

E-assessment in The Netherlands, innovations for the 21st Century

Developments since 2002 concerning computer based and computer assisted National Examinations for Dutch secondary education - a report

Jan Wiegers, Cito Institute for Educational Measurement, The Netherlands

Jan.wiegers@cito.nl

Summary

Cito, the Dutch National Institute for Educational Measurement, has invested much time and energy in research and development of innovative e-assessments in the past 10 years. From the beginning of its involvement in e-assessment, Cito has taken the position that computer use in assessment should contribute to the improvement of the quality of assessment in terms of validity, effectiveness and efficiency and furthermore, that the needs of the test taker should be leading, not the technology. Cito has never limited its efforts to one specific type of computer use. Instead a number of innovative new formats that provide versatile assessment options for modern education, learning and training were developed: computer supported assessments, computer based tests, computer adaptive tests and web based tests.

Cito is expert in the field of valid, reliable measurement of learning performance. On instruction from the government, Cito develops the national examinations in the Netherlands for preparatory intermediate vocational education, higher general secondary education and pre university education. As an expertise centre, Cito also does research in and offers advice for modernizing national examinations. The examinations in the Netherlands are the responsibility of the minister of Education. Various parties collaborate on creating the examinations. As stated above, Cito is responsible for designing examination questions. Annually, at the end of secondary education, some 200,000 students in 700 schools take part in the national examinations. Each year, Cito designs more than 500 different tests for all subjects of the various types of education. In most cases the questions are presented to students on paper. However, more and more opportunities arise to administer examinations by computer instead of on paper. Schools¹ are acquiring adequate ICT infrastructure and related knowledge to use this. Computers play an increasingly important role in education. And outside the schools as well, students are making use of ICT and computers in a growing number of situations.

This lecture will focus on the use of computer based testing (CBT)² in the national examinations in secondary education in The Netherlands (in 2009 about 17% of all examinations growing to about 30% in 2014), the goals, the conceptual framing and the process of introduction. Essential part of the introduction is the acceptance of the innovation by students, teachers and principals. Results will be presented.

1. A decade of computer based assessments

Ten years ago, Dutch school leaving examinations were 100% paper-based. Computer use was limited to data processing and test- and item analyses.

Today, we administer and distribute computer based tests and assessments on demand. We moved to computerized adaptive testing and digital student monitoring systems, support online submission of test scores, and collect and process data to generate reports with a single

¹ In the Netherlands the national exams are taken in the school and under the responsibility of the Inspectorate.

² During a CBT examination, all questions are successively displayed on screen. The candidates answer all the questions on a computer. The scoring is also done on a computer. Answers to closed questions are automatically scored, while answers to open-ended questions are scored by a teacher.

mouse-click. Computerized item and test management allows us to store and retrieve items and tests anytime, anyplace. We facilitate downloads of examination papers, marking schemes and results, use networked test construction, an item banking system and a single login test portal. In fact, computer-use covers the entire workflow at Cito from the first steps of item production to reporting.

The first experiments with innovative computer based assessments within the setting of National Examinations started in the late 1990's. Our activities then were already focused on the development of assessments that would go beyond computer based delivery of paper based tests. In 1998, the National Examination Board commissioned a computer based version of the school leaving examination for basic level foreign languages. This computer based examination should include: listening, reading, writing and speaking. The first prototype (a digital "language village" with a simulated holiday trip to villages in France, England or Germany) met most of the technical specifications and requirements. In the final version the speaking component had to be skipped due to limitations of the commercial software that we used and students had to register their answers on paper because the risk of errors in the answer database could not be eliminated. Despite of these drawbacks the digital language village was released as an official examination.

In 2002, the National Examination Board decided to start the "Compex" (Computers and Examinations) project to coordinate different initiatives in the development and implementation of innovative computer use in examinations. Compex examinations were initially taken by small groups of students at a limited number of schools. All students from other schools sat for the paper based version. The aim of the Compex-project was to build up expertise in computer based assessment and introduce computer use in the national examinations.

Compex examinations have been developed as stand alone, Internet based and local network applications, using almost every suitable software package that could be handled by the computer systems in schools. After five years, the focus of the Compex project shifted from innovative computer based assessment tasks towards the development of a standardized infrastructure for the delivery of the computer based assessments that could handle non-linear, interactive, open-ended assessment tasks leaning heavily on use of multimedia.

2. Added value of computer-based testing

Offering a computer-based national examination is a means to an end, not an end in itself. Digitization of the national examination should yield demonstrable added value. The aim is always to achieve the best possible mix of added values (in terms of content, educational importance and organization), in which the added value in terms of content should prevail. Research and the experiences at schools have made quite clear when computer-based examinations have indeed added value.

A great many pilot studies and projects have been carried out over the past few years. They went so well that computer-based administration is now normal for a number of national examinations.

Added value is determined by the following aspects:

- Other skills

Compared to paper-based examinations, computer-based national examinations enable the testing of other skills required in the examination programmes, for example, the skills required to make calculations on a spreadsheet in economics, to analyse animal behaviour on the basis of video footage in biology, and listening and viewing skills in language programmes.

- Alignment to further education

As a result of the digitization of national examinations, the examinations are better aligned to further education. The computer plays an important part in the students' future study and professional careers. This is evident from various examinations in the vocational subjects in preparatory secondary vocational education. For example, the examinations in Building Techniques involve the use of a CAD (computer aided design) program.

- Alignment to students' life-world

Computer-based national examinations are better aligned to students' life-world. Students are used to spend much time behind their computers, which form part of their natural environment. They often indicate they prefer to be shown one question at a time and to type in their answers rather than writing them down on paper. They also notice that computers play an increasingly important part at school.

- More attractive examinations

Students find computer-based examinations more attractive. Questions can be presented in a better way and new elements can be added, such as audio and video excerpts and animations. This is the case for most examinations. Particularly in language and art examinations, where the representation of sounds and images is an essential component, computer-based administration proved very successful.

- More flexibility

The use of computers provides more room for flexibility with regard to national examinations. Digitization makes it much simpler to keep assignments confidential so that they can be used more often and in various versions of an examination. Schools can thus offer students several opportunities to take examinations at different points in time. This proved a great success in the basic vocational programme at the secondary vocational education level, where, thanks to the computer, the general subjects are flexibly examined at 200 schools.

- Reduced workload

The workload of teachers is reduced as the computer takes over part of the scoring.

3. Conditions for digitization

Nowadays, computers feature prominently in national examinations. For instance, it is unconceivable that art examinations would again be administered on paper. That does not mean, however, that all examinations can be digitized. The past few years have shown that it becomes worthwhile to administer examinations on computer only after a number of preconditions have been fulfilled. These preconditions and the ways in which we will put these in place in the coming years are stated below.

- Added value:

Compared to paper-based examination, computer-based examination should have added value (preferably in terms of both content and educational importance).

Generally, schools are asked for their opinions with regard to each examination, particularly in the period in which the opportunities for digitization are explored and tried out. Has a computer-based version more value than a paper-based examination? And is it a better testing tool?

- Level of support:

There should be sufficient support at schools for offering computer-based examinations. This requires that teachers and the school management acknowledge the added value of such examinations, that the schools are well informed, and that they can prepare themselves properly (e.g. by investing in ICT applications) and are confident that they can organize such examinations securely. The level of support should be checked frequently both at schools that are enthusiastic about computer-based testing and at schools that have adopted a more critical stance.

- Financial feasibility of examinations:

It is regulated by Dutch law that the Dutch government pays all the costs involved in the national examinations. Computer-based examinations are currently more expensive than paper-based ones. It is therefore essential that the government remains prepared to continue to pay the construction of computer-based examinations.

- Technical feasibility of examinations:

In the Netherlands, schools are free to choose their own computer systems. National examinations are not administered at a central location, but at all these schools. This situation therefore limits the number of possibilities as far as the technical feasibility of CBT is concerned as schools using *TSS*, *Linux* or *Mac* should also be enabled to administer computer-based examinations.

Examinations often involve the use of texts. In preparatory secondary vocational education, the rule is that students should not have to scroll through texts, which may create problems for language examinations. Particularly schools for senior general secondary and pre-university education are critical where scrolling is concerned. The main reason why CBT has not yet been introduced for language examinations in pre-university education is that long texts are absolutely essential in the examinations at this level of education.

- Organization of examinations

Schools should be able to organize the examinations, which means, among other things, that they should have a sufficient number of computers at their disposal. This is the reason why CBT is not yet offered in all study programmes that encompass subjects in which many students need to take the examinations simultaneously (e.g. Dutch language).

- Information

Schools need to be informed properly on the confidentiality and security specifically required for the administration of computer-based examinations.

- As far as possible, CBT examinations need to be accessible to all examination candidates, including candidates with impairments.

4. New actors, new roles

Compared to paper based tests, developing computer based assessments brings more and different fields of expertise into the process of test development than just item writing skills and psychometrics. A team developing computer based assessments might include multimedia designers and programmers, animators, interaction and interface designers, graphic designers, audio technicians and cameramen. In our experience these specialists need to be actively involved in the process of test development from start to finish.

It is vital for a successful production of computer based assessments these different fields of expertise are included in teams of developers rather than transfer the content of the assessment from one department to another. In contrast to paper based test, development of computer based assessment imply a dynamic and interactive production process and not the linear workflow that is typical for paper based tests.

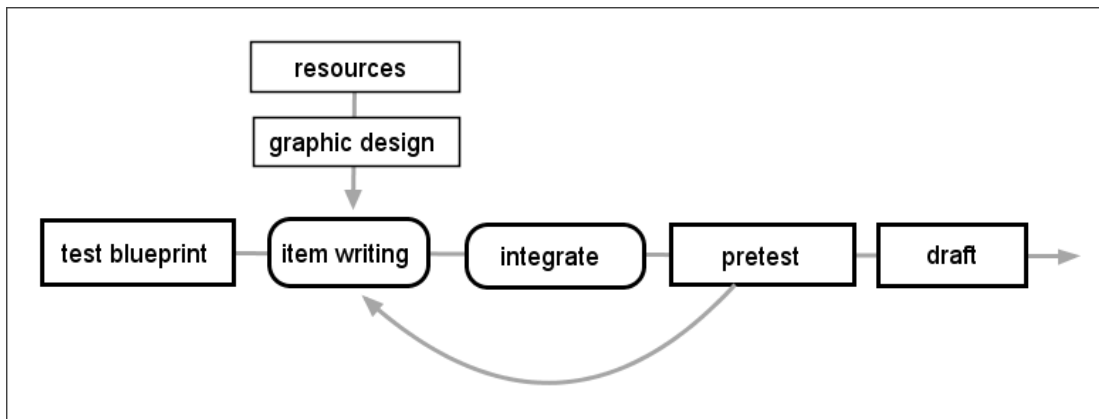


Figure 1: basic model of paper based test development

Figure 1 shows a basic model of the development of a paper based test, ending with the draft version of the test to be submitted to a government body for approval as a national examination.

In the case of innovative computer based assessments, a test blueprint or table of specifications is too limited to serve as a design brief for the multidimensional processes that guide the development. Such blueprints have to be translated into a narrative, a scenario and finally a detailed script, describing not only the required and expected thinking processes and actions of the test taker, but also the possible sequences of events and actions that are expected.

And the analogy with movie making does not stop here, managing the development of computer based assessments, combines major characteristics of the roles that producers and directors have in the case of a movie.

Innovative computer based assessments add a second new dimension to test development. In order to guide the work of graphic and interaction designers, it is often necessary to visualize the content of a computer based assessment in a series of screen designs similar to making of a storybook in movie making.

Figure 2 shows a basic model of the development process of an innovative computer based assessment.

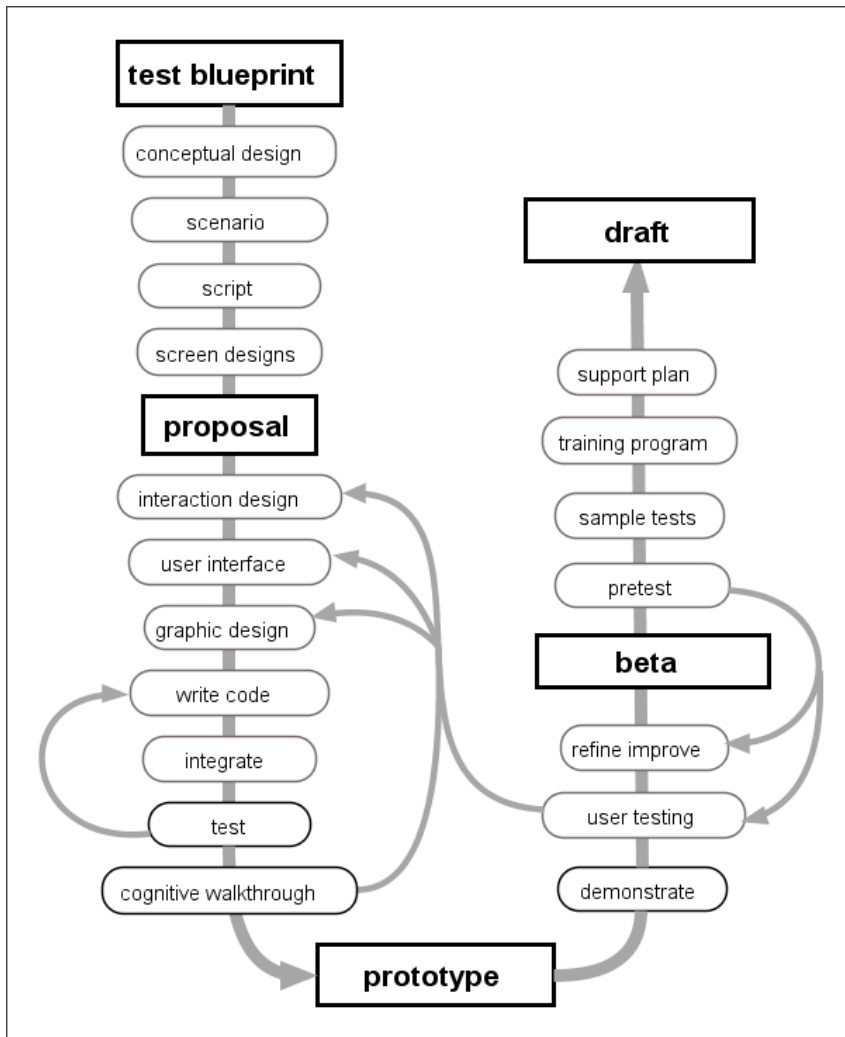


Figure 2 basic model of the development of computer based assessments

The first part of the process is of a creative nature. Test developers, item writers and designers try to imagine narratives that describe a task (a ‘quest’) that includes as much of the knowledge and skills included in the table of specifications or test blueprint as possible. Narratives like this need to be translated into scenarios and at the end in very detailed scripts describing sequences of events, situations and interactions (‘dialogues’).

A proposal, describing the assessment in detail, including assignments, tasks, questions and answers, is then translated into a computer application by designers, animators and programmers. This proposal might include first prototypes of parts of the assessment task or previously developed prototypical applications used in other computer based assessments.

After a series of technical and other tests, the application is then demonstrated as a prototype to the commissioning governmental body. After approval by the commissioning body, a next series of tests, including testing the application with potential users of the test, leads to a beta for testing in the school environment, mostly by using sample tests in the context of a school examination. A support plan and a training program for system engineers and examination administrators are an essential part of a large scale computer based assessment and should therefore be part of the draft of the final examination that is submitted for final approval.

Paper based tests seem to follow a more or less linear development process. Development of computer based assessments implies a process that on the one hand needs to be as rigid as a software development and on the other hand should leave enough room for creative contributions from all the parties involved. Managing this process means finding a delicate balance between control and creation. It also means establishing a common working culture for professionals with different ideas about creativity, standards and control.

Crucial in this process is the development of design standards, the establishment of common design rules and procedures for quality control.

Before the draft of an innovative computer based assessment is submitted for approval as a national examination, the Board of Examiners has at least reviewed three earlier versions of the assessment task: the proposal, the prototype and the beta version. Planning 'go - no go' decisions at these points in the development process is essential for the success of the project. Developing computer based assessments is expensive and most of the time funds are limited, and therefore the commitment from the governing body from beginning to the end is vital. The decision about the prototype is the most important moment in the development process and there are three kinds of testing routines related to this phase in the development of a computer based assessment task.

- Before an innovative assessment is released as a prototype, a series of tests, so called 'cognitive walkthrough', is carried out by experts with special focus on interaction design, interfacing and ergonomics. In a 'cognitive walkthrough', the sequence of actions refers to the steps that an interface will require a user to perform in order to accomplish some task. The experts step through that action sequence to check it for potential usability problems.
- A technically reliable delivery of computer based assessments requires a thorough testing routine on a series of different configurations and under all of the conditions imaginable during a national examination, long before the actual examination takes place. Every change in the application, makes it necessary to go through the testing routine again. Not only changes in the program code require a new series of tests, changes in the content of innovative test-items can also lead to an unexpected error message, computers crash or loss of data.
- It is vital for a successful delivery that a second series of technical tests is done at every school. For this purpose schools receive a sample test, with versions of the assessment that equal the actual examination technically but with a different content. Running these kinds of tests well in advance of the examination period also helps to guarantee that students are enabled to complete computer based assessments without problems.

5. CBT implementation procedure

Before a CBT examination can be offered as compulsory, a meticulous implementation procedure will need to be followed. This procedure covers several examination years. During this procedure, it is investigated whether the examination in question lends itself for computer-based administration, and the schools involved are given time to get used to the idea and to prepare themselves for the CBT examination.

Year 1 – exploration

This stage is required for examinations that include special items or have been given a new structure. Here, a sub-examination is constructed using relevant items. This sub-examination is then submitted to the schools for their opinions and to be administered to the students. This sub-examination is also used as an example for experts (e.g. teachers' associations) as a major

objective at this stage is the determination of the exam's added value by those working in education. Another objective is to determine whether it is financially and technically feasible to offer the examination on computer (for several years). If, at this exploration stage, it becomes clear that these two objectives cannot be achieved, computer-based administration will be cancelled.

Year 2 – experiment

Once the exploration stage has been successfully completed (and the added value as well as the financial and technical feasibility of the computer-based examination have been determined), a full-scale examination will be constructed. This examination will be submitted to a small number of schools to be administered outside the official examination periods. Depending on the objectives of the experiment, the schools will be asked to either include or exclude the results of this examination in the calculation of their students' final results. In experiments in which psychometric data are collected, for example, it is crucial that the students take the examination "as if for real". In experiments designed to test the software or the user-friendliness of an innovative assignment, this will not always be required.

Year 3 – examination test

A full-scale examination is constructed (which may contain elements of the experiment conducted in the previous year), which is administered by a small number of schools in the official examination period. Partly as a result of the additional tension caused by the examination conditions, all aspects can now be tested (examination, software, organization).

Year 4 – pilot study

A full-scale examination is administered in the official examination period at a reasonably large number of schools (20-40% of the total number). Robustness, added value, user-friendliness and psychometric accuracy are established beforehand. The pilot study is particularly conducted to investigate whether further implementation is advisable. If there are several variants of an examination, this will be the first year in which these will also be used (several administration opportunities are scheduled for May).

Year 5 – transitional year

The examination will now be offered at all the schools. In this year, schools will be encouraged to administer the examination on computer, although this will not yet be compulsory. This is the last year in which a paper-based back-up will be provided.

Year 6 – compulsory computer-based administration

No paper-based back-up will be provided in this year.

The implementation procedure described above constitutes the starting point. If there is reason to do so, this procedure may be deviated from. To assess whether a certain stage can be completed positively, the following criteria have been formulated:

The exploration stage has been completed positively if:

- The computer-based examination has sufficient added value in terms of content and/or educational importance according to the client (also based on the opinions of those working in education).
- It is technically and financially feasible to produce an examination with sufficient added value each year.
- There is sufficient technically tested material to construct a full-scale examination.

The experimental stage has been completed positively if:

- The examination has sufficient added value according to the participating schools.
- The administration of the examination has not brought to light any major technical or organizational problems.

The examination test stage has been completed positively if:

- The examination has sufficient added value according to the participating schools.
- The schools did not encounter any problems during the administration of the examination in the official examination period.

The pilot study has been completed positively if:

- The examination has sufficient added value according to the participating schools.
- The organization of the examination went well as far as Cito is concerned.
- There was adequate technical support.

The transitional year has been completed positively if:

- More than 80% of the schools have registered themselves for the computer-based examination.
- Measures have been taken with regard to schools that could not participate in the optional year for technical reasons (e.g. non-compatible operating systems).
- Sufficient measures have been taken to properly inform and provide technical support to all schools.

6. The schools' experiences with the implementation of CBT

At a conference convened in late 2006, teachers were asked whether they considered their subjects suitable for CBT and whether they thought this would have added value compared to paper-based examination.

This resulted in a list of examination subjects checked and modified by the Central Examination Board (CvE) and Cito. A number of subjects were excluded, for example, because the student groups at the schools were too large or because implementation was technically not really feasible. Other subjects were included because there would be sufficient added value according to CvE and Cito.

The added value of a number of examinations (including language and art subjects) were proven, and the implementation of CBT in these subjects was therefore initiated. Other subjects still required further investigation before CBT could be implemented.

There are a number of subjects in which CBT can be implemented quite easily. The CBT examinations in these subjects are therefore the first that the schools will become acquainted with. These subjects are typified as follows:

- The added value of the CBT examinations in these subjects is not disputed.
- The production of the examination is technically simple (the computer applications require few additional costs and do not result in extra work).
- The examination subject is offered preferably at a large number of schools so that many schools become acquainted with CBT examination when implemented nationwide.
- The number of students taking examination in the subject is preferably not too large so that the organizational burden of this first CBT examination is not too large for the school.

The following subjects were defined as suitable for relatively easy CBT development:

Table 1: Overview of subjects for which CBT can be developed relatively easy and that also are widely supported by those who work in education.

Schooltype	Subject	% schools offering the exams	First year experiment with CBT-examination

Vmbo ³	Biology	100%	2008
Havo ⁴	German language	100%	2007
Havo	Art	30%	2007
Vwo ⁵	Art	30%	2007
Vwo	Music	13%	2008

These subjects are particularly those that are part of language and art examinations, because their added value is indisputable. Biology in preparatory secondary vocational education is also one of these subjects, given the rather good experiences with this subject in the Complex project and the simplicity with which Cito expects to convert this examination into a CBT format. It is still difficult to find subjects which are suitable for this purpose and are taught at all schools for pre-university education. This means that it will still take some time before these schools will be able to administer computer-based examinations.

Since 2006, annual surveys are held among teachers, students and managers of the participating schools after the examinations in order to record their experiences. Below, you will find a brief summary of the results of the past few years.

Table 2: Confidence in the new CBT examinations as measured among examination secretaries, teachers and system managers

	Pilot (small scale)		National Introduction (large scale)	
	2006	2007	2008	2009
Confidence in next year CBT examinations (examination secretary)	89%	97%	47%	71%
Confidence in next year CBT examinations (teacher)	76%	79%	50%	66%
Confidence in next year CBT examinations (system manager)	51%	60%	14%	48%

In 2008, there were various software problems at a number of schools. Confidence in the examination software was largely restored in 2009 after a decrease in the previous year. The level of confidence is, however, not yet as high as 2-3 years ago. This probably has to do with the following:

- In 2008, nation-wide implementation was initiated, and many schools took part for the very first time (the group of participants in 2006 and 2007 consisted mainly of a small number of 'pioneers' that were already enthusiastic to begin with)
- In 2008, a transition was made to a new software version for which the schools' intranet servers are used (more secure, but the communication with the intranet servers was a new aspect and occasionally led to malfunctions).

Table 3: Overview of the opinions of examination secretaries, teachers, system managers and students in response to the question as to whether they would like to go back to paper-based examination

³ School for preparatory secondary vocational education

⁴ School for senior general secondary education

⁵ School for pre-university education

	Pilot (small scale)		National Introduction (large scale)	
	2006	2007	2008	2009
Preference for paper-based examination (examination secretary)	3%	4%	12%	9%
Preference for paper-based examination (teacher)	17%	18%	27%	24%
Preference for paper-based examination (system manager)	-	-	-	20%
Preference for paper-based examination (student)	12%	14%	13%	6%

Only 6% of the students preferred paper-based examinations. A large majority of the examination secretaries, teachers and system managers was also positive about computer-based examination.

Some of the teachers still had some doubts. To them, automatic scoring and the flexible administration to students does not always make up for the amount of extra work and the tension involved. A recurrent complaint is that the assignments made by the students at confidential examinations cannot be inspected. The teachers also do not always like to do the scoring on screen, and some still ask for scoring instructions on paper, but these cannot be provided for confidential examinations. The importance of confidentiality therefore needs further explanation at a number of schools.

Table 4: Overview of the opinions of examination secretaries and system managers about the sustainability of CBT examinations

	2007	2008	2009	2010	2011	2012	2013
The examination secretary expect the computer-based administration to run smoothly next year.	97%	47%	71%	94%	92%	95%	98%
The teacher expects the computer-based administration to run smoothly next year.	79%	50%	66%	97%	95%	90%	-
The system manager expects the computer-based administration to run smoothly next year.	60%	14%	48%	94%	92%	80%	94%

In 2013 98% of the examination secretaries expect the CBT examinations to run smoothly the subsequent year. Over the last four years, a vast majority is positive.

The experiences so far make clear that schools at which computer-based examinations are held for the very first time need to be given ample support. Practical experience has shown that this need for support decreases steeply over the course of time.

7. Recapitulation

Ten years of developing computer based assessments have taught us that:

- Managing the development of computer based assessments is a mix of classical (product oriented) project management and process based (vision oriented) management. Both need careful monitoring during the entire development process.
- Members of multidisciplinary development teams have different training, ways of working and culture and the roles and boundaries between the different disciplines change and shift

constantly. This not only affects how team members work together as a team but it can also cause confusion about responsibilities and decision making.

- The basis for a successful development of innovative computer based assessment is a precise script describing the actions the test taker does or is supposed to do (mental operators as well as physical actions) to complete the assessment task.
- Before introducing computer based assessments in national examinations, experiments and pilot projects are to be carried out to gain the necessary experience with the use of software and organizational aspects (examination and administration process).
- Crucial in this process is the development of standards, the establishment of common rules and procedures for quality control
- Close cooperation with management, teachers and IT staff in the schools is a prerequisite for managing the quality of computer based assessments.
- Computer based assessments enables a more flexible organization of national examinations. Schools can then offer more examination-moments, or organize the examinations in different periods than is the case in the current situation.
- In general, the reactions of schools are favourable. Students also favour computer based examinations. In most cases they find the design of the questions on computers more appealing than that on paper.
- More pilot projects will be carried out in the next years. The use of computer based testing in the national examinations in secondary education in The Netherlands is expected to grow to about 37% in 2011.

8. References

- College voor Examens (2009). The computer in the Dutch national examinations 2009-2013: -definitely digital.
- College voor Examens (2009). Evaluation of digital examinations.
- College voor Examens (2013). Evaluation of digital examinations.
- Hermans, P.J.H., (2010). Developing an e-assessment environment for the 21st century.
- Martinot, M. J. (2009). Examinations in Dutch secondary education - Experiences with Cito -Tester as a platform for computer-based testing.
- Wiegers, J. (2013). E-Assessment: 10 Years experience of the Dutch Educational System. A -Lecture presented to the Management of the Joint Admissions and Matriculation Board, Abuja, Nigeria.