Mega-Trends: Technology in Testing

Marten Roorda
Chief Executive Officer
CITO, The Netherlands

In a time where test publishers must make big technology investments, we must consider what the trends are in computer-based testing, and where they are leading. Speaking on behalf of CITO (Europe’s market leader from The Netherlands), this presentation offers a European view on the future of testing.

In order to present a future vision, it must first be established what are the benefits of using ICT in testing and assessment. From these benefits the developments should evolve. Then we must distinguish between target groups, because different markets will have different needs. In the next part of this presentation a list of mega-trends is put forward, looking at a wider range of changes in society. After a brief assessment of the state of the art in computer-based testing, a comparison between the American and the European situation is made. Given the mega-trends, we can imagine a futurology of testing and assessment. If this future would be the true scenario, it could lead us to investment decisions, hoping – at the same time – that open sources, sharing and setting standards world-wide will make the burden of investment less hard to carry.

(Bron: Internetsite ATP 2004 Annual Conference)
Mega-Trends: Technology in Testing

In the American masterpiece ‘2001: A Space Odyssey’ a television network conducts an interview with the crew of the Jupiter Mission which consists of five men and a computer. The sixth crew member is described as ‘the latest result in machine intelligence: the HAL 9000 computer, which can reproduce most of the activities of the human brain and with incalculably greater speed and reliability’. Hal is also interviewed and exhibits a very human way of speaking. When his great responsibility with respect to the mission is pointed out to him, Hal shows that he is fully confident. ‘We are, by any practical definition of the words, foolproof and incapable of error.’ In response to the question whether he never becomes frustrated by his being dependent on people, Hal replies extremely correctly: ‘I enjoy working with people.’ Those who have seen the film, know that Hal has adopted another typically human trait here: lying. During the journey to Jupiter crew member Dave has a confidential talk with Hal. Dave has drawn and shown some sketches, about which Hal comments: ‘I think you improved a great deal.’ Then Hal asks Dave what he calls a personal question, namely what his feelings about the mission are. But because the human crew member knows the computer wants to perform an unnoticed assessment, he evades this directness, saying: ‘You’re working up your crew psychology report?’ Hal replies: ‘Of course I am. Sorry about this. I know it’s a bit silly’, and changes the topic of discussion. By then however Hal already knows enough, not by an examination of fifty questions but by a casual psychological approach. In the movie this leads to a rejection of all human candidates, with all the associated consequences. Of course things will go wrong when you make computers the boss. It is incredible what an apt vision and representation of the future Stanley Kubrick and Arthur C. Clarke came up with, especially when you realize that this film was made even before the first moon landing. Still the question remains how far away the future is, because many of the predictions in the film have not come true so far. For a certain image of the future you can say chances are that it will occur, but the question is when. Futurologists generally have a tendency to project that future too soon, even though developments are already going as fast as they are. It is not insignificant that the future of A Space Odyssey, considering the year 2001 in the title, is already yesterday's future. The reason for the exaggerated optimism is that futurology is often already bound up in the present. We often find ourselves saying (do catch yourself at it sometimes): everything is already technically possible. Restrictions are imposed by earthly matters, like a lack of money, changes in politics or society, and the fact that progress is not determined by the available technology but by the demand of its users, which cannot be planned ahead.

Although my origins lie in the publishing industry, I have now spent about a year and a half in the world of testing and assessment, as CEO of the Dutch test development company, Cito. As a relative novice I may be able to offer a fresh view on the business and give some insight into my strategic vision. This vision is inspired purely by business economics, translated into attainable goals and many of the trends I will point out can definitely already be felt and noticed in our organization. Innovation is an economical imperative for us and at the same time second nature. We do it out of curiosity, but also to save our future. Innovation is necessary for the continuity of our business and that is why we invest in it heavily. It is a journey into unknown space, a world ‘not as we know it’, but it is certainly not science fiction.

In order to extend a line into the future, we first have to know the starting point, which of course lies in the present. You probably know very well what all is available in the realm of test technology, so I will limit myself to a very brief overview, just to demonstrate a point. The achievements lie mostly in the following areas.
1. Item production. Various software applications are available which can be used to construct or generate items in an automated way. A special protocol for test purposes has even been developed within the IMS standard: the Question and Test Interoperability.

2. Item and test management. There are systems for managing item banks and composing tests. Adaptive tests are an example of a technology perfected by Cito, among others. In our institution there are already many applications of this, as was described earlier in this conference in our contribution "Dutch treat", by Cor Sluijter, Theo Eggen and Willem Rosier.

3. Test administration. Computer-based testing is employed more and more, both in education and in the business world. Here also some standard administration systems have been made which are carried out both in and out of a high stakes environment. In worldwide test center networks the highest degree of administration reliability and security can be achieved. The Internet offers additional opportunities for distribution and administration, especially in cases of ‘low stakes’.

4. Test processing. There has been an enormous increase in the number of possibilities for automatic processing, analysis and reporting of computer tests. There are even systems for automatic scoring of open questions and essays.

The advantages of the currently available test technology are crystal-clear.

1. It is faster. That goes for the test process as a whole, starting with production. The administration itself is also more efficient and as soon as the tests are made, they have been handed in and are available in the right place. Scoring can largely be done in an automated way and is therefore finished quickly. The results can also be delivered instantly, together with the item analyses made. Thanks to all of this, feedback from test or exam to the learning process can be delivered more rapidly.

2. It is more valid and reliable. The use of ICT in tests can lead to better quality in the instruments. That quality may be translated into two areas. On the one hand there is the content aspect: each item can be made more authentic by using multimedia and interactive technologies. On the other hand there is also the psychometric aspect: by putting the item response theory into practice, more data becomes available per item. This can be employed for instance to make more reliable tests but also for more advanced technologies like adaptive testing.

3. It is available more easily. This argument is as practical as it is well-known. Computer-based tests can be administered at any desirable moment. The use of an item bank makes flexible, asynchronous administration a piece of cake. Any distance can be bridged. That is why all employees of a multinational corporation or all military personnel in NATO can take the same test throughout the world. Scorers also can do their work anywhere.

4. It is more cost-effective. And that is music to our ears as test developers, because is not that what it is all about? By the use of item banks we do not have to produce full tests time and again. We save on printing and dispatching costs. All of a sudden we have larger ranges of distribution. Even a modest test development company in a small country can now supply to the entire world, even though a fair amount of research, translation and localization will have to be done. New business models are emerging, where licenses are used more and solutions are found for unauthorized re-use of tests, by working as a service provider and billing via pay-per-view. Cross-selling is possible much more easily this way as well.

By now this is beginning to be a reassuring and also well-known story, something to make you lean back in satisfaction: great achievements and great advantages, what more do you want? Unfortunately I now have to add two critical notes.

1. A considerable part of the technological investments in our sector is going towards perfecting traditional test processes. Many test companies really consider it safer and easier to keep all facilitating and logistic functions on their own premises, starting with contact with
the client by telephone, processing most applications by hand, including email. A lot of money is put into call-center systems. It continues with designing paper tests via desk-top publishing and next with processing the administered tests, machines scanning forms at break-neck speeds. Last but not least, electronic platforms and electronic scans make sure that answers to open questions are scored extremely efficiently. Summarizing, a whole lot of technology is employed to make paper-based testing run perfectly. That makes it more difficult to change to computer-based testing. Every company has to ask itself consciously whether support and logistics really are a primary company process. If not, it would be better to spend the investment sums on higher-quality issues, things that will help acquire more expertise in the core business. If non-primary company processes demand too high an investment, it might be advisable to outsource these processes.

2. When I listed the four achievements and the four advantages of test technology, you will have noticed that all achievements relate to production, management and processing, in other words the domain of the providers in the market. And all the advantages are also primarily the advantages of the test companies. If I were to approach the matter negatively, I could say that the suppliers and administrators of tests have primarily considered their own interests and have forgotten about the users and candidates. And where there is a win-win situation, we forget to sell the advantages to the clients. One would be right in arguing that our market is mostly supply-driven, which is characteristic for a ‘technology push’, and it is reinforced by the dependence on government funding. It is a tough lesson from economic history that innovation, economic renewal or growth work best based on a technology pull, when the market itself demands something. The recent collapse of a large part of the commercial Internet activities has made that clear once again. The future can be imagined, but it cannot be made.

Let us look at test-technology innovation from the end-user's point of view for a change. The drawbacks of test technology that can be discerned from this angle can turn into advantages when more attention is paid to this.

1. The employability is limited. The fact that only a limited number of item and test types have been automated successfully will remain an important hindrance for the foreseeable future, unless additional investments are made here. It may be true that by now we can let the computer correct essays and short answers to open questions fairly reliably. But the answers do have to be identical in form and the correction does not reveal what the candidate's real insight is. The more important thing however is that there is a movement more interested in competence than only in knowledge. Simultaneous with the digitalization of tests we will therefore also have to extend our work area to assessing all kinds of skills and to ways of making work attitude and attitudes in general visible. The technology for this will be a long time coming.

2. The individualization is limited. Computer-based testing initially decreases the personal involvement of test subjects and examiners, and increases the distance. It is an impersonal instrument that will not lead a priori to the same motivation in those concerned. We test people because we want to know what level they have or to find out whether they meet a certain standard required for a certain profession or advanced training. That level and that standard are determined by people. Testing always has to be objective, but in fact it is the objectivation of the human measure. One of the major advantages of ICT for testing however is that we can individualize. It is easier to link up with the candidate's individual learning path and it is even possible to give a different test to each person. Additionally ICT offers a unique possibility to monitor people's development and to stimulate it specifically. And the impersonal aspect of computer exams is offset by the fact that they enable you to test much more authentically and attractively.
3. The infrastructure is limited. Paper tests do not often disappoint us. It is a thoroughly tried system, in which all failings and flaws have been remedied. With computer tests there is a chance of system failure and the use of Internet presents problems regarding security and speed. On the client's side there is a range of computers and networks, which often makes things we would like impossible, and even if the hardware is sufficient, proper test conditions have to be ensured, like a separate computer room. But even if we can arrange for distribution and administration well, there is the problem that the industry will have to invest large sums during the next years to get computer-based testing where we want it to be. Looking at the cost per test administration the computer will win quickly, particularly when test volumes are large, but then we are skipping over the technological innovation that has to be financed first.

In any change in test technology the users' advantage should come first. If we assume that in principle technically everything is possible, we should not base our reasoning on technology but on objective, on what we want to achieve or enable. In order to sketch the point of departure I would like to ask the members of the audience here to think of the tests you yourselves have taken, your own exams in the past. Did you look forward to taking the exam? Was it fun to take that test? Did you study for the exam or for life? For the certificate or for the knowledge? How many of the facts you learned do you still remember? And when have you been in a position to use the tested knowledge?

In order to predict trends in test technology we have to assume that the computer will become a real companion for the learning individual, that testing will take place much more along the way, often without the candidate noticing, and that tests will simply become much more fun. Test situations will be realistic and challenging. When a proven level is present, the relationship between tester and test subject will not be broken: a person's development can be monitored throughout their lives. Whether it be new-borns subjected to an APGAR test, or toddlers studied for deviations in development, whether the issue is high-school or university students and their schooling development, whether it is the competence development of employees in companies, testing is not only about measuring human potential but also about measuring human development.

For the following eight mega-trends in test technology I will base myself in each case on the advantage of the innovation to the candidate. I assume that we as providers will benefit anyhow. Demand-orientation will cause a true revolution in the world of testing and assessment. The fact that this will happen unavoidably is shown by the innovations that are manifesting themselves in education and training and which we should tie up with. Here are the eight trends.

1. Learning and testing become one. In the future, e-learning and computer tests will no longer be able to exist separately from one another. When I worked in corporate trainings, I saw the emergence of e-learning and contributed to it myself. The delay of the breakthrough did not only result from the economic circumstances. A ‘technology push’ made sure that the investments in e-learning mainly went to infrastructure. Many companies and institutions had wonderful, very expensive learning management systems, but could not continue with them because they could not afford investments in content anymore. The most important advantage that was noticed was that of volume and distribution: larger groups of entrants, anywhere in the world, could participate in e-learning. Of course this is a reversed argument that we ought to be aware of, also in the testing world: since the investment is too expensive, the highest possible number of people have to share in it. Of course e-learning only has a right to exist when it can offer more than learning in a classroom or from a book. The real added value only appears when e-learning can adapt to the pupil, student or employee. Continuous assessment of someone's level and a diagnosis of the learning need and stumbling blocks is needed to adjust the supply to the individual. The only method to do so is a form of continuous
assessing, testing and measuring. That takes place partly via short response, questions between the lines or possibilities for the participant to supply input, and partly also by interim tests. It is unavoidable that learning and testing will be put together into one ICT system, blurring the boundary between the two. When we talk about the integration of testing and learning, we can also think about learning while testing or testing while learning. Questions about certain content are repeated until they are answered flawlessly and until all subject matter has been dealt with. Thus the lesson material is presented in the form of questions.

2. Competence development. In the future the whole of human capacities will be assessed and monitored much more. Competence measurements are used to determine which part of the human potential has been transformed into capabilities in a specific context. In my view, there have been three periods in the history of testing and examination. Let us look at them for a moment. In prehistoric times initiation rituals were used mostly, determining whether a member of the tribe would be allowed to belong to the group of adults. Apart from some skills the focus of the tests would be to measure attitudes, like: obedience, courage and survival potential. When more and more specializations developed after prehistoric times, trades and professions had to be protected by means of trying a person's capabilities, their skills being assessed primarily. A good example of this is the so-called masterpieces that had to be presented to the Medieval guilds. In current times a very large part of the examinations is administered by means of tests, assessing primarily the candidate's knowledge and understanding. In competence measurement all the properties from earlier periods are joined together: knowledge, skills and attitudes, in other words everything a person needs in order to function well in a specific context. This may be a certain position or profession, but also a potential advanced training. Systems for personnel certification already point in that direction. These systems check regularly whether an individual in a responsible job possesses enough competence to practice this profession. Through Cito several certification routes are in operation, for positions like real-estate agents and project managers. It is also interesting that we have computer-based assessment instruments for teacher competence, with a solid psychometrics foundation. My colleague Erik Roelofs described those earlier during this conference. ICT will be given an important role in joining different assessment criteria, each of which is an operationalization of a range of variables. Digital portfolios and data warehousing will help us store various data which may be important in an assessment.

3. Monitoring development. In the future tests and assessments will constitute an integral part of an electronic monitoring system in which the learning achievements are charted, plotted against the desired or expected standard. It will no longer be desirable to use one-off tests that are separate from the learning process or the learning development. Regular measurements with 'low stake' tests within the system will be on the same psychometric scale with 'high stake' tests that determine whether the person has completed the current level successfully. In the Netherlands Cito supplies ninety percent of all primary schools with a student monitoring system, which by now has been digitalized and whose tests are currently being transformed to computer-based ones. This enables schools to identify deviations in their students' development, they can use the data to work on quality improvement, but they can also generate reports from the system to their boards of directors and to parents. In 2004 we will adapt the basis of the monitoring system for application in other markets, like secondary education, higher education and even for monitoring competence development in the business world. I expect that all of these target groups will get their own monitoring systems and that when a person moves from one group into another, for example from primary to secondary education, the data go with the student, so that everybody gets a 'personal file' or a portfolio which is added to and changed as long as you live. It seems likely that the monitoring system will eventually constitute the backbone for the learning and testing modules I described in the first mega-trend. That way a combined learning, testing and monitoring system will come into
being that accompanies a person throughout their lives. Particularly during pre-school years development will be monitored and stimulated more intensively. For this, Cito has launched the successful program Pyramid which offers a stimulating learning and play program for ages three to six, with many testing and observation moments, linked to a proper monitoring system. At Cito we consider it quite normal that toddlers and pre-schoolers are tested at very young ages. In Pyramid and in the monitoring system we already have pre-schooler tests that we administer entirely via computer and adaptively as well. Pyramid has now been introduced successfully into Germany and in the current year Cito will also introduce the program on the American market, from our newly opened office in Montclair, New Jersey.

4. Authentic and realistic. In the future tests and assessments will much more resemble the real situation about which they have to make predictions. This will increase the predictive validity. ICT yields powerful instruments to stimulate these situations, part of which may have really happened and part of which has been made up. The advantages are that the assessment becomes much more valid and that the candidate is presented with an interesting and challenging test. The line between practicing and testing will also be drawn less clearly. It may well be that someone, after a number of practice sessions with a simulation, is told they now know how to do it well enough, and therefore have passed. A simulation may also be linked to a real situation. For example in my former job I invested in an educational laboratory for measuring and control engineering that really works and is controlled from the PC. By making computer control reachable via the Internet, students can control the laboratory from their own work station and read off all gauges. In simulations, the reality level can be improved by using ‘virtual reality’, in which the person involved actually seems to find themselves in the simulated environment. Another step is made when a game element is added to a simulation, so we can use the term ‘gaming’. This method has been employed conventionally for quite a while to improve the performance of teams. In electronic versions ample possibilities can be built in to reach a good assessment of the participants. Moreover: the final score of the game may be the result of the test. The possibilities are numerous and the game can be played alone or with several participants. In the game, participants can set up companies and make profits or incur losses as a result of decisions they make. They can participate in traffic and try to reach the final destination without violations or accidents. Just like in the CD-ROM’s for our children, while playing they can attain higher ‘levels’ in order to finish after they have reached the highest ‘level’. A good example is the game ‘America’s Army’, which is now employed for recruiting personnel for the American military. A recruitment officer was quoted as saying: ‘Youths who play computer games have no trouble processing a lot of visual information on a monitor. That makes them cut out for the new type of troops that the future army will be made up of.’

5. ‘Stealth testing’. In the future tests and assessments will take place much more often without the subject noticing. During an electronic learning process, data are stored about how someone goes through the modules and between which modules they switch. Apart from this, the ‘log files’ also contain data about the speed of information acquisition, results of casual and direct tests and other measurements, from which an assessment is formed of a person's capacities and competence. The same type of assessment can take place during simulations or ‘games’. The computer can substantiate its assessment with observations which we do not experience as typical test elements, like the way in which you have reacted or the degree of self-assurance that has been exhibited. It is even imaginable that all activities performed by an employee on the computer in a normal work situation are used to form an assessment of the person's current level and if desired, to propose a training plan. Of course we do have to realize that at some point a limit will be reached, when the computer becomes a ‘big brother’, invading people's personal integrity too much.
6. Computer becomes companion. In the future the computer will act much more often as a single user's personal company, partly playing the role of an assistant and partly also that of an assessor. Keyboards will retreat into the background and spoken contact between human being and computer will take its place. The computer will have more anthropoid traits and one will be able to have entire conservations with it. This will land us in a situation that has already been described in many a science-fiction series or film, like in Star Trek, Blake’s Seven and – indeed – also in ‘2001: A Space Odyssey’. In some learning environments today, already so-called ‘intelligent agents’ are active, which adopt the role of electronic coach and guide the student or trainee through the learning process. The ‘agent’ can also act as a navigator between the monitoring system, external sources of information, the education or training institute and other participants in the environment. Such a personal assistant can even be completely independent of learning and test systems and represent the interests of the candidate. Such systems make use of artificial intelligence and can expand their knowledge of the world as well as of the student continually. Computer companions can be a friend, they can be an advisor whom you ask for a pointer regarding further studies, but also for your next vacation destination. The computer program that is your assistant also has to be able to assess you well, because you tell it a lot. On the other hand these ‘tutor agents’ can only help the student properly when they have enough information from that student at their disposal. That defines the relationship between ‘tutor’ and student as a relation of trust.

7. Just-in-time-tests. In the future the test moment will be as close as possible to the moment of use, and the test location close to the location of use. Of course we have to acquire certain basic skills and basic knowledge at the beginning of our lives or our careers. That is why those are tested and examined in initial education. Apart from that however, in order to function well in a profession or position, we need a lot of ready knowledge and skills. In contrast to today, soon the readiness of these will be tested more often when we need it. For it would be strange to assume that someone who was granted a diploma years ago, still has command of the relevant skills. And it is a waste to carry unnecessary knowledge around with you. For difficult, responsible jobs therefore people will be tested just ahead, ‘just-in-time’, immediately prior to hiring. Since it can be difficult to organize this and since these kinds of situations happen a lot, in the future we will see more and more often that all kinds of professions and trades will be certified. A means to test just-in-time and on the spot will emerge by the introduction of various types of mobile, wireless technology. It is expected that a broad range of technological services will become available in small personal instruments. During the development of e-learning I already looked at the possibilities of m-learning, mobile digital learning. This was to be linked to mobile portals with information services. Publishing companies are currently working hard on these types of service, which should make it possible to provide tailor-made information for example to medical specialists with handheld computers at the patient's bed, from symptoms up to and including medication. Just before an unusual operation this doctor could take a quick test via the same system. Of course one needs little imagination to think of a link with the personal assistant I described in the sixth mega-trend. In stead of the profession of doctor I could easily fill in other, less prestigious professions, like maintenance engineer, ‘operator’ or sales representative. One example is the European project of M-learning which uses computer games and innovative teaching materials for mobile telephones and ‘personal digital assistants’ to try and stimulate young people who do not feel like learning.

8. Globalization. In the future, testing and assessing will be part of the international domain much more and this will be subjected to world-wide standards. There are clear indications that the exchange of knowledge and information will take place in an entirely different way. That is already taken into account in the innovation of the current Internet to Internet 2. For instance experiments have been done with a ‘science grid’, that independently of the location
can make a large number of servers and computers work together, combining an impressive
collection of data and great calculating power. In the Netherlands currently the measuring
data of all radio telescopes are being combined via the Internet, resulting in an immensely
powerful radio telescope. Here and there information databases and knowledge banks are
being made medium-independent and they will be linked to expert systems that are available
everywhere and always. Simultaneously via the Internet ‘communities’ are coming into
existence in all sorts of areas, in which professionals from all corners of the world share their
know-how. Within these virtual communities the participants are sometimes very capable of
assessing each other properly. ‘Peer reviewing’ is already taking place for the evaluation of
scientific articles, but in the future it will occur especially in virtual communities. A parallel
movement is to bring international educational systems more into line. In the long run for
example most European countries will adopt the bachelor-master system for higher education.
A student will then be able to switch from one university to another effortlessly. The
examination and test standards will be set much more on the basis of international
qualifications in the future. Please consider moreover that young people these days can find
their way to the Internet easily to play an ‘adventure game’, a game of chess or bridge with a
number of completely unknown fellow-players from various countries. The conclusion to be
drawn from trends like these is that testing in the future will become partly self-organizing
and will take place literally and figuratively above our heads. As test companies we can keep
up by participating in the development of international standards, both technically and
content-wise.

Of course we can try to look even further into the future by searching for instance for the
cosmic antenna to register the cognitive resonance that Rupert Sheldrake believes is present.
One day we will certainly also develop the technology to download and upload brain capacity,
including knowledge, skills and attitudes. But such inventions threaten the roots of testing and
therefore I consciously limited my argument to developments of which signals can already be
noticed now. Which is why I do not take these possibilities into account.

There is a clear difference in points of departure for these trends when we compare the US
and Europe with one another. In order to do that I am forced to make generalizations, because
the situation in the US sometimes differs strongly between states, and in Europe the national
differences are even bigger. Still it can be interesting to make a comparison in broad outline.
The educational and training system in the US is something that commands respect in Europe:
it is ambitious, offers a lot of room for the best and leads to great achievements in the fields of
science and economy. But the future is not the present and we cannot extend the current line
just like that. I will now indicate seven differences with Europe which contribute to the degree
in which the mega-trends indicated can be responded to.

1. In Europe most school systems aim at a high average level, which stimulates lesser and
deprived students as well. The system in the US focuses more on students who perform best.
This results in more individual peak performances, but yields a narrower over-all basis.
2. In some European countries the quality differences between schools are not very large. In
the US being in the right school largely determines your success in later life. Often money
plays a role. A good education is affordable to a limited group.
3. In many European countries there is a national curriculum. For example in the Netherlands
every type of secondary education is concluded by exams, all made by Cito. If you pass it,
your diploma will have the same value and effect throughout the country and you almost
automatically have access to the corresponding advanced education, also higher education. In
the US these types of transitions are not matter-of-course.
4. Many tests in the US are admission tests, for instance knowledge tests, language tests or
learning aptitude tests for higher or advanced education. Many of these tests have become a
phenomenon in themselves and are fairly separate from schooling or training that are taken
before or after. A considerable trade has even sprung up in readying people for such tests.
That also goes for assessments, psychological tests and intelligence tests which are studied
and practiced for extensively. In Europe there is often a final test rather than an admission test
and practicing is done less often.
5. The ‘war on testing’ plays a larger role in the US as an interfering factor than it does in
Europe. ‘Evaluation apprehension’, teachers being shy of assessing, has a more inhibiting
influence in the US than in Europe, perhaps also under the influence of the unions, so that
generally Europe is more test-minded.
6. I said earlier that in the US so much technology was put into paper-based testing that this
has been perfected very much. That has made the transition to computer tests more difficult:
the law of diminishing returns. Due to all the differences between the countries in Europe that
is not such an issue and the typical volume arguments do not get a chance. In some cases in
the US computer tests have reverted back to paper.
7. In computer tests in most other European countries serious attention is paid to the quality of
the test. In the Netherlands for instance there is COTAN, a commission set up by the
professional organization of Dutch psychologists, which publishes comments on all tests used
in the Netherlands, whether they are asked to do so or not. The Dutch Ministry of Education
only allows the use of tests in educational institutions that COTAN has evaluated at least as
passable. We consider the psychometric foundation for a test extremely important. My
impression is that the approach in the US is often inspired more by the consideration of
liability, in other words what conclusions can be derived from a test without the publisher
having to fear claims. Such an approach in principle is more negative.
These seven differences are intended as an appeal to create points of departure that will offer
the best opportunities to the trends and developments in test technology, in order to maintain
the prominent position of the US. To that end it is important to create a broad base, well-
flowing learning lines, national standards and a positive test culture.

How close the future can be will depend of the level of investments and their focus. There are
two main sources of financing for test innovation. The first source is via subsidies.
Governments and international organizations often grant project-based subsidies to stimulate
educational innovation. It is fairly rare that fundamental infrastructure innovations are
supported by government funding. The second source of funds is private investors, the test
companies. It is obvious that the latter group will have to supply the most important part and
will have to take the initiative. The drawback however now is that the efforts are mostly
product-oriented and that there is little cooperation. Let us take a quick look around the world
to see where the largest innovations are to be expected.
1. Eastern European markets. The former Soviet republics, the former Easter Block countries
and the republics of the former Yugoslavia constitute an interesting and enthusiastic market
for Western test companies. Since many of the projects are subsidized by the World Bank,
OECD and the European Union educational reform is mostly the focus and there are no
commercial initiatives.
2. Emerging markets. The countries in the Middle East and the Far East constitute a group
that is booming. In some countries the test culture is rooted in a strong, ancient tradition, like
in Japan and China. But there are no commercial parties in place with sufficient financial
strength for major investments. Moreover the changes are happening slowly. In African
countries no fundamental innovation is taking place yet and there are no strong private test
institutions either.
3. Western European markets. In the Western European countries there are mainly non-profit
exam organizations, both on the education side and the training side. These organizations
usually have a narrow financial position and therefore cannot invest much. If they have been privatized by the government, they often have to link up with a larger company. The exceptions to the rule are two large Anglo-American combinations (Pearson and Reed Elsevier) and of course also Cito, which is a fully commercial enterprise and is currently investing ten percent of its annual turnover in innovation of test technology and market development. There are no other large test companies in Western Europe.

4. The American markets. Particularly in North America the market for tests and assessment has been developed very much commercially. Besides the two Anglo-American combinations Pearson and Reed Elsevier, ETS and Prometric are the major parties.

All in all the conclusion is that the large providers who have the ability to make proper investments can be counted on one hand. Additionally the total market for making and administering tests is relatively small if you compare it with other industries and sectors. Of course there are adjacent industries, like the ICT sector that is also present here, which can play a role in the innovation of software, hardware and infrastructure.

As the five largest test companies we should combine investments together with some ICT companies, based on open standards and ‘open sources’, so we can go beyond merely improving efficiency: developing a new way of testing which will attract the candidates –who are what it is all about. The technology will then be demand-driven rather than supply-driven. We can pattern ourselves on other industries, like consumer electronics, where a group of competitors jointly developed the new standards like CD and DVD. Cito will be more than happy to participate in such a scheme.

There is no doubt that my predictions will come true; it only remains to be seen when. If we take up the challenge together, it will sooner be tomorrow than one day.

*Marten Roorda*

*Palm Springs*

*February 4th 2004*