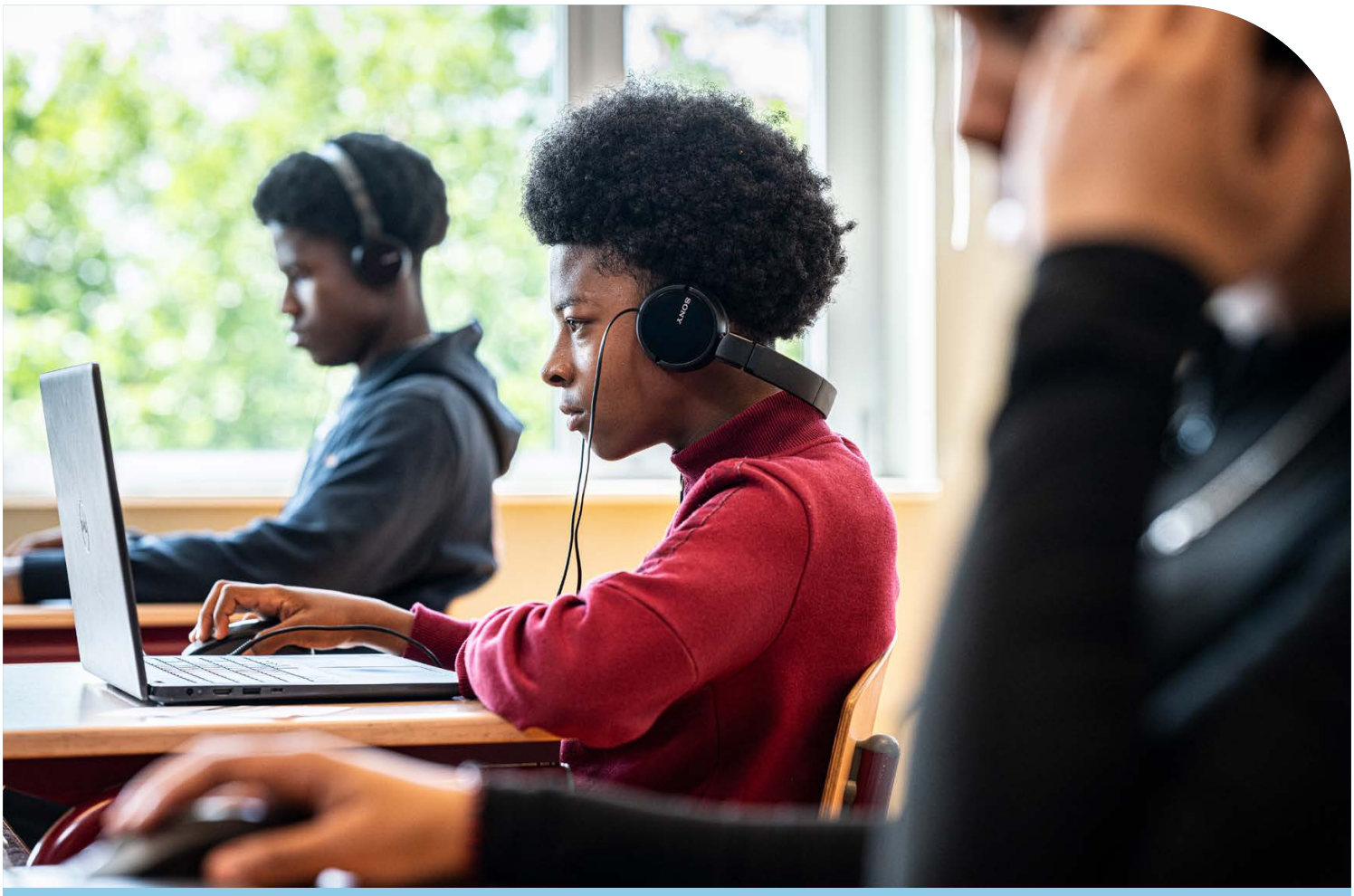
Government of the Netherlands



Sub-report

Digital examinations in Europe

An international comparative perspective

Programme line 1
Exploring new technological solutions

Digitalisering centrale examens

Digitalisering centrale examens

The Dutch programme for the digitalisation of centralised examinations is an initiative of the Board of Tests and Examinations (CvTE), the Education Executive Agency (DUO), Stichting Cito, and the Ministry of Education, Culture and Science (OCW).

Through the programme, the examination supply chain partners explore whether, how, and for which subjects and subject components the digitalisation of centralised examinations offers added value. Because this could make it possible to examine new skills in the long term and ensure that centralised examinations are more in keeping with students' experiences.

For more information about the programme and projects, see digitalisation of centralised examinations | CvTE (in Dutch).

There are three programme lines:

Programme line 1

Exploring new technological solutions for examinations and the examination supply chain

'What is technologically possible and desirable?'

Increasing knowledge of technological solutions and their assessment-related and ethical impact on the examination supply chain. This programme line focuses on learning. We look at the long term.

Programme line 2

Investigating ICT solutions for substantive innovations

'What is technologically possible and desirable?'

A practical manifestation of the current needs of examination and syllabus committees with regard to the design of the centralized examination and the use of digital examinations. We look at the medium term.

Programme line 3

Further development of supply chain infrastructure (ecosystem)

'What is technologically possible and desirable?'

The focus is on research into a modernised and future-proof ecosystem and on opportunities to optimise the existing landscape. We look at the short term.

Contents

Summary 4

Introduction 7

Method 10

Findings 14

- **Centralised examinations abroad** 15
- **Digitalisation - how long does it take?** 17
- **Reasons for digitalisation** 17
- **The assessment platform as the hub** 18
- **Subject-specific implementation route** 20
- **Platform availability as a key success factor** 21
- **Low-stakes pilot schemes** 21
- **Mathematics: an international challenge** 22
- **Expectation management and communication** 23

Discussion 25

- **Why digitalise, and in what order?** 26
- **Test early, often and frequently. Fail early, fast and safely.** 27
- **If you want to go fast, go alone; if you want to go far, go together** 28
- **An agile ecosystem: a balance between the supply chain and the market** 28
- **AI as a controversial issue** 29
- **Limitations of the study** 29

Conclusion 30

Literature 32

Appendices 34

Summary

The programme for the Digitalisation of Centralised Examinations (DCE) positions the Netherlands within a broader European movement in which digital technology is radically transforming the way examinations are designed, administered, marked and evaluated. This international comparative study has been designed to inform the DCE programme by examining how other European countries are digitalising their examination supply chain, the choices being made in this regard, and the lessons - and pitfalls - that may be relevant to the Dutch context.

This report focuses on centralised examinations and other forms of standardised assessment in Europe. Through semi-structured interviews with representatives from twelve countries, supplemented by desk research, insight was gained into platform choices and implementation strategies, the conditions for responsible digitalisation, and the challenges and success factors encountered.

One of the ways in which the Netherlands stands out in Europe is that digital centralised examinations are already in place for VMBO BB/KB, with flexible administration times, while other streams still use paper-based examinations. In countries where high-stakes examinations are conducted entirely online, they are usually held at fixed, centralised times. Countries also vary considerably in terms of the number of examination subjects per pupil, the size of cohorts and the weighting of the centralised examination in the decision on the award of a qualification.

As regards **the pace and drivers of digitalisation**, the comparison shows that many countries have only been undertaking large-scale digitalisation since the start of the current decade, often accelerated by the COVID-19 pandemic. The most frequently cited drivers are improving the quality of the assessment process, efficiency, logistical improvements and inclusivity/accessibility. Countries regard equality of opportunity, transparency and inclusion as powerful, socially recognised drivers of change. It is striking that the argument of 'assessing broader or more authentic skills' is less prominent internationally than it is in the Dutch ambition.

In virtually all countries, the **digital assessment platform** plays a central role. The interviews reveal two main routes: (1) develop an in-house platform (with the government as the owner) or (2) procure an existing commercial platform (with the government as the customer). Both routes have clear trade-offs. In-house management can facilitate autonomy, integration and open source, but requires a high level of development and management capacity. Procurement via a platform offers varying degrees of influence over the roadmap and requires clear agreements on governance, tendering risks and migrations. On the other hand, it offers speed and scalability. At the same time, it heightens sensitivities regarding public responsibility and dependence, and criticism of the platform often ends up being directed at the examination supply chain.

A second key theme is **the implementation route**: in what order are subjects digitalised? Countries are adopting a variety of strategies: starting with 'small' subjects, beginning with large cohorts as a stress test, digitalising all subjects in the same year to avoid hybrid overlap, or opting for a conscious mix of subjects (languages, science subjects, social sciences). Every strategy has its advantages and risks; the choice depends on risk appetite, available infrastructure, the political context and the technical complexity involved.

Many countries emphasise that **the availability of the platform** outside the examination period is a key factor for success. If pupils and teachers are able to use the platform throughout the year for practice tests and low-stakes applications, bugs, security risks and infrastructural limitations are more likely to come to light, and schools will develop a routine.

In a broader sense, practical examples from abroad show that digitalisation is almost always accompanied by teething problems. A flawless implementation is unrealistic. Success therefore depends heavily on the testing strategy, robust contingency plans, and transparent management of the expectations of schools, pupils and the wider community.

In this context, **the lack of low-stakes 'pilot schemes'** in the Netherlands is a point of concern. In several countries, there are standardised assessments that have little or no impact on pupils, allowing the ecosystem (platform, logistics, infrastructure and processes) to learn and improve under relatively safe conditions. These pilot schemes also help to bring about the necessary cultural shift and foster acceptance. The report therefore emphasises the importance of structured training and experimentation environments that simulate realistic production conditions.

A **recurring international challenge is the digitalisation of mathematics** and related science subjects. Questions regarding formula entry, the display of intermediate steps, validity (are you assessing mathematics or IT skills?) and acceptance by further education institutions make this a sensitive issue. Some countries are delaying the digitalisation of mathematics or opting for hybrid formats (partly digital, partly on paper), while Finland administers all mathematics examinations digitally and Flanders is encouraging the private sector to develop new functionalities or integrate existing digital applications. Mathematics is often a drawn-out problem that requires explicit attention, collaboration and iterative learning.

Finally, the report highlights **the importance of communication and governance**. Once the research phase is over, digitalisation takes on the distinct characteristics of an ICT project. Experience from other countries shows that trust can quickly be lost in the event of incidents or where there is a lack of transparency regarding problems. Early and wide-ranging communication through accessible channels, a clear and consistent message from partners across the supply chain, transparency regarding objectives in terms of fairness and equal opportunities, and a realistic explanation that start-up phases are inevitably accompanied by errors are recommended.

When it comes to AI, many countries are still in an exploratory phase and there is **limited practical experience of using AI in high-stakes assessment**. International cooperation, collaborative learning environments and the timely development of frameworks for the responsible use of AI can play a key role in preventing fragmentation and duplication of effort. This is very much in line with the interviewees' desire to forge closer ties in future and to engage in accessible dialogue on the complex issue of digital assessment.

In summary, the study shows that there is no single, universally best approach to digital centralised examinations; choices depend on the context. However, there is broad international consensus on a number of strategic issues that need to be addressed in good time. While these insights do not provide the DCE programme with ready-made solutions, they do offer concrete starting points for underpinning Dutch policy decisions in an informed and transparent manner.

Introduction

The programme for the Digitalisation of Centralised Examinations (DCE) places the Netherlands at the centre of a joint European movement in which digital technology is radically transforming the way examinations in secondary education are organised, designed, administered, marked and evaluated. The Dutch education system and examination system have unique features that set us apart from other countries. At the same time, there are also many similarities, some of which stem from joint European agreements on the value of qualifications at the end of secondary education. Against this background, it makes sense, at the start of the DCE programme, to look beyond our borders and investigate which solutions, working methods and safeguards work in other countries and why.

However, such a comparison requires a critical eye: differences in culture, educational culture, subject content, assessment methods, governance relating to examinations, and ICT infrastructure in schools mean that blindly copying successful approaches from abroad is not necessarily a fruitful strategy. We should therefore aim not to identify and directly copy approaches from other countries, but rather to gain an understanding of which characteristics are comparable, which are fundamentally different, and what this means for the extent to which policy choices made in other countries can be adapted to the Dutch context (Steiner-Khamsi, 2014). We can only achieve such an understanding if we look beyond superficial characteristics such as exit profiles and the number of examination candidates in our comparison, and instead also take into account differences in educational culture, motivations and informal dynamics within the education sector. The aim of this international comparative study into the digitalisation of examination supply chains in Europe is to develop a nuanced understanding of the similarities and differences, so that this analysis can inform policy decisions in the Netherlands.

The Dutch education system and examination system

Developing an understanding of the differences between the Netherlands and other countries begins with an understanding of the characteristics of education and assessment in the second stage of secondary education in the Netherlands. First and foremost, the Dutch secondary education system is characterised by a wide variety of streams (VMBO BB/KB/GL/TL, HAVO, VWO, PRO, VSO), each usually with its own learning objectives and examination syllabuses, with further options for choosing between different profiles within each stream. What is unique here is that there are pathways that have been using largely digital examinations for years on the Facet platform managed by the Education Executive Agency (VMBO BB/KB), while other pathways conduct all examinations using pen and paper (VMBO GL/TL, HAVO, VWO).

Dutch pupils sit final exams in a relatively large number of subjects at an individual level, particularly when compared with countries where the decision on whether to award a qualification is based on three or four core subjects (Van der Molen & Feskens, 2025). Furthermore, the combination of school-based examinations (SE) and centralised examinations (CE) plays a decisive role in the Netherlands in determining the content covered in the centralised examinations, with the result that skills such as writing, viewing and listening play a limited role in the CE. Another characteristic of the Dutch approach is that, on the one hand, central coordination, logistics and standardisation are handled by the Board of Tests and Examinations (CvTE), the Education Executive Agency (DUO) and Stichting Cito, while, on the other hand, teachers play a prominent role in devising questions and assessing their own pupils.

Finally, the Netherlands also stands out for its multi-layered governance structure, involving several parties in the examination supply chain (Ministry of Education, Culture and Science, CvTE, DUO, Stichting Cito). Taken together, all these characteristics paint a clear picture of why adapting foreign solutions to the Dutch context requires more than simply identifying superficial differences.

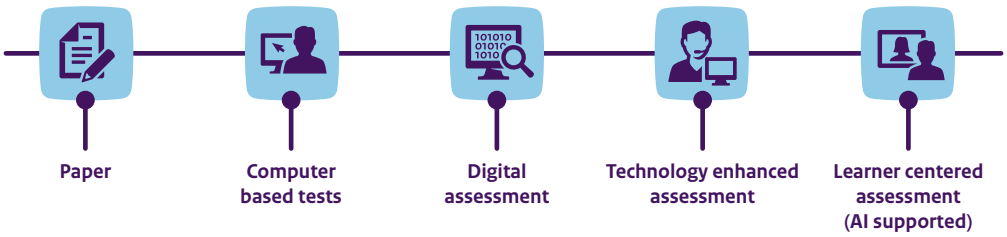
The value of international comparative research in a changing world

The overarching aim of the DCE programme is to develop examinations that are more flexible and authentic, and that assess a broader range of skills. At the same time, we are striving to create an agile examination supply chain that operates efficiently and effectively, takes the pressure off schools and guarantees safe and reliable assessments. This calls for a fresh perspective on what a desirable future for the examination ecosystem in the Dutch context entails.

International examples provide an important source of inspiration, but need to be carefully adapted to our specific context. It isimportant that we look not just at individual solutions or tools, but at all stages of the assessment cycle and how they work together toachieve the desired outcome, with or without the use of digital technology. We endeavour to take account of each individual phase ofthe assessment cycle, not only within the DCE programme as a whole, but also within this international comparative study.

Furthermore, the rise of artificial intelligence (AI) compels us to reflect on previous design choices in the structure of the examination supply chain, and on the question of which processes we want to automate, support or, conversely, keep in human hands. This means that digitalisation should not be viewed merely as technical modernisation, but as an opportunity to focus on the pupil and the skills to be assessed, and to provide scope for innovative solutions. In line with standard models relating to different generations of digitalisation in education and assessment (Bennett, 2015), we therefore refer to different stages of digitalisation in the DCE programme: from simply 'digitalising' paper documents to fully personalised and AI-supported experiences for pupils (see Figure 1).

Figure 1 – Stages of digitalisation identified within the DCE programme, based on comparable assessment digitalisation models (Bennett, 2015).



An international comparative perspective helps us to identify promising approaches and alternative options that have not yet been explored within our own system, while remaining open to ideas that relate to all phases of the assessment cycle and all stages of digitalisation. At the same time, looking at other countries makes us aware of potential pitfalls and of how contextual differences can influence implementation and effectiveness. Against this backdrop, the overarching objective of this study is as follows:

Objective: *By reviewing and describing how other European countries are shaping the digitalisation of their examination supply chain, we can provide the DCE programme with insights into existing and future projects abroad, suggest promising project ideas or warn of potential pitfalls.*

Method

This 'Method' section describes the steps taken to carry out this project. Some aspects of the research questions have been refined compared to the original project plan set out in the programme plan. A brief explanation is also given of the report's structure and where specific information can be found within the report.

The project began with four broadly formulated research questions:

- *Which countries in Europe are trialling digital examinations and AI?*
- *What are the legal and ethical frameworks in these countries?*
- *What lessons can we learn from their approach?*
- *How do these findings align with our national programme priorities and supply chain partners?*

These questions have been refined in consultation with the project team to provide a clearer picture of the context in which certain countries are shaping their digitalisation agenda. The decision was taken to focus on centralised examinations and standardised assessments.

The main questions and sub-questions used during the project are as follows:

Main question: What are the different European approaches to digitalisation in examination supply chains, and what insights from other countries can be relevant to the Dutch programme for the digitalisation of centralised examinations?

Sub-question 1: How is secondary education, and in particular the examination system and the examination supply chain, structured in the European countries studied?

Sub-question 2: To what extent do other European countries have a digitalisation strategy relating to centralised examinations or other forms of standardised assessment? How have these strategies been developed, and which parts of the examination supply chain have been digitalised? If there is not (or not yet) a digitalisation strategy in place, why?

Sub-question 3: How do the countries studied view their current situation with regard to the digitalisation of examinations, and what plans are in place to change or further develop this situation in the future?

Sub-question 4: What lessons can be learned from the experiences of the countries studied in terms of platform selection, technical implementation, and educational and societal integration?

Sub-question 5: What ethical, legal and educational conditions are regarded within the countries studied as prerequisites for the responsible use of digitalisation - and, in particular, the use of AI - within the examination supply chain?

To answer these questions, we have drawn up an interview guide (see Appendix 1). This guide was followed in all interviews, but was applied in a semi-structured manner, with the focus of each interview varying from country to country.

Both existing and new contacts were approached to take part in the project from the network within Stichting Cito and the communities associated with the international partnerships and conference organisations FLIP+ and AEA Europe. We then gauged interest in participating in the project through an initial round of no-obligation introductory meetings (see Figure 2). The interested organisations were then contacted and sent a consent form to confirm their participation in the project and their consent to the processing of their data (see Appendix 2). In the end, discussions were held with representatives from twelve countries (see Table 1).

Figure 2 – Visual overview of countries approached

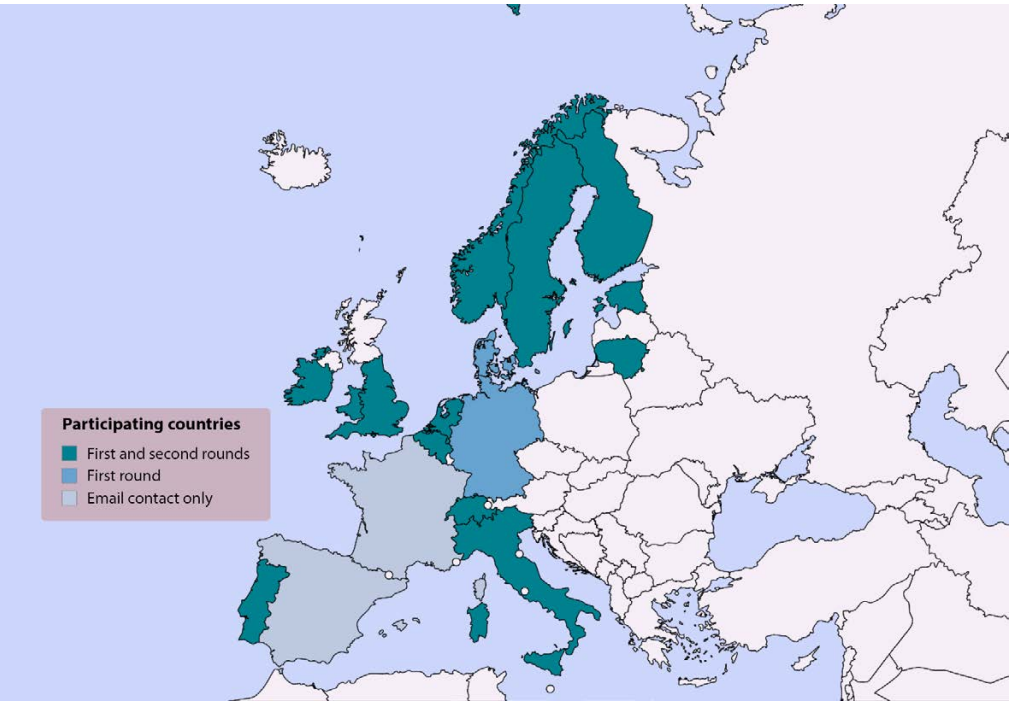


Table 1 – Participating countries and organisations

Country	Organisation
England	Ofqual
Estonia	Education and Youth Board of Estonia (HARNO)
Finland	Matriculation Examination Board
Ireland	State Examination Commission (SEC)
Italy	Invalsi
Lithuania	National Agency for Education (NSA)
Norway	Norwegian Directorate for Education and Training (Udir)
Portugal	Institute for Educational Assessment (IAVE)
Flanders	Ministry of Education and Training
Wales	Qualifications Wales
Sweden	Swedish National Agency for Education (Skolverket)
Switzerland	Institut für Bildungsevaluation (IBE)

All interviews were recorded in Microsoft Teams, automatically transcribed, and cleaned up where necessary. The transcripts were submitted to the interviewees for approval. Based on the interviews, supplementary documents or presentations were occasionally made available to the project team to support the topics discussed. Following approval, the relevant information was extracted from the transcripts, discussed within the project team and grouped into clusters. The most important findings from these clusters were selected for inclusion in the findings section of this report.

For the sake of clarity, the decision has been made not to include the responses to sub-questions 1-3 and the other findings from the interviews in this report. Appendix 3 contains a list of background information and other results that can be obtained from the authors. The education and examination system has already been described comprehensively and in detail in various reports and on various websites. For further information, we refer the reader to the Eurydice website, the publication The Structure of the European Education Systems 2023/2024, and the OECD's Education GPS.

Findings

This chapter sets out the main findings that emerged from the interviews in response to sub-questions four and five. These are all complex issues that require sufficient attention within DCE. The findings described below do not offer ready-made solutions, but they compel us not to shy away from complex questions and encourage an ongoing discussion about what is meaningful and important at this stage of the programme. They also offer a glimpse of what the future might hold for us.

The chapter begins with a number of descriptive findings regarding the differences and similarities in centralised examinations, the pace of the implementation processes and the reasons for digitalisation. We then describe scenarios of how other countries are digitalising their examination platforms, the order in which subjects have been digitalised, and a number of key factors in the success of their digitalisation agendas.

Centralised examinations abroad

The Dutch education system and its associated examination supply chain are just one of the many ways in which education and assessment can take shape. In Table 2, we describe a number of characteristics of the examination supply chain of the partners interviewed compared to the Dutch system. What is striking is that the Netherlands is the only country that makes a distinction at pathway level when it comes to centralised examinations, with VMBO BB/KB students sitting digital examinations while students in the other pathways do not. The flexibility in terms of when examinations can be taken that is associated with these digital examinations is not found even in the only two countries (Finland and Norway) that have already definitely introduced an all-digital examination process. There, examinations are held at set times. Another area in which the Netherlands stands out is that it, along with Ireland, has a high number of subjects in which examinations must be taken. It is also worth bearing in mind the size of each cohort of examination candidates. It is impossible to compare England and Estonia at system level without bearing in mind that the English system caters for 80 times as many pupils as the Estonian system. The relative weighting of the centralised examination, which varies between 20% and 100%, may also influence the difficulty involved in innovating systems or taking political decisions.

Table 2 – Overview of statistics and characteristics of centralised examinations at the end of secondary education in the countries included in this study. Sources: Eurydice, websites of national institutions, and interviews with contacts.

Country	Number of candidates	Number of subjects *(average per pupil)*	CE share + pass rate	CE administration	Decision on the award of a qualification *(including the lecturer's role)*
The Netherlands	**180,000-200,000**	**6-8**	**~50% CE; ~90% pass**	**Digital + flexible timing (VMBO BB/KB), paper-based + fixed timing (other)**	**Examination Board, based on CE and SE grades; the pupil's teacher acts as the assessor for CE and SE.**
England	800,000-850,000	3-4	~100% CE; ~97% pass	Mostly on paper + at a fixed time	Examination boards award grades; Ofqual regulates; the teacher's role is limited.
Estonia	10,000-13,000	3	~60% CE; >90% pass	Mostly on paper + at a fixed time	Board of teaching staff, based on CE, SE and mandate.
Finland	95,000-150,0001	6	100% CE; ~90% pass	Entirely digital + at a fixed time	The central national examinations board makes the decision; the teacher acts as the provisional assessor, and the external examiner as the final assessor.
Ireland	60,000-65,000	7	100% CE; 'failing' not possible	Entirely on paper + at a fixed time	Design, assessment and decision-making at central level.
Italy	500,000-550,000	3 (2 written, 1 interview)	~60% CE; ~99% pass rate (for Maturita examinations following admission)	Entirely on paper + at a fixed time, flexible interviews	The class council determines eligibility for the Maturita, based on attendance, SE grades, standardised assessments and compulsory assignments; the examination board, which includes the pupil's own teacher, makes the final decision.
Lithuania	25,000-30,000	3-4	~100% CE (for the Matura certificate); ~90% pass	Mostly on paper + at a fixed time	Design, assessment and decision-making at central level; teachers can act as assessors.
Norway	60,000-70,000	4+ (at least 3 written, 1 other format)	~20% CE; ~90% pass	Languages mainly digital, science subjects not digital, always at fixed times	Teachers play a very significant role in the decision on the award of a qualification, due to the high proportion of SE, CE is organised entirely at central level.
Portugal	90,000-100,000	3	30% CE (for a limited number of subjects where a CE is available); ~90% pass	Entirely on paper + at a fixed time	The class council decides on the award of a qualification; teachers play a major role in the SE and as members of the class council, while the CE is organised entirely centrally.
Flanders (Belgium)	225,000	4	0% CE; pass N/A	No centralised examinations	Every school has its own graduation procedure. The Flemish standardised assessments must not play a decisive role in this decision.
Wales	30,000-35,000	3-4	Varies by subject, ~60% CE; ~97% pass	Mostly on paper + at a fixed time	Examination boards award grades; Qualifications Wales regulates; the teacher's role is limited.
Sweden	95,000-100,000	N/A	0% CE; pass N/A	No centralised examinations	The teacher proposes a decision on the award of the qualification based on the SE and grades in standardised assessments; the headteacher has final responsibility; teachers themselves play a relatively significant role.
Switzerland	80,000-90,000	5+ (at least 5 in writing, possibly also orally)	~40%-50% CE (depending on the canton); >90% pass	Mostly on paper + at a fixed time	Examinations are organised and regulated at cantonal level, but recognised at federal level; teachers play a major role in the SE, in marking the CE, and in deciding on the award of qualifications.

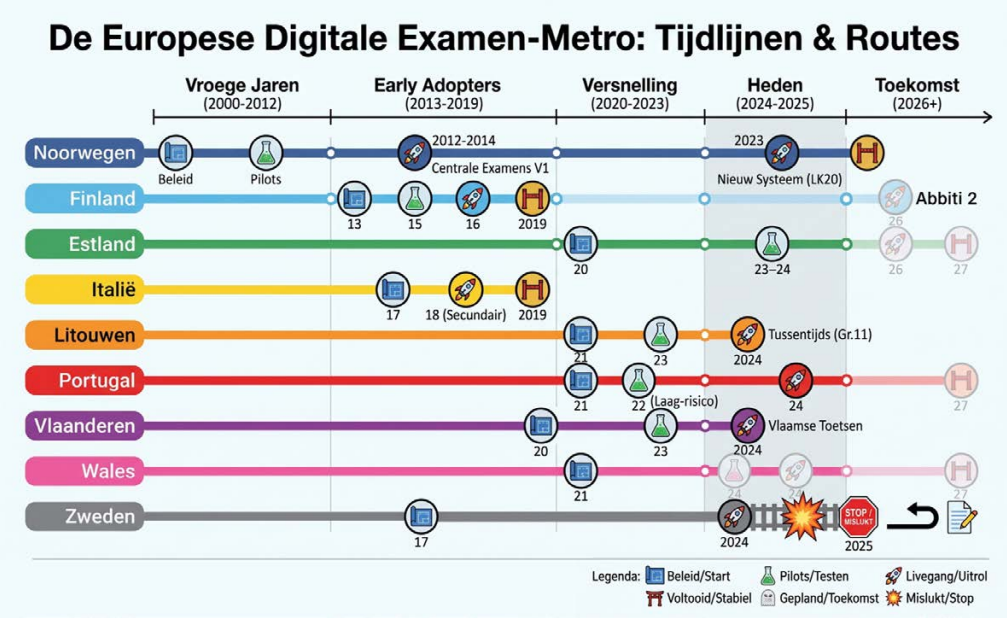
1 These figures represent the total number of examinations taken, not the number of candidates. The relatively large fluctuation in the number of examinations taken in Finland can be explained by the fact that, as there are both a spring and an autumn sitting, differences arise not only from smaller or larger cohorts, but also from the number of resits and the relative preference for the time of year in which the examinations are held.

Digitalisation - how long does it take?

To gain an insight at a meta-level into the progress of our interviewees' digitalisation processes, Figure 3 presents a timeline subway map based on the interviews, supplemented by desk research. An important addition to Table 2 is that this overview includes not only existing high-stakes centralised examinations, but also centralised examinations currently being piloted and *Standardised Assessments*, provided they have been digitalised. This means that timelines can be plotted not only for Norway and Finland, but also for other countries.

Norway is undoubtedly one of the pioneers in the field of digital assessment, having introduced its first system in the early 2000s. Finland was one of the first neighbouring countries to follow suit. However, it was only at the start of this decade that the digitalisation of examination supply chain actually really began to gain momentum. These countries are therefore now at the start or end of their pilot phases and are preparing for a roll-out in the near future. For the Netherlands too, the digitalisation of high-stakes examinations is of course not a new development. As early as 2009, plans were drawn up for Compex (computer-based examinations), which subsequently led to the piloting and roll-out of digital examinations, initially using Examentester and, since 2015/2016, the introduction of Facet.

Figure 3 – Timelines of international digitalisation projects (figure generated using Gemini and edited by the authors)



Reasons for digitalisation

The COVID-19 pandemic accelerated the transition to digital assessment and examination supply chains in Lithuania and Portugal. The sudden closure of schools led to substantial investment in digital infrastructure and digital examination systems in these countries.

Particular emphasis is placed on improving quality and organising the marking process in a fairer and more efficient manner. Lithuania, which experienced the longest education lockdown in Europe, is now gradually moving towards high-stakes assessment following successful pilot schemes involving low-stakes assessment in primary and secondary education. The marking process is now faster, making the progression to universities more efficient. It is also hoped that digital marking will help to alleviate the teacher shortage to some extent. In Portugal and Sweden, digitalisation has facilitated double marking and anonymous marking to improve the quality of marking. However, the reduction in workload that had been expected in Sweden failed to materialise, as administrative tasks within the digital platform offset much of the time saved.

Italy focused explicitly on efficiency and logistics with the digitalisation of the INVALSI assessments. The switch to centralised marking has sped up and standardised the process, which not only saves time but also provides more reliable data for research. The digital chain has significantly reduced the costs and errors associated with paper-based distribution and local marking.

Inclusivity is an important and increasingly central motivation. Estonia, Italy, Norway and Flanders emphasise that digital platforms can much more easily incorporate adjustments for pupils with dyslexia, ADHD, colour blindness and other common types of neurodivergence. Applying the WCAG 2.0 and Universal Design Assessment principles makes it possible to provide standard accessibility for almost all pupils: Norway makes listening tests accessible to pupils with hearing impairments by providing videos with subtitles, and offers British Sign Language as a subject for reading and writing. Flanders provides a standard read-aloud mode. It is crucial that inclusivity is embedded right from the design stage, rather than being added later as an extra feature. However, one remaining limitation persists: pupils with severe visual impairment or who are blind still do not benefit sufficiently, as many digital question types are highly visual in nature. The next question for these countries was, of course, how they would go about delivering their digital examinations.

The assessment platform as the hub

The digitalisation of the examination supply chain is often synonymous with the digitalisation of the assessment platform. Here we have identified two different scenarios through which we can facilitate the sitting of digital examinations:

1. Developing or commissioning an in-house assessment platform
2. Purchasing from a market participant

Both scenarios have their strengths and weaknesses. Each scenario is discussed below. It is important to note that this is not an exhaustive analysis of the available platforms.

1. The government as owner

Portugal has developed an open-source platform in-house, as this was funded by the EU and licences cannot be purchased for such software. Furthermore, the available solutions did not offer sufficient integration with the government's own scoring platform or the required backend features. The platform actively blocks all other applications, including AI tools. The platform can run offline on a local server at the school, as reliable internet access is not available everywhere in the country. Schools often lack the necessary expertise in this area and therefore turn to online resources.

In Estonia, the platform (EIS) is closely integrated with the rest of the public sector. Link healthcare and other government-related services to the same platform. In practice, it proved neither feasible nor desirable to forego this integration. In 2012, when development began, Estonian policymakers believed that there were still not enough high-quality, reliable commercial providers of assessment platforms. You log in to the system using your ID card. Before sitting high-stakes examinations, a student must have their ID verified and must sit the examination only on a pre-registered device. The assessment environment is available all year round, but in practice teachers rarely use this functionality. Pupils can only sit examinations if their teacher grants them access.

In Finland, the Matriculation Examination Board owns and manages its own Abitti system, but engaged external expertise through public tenders during its development. The platform is currently undergoing its first major overhaul. In both the original Abbiti and Abbiti 2, all examinations run locally on school servers. No internet connection was or is required during the examination, and every school had to run a server and a backup server. No pupil data was stored on personal devices. Abbiti 2 differs from the first version in that it features a server-level internet connection to the Matriculation Board, enabling support to be provided during the examination. Finland also appears to have the most comprehensive contingency plan. Back-up laptops and batteries are mandatory, and a school must be able to set up a wired network within a reasonable time if wireless connections fail.

2. The government as customer

In this model, the government opts for a market-based solution and pays licence fees for the use of the platform. In its tender, the government sets out what is expected of the assessment platform (mandatory and desirable functionalities) and what is expected of the collaboration with the market operator. Once the contract has been awarded, technical responsibility lies outside the public sector, and the maintenance costs are borne by the commercial partner. In this model, every new tender round poses a risk; the existing ecosystem may have to migrate to another provider. A transition period can be contractually agreed for this purpose, but the consequences of a full system migration are significant.

Within this model, there is a certain bandwidth in the relationship between the government and the market operator. The Flemish government opted for a software-as-a-service (SaaS) application to administer examinations. As customer, they have a say in, but no control over, the platform's development roadmap. The requirements for the Flemish standardised assessments are considered in close consultation. The Flemish Government may commission bespoke solutions provided that they fit within the market operator's roadmap. However, the market operator retains the right to reject requests if it considers them to be inconsistent with its strategy. The market operator also retains the right to make independent decisions to procure certain functionalities from the market rather than developing them in-house. The GeoGebra plug-in for mathematics and arithmetic is an example of this.

New features developed at the request of the Flemish Government (tailor-made solutions) are automatically also made available to other customers. At the same time, features requested by other customers are also available to the Flemish Government. Where the partner sees a general benefit in feature requests, it also develop these at its own expense. The associated costs may therefore be factored into increased licence fees.

It is important to set out clear arrangements regarding the collaboration within the tender. In Flanders, a high-level steering group has been set up for this purpose, comprising representatives from the government and the market operator.

If, as a customer, you choose to purchase a licence from a commercial provider rather than putting the project out to tender, this does not mean that you have absolutely no direct say in the platform's development roadmap. Countries including Lithuania, Sweden and Italy have stated that functionality has been developed based on the requirements they discussed. Although it is not possible to fulfil every request (or in good time), they would never have been able to develop such a platform themselves. In this case, it seems to be less of a partnership and more of a traditional customer relationship. However, in the public debate, people often fail to realise that this is a purchased platform, and it is the authorities themselves who have to deal with all the criticism regarding its capabilities and limitations. Even though the examination supply chain shares the public's frustration. Lithuania is also unhappy that the decision to opt for TAO is causing public finances to drain resources from its own economy.

Alongside the choice of platform, deciding the order in which subjects are to be digitalised plays a key role in shaping the digitalisation agenda.

Subject-specific implementation route

Lithuania introduced all subjects in the same year. In hindsight, they might have preferred to have done things differently, and they advise the Netherlands to start with the subjects that have the fewest candidates, so as to gain experience. In England, Ofqual put forward a proposal in 2026 whereby every exam board would be permitted to begin digitalising two of their 'smaller' qualifications for up to 100,000 pupils, which, as Table 2 shows, is still roughly eight times the total number of Estonian candidates. Initially, each exam board is likely to continue offering qualifications in both paper and digital formats, as there will be a market for both. The questions must not be the same, and pupils must not use their own devices. This proposal was not so much drawn up to facilitate digitalisation, but rather to establish rules for a cautious and phased roll-out of digital qualifications, as there is currently no framework in place from Ofqual that prohibits exam boards from offering digital qualifications. Ofqual does not envisage a complete transition to digital qualifications and continues to regard pen and paper as the standard.

Norway and Estonia, on the other hand, chose to digitalise subjects with the most candidates first. In Norway, one of the reasons for this was that the examinations are set by the universities, which meant that the largest universities were involved in developing the first cohort of subjects.

One vision, one message, one agenda

In Wales, the approach to curriculum reform and the associated digitalisation is very similar to the way we collaborate within the chain in the Netherlands. Every four weeks, during the design and implementation of the new curriculum, all the relevant bodies meet to discuss feedback on the programme and coordinate their response to the public's reaction to relevant topics. They organise a joint webinar and work together to produce and publish materials. This ensures that schools, teachers and pupils receive the same answer to the same question from every organisation, thereby preventing confusion in the debate.

Finland and Wales have opted for a balanced approach, implementing subject packages that divide subjects into languages, science subjects and social sciences. Table 3 shows the implementation route followed in Finland. The idea was that each type of subject has its own challenges and that experience across the different subject types was to be gained as quickly as possible. Finland has two assessment periods each year, which meant that new subjects could be introduced on two occasions during each calendar year.

Table 3 – The order in which the various subjects in which Finnish pupils can sit their final examinations have been implemented in digital form.

Autumn '16	Spring '17	Autumn '17	Spring '18	Autumn '18	Spring '19
German	French	Swedish	Biology	Physics	Mathematics
Geography	Society	Worldview	Spanish	Chemistry	
Philosophy	Psychology	Finnish	English	FT2/ZT2	
		History	Italian	Sami	
		Health	Latin	Russian	
		Religion	Portuguese	Mother tongue	

One of the next decisions to be made once it is known which subjects will be transitioning from paper to digital and at what point in time is whether or not to make this platform available to schools, teachers and pupils for the entire academic year.

Platform availability as a key success factor

Many countries use their low-stakes examinations to familiarise secondary school pupils with the assessment platform. Finland emphasises that the key factor for successful implementation is likely to be the opening up of the platform. This means that teachers and pupils have access to Abitti throughout the year to develop and sit tests. This approach provides developers with early feedback on features, potential security risks and bugs that would otherwise only come to light during the high-stakes examination.

This ensures the technical reliability of the platform and means that pupils are well prepared for their high-stakes assessment. Abitti is used solely by students aged 17 to 20 for centralised examinations, which means the platform can be tested specifically among young adults and does not need to be suitable for young children as well. The platform is updated monthly, meaning it functions more as a market solution than as a measuring instrument that is dusted off twice a year.

As soon as the Flemish standardised assessments were introduced, Flanders began integrating the platform into schools. Placement tests have been developed to enable all pupils, in conjunction with comprehensive introductory and instructional videos, to complete the onboarding process on the platform.

In Lithuania, the assessment platform is also available to teachers and pupils all year round. Here, pupils start using the assessment platform from the age of 10 onwards, in preparation for the high-stakes examinations they will sit at the age of 15-16. However, the roll-out in these classes was too abrupt, as all subjects were introduced all at once. As a result, they realise in hindsight that they should have devoted more time and energy to the implementation process, and could have introduced the modules in a more phased manner. The curriculum and the assessment process are now being digitalised in parallel, ensuring that pupils are not faced with any surprises during the assessment.

In contrast, in Sweden the platform was only available during the assessment period. This was one of the reasons why major problems arose during the assessment period, leading to the digitalisation programme being temporarily suspended. In Wales, with its various commercial providers of qualifications and the freedom schools have to design their own curriculum, the government does not wish to use a single assessment platform. It depends on the agreement between the school and the exam board as to whether and how the platform is available throughout the year. In Switzerland, IBE, as a commercial organisation, makes its platform available to customers in whatever way the school wishes, but in this respect is difficult to compare to the government bodies responsible for organising and conducting high-stakes examinations.

Making the platform available can be relevant in determining the level of acceptance within the education system. Furthermore, the interviews reveal that other countries often move from low-stakes to high-stakes implementation, in which case a secure test environment is important.

Low-stakes pilot schemes

Experience in other countries shows that the digitalisation of examinations is almost always accompanied by teething problems, software errors and logistical challenges. That is why it is crucial to manage internal and external expectations effectively, to involve a wide range of stakeholders, and to provide clear information about the development and implementation process. In addition, a thorough risk analysis is required, setting out the likelihood and impact of various incidents on the entire examination supply chain, so that robust contingency plans can be drawn up for all risks.

Many of the participating countries have an advantage over the Netherlands in that they have Standardised Assessments that are used solely by the government to collect data on the state of education and provide schools with an insight into their pupils' attainment levels. For the pupil themselves, however, the result of this assessment has no consequences and it is therefore a low-stakes assessment. The consequences of errors in this context are considerably less serious than when pupils are unable to sit their final examinations or their examination and assessment records go missing. These low-stakes assessment opportunities are excellent pilot schemes for the education ecosystem to learn how to navigate the digital transformation. The platform can be used for practice: teachers, pupils and examiners can run through their logistical processes, and the infrastructure can be mapped out and stress tested. Last but not least, it supports the cultural shift whereby the education sector can become accustomed to the idea of digital assessment and recognise the benefits of a wider range of question types, faster marking and streamlined administrative processes.

Estonia, Finland, Italy, Lithuania, Portugal, Flanders and Wales have all opted for an approach based on low-stakes assessments. In Estonia, Year 12 pupils have, in fact, been sitting both the paper-based and digital examinations side by side for several years now, in order to identify similarities and differences and to optimise the digital examination through an iterative process. In 2026, the final pilot scheme in the lower years is expected to involve 10,000 pupils studying Estonian and English. In England, where there are currently no plans to introduce high-stakes digital examinations, Ofqual nevertheless encourages small-scale trials to explore the possibilities and limitations.

With regard to the implementation issues discussed during the interviews, it is important to include a realistic disclaimer: regardless of how long or short the preparation period may be, unforeseen circumstances will always arise during pilot phases and during the final, high-stakes roll-out — ranging from minor teething problems to more serious technical or logistical issues. A flawless implementation is unrealistic: the degree of uncertainty must be commensurate with the examination supply chain's risk appetite.

Even when low-stakes pilot schemes are facilitated, this does not necessarily mean that every subject can be automatically digitalised. Other countries still see this as a major challenge, particularly in the sciences.

Mathematics: an international challenge

It is clear that there was a good reason we launched a project within DCE specifically for mathematics at an early stage: in every country, this subject plays a distinct, often sensitive role in digitalisation. The experiences reveal a wide range of choices and recurring concerns. And the results appear to apply by extension to other science subjects as well.

In Norway, mathematics was initially considered one of the top three subjects to be digitalised, but was ultimately dropped. There was intense pressure from the universities that it was no longer possible to demonstrate digitally that school-leavers were able to meet the strict entry requirements. The lack of technical capabilities for using formulas and the inability to visualise the thought processes a pupil goes through proved to be insurmountable obstacles. There is also a great deal of scepticism in Estonia and Lithuania: Estonian maths teachers fear that complex skills cannot be assessed adequately, and pupils say they prefer paper. International developments are being closely monitored in the search for a suitable solution. In Lithuania, there is an ongoing debate as to whether digitalisation measures genuine mathematical competence or primarily ICT skills. Following numerous discussions with providers of mathematics and arithmetic software, Wales has provisionally decided not to digitalise mathematics for the time being, as there is no standardised, widely accepted solution for inputting mathematical formulas and functions.

Finland is at the other end of the spectrum: maths there is fully digitalised. That decision followed a national decision to digitalise all examinations, despite a definite lack of unanimous public support for the digitalisation of this subject. Although there were initial concerns about a loss of measurement capabilities, this debate died down once the system came into use. Flanders is also actively committed to developing functionality to digitally assess complex mathematical skills among final-year pupils; the platform is being enhanced with new features, and pilot schemes will determine in the coming period which aspects are practical and user-friendly.

Other countries are opting for hybrid models or phased approaches. In Portugal, a hybrid model has been introduced: multiple-choice questions are completed digitally, while open-ended questions are answered on paper and are subsequently digitalised and marked. Sweden currently administers and marks examinations on paper, but allows teachers to enter results digitally into the system used for other digital examinations.

Whether or not mathematics is to be digitalised in the (near) future, the digitalisation of the examination supply chain is a major change. As with any major change, this process also benefits from timely, targeted communication that engages the relevant stakeholders and society at large. It must be possible to explain the how and why in a transparent manner.

Expectation management and communication

Following the research phase, the digitalisation of centralised examinations takes on the characteristics of an IT project to a large extent. According to Ofqual, digital examinations are more complex than paper-based ones: in their view, a flexible system always entails risks. These risks can be analysed and mitigated in advance, but cannot be completely ruled out. Ofqual therefore emphasises the importance of planning for potential incidents and putting appropriate control measures in place. It is wise to be open both about the known risks of a digital approach and about the limitations and errors inherent in the current paper-based process. Transparency helps to manage expectations realistically and prevents unrealistic demands or assumptions from leading to disappointment later on.

This requires a well-thought-out communication strategy: involve relevant stakeholders and keep the wider public informed about the changes as they unfold. Presenting a ‘perfect’ off-the-shelf solution is usually counterproductive, as experience from other countries shows that this can undermine trust and create unnecessary pressure. Sweden and Norway found that errors, or a failure to communicate such errors in a timely manner, can quickly undermine the trust of permanent stakeholders in particular, such as schools, government departments and interest groups. Pupils generally experience less of an impact in the short term, but for the sector as a whole, loss of trust has long-term consequences and can hinder innovation. In Sweden, where people were already acutely aware of the need for clear communication during the interview, the situation still ended up blowing up later.

Real-world examples show that even after extensive testing, minor problems can still arise: software is constantly evolving and is rarely ‘finished’. In Flanders, for example, during the first two years of the assessment programme, a very small group of pupils (n=125) encountered problems when calculating proficiency scores. Thanks to timely and clear communication, and because these were low-stakes assessments, the consequences were limited. This highlights that careful communication, robust contingency plans and an iterative approach are essential when introducing digital examinations.

Communication tips from abroad

It will come as no surprise when we say that there needs to be frequent, careful, timely and wide-ranging communication regarding the Dutch digitalisation strategy. Nevertheless, the various countries provide a number of guidelines that are relevant to our programme:

- A great deal of time and energy goes into translating government policy for implementation in individual schools. Guidelines with concrete checklists still need to be applied by people in a gymnasium in their own context. Show that you understand this and offer support where necessary.
- Explain why the transition brings benefits. That also means communicating about what is not going well at the moment, otherwise you will inevitably run into the 'never change a winning team' mentality.
- Start communicating via popular forums as early as possible, not just in trade journals and policy papers. Communicate in an accessible way about plans, projects, studies, results and specific implementation steps. Allow stakeholders to grown alongside the process.
- Communicate transparently about the aims of the change in terms of equal opportunities and fairness.
- Communicate with humility, patience, good communication skills and a willingness to embrace change, particularly when dealing with teachers, some of whom may instinctively be opposed to innovation or digitalisation.
- How does the digitalisation of centralised examinations compare to the digitalisation of society as a whole? It is perfectly possible to explain that digital centralised examinations are not a factor in young people's smartphone addiction. One presents an opportunity; the other is a problem.
- Be prepared to explain that you will need a number of start-up cycles during which there will be minor issues. The supply chain will do its utmost to minimise these issues or resolve them as quickly as possible.
- Counter negative emotions in the public debate as quickly as possible, otherwise they will dominate the discussion. Engage with people on social media and create a space for discussion about the pros and cons.
- Put the innovation of digital examinations into perspective. Everyone has been using computers for well over 20 years, but when it comes to examinations, it is suddenly seen as either impossible or highly innovative.
- If the system crashes, it is stressful for teachers and invigilators. Be honest about the potential problems that may arise and what the solutions are.

Discussion

A clear key message emerges from the findings: there is no single best way to digitalise centralised examinations. After all, many choices are highly context dependent. At the same time, there is broad consensus on a number of complex issues that it is essential to address in a timely manner: the choice of assessment platform, the order in which subjects are digitalised, the technical and pedagogical handling of formulae and calculations (with mathematics being a recurring bottleneck), and the organisation of realistic and safe experiments while not everything has yet been digitalised. Identifying these issues, and how other countries deal with them, is in itself a valuable outcome. In this chapter, we aim to provide practical guidance for the Dutch context regarding these complex issues, by linking insights and lessons from neighbouring countries to Dutch practice, which is sometimes unique. We do this by presenting possible starting points for future projects as trial balloons at the end of each topic discussed.

Why digitalise, and in what order?

From our discussions on the reasons behind digitalisation, it is striking that the ability to assess a broader range of skills plays only a very limited role. While there is frequent discussion in the Netherlands about the added value that digital assessments can offer in identifying new content in innovative ways, we do not hear this line of reasoning cited as a key driving force in other countries. Nor does a closer alignment with current educational practice and the professional field, where digital resources are already widely used, usually appear to be a primary motivation for digitalisation. One explanation for this is that the first step is often seen as the biggest one. As the substantive and learner-centred benefits of digitalisation generally only become apparent much later, it is easy to lose sight of these potential advantages amidst the initial enthusiasm for digitalisation. It seems that the current DCE programme strikes a good balance in this regard, with programme lines that address both specific infrastructure needs and substantive issues for the future.

At the same time, it was certainly pointed out in interviews that equality of opportunity, transparency and inclusion are issues affected by the digitalisation of assessments and are of great social importance. Both as a starting point for change and as a basis for clear communication, these are issues that deserve to be high on our agenda. In the Netherlands, we seem to have a relatively wide range of reasons for digitalisation. This is one of the strengths of the Dutch education system and the examination supply chain, but we must ensure that an important issue such as equality of opportunity is not overshadowed by a fragmentation of attention.

The reasons for digitalisation are closely linked to the decisions taken in each country regarding the order in which digitalisation is carried out. In our interviews, we therefore observed a wide variety in terms of which subjects are digitalised first and which last. In practice, there are broadly speaking four logical deployment strategies, each with their own advantages and disadvantages. For example, you could start with the low-hanging fruit, such as smaller language subjects with lots of multiple-choice questions. This means that the initial implementation steps are relatively straightforward and safe, but you do run the risk of putting off dealing with problematic cases. Some countries choose to start straight away with a major subject, such as the mother tongue. This immediately provides a good stress test for the system and highlights the benefits of large-scale digitalisation. At the same time, you increase the risk of large-scale logistical and operational problems. A third option is to digitalise all the subjects in one go. This strategy helps to avoid a lengthy overlap between paper-based and digital workflows. It should be clear, however, that this is a strategy that is extremely demanding in technical, organisational and substantive terms.

Finally, there are countries, including Finland and Wales, which opt for a subject-based phased approach. This could include carrying out annual digitalisation projects across a carefully considered mix of language, science subjects and social sciences. In the Netherlands, where substantive arguments carry considerable weight, it is tempting to explicitly include subject content in the timeline and to link this to the updating of examination syllabuses. Nevertheless, this strategy calls for a degree of caution. We see, for example, that in other countries the most challenging areas of digitalisation, particularly mathematics, are often addressed later in the implementation process. If such subjects are addressed at an early stage in curriculum reform and the updating of examination syllabuses, this does not necessarily mean that it makes sense to take the first steps towards digitalisation in those subjects at that stage.

In any case, based on the principle of constructive alignment (Biggs, 1996), it is advisable that the teams responsible for updating the examination syllabuses and the DCE programme understand one another and are aware of each other's plans and concerns. For the time being, it is still too early to identify a single deployment strategy for digitalisation as the obvious choice. It therefore makes sense to first draw up and cost various deployment strategies, while regularly seeking to align the two programmes, so that policy decisions can be made in a more informed and transparent manner.

Test early, often and frequently. Fail early, fast and safely.

In the Netherlands, there is no obvious low-stakes alternative to the centralised examinations with which to experiment as part of the DCE programme. Unless some of the school examinations can serve as a pilot scheme, we would be well advised to organise alternative solutions that allow us to learn in a practical setting without any mistakes having large-scale negative consequences. The countries that are further along in the roll-out process are acting as a canary in the coal mine: not a single country reported a completely flawless go-live. In the Netherlands, too, we know from previous attempts to digitalise examinations that unforeseen challenges always arise. It is important to realise that organising the administration of digital examinations is primarily an IT project, rather than an educational one. Traditional educational questions regarding content, validity and reliability remain key, but the organisation of a digital assessment nevertheless raises primarily technical and logistical issues for the supply chain.

This means that it is not a given that we will be able to continue to approach the organisation of the examination supply chain in the same way in order to ensure a successful digital examination for all streams. The way in which Finland updates and tests the Abitti assessment platform in practice, while maintaining a degree of flexibility, brings significant benefits. This makes it easier to organise quality monitoring of the platform, provides a clearer picture of the suitability of infrastructure at schools, and offers more points of reference for monitoring and assessing the competence of schools, teachers and pupils. This flexibility has its advantages, but also its disadvantages. Software is never finished: the architecture requires constant updates, which must be tested each time and which, to a greater or lesser extent, entail a risk each time. In a sense, therefore, such an approach also requires a different risk appetite on the part of the examination supply chain.

An Ofqual study into digital examinations, in which they spoke to representatives from countries including Finland and New Zealand, provides further justification and support for this point:

"In all cases those leading their implementation had a clear understanding of the potential barriers in their local context and detailed risk management mitigations and countermeasures. There was also a clear risk appetite and an understanding there would be some learning by doing so not everything would be perfect the first time. The latter point was particularly evident in New Zealand and Finland where there was an expectation and an acceptance by sponsors that things would go wrong but there were robust plans to deal with these issues arising, and an understanding that there are also failures in existing arrangements." (Online and on-screen assessment, Ofqual 2020)

The quote above highlights, on the one hand, the importance of the aforementioned risk appetite, but, on the other hand, also the importance of drawing up and developing robust plans to tackle, in a timely and appropriate manner, the problems arising from the assumption of risks. During our interview, Ireland emphasised that many problems only become apparent after implementation, even though they had been taken into account during the preparatory phase. We will therefore also need to have the courage to make the transition from test environments to production environments during the programme. This will enable us to identify any vulnerabilities at an early stage and allow us to work on solutions through the DCE programme - solutions which, in many cases, would require too much time and money to implement within standard production processes. All the more reason, then, to test early and often.

If you want to go fast, go alone; if you want to go far, go together

One of the most tangible outcomes of this project is the clear desire among almost all those interviewed to maintain close ties with the Netherlands and with one another. Existing conferences (such as AEA-Europe and FLIP+) go some way towards meeting that need, but do not always offer the accessible, frequent exchanges that organisations need right now. The openness with which countries shared their challenges and mistakes, and their willingness to answer questions even from a distance, is an asset that must be actively nurtured. By maintaining regular contact, for example through targeted working visits, thematic meetings or online webinars, this report will not be an end in itself, but rather the starting point for continuous learning at European level.

For the Netherlands, there are a number of specific cases where it makes sense to join forces with other countries. For example, several countries are currently working to address the challenges surrounding digital assessment in mathematics (for example Estonia, Finland, Norway and Wales). There is also plenty to learn from one another in other areas, such as: the structure of the assessment process (own teacher, another teacher, centralised), permitted devices (BYOD, school laptops, government-issued laptops), authentication and login (separate login per system, MFA, SSO), and the provision of practice environments (standardised assessments, school examinations, school-specific assessments). Maintaining accessible contact with other countries gives us the opportunity to broaden our knowledge of such topics in a pragmatic and cost-effective way. At the same time, this external perspective may also prompt us to reflect on fundamental choices regarding the balance between validity, reliability, transparency and practicality in the Dutch centralised examinations.

An agile ecosystem: a balance between the supply chain and the market

The decision on how commercial parties fit into the examination supply chain ecosystem is not merely a technical or financial one, but is directly linked to issues discussed earlier: choice of platform, risk appetite, scope for experimentation and deployment strategy. Vendor lock-in is a real risk, both with external suppliers and with seemingly internal solutions, and can lead to reduced innovation, high costs and undesirable washback effects on educational practice. At the same time, commercial providers often offer the scalability, development speed and specialist expertise that the supply chain requires. What is important, therefore, is not whether the market plays a role, but how that role is shaped. Transparency regarding ownership and responsibilities, clear exit routes, interoperability through standards, and robust data and continuity requirements should be standard components of any decision in this regard.

Our interviews have shown that each country chooses its own approach to cooperation with market operators. Which route is best suited to the Netherlands depends on what we prioritise from an assessment expert's perspective, but also, undoubtedly, on our willingness to accept risk and the operational capacity within our own examination supply chain. Looking back at our previous recommendations, it is clear that only by making clear choices and carrying out early testing can you effectively mitigate the challenges involved in working with market operators, while at the same time allowing those operators the scope to contribute to high-quality assessment.

AI as a controversial issue

The interviews conducted for this study reveal that many countries are still only making limited progress on issues and policies relating to AI and assessment. None of our interviewees are explicitly focusing on developing AI expertise, and thinking about AI is still in its infancy almost everywhere. This means that the Netherlands could soon become, or perhaps already is, a frontrunner when it comes to knowledge of AI and assessment. At the same time, it is fair to say that there is still considerable uncertainty regarding the effects, acceptance and feasibility of AI applications in the field of assessment.

The AI Regulation ensures that issues relating to AI and assessment are not only substantive and ethical in nature, but also legal. This calls for timely and structured consideration of the preconditions we set within the examination supply chain for the use of AI, including in the areas of autonomy, privacy and copyright. Based on the interviews we conducted, we can conclude that every country is still finding its way in this regard. That is precisely why cooperation at European level is crucial. Common standards and shared insights prevent us from having to constantly reinvent the wheel in this ever-changing world, while of course allowing scope for differences in national interpretations and educational cultures.

Given the lack of extensive experience with AI elsewhere in Europe, the trial balloon set out below does not so much stem directly from our interviews, but rather builds on existing practices relating to the responsible use of AI (regulatory sandboxes2), and previous experience with international cooperation on assessment (FLIP+ item library3).

Limitations of the study

The semi-structured interviews that form the basis of this study are, in a sense, snapshots, often involving one or two individuals from a single organisation within a country's examination supply chain. We do not claim to provide a comprehensive or exhaustive analysis of every digitalisation process, every examination supply chain or every education system. The answers given are coloured by the perspectives and interests of the interviewees and do not always reflect the full national picture. Furthermore, interviews lasting 'only' ninety minutes cannot always do justice to the complexity of education and all the subtle cultural and systemic differences between the Netherlands and other countries. Nevertheless, we believe our findings offer valuable insights. In Europe, we often face the same complex issues and regularly find similar solutions in our search for a way forward. And, as has already been emphasised, the international network that has been built up and the clear willingness to cooperate may well prove even more valuable in the future than the insights gained at present.

2 https://www.rdi.nl/documenten/2025/03/25/vormvoorstel-regulatory-sandbox-ai-verordening

3 https://www.itemlibrary.org/login

Conclusion

The digitalisation of centralised examinations is not a straightforward process with a single, universal formula for success. There is no right or wrong. This project demonstrates that the examination supply chain in the Netherlands is faced with a series of complex, interlinked issues ranging from the choice of platform and implementation strategy to digital assessment in mathematics and facilitating safe experiments. These issues cannot be viewed in isolation from broader themes such as equality of opportunity, inclusivity, transparency and public support.

Although it may occasionally be tempting to put off dealing with difficult issues, experience from other countries shows us that tackling the biggest challenges in a timely manner is precisely what lays the foundations for a stable and future-proof examination supply chain. By starting this early on, we create scope to learn, test, fail and improve without the impact of mistakes becoming unnecessarily severe.

The Netherlands has a great deal to gain by actively maintaining and utilising the international network it has built up. The openness and willingness to collaborate demonstrated by partners offers a unique opportunity to share insights and solutions, develop common standards and help one another tackle new challenges, such as the use of AI. This network can act as a catalyst for innovation, quality improvement and risk management, provided we maintain it on a structural basis.

The key message is clear: to successfully implement the digitalisation of centralised examinations, we must not put off tackling the complex issues, but address them in phases, transparently and in collaboration with our partners. In this regard, it is important to maintain a balance between content, technology, policy and practice, while ensuring the supply chain remains agile. With an active, learning-centred and internationally connected approach, we can develop an examination supply chain that is ready for the future.

Literature

Steiner-Khamsi, G. (2014). Cross-national policy borrowing: Understanding reception and translation. Asia Pacific Journal of Education, 34(2), 153-167.

Van der Molen, P. & Feskens, R. (2025). *Examenstelsels – Een vergelijking van de examenstelsels in Belgie, Frankrijk, Duitsland, Finland en het International Baccalaureate* (Examination systems –
A comparison of the examination systems in Belgium, France, Germany, Finland and the International Baccalaureate). Published by Stichting Cito.

Bennett, R. E. (2015). The changing nature of educational assessment. Review of research in education, 39(1), 370-407.

Eurydice website. https://eurydice.eacea.ec.europa.eu/ last visited on 26-03-2026

Biggs, J. (1996). Enhancing teaching through constructive alignment. *Higher education*, *32*(3), 347-364.

Ofqual research on on-screen assessment: https://www.gov.uk/government/collections/research-on-on-screen-assessment

Appendices

Appendix 1 – Interview guide

Appendix 2 – Consent Form

Appendix 3 – Background information available on request from the authors

Appendix 1

Interview guide (in English)

Interview Guide Second Round EU Partners:

Italicized are a few less relevant questions, only ask if time permits.

General:

- Is there a digitalization agenda, and what does it look like?
  - Planning; what has happened, what still needs to happen, what has been reversed?
    - If there was a specific order of implementation (per subject or educational level), what was the reasoning for choosing this route?
  - From which direction did the desire for digitalization come?
    - Political
    - Societal
    - Educational professionals
- What are the main considerations for digitalizing within high-stakes or other forms of standardized examinations?
- Where in the examination chain is this digitalization located? And where do you see specific risks and opportunities for the use of AI?
  1. Construction of items and/or source material
  2. Assembly
  3. Administration
  4. Grading
  5. Psychometric analyses
  6. Reporting
  7. Evaluation
  8. Miscellaneous

Implementation:

- What do you identify as unique points of your education system, and do these influence the possibilities for digitalization?
- What are the biggest challenges of your education system, and do these influence the possibilities for digitalization?
- *What can we learn from low-stakes digitalization processes?*
  - *Do you see this as a testing ground for high-stakes examination? And what evidence would you need to take the step toward high-stakes?*
  - *What differences exist at the process level between high- and low-stakes exams to remain cautious?*
- What were the main bottlenecks for a proper rollout of digital central exams?
  - Looking back, how would you have done the implementation differently?
- What are the main forms of resistance you encountered, and how did you deal with them? Break down by:
  - Societal (parents and students)
  - Professional (teachers and schools)
  - Within your own organization
  - Political

Technology:

- Which platform is currently being used? What were the considerations for this choice?
  - Commercial or in-house?
  - BYOD?
  - If not yet applicable, are there any plans on this subject?
- Can students and teachers also use the (mathematics) examination environment during the rest of the year? During class, can the teacher create learning material with it, can the student practice independently in this environment?
- What about the platform is already satisfactory, and what would you like to see change over time?
- *If you had to make the choice for a platform again, what would your choice be?*

Ethical and legal:

- Have you had to make major legal decisions, make adjustments to laws and regulations, and if so, what were the major pitfalls and what went smoothly?

Closing questions:

- *What in the digitization process has been your low-hanging fruit on which there was (almost) no discussion it was a step in the right direction?*
- Suppose you have five years and a large project team, what would you start doing on day one?
  And what would you immediately start researching because you have the resources for it?
- Are there matters not yet discussed that you would still like to share with us?

Appendix 2

Consent form (in English)

Participant information 'Digitizing high-stakes exams in the EU'
Dear participant,

Thank you for your interest in taking part in this study. To participate in an informed manner, it is important that you gain insight into the design and procedure of this research. In this letter we inform you about the purpose of the study, how the study will be conducted, and how your data will be handled. We ask you to read this information carefully. If you decide to take part, we will assume that you agree to the conditions described.

What is the aim of the study?
This study focuses on exploring and inventory the extent to which central examinations or other forms of examinations in EU countries have been digitised. Among other things, we want to record:

- whether, and to what extent, digital test administrations are used
- whether other processes within the examination chain have been digitised
- whether and how AI is used within the examination chain
- what the experiences are with implementing these digitisation steps
- or what motivations exist for not digitising

This research is being carried out by the CitoLab department of Stichting Cito Instituut voor Toetsontwikkeling. The results will be used to inform choices for projects within the Dutch Digitisation Programme for Central Examinations.

What does participation in the study involve?
You give permission to take part in an interview with researchers from CitoLab, in which you will answer questions about the state of your country's digitisation programme. The interview lasts approximately 90 minutes. No later than two weeks before the interview, a questionnaire will be sent by email in which topics for the interview, as well as an additional set of questions, are posed. You may prepare the interview questions, and we would appreciate it if you answered the additional questions in writing. If you yourself are not able to answer these questions substantively, a referral to a (direct) colleague or relevant organisation would be helpful.

What data will be collected from you?
A video and audio recording of the interview will be made using the AI function within Microsoft Teams, also capturing your name, position, and organisation, so that the interview can be accurately processed and analysed. The cleaned transcript will then be shared with you for approval. The collected data will be used solely for research purposes and will be processed in a manner that safeguards your privacy. In internal communications we reserve the right to use your name and organisation, but for all external communication outside the Dutch Digitisation Programme for Central Examinations we present only a general picture of the organisations from which participants originate. In doing so we ensure that no participant can be individually identified. Your name or personal data will not be included in external reports.

How are the data stored and processed?

- The video and audio recording is stored in a secure environment. What is said in the recordings is converted into text, in which no personal data traceable to you appear. The recordings will be deleted immediately after conversion to text.
- The anonymised text files will be retained for 10 years.
- Only the involved researchers of CitoLab have access to the recording.
- The anonymous data is accessible to all staff of CitoLab.
- We use anonymised data to communicate about the study in reports, (scientific) articles and other knowledge outputs.
- Your privacy remains protected. Confidential information or personal data will not be disclosed in any way, unless you indicate that your name may be mentioned as a participant in the study.

Voluntariness

Participation in this study is entirely voluntary. You may stop your participation at any time during the interview without providing a reason. In addition, you may withdraw your consent for the use of your data within two weeks after the interview. Withdrawing consent will have no consequences for you.

By taking part in the interview, we assume that you agree to the use of the information shared for research purposes as described above. If you object to (certain aspects of) the use of your data, you can indicate this before the start of or during the interview.

If you decide to stop participating in the study, or if you have questions or complaints, please contact the project leader of the study, Teun Aalbers, teun.aalbers@cito.nl.

Consent Form

☐ **Participant Statement**
I have read and understood the information about the study. I understand that I may stop my participation at any time during the study. I understand how the study data will be stored and for what purposes they will be used. I consent to participate in the study as described in the information above.

☐ **Optional** I give explicit permission for my name to be mentioned in the external communication.

Name interviewee	Name interviewer
Place	Place
Date	Date
Signature	Signature

Appendix 3

Background information available on request from the authors

If, after reading this report, you require further information, for example regarding the responses to sub-questions 1 to 3 or the other findings under sub-questions 4 and 5, please contact project leader Teun Aalbers at teun.aalbers@cito.nl.

Imprint

This publication of the programme for the digitalisation of centralised examinations was produced by programme partner Stichting Cito.

For more information about the programme, please visit https://www.cvte.nl/onze-themas/digitalisering (in Dutch).

Utrecht, November 2026

© 2026 CvTE, Utrecht. All rights reserved.