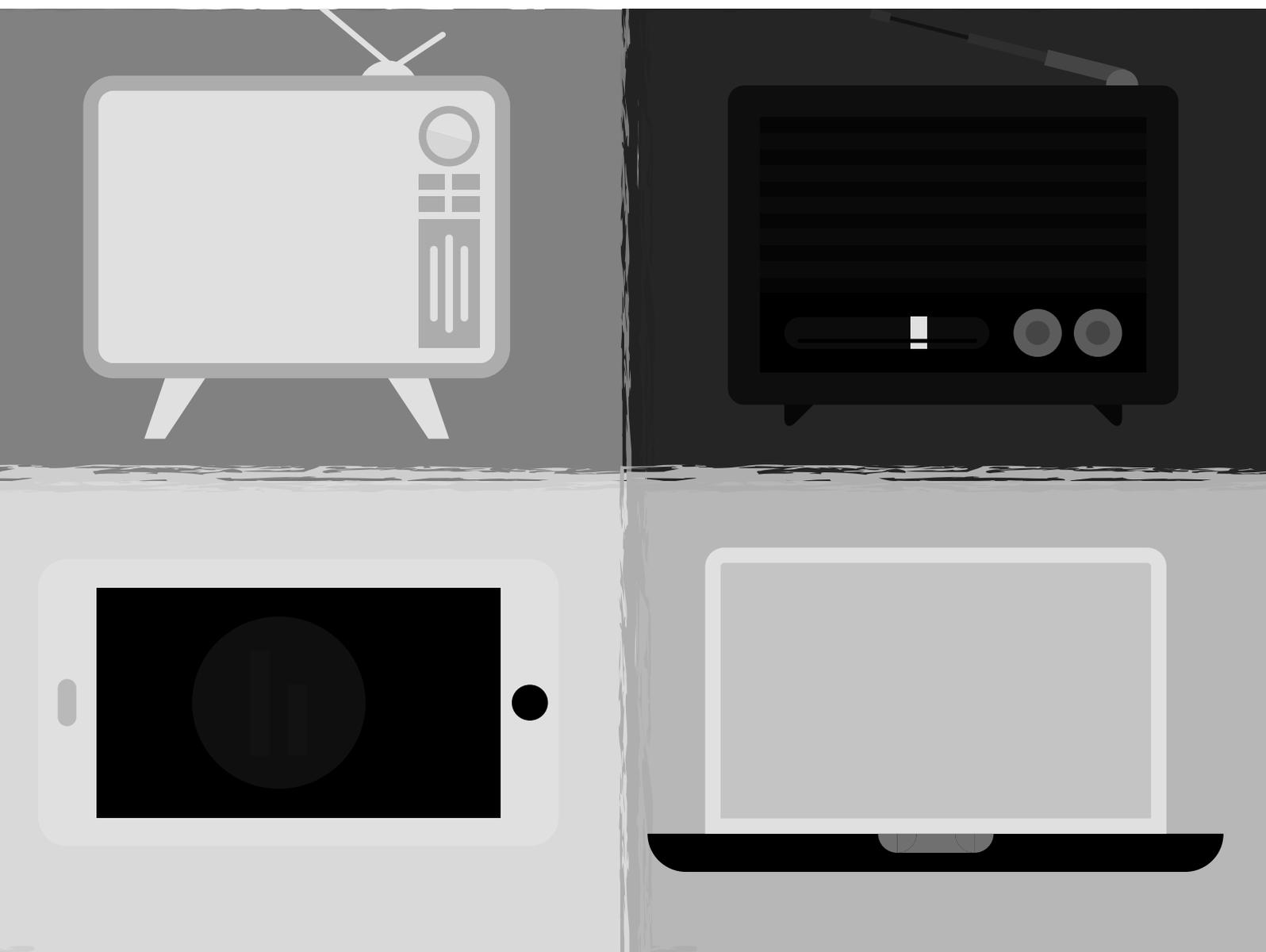


Saudi Arabia: Moving high-stakes examinations online with AI technology

Abdullah Al-Qataee, Nayyaf Aljabri, Laura Gregory,
Amira Kazem



Saudi Arabia: Moving high-stakes examinations online with AI technology

Authors: Abdullah Al-Qataee, CEO, National Center for Assessment, Education and Training Evaluation Commission (ETEC); Nayyaf Aljabri, Director General, Research and Innovation, ETEC; Laura Gregory, Senior Education Specialist, World Bank; and Amira Kazem, Senior Operations Officer, World Bank.

Type of intervention: governmental

Website: www.etec.gov.sa

General description

When Saudi Arabia's schools closed on March 9, 2020 due to the COVID-19 pandemic, the Education and Training Evaluation Commission (ETEC) rapidly brought forward plans to convert its high-stakes Standardized Achievement Admission Test (SAAT) – the Attehseeli – from a paper-and-pencil format to an online format. This move was possible due to investments made over previous decades in infrastructure and expertise for assessments, plus careful planning and communication for the new system and its roll-out. In addition, the switch to online examinations was enabled by new investments in some of the latest artificial intelligence (AI) technology.

ETEC is an autonomous governmental entity responsible for evaluating, assessing, and accrediting education and training in Saudi Arabia, reporting directly to the Prime Minister. Within ETEC, the National Center for Assessment (QIYAS), which has operated for almost two decades, is responsible for developing and implementing over 90 standardised and professional tests for the public and private sectors, having over 1 500 test models and an item bank of over 230 000 questions.

The coronavirus crisis led schools and universities to rapidly transition to a distance-learning mode, via the Internet, television or radio. This series documents some country initiatives that ensured education continuity for all using technology and provided support to teachers, students, and their families.



hundrED

The announcement to move the SAAT online was received with apprehension by students and universities, mostly because they were concerned about potential unfairness, cheating, and technical problems. ETEC put in place measures to (1) prepare students and universities, including a mock SAAT one week before the final SAAT; (2) provide options for every student to sit the SAAT including online or in QIYAS computerised testing centers; and (3) guard against abuses of the system with automated proctoring through AI technology. ETEC's communication of these measures was wide-reaching and pre-emptive, leading students and universities to feel more assured of the new processes, and enabling successful administration of the SAAT on June 8-9, 2020, just 4 weeks behind the usual schedule. At that time, around 218 000 students downloaded the SAAT at home on their computers, completed the examination offline with camera monitoring, and uploaded their responses within the required time, with AI technology flagging potential violations for human review. A further 117 000 students went to an ETEC computerised testing center or mobile testing unit to sit the online version of the SAAT.

This experience has encouraged ETEC to accelerate its previous timelines for moving other tests and assessments online, making better use of latest technologies, and thereby bringing efficiency savings in the long term. In the short turnaround time brought about by the COVID-19 pandemic, much has been learned by ETEC and by the students and universities in Saudi Arabia affected by this switch to online assessments, which will help to further enhance the online examination system in the future.

Main problems addressed

The SAAT is one of three requirements for university entrance in Saudi Arabia, along with the secondary school grade point average (GPA), which is calculated by schools across the three grades of secondary school, and the General Aptitude Test (GAT) known as "Qudorat." Each university decides its own way of weighing these three scores, but universities tend to give SAAT greater weight for several reasons. Firstly, the SAAT correlates with university grade point average more than the General Aptitude Test and the secondary school grade point average. Secondly, compared to the secondary school grade point average, the SAAT has the advantage of being standardised, overcoming the possible problem of grade inflation and harmonisation for grades awarded by class teachers across different schools in secondary education. Finally, the SAAT is curriculum-based, unlike the General Aptitude Test. The science-track SAAT tests knowledge of mathematics, chemistry, physics, and biology based on curriculum content in the three grades of secondary school. The arts-track SAAT tests knowledge of religious education, Arabic language, and social studies over the three grades of secondary school.

The main problem addressed by this solution was to *allow students to graduate on time and generate fair and comparable information about students' academic achievement for the university entrance and selection process*. With the closures brought about by the COVID-19 pandemic, few options were available to allow final grade secondary school students to complete their studies and apply for university and college entrance. Relying on the teacher-awarded secondary school GPA to measure achievement was unsatisfactory given that it is not subject to a moderation process for standardising grades across schools, and because grade inflation may not allow to discriminate between students. Delaying to the next school year would have caused significant disruptions for students, schools, and universities. Given that ETEC had already begun investigating a move from paper-and-pencil to online testing, a decision was made to seize this opportunity to accelerate the move online for the SAAT. As a high-stakes and wide-ranging examination, this required careful planning for well-managed change, affecting large number of students.

Several other related problems needed to be addressed to make the move to online SAAT possible. One key challenge was to ensure that all students had access to the SAAT, regardless of their location, their digital skills, or the types of devices, software, or connectivity they had access to (if any). Another challenge was to ensure the safety of students and their data, while also embedding measures to verify the identity of the students and to identify any violations of expected behaviors during the SAAT. Advanced technologies, including AI, are now available to tackle these issues. Procuring these technologies with

efficiency, flexibility, relevance, and cultural sensitivities was a key consideration for ETEC as it moved quickly to prepare for the new mode of SAAT.

Mobilising and developing resources

As one of the six biggest assessment centers in the world, QIYAS is well established, with testing centers across the country, mobile testing units, advanced security systems, and a vast array of specialist staff and training facilities. These investments made over the last two decades enabled ETEC to respond rapidly to the challenges brought about by the COVID-19 pandemic.

As another important condition that made the examination possible, the Ministry of Education provided remote learning opportunities on multiple platforms so students could continue to study for their examinations. These included an online portal, YouTube channel, and satellite TV channels.

A key decision to make was whether to use live proctoring or automatic proctoring of the online SAAT to guard against student cheating. Live proctoring involves trained proctors monitoring examinees' behaviors through live video and audio, such as their eye movements or interruptions. This is the more expensive option and requires continuous Internet connection throughout the examination period. Automated proctoring involves recording the examinee through video and audio and using AI technology to flag suspicious activity. The AI proctoring can take place live or be activated once a recording is uploaded, thereby minimizing connectivity requirements during the examination. The AI technology identifies patterns associated with cheating such as sounds and voices, facial recognition to ensure the authenticity of the examinee or identify other faces in the room, identification of objects not permitted, patterns of eye movement, and so on. Any violations flagged by the AI proctoring can be reviewed by humans for decision and action.

A new examination platform needed to be procured by ETEC for the SAAT to go online. The following selection criteria were developed by ETEC prior to reviewing the solutions available on the market.

1. *Encryption.* The platform should follow the international encryption standards AES256 to exchange data with examinees; use encrypted email protocols for sending passwords; and encrypt content after uploading of answers and expiration of the examination.
2. *Compatibility.* The platform should support Microsoft Windows and Apple Mac operating systems.
3. *Device activity.* The platform should disable access to all files, programmes, and networks during the examination, as well as close all video and audio data access ports to, and from, the examinee's device; detect and close any unlawful access to the examinee's device and any transmission points before starting the examination; and use time restricted downloading, starting, monitoring, and uploading of the examination with the lowest possible Internet data consumption.
4. *Examinee's behaviour.* The platform should prohibit the taking of screenshots and monitor common examination violations, and use AI technology to monitor the examinee's behaviours during the examination and flag exceptions for later investigation.
5. *Cybersecurity.* The platform should ensure privacy of download and upload for each examinee; apply technical standards for cybersecurity in protecting and implementing the examination; and apply quality standards in data governance, storage, and distribution.

Around five of the most common platforms were compared and found to have similar features, though a couple of them did not allow large numbers of examinees to take the examination in each session, or did not allow offline taking of the examination, or did not include AI proctoring and monitoring.

To ensure equitable access for all students, ETEC prioritised the platforms that had minimum technical requirements and for which Internet connection was required only for upload and download. In addition, ETEC made arrangements for a version of the SAAT to be available at computerised examination centres across the country. While students previously sat SAAT using paper-and-pencil in halls with in-person

4 | SAUDI ARABIA: MOVING HIGH-STAKES EXAMINATIONS ONLINE WITH AI TECHNOLOGY

proctors, the computerised centres already existed for other ETEC testing services. At the point of registering for the SAAT, students were asked whether they wanted to take the examination at home or in one of the centers (with safety precautions for COVID-19 in place). Around two out of three students elected to take the SAAT online at home. The mobile testing units were moved to various locations more frequently than usual to make sure that students in remote areas had access to the computerised SAAT if needed.

Existing ETEC proctors were trained to take on new roles, including investigating violations raised by the AI proctoring. With state-of-the-art training facilities and trainers on hand, ETEC was able to mobilise all training quickly and at little additional cost.

Some changes needed to be made to the design of the SAAT. For example, the length of the SAAT was reduced by 50 percent to minimise time without losing test reliability. An analysis undertaken by ETEC using previous years' data found that the shortened SAAT remained highly reliable. In addition, the test questions were amended so that examinees did not need to use "scratch" paper, since paper for note-taking was not allowed (to minimise opportunities for cheating). Finally, the examinations were adapted to remove content from the sixth term of the final year of school, since that term was affected by COVID-19-related school closures.

In terms of regulations, the only changes that were needed to implement the new online SAAT were those related to student privacy and disclaimers. All other regulations remained as they were.

Fostering effective use and learning

Students needed to be prepared with information about the necessary conditions for completing the examination online at home. Clear communication of the requirements and expectations was crucial for a successful move to online examinations. ETEC prepared and disseminated a brochure telling students the minimum requirements for equipment—including Windows 10, a stable Internet connection, and video and audio capabilities, for example. The brochure also informed students that the examination would be recorded to check for cheating, and described what they should and should not do during the examination. The information disseminated to students made it clear that they needed to take these requirements seriously or else invalidate their results.

Naturally, students were anxious and concerned about the new processes and had questions. ETEC provided several ways to answer their questions, including (1) a telephone line operating from 8am to 8pm, (2) an interactive web chat feature, (3) ETEC's Twitter account, and (4) a WhatsApp number for texting. Brochures, infographics, and videos were prepared to help explain all of the processes and requirements.

To pilot the new system and to give students a chance to practice the new examination conditions, a mandatory "mock" SAAT was held one week in advance of the final SAAT (between May 29 and June 3, 2020). Without a mock examination experience, students may have inadvertently invalidated their final examination result. For example, they may not have taken seriously the need for a quiet place with no interruptions, which would have been flagged by the AI proctoring as a violation. Students were sent a detailed report about their performance in the mock SAAT both in terms of their manoeuvring through the platform and their behaviors during the examination (whether there were any violations). The report provided advice to the students on what they needed to do differently if violations had been noted. Therefore, the mock SAAT was considered an important learning experience for the students in sitting this type of examination. In addition, ETEC provided an online portal with access to previous tests and learning materials for basic subject matter content for students to use in their preparation for the SAAT (<https://elearning.etc.gov.sa>).

The experience of rolling out the mock SAAT also led to the identification of improvements that needed to be made by ETEC before the final version, including to address an initially slow identification verification process and to provide clearer information to students to avoid technical problems. Those students who had technical problems with the mock SAAT (such as not accepting the invitation or not downloading the

test or not uploading their answers) were contacted directly by SMS and email. The mock SAAT period was extended for a day and exemptions were made to allow those students experiencing technical problems to take the SAAT at the designated center.

Universities needed to be kept informed about the changes to the SAAT, so ETEC arranged informational meetings with university management. At these meetings, ETEC explained all aspects of the new online system and presented results of an analysis of the SAAT results compared to previous years, including simulations of various weights for the GAT, secondary school GPA, and SAAT. Sharing this information allowed any concerns to be pre-empted and ensured that the universities felt confident in the reliability of the new online system. ETEC made it clear that universities had the right to decide on the weight to be given to the SAAT. If any university felt uncomfortable with the new online SAAT, it could reduce the weight of the SAAT in its selection criteria. However, ETEC explained to the universities that the SAAT would be implemented because it was the only available standardized curriculum-based admission examination.

Implementation challenges

The situation caused by the COVID-19 pandemic called for urgent and decisive actions on the high-stakes examinations and university entrance criteria for students graduating from school in 2020. Under normal circumstances, some of the choices that were made in rolling out the new online SAAT may have been a bit different; however, most of the implementation challenges would have existed regardless of the timing. Any reform of high-stakes examinations is highly contentious and likely to raise understandable concerns from students, parents, and universities. As a result, the implementation challenges outlined here and their solutions are not restricted to the COVID-19 response alone.

Making online proctoring acceptable. After making the decision to move the SAAT online, the first implementation challenge related to the use online proctoring. With most people being unfamiliar with this type of technology, efforts were needed to alleviate any fears over data privacy, especially with regard to the video-recording of students at their computers for the duration of the examination. QIYAS took two pre-emptive measures to allay these fears: (1) ETEC involved the government agency responsible for regulating data control, and (2) ETEC communicated clearly with students on who would have access to their data, how it would be used, and how it would be stored. Ultimately, significant public concerns about cheating meant that there was general acceptance that the online proctoring offered more advantages than disadvantages. After students experienced the mock SAAT and received reports warning of any violations they had made, it appears that there were far fewer concerns expressed about the online proctoring, possibly because students could see its advantages in terms of fairness and tackling cheating.

Selecting an online platform. The second implementation challenge related to the selection of an online platform for the examination. With more time, ETEC may have been more stringent in terms of the selection or adaptation of existing platforms. However, given the limited time, ETEC prioritised the most important elements, which were fairness, robustness, and security. Some issues arose during the mock examinations that needed to be addressed before the final examinations. These included the handling of the identification verification process during registration – the process was slow and the image recognition sensitivity needed to be increased. In addition, better links between the registration process and the platform were needed. ETEC has become in a better position to make or request such changes, having gained significant knowledge of their needs and the platforms' abilities.

Informing students about new examination rules. A third implementation challenge was ensuring students fully understood the technical and behavioural requirements of the online examination. To overcome this, ETEC put significant efforts into relaying clear messaging through brochures, videos, social media, and “help desks” (telephone, website, and WhatsApp). This approach proved to be generally successful, with most students able to complete the online mock SAAT as intended. However, some students still ran into technical difficulties during the final SAAT and were contacted through SMS and email and rescheduled to take their test at a designated ETEC computerised center.

Lastly, while training of staff to review the reports of the AI proctoring and investigate flagged violations could be an implementation challenge in other settings, ETEC had staff and training facilities ready and only needed the training materials to be developed, so this was not considered an implementation challenge.

Monitoring success

Implementation of the new online SAAT was monitored by ETEC throughout the process – from the mock SAAT to the final SAAT and beyond – in terms of students' adherence to the technical and behavioural requirements, the students' results, and their attitudes toward the new format.

Quality procedures were put in place for monitoring violations of the online SAAT. These processes included (1) an AI proctoring report, (2) human review, (3) a violation audit, (4) statistical assessment reports, and (5) an examination review committee for regulations-based decision-making. During the SAAT, the AI proctoring report found that 88 percent of examination submissions had no flags, while 10 percent had potential violations that needed re-examination, and 2 percent had explicit violations.

All indications suggest that students coped well with the new online SAAT, and that students generally considered the experience fair, and that their level of trust in the system increased. Before the mock SAAT, students had generally wanted the examination to be waived this year. Attitudes toward online testing were measured before and after the mock and final SAAT. Results show that only 28 percent of students felt confident about taking an online examination before the mock SAAT. However, this rose to 31 percent after the mock SAAT and 38 percent after the final SAAT. Likewise, before the mock SAAT, 36 percent of students felt that their home environment was comfortable, calm, and helped them to focus on the examination, but this rose to 43 percent after the mock SAAT, and 48 percent after the final SAAT.

To ascertain the reliability of the new format, ETEC analysed the results of those students who had taken the SAAT last year and this year. The results were generally comparable, indicating that the new format produced results similar to the previous paper-and-pencil format.

Further monitoring and evaluation of the experience of moving high-stakes examinations to an online format with AI technology will take place in Saudi Arabia in the coming months and years. However, the results to date suggest a successful implementation without loss of fairness, reliability, or quality.

Adaptability to new contexts

Saudi Arabia was able to make this rapid change from paper-and-pencil high-stakes examinations for university entrance due, in part, to investments made in previous years. In particular, the infrastructure was available in terms of widespread connectivity and high levels of ownership of digital devices and familiarity with using the Internet among young people. For countries without this type of infrastructure and experience, the move to online examinations would require greater investments and time.

Having computerised testing centers with trained proctors across the country and mobile units for very remote locations for equitable access also enabled Saudi Arabia to have a successful experience with the online SAAT. In other contexts, variations could be made such as using computer facilities in universities and in-person proctoring as per traditional examinations for those who are unable to access the required technology at home.

While the costs were high, the investment is a long-term one – the high initial cost is expected to lead to lower overheads in the longer term. As a result, the online SAAT is now a part of Saudi Arabia's strategic plans for digital transformation, and there are plans to expand remote and computer-based testing for other high-stakes assessments to increase convenience for the examinees, produce faster results, and realise cost effectiveness for the education system.

Key points to keep in mind for a successful adaptation

To move high-stakes examinations online:

1. Decide the most suitable type of online examination, leveraging the existing infrastructure to the extent possible, and bearing in mind the technological requirements (which should be minimised as much as possible) and whether live or automatic proctoring is preferred.
2. Assess the required legislative or regulatory changes, which should be minimised as much as possible.
3. Develop criteria for selecting the online platform and carefully compare solutions available on the market. Be sure to request changes for the local situation as needed.
4. Provide alternative arrangements for students who do not have the required technology or who have special needs or are in remote locations.
5. Adapt the content of the examinations to meet the requirements for online implementation and to reflect the curriculum coverage for the year.
6. Provide help desks in various forms (telephone, online, and text, for example).
7. Publicise the experience as much as possible through effective communication with all those affected by the change – mostly students, proctors, schools, and universities – to keep them well informed with guidance, rationales, evidence, and precise scheduling information.
8. Provide the necessary training for all relevant stakeholders, including proctors and markers.
9. Give students a chance to practise the new format and to test the new system through a mock examination.
10. Expect problems and challenges along the way, and apprehensive students or negative public responses at times, but address them, continue to communicate well, monitor closely, and share information to allay fears and raise trust in the new system.

Acknowledgements

Dr. Husam Zaman, President of ETEC; and ETEC staff involved in the initiative

Please cite this document as: Al-Qataee, A., Aljabri, N., Gregory, L. and Kazem, A. (2020), *Saudi Arabia: Moving high-stakes examinations online with AI technology*, Education continuity stories series, OECD Publishing, Paris.

These papers should not be reported as representing the official views of the convening organisations OECD, the World Bank, or their corresponding member countries, HundrED, and Harvard Graduate School of Education. The opinions expressed and arguments employed are those of the authors.

Comment on the series is welcome, and should be sent to edu.contact@oecd.org.

This document and any map included herein are without prejudice to the status of or sovereignty over any territory, to the delimitation of international frontiers and boundaries and to the name of any territory, city or area.

Except where otherwise noted, content in this work is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 3.0 IGO (CC BYNC-SA 3.0 IGO). For specific information regarding the scope and terms of the licence as well as possible commercial use of this work or the use of PISA data please consult Terms and Conditions on www.oecd.org.