Practical Assessment, Research & Evaluation

A peer-reviewed electronic journal.

Volume 13, Number 2, February 2008 ISSN 1531-7714

Practical Aspects of Task Allocation in Design and Development of Digital Closed Questions in Higher Education

Rob Hartog, *Wageningen University, The Netherlands*
Silvester Draaijer, *Vrije Universiteit, The Netherlands*
Luuk Rietveld, *Delft University of Technology, The Netherlands*

For projects on design and development of digital closed questions in higher education the task structure is analyzed. Based on fifteen small to mid sized projects in four universities, a practical set of tasks is defined and practical aspects of task allocation are described and discussed. Ten design and development scenarios are presented. Based on time registrations in the projects and on a few experiments, estimates are given for the most important cost categories in the budgets for the scenarios.

The increased availability of Learning Management Systems and facilities for computer-based assessment (CBA), induce more and more teachers in higher education to invest in the design and development of pools of digital closed questions. A closed question is any fixed response item that can be administered by a computer. Digital closed questions are being developed for computer-based assessment but also for use as activating learning material (ALM). In practice, several hybrid roles for closed questions can be distinguished.

To take full advantage of innovative closed questions, considerable knowledge is required, regarding question design, educational measurement and multimedia development. In addition, a set of practical skills is needed with respect to question editing and entry, image processing and management of questions and pictures. Therefore, design and development of question pools in higher education is often a matter of teamwork in projects. The number of students that will use the questions resulting from such a project will generally be much lower than the number of participants in a nationwide or large scale test or exam. Because the costs per student tend to determine what budget is acceptable, smaller numbers of students in practice correspond to smaller project budgets. Thus, realistic budgets for the design and development of a set of questions in higher education are much lower than budgets for large scale tests. If there would be strict quality criteria for digital closed questions to be used in a CBA or ALM role and if these quality criteria would be widely accepted, reality might be different. Currently, quality is de facto an implicit derivative of the quality of the design and development team and their working procedures.

This article focuses on small to mid sized projects for the design and development of closed questions in higher education with no explicit quality criteria. These projects were projects in a larger program (Hartog, 2007) aimed to develop a methodology for the design and development of digital closed questions. One of the aims of this program was to identify what aspects design and development of digital closed questions for different roles (ranging from pure ALM role to pure CBA role) may have in common. The first author was supervisor of this program and took part as educational technologist in seven of the projects. The second author took part as educational technologist in two of the projects. The third author took part as subject matter expert in two of the projects. The program also involved a number of projects on educational measurement issues related to innovative closed questions and interoperability. The results of these latter projects fall outside the scope of this paper (see Hartog, 2007).

The article describes the most common classes of human resources, defines and discusses the tasks and matches these tasks to possible functions that might be defined within the university. Suggestions are given to prevent waste of efforts. Furthermore, the article presents a set of scenarios and corresponding budget templates. For a number of entries in these templates, cost estimates are given.

METHOD

Data for this article were collected from fifteen small to mid sized projects in higher education in which closed questions for learning goals and objectives in natural sciences, engineering sciences and social sciences were designed and developed. On average in the projects, about one third of the developed questions were the common type multiple choice questions. About two thirds of the developed questions make use of other question types such as multiple-correct, matching, ranking, hot-spot, drag-and-drop. Examples of innovative use of question types are presented in (Hartog, 2007).

The aim of the projects was twofold: first to design and develop pools of digital closed questions and second to develop design requirements, design guidelines and design patterns for new design and development projects in higher education. As such, the projects can be classified as developmental research projects (Richey, Klein, & Nelson, 2004).

Table 1 presents an overview of these projects. In the table, the case, course level, course subject, number of students, role of the questions, the authoring software, the set-up of the development team and the average design and development time per question are listed.

Progress and experience was reported at regular time intervals. Each project was evaluated and attempts were made to use experience in the form of requirements, guidelines and patterns in the next project. The most tangible results of the projects were more than 2000 questions and about 30 design patterns.

At regular time intervals, initially every three months, progress in terms of newly developed questions was reported. For reasons of accountability, the time invested by every person in the project was registered. Furthermore, observations were reported as to inefficiencies, problems and issues that were recognized as important. From now on, the term 'case study' will be used to refer to the body of qualitative and quantitative data and the corresponding analysis of a project.

Analysis of the collected quantitative data (numbers of questions, designed and developed and corresponding time registrations) and qualitative data (observations, descriptions of working procedures) revealed a common task structure. This was a basis on which ten scenarios were developed for small to mid sized projects for design and development of closed questions in higher education. The next sections describe resources and roles of team members, tasks and options for allocating this task and issues related to the costs of this task. In particular, one section presents a budget template for each of the scenarios.

DESCRIPTION OF THE DESIGN AND DEVELOPMENT CONTEXTS

In this section, resources that are needed for question design and role descriptions of team members within a design and development project are described.

Question authoring and delivery environments

First of all, without any hard- and software system, there would not be any project for digital question design. Probably every institute for higher education now has a Learning Management System and sometimes a dedicated computer-based assessment system. These systems offer support for authoring questions and for managing questions, pools of questions and exams.

In eight of the fifteen case studies, Blackboard version 6.0 (2006) was used as Learning Management System. In one of the projects N@tschool (ThreeShips, 2007) was used as Learning Management System. Most of the Learning Management Systems offer support for 'quizzes' and 'tests' that primarily contain closed questions. In four of the fifteen projects Questionmark Perception version 3.x (2002) was used. Finally, in two projects, a Questions and Test Interoperabilty v2.x (QTI 2.x) delivery system was used. For these two projects, the questions were edited directly in QTI 2.x XML templates.

Instructors use Learning Management Systems and computer-based assessment systems to present 'quizzes' to students and for summative assessment in regular courses. These systems provide a number of new question types, or seemingly new question types, which are often referred to as 'innovative'. A number of these types involve the use of multimedia.

Different functions and different competencies

The case studies revealed four specific functions within a design and development project. These functions are:

- Subject Matter Expert;
- Assistant SME;
- Educational Technologist;
- Rendering Specialist.

The Subject Matter Expert usually is a professor or associate professor. The professor is also the principally responsible person for the content of a course, the learning goals and for the development of questions. Alternatively, the subject matter expert may be an invited speaker, for instance from industry.

Educational Technologists are the designated persons to provide knowledge and skills with respect to the design and development of closed questions, the possibilities and limitations of the available authoring and question delivery

Table 1 Overview of 15 small to mid sized D&D projects

Case	Course level	Course subject	Number of students per year (about)	Role of the questions	Software	Development team	Average D&D time/ question (in minutes)
WU1	Master	Food Safety (Toxicology/Food Microbiology)	30	summative	QM	SME and ASME	160
WU2	Master	Food Safety Management	30	activating	Bb	SME and ET	150
VU1	2nd year	Heart and Blood flow (physiology, ECG measurement and clinical ECG interpretation)	300	diagnostic and summative	QM	SME and ET	220*
VU2	3rd year	Special Senses (vision, smell, hearing, taste, equilibrium)	300	summative	QM	SME and ET	80
TUD1	3rd year	Drinking water treatment	30	activating	Bb	SME and ASME	85
WU3	Master	Epidemiology	100	summative (open book)	QM	SME and ASME	130
TUD2	3rd year	Sanitary Engineering	50	activating	Bb	SME and ASME and ET	95
WU4	Master	Food Toxicology	100	summative	QM	SME and ASME	130
WU5	Master	Food Micro Biology	40	activating	Bb	ASME	80
WU6	Master	Advanced Food Micro Biology	30	activating	Bb	ASME	130
WU7	Entry Master	Food Chemistry (general introduction module for candidate students)	open self test WWW	diagnostic	QTI delivery	SME = ET	120**
WU8	Entry Master	Food Toxicology	open self test WWW	diagnostic	QM	SME and ASME	120

Table 1 Overview of 15 small to mid sized D&D projects

WU9	Master	Sampling and Monitoring	30	diagnostic (self -)	Flash	SME and ASME and ET and Flash programmer/ design patterns used	80**
WU10	Master	Food Safety Economics	30	summative (not open book)	Bb and on article	SME and assistant and ET/design patterns used	***
FO1	1st year	Curriculum: General Sciences	30	Diagnostic-'plus'	N@t-school	SMEs and Rendering Specialist	160**

Note. WU = Wageningen University, VU = Vrije Universiteit, TUD = Delft University of Technology, FO = Fontys University of Applied Science, QM = Questionmark Perception, Bb = Blackboard LMS, QTI = Question and Test Interoperability 2.0 format, N@tschool = N@tschool LMS, SME = Subject Matter Expert, ET = Educational technologist, ASME = recently graduated student or student-assistant with subject matter expertise but not at SME level, RS = Question Entering and Picture Processing Specialist.

* Time included extensive training sessions of SME with ET, aiming at using other than MC questions.

** For a number of questions only time for design in Word was registered. For those questions average of 20 minutes/question for RS was added.

*** Time registration included too many other activities for which correction was not possible.

environments. The educational technologist is assumed to have broad knowledge from the field of instructional design and educational measurement. Typical sources used in the projects were (Scalise & Gifford, 2006), (Haladyna, 1997, 2004) and (Bull & McKenna, 2004). In addition, the educational technologist has to play an important role in the project definition and project set-up. On that basis, educational technologists are to provide design guidelines and present design patterns. A separate parallel project was defined for investigation of issues with respect to educational measurement. Insofar the design and development projects encountered questions with a strong educational measurement component these questions were passed on to this parallel project.

For the actual design and development none of the teams incorporated an educational measurement specialist. For the design and development of questions for the ALM role this was not deemed relevant because measurement was not the primary function of those questions. For the design and development of questions for the CBA role, the combination of the educational technologists, access to literature and the link to the parallel project was considered sufficient.

An important aspect of design and development projects in higher education is that many of the relevant learning objectives cannot be understood or grasped by those team members who are not subject matter experts. This puts a tension on the position and possibilities of the educational technologist within such projects. The educational technologist costs less per hour than a subject matter expert.

The assistant of the subject matter expert has some subject matter knowledge but cannot be considered an expert. Often, the assistant of the subject matter expert is an almost or just graduated student within the relevant discipline. The subject matter knowledge of the assistant is greater in comparison to the subject matter knowledge of the educational technologist. The assistant however, will usually not have any specific question design and development competence. In most universities, the typical assistant subject matter expert will be hired just for the project. The assistant is always considerably cheaper than the educational technologist. For the majority of the small to mid sized projects, an assistant was appointed to contribute to the design and development of the questions.

Above, the term 'rendering specialist' refers to the question entry and picture processing specialist or service. This specialist (or pool of specialists, for example within an institution's audio-visual services department) is someone who is proficient with desktop computers and has a lot of routine with question entry and elementary picture processing tasks. Thus, the productivity of the rendering specialist can be very high as long as his tasks are well defined. In practice the latter implies that the questions to be entered are available in a very clear format (for example MS-Word documents with sufficient annotation) and that the entry task can be completely outsourced to such a team member. In one of the case studies a rendering specialist as defined above, did most of the question entry and picture processing work. The rendering specialist is not necessarily cheaper than the assistant of the subject matter expert.

The four functions do not imply that every design and development project necessarily involves each function. For example, it is imaginable that a subject matter expert decides to fill a question bank without support, just using a readily available authoring environment. Furthermore, a subject matter expert can perfectly well realize a question bank by appointing a subject matter expert assistant or a rendering specialist and delegate work to them (for example a help-desk employee). The scenarios that are presented below, take these set-ups into account. Because the authoring environments are usually considered as overhead costs, accounted for within institution wide budgets, human resources are the most dominant factors for the costs of a project. In Table 2, the roles, competencies and relative costs of the team members of mid sized question design and development project are listed.

PRACTICAL TASK ANALYSIS

Because of similarity in used question types and software tools, design and development of closed format questions for both Computer Based Assessment and Active Learning Material always includes a number of common tasks. In this section, the tasks in mid sized projects on the design and development of closed questions are described. The tasks cannot be mapped one to one to phases in a project because tasks may overlap considerably. The practical task analysis has been carried out from the perspective of actual design and development of innovative questions. Furthermore, we have tried to highlight what design and development of digital closed questions for different roles have in common and what the differences are. A task analysis primarily focused on the delivery of a complete assessment would have resulted in a different set of primary tasks.

Defining the Project

Every project requires that some effort is invested in assessing the context of the project and the context of the project results. On that basis, a realistic project plan and a corresponding budget can be defined and financial means for the project can be acquired. The project plan should

Table 2 Roles, competencies and costs for question design and development

	SME	ET	ASME	RS
Cost/hour	High	Medium	Low	Very low
Subject Matter Knowledge	High	Low	Medium	None
Question design and development Knowledge	Low	High	Low	Low
Educational Measurement	None	Medium	None	None
Knowledge of Authoring environment	Low	High	Low	Medium
Routine with the Authoring environments and other computer tools	None	Medium	Medium	High

Note SME = Subject Matter Expert; ET = Educational Technologist; ASME = assistant of the Subject Matter Expert; RS = Rendering Specialist

specifically describe the intended output of the project, the role of the questions, available resources and a deadline. In the case studies, these variables have shown to be important determinants for the quality and success of the project in terms of effectiveness and efficiency. Determining the available resources at the start of the project implies:

- which Learning Management System or Computer Bases Assessment system will be used (or is available);
- whether there are well defined learning objectives or previous questions or exams available;
- what learning materials are available;
- whether there are design patterns available;
- what the number and type of questions to be designed and developed should be.

For the CBA role, the case studies showed that in practical contexts a tangible and sensible goal is a pool of questions that is sufficient for four exams and a trial exam. The reason for this is that subject matter experts generally will need several equivalent exams for a few successive cohorts of students. Furthermore, subject matter experts need a trial exam which shows the students what to expect for an upcoming digital exam. When an exam contains about sixty closed questions this implies that about sixty clusters of five equivalent questions need to be designed and developed. In fact, the first question that is designed for such a cluster should be a good operational definition of the detailed learning objective in this cluster. For exams with sixty questions, about three hundred questions will have to be designed and managed. This requires that the clusters are labeled. Part of defining the project should involve a conscious decision with respect to the composition of the development team.

Setting Up the Project

Given an approved project, the project plan can be worked out in more detail.

A development team can consist of one or more subject matter experts, an educational technologist, an assistant of the subject matter expert and a rendering specialist. For reasons of cost efficiency and because the time of most subject matter experts is scarce, it should be the intention in a project to delegate as much as possible of the subsequent tasks to educational technologists, assistants and rendering specialists. For example, for entering questions in a CBA system, a subject matter expert or educational technologist is actually too expensive. Such work is more appropriate for an an assistant or a rendering specialist. Also, the assistant or rendering specialist will often have more routine in question entering and picture processing and therefore can execute the task more quickly.

Analysis of the case studies highlighted the fact that the diary of many subject matter experts seldom displays empty time slots. Given their crucial role in setting learning objectives, providing inspiration and validation of questions, everyone involved should be prepared to dynamically adapt the project agenda to availability of the subject matter expert. In order to avoid frustration and delays in project progression it is advisable to set as soon as possible due dates for delivery of specific batches of questions (e.g. for specific topics, learning objectives or question types) in different stages of completion. These stages are typically characterized by a draft version, a revised draft in some intermediate form of representation, and a final version in the authoring environment.

The project set-up will almost always imply some training. More specific, for computer-based assessment, a subject matter expert and an assistant must be made familiar with elementary knowledge on educational measurement and question design. Typical resources for such training are (Bull & McKenna, 2004; Frary, 1995; Haladyna, 1997, 2004; Kehoe, 1995a, 1995b)

Furthermore, subject matter experts and assistants must be made aware of the possibilities and limitations of digital questions and the specific software application. In the projects a number of fundamental problems with response processing for innovative question types were identified. These problems were passed on to an educational measurement project within the program and are beyond the scope of this paper (Hartog, 2007).

In the fifteen projects, no training material was found that is adequate for subject matter experts or assistants as defined in this article. Most of the knowledge that an educational technologist had readily at hand is based on handbooks listed above. However, examples in these handbooks stem from secondary or vocational education, or from disciplines that had not enough in common with the disciplines in the projects. Also, presenting requirements as to correct grammar and clear formulations in the form of guidelines was not appreciated. In the case studies, some training material was developed for the subject matter experts in the form of about sixty design guidelines and a set of design patterns (Hartog, 2007). Experience in the case studies suggested that design patterns were more helpful than design guidelines.

For Activating Learning Material, elementary knowledge on learning and instruction is necessary. In a number of projects use was made of (Keller, 1983; Smith & Ragan, 1993) and (Merriënboer, 1997). In a later stage also (Fenrich, 2005) came in view.

Collecting and defining learning objectives

When designing and developing questions for the CBA role, it turned out to be an effective approach to define a label for each cluster of five equivalent questions in an early stage of the project. Such a label can be denominated a 'bucket' for which questions need to be designed. If there is a list of detailed learning objectives available, this will reduce the effort needed for this task. However, in the case studies, there was seldom an adequate list and when there was such a list it allowed for too many interpretations. Often, an even more specific subject matter denomination was necessary up to the level of specific micro-subjects within a course. Previously developed sets of (mostly open) questions, assignments or learning materials, such as presentations or lecture notes could partly be used to extract and define the learning objectives up to micro level.

The task of 'labelling clusters' is irrelevant if the project aims at questions for the ALM role. In such a case, the team should make an ordered list of detailed learning objectives for which learning can be supported by closed questions. The ordering should be based on a quick cost/benefit analysis. This cost/benefit analysis should identify for which learning objectives it will require relatively little effort to develop motivating closed questions with a high expected impact.

When the project focuses on CBA, the assistant was usually able to extract a part of the list of learning objectives from overall learning goals in combination with learning materials such as slide presentations, textbooks and from previous exams. To some extent, the educational technologist was able to coach the assistant. However, the subtask of defining a set of labels ('buckets') could never be completed without involving the subject matter expert.

In case of a project for the design and development of questions for the ALM role, the assistant was usually able to indicate some pieces of learning material that are – at the start of the project – insufficiently complemented by activating learning material. Exam results of previous cohorts also pointed the assistant towards learning objectives that call for additional activating learning material. When the subject matter expert becomes familiar with the real possibilities of innovative closed questions, it will be the subject matter expert who can best identify those learning objectives which are likely to have a low cost/benefit ratio. The case studies showed that many subject matter experts need some training in order to become familiar with the real possibilities of innovative closed questions.

Design and intermediate representation of questions

Ultimately, design and development of a closed question implies that a micro learning objective is represented in the form of a closed question. Assuming that the designer(s) has/have such a learning objective in mind, a first idea of a question (or cluster of questions) must be generated. The remainder of the design and development of a closed question will then involve:

- deciding on the exact interaction type
- including a case
- deciding on including of media
- authoring the text-based components of the question.

How the first draft of questions comes about depends on the knowledge and skills of the subject matter expert assistant or the skill of the educational technologist to inspire them. The initial training may help in this process.

In three case studies, design patterns proved to support both the generation of ideas for questions as well as decision making for the used interaction types (Hartog, 2007). Design patterns can form a powerful tool to let subject matter experts and assistants see the possibilities of digital closed questions and also reduce costs through a more effective generation of first draft questions.

The case studies revealed that the first drafts of questions are usually laid down in MS-Word documents with annotations on specific detail: the intermediate representation of questions. The relatively easy method of creating, editing, revising and sharing MS-Word documents is the principal reason for that approach. Another reason is that subject matter experts are familiar with a standard text editor but would have to invest considerable time in learning to use a question authoring environment. Often email-communication was used to share information. The case studies revealed that such communication is very sensitive to problems with versioning.

Including media in questions may involve designing or finding a picture or designing or finding an audiovisual object.

The design task requires deep subject matter knowledge and understanding. This implies that the subject matter expert and the assistant must do most of the work. The educational technologist can provide inspiration in terms of design patterns and by suggesting guidelines. The extent to which the subject matter expert can delegate the design task to the assistant depends very much on the subject matter knowledge of the assistant, on the availability of learning materials and on the question design competence of the assistant. Within the fifteen case studies, the output of assistants in terms of quantity and quality differed widely.

The aggregation level of the case study data is not adequate for determining the costs that are involved in this part of the design process. However, the fifteen case studies highlighted many sources of inefficiency. This resulted in the following lists of don'ts in order to keep the costs within limits.

- Don't search for a specific picture, only use readily available materials.
- Don't make drawings or pictures, but if you do, use them for more than one question
- Don't develop case-based questions, but if you do, make sure it is a fertile basis for a number of questions.
- Don't start by default making traditional MC questions; do invest some time in starting with different types that do not require developing distracters.
- Don't design and develop instances of innovative question types for assessments unless scoring is adequately supported by the available CBA system and the rationale for the scoring rules used by the available system is transparent to faculty and students.
- Don't write extensive feedback.
- Don't let the assistant develop questions for which no design patterns or examples exist.

It is important that the subject matter expert has contact within short time intervals with the assistant. This will prevent that the assistant invests much time in designing questions that later turn out not to conform to the learning goals and have to be discarded. The case studies confirmed that it is difficult to represent detailed learning objectives in some form other than the question itself. In some of the projects the assistant would, based on an initial formulation of a learning objective in natural language, design questions which were completely of the mark. Furthermore, it also occurred that at the end of the course period detailed learning objectives for which questions already were developed, had to be removed from the list of detailed learning objectives. One of the reasons for this was for instance that guest lecturers tended to change ad hoc the content of their lectures.

All in all, development efforts that lead to questions that are useless increase the average development effort per useful question. It is believed that this is one of the factors leading to gross underestimation of design and development efforts.

Validating questions

When a first draft of a question has been made, the question will have to be validated, checked and revised. Validating the first draft involves more than just answering the questions and checking if the answer is 'correct'. It also involves checking for errors and ambiguities in the question formulation. Most of all, the validator has to check if the question really measures (i.e. operationally defines) a learning objective (in case of CBA role) or stimulates the intended action and line of reasoning (in case of ALM role).

The case studies made clear that it is not enough to point out problematic issues within a question. In the type of small to mid sized design and development projects which are the subject of this article, validators cannot restrict themselves to indicating which questions are not good enough. In practice, the validator is actually co-designer. Thus (s)he has to provide a handle for improvement of the question or suggest a completely different approach with respect to the learning objective. Consequently, in most case studies the validation task overlapped with the task of intermediate design. This obscures good quality control. However, a more strict separation of formal validation and actual design and development would require a larger investment and a different type of projects.

The lecturer or professor who is responsible for the course and for the corresponding assessment will have to validate questions drafted by the assistant. Alternatively, when the subject matter expert has drafted the questions, an assistant and in some cases the educational technologist can check many aspects of the question such as consistency, phrasing, choice of terminology, et cetera.

Validation can often be supported by data if the questions have been used by students in previous exams or by previous cohorts. Analysis of data often points toward 'suspect' questions. However, such analysis falls outside the scope of this article. In the budget templates below we therefore refer to 'ex ante' validation.

Revising questions

In practice, many first draft questions were revised or discarded on the basis of the validation results. Often, second drafts were made and needed to be validated again and discussed again. This process results in several versions of questions and pools of questions. The case studies showed that the teams had difficulties in managing versions of questions and keeping track of which question had what qualities.

The revision task is primarily a task for the subject matter expert and assistant. From the case studies, it became apparent that the delegation of the design task and the revision task to the assistant will always induce some waste of efforts.

Image processing

In the case studies, a considerable number of images have been used. Even though the images were already available in digital format, they still had to be processed. This involved operations like: changing the format of the image, resizing, clipping, deleting part of the image, replacing part of the image, inserting text in the image, indicating hotspots. These operations require routine with an image processing application. Some of these operations also require routine with the question editor of the Learning Management System or CBA system. Most of this work does not require subject matter knowledge and at first sight a rendering specialist would seem the most appropriate person to execute this task. However, the case studies do not provide sufficient information to arrive at a decision rule about to what extent image processing should be delegated to a rendering specialist.

Realization in CBA or learning management system

In this task, the finalized draft questions are entered into an authoring environment. This includes at least: calling up the system, initiating a new question, copying text and images into the stimulus, choices, distracters, and also formatting, layout and setting scoring rules. Furthermore, this task implies question pool management. This task requires routine with the authoring environment, file management and with picture sorting and selection tools and often still requires picture resizing operations as well. In general, this task should be delegated to an assistant or to a rendering specialist.

In the practice of the fifteen projects, it was not standard procedure to check the final version of the question in the system and to check aspects such as final lay out quality et cetera. In case it is really necessary for the subject matter expert or assistant to validate the questions on screen and to send the comments back to the rendering specialist cost savings might be negligible. Therefore, for the type of these small and mid sized projects, it is deemed better to train the rendering specialist and make this person fully responsible for the final version.

The costs for entering a validated question into an authoring environment are based on the type of question that is entered and whether media is to be included or complex scoring rules need to be entered. In order to estimate how much time this would require by someone who is very proficient with authoring tools, a benchmark set of questions was entered by three proficient persons in Blackboard, Questionmark Perception and by means of editing QTI2 conformant questions in XML. Table 3 lists the results.

In practice, the task of entering questions in an authoring environment took always much longer than the figures in Table 3 suggest. In the case study in which this task was performed by a dedicated rendering specialist, the average question entry and picture processing time of almost 25 minutes was recorded. The order of magnitude was confirmed by data from two other projects apart from the case studies.

While the time registrations in other case studies are not detailed enough in order to provide more quantitative data, many time-consuming actions related to the task of question entry were mentioned. Examples are: looking up missing details, rearranging materials, rearranging desktop settings, interpreting meta information scribbled by the subject matter expert, adjusting picture sizes, moving files around, making mistakes and repairing mistakes, previewing the question, system failures and so on and so forth.

Table 3 Benchmark test for entering a set of standardized questions into different authoring environments.

Question-type	time in minutes		
	Bb-expert	QM-expert	XML-editor
multiple choice	2	3	9
multiple choice with image	5	7	14
multiple answer	3	5	11
multiple answer with image	7	8	19
fill in blanks	4	12 [1]	17
fill in blanks (numeric) with image	3	7	22
matching	3	4	5
matching with image	6	6	13
pull down	5	20 [2]	19
pull down with image	5	8 [3]	16 [4]
ranking	2 [5]	7	8 [4]
ranking with image	4 [5]	8	30 [6]
drag and drop	5	10	24
hotspot	3	5	9
select a blank	3 [7]		15
select a blank with image	4		17

Note. Bb – Blackboard, QM – Questionmark Perception v 3.x, XML-editor – a person familiar with XML editing who edited two sets of 80 questions in QTI2 (QTI = Question and Test Interoperability).

1 – Time to enter without modifying the outcome definitions to give a score for partial correct answers: about 6 minutes 2 – On the basis of an existing question, used as template 3 – Table inserted as 1 image 4 – Implemented as select a blank 5 – Implemented as matching 6 – Implemented as drag and drop 7 – Implemented as fill in blanks

Additional CBA-related tasks

This article is based on the assumption that design and development of closed questions can be discussed as a distinct cluster of tasks. The complete process of computer-based assessment involves several other tasks. These tasks are not directly related to the actual design and development of questions. Strictly speaking, they do not fit the scope of this article. However many subject matter experts in higher education are interested in some indication of the point were computer-based exams become more cost efficient than 'traditional' exams. For this reason, also *organization of exams* (including configuring the exams and organizing exam sessions) and *processing of exam results* (psychometric test analyses and score interpretation and grade curving) have been included in the budget templates below.

Additional management and communication within the team

For projects in general, management rather than communication is usually defined as a separate task. In this article, the communication within the team is defined as a separate task because the cost for communication grows when more people are working in a project. The main factors that currently contribute to communication costs are threefold. Firstly, the fact that subject matter experts have in general few timeslots available for face to face communication. Secondly, a lack of subject matter expert-friendly support for workflow, collaborative design and version control. Finally, the challenge to optimize the workload of the rendering specialist whose capacity will be shared among different projects.

TEN SCENARIOS

In this section, we present ten possible scenarios for design and development projects of closed questions. The authors believe that these scenarios cover the various set-ups of small or mid sized projects for design and development of sets of closed questions within higher education. The scenarios are intended to support initial planning and setting a budget for the project.

Table 4 describes scenarios for projects that focus on questions for a Computer Based Assessment role. Table 5 describes the scenarios for projects that focus on questions for an Activating Learning Material role. Table 4 also supports a structure for comparing the costs of a written exam, based on open questions with the costs of a digital exam, based on closed questions. Both sets of scenarios are ordered from maximum support for the subject matter expert to minimal support for the subject matter expert. The tables provide a comprehensive overview of relevant tasks within a project for the design and development of closed questions, the allocation of these tasks and the amount of time required. Such tables have not been found in literature yet and are believed to form an important tool for anyone involved in mid sized question design and development projects.

Both templates assume that a project is set up to design and develop a pool of about 300 questions. For the CBA role, this can reflect the design of 60 clusters of 5 equivalent questions. The tables highlight the cost structure and the structural consequences of reallocation of tasks. The time values in the table are estimates based on the time registrations in the fifteen projects. However, the reader can easily insert other values for certain parameters. Some of the scenarios imply independent choices, for instance, the percentage of questions that will include a picture or the amount of training to be provided for the assistant. Parts of the data are contextual data depending on the institution and often also on the country where the institution resides. The costs/hour of a subject matter expert vary widely across different countries in the world and so do the costs of the other specialists. Another example of an estimate that may vary widely for different projects is the ratio of the time for question entry needed by a subject matter expert and the time needed for this task by a rendering specialist. In the tables, this parameter is set to 1.5.

Apart from these project specific parameters, the last column in

Table 1 contains the average Design and Development time per developed question. This value is based on an analysis of the time sheets of every employee in each of the projects that were used for the case studies. The overall conclusion was that average design and development times were up to 2 hours. Based on experience in the case studies it is believed that in a budget for a design and development project, this time should not be set lower than two hours per question for projects. Based on experience in the case studies this average overall time is divided over different subtasks. Notice that the difference in the time between questions for the ALM role and questions for the CBA role is mainly due to the necessity to provide feedback in the former.

The budget examples presented in table 4 and 5 make clear how cost efficiency gains might be realized by reallocation of tasks. For instance, with the current settings of parameters and values the budget templates suggest that the average design and development time without support will be relatively low. However, for many institutions it is likely that the costs will be higher. The actual efficiency gains for any institution can only be determined by inserting the actual data in the cells.

CONCLUSIONS AND DISCUSSION

From fifteen small to midsized projects on design and development of innovative digital closed questions for natural and engineering sciences in higher education quantitative and qualitative data were collected. Analysis of these data from the shared perspective of computer-based assessment and activating learning materials led to a practical task structure for such projects. For a number of these tasks this analysis has led to practical advice, which has been described in the respective paragraphs.

Based on the case studies the options to delegate tasks to an assistant of the subject matter expert, to an educational technologist and to a rendering specialist have been described. For defining, planning and budgeting such projects good estimates for an average design and development effort of closed questions, typical for a university context, are important. However, such estimates could not be found in literature. Communication with colleagues in higher education as well as some initial experiments always seemed to point to 'about half an hour per question' as a good estimate. Time registrations within the projects have resulted in more empirical cost estimates for some of the tasks and the average total design and development time per question. On average, the latter was close to two hours per question.

Based on reports produced within the projects, sources of inefficiency were identified and a number of 'does and don'ts' are formulated. It is concluded that efficiency improvements, which are mainly based on division of labor, tend to increase the need for communication between the subject matter expert and the other members of the team. Realizing efficiency gains requires adequate control of this communication process. It is suggested that an educational technologist takes the specific responsibility

Table 4 Five scenarios and corresponding budgets for the development of 300 questions for the CBA role

	Max Support									Max Support No RS							Max Support only ET					Max Support only ASME					No Support		
Design&Development of Pool of 300 questions and execution of 4 exams	**SME**		**ET**		**ASME**		**RS**		**Total**	**SME**		**ET**		**ASME**		**Total**	**SME**		**ET**		**Total**	**SME**		**ASME**		**Total**	**SME**		**Total**
	hr	**€**	**hr**	**€**	**hr**	**€**	**hr**	**€**	**€**	**hr**	**€**	**hr**	**€**	**hr**	**€**	**€**	**hr**	**€**	**hr**	**€**	**€**	**hr**	**€**	**hr**	**€**	**€**	**hr**	**€**	**€**
1. making project plan, defining budget	1	90	14	980					1070	1	90	14	980			1070	1	90	14	980	1070	4	360	20	1000	1360	4	360	360
2. set-up team/allocate people to tasks	1	90	4	280	1	50	1	30	450	1	90	4	280	1	50	420	1	90	4	280	370	4	360			360			
2a.setting up communication in the team	2	180	2	140	2	100	2	60	480	2	180	2	140	2	100	420	2	180	2	140	320	2	180	2	100	280			
2b. training	4	360	16	1120	24	1200			2680	4	360	16	1120	24	1200	2680	4	360	16	1120	1480	4	360	24	1200	1560	8	720	720
3. labeling clusters of five	4	360			16	800			1160	4	360			16	800	1160	8	720	8	560	1280	8	720	2	100	820	12	1080	1080
4. design/intermediate representation 1)					150	7500			7500					150	7500	7500	120	10800	30	2100	12900	30	2700	120	6000	8700	120	10800	10800
5. ex ante validation 2)	50	4500							4500	50	4500					4500	50	4500			4500	50	4500			4500	50	4500	4500
6. improving and/or replacing questions			25	1750	25	1250	25	750	3750			25	1750	50	2500	4250			75	5250	5250			75	3750	3750	25	2250	2250
7. image processing							75	2250	2250					75	3750	3750			125	8750	8750			125	6250	6250	150	13500	13500
8. entering in CBA system							100	3000	3000					120	6000	6000			130	9100	9100			120	6000	6000	150	13500	13500
9. definition of 4 exams & 1 trial exam	8	720			16	800			1520	8	720			16	800	1520	8	720	16	1120	1840	8	720	16	800	1520	12	1080	1080
10. organization/execution of 4 exams			8	560	64	3200			3760			8	560	64	3200	3760			64	4480	4480	4	360	80	4000	4360	64	5760	5760
11. processing of results of 4 exams 3)	2	180	12	840	12	600			1620	2	180	2	140	12	600	920	2	180	12	840	1020	2	180	16	800	980	8	720	720
12. additional communication within team	8	720	4	280	8	400	8	240	1640	8	720	4	280	4	200	1200	8	720	4	280	1000	8	720	4	200	920			
characterizing (4 + 1) ***closed*** question based exams for 100 students							**hr**		**€**					**hr**		**€**			**hr**		**€**			**hr**		**€**	**hr**		**€**
total budget costs									**35380**							**39150**					**53360**					**41360**			**54270**
total costs per student per exam 4)									88							98					133					103			136
D&D time per question 5)							1.9							1.9					2.0					2.0			1.7		
D&D costs per question 5)									95							110					153					115			156
characterizing (4 + 1) ***open*** question based exams for 100 students							**hr**		**€**																				
D&D costs of (4 + 1) exams (2 hr/exam)							10		900																				
D&D costs per student per exam 4)									2.25																				
manual scoring and marking per student							1		90.00																				
total costs per student per exam									9225																				
total budget costs 4)									**36900**																				

1) e.g. MS Word
2) of intermediate representations/ in case SME makes questions + validation by a colleague
3) independent of number of students
4) excluding trial exam
5) excluding task 9,10 and 11

Table 5 Five scenarios and corresponding budgets for the development of 300 questions for the ALM role

	Max Support									Max Support No RS							Max Support only ET					Max Support only ASME					No Support		
Design&Development of Pool of 300 questions for ALM role	**SME**		**ET**		**ASME**		**RS**		**Total**	**SME**		**ET**		**ASME**		**Total**	**SME**		**ET**		**Total**	**SME**		**ASME**		**Total**	**SME**		**Total**
	hr	**€**	**hr**	**€**	**hr**	**€**	**hr**	**€**	**€**	**hr**	**€**	**hr**	**€**	**hr**	**€**	**€**	**hr**	**€**	**hr**	**€**	**€**	**hr**	**€**	**hr**	**€**	**€**	**hr**	**€**	**€**
1. making project plan, defining budget	1	90	14	980					1070	1	90	14	980			1070	1	90	14	980	1070	4	360	14	700	1060	4	360	360
2. setup team/allocate people to tasks	1	90	4	280	1	50	1	30	450	1	90	4	280	1	50	420	1	90	4	280	370	2	180	4	200	380	1	90	90
2a.setting up communication the team	2	180	2	140	2	100	2	60	480	2	180	2	140	2	100	420	2	180	2	140	320	2	180	8	400	580	2	180	180
2b. training	2	180	4	280	8	400			860	2	180	4	280	8	400	860	2	180	4	280	460	2	180	8	400	580	2	180	180
3. matching of objectives and questions	4	360			16	800			1160	4	360			16	800	1160	8	720	4	280	1000	8	720	4	200	920	12	1080	1080
4. design/intermediate representation*					150	7500			7500					150	7500	7500	120	10800	30	2100	12900	50	4500	100	5000	9500	125	11250	11250
4a. authoring presentational feedback					75	3750			3750					75	3750	3750	40	3600			3600			75	3750	3750	40	3600	3600
4b. authoring interactive feedback**									PM							PM					PM					PM			PM
5. ex ante validation***	50	4500							4500	50	4500					4500	50	4500			4500	50	4500			4500	50	4500	4500
6. improving and/or replacing questions			25	1750	25	1250	25	750	3750			25	1750	50	2500	4250			75	5250	5250			75	3750	3750	25	2250	2250
7. image processing							75	2250	2250					75	3750	3750			125	8750	8750			125	6250	6250	150	13500	13500
entering in CBA system							100	3000	3000					120	6000	6000			130	9100	9100			120	6000	6000	150	13500	13500
9. providing access to students					16	800			800					16	800	800			16	1120	1120			16	800	800	20	1800	1800
12. additional communication within team	8	720	4	280	8	400	8	240	1640	8	720	4	280	8	400	1400	8	720	4	280	1000	8	720	4	200	920			
total budget costs									**31210**							**35880**					**49440**					**38990**			**52290**
D&D time per question (hr)							2.1							2.1					2.1					2.2			1.9		
D&D costs per question (€)									101							117					161					127			168

Note * e.g. in natural language in MS Word, ** in the ALTB project no data about authoring interactive feedback have been collected, *** of the intermediate representations to support and manage this process. In addition, the need for subject matter expert-friendly computer-based support of workflow management, version control and collaborative design was identified.

In order to support planning and budgeting of future projects, two sets of reference scenarios and budget templates for mid sized design and development projects have been developed. The reference scenarios and corresponding budget templates cover the most likely practical contexts for such projects and highlight for which tasks efficiency gains might be realized and what consequences of labor division are possible.

The scenarios presented in this paper highlight that design and development of digital closed questions for different roles ranging from the role of activating learning material to the role of questions for computer-based assessment have a number of aspects and tasks in common. Clearly, the design and development of complete assessments using innovative digital closed questions involves a need for deep knowledge and understanding of educational assessment theory. However, experiences in the projects showed that when detailed educational measurement knowledge needs to be acquired during project, it can lead to a frustration and waste of effort. In the program on which this article is based, expertise on educational assessment was clustered in a special project within the program. This project falls outside the scope of this article.

Experience in the fifteen projects suggests that educational assessment expertise that goes beyond the expertise that can be expected of an educational technologist concerns primarily two forms of experience. The first form implies understanding the possibilities and limitations for assessment of innovative question types that are available in the learning management system or computer-based assessment system at hand. This implies knowledge of theory of educational assessment combined with detailed knowledge of the system used for educational measurement. The second form implies all knowledge that is directly related to complete assessments. The educational technologist often lacks these two forms of knowledge. This will make it necessary to involve an educational assessment expert.

Subject matter experts and assistants with subject matter knowledge need training with respect to design of digital closed questions for both roles of questions, the role to function as activating learning material and the role to function within computer-based assessment. Therefore, the next step is to develop a workshop for subject matter experts and assistants with subject matter knowledge. Initial experience with the design patterns developed in the case studies suggests that these design patterns might form the core of the training material for assistants.

REFERENCES

Blackboard. (2006). Blackboard. Retrieved march 04 2006, 2006, from http://www.blackboard.com/

Bull, J., & McKenna, C. (2004). *Blueprint for computer-assisted assessment*. London: RoutledgeFalmer.

Fenrich, P. (2005). *Creating instructional multimedia solutions: Practical guidelines for the real world*. Santa Rosa, California: Informing Science Press.

Frary, R. B. (1995). More multiple-choice item writing do's and don'ts. *Practical Assessment, Research and Evaluation*(4), 11.

Haladyna, T. M. (1997). *Writing test items to evaluate higher order thinking*. Needham Heights: Allyn & Bacon.

Haladyna, T. M. (2004). *Developing and validating multiple-choice test items* (Third Edition ed.). London: Lawrence Erlbaum Associates.

Hartog, R. J. M. (Ed.). (2007). *Design and development of digital closed questions: A methodology for midsized projects in higher education*. Utrecht: SURF Foundation.

Kehoe, J. (1995a). Basic item analysis for multiple choice tests. *Practical Assessment, Research and Evaluation, 4*(10).

Kehoe, J. (1995b). Writing multiple-choice test items. *Practical Assessment, Research and Evaluation, 4*(9).

Keller, J. M. (1983). *Development and use of the arcs model of motivational design*. Enschede: Technische Hogeschool Twente.

Merriënboer, J. J. G. v. (1997). *Training complex cognitive skills: A four-component instructional design model for technical training*. Englewood Cliffs, NJ: Educational Technology Publications.

Questionmark. (2002). Questionmark. Retrieved 19 jan 2007, 2007, from http://www.questionmark.com/

Richey, R. C., Klein, J. D., & Nelson, W. A. (2004). Developmental research: Studies of instructional design and development. In D. H. Jonassen (Ed.), *Handbook of research on educational communications and technology* (2e ed., pp. 1099 - 1130). Mahwah NJ: L. Erlbaum Associates.

Scalise, K., & Gifford, B. (2006). Computer-based assessment in e-learning: A framework for constructing "Intermediate constraint" Questions and tasks for technology platforms. *The Journal of Technology, Learning and Assessment., 4*(6).

Smith, P. L., & Ragan, T. J. (1993). *Instructional design*. USA: Macmillan Publishing Company.

ThreeShips. (2007). N@tschool. Retrieved 8 februari 2007, 2007, from www.threeships.nl

.

Acknowledgements

The authors would like to thank Kees Oomen for his support in data collection and analysis.

This work has been made possible with support of the SURF Foundation. The SURF Foundation is the higher education and research partnership organization for network services and information and communications technology (ICT) in the Netherlands. For more information about SURF Foundation: http://www.surffoundation.nl/

Citation

Hartog, Rob, Draaijer, Silvester, & Rietveld, Luuk (2008). Practical Aspects of Task Allocation in Design and Development of Digital Closed Questions in Higher Education. *Practical Assessment Research & Evaluation*, 13(2). Available online: http://pareonline.net/getvn.asp?v=13&n=2

Authors

Rob Hartog [rob.hartog [at] wur.nl]
Wageningen University, Wageningen, The Netherlands

Silvester Draaijer [s.draaijer [at] ond.vu.nl]
Vrije Universiteit, Amsterdam, The Netherlands

Luuk Rietveld [L.C.Rietveld [at] TUDelft.nl]
Delft University of Technology. Delft, The Netherlands