The docimology research showed that the standardized testing is an efficient evaluation form of the learners’ competencies provided to meet qualities of validity, reliability, sensitivity, diagnosticity, equity, practicability, communicability and authenticity. When standardized tests meet these qualities, they are appreciated by trainers, learners and by heads of training institutions. Currently, the construction and the quality management of standardized qualifying evaluations is a worry for the higher/university education in many European countries which have to face a still growing number of students. The training professionals and the human resources managers of companies are often constrained to evaluate staff competencies acquired outside the schools. Moreover, these concerns, regarding the quality of the tests, are also in relation with a worldwide movement of the introduction of quality management in education and training activities. Our research aims at facilitating the creation of standardized tests to make them more reliable, more efficient and quicker to set up, in order to satisfy at best the actual requirements of the education/training/human resources world. Our objective is to create an electronic platform of construction and quality control in standardized testing entitled “e-C&QCST (electronic Construction & Quality Control in Standardized Testing) build in an “Open Source” policy for a web based utilisation.

**Keywords:** Assessment; Standardized testing; Computer application

The standardized testing is an evaluation form of the learners’ competencies which - when it meets the qualities of validity, reliability, sensitivity, diagnosticity, equity, practicability, communicability and authenticity (Gilles & Leclercq, 1995; Castaigne, Gilles & Hansen, 2001; Gilles, 2002; Gilles et Lovinfosse, 2004) – is appreciated by the trainers, learners and the heads of training institutions.

The preoccupations concerning the quality of standardized evaluations of learners’ competencies fit into a movement which has been reaching the Belgian and European institutions in the early 90’ (Kaufman & al., 1995; Segers & Dochy, 1996). This movement itself is part of a worldwide trend of introducing the quality management in the education and training activities spheres (Ramsden, 1991; Nightingale & O’Neil, 1994; Zink & Schmidt, 1995).
The team of the unit entitled “Système Méthodologique d’Aide à la Réalisation de Test (SMART)” of the University of Liège helps academic teachers and trainers of private companies to construct assessments tools for measuring students’ or trainees’ competencies. The SMART proposes a quality approach for the development of standardized tests for educational or training purposes. This quality approach is based on four fundamental principles which to be taken into account in the creation of evaluations (Gilles, 2002). This approach is inspired by the recommendations of the international standard ISO 9004 2 (Quality management and quality system elements - Guidelines for the services). In the SMART context of docimologic engineering, the quality approach suggested allows to focus the actions on four types of concerns to be undertaken in the creation of learner’s competencies tests. At the heart of our approach, the central principle is related to the needs of the developers, of the teachers/trainers, of the students/trainees and of the persons in charge of the evaluation tools which have to be conceived and developed. It is the central pole of our approach and also the point of convergence of three other poles: the pole of the models used as references for the development of the assessment tools, comprising the model entitled CCQCST “Cycle of Construction and Quality Control for Standardized Testing” (briefly exposed later); the pole of human and material resources to mobilize and the pole of the quality standards which have to guide us in our actions.

**Figure 1.** SMART’s quality approach in standardized testing construction

In this context, our research aims at setting up methods, procedures and tools for the quality management of standardized tests. The spreading and the worldwide utilisation of our research results are possible today thanks to interfaces integrated in the software we propose. They are indeed created to work in a web based environment. This is the idea at the base of our research and development project entitled “electronic Construction and Quality Control...”
in Standardized Testing (e-C&QCST). The e-C&QCST platform is developed in an “Open Source” policy.

Our contacts with private institutions within the context of our research confirm our hypothesis of a large market which opens for the e-C&QCST platform and for the service companies which would ensure a coaching and a follow up of the users. Furthermore the idea of a quality label for standardized testing (Gilles, 1998) will be proposed in the perspective of a qualitative docimology in adequacy with the trainers’ and trainees’ needs. This "labellisation" perspective seems to meet some expectations in the training world. In this economic valorisation perspective of the e-C&QCST research, we propose to structure the supply of services in the following way:

[1] Audit of the standardized testing systems of a client company/organization to measure the adequacy of the human and material resources, the needs of the client, his objectives and the expected results;

[2] Training and coaching of the users of the e-C&QCST platform within their company/organization context in order to facilitate the appropriation and integration of the platform in the training and evaluation system;

[3] Docimological services for the improvement of the standardized tests of a client company/organization. These kinds of services should be facilitated by the use of the e-C&QCST platform which permits to a person in charge of the development of a standardized test to give partial access to subcontractors for precise tasks;


Method

The modules of the e-C&QCST platform

The idea is to formalize, adapt or create docimological procedures in order to control and quantify the quality of standardized tests of learners'/trainees’ competencies. For this purpose, we will use the Cycle of Construction and Quality Control in Standardized Tests (CCQCST) model that we have proposed (Gilles et Leclercq, 1995; Gilles, 2002). The 8 realisation steps of this model will be programmed in the web based e-C&QCST platform. This platform is being created with the free programming language PHP and the free database manager MySQL. We also use the free server Apache.

Since the middle of the nineties, this eight steps have been constantly improved by the SMART team.
The 8 steps of the CCQCST model related to the 8 modules of the e-C&QCST platform are:

[1] ANALYSIS: The course/training or competencies description must be analysed in order to list out the points to evaluate (PE), their priorities (P) and the categories of performances (CP) of the learners/trainees. The end product consists of a test specification table where the points to evaluate and their priorities are crossed with the performances categories (PE x CP).

[2] DESIGN: In the second step of construction the developer must design the multiple choice questions modalities and make decisions about the importance of each modality. The end products are of two types. On one hand, there must be an association of categories of performances of the previous step with the modalities of the questions (CP x MQ) defined at this step (i.e. a MQ could be a type of MCQs with confidence degrees). On the other hand, the binomials (PE x CP) and (CP x MQ) must be combined to obtain trinomial ones (PE x CP x MQ). Finally, from the list of the trinomials objects the ones which should be linked to questions during the test must be selected.

[3] QUESTIONS: In the third step, the test developer selects or writes the questions which are linked to the trinomials objects selected for the test at the former step of the construction. The questions are created or selected in a secured item bank database.

[4] TRAINING: At the fourth step, the person in charge of the development of a test can choose to propose “blank tests” (paper-pencil or web based) in order to train the evaluated persons on the modalities of the questions previously defined.

[5] TESTING: In the fifth step, the questions created or selected at the third step are put in the standardized test which is proposed to the people who perform the test (paper-pencil or web based). In certain cases, it is possible to perform online exams - in a secured environment - or to link the questionnaires to automated responses acquisition devices such as optical marks reading.

[6] CORRECTION: At the sixth step, the standardized test developer carries out the correction of the tests. Treatments are designed to check the quality of the questions before the final version of the results of the evaluated persons is validated in one form or another.

[7] FEEDBACK: Personalized and tailored feedbacks with for example online access are proposed on the basis of the final version of the results proposed at the previous step.

[8] REGULATION: This is the eighth and last step of the construction of the standardized test. Here, the test developer collects and analyses the opinions of the all the people who intervened in the test construction process. The aim is to improve the next standardized test which will be created. This explains and justifies the C of the term “Cycle” in the acronym “CCQCST” of our model.
For each construction step defined above, quality control procedures are also considered within the framework of the research and development of the e-C&QCST platform.

**Quality criteria for standardized testing construction**

The docimological procedures used in the 8 modules of the e-C&QCST platform are built in reference to a series of quality criteria that are used by the SMART team in standardized testing construction (Gilles, 1998, 2002; Gilles et Lovinfosse 2004):

- **Validity**: candidates’ scores must reflect what the trainer/teacher wants to measure. The idea is to ensure solid inferences with the categories of performances (CP, construct or theoretical validity) and to cover the important aspects of the learning/training, what we called the points to be evaluated (PE, content validity);
- **Fidelity**: an evaluated person must essentially obtain the same score if his/her copy is corrected by different persons in charge of the correction (inter-judge corroborating) and/or at another moment by the same person or the same team (intra-judge consistency);
- **Sensibility**: the candidate’s competencies measures must be precise and reflect subtle phenomena;
- **Dagnosticity**: tests results must ensure a precise diagnostic of the learning difficulties (and ideally of their causes) and of the competencies which are mastered and those which are not;
- **Practicability**: the feasibility in terms of time, human and material resources must be ensured;
- **Equity**: each candidate must be treated in a fair way, as far as possible in the same way;
- **Communicability**: non confidential information concerning the process progress must be communicated and understood by the people who perform the test;
- **Authenticity**: the questions must be in relation with situations which make sense for the people who perform the test, relevant in comparison with the real world and able to enhance transfers in contexts and practices that can be met in life.

**Roles’ distribution in the research and development team**

The SMART team, who is the promoter of this e-C&QCST project, joined for the needs of this research the team of Management of Competencies of the High School of Management at the University of Liege. The General Direction of Technologies, Research and Energy of the Walloon area of Belgium supported financially this R&D project since the end 2003 for a period of three years.

The roles in the research group are distributed in the following way. The SMART team is divided into two subgroups. The first subgroup - two part time scientists: one specialized in docimology (he coordinates the research) Gilles, J.-L., Piette, S.-A., Detroz, P., Tinnirello, S., Pirson, M., Dabo, M. et Lê, H. (2005). The electronic Construction and Quality Control in Standardized Testing platform project (e-C&QCST), European Association For Research on Learning and Instruction (EARLI) 11th Biennial Conference of EARLI. Nicosia: University of Cyprus
and the other in ergonomics - conceives and formalizes the docimological procedures used within the framework of the CCQCST model and writes the project specifications of the e-C&QCST platform. The second subgroup (three data processing specialists, two full-time and one part time) programs the project specifications of the eight modules, starting with the first and finishing by the 8th. The Management of Competencies team (one half-time economist and a 10 percent full-time equivalent scientist specialized in human resources management) organizes the contacts with the companies/organizations which are partners in this project. In particular this Management of Competencies team submits the scenarios of the docimological procedures to the companies/organizations in order to validate the procedures and accompanies the partners during the experimentations of the eight modules. These tests of the modules in the world of the companies as well as other tests realized in the academic world bring readjustments of the project specifications. They are discussed within the research group.

**Research and development hypotheses**

This project covers various conceptual frameworks to which we attach research and development hypotheses. We identified eight conceptual frameworks and propose these principal hypotheses:

(A) Docimological framework: the quality of standardized tests could be optimized by combining different tests construction procedures at the eight steps of the CCQCST model which will make it possible to guarantee a high quality level at the eight criteria we have defined (validity, fidelity, sensitivity, diagnosticity, feasibility, equity, communicability and authenticity). The requirements in standardized testing of the companies’ world are sufficiently close to those of the academic world so that it is possible to create an e-C&QCST platform which can be appropriate in both cases.

(B) Technological framework: an electronic platform covering all 8 steps of the CCQCST model helps to improve the quality when assessment developers create standardized tests.

(C) Framework of ergonomic aspects: it is possible to develop an electronic platform which makes it possible to implement the Cycle of Construction and Quality Control in Standardized Tests (CCQCST) model in an intuitive way.

(D) Framework of quality management: quality controls are possible in link with our quality criteria at the different steps of the CCQCST model. The use of these quality controls increases the quality of the standardized tests. These quality controls could be used in a quality label policy.

(E) Economical valorisation framework: there is a market for subcontracting services in the field of an economical valorisation of the e-C&QCST platform. There is also a market for the development of a quality label related to the standardized testing.

(F) Users framework: it is possible to develop a community of users of the e-C&QCST platform who could develop interactions by sharing practices, by exchanging tests... Those interactions could improve the...
competencies in quality standardized tests construction of tests developers. The use of the e-C&QCST platform by companies/organizations changes the vision of the standardized testing and contributed to the emergence of the complementarities and the responsibilities of the roles of the persons who are involved in test creation.

(G) Linguistic framework: a broad diffusion of the e-C&QCST platform implies a translation of the interfaces in most languages of the European Union.

(H) Experimentations framework: a collaborative design with a sample of companies/organizations, future potential users, implied in each construction step (the eight modules of the e-C&QCST platform), makes it possible to check the adequacy of the procedures suggested with the users’ needs. Collaborative design makes it also possible to consider different uses from those initially envisaged and to progressively test the software ergonomics while the product is constructed.

In this article, we will discuss first results concerning the hypotheses from the frameworks A, C, E and F. At this stage, at middle of the research, it is still too early to learn all the useful lessons concerning the others frameworks hypotheses.

**First results**

*Companies and organizations collaborating in this research*

As we saw above, the SMART created the CCQCST model and made use of it since middle of the nineties within the framework of the help provided to higher education teachers. Former studies showed the effectiveness of this model in the academic context (Gilles & Leclercq, 1995; Castaigne & Gilles, 2000; Gilles, 2002; Gilles & Lovinfosse, 2004). But this model was used by few trainers of private companies and organizations. This is the reason while early in this project we proposed collaborations to private companies or organizations.

We contacted the EPSILON association to help us in this task. This association proposes to all the private companies or organizations of Belgium to share their practices in relation with the development of the competencies of their collaborators. With the help of the EPSILON staff, we defined sample characteristics of the companies/organizations to be approached in order to propose to them to join the e-C&QCST project partnership. The starting idea was to meet people in charge of standardized tests in widely different companies and organizations. We knew that very small companies or organizations had few chances of being potential users of the e-C&QCST platform (excepting those whose trade is in training or in testing).

The characteristics of the companies and organizations selected to create the partnership in this R&D project were related to:
- The type of company (public, private, mixed);
- The size of the company or the organization and its network (membership or not of a group, great number of subcontractors...);
- The existence of an internal training service;
- The stakes of the evaluations (attesting, recruitment, mobility, career management...).

At the end of the selection process, we started collaborations with ten companies and organizations. These are:
- The FORTIS Bank;
- The AXA Bank and insurance group;
- The Training Institute of the Belgian’s Federal Administrations (IFA);
- The Pizza Belgium company;
- The communal association of economic development of the Belgian province of Luxembourg (IDELUX);
- The FOREM (Belgian community and regional office of the vocational training and employment);
- The PHONE LANGUAGE company;
- The TNT Express Worldwide Euro HUB;
- The UNIQUE interim company;
- The “Consortium de validation des competences”, a consortium of the Belgian french speaking community, the Walloon area and the Brussels area (Its mission is to create centers of validation of competencies in connection with the continuous vocational training)

**Interest in the CCQCST model and its docimological procedures**

One of the first activities proposed to our partners consisted in presenting in details the 8 steps of the cycle of construction and quality control for standardized testing. Many partners expressed a great interest for the procedures we proposed. Four partners (AXA, FOREM, FORTIS Bank and IFA) requested immediate implementations of our docimologic procedures.

These requests for immediate docimological subcontracting services were carried out by members of the SMART team in parallel with the R&D project activities. This situation gives us the chance to try in the companies’ context the standardized testing construction procedures initially created for the academic world.

Among the types of docimological services requested by the four companies and organizations, we mention:
- Training the teams in charge of the construction of standardized tests to the CCQCST model and its procedures for all four companies and organizations;
proposed at the 3rd step entitled “QUESTIONS” of the CCQCST model (for AXA and IFA);
- Accompaniment and follow-up of the trainers in charge of the realization of standardized tests which attest the trainees’ competencies (for IFA). These kind of activities are proposed at the 1st step “ANALYSIS”, the 2nd step “DESIGN”, the 3rd step “QUESTIONS”, the 4th step “TEST” and the 5th step “CORRECTION” of the CCQCST model.
- Reliability analysis of the results obtained at standardized tests. Proposals for an improvement of the reliability on the basis of the improvement of the MCQs and their solutions (for FOREM). This kind of activity concerns the 6th step “CORRECTION” of the CCQCST model.

On the one hand, these collaborations permit life tests of our docimological procedures. On the other hand, they give us the opportunity to try and improve an offer of docimological subcontracting services as a private company which valorises research results would have done it.

**Relevance of the CCQCST model in the partners’ context**

As we have just said, as soon as we explained the details of the CCQCST model to the companies and organizations which are partners in this research, four of them requested immediate docimologic services. The tasks requested were in direct connection with the CCQCST model we use for the development of the e-C&QCST platform. This experience provided us information about the relevance of our model in contexts of companies. It appears that our model is well appreciated.

We also interviewed persons in charge of creating standardized tests in the ten participating companies and organizations. They were asked to describe the actions that they realized during the construction of their standardized tests. A series of complementary questions were asked in order to determine certain aspects of the process, such as: collaborators concerned, temporal dimension, importance of the quality of the tests, insertion of the tests in other processes...

The documents created within the framework of the subcontracting docimological services offered by the SMART and the result of the interviews were discussed within the research group in order to analyze eventual readjustments of specifications of our model. It comes out that the CCQCSTS model and its procedures which were initially realized on the basis of docimologic researches and from former experiments with higher education teachers is relevant and well accepted by our partners coming from the private companies and organizations world. The model requires only small readjustments that we presented for validation to the ten private companies or organizations at a workshop meeting. The principal readjustment consisted of a clarification of the roles of the standardized tests developers. The difficulty lays in the fact that this definition of the developers’ roles had to be extended so as to be used in very different contexts and that their roles should be

assigned to a series of collaborators of one person in charge for the
development of the standardized test (the head of the project) within the
e-C&QCST platform. In other words, a standardized test construction is more
often considered as a team policy in the world of the companies than in the
academic world.

Users’ roles within the e-C&QCST platform

We propose seven roles within the e-C&QCST platform. Some of them can
be distributed to one or more people. The first role is just a technical support
for the installation of the software but the six other roles are directly
concerned with the development of the standardized tests. The seven roles are:
- Role 1: “INSTALLER”. The installer installs the e-C&QCST software on a
server and after the installation, defines an administrator of the platform.
- Role 2: “ADMINISTRATOR”. The administrator encodes the persons in
charge of the test construction.
- Role 3: “TEST SUPERVISOR”. The test supervisor defines and encodes
the parameters of the platform. He encodes the tests developers, the
collaborators or the subcontractors for the tasks related to the construction
of a standardized test. The test supervisor has a right of access to all
information and can directly intervene on all the subcontracted actions. He
can also define a person who must give his opinion about the quality of the
productions and deliver a quality label if the level of quality is sufficient.
The test supervisor must also define the public of evaluated persons. In an
academic context, the test supervisor is generally the teacher in charge of
the assessment of his students.
- Role 4: “TEST DEVELOPER”. One or more tasks related to the
construction of the standardized test for every step of the CCQCST model
are distributed to one or more developers by the test supervisor. The test
developer can be a collaborator inside the test supervisor’s company or
institution.
- Role 5: “QUALITY SUPERVISOR”. The quality supervisor has a right of
glance without direct intervention on the productions of which he examines
the quality. The results of the examination of the quality of the developers’
productions are put into the system. The quality supervisor can subcontract
all his tasks or a part of them. The results are accessible to the test
supervisor who can decide to make this information known to the
collaborators and/or to the subcontractors and/or to the persons who are
evaluated.
- Role 6: “QUALITY CONTROLLER”. One or more tasks related to the
attribution of a quality label of the evaluation can be attributed to one or
more persons by the quality supervisor.
- Role 7: “ASSESSEES”. They are the persons who are evaluated (i.e.
students or trainees). The test supervisor can give them partial access to the
information pertaining to some parts of the eight steps of the construction.

The electronic Construction and Quality Control in Standardized Testing platform project
(e-C&QCST), European Association For Research on Learning and Instruction (EARLI)
11th Biennial Conference of EARLI. Nicosia: University of Cyprus
These hypotheses concerning the roles related to the framework of the e-C&QCST users (G framework) will be put to test within the experimentation of the first modules which is presently beginning.

**The SAFES project**

The immediate docimologic interventions requested by four companies and organizations of this project constitute an element in favour of the hypothesis of the existence of a market for subcontracting services in connection with the use of the e-C&QCST platform, an hypothesis of the economical valorisation framework (E).

Currently, the Walloon area encourages the economic valorisation of results of researches by supporting the creation of spin-off companies. Being given the success of the subcontracting service requested by four companies and organizations of this project, the SMART team filed in a project within the framework of program FIRST SPIN-OFF of the DGTRE. Our project entitled “Smart Advices For Evaluation Systems (SAFES)” (Gilles & Tinnirello, 2005) was well ranked by an independent scientific committee and accepted so that we are currently starting a two years feasibility study concerning a project of a spin-off company which would propose its services in the realization of assessments and standardized tests. At the end the two years, the creation of this spin-off company should coincide with the finalisation of the e-C&QCST platform so that this tool could be used by the spin-off team.

The SAFES project will allow us:
- To create a method of audit of the systems of testing;
- To conceive the procedures to be implemented within the framework of a quality control mark of the standardized tests;
- To define a typology of services which will be proposed to the future customers of the spin-off company;
- To realize a market research based on the methodologies of desk research and of field research.

**First interfaces of the e-C&QCST platform**

The general ergonomics of the platform is as follows:

We are actually testing with each of our partners the ergonomic choices we made for the interfaces. The first results concerning the intuitive utilisation of the platform are positive (hypothesis from the ergonomic framework [C]).

A positive effect of the meetings with the partners concerning the project specifications was to bring a clarification of the management of the collaborators and subcontractors intervening in the creation of the tests. We introduced the possibility for the test supervisor to create entries in the platform for various developers with reading or/and writing rights for the construction steps of a standardized test (hypothesis concerning the user’s framework [F]).

---

Figure 2 shows the interface which enables the person in charge of the standardized test to define roles and rights in reading and writing within the framework of a series of tasks envisaged at the analysis module. In the example, the person in charge of the test defined a collaborator who is a member of the developers team and a data processing specialist (so called “informaticien” on the screen). The person in charge of the test granted him rights of reading and of writing for the task “List PE” which consists in drawing up the list of the points which are likely to be evaluated during the test (for example the contents list of a computer course). On the screen we can see that another person intervenes within the framework of the docimologic quality label. She must here control and check the quality of the work which is to be completed by the service provider in charge of creating the list of PE. She has only the right in reading and not in writing on the productions “List PE”. We mentioned this short example explicitly to fix the ideas about some possibilities to define the roles. Entering other persons in other categories of roles for other tasks is of course possible. Our hypothesis is that this kind of functionalities should contribute to a better coordination and visualisation of the standardized tests developers’ and user’s roles.

Actually, the first module entitled “ANALYSIS”, concerning the course/training or competencies description analysis, is finished. The productions of this module consist in a list of the points to evaluate (PE), of their priorities (P) and of the definitions of the categories of performances (CP) of the learners/trainees. The figure 4 presents an example of the end product which is represented by a test specification table where the points to
evaluate and their priorities are crossed with the performances categories (PE x CP).

**Figure 4.** End product of the 1st module ANALYSE: specification table where the points to be evaluated (PE) are crossed with the categories of performance (CP).

### Conclusions

We have defined six phases in the research and development planning of our 38 months project starting November 2003 and which is planned to end in December 2006:

- Phase 1: user’s needs analyses, research frameworks, hypotheses and objectives definition, docimological procedures creation (months 1 to 8);
- Phase 2: conceptualization and preparatory project of each module of the e-C&QCST platform (months 5 to 17);
- Phase 3: writing of the project specifications for each of the 8 modules of the e-C&QCST platform, prototypes validation (month 8 to 32);
- Phase 4: experimentation in real size conditions of the platform modules (months 22 to 33);
- Phase 5: analyses of the experimentation results, improvement of the modules, final version creation and translation of the interfaces in the principal languages of the European Union (months 27 to 36);
- Phase 6: writing of the final report and spreading of the final results (months 35 to 38).

Actually we are in the middle of the 3rd phase of the project specifications and at the beginning of the 4th phase of the experimentation of the modules. The 1st module is being experimented with our research partners in academics’ and companies’ contexts.

The first lessons we can draw at this stage of the project, concerns the relevance of our CCQCST model, the users roles, the ergonomics of the platform and the economical valorisations perspectives of the results of this R&D project.

Concerning the relevance of the CCQCST model, we can say that it is well appreciated in the context of the ten companies and organisations which are our partners in this project.

Concerning the users’ roles, the seven categories we defined were validated by our partners during a workshop. This clarification should contribute to highlight the complementarities and the responsibilities of the persons who are involved in standardized test creation.

Actually, we experiment the first module entitled “analysis” within the contexts of the ten companies and organization. The first results show that the ergonomics of the interfaces facilitate the production of specification tables related to the first step of the standardized tests construction.

Concerning the economical aspects, the research group is encouraged to create a spin-off company at the end of this project. The immediate docimological interventions requested by four of the ten companies and organisations partners in this project constitute an element in favour of the hypothesis of the existence of a market for subcontracting services in connection with the use of the e-C&QCST platform. The acceptation of the SAFES project in connection with these aspects gives us the possibility to make a feasibility study in this direction during the two next years.

***

References


