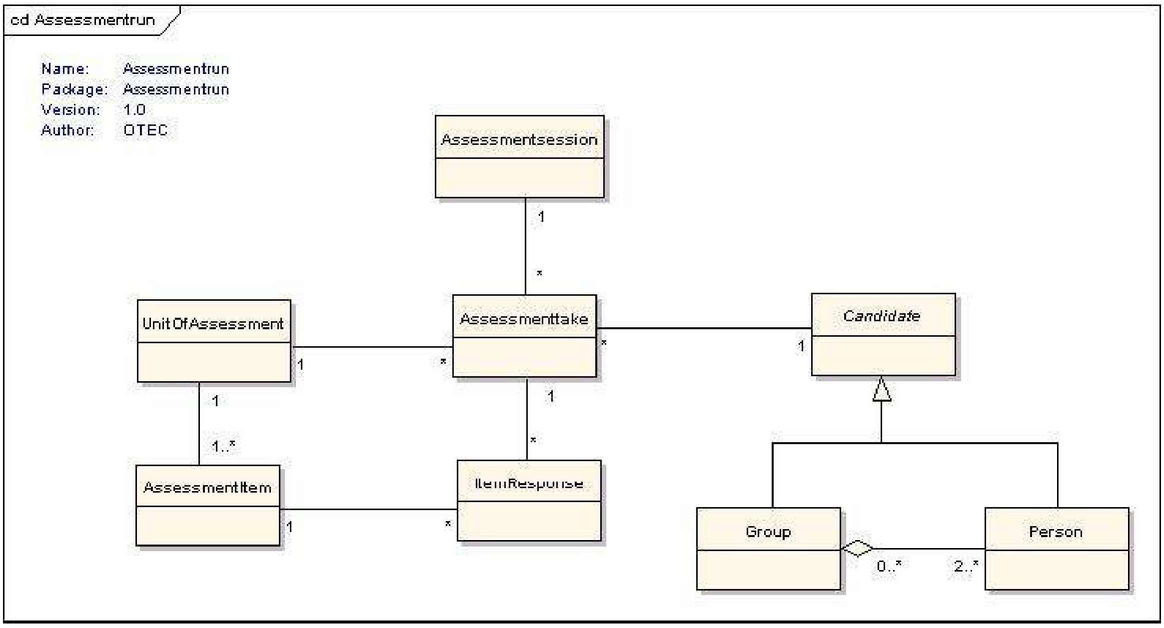

Figure 2. A UML class diagram modelling Assessmentrun

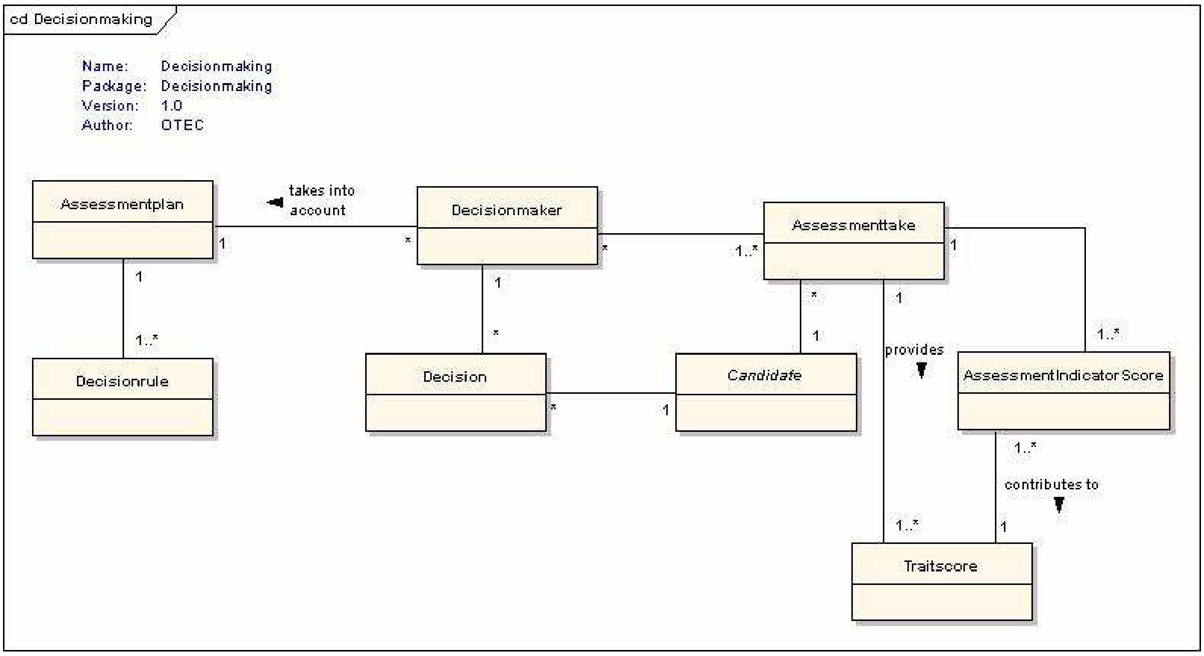

Figure 3. A UML class diagram modelling Decision making

To allow delivery of the assessment to the candidate an *Assessment run* is needed which presents a description of the candidates, sessions, and the output in the form of item responses. Processing of the results is taken care of by *Response rating* which contains the definition of rubric scores, indicator scores and trait scores. Finally, the scores are interpreted during the process of *Decision making.* Examples of classes are candidate, decision rules and the decision.. As an example, three sub models are depicted below. For the complete model we refer to Hermans et al. (2005) and the description in Joosten-ten Brinke et al. (submitted). The model can be downloaded at http://hdl.handle.net/1820/308.

One of the characteristics of the approach is *on-the-fly* UML modelling. The experts' input would be translated immediately into a tentative UML model that the experts would then discuss and amend whenever they felt that was necessary. In between sessions, the UML modeller would further refine the amended, provisional model into a complete (though partial) UML class diagram. This version would be presented to the experts at the start of the next session for final comments and adaptations.

In order to participate in this modelling process, the experts had to be able to comprehend basic diagrams in UML and to perform basic modelling activities. To that end they received a brief introduction to the concepts of UML class diagrams during the first session. *On-the-fly* modelling turned out to be very tedious and was replaced by modelling outside the expert session and the expert session was then used as a panel to evaluate the model.

Method

The investigation reported here is directed at the assumptions underlying the expert-driven, model-centred and consensus-based methods applied in this (and previous) cases. Associated with the expert-driven methods are assumptions regarding the organization of the sessions, and the domain knowledge brought forward during these sessions. It is assumed that (1) the experts can help organize the sessions, that is, they can define, understand and maintain the overall goal as well as the session goals (keep the agenda). As far as the domain knowledge of the experts is concerned, it is assumed that (2) experts can and will express their knowledge of the domain; (3) this knowledge is state of the art; (4) this knowledge is correct and sufficient to create an educational model of the domain, thus excluding the need to use other sources.

Associated with *model-centred* methods are assumptions related to the use of the notations of the models in the sessions. In this case, it is assumed that experts (5) can read and understand basic UML notations, and (6) can translate their knowledge into a UML notation.

Assumptions associated with consensus-based methods are that experts can reach agreements on (7) the organizational matters discussed above; (8) the objects and relations in the domain, and (9) on a UML model for the domain. Finally, it was assumed that (10) using UML would improve consensus building, because it would offer an unambiguous representation of the domain, in contrast to natural language.

Not all assumptions enumerated above were probed. In particular no assumptions have been tested that relate to the domain knowledge of the experts (assumptions 2, 3 and 4). These assumptions were addressed in subsequent phases of the project where the model was tested against on assessment scenarios and models.

Data were gathered at different times during the process: (a) session minutes were made, (b) after each session a questionnaire was administered with questions pertinent to the assumptions, (3) after the session series semi-structured interviews were held with experts and moderators covering, among others, questions relating to the assumptions.

The questionnaire

The questionnaire contained ten statements that were rated on a five-point scale. Questions related to assumptions 1 and 2 asked participants to rate their understanding of the overall goal of the sessions, as well as whether the agenda and assignment for the sessions were clear. Other questions related to these assumptions as well as to assumptions on consensus building (7) asked participants to respond to the clarity and appreciation of work methods used, and to rate the extent to which they could contribute to the sessions and their appreciation of the collaboration in the team.

Other questions were related to the assumptions behind the model-centred approach (5 and 6). After the first session participants were asked to rate the clarity of the introduction to UML. In subsequent questionnaires participants rated their ability to read and explain UML models.

Results

In this section we report the results obtained from the experts. Wherever appropriate these results will be compared to those obtained from other participants in the sessions.

Assumptions behind the expert-driven nature

Table 1 summarizes the ratings of the understanding of the overall goal of the sessions. All medians are above the neutral point on the scale. Minimum and maximum ratings indicate that experts' opinions on this issue were diverging until session 4. During the final sessions they are more in agreement.

Table 1. Understanding of overall goal of the expert sessions

Session #	1	2	3	4	5	6	7
N	5	6	5	4	4	4	5
Mdn	4.0	4.0	4.0	4.0	3.5	4.0	4.0
Min	4	2	1	1	2	4	4
Max	4	5	5	4	4	5	5

As noted above, the overall agenda for the sessions was dropped. Perhaps this led some experts to be less secure on the overall goal. In their interviews experts indicate that maintaining the overall agenda could have improved the process.

After each session participants rated the clarity of the agenda and the assignment for the session and their understanding and appreciation of the methods that were used. The results are presented in Table 2. Overall, the session agenda's were rated as clear, but among the experts opinions vary. The ratings of the clarity of the assignments show a similar pattern as in Table 1 with differences among experts during the first three sessions.

Table 2. Session evaluation by experts

Session #	1	2	3	4	5	6	7
The agenda for the session was clear							
n	5	6	5	5	3	3	4
Mdn	4.0	3.0	4.0	4.0	4.0	4.0	4.5
Min	3	2	3	3	4	3	4
Max	5	5	4	5	5	4	5
The assignment for the session was clear							
n	5	6	4	5	3	4	4
Mdn	3.0	3.5	3.5	4.0	3.0	2.5	4.0
Min	2	2	2	3	3	2	3
Max	4	4	4	4	4	4	5

The moderators provided rates that in general were slightly lower than those given by the experts. They found the agenda of session 3 not clear (*Mdn*=1) and during that session their rating on the clarity of the overall goal was low (*Mdn*=2).

Assumptions behind the model-centred nature

Five experts provided a rating on the clearness of the introduction to UML (*Mdn* = 4.0, min=2, max=5). We also asked the experts to rate the extent to which they understood why UML was used in the project. Five experts rated their understanding (*Mdn*=4.0, min.= 2, max. = 4). Two experts provided the minimum ratings on both scales.

Table 3 presents the experts' rating of their ability to read the UML diagrams. There is a constant drop in the median rating until session 4, from whereon the experts start to rate their ability somewhat higher. Note that the range of ratings here is high until session 4 and the minimum ratings are very low.

Table 3. Ratings of the experts of their ability to read the UML models

Session #	1	2	3	4	5*	6
N	5	5	5	4	--	5
Mdn	4.0	3.0	2.0	3.5		4.0
Min	1	1	1	3		3
Max	4	4	4	4		4
* Due to an error in the forms used no data were collected in this session						

From session two onward the educational model for assessment was developed and after each sessions experts rated their understanding of the educational model, as well as their ability to explain the model to colleagues. Table 4 reports these ratings. Minimum ratings for understanding the model are at the neutral point. When the experts have to consider whether they could explain the model, the median rate remains positive, but the high range of the rates indicates important differences between experts.

Table 4. Ratings of understanding of the educational model developed

Session	1	2	3	4	5	6	7
	Understand educational model developed in current session						
n	n.a.	5	5	5	4	4	5
Mdn	n.a.	4.0	4.0	4.0	3.5	5.0	4.0
Min	n.a.	3	3	2	3	4	4
Max	n.a.	5	5	4	4	5	5
	Can explain model to colleagues						
n	4	6	5	5	3	4	5
Mdn	4.0	4.0	4.0	4.0	4.0	3.5	4.0
Min	3	1	1	2	2	3	2
Max	4	5	5	4	4	4	4

The ranges of the rates presented above indicate that on several measures a majority of experts gave favourable ratings, whereas a minority gave unfavourable ratings. Closer inspection of the data revealed a pattern where two experts gave low ratings on the clarity of the overall goal of the sessions (Table 1), capability of reading UML (Table 5) and the capability to explain the model to a colleague (Table 6). On other scales with minima below 3 they were found to have rated these scores. Inspection of the session notes indicated that the contributions of these two experts were diminishing over the course of the sessions. This was confirmed in their post-hoc interviews. They indicated that they needed more time between sessions to prepare themselves. Although more experts reported this, these two experts indicated that the highly abstract and technical nature of most of the discussions made them feel they had little to add.

Assumptions behind the consensus-based nature

In the session notes as well as in the interviews consensus building emerges as a very tedious process. Experts rated the clarity of and their appreciation for the work methods and the collaboration with others in the session. As a proxy to consensus building we asked them to rate to which extent they could contribute to the session's work. The results are presented in Table 5 and Table 6. The clarity of the work methods was rated consistently above neutral for all sessions and with little differences only between the experts. The appreciation of these work methods, although in general positive, shows differences between the experts.

Some contrasts between experts and moderators seem to occur as well. The moderators rated the clarity and their appreciation of the work methods of sessions 3 and 4 lower ($Mdn \leq 2$) than the experts. The median rating of the appreciation of the contribution to the session is at moderate high level. The range of ratings varies over the sessions. Whereas experts sometimes express that they could not contribute much to a session, they all rate the collaboration in the team as positive.

Table 5. Clarity and appreciation of work methods

Session	1	2	3	4	5	6	7
The work methods to be used were clear							
n	5	5	5	5	4	4	5
Mdn	4.0	4.0	4.0	4.0	4.0	4.0	4.0
Min	2	3	3	3	4	3	3
Max	5	4	4	4	4	4	5
Appreciation of work methods							
n	5	6	5	5	4	4	5
Mdn	4.0	5.0	4.0	4.0	4.0	4.0	3.0
Min	4	2	2	3	3	4	1
Max	5	5	4	4	4	4	4

Table 6. Appreciation of contribution and collaboration during the sessions

Session #	1	2	3	4	5	6	7
Contribution to session							
n	5	6	5	5	4	4	5
Mdn	4.0	4.0	4.0	4.0	4.0	4.0	4.0
Min	4	2	4	3	3	4	2
Max	5	5	4	5	5	5	4
Collaboration with others							
n	5	6	5	5	4	4	5
Mdn	4.0	4.0	4.0	4.0	4.5	4.5	4.0
Min	4	4	4	4	4	4	4
Max	5	5	5	5	5	5	5

The moderators here, as well, tend to rate slightly lower than the experts. In sessions 4 and 7 their satisfaction with the collaboration with other group members drops below neutral (*Mdn*=2). The session notes as well as the interviews contain several statements in which experts complained about having insufficient time to prepare for the next meeting. Although the median ratings of the experts seem to indicate that they were satisfied with the techniques used and with the way of working that emerged, some experts felt more and more reserved to make a contribution. In their interview they indicated that this was due to the technical nature of the discussions. This is likely to be related to the centrality of the UML representation of the model.

Discussion

As a first step towards a methodology for building educational models, we took a closer look at the assumptions underlying our current practice that we typified as being expert-driven, model-centred and consensus-based. What we reported on is the perception of the participants of phenomena related to these assumptions, for example their reported capability in interpreting the educational model. These perceptions are often corroborated by the interviews or session notes.

The results presented here indicate that most assumptions were only partly met, certainly when we consider consensus building simultaneously with the two other characteristics. For example, experts had no problem to define and understand the overall and session goals, but this neither ensured that the goals were shared, nor that the goals or agenda were maintained. It took three sessions before all participants had a clear and corresponding view on the goals. Inspection of several other ratings demonstrate that behind the favourable median ratings are ranges that indicate that this was not a shared perception among experts.

It was assumed that experts could read basic UML models, that they could translate their knowledge to the UML models and test the presented models on their knowledge. However, on-the-fly modelling turned out to be a cumbersome process. The experts' rating of their ability in UML dropped steadily and even the increase in session 4 (see Table 3) may very well be the result of the absence of one of the experts who rated all the questions related to

UML low. In our interpretation on-the-fly modelling assumes more UML knowledge and experience than the experts could possibly bring to the sessions.

The assumption that UML would support consensus building did not hold. The session notes and interviews clearly show that consensus building was a very hard process. Yet, the ratings of the collaboration and contribution to the sessions do not indicate that there were conflicts within the team. In our interpretation consensus building was hampered by the interpretation of the UML model and differences of opinion on the importance of modelling particular aspects of the domain. This is supported by an observation made during the debriefing session. Here an outline of the model was presented without using UML notations. All experts agreed to this model, but they also indicated that opening up details would inevitably lead to new discussions.

In the sessions it became clear that the approach was based on a number of assumptions that proved wrong. The first sessions were spent on topics brought forward by the experts. Progress in modelling was extremely slow. It then became clear that not all work could be done in the sessions and that more preparation was needed. Then however experts reported that they did not have enough time to prepare between the sessions. During the fifth session the group decided that any further modelling would be done outside the group of experts, who would continue to function as a panel to discuss the resulting model. During the seventh session the moderators decided to take some time to recapitulate and model the final version by themselves (in fact, dropping their role) with the modeller and using the experts' input thus far. The eighth session was used as a debriefing session to present the final outcome and here the experts indicated they all were satisfied with that version of the model. The debriefing session also provided the opportunity to make suggestions for future modelling processes.

One might argue that the methods used both overrated and over-asked the experts. As far as modelling was concerned, the experts were clearly over-asked. On-the-fly modelling pre-supposed UML capabilities that our experts did not possess, or could have developed before or during the sessions. On the other hand, experts were overrated. They were the only source of knowledge used to structure the domain and to define the components of the model. Only in later stages the model was compared and contrasted to other models, as reported in Joosten-ten Brinke et al. (submitted). The expert-driven approach also led to deadlock whenever consensus could not be reached

Translating our experience into recommendations for future projects, we come to the following six points:

1. Establish a baseline reference for the project by collecting relevant sources and identifying the domain knowledge to be collected.
2. Avoid knowledge elicitation in large groups.
3. Separate modelling and knowledge elicitation.
4. Cater for expressing different views on the domain.
5. Use expert panels in combination with case descriptions to test and validate the model.
6. Define ownership and responsibilities. Who will own the model? Who will collect the domain knowledge? Who is responsible for modelling? How and by whom will the model be tested and evaluated? In order to avoid diffuse tasks and responsibilities these need to be defined for all persons involved in an educational modelling project.

These guidelines are no guarantee for success. Fortunately, failing to meet them has not caused a failure either. The model for assessment that was ultimately produced by the project was welcomed in a positive way by several independent reviewers and experts in assessment.

References

Abdullah, M. S., Benest, I., Evans, A., & Kimble, C. (2002). Knowledge Modelling Techniques for Developing Knowledge Management Systems. In *Proceedings of the 3rd European Conference on Knowledge Management*, 15-25, retrieved December 28, 2006 from http://www.cs.york.ac.uk/mis/docs/ECKM2002.pdf.

Arlow, J., & Neustadt, I. (2002). *UML and the Unified Process*, London: Pearson Education.

Booch, G., Jacobson, I., & Rumbaugh, J. (1998). *Unified Modelling Language User Guide*, Boston, MA: Addison-Wesley.

Cooke, N. J. (1994). Varieties of knowledge elicitation techniques. *International Journal of Human-Computer Studies, 41* (6), 801-849.

Gorissen, P., & Tattersall, C. (2005). A Learning Design Worked Example. In: Koper, R. & Tattersall, C. (Eds.), *Learning Design: A Handbook on Modelling and Delivering Networked Education and Training*, Berlin-Heidelberg: Springer Verlag, 3-20.

Hermans, H., van den Berg, B., Vogten, H., Brouns, F., & Verhooren, M. (2002). *Modelling test-interactions,* Series/Report no.: OTEC2002/25, Heerlen: Educational Technology Expertise Centre, Open University of the Netherlands.

Hermans, H., Burgers, J., Latour, I., Joosten-ten Brinke, D., Giesbers, B., van Bruggen, J., & Koper, R. (2005). *Educational Model for Assessment version 1.0*, Heerlen: Open University of the Netherlands & Citogroep, retrieved December 17, 2006, from http://hdl.handle.net/1820/308.

IMS ePortfolio (2005). *IMS ePortfolio Specification. Version 1.0. Final Specification*, IMS Global Learning Consortium, retrieved November 15, 2006, from http://www.imsglobal.org/ep/index.html.

IMS LD (2003). *IMS Learning Design Specification. Version 1.0. Final Specification*, IMS Global Learning Consortium, retrieved November 8, 2006, from http://www.imsglobal.org/content/learningdesign/.

IMSLIP (2005). *IMS Learner Information Package. Version 1.01*, Final Specification IMS Global Learning Consortium, retrieved November 15, 2006 from http://www.imsglobal.org/profiles/.

IMS QTI (2006). *IMS Question & Test Interoperability. Version 2.1. Public Draft Specification*, IMS Global Learning Consortium, retrieved November 2, 2006, from http://www.imsglobal.org/question/index.cfm.

IMS Rubric (2005). *IMS Rubric Specification. Version 1.0. Final Specification*, IMS Global Learning Consortium, retrieved November 15, 2006, from http://www.imsglobal.org/ep/epv1p0pd/imsrubric_specv1p0pd.html.

Jacobson, I., Booch, G., & Rumbaugh, J. (1999). *The Unified Software Development Process*, Boston, MA: Addison-Wesley Professional.

Joosten-ten Brinke, D., van Bruggen, J., Hermans, H., Burgers, J., Giesbers, B., Koper, R., & Latour, I. (in press). Modeling assessment for re-use of traditional and new types of assessment. *Computers in Human Behaviour*.

Koper, E. J. R. (2001). *Modeling Units of Study from a Pedagogical Perspective: The pedagogical meta-model behind EML*, (OTEC working paper), Heerlen: Educational Technology Expertise Centre, Open University of the Netherlands, retrieved November 8, 2006, from http://learningnetworks.org.

Manderveld, J., van den Berg, B., Hermans, H., Koper, R., & Brouns, F. (2001). *EML models of IMS Learning Design use cases*, (Series/Report no.: OTEC2001/17), Heerlen: Educational Technology Expertise Centre, Open University of the Netherlands.

Tattersall, C., Manderveld, J., van den Berg, B., Es van, R., Janssen, J., Waterink, W., & Bolman, C. (2003). *ROMA: Road Mapping*, (LTD project plan), Heerlen: Educational Technology Expertise Centre, Open University of the Netherlands, retrieved November 18, 2006, from http://hdl.handle.net/1820/86.

W3C (2004). *OWL Web Ontology Language Overview*, retrieved November 13, 2006, from http://www.w3.org/TR/2004/REC-owl-features-20040210/.