



# Developing a school evaluation framework to drive school improvement

## Introduction

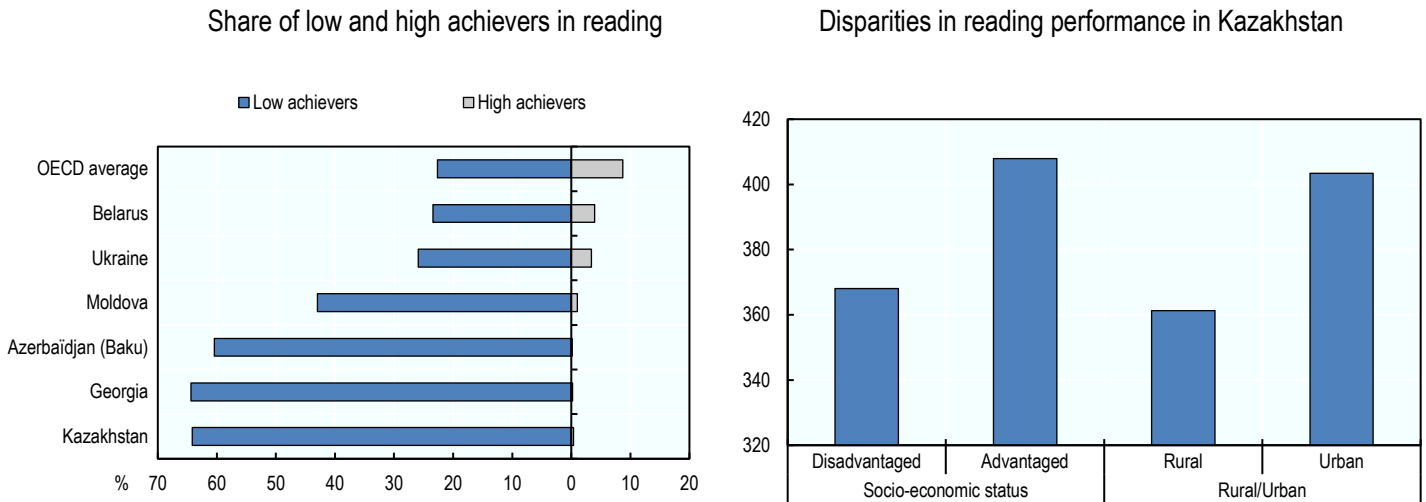
Since achieving independence in 1991, Kazakhstan has undergone rapid development and emerged as a regional economic leader. Recent ambitions, expressed notably in the long-term strategy Kazakhstan 2050, aim to strengthen and diversify the economy in order to position the country as a global leader (Republic of Kazakhstan, 2016<sup>[1]</sup>). A key priority in this strategy is to develop the knowledge and professional skills of Kazakhstan's population, 28% of whom were below the age of 15 in 2018 (compared to an OECD average of 18%) (World Bank, 2018<sup>[2]</sup>). Kazakhstan has already made tremendous progress in providing access to all levels of schooling. Today, enrolment in primary and lower secondary education is nearly universal. Moreover, almost all graduates from lower secondary school continue to either general upper secondary school or vocational studies and roughly half of Kazakhstanis between the ages of 25 and 34 now hold a tertiary degree, which is greater than the OECD average of 41% (see Annex A).

Having achieved high levels of educational access, Kazakhstan is now turning its attention towards improving educational quality. To understand progress in this area, Kazakhstan benchmarks its educational performance against those of leading economies through international surveys, such as the OECD's Programme for International Student Assessment (PISA). Results from PISA 2018 reveal that the average Kazakhstani student scored around 100 points below the OECD average in reading and around 64% of Kazakhstani students were unable to achieve a baseline level of reading proficiency needed to participate fully in society (OECD, 2019<sup>[3]</sup>). This share of low performers is much higher than the OECD average (23%) and one of the highest among PISA participating countries in the OECD Eurasia Competitiveness Programme (Figure 1)

Results from PISA 2018 also show large degrees of inequity in Kazakhstan. Factors such as socio-economic background and, in particular, school location can influence students' performance (Figure 1). Whether the schools of Kazakhstani students are in rural or urban communities explains a greater share of student variance in reading performance (6.7%) than across OECD countries (4.5%). For instance, students in Nur-Sultan city, the national capitol, scored 428 on average, compared to 344 for students from Atyrau, a comparatively more rural region (Figure 2). These findings can be partially explained by a national focus on developing a cadre of very high-achieving students combined with a lack of adequate attention to improving education provision in marginalised areas. In 2008, the government established the Nazarbayev Intellectual Schools (NIS), a network of 20 high-performing schools to which entrance is highly selective and competitive. While students from these schools achieve impressive outcomes, the pedagogical initiatives they have incubated are difficult to scale and not always well adapted to schooling environments in all parts of the country. Meanwhile, students in areas such as Atyrau struggle to achieve

basic minimum standards, influencing their chances of attending tertiary education and finding good employment.

**Figure 1. Reading performance in Kazakhstan in PISA 2018**



Note: The 13 countries included in the OECD Eurasia Competitiveness Programme are Afghanistan; Armenia; Azerbaijan; Belarus; Georgia; Kazakhstan; Kyrgyzstan; Mongolia; Republic of Moldova; Tajikistan; Turkmenistan; Ukraine and Uzbekistan. Only countries with PISA data are included in the figure.

Source: PISA 2018 Database.

**Figure 2. Regional differences in reading performance**



Note: Not depicted are the cities of Nur-Sultan and Almaty, which have special administrative status. Nur-Sultan scored 428, while Almaty scored 424

The challenge of widening inequalities in Kazakhstan is compounded by demographic trends that are straining the system's capacity to provide a quality education for all students. Rapid urbanisation has created overcrowded schools in cities across the country. As of 2018, over 6% of students attended schools that operated in triple shifts (IAC, 2019<sup>[4]</sup>). Meanwhile, achieving universal access to education in a large country with many remote communities has created an extended network of small rural schools that face challenges related to poor infrastructure and staff shortages (IAC, 2019<sup>[4]</sup>; OECD/The World Bank, 2015<sup>[5]</sup>). Particularly representative of these circumstances are “ungraded schools”, which do not have enough students to form full classes of separate grades. As of 2018, around 41% of public schools were ungraded schools, though they only enrolled 6% of the student population (IAC, 2019<sup>[4]</sup>).

To develop the sustainable and knowledge-based economy that Kazakhstan envisions, the government needs to create systems and instruments that help it understand how all students are performing and how they can be supported in their learning. This OECD country review examines four educational policy areas (see Box 1) that Kazakhstan can focus on in order to improve the outcomes of all students.

### **Box 1. The OECD's review of education evaluation and assessment policies in Kazakhstan**

This policy perspective is one in a series of four that draw on an OECD knowledge-base created through reviews of evaluation and assessment policies in over 25 education systems. To complete this review, the Ministry of Education and Science of Kazakhstan (hereafter, the ministry) and the OECD review team chose a specific policy issue within four broad areas of evaluation and assessment (student assessment, teacher appraisal, school evaluation and system evaluation). The selected issues are:

- Strengthening national examinations in Kazakhstan to achieve national goals
- Raising the quality of initial teacher education and support for early career teachers in Kazakhstan
- Developing a school evaluation framework to drive school improvement
- Developing a national assessment that supports Kazakhstan's education goals

The review of these policy issues was based on national information that Kazakhstan provided to the OECD, background research and a visit to different parts of the country in November 2019. During the visit, a team of OECD staff met with key actors across the education system to discuss the policy issues. This evidence formed the basis of the policy perspectives, each of which provides actionable recommendations based on insights from international practices to help Kazakhstan strengthen student learning while making learning outcomes more equitable.

### ***The importance of school evaluation***

School quality is essential to facilitating strong and equitable student learning. Schools that are well run and focused on what matters most can lift the performance of students from diverse backgrounds. On the other hand, schools that do not provide adequate conditions for teaching and learning will struggle to help their students progress and can allow achievement gaps to widen.

In order to ensure quality schooling, countries develop school evaluation systems, which set standards for good schooling and processes for evaluating schools against those standards. These systems help hold schools accountable for their operations and performance and support them to improve (OECD, 2013<sup>[6]</sup>). To function effectively, school evaluation systems require information that is relevant to understanding the quality of teaching and learning in schools. This information needs to be fed back to schools in ways that can help them understand what they can do to improve. The information also needs to be fed up to policy-

makers, so that schools have the support and incentives they need to meet standards and assume progressively more responsibility for their own quality assurance.

School evaluation systems that are able to accomplish these tasks are trusted by schools to help them improve their practices. School evaluation systems that are poorly designed can turn the evaluation process into a punitive, compliance-based exercise that is unrelated to schools' core activities. In these cases, schools are less likely to engage meaningfully with the system and will not receive the type of formative feedback they need to improve their practices.

## Schools in Kazakhstan

In Kazakhstan, schooling is compulsory until the end of lower secondary education, which, given the size of the country, results in an extensive school network. In 2018, over 7 000 schools were in operation, but students and schools are not distributed equally around the country. Schools in urban areas represent roughly 25% of all schools but enrol 54% of all students (IAC, 2019<sup>[41]</sup>). This unbalanced enrolment creates different challenges related to schooling. In urban areas, many schools are overcrowded, with 4 826 schools operating in double shifts and 128 operating in triple shifts as of 2018. According to PISA 2018 data, over 97% of schools from large cities (a community over 1 000 000 people) have more than 15 students in their classes.

In rural areas, schools are smaller and struggle with adequate resourcing. Over 30% of schools from villages in Kazakhstan (a community with fewer than 3 000 people) have fewer than 15 students in their classes. Roughly twice as many principals from villages in Kazakhstan reported that instruction is hindered a lot by inadequate infrastructure than principals from large cities. Similarly, almost three times as many principals from villages reported that instruction is hindered a lot by inadequate materials than principals in large cities (OECD, 2019<sup>[77]</sup>).

Given the size of the Kazakhstani school network, many schools in rural areas do not have enough students to form full classes with separate grades. These schools are referred to as “ungraded schools” and account for over 41% of all public schools in Kazakhstan, though only 6% of students (IAC, 2019<sup>[41]</sup>). Ungraded schools face a particular set of challenges related to their size and isolation—60% do not have classrooms with suitable chemistry and biology equipment, 27% do not have gymnasiums, 26% do not have cafeterias and 13% do not have high-speed internet connectivity. Part of the reason these challenges arise is that 35% of ungraded schools are in “adapted buildings”, meaning their facilities were never intended to act as schools (MoES, 2019<sup>[8]</sup>). Ensuring school quality in Kazakhstan means that these very different environments, and how they affect student outcomes, need to be taken into consideration when evaluating schools and offering suggestions for improvement.

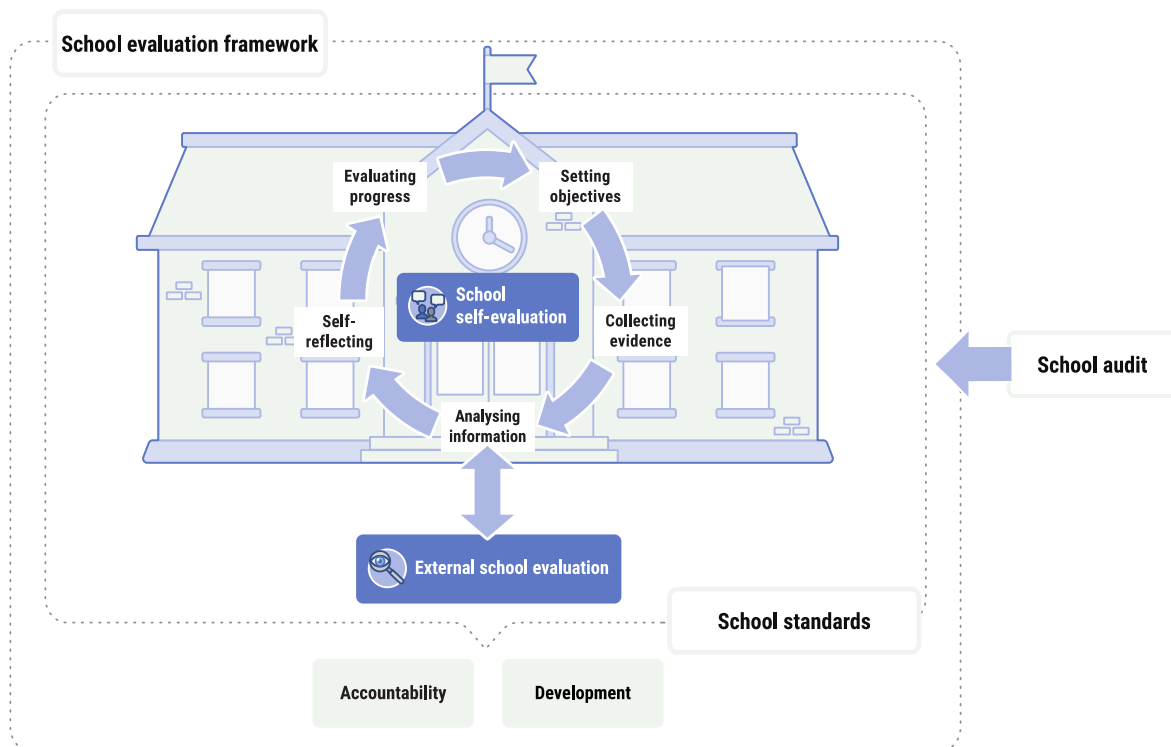
## School evaluation in Kazakhstan

### *Key features of school evaluation*

A country's school evaluation system is set out in a school evaluation framework, which is a set of explicitly related policies that cover how school quality will be determined and evaluated. A key component of a school evaluation framework are school standards, which define the vision and dimensions of quality schooling in a country. Typically, schools are then evaluated against the standards through two methods, which both use a variety of evidence and information to determine school quality. The first is external evaluation, where individuals from outside the school evaluate the performance of the school. The second is self-evaluation, where staff from the school evaluate the quality of the school's practices (OECD, 2015<sup>[9]</sup>).

Separate from school evaluation is school auditing, which examines whether schools are complying with statutory regulations. While school evaluation and school auditing can examine common elements, and therefore need to be co-ordinated, their fundamental purposes and focuses differ—school evaluation serves a primarily educational function, while school auditing serves a more administrative function. Figure 3 illustrates these different components of school evaluation and the remainder of this section discusses the components in greater detail.

**Figure 3. School evaluation**



### *School evaluation framework*

Most OECD countries have developed comprehensive school evaluation frameworks that outline the key components of their school evaluation systems (OECD, 2013<sup>[6]</sup>). School evaluation frameworks typically define and describe:

- Purposes of school evaluation
- School standards that explain what is a quality school
- School evaluation processes and how those processes relate to each other
- Responsibilities for conducting external and self- evaluation
- Evidence used to evaluate schools
- Use of school evaluation results
- Capacities needed to conduct school evaluation

Kazakhstan does not currently have what is understood to be a school evaluation framework in most OECD countries. Instead, determinations of school quality have historically relied on school audit processes (these are described later) that focus on compliance with administrative regulations instead of the quality of educational practices and outcomes.

However, over the past five years Kazakhstan has been developing a potential school evaluation framework, the “School Review”, which focuses on assessing teaching and learning in schools and supporting schools to improve. Development started in 2015 when, as part of a World Bank supported school improvement project, Kazakhstan developed and piloted new School Review standards that focused on teaching and learning instead of compliance with regulations. These standards were completed in 2016. However, in October 2016, the World Bank school improvement project was suspended and School Review was never completed. Although School Review continues to be discussed as a potential school evaluation framework, it has not yet been approved by the Parliament and at the time of completing the policy perspective, its development and implementation have been put on hold. The remainder of this section discusses the different proposed components of School Review.

### School standards

School standards define, in the context of a specific country, what constitutes a quality school and act as the reference points for external and self-evaluation procedures. School standards usually comprise several different areas, or domains, such as school processes, students’ outcomes and school infrastructure and resources (OECD, 2013<sup>[6]</sup>). Domains are then further defined through criteria and descriptors, which indicate in more precise terms what quality looks like in a given area. Some countries also provide illustrations of different degrees of quality, which can help both evaluators and schools visualise the difference between excellent and very good, or satisfactory and weak. Experience in conducting school evaluation has led many countries to reduce the number of school standards in order to focus more on what matters most and avoid overburdening schools and evaluators with too many criteria, which risks turning evaluation a check-box exercise.

Kazakhstan’s School Review standards are categorised into six domains—leadership, teaching, learning outcomes, resources, care and support and community relations. The domains are associated with 41 total criteria that each have a weighted coefficient. Criteria are further disaggregated into 164 descriptors that describe the practices associated with four rating levels: excellent, good, satisfactory or unsatisfactory. The domains and criteria of the standards are shown in Table 1.

**Table 1. School review standards for primary and secondary education**

Domain	Criteria
Quality leadership and management	Analysis and self-assessment of activities
	Understanding the mission and strategic development plan
	Distribution of responsibilities and teamwork
	Professional development and motivation of teachers
	Monitoring and evaluation of the quality of teaching and education activities
	The moral and psychological climate in the organisation of education
	Compliance with laws and regulations
	Planning and organising the current activities of the school
	Resource usage
Teaching quality	Accurate and in-depth knowledge of the subject
	The quality of lesson planning and educational process management
	Organisation of the learning and cognitive activity of students
	The ability to engage and motivate students
	The ability to reliably and objectively evaluate

Domain	Criteria
	Ability to organise feedback
	Use of educational resources
	Methodology for organising homework
	Favourable psychological climate in lessons
	Communicative competences
Learning outcomes	Student growth according to the results of internal assessments over a period of time
	External assessment results (UNT, EAAA, TIMSS, PISA, etc.)
	Results of testing (cross-sectional)
	Results in intellectual contests / competitions
	The results of additional educational services (sports, cultural, creative)
	Employment of graduates
	Personal development
Resources	The condition of the building and the availability of the necessary premises
	Logistics
	Library resources
	Adequacy of teaching and support staff
Care and support	Ensuring the safety of students
	Health care
	Creating conditions conducive to the individual development of students
	School self-government development
	Psychological and social support for students
	The development of the personality of students through extracurricular activities
The quality of interaction with parents and the local community	Career guidance
	The interaction of parents and the local community with the organisation of education to improve the quality of the educational process
	The interaction of parents and the local community with the organisation of education to increase the motivation of students
	Informing parents about the educational achievements of children
	Organisation of the activities of the parents' committee

Source: Minister of Education and Science of the Republic of Kazakhstan (2016<sup>[10]</sup>), On the approval of evaluation criteria for educational institutions. Order of the Minister of Education and Science of the Republic of Kazakhstan dated February 2, 2016 No. 124.

Notes: Unified National Test (UNT), External Assessment of Academic Achievement (EAAA), Trends in International Mathematics and Science Study (TIMSS), Programme for International Student Assessment (PISA)

### External evaluation

External evaluation refers to the evaluation of a school's quality, according to school standards, by an external body (OECD, 2013<sup>[6]</sup>). Effective external evaluation is led by an independent agency and carried out by teams of qualified evaluators. Typically, when external evaluations occur, evaluators first perform desk research on the school using information ranging from central data, audit reports and the school's self-evaluation reports. Evaluators then perform a school visit, during which they interview school staff and observe classrooms in order to collect more information about the school. Finally, the external evaluation team will express their views about the school's quality and their recommendations for improvement. Usually, results are communicated through a formal, written report, all or most of which is made public.

An important consideration of external evaluation is whether all schools should be subject to it, on the same basis, within a given time period. Such universal coverage helps ensure school quality around the country. However, sending evaluators to all schools is resource intensive. To address these trade-offs, many OECD countries have established differentiated external evaluation procedures, whereby schools are evaluated according to different schedules and, depending upon the results of their evaluations, receive different types of support (OECD, 2013<sup>[6]</sup>). A key component of differentiated external evaluation is determining when and how frequently schools are evaluated. Many countries employ a data-based risk



assessment, in which central data are reviewed to detect schools that are most in need of support. Based on regular risk assessments, some schools might be externally evaluated quite frequently, while others not for several years.

In Kazakhstan, School Review outlines differentiated external evaluation procedures. How schools were rated during their previous external evaluation would determine when their next external evaluation occurs, though how schools would be prioritised for their first external evaluation has not been determined. The aim of School Review is that all schools receive at least one external evaluation every five years. Three types of trained and certified evaluators (experts) from around the country (school leaders, teachers and psychologists) would be selected to join evaluation teams. The size of evaluation teams would be based on school size (roughly one expert per 300 students, with a minimum of three) and school location (evaluators would be chosen from outside the school's region). The exact duration of an external evaluation visit has not been determined.

### **Self-evaluation**

School self-evaluation refers to the process by which schools review their own policies and practices and monitor their performance to encourage reflection, goal setting and inform school development plans (OECD, 2013<sup>[6]</sup>). Many OECD countries, such as Austria, the Czech Republic and France, require schools to undertake self-evaluations at least once every two years (OECD, 2015<sup>[9]</sup>). Furthermore, many countries now use the results from self-evaluations to feed external evaluations, with evaluators reviewing self-evaluation results as part of external evaluations.

In Kazakhstan, schools were required to conduct self-evaluation annually and to produce a self-evaluation report as part of the previous school attestation process (OECD/The World Bank, 2015<sup>[5]</sup>). Although self-evaluation standards, including indicators and descriptors and a self-evaluation template, were available, schools were provided with little guidance on how to carry out the process. Since the audit process recently shifted away from school attestation and towards a preventative control and unplanned inspection model (see section on Audit, below), self-evaluation is no longer purposeful and its importance is no longer signalled in policy.

### **Use of evaluation results**

How school evaluation results are used is central to the overall effectiveness of a school evaluation system. Following external evaluations, schools are typically given a report that summarises the findings of the evaluation team and rates the school according to the school standards. Countries vary in terms of the consequences associated with different ratings. In case of exceptional performance, a school might receive, beyond the public recognition that this implies, less frequent evaluations or more flexibility in how they organise their curriculum or staff time. In some countries, such as Serbia, schools rated excellent are given a system leadership role in the form of mentoring poor performing schools (Maghnouj et al., 2019<sup>[11]</sup>). For schools given low ratings, the first response is usually enhanced oversight and support, such as more frequent external evaluations and additional guidance from identified support bodies. More punitive consequences, such as sanctions or school closure, are usually avoided, or seen as the very last response to persistent underperformance (OECD, 2013<sup>[6]</sup>; OECD, 2015<sup>[9]</sup>). This type of approach reduces the negative pressures around school evaluation and can create more openness from schools to engage with school evaluation as an exercise in self-improvement.

According to School Review, schools would be assigned one of four ratings following their external evaluation, excellent, good, satisfactory or unsatisfactory, which correspond to a school's weighted overall score vis-à-vis the School Review criteria. Higher rated schools would receive less frequent reviews while lower rated schools would receive more frequent reviews (potentially every year), support from the Committee and funding to help themselves improve.



This proposed system for school support, whereby a single organisation supports schools based on the results of external evaluation, would represent a significant change in Kazakhstan. At the moment, several agencies are responsible for providing support to schools. They include the methodological cabinets (see further description below), Orleu (a teacher training centre), the NIS Centre of Excellence and the National Academy of Education. Moreover, in the absence of a school evaluation framework to focus and co-ordinate efforts, each support agency has had its own support agenda, with its own methods of identifying schools and its own ways of providing the support.

### *Audit*

A school audit system is different in purpose from school evaluation and focuses primarily on a school's compliance with regulations, rather than the quality of teaching and learning at the school. Because auditing requires different expertise than evaluating schools, audits tend to be conducted by different agencies than those that oversee school evaluation. This separation can also help the school evaluation body engage schools in a more open conversation on educational quality and be disassociated from the potential punitive consequences and pressures of an audit. Nevertheless, despite being separate, audits are usually co-ordinated with school evaluation. For example, audits and school evaluation visits should not occur close together to avoid overburdening the school. School evaluation can also use information produced by an audit to inform its assessments of school quality, particularly with respect to standards that focus on school infrastructure, which avoids schools having to provide the same data repeatedly.

Until 2016, Kazakhstan had a process called school attestation that attempted to measure school quality, but it remained mainly focused on schools' compliance with regulations and education standards, making it akin to a school audit. After a change in the national Entrepreneurial Code (see Key agencies of school evaluation, below), and partially motivated by resource limitations, school attestation was recently discontinued and changed to a two-pronged audit approach, preventative control and unplanned inspection. Both are carried out by the Committee and are described below.

### **Preventative control**

Preventative control is conducted via two-steps. First, an annual risk assessment is performed based upon data from the National Education Database (NED), documents provided by schools and, if applicable, results from a school's previous preventative control visits. Risk assessment criteria, among other areas, look at teachers' compliance with certification requirements, compliance with the national curriculum, quality of infrastructure and the amount of resources provided. Based upon the evidence collected, schools are assigned a score from 0 to 100, with higher scores indicating a greater degree of risk. Schools that receive 61 points or higher proceed to the second step of preventative control, which is a school visit. After the visit, the Committee provides the school with a report that includes recommendations that the school should adopt within a fixed period of time. If the school does not remedy its issues within this timeframe, the school may be subject to sanctions, potentially leading to school closure.

### **Unplanned inspection**

Unplanned inspection is triggered by a school emergency. When an incident is reported, the regional Committee office sends a representative to visit the school in order to verify the violation. If confirmed, the school must address the violation or be subject to sanctions.

The school-level data generated by preventative control and unplanned inspection are stored in the Committee's regional offices. Twice a year, the Committee's regional offices share the information with the central Committee in the form of aggregate reports of their regions.

## **Key agencies of school evaluation**

### *The Quality Control Committee in Education and Science*

Kazakhstan has an intricate system of “control” that is governed by a national Entrepreneurial Code. In education, the Quality Control Committee in Education and Science (formerly the Committee for Control) fulfils this role and is responsible for several activities. The Committee currently oversees preventative control and unplanned inspection, and had previously been responsible for school attestation. It has also been leading the development of the potential School Review system, according to which it would be charged with providing support to schools and teachers in response to School Review results. Finally, the Committee oversees the administration of Kazakhstan’s national assessments and examinations.

In addition to its central office in Nur-Sultan, the Committee operates out of its 17 regional offices. There are roughly 200 staff working for the Committee, of which around 150 work in the regional offices. Individuals at regional offices are responsible for conducting preventative control and unplanned inspection. Evaluators at regional offices are required to have a degree in education but there is no requirement to have teaching experience.

### *Methodological cabinets*

Rayon and oblast akimats (local and regional governing bodies, respectively) contain Departments of Education, within which are methodological cabinets that are staffed by education specialists and are responsible for providing support to schools. Each methodological cabinet develops annual support plans and monthly calendars that list the activities available. For example, methodological cabinets organise trainings, conferences and workshops to support schools around using the new curriculum. Each methodological cabinet has its own criteria for determining which schools need support and what type of support to provide.

## **School leaders**

School leaders have an essential role in carrying out effective school evaluation and, more generally, in leading schools toward improvement. Fulfilling these responsibilities requires school leaders to have the capacity not just to administrate their schools, but to be effective instructional leaders who can direct improvements in teaching and learning.

Pre-service preparation is key to building the capacity of school leaders to perform their roles. Nearly 58% of principals on average in Teaching and Learning International Survey (TALIS) participating countries reported that they received training in instructional leadership before taking up their positions (OECD, 2019<sub>[12]</sub>). Principals’ initial training must also be complemented by opportunities for continued professional development once they are in their posts. These opportunities include formal training programmes offered by the government, or informal ones where principals work together to examine practices and acquire new knowledge (DuFour, 2004<sub>[13]</sub>).

In Kazakhstan, school leaders have diverse administrative and educational responsibilities, including selecting teachers and support staff, approving the management structure, deciding on the dismissal of teachers, fostering pedagogical improvement and providing school-based professional development to teachers (OECD/The World Bank, 2015<sub>[5]</sub>). However, how they are selected is mainly reflective of their administrative function, with their educational qualifications and their knowledge of national laws being important selection criteria (OECD/The World Bank, 2015<sub>[5]</sub>).

The vast majority of lower secondary principals are experienced teachers, which makes them experienced at building relationships with school faculty. Nevertheless, principals are generally not prepared to be the strong instructional leaders that are needed to catalyse school improvement. In TALIS 2018, roughly 75%

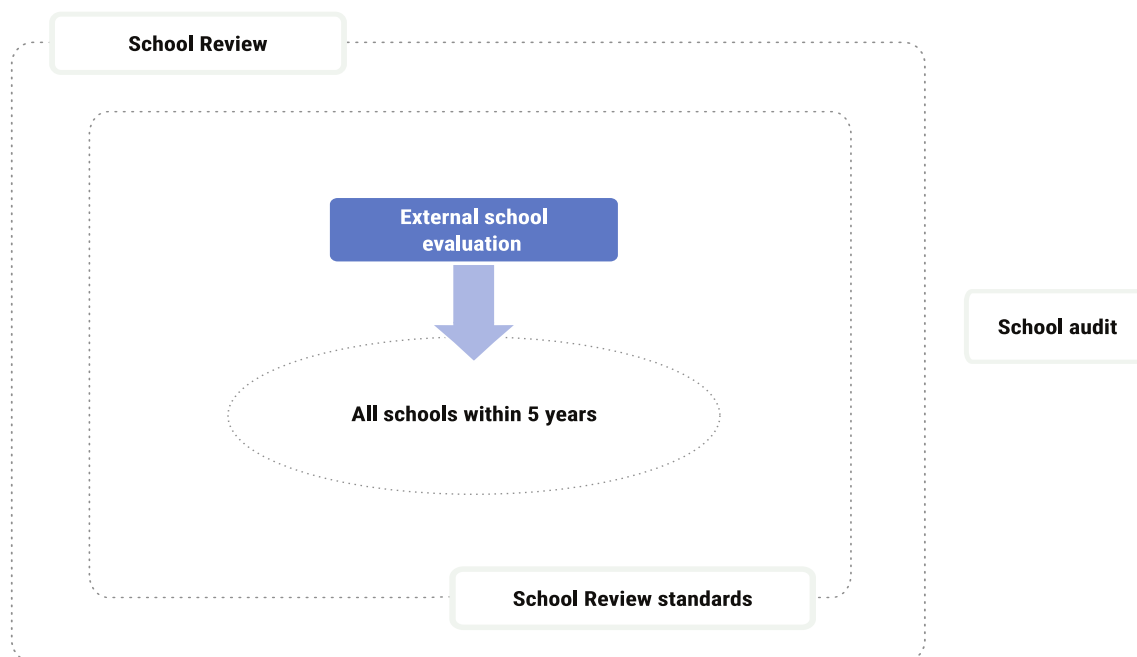
of principals reported that they did not receive formal training in instructional leadership before becoming principals (OECD, 2019<sup>[12]</sup>). In fact, there is only one master's level programme in educational leadership in the country (at Nazarbayev University). Once in their posts, there are continuous professional development opportunities for principals, such as short-term training provided by Orleu and a 9-months programme proposed by the NIS Centre of Excellence (Tashibaeva, n.d.<sup>[14]</sup>). Overall, however, principals' opportunities of professional development remain limited, especially in rural areas.

## Review of the context

The two primary purposes of school evaluation are to hold schools accountable for the quality of schooling that they provide and to help them improve their teaching and learning practices. To achieve these purposes, a school evaluation system should strive to meet certain criteria. Its school standards should focus on what matters most for establishing a positive educational environment and producing good student outcomes. Its procedures around external and self-evaluation need to be efficient for central authorities (i.e., not overburden evaluators with visits) and schools (i.e., not to overburden staff with paperwork). Finally, and importantly, the entire school evaluation system must be meaningful to schools and not just a compliance exercise. Therefore, the actors involved must have the capacity to fulfil their roles and schools must trust and want to engage with evaluators and providers of support. The extent to which school evaluation in Kazakhstan fulfils these purposes and meets these criteria forms the basis of the analysis and recommendations found in this policy perspective.

Kazakhstan has never had a school evaluation system that focused centrally on improving teaching and learning. Its audit systems, while comprehensive, have always focused on compliance and have not provided schools with sufficient support to improve their educational practices. Recognising this need, the Committee has been developing School Review as a potential school evaluation model, which would be separate from its own audits and intended to help schools improve the quality of education that they provide (Figure 4). However, further development of School Review is currently on hold and its implementation in the future remains uncertain. The OECD strongly supports continued development of School Review, and situates the analyses in this policy perspective in accordance with what has already been created for School Review. Nevertheless, in consideration of the Kazakhstani context, the OECD has also identified some potential risks regarding the potential School Review.

Figure 4. School Review



One key issue regarding the potential School Review is that the Committee would be responsible for overseeing it. The Committee has always been an agency that checks schools against compliance and delivers punitive sanctions. Schools regard the Committee with anxiety and do not like interacting with it for fear of reprisal. Therefore, the Committee would not be well-positioned to help schools focus on the quality of their teaching and learning practices and lead a school evaluation system about support and improvement.

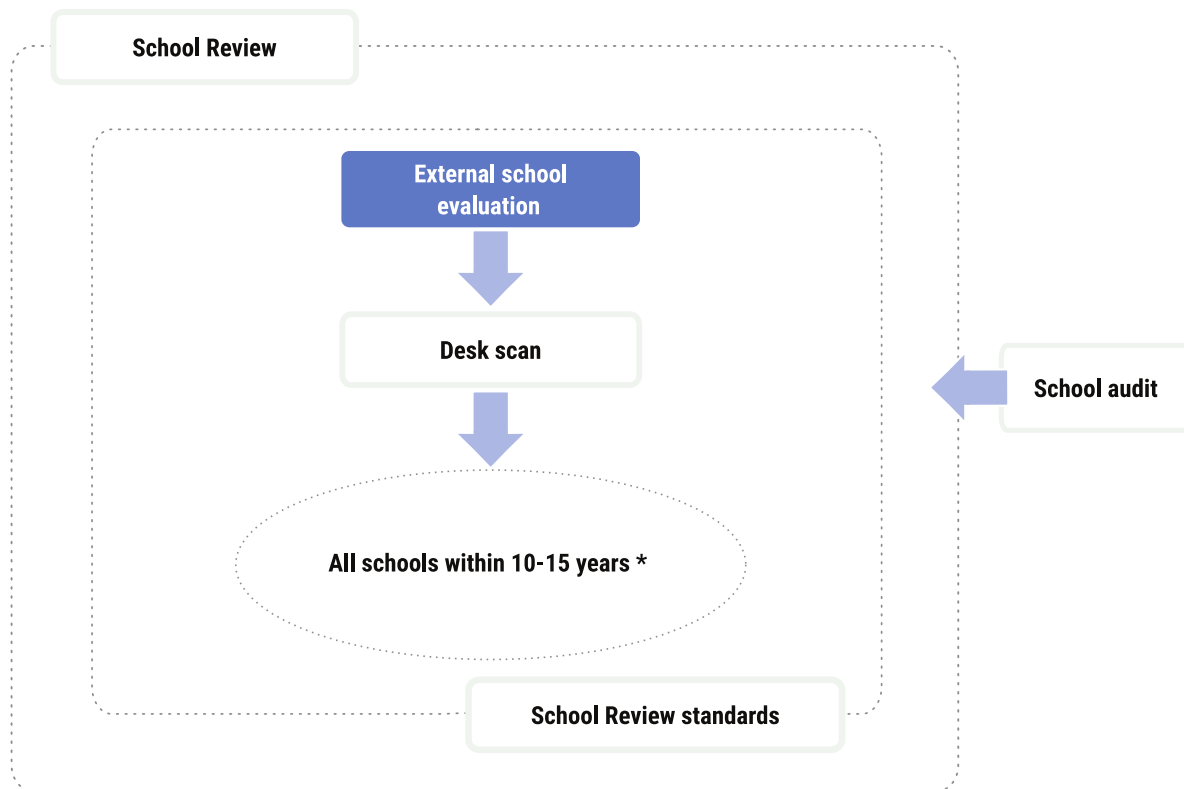
Another issue concerns the school standards that would be part of the potential School Review. While they are rigorous and set high expectations, they are not focused enough on the main quality concerns in Kazakhstani schools. Many schools in Kazakhstan struggle to provide the fundamental conditions needed for students to learn and cannot meaningfully engage with standards that do not consider their contexts. Evaluating such schools against many of the proposed School Review standards would not be a constructive exercise for them and would not produce recommendations that would help these schools improve.

The approach to implementing the potential School Review also raises concerns. While the ambition to visit every school is positive, this goal might not be feasible or efficient for the Kazakhstani school network, which extends to far, remote areas. Officials had some difficulty visiting all schools under school attestation because of resource constraints. Re-introducing this expectation through School Review also would re-introduce the same challenges. Additionally, given the acknowledged achievement and resource gaps between Kazakhstani schools, comprehensively evaluating all schools would risk that schools that need the most support would only receive limited assistance, which would exacerbate inequalities.

After external evaluation is complete, schools would be identified for support based upon their evaluation results, a strategy that would represent a significant improvement compared to the current environment in which support is offered in an un-coordinated manner that is not always based on clear evidence of school needs. However, School Review identifies the Committee as the agency that would provide the support. As an auditing agency, the Committee is not staffed by individuals with an educational support background. They would have great difficulty providing the educational support that schools in Kazakhstan need.

Finally, the potential School Review is centred around external evaluation. There is little mention of how schools would be expected to review and improve themselves in between external evaluation visits, which, given the size of the Kazakhstani school network, could represent a period of several years. Furthermore, any incentive to engage in self-improvement will necessitate that schools have the capacity to do so. Since schools have had little experience engaging in these activities, it is unlikely that they have such capacity in the immediate future. Figure 5 illustrates the OECD’s recommendations regarding Kazakhstan’s School Review model. The specific recommendations represented in the figure will be discussed in the remainder of this policy perspective.

**Figure 5. School Review with OECD recommendations**



Note: \* Every year, a sample of schools identified as the most in need of support are evaluated in addition to a representative sample of schools identified as having different quality levels by the desk scan process.

### **Recommendation 1. Establish a new school inspectorate to manage School Review**

The development of School Review originated from Kazakhstan’s desire to shift the understanding of school quality from complying with legal measures to fostering teaching and learning. It is imperative, therefore, that the public perception of the agency that would be responsible for managing School Review help reinforce this new understanding of school quality. In addition, the agency itself should be well-positioned to execute the necessary responsibilities if Kazakhstan decides to reconsider the development and implementation of School Review in the future. From this perspective, the Quality Control Committee may not be ideally placed to manage School Review. Its perception as an agency that ensures legal

compliance might be a barrier to advancing an overall understanding of school quality that is based on teaching and learning. Furthermore, because of the Committee's historical role of delivering sanctions, schools might be hesitant to engage with it openly. Finally, the Committee has several responsibilities and does not have the capacity to take on extra ones.

### 1.1. Create an independent school inspectorate

#### *Evidence*

#### **Promoting a new national vision of school quality will be difficult so long as School Review is associated with the Committee**

In modern education systems, school evaluation is underpinned by the notion that school quality is about how well a school is able to help students learn and develop. In many countries, however, heavy regulatory frameworks have encouraged schools to focus their attention on compliance with national legislation, with less impetus and space given to how they can develop their teaching and learning practices. In these situations, introducing school evaluation not only means introducing new school standards and procedures that align with a more modern understanding of school quality, but helping to reinforce that understanding itself.

Internationally, a growing number of OECD and other European countries have set up independent school evaluation bodies that are responsible for school evaluation (OECD, 2013<sup>[6]</sup>). These bodies have a stated responsibility for improving the quality of schooling and employ or contract staff with professional expertise in education and, in most cases, hands-on experience (e.g., principals or teachers). The functions of these bodies, and the skills and resources they require, are different from those of school audit agencies. For these reasons, in most countries they are institutionally separate, which helps maintain the difference and integrity of the two functions, and can help build broad understanding of how improving teaching and learning differs from simple administrative compliance.

Kazakhstan has a strong system of school audits, but the country has recognised that this system evaluates schooling through a narrow lens of legal compliance and not teaching and learning. To address this situation, Kazakhstan is developing School Review as a potential model to evaluate schools according to a more modern and educationally relevant understanding of school quality. However, School Review would still be under the auspices of the Committee. Given the mandate of the Committee as an audit agency and its historic role as legal regulators, there is significant risk that School Review would be seen as yet another compliance check and not a new way of understanding and evaluating school quality.

#### **Schools' interactions with the Committee are characterised by anxiety, which might discourage them from engaging with School Review and distort their behaviour**

International evidence shows that, in order for school evaluation to be effective, schools must regard the responsible agency positively and trust its judgments and intentions (Mcnamara and O'hara, 2006<sup>[15]</sup>). Since the legal compliance checks under the school audits are a demanding and stressful process, schools are less inclined to open up to the audit agency. Moreover, they might even adapt their procedures solely to satisfy audit requirements, which can distract them from their core responsibilities of enhancing teaching and learning.

Kazakhstan recognises that the Committee is regarded with anxiety by schools. National data on TALIS 2018 and interviews carried out by the OECD review team revealed that school attestation visits from the Committee were a high source of stress on schools. Interviews with principals and teachers also revealed that they feel a lot of pressure when engaging with the Committee and its audits. To work around this perception for School Review, the Committee is considering identifying, training and certifying a cadre of



external experts to evaluate schools instead of sending Committee staff. Nevertheless, the external school evaluators would still be associated with the Committee and its reputation as a punitive organisation. There is, therefore, risk that schools would regard the evaluators and School Review similarly to how they view the Committee's audits, and would therefore be reluctant to engage with School Review for development purposes.

### **The Committee lacks the capacity needed to successfully implement School Review**

The Committee fulfils multiple tasks as the primary agency responsible for quality control in Kazakhstani education. Given the size of the country's school network and the limited number of staff in the Committee, it is unsurprising that capacity at the Committee is a concern. In fact, the difficulty the Committee encountered in auditing every school in the country was one of the reasons that Kazakhstan moved from school attestation to risk-based preventative control. While this change temporarily relieved some of the pressure on Committee staff, the limited available staff coupled with their various responsibilities would make the management of School Review a challenge. Even though the Committee would not use its own staff to conduct external evaluations as part of School Review, it would still be responsible for managing School Review overall, including hiring evaluators, training evaluators, and supporting schools. The already overburdened Committee staff will not be able to assume these new responsibilities effectively.

#### *Recommended actions*

### **Establish an independent body in charge of external school evaluation**

Kazakhstan should create an independent school inspectorate (hereby referred to as “the School Inspectorate”) that is separate from the Committee and the ministry. This Inspectorate would then become the sole body responsible for overseeing the potential School Review. Separating School Review from the Committee in this manner would help reinforce the new understanding of school quality espoused by School Review and distinguish it from previous understandings of school quality that are focused on statutory compliance. Furthermore, it would help develop a distinct mission, instruments and capabilities for the newly created body in charge of external school evaluation.

The independence of the School Inspectorate is critical to its success as it ensures that the School Inspectorate's evaluations of school quality would be impartial. Knowing this, schools would be more trusting of the School Inspectorate's judgments and recommendations and want to engage with it in a way that they would not with the Committee. To maintain its independence, the School Inspectorate should be led by a professional who is appointed for a specified duration following a competitive process on the basis of their competencies.

### **Establish oversight over the School Inspectorate**

Given the powerful influence that an inspectorate has on schools' work, Kazakhstan should create an independent advisory board composed of respected education professionals and governance experts to oversee the work of the Inspectorate and hold it accountable. While such a board would help to maintain the inspectorate's independence and integrity, it could also play a key role in monitoring of the work of the Inspectorate to ensure that it is focused on school quality and improvement and does not deviate towards the compliance-oriented audits carried out by the Committee.

The OECD also recommends that Kazakhstan establish a regular reporting channel between the Inspectorate to the Minister of Education and other members of the government. The Inspectorate could produce an annual public report that, similar to Ofsted annual reports, would focus on its governance, finances and performance of the Inspectorate against key objectives (Ofsted, 2018<sub>[16]</sub>). This annual report

that would allow the government to monitor the quality of the work of the Inspectorate and act as a reference point during government debates.

### **Ensure that the School Inspectorate is sufficiently financed**

The School Inspectorate should have a dedicated, sustainable funding stream that it can use to develop and refine School Review procedures and tools and to provide training to evaluation teams. With its own budget line, the functioning of the School Inspectorate would not be tied to other budget items, which would help establish its legitimacy and ensure its continuity.

## **1.2. Maintain the Committee's audit responsibilities, but co-ordinate audit activities with School Review**

### *Evidence*

With the establishment of the School Inspectorate, there is a question of what the Committee's role would then be with respect to school oversight. Some former Soviet countries, such as Azerbaijan, abolished the control system after independence on the grounds that it lacked societal trust and integrity and integrated many compliance functions into its school evaluation system. Other transition countries, such as Romania and Serbia, have introduced a new school evaluation framework alongside existing control systems (Kitchen et al., 2017<sup>[17]</sup>; Maghnouj et al., 2019<sup>[11]</sup>). While this configuration creates more space to focus evaluation on educational practices, it also raises challenges related to the administrative burden on schools and the reinforcement of a compliance-based culture.

In the presence of two systems for school evaluation and school audit, streamlining data collection is essential to avoid duplicating efforts. Audit results contain information about a school's physical characteristics, availability of resources, and compliance with regulations that can help determine the extent to which school environments are safe and conducive to learning. Knowing this information would help evaluators form contextualised judgments about the school. For these reasons, in some countries, such as New Zealand, a school's audit report is given to evaluators for review before a school is externally evaluated (Education Review Office, 2019<sup>[18]</sup>). Thus, schools do not have to submit the same information twice and evaluators do not have to waste valuable time during school visits examining school conditions that have already been reviewed as part of the school's audit.

In Kazakhstan, if School Review is introduced, school audits would continue alongside it. However, there are no plans for how the two activities would be co-ordinated. A school could, in theory, be audited and undergo a school evaluation visit at the same time, which would distract them from helping their students learn. This confluence would further lessen the distinction between audits for compliance and school evaluation for improvement. Finally, examining school audit data is not part of external evaluation procedures, which contributes to evaluators' judgments being decontextualised from the schools' environments.

### *Recommended actions*

#### **Maintain school audit as a separate function from school evaluation, but lessen its frequency**

In Kazakhstan, many schools are far and remote, which limits the oversight that central authorities have over their activities. Moreover, there are strong concerns regarding the fundamental operations of schools, including their basic physical infrastructure and the provision of educational materials. The OECD agrees that the Committee should continue fulfilling some of its audit functions so all schools receive some degree of oversight. However, the OECD recommends that the weight of school audits be lessened in order to

reduce their burden on schools. For instance, the Committee can consider changing the frequency of preventative control (i.e. conducting the risk assessment) from a one-year cycle to a two-year cycle, or consider performing the risk assessment for half of all schools each year. Furthermore, Kazakhstan should consider reducing the number of criteria used for preventative control school visits, which is currently 45. For instance, some of the criteria around teacher performance can be reduced, given that School Review will be looking carefully at teaching and learning.

### **Co-ordinate school evaluation and school audit visits**

If Kazakhstan decides to further develop and introduce School Review, this review recommends that school audits visits and School Review visits should be co-ordinated so schools do not have to prepare for two consecutive evaluation visits. The OECD recommends that, as far as is possible, School Review visits and preventative control visits not occur during the same school year. Unplanned inspections could still occur in an ad hoc manner in response to emergencies, though Kazakhstan should consider if it is necessary to conduct a preventative control visit to a school that recently received an unplanned inspection.

### **Make audit data and reports available and use them as part of external evaluation**

As audits identify schools with critical needs and generate crucial information about the school environment, the OECD suggests that external evaluation should use information generated from audits to help determine school quality. Providing Kazakhstani evaluators with the most recent school audit reports (from school attestation, preventative control, and unplanned inspection) would help them understand the environment of a school and its community so they can make more contextualised judgments about the school's performance. Furthermore, where appropriate, the audit information itself can be used directly to evaluate schools against school standards. This process would streamline data collection across school audit and school evaluation and avoid overburdening schools.

Audit data are currently stored in the regional offices of the Committee and are not accessible to other agencies. Although the regional offices report bi-annually to the central office, the reports are of aggregate indicators and the school-level microdata are not shared. To make audit data available to more stakeholders, Kazakhstan should consider allowing access to non-sensitive data to the School Inspectorate, or link the databases of the regional Committee offices with NED.

## **Recommendation 2. Revise School Review standards to focus on the main challenges in Kazakhstani schools**

Kazakhstan's potential School Review standards benchmark positively against school standards in OECD and other European countries. The fact that they were also piloted before they were finalised suggests that they were created according to rigorous procedures. As part of this review, the OECD performed an in-depth analysis of the standards and determined that, while very comprehensive, the School Review standards can still be updated to reflect Kazakhstan's current context and the latest international practices. In particular, the diversity of school environments in Kazakhstan means that some standards are focusing on criteria that are not crucial to understanding the teaching and learning that are occurring in a school, or are not flexible enough to be relevant to all contexts. In these cases, schools will not only have a hard time meeting the standards, but, more importantly, evaluating these schools against the standards will not produce relevant information about the what the schools can do to improve.

## **2.1. Develop the School Review standards to better reflect the education context in Kazakhstan**

### *Evidence*

The objective of school standards is to communicate national expectations for school quality. Standards need to be ambitious and give direction and impetus to school improvement. They should be based on evidence about what makes for effective schooling and also be aligned with the curriculum and teaching standards, so that the school evaluation system supports the national education agenda. However, they, and the criteria they set, also need to be accessible to schools and take into account the complete range of schooling environments.

The School Review standards cover some of the main factors that research suggests are important for school quality such as teaching practices, school planning and management, students' outcomes and progress (OECD, 2013<sup>[6]</sup>). Another positive element is that the measurement of learning outcomes is reasonably balanced, with national and school assessments being examined, as well as student performance over time instead of just one year. Self-evaluation is also addressed in the School Review standards, which signals the importance of this activity. However, some parts of the standards do not focus attention on the most critical components of schooling in Kazakhstan, and others might not be accessible for all schools.

### **School review standards do not focus strongly on staff capacity**

The School Review standards lack focus on developing the capacity of school staff, in particular school leadership, which is a well-documented challenge in Kazakhstan. For instance, the “school management” domain has the criterion “professional development and motivation of teachers”, and the “teaching quality domain” has the criterion “accurate and in-depth knowledge about the subject”. With respect to school leadership, though, there are no criteria that focus on their capacity. The standards do not check if principals have the knowledge needed to be instructional leaders, or if they are developing themselves to be more effective change agents.

Similarly, to the extent that the standards focus on capacity, the focus is on knowledge of their practice and their field (e.g., teaching a specific subject). Other types of capacity, however, are also important to for effective schooling, such as the capacity of the school leadership to organise professional development for teachers and to carry out self-evaluation. These criteria are not present in the School Review criteria. Another important criterion is capacity to align school activities with national aims. For example, new teacher standards were introduced in 2017 (see the policy perspective on initial teacher education). It is important that those standards act as a reference for schools, such as during teacher appraisal and for teachers' professional development activities. There is, however, no criteria in School Review about integrating national standards in the school.

### **Learning criteria and descriptors are narrow and de-contextualised**

The criteria and descriptors for the domain learning outcomes primarily consist of results in competitions and on summative tests, including one specifically administered by evaluators during a school visit (the “cross-sectional test”, see Table 1). This focus suggests that learning is narrowly understood to be the acquisition of academic knowledge. Other than the criterion “personal development”, the learning outcomes domain provides very limited coverage of all the skills students need to achieve in school and in life, such as socio-emotional skills and participatory citizenship. These criteria feature prominently in international standards. For example, New Zealand's school evaluation standards include, as criteria, “Socially and emotionally competent, resilient and optimistic about the future” and “Participates and

contributes confidently in a range of contexts – cultural, local, national and global” (Education Review Office, 2016<sup>[19]</sup>).

Moreover, the descriptors that do focus on learning outcomes look exclusively at raw performance and not the circumstances of the schools or the value they add to student learning. For instance, the descriptors for the criterion “External evaluation results” indicate that schools must have assessment results that are higher than averages at the national, regional and local levels in order to be rated as “Excellent”. Similarly, their students must win national and international competitions. Such descriptors are likely to disadvantage schools with challenging contexts (e.g., rural and ungraded schools) even if those schools are performing admirably given the challenges they face.

### **School review standards lack flexibility and relevance in some contexts**

Although Kazakhstan’s school standards are rigorous, the way that some are phrased might be too rigid for them to be inclusive of all Kazakhstani school environments. For example, the “library resources” criterion in the “resources” domain aims to evaluate schools’ educational resources, which is positive. Nevertheless, the way it is written limits its descriptors to focusing on a dedicated library (descriptors for this criterion refer to a library fund and library stock). Small remote schools might share resources or provide them through community donations, but might not have the needed space or funding to construct a conventional library. If evaluated according to these descriptors, those schools would be unfairly penalised.

There are also questions regarding whether some descriptors are relevant to the quality of teaching and learning in some school contexts. In the same “resources” domain, descriptors for the “Building conditions” criterion refer to a modern dining room. In some schooling environments, such as large crowded schools, a designated dining area can help support student safety and health. In other environments, though, such as very small schools, students would not be significantly impacted if they do not have such an area. In these cases, evaluating and reporting on the quality of the school’s dining room does not communicate the school’s main challenges. Its actual needs would still remain unknown centrally, which would prevent the government from providing meaningful support to it.

### **Equity and inclusion do not feature prominently in the standards**

Given the noted learning disparities in Kazakhstan, one would expect the School Review standards to focus heavily on student equity and inclusion. However, the standards do not suggest that equity is a primary focus. No descriptor in the learning outcomes domain mentions the performance of student sub-populations, nor is there a descriptor in “quality leadership in management” that deliberately looks at school practices that support equity and inclusion. The absence of these standards contrasts with international school standards. The Scottish school evaluation framework (Box 2) has five criteria related to school leadership, and one is explicitly about managing resources to promote equity.

### **A narrow, quantitative rating system could distort school behaviour**

The rating system of the potential School Review standards calculates a single, overall score using a system of coefficient weights. This quantitative approach is rigorous and transparent, but could push schools’ focus away from the quality of practices towards the score the school receives. Schools might, for instance, ignore criteria that it thinks are harder to achieve for them and concentrate on maximising their ratings in other criteria to attain the highest possible score.

This recommendation has already discussed that the Kazakhstani School Review standards do not emphasise staff capacity, are not always applicable to all environments, and focus more on overall summative outcomes instead developing equitable education opportunities. Given these circumstances, introducing a quantitative scoring system could incentivise schools to focus on improving the test and/or

competition results of their highest performing students and not on strengthening staff capacity to improve the learning of all students.

### *Recommended actions*

#### **Introduce criteria about staff capacity and equity and eliminate inequitable tools**

The OECD recommends that Kazakhstan revise the potential School Review standards to focus more strongly on the development of staff capacity. For instance, school leadership capacity should be explicitly included, as well as their participation in professional development. Furthermore, the extent to which school leaders and teachers understand and apply the new teacher and school standards should factor into the overall evaluation of the school.

School Review standards must also focus more strongly on equity. The standards should include clear criteria addressing the commitment and strategy of the school to ensure equity and inclusion. They should also highlight teaching practices to adapt instruction to the different characteristics and needs of students, and in particular to vulnerable subgroups of students such as socio-economically disadvantaged students. The “learning outcomes” domain should have criteria that focus on the performance of different student groups within the school, and not just the entire school.

Sources of evidence that are inherently inequitable should not be included in the standards as criteria or sources of evidence. These include student performance in competitions and de-contextualised results on summative tests, such as the cross-sectional test. The policy perspective about the national assessment recommends that the EAAA be administered on a census basis, which will enable the production of school-level results. These results, reported in a contextualised manner, can be used as a source of evidence. With school-level EAAA results, it would no longer be necessary for evaluators to administer a separate school-level test as part of School Review, and that test should be eliminated.

#### **Make school standards more flexible and relevant in consideration of all schooling environments in Kazakhstan**

Kazakhstan should improve the flexibility and relevance of the School Review standards, especially the criteria and descriptors in the “resources” and “care and support” domains. One step would be to make the language more flexible so it can be applicable to more schooling contexts. The “library resources” criterion, for instance, could be changed to “access to educational materials” so schools are not expected to have a dedicated library in order to meet the criterion. Descriptors for this criterion, instead of referring to a library fund and library stock, could instead refer to a school’s efforts to provide access to more and better educational resources.

In cases where descriptors cannot be revised to consider all possible environments, Kazakhstan can consider including benchmarks within a descriptor that are specific to certain school circumstances (e.g., ungraded, multi-shift or very remote schools). An example of these descriptors could be those that refer to internet access, which are under the “Logistics” criterion of the “resources” domain. Access to the internet is vital to excellent schooling and it is difficult to write a descriptor that flexibly interprets connectivity. However, some schools might have limited or even no internet access because of the lack of connectivity in their surrounding communities. To avoid evaluating these schools based on something they cannot control, the connectivity descriptor can have a different benchmark for these schools, such as whether they make efforts to collaborate with schools that do have internet access.

Since these descriptors will be used to judge school quality, it is important that they be revised with the schools in mind. School descriptors in New Zealand and in Scotland were designed in this manner, with the input of practitioners, and serve as good examples of school-centric descriptors (Education Scotland, 2015<sub>[20]</sub>; Education Review Office, 2016<sub>[19]</sub>). In Kazakhstan, principals and teachers from different



schooling environments should be asked to provide input into the revision of the descriptors to ensure that it makes sense to them and can be used by them when evaluating the quality of their own schools.

### **Diversify the learning outcomes standards**

In addition to students' academic outcomes, Kazakhstan should consider including criteria that relate to students' socio-emotional skills, engagement in class and their approach to learning. Descriptors related to these criteria should not focus solely on quantitative measurement, such as tests, but how schools can help students develop these skills. Evaluators would then use these descriptions to form an evidence-based, qualitative judgment of school quality in these areas. In the Scottish school evaluation standards, a descriptor of the criterion "Overall quality of learners is:

*"Overall, our learners are successful, confident, exercise responsibility and contribute to the life of the school, the wider community and as global citizens. They are personally and socially adept and have achieved a range of skills and attributes through a wide range of activities. As they move through their learning pathways they take increasing responsibility for ensuring they continue to add value to their achievements"* (Education Scotland, 2015, p. 51<sup>[20]</sup>).

### **Shift the rating system towards a qualitative approach based on core and secondary criteria**

Kazakhstan should eliminate the quantitative rating methodology that produces a single, overall score for a school, including coefficient weights across School Review criteria. In its place, the OECD recommends that Kazakhstan develops a qualitative approach to evaluating schools. The aim of this approach is to ensure that all criteria and domains are taken seriously by schools and reviewed carefully by evaluators. A school should not be able to, for example, have a very low rating in an entire domain and still be considered an excellent school because it has high ratings in other domains.

Implementing a qualitative approach is associated with several challenges. While potentially more fair and nuanced than a purely quantitative approach, evaluators will need to be trained effectively and provided with high-quality tools to ensure that their judgments of schools are reliable and comparable. Recommendation 3.2 provides further guidance about how to ensure the integrity of a qualitative approach to school evaluation.

Since schooling environments differ substantially in Kazakhstan, the government should consider establishing two tiers of criteria within the School Review standards according to their importance to school quality and relevance to different schooling contexts. Core criteria would be considered essential to schooling in all environments and mandatorily evaluated in all schools. These might include criteria related to basic resources, staff capacity and learning outcomes. Secondary criteria would be considered those that influence school quality in some contexts more than others, and thus would only be evaluated in some schools (e.g., dining areas). Importantly, achieving the minimum standards of satisfactory schooling should be within reach for all schools through only meeting core criteria.

Whether secondary criteria would be evaluated at a particular school would depend on evaluators' judgments of that criteria's applicability. If a criterion is determined to be inapplicable to a school, the school's evaluation report would indicate that this criterion was not evaluated, not that the school failed to meet standards. In this manner, schools will not be unfairly punished for not meeting standards that are not applicable to their contexts, but schools for which the standards are applicable can still demonstrate excellence in those areas.

## 2.2. Reduce the number of School Review standards and criteria to enable a more in-depth evaluation of essential quality dimensions

### Evidence

As is the case with all school evaluation components, it is important that school standards enable the evaluator to focus on the core components of effective schooling and to have a good understanding of actual practices in schools. However, too many standards and criteria can turn school evaluation into a superficial check-box exercise where evaluators have limited time to consider a long and fragmented set of criteria that do not all focus the essential dimensions of school quality. Moreover, having too many standards can be cumbersome to schools and motivate them to regard school evaluation as inconvenient rather than engage with it meaningfully. Internationally, as countries have gained experience and schools become more familiar with the process, they tend to simplify school standards. The school standards in Scotland, for example, have only three domains (Leadership and Management, Learning Provision and Successes and Achievements) and 15 total criteria.

Although Kazakhstan’s school evaluation system is much newer than Scotland’s, the 41 criteria in Kazakhstan’s potential School Review standards is still disproportionate. More importantly, the criteria are less focused on the teaching and learning that occurs within the school and less flexible (see Box 2 for a comparison of the criteria for the domain “school leadership and management”). The long list of fragmented criteria might prevent evaluators from meaningfully evaluating school quality and using School Review as an exercise for improvement. From the school’s perspective, the criteria might be burdensome to fulfil and distract them from focusing on helping students learn.

### Box 2. School management criteria in Scotland and Kazakhstan

The table below presents a comparison of the criteria measuring the quality of leadership and school management across the School Review standards in Kazakhstan and the school evaluation framework in Scotland, “How good is our school?”.

Domain	School evaluation criteria in Scotland	School evaluation criteria in Kazakhstan
School leadership and management	<ul style="list-style-type: none"> <li>• Self-evaluation for self-improvement</li> <li>• Leadership of learning</li> <li>• Leadership of change</li> <li>• Leadership and management of staff</li> <li>• Management of resources to promote equity</li> </ul>	<ul style="list-style-type: none"> <li>• Analysis and self-assessment of activities</li> <li>• Understanding the mission and strategic development plan</li> <li>• Distribution of responsibilities and teamwork</li> <li>• Professional development and motivation of teachers</li> <li>• Monitoring and evaluation of the quality of teaching and education activities</li> <li>• The moral and psychological climate in the organisation of education</li> <li>• Compliance with laws and regulations</li> <li>• Planning and organising the current activities of the school</li> <li>• Resource usage</li> </ul>

Source: Education Scotland (2015<sub>[20]</sub>), How good is our school? 4th edition. [https://education.gov.scot/improvement/documents/frameworks\\_selfevaluation/frwk2\\_nihedithgios/frwk2\\_hgios4.pdf](https://education.gov.scot/improvement/documents/frameworks_selfevaluation/frwk2_nihedithgios/frwk2_hgios4.pdf) (accessed 27 January 2020).

A related issue is that some of the School Review criteria more closely resemble sources of evidence. For example, student employment is better understood as a data point than a criterion that can be applied to all schools. Similarly, student results on tests are not criteria in and of themselves, but evidence that can be used to evaluate a broader criterion, such as student learning. Including these evidence sources as

criteria makes the number of criteria unnecessarily long and can also compel schools to narrowly focus on improving their results on these measures, which can distort their teaching and learning practices.

### *Recommended actions*

The OECD recommends that Kazakhstan revise its school standards so they are fewer and broader and emphasise the core aspects of school quality, which would help make school evaluation a meaningful exercise. Kazakhstan should consider removing fragmented criteria that do not directly relate to teaching and learning, such as “resource usage” and “compliance with regulations”, which can be captured through school audits and referenced during external evaluation. Some criteria overlap can be merged, such as “distribution of responsibilities and teamwork” and “planning and organising the current activities of the school”.

Introducing some of the standards suggested in Recommendation 2.1 can also help reduce the number of overall criteria by consolidating them into fewer but more overarching criteria. For example, it was recommended that Kazakhstan add criteria about the extent to which teachers’ practice is aligned with national teacher standards. Such a criterion might be called, “Alignment of teacher’s skills and competences with the national teacher standards” and could encapsulate several current criteria, such as subject matter knowledge, lesson planning, assessment and motivation (these are all covered in the teacher standards).

Another way of reducing the number of criteria is to remove some of those that are actually sources of evidence and replacing them with higher-level criteria that are measured using those sources of evidence. As many of the criteria in the “learning outcomes” domain fall under this category, Kazakhstan should consider introducing, as higher-level criteria, the different types of skills and attitudes students are expected to develop that were suggested in Recommendation 2.1. Some current criteria (e.g., employment rates), could be reclassified as sources of evidence that are consulted to determine if schools have met standards in the higher-level criteria.

### **Recommendation 3. Create an external evaluation approach that is efficient in the Kazakhstani context and considers all relevant information**

According to the potential School Review, all schools would be evaluated at least once within a five-year period and how frequently schools would be evaluated would be based their previous rating. The OECD supports this differentiated approach as it would more strongly support schools that need it the most. However, given the size of the Kazakhstani school network, it is doubtful that all schools can be meaningfully evaluated within such a short timeframe. Furthermore, there is lack of clarity regarding how schools would be selected for their first external review. A final concern is that the external evaluation process would not be informed by important sources of available information, such as school audits. Disregarding this information would result in less accurate evaluation judgments and increase the burden on schools to produce the information themselves.

### **3.1. Aim to evaluate all schools over a long timeframe and use a risk-based, desk scan to direct support to schools that need it most**

#### *Evidence*

#### **Externally evaluating all schools over a narrow timeframe would stretch limited resources and might draw support away from schools that need it most**

Several well-established school inspectorates, such as Ofsted in the United Kingdom and the Inspectorate of Education in the Netherlands, have school evaluation frameworks that require all schools to be externally evaluated within a certain time period (SICI, 2009<sup>[21]</sup>; SICI, 2018<sup>[22]</sup>). This model is appropriate for these countries because they are geographically small, making schools easier to visit. These inspectorates are also among the most mature in the world, having supported their countries' schools for several decades. There is little concern that a large number of schools in these countries are struggling to meet basic minimum standards and are in need of intensive support.

In Kazakhstan, however, it would be very challenging for a new school inspectorate to meaningfully evaluate more than 7 000 schools in only five years. The country is large and, while school attestation was in effect, the Committee struggled to visit all schools in a given timeframe. This challenge would certainly be magnified with an external evaluation process that requires more qualitative evaluation and judgment than was the case with school attestation. Additionally, effectively evaluating schools requires a cadre of highly qualified external evaluators. In Kazakhstan, this position has not previously existed and it will be difficult to quickly create such a cadre.

Finally, there is widespread understanding that many schools in Kazakhstan are in need of critical support to meet basic standards. These could be urban schools that are overcrowded and operating three shifts, or rural schools lacking in vital infrastructure. In this type of environment, trying to evaluate all schools externally is less advisable. Not only would such comprehensive external evaluation be logistically difficult to accomplish, but it would draw support away from the schools that need it the most.

#### **The criteria for determining which schools receive an external evaluation are limited and do not identify which schools should be evaluated first**

An important element of a differentiated school evaluation approach are the school-level indicators that are reviewed to determine when and how schools should be externally evaluated. Information that is commonly used to construct the criteria include the results of previous external evaluations, administrative data and performance on national assessments. Since the indicators are all constructed using centrally available data, they can be regularly reviewed via desk scans (this process is commonly referred to as a “risk assessment”, but is called a “desk scan” here to avoid confusing it with the risk assessment that is part of preventative control, which is a similar procedure). Schools then receive external evaluations based upon how their indicators compare to pre-determined thresholds.

Internationally, the indicators that comprise a desk scan focus strongly on outcomes in order to reinforce the notion that school evaluation is about school quality and improvement (Ehren and Shackleton, 2016<sup>[23]</sup>). Components of equity, such as the performance of different student groups, are often included in desk scans to support education for all students.

In Kazakhstan, the criteria that would be used to determine differentiated school evaluation raise some concerns. First, there is currently no plan that determines which schools will be evaluated first before schools have external evaluation ratings, which risks that the neediest schools would not be those that are first evaluated and supported. Second, relying solely on previous evaluation ratings to determine future evaluation frequency assumes that schools encounter minimal changes between evaluations. However, this is not always true. A school that receives an “excellent” rating could encounter changes in student

intake resulting from population patterns or the consolidation/closing of neighbouring schools. The current approach would not identify these changes for quite some time, and the school would not receive needed support in a timely manner.

### *Recommended actions*

#### **Evaluate all schools over a reasonable timeframe and use a desk scan to help determine which schools are evaluated each year**

The OECD recommends that Kazakhstan evaluate all schools over 10-15 years with a sample of about 700 schools being externally evaluated every year. The sample would be divided into two groups of schools. The first group, which would be the smaller group, should be representative of different levels of quality schooling so evaluators are able to continuously calibrate their judgment vis-à-vis the complete range of schooling quality. The second group, which would be the larger group, should be comprised of schools that are most in need of support, to ensure that important resources are directed to where they would be most helpful.

To determine which schools should be part of the sample, Kazakhstan should introduce a desk scan to continuously collect data about schooling indicators. While there would likely be overlaps in terms of the data needed for the desk scan and the data used for preventative control's risk assessment, the criteria would be different. The desk scan would aim detect schools with a range of educational processes and outcomes, not just those that fail to meet compliance. The two aforementioned external evaluation school groups would be formed based upon these results. The first desk scan would determine which schools should first receive external evaluations. Through this method, Kazakhstan would have a regular understanding of which schools are most in need of support and be able to prioritise support to them in a timely manner.

#### **Develop indicators that form the basis of the desk scan**

The first step to creating an effective desk scan is to identify relevant indicators that will be used as part of the scan. In consideration of the Kazakhstani context, the focus of the School Review standards, and based on the experiences of other countries, the following could be indicators that are included:

- Student outcomes. Results from the EAAA, gathered from the National Testing Centre's (NTC) database, can be reviewed, which will necessitate that the EAAA is administered on a census-basis (see policy perspective on the national assessment).
- Student profile. The desk scan should reflect the context in which a school is operating. Using student demographic data will highlight schools where there is a concentration of students at greater risk of low performance given their backgrounds. Relevant information about student demographics include age, grade, immigrant status and socio-economic background (as measured, for instance, by the share of students from families who receive financial support from *Vseobuch* or from local/regional budgets and the share of students whose mother tongue is not the language of instruction).
- Staff. Data from NED can be used to provide information on the share of teachers in each school by age, gender, teaching status, and participation in professional development.
- Physical infrastructure and resources. Because physical infrastructure is a key concern in Kazakhstan, indicators that capture basic student health and safety should also be included in the desk scan as these factors affect the quality of schooling.

All of these data can be acquired through central databases, but the data related to physical infrastructure and resources represents a unique case in the Kazakhstan context. In some systems, information about basic infrastructure is available through the same national databases as student and teacher data (Li

et al., 2019<sup>[24]</sup>). In Kazakhstan, while NED includes some data that related to physical infrastructure and resources, much of this information is gathered during school audits and stored in regional offices of the Committee. As suggested in Recommendation 1.2, school-level audit data should be made more available, which would allow it to be used as part of the desk scan.

**Determine the thresholds of the desk scan in consideration of the external evaluation capacity of the recommended School Inspectorate**

After the indicators for the desk scan have been finalised, it will be necessary to set quality thresholds to determine which schools should be externally evaluated. Since the sample will represent the full range of quality schooling in Kazakhstan, a series of thresholds should be established to reflect different levels of schooling (e.g., excellent, good satisfactory, and in need of improvement). Box 3 describes the school evaluation system, including desk scan indicators and criteria, in the state of Alaska, United States. As a large and sparsely populated state with many small schools, Alaska has developed procedures to understand schooling across the state while focusing resources on schools that need it most.



### Box 3. School evaluation in the state of Alaska, United States

In the United States, each state monitors school quality according to regulations set out in the federal Every Student Succeeds Act (ESSA). According to this law, each state must regularly collect information via a desk scan according to five broad indicator areas, which states can interpret for their own contexts:

- Academic achievement, as measured by proficiency on the annual assessments
- Student growth
- High school graduation rates
- Progress of English language learners toward proficiency
- At least one indicator of school quality or student success (SQSS). It may include, for instance measures of student engagement, educator engagement, student access to and completion of advanced coursework, postsecondary readiness, school climate and safety.

All indicators must be measured annually for all students at the school and relevant subgroups. The subgroups include gender, racial and ethnic groups and economically disadvantaged students. At the state-level, the school quality indicators of Alaska are shown in Table 2.

**Table 2. Alaska school accountability indicators**

ESSA indicator area	Alaskan indicator(s)
Academic achievement	<ul style="list-style-type: none"> <li>• Percentage of students scoring at the proficient or advanced achievement levels on the statewide assessments in English language arts for all students and all subgroups</li> <li>• For schools that serve Grade 12, the Academic Achievement indicator also includes student growth, as measured by the annual statewide English language arts and mathematics assessments</li> </ul>
Student growth	Academic growth on the statewide assessments in English language arts and mathematics for students in Grades 4-9.
High school graduation rates	Four- and five-year adjusted cohort graduation rates for the all students and for all subgroups present in a school
Progress of English language learners	Percentage of eligible English that meet the definition of making progress in achieving proficiency in English, as measured by the state English Language Proficiency assessment
Indicator of school quality or student success	<ul style="list-style-type: none"> <li>• Chronic absenteeism for all students and subgroups</li> <li>• Grade 3 English language arts proficiency for all students and subgroups</li> </ul>

Based upon the information collected via the desk scan, each school will have an accountability index value calculated. The schools determined to be struggling the most (according to their accountability index values) will receive comprehensive support and improvement (CSI). Some of the remaining schools will receive targeted support and improvement (TSI). Schools that do not receive CSI or TSI, which are considered to have met basic minimum standards, will receive “universal support”.

Schools receive CSI if they meet one of the following criteria:

- Low-income schools that are in the bottom 5% of all such schools
- High schools with a four-year adjusted cohort graduation rate less than or equal to 66.6%
- Schools that have met the TSI eligibility criteria for the same subgroup for three consecutive years.

Schools receive TSI if they:

- Have one or more underperforming subgroups.

Schools that are eligible for CSI or TSI will be visited by the State Department of Education, who will implement a series of interventions. These include an external review of the school's data, evaluation of school improvement practices, and the guided development of school improvement plans. Schools that receive "universal support" will not be visited.

Small schools follow differentiated accountability procedures. Three years of data are aggregated for these schools in order to generate their accountability index values. If a school does not have enough students to generate subgroup-level indicators, even after aggregation (and thus cannot generate an accountability index value), a special small school performance review will be conducted. Schools are then eligible to receive small school CSI or TSI that is adapted to their contexts.

Sources: US Department of Education (2018<sub>[25]</sub>) Elementary and Secondary Education Act of 1965 as amended through P.L. 115-224, Enacted July 31, 2018.

<https://legcounsel.house.gov/Comps/Elementary%20And%20Secondary%20Education%20Act%20Of%201965.pdf> (accessed 30 January 2020);

U.S. Department of Education (2018<sub>[26]</sub>) Alaska Consolidated State Plan.

<https://www2.ed.gov/admins/lead/account/stateplan17/akconsolidatedstateplanfinal.pdf> (accessed 30 January 2020)

In developing its desk scan thresholds, it is important that Kazakhstan consider not only the definitions of quality schooling according to the School Review standards, but also the capacity of the recommended School Inspectorate to conduct external evaluations. If the thresholds are too high, then a large number of schools will be labelled as in need of support and the School Inspectorate might not have the capacity to evaluate all those schools, which diminishes the utility of School Review. If the thresholds are too low, then very few schools will be labelled in need of evaluation/inspection, which will discourage them from taking School Review seriously.

To arrive at appropriate thresholds, Kazakhstan should first determine the external evaluation capacity of the School Inspectorate. How many external evaluations can the agency conduct in a year, and would the number vary depending upon where the schools are? Previous data from school attestation can be consulted to help determine these figures. Kazakhstan can then adjust the thresholds of the desk scan until the number of schools in need of support matches the capacity of the School Inspectorate (see Recommendation 3.3 about piloting School Review).

### **3.2. Provide evaluators with the guidance and tools they will need to reliably evaluate schools externally**

#### *Evidence*

External evaluation requires evaluators to know how to use the evidence they gather in order to make well-informed judgments about school quality. In many countries results on standardised tests help inform judgments of school quality. However, results in national assessments and examinations and socio-economic status are highly correlated. Therefore, "raw" student test data would partly reflect students' socio-economic characteristics and not necessarily the contribution of the school to students' learning (Faubert, 2009<sub>[27]</sub>). Given this risk, evaluators need to be provided with guidance on how to use students' performance data in a balanced manner, and especially on how to contextualise the data. Contextualising students' performance data can be done by providing school-level information on the students that participated to the assessment or examination (e.g. students' socio-economic background) or applying

statistical adjustments on students' performance data to account for the school context (OECD, 2008<sup>[28]</sup>). The policy perspective on the national assessment provides further information about how reporting around national assessments can support this type of contextualisation.

As mentioned in Recommendation 2.1, reviewing results on standardised tests provides only a limited understanding of learning outcomes and school quality. Equally meaningful are the qualitative judgments of expert evaluators on how schools are helping their students learn. To help evaluators form their judgments, school evaluation frameworks often require a range of qualitative data collection activities during school visits, such as classroom observations and interviews with school staff. Given the difficulty of reliably evaluating subjective information, countries develop significant guidance to aid evaluators, such as rubrics for classroom observations to ensure fairness and consistency (OECD, 2013<sup>[6]</sup>).

If School Review is introduced, it would ask evaluators to use a broad range of evidence to form their judgments about school quality. These include reviewing standardised test scores, conducting classroom observations, and reviewing school documentation. It is unclear, however, what resources would be developed that guide evaluators on how to properly gather evidence at schools and use the evidence they collect. Without such support, Kazakhstan's evaluators will have a more difficult time forming accurate and fair judgments about school quality.

### *Recommended actions*

#### **Develop descriptors to help evaluators assess a diverse set of learning outcomes**

If Kazakhstan decides to further develop and introduce School Review, its evaluators will need support in evaluating the diversity of learning outcomes that this policy perspective recommends be incorporated in the School Review standards. The first step in providing this guidance is to develop relevant and flexible descriptors that are associated with the criteria in the learning domains outcome. These descriptors should describe what evaluators are expected to observe for each type of outcome at various levels of school quality. For example, the New Zealand school standards contain a criterion that student be confident in their identity, language and culture. Descriptors for this criterion include, "Students value diversity and difference and represent and advocate for self and others" (Education Review Office, 2016<sup>[29]</sup>).

#### **Create tools that guide evaluators in the collection and use of evidence**

A second step in supporting evaluators is to develop tools that help evaluators conduct school visits and form judgments. These tools include:

- An evaluation manual

Evaluators must be given clear instructions to help them do their jobs properly. Internationally, these instructions are typically contained in an evaluation manual. For example, Ofsted's school inspection handbook which tells evaluators the steps that must be followed, documents collected and how to review the collected evidence to rate a school (Ofsted, 2019<sup>[30]</sup>).

Kazakhstan's evaluation manual should be developed in consideration of the specific challenges that evaluators are likely to face. In particular, the manual should emphasise a balanced use of evidence without overly relying on test scores. The manual should recommend that, in addition to summative test results, evaluators use classroom observations and an analysis of students' work to evaluate learning outcomes.

Regarding the test scores themselves, the evaluation manual should emphasise the need to contextualise school-level results on the EAAA and UNT. Evaluators should not only review the raw results data, but also consider the school's location and the share of students that receives Vseobuch support in order to achieve a more contextualised understanding of the school's performance. Evaluators could as well use

demographic benchmarks to compare the school results with schools that have similar settings (see policy perspective on the national assessment).

- A classroom observation protocol

The OECD recommends that Kazakhstan develop a classroom observation protocol that indicates the purpose of classroom observations, the conduct evaluators need to adopt and what they need to observe during the classroom observations. In the protocol, there should be a set of qualitative measures that help evaluators evaluate teaching practice in a structured way. To develop such an instrument, the Inspectorate might draw on the classroom observation indicators developed by the International Comparative Analysis of Learning and Teaching (ICALT), which are based on teaching and learning practices with a proven impact on student learning (OECD, 2013<sup>[6]</sup>). Specific considerations for inclusion in the classroom observation protocol is how evaluators can look for indicators related inclusion and equity in the classroom, teachers' use of criterion-based assessment, a focus on students' engagement in class and on the development of a positive attitude to learning.

- A reporting template

The OECD recommends that Kazakhstan's School Review reporting template instruct evaluators on how to use the descriptors and consider a school's context when evaluating it. The template should help evaluators not to use the descriptors as a checklist, but to adopt a "best fit" approach that relies on the professional judgment of the evaluation team. Since audit data contain crucial information about the school's environment, the template should emphasise the need to use school audit data to help form contextualised judgments about the school (Recommendation 1.2). To further encourage evaluators to rely on this data, the template could ask the evaluators to describe the school and include key school-level data based on audit data. In New Zealand, each report begins with a description of the school context and contains a standardised list of school-level indicators, which include school size, ethnic composition and whether schools have students with disabilities (Education Review Office, 2020<sup>[31]</sup>).

Moreover, if Kazakhstan decides to divide the School Review standards into core and secondary criteria, the template need to provide guidance on when and how to apply these criteria during the evaluation and rating processes. The template could, for instance, provide examples of cases in which specific criteria may not apply, and how this would and would not affect a school's final rating.

### **3.3. Pilot and refine School Review**

#### *Evidence*

Introducing a school evaluation system can create confusion and anxiety among schools and the general public. Schools will receive ratings following external evaluation and will need to understand what those ratings mean and trust that the ratings are accurate. Given the importance of having a relevant and reliable school evaluation framework, it is important that countries first pilot their framework and, based upon the results of the piloting, refine it before formally introducing it.

Kazakhstan's understanding of school quality has historically focused on compliance and quantitative measures of success, not on the extent to which the school improves student learning. Introducing a significant reform, such as School Review, would make schools anxious and they would be understandably sceptical about the ratings that they receive. For example, a school that is legally compliant, succeeds in competitions, but has inequitable student achievement might wonder why it received a low rating. In this kind of environment, making sure that School Review is reliable, trustworthy and widely understood will be imperative to ensuring its success.

*Recommended actions*

Creating a new and independent School Inspectorate will help build trust in School Review (see Recommendation 1). Nevertheless, it will also be important that the potential School Review framework itself is rigorous and considerate of schools' contexts. Several recommendations have made suggestions about how this can be done, such as through revising the existing school standards and giving evaluators information from school audits. If Kazakhstan decides to pursue the development of School Review and incorporate these recommendations into the school evaluation framework, Kazakhstan should not immediately implement it, but should first pilot the framework, similarly to how it piloted the initial school standards in 2015. Based on the results of the test, Kazakhstan should then refine the framework to ensure its reliability before ultimately implementing it.

**Piloting the school evaluation framework**

Based on international practices, the OECD recommends the following steps for piloting Kazakhstan's school evaluation framework:

- Select a sample of schools to be part of the pilot. Since School Review has not been conducted, it will not be possible to identify a representative sample of school quality. However, demographic variables can be used to improve the sample's representativeness. For example, schools can be selected based upon their region, if they are in rural or urban areas, how many shifts they have, and their performance on the EAAA.
- Perform the desk scan on all schools.
- Externally evaluate all schools, regardless of their results on the desk scan.

**Revising the school evaluation framework**

After all schools have been rated, the results should be reviewed. The OECD suggests that Kazakhstan engage focus groups of teachers and principals from different school types in reviewing the school evaluation framework and its components vis-à-vis the results of the pilot study. Questions that should be asked during the review include:

- Are the school standards relevant? Did they encourage evaluators to look at the most important aspects of school quality? Are there important aspects that are missing?
- Are descriptors relevant to determine the practices schools need to have to be rated as satisfactory for each criterion? Do they define minimum standards that are within the reach of all schools?
- Is the desk scan valid? Were the ratings of the desk scan consistent with the judgments of the evaluators following external evaluation?
- Are external evaluation tools (e.g., classroom observation protocols, evaluation grids) effective? Were they easy to use? Did they collect important information about teaching and learning at the schools?
- Are external evaluation tools sufficient? Did evaluators have the information they needed to form valid judgments? What further information could help them form better judgments?
- Are school ratings reliable, fair and transparent? Do evaluators agree on how schools should be rated according to the standards and using the tools that they have?
- Is the distribution of school ratings appropriate? Are any categories under or over-represented? Is there sufficient capacity to support the number of schools identified as in need of further support (see Recommendation 4)?

Based upon the results of the pilot, the school evaluation framework should then be refined.

## Recommendation 4. Build capacity for school improvement

Without a school evaluation framework, Kazakhstan has historically provided support to schools in a non-strategic manner. Some schools, particularly isolated, rural ones, were too difficult to reach and did not receive support, despite being those most in need. The 2020-2025 State Program for Education and Science Development (SPESD) aims to bridge to the national achievement gaps by targeting support towards schools with low performance and socially vulnerable categories of student (Republic of Kazakhstan, 2019<sup>[32]</sup>). Developing this objective is very positive because it helps identify which schools are prioritised for support, but how the objective will be achieved is unclear. Historically, there has been no national programme to support school improvement, but various, un-coordinated improvement initiatives carried out by different agencies across the country. If school review is introduced, it could help identify which schools are most in need of support. However Kazakhstan would need to determine how exactly to support those schools, which organisation is best positioned to do so and build the capacity of this organisation. Moreover, because not all schools would receive support following external evaluations, Kazakhstan would also need to focus on how to encourage schools to improve themselves and each other.

### 4.1. Establish a strategy and select a suitable agency to provide effective support to schools that need it most

#### *Evidence*

Providing relevant and effective support to schools is especially important in Kazakhstan because of the widely acknowledged gaps in school capacity and the inconsistencies in past support methods. Achievement disparities have widened over time and a significant, concerted investment in supporting struggling schools is needed to reduce those disparities. The potential School Review framework is oriented towards helping schools to improve, but there are gaps related to support provision

#### **There is no co-ordinated strategy about how to support the neediest schools**

Research on school support highlights the importance of the fit between a school's challenges and the responsiveness, intensity, stability and timeliness of the support provided to improve those schools (Boyle et al., 2000<sup>[33]</sup>). While the same indicators could identify two schools as struggling, the reasons those schools are struggling could differ considerably (e.g., lack of teacher capacity on one hand, or inadequate materials on the other). The support that they receive would need to be tailored to their individual contexts, and should be based on a national understanding about what type of support is most effective in those types of contexts. In Colombia, for example, the "Let's All Learn" programme aimed to strengthen teaching skills in rural schools and supported over 90 000 teachers in over 4 000 schools (OECD, 2016<sup>[34]</sup>)

In Kazakhstan, School Review, if it is introduced, would identify schools that need support, but those schools will need different types of support that are relevant to their contexts (e.g., ungraded schools will need different assistance than triple shift schools). However, at a national level, there is currently no clear improvement strategy about how these different types of schools should be supported.

#### **Neither the Committee nor its evaluators have the capacity to fulfil a school support function**

Determining which agency should provide support is a pivotal decision. The proposed School Review expects the Committee, through its regional offices, to support schools following external evaluations. As mentioned previously, the Committee is already overstretched, its staff do not have a background in providing support to schools, and schools regard the Committee anxiously. The Committee will not be able to provide the support to schools that they critically need.



Furthermore, School Review evaluators would not be well-positioned to provide support to schools. First, having the same individuals, especially if they are contracted, both evaluate schools and provide support to them might create conflicts of interest between those two functions. Second, the evaluators themselves are often principals and teachers who must return to their permanent posts once an external evaluation has concluded.

**The NIS Centre for Excellence is highly qualified but does not have a role in the School Review framework**

The NIS Centre of Excellence is a high-capacity organisation and experienced with providing support to schools. However, in spite of its regional offices, it does not have a presence in all areas of the country. According to School Review, any school in the country could be selected to receive direct support and it is likely that schools in difficult to reach areas of the country will be those identified as most in need of support. The NIS Centre for Excellence is not well situated to help these schools. Nevertheless, the organisation has valuable experience in helping schools develop. It leads several online training courses and trains “leading schools” to build the capacity of struggling schools. Despite its high capacity, the NIS Centre for Excellence has not been given a formal role in the potential School Review framework.

**Methodological cabinets have a wide reach and are experienced with school support, but do not have a clearly defined role in the School Review framework**

Internationally, the role of middle tier bodies, such as school districts and local authorities, is evolving towards a strengthened role in building the capacity of schools and sharing knowledge (Munby and Fullan, 2016<sup>[35]</sup>; Hargreaves et al., n.d.<sup>[36]</sup>). These bodies are well-suited for this role because they are located throughout the country and many of their staff come from educational backgrounds. For instance, the responsibilities of local and regional oversight bodies in England, Poland, Ontario (Canada), Hong Kong, South Korea and Wales have been redefined to be more educationally supportive of schools (Mourshed, Chijioke and Barber, 2010<sup>[37]</sup>; OECD, 2018<sup>[38]</sup>).

In Kazakhstan, the organisations that have both experience with providing school support and a nationwide network are also middle tier bodies - the methodological cabinets. These bodies are based in akimats and have well-established relationships with schools. OECD interviews with principals and teachers revealed that the methodological cabinets are spoken of positively and relied upon as sources of pedagogical expertise, though the level of expertise varies between akimats. Despite their experience and position, however, the role of methodological cabinets has not been clearly defined in the School Review framework.

*Recommended actions*

**Use the first external evaluation results to create a school improvement programme**

Kazakhstan should create holistic improvement programme to support schools across the country. To create this programme, Kazakhstan should analyse the results from the first cycle of external evaluations to understand the specific, local challenges that the neediest schools are facing. Based upon this analysis, Kazakhstan can determine what typologies of support need to be developed to help different schools improve. Some initiatives could focus, for instance, on improving education in ungraded schools, or overcrowded urban schools. These methods would form the basis of Kazakhstan’s overall school improvement programme and be implemented in schools following School Review external evaluations.

### **Establish the methodological cabinets as the primary bodies responsible for school support and build their capacity**

If Kazakhstan decides to further develop and introduce School Review, the OECD recommends that methodological cabinets, and not the Committee, be responsible for providing school support according to the results of external evaluation. However, while the cabinets are familiar with supporting schools, they are not familiar with supporting them in a systematic manner that is guided by central standards. Therefore, their capacity to provide this kind of support will need to build.

Methodological cabinets need to be provided with the standards, descriptors and evaluation guidelines that evaluators receive. Co-ordination between the methodological cabinets and ministry agencies will also be essential in order to ensure that cabinets receive the school evaluation reports and information needed to plan their work and provide effective support to schools. For instance, a debriefing meeting between the local methodological cabinet and the evaluation team could be organised following the evaluators' school visits in the rayon.

Kazakhstan will also need to help methodological cabinets understand how to support their local schools. To achieve this objective, the Ministry of Education should develop national guidance that clearly sets out the roles of the methodological cabinets in relation to School Review and what procedures they should follow to support schools. For example, methodological cabinets can help schools develop school improvement plans based upon the principles outlined in the national school improvement programme. The cabinets can then help schools monitor the implementation of their plans through follow-up school visits. Since methodological cabinets have strong pedagogical expertise, they could also deliver training directly to teachers and school principals.

### **Leverage the expertise of central agencies in developing the school improvement programme and supporting the methodological cabinets**

Kazakhstan has several national agencies (e.g., the NIS Centre of Excellence and Orleu) with expertise in school support, but these agencies do not have a formal role in School Review. While they are not well-positioned to be the primary support providers, it would be valuable to leverage their expertise in designing efforts to support schools. As a first step, these two organisations should have significant roles in helping Kazakhstan create its school improvement programme.

In addition to designing school support initiatives, central agencies can provide support to methodological cabinets, who will inevitably encounter challenges in their new roles. Kazakhstan should consider creating channels of communication between the methodological cabinets and key educational support bodies, such as those already mentioned and initial teacher education providers. This arrangement would allow cabinets to consult the expertise of those institutions and be aware of new education reforms in order to better support teachers and schools. Where appropriate and available, the methodological cabinets could orientate teachers and principals towards trainings delivered by central agencies.

## **4.2. Develop a system of peer learning across schools**

### *Evidence*

Peer learning between schools is an effective mechanism to incentivise school-led improvement (Pont, Nusche and Moorman, 2008<sup>[39]</sup>). Peer learning can take many forms, from direct and regular contact, to online interactions, to irregular engagement through conferences and workshops. Box 4 provides an example of a peer-learning network that was implemented in Serbia, which offers a good example of how school evaluation can be used to establish school partnerships.

#### Box 4. The SHARE programme in Serbia

The SHARE project, a joint project of UNICEF and Serbian education bodies, is the first initiative in Serbia aiming to create learning communities and peer learning between schools. SHARE aims to improve the quality of education by developing horizontal learning between schools and developing schools' and teachers' agency to learn and lead change in the education system. The initial phase of the project took place between 2015 and 2017 with 20 schools, 1 080 teachers and 12 665 students participating across Serbia. The project paired 10 schools that performed very well in the external school evaluation (score of 4), known as “model schools”, with 10 schools that performed weakly (score of 2 or 1), known as “SHARE schools”.

The project used a reflective approach combining classroom observation and feedback on observed practice. Following the selection of participating schools, classroom visitations are planned to support reflective practice. During this step, teachers, school principals and support staff from the SHARE schools observed between 10 to 15 hours of teaching at the model schools.

Based on a pairing system, the majority of discussions between schools focused on classroom management, lesson planning, teaching techniques, student support, teamwork and preparing for external evaluation. To give constructive feedback during these peer-to-peer sessions, staff in the model schools received training on how to articulate, document and share their success with their paired schools. During the final school visits, SHARE schools were also given the opportunity to present their experience and examples of best practices, thus motivating self-reflection.

The SHARE project initiated and established mutual exchange of knowledge and best practices between schools. It provided schools with hands-on experience through its peer-to-peer learning component. In addition, as a way to enhance the sustainability and long-term benefits of the project, a learning portal was created and shared amongst educators in Serbia. Moreover, 100 practitioners were trained to provide support for quality improvement in low performing schools, creating a network of facilitators who have been integrated into the ministry of education as educational advisors linked to school administrations around the country.

The first phase of the project had a positive impact on the 20 participating schools and show scope for growth and scaling up. A majority of participating schools have seen an improvement in six out of seven areas of quality measured by the external school evaluation. This improvement was mostly seen in the areas of teaching and learning, school ethos and organisation of work and leadership. More broadly, the project introduced participating staff to the concept of horizontal learning and encouraged teachers to work together without the fear of being judged by their peers. It also allowed them to practice new teaching methods and play a more active role in shaping their classroom and school practices.

Sources: UNICEF (n.d.<sup>[40]</sup>), Dare to Share: Empowering Teachers to be the Change in the Classroom; European Commission (2017<sup>[41]</sup>), Networks for Learning and Development across School Education, Directorate-General Education, Youth, Sport and Culture, Schools and Multilingualism, Brussels.

Peer learning could be especially beneficial in Kazakhstan, since many schools might not regularly receive an external evaluation. The national appetite for peer learning is strong, but formal procedures are underdeveloped. Teachers communicate frequently on social media, but without government facilitation. In the past, there have been mechanisms to encourage peer learning, but they have been sporadic and not always available to all schools. For example, the methodological cabinet from the Nur-Sultan akimat created a programme that grouped schools into clusters of five schools for peer support purposes. Each cluster had two high-performing schools, one average performing school and two low performing schools. However, this programme was not scaled outside of the akimat. An online website, <http://smk.kz>, allows

teachers to communicate with each other, but is only available to teachers who participated in specific training modules. Kazakhstan has also resource centres, where students from small schools spend time in larger schools (OECD/The World Bank, 2015<sup>[5]</sup>). However, this initiative focuses on giving students access to greater resources, not on encouraging peer learning between schools.

### *Recommended actions*

The OECD recommends that Kazakhstan facilitate the development of peer learning across schools as part of the national school improvement programme. Kazakhstan should develop and pilot several peer learning strategies and select a few for nationwide scaling based upon the results. One such strategy is the formation of school clusters, as was done in Nur-Sultan but not scaled. If Kazakhstan decides to further develop and introduce School Review, external evaluation results could be used to create clusters composed of schools with different ratings, but similar characteristics (e.g., similar sizes and community demographics). These schools can then share practices with each other about how they addressed common challenges. Alternatively, school staff could visit other schools within the cluster to observe teaching and learning practices. Schools that already belong to resource clusters can begin learning from each other even before School Review would begin. In situations where schools are too distant to meet regularly, Kazakhstan can consider creating internet-based platforms to facilitate interaction, such as the National Improvement Hub in Scotland (Education Scotland, n.d.<sup>[42]</sup>).

Methodological cabinets in akimats can support peer learning by organising regular events and workshops that bring schools together and allow them to exchange ideas. Akimats are already familiar with convening schools because they do so to host competitions such as the International Olympiads. When schools are not able to attend the event, a videoconferencing system should be set up to allow virtual participation. While inadequate infrastructure will not allow all schools to remotely attend, videoconferencing will nevertheless help fill the gap.

Defining incentives to motivate schools to participate in peer learning is essential to ensuring that the activities are meaningful. Currently, schools in Kazakhstan are recognised for high performance, in particular their students' success in international competitions, which prevents most schools from being able to receive national recognition. The OECD recommends that Kazakhstan's approach to rewarding schools be broadened to include schools that have demonstrated improvement. For high-performing schools, what should further distinguish them is how they are able to share their expertise with other schools. This type of recognition scheme would incentivise excellent schools to engage in peer learning, and weaker schools to work with excellent performing schools to improve themselves. This type of incentive has shown to be effective in many education systems, such as in the state of Ceará in Brazil, where the high-performing schools receive additional funding to work with low performing schools to help them improve (Bruns, Evans and Luque, 2012<sup>[43]</sup>).

## **4.3. Support school leaders to lead school improvement**

### *Evidence*

The most vital ingredient to successful school-led improvement is the capacity of school leadership. It is the principal's responsibility to define school goals, observe instruction, support teachers' professional development and collaborate with teachers to improve instruction (Schleicher, 2015<sup>[44]</sup>). Even well-developed systems of school evaluation and peer learning will struggle to succeed if principals do not have the capacity to use those systems to improve their schools.

School leaders in Kazakhstan are selected to be administrators, not to be instructional leaders in their schools. There is only one master's programme in the country in education leadership, suggesting that most principals have not had formal preparation to assume their roles. Once in their posts, professional

development opportunities are provided in a non-systematic manner and are difficult to access for principals from rural schools. Those professional development courses are not always designed based on principals' development needs and can be disconnected from schools' daily practices (OECD/The World Bank, 2015<sup>[5]</sup>).

### *Recommended actions*

Although school leadership is not the focus of this policy perspective, the capacity of school leaders is integral to school improvement, especially in the context of Kazakhstan where it will take a long time for some schools to receive an external evaluation if School Review is introduced. Therefore, this specific recommendation makes several broad suggestions that about how school leaders can be developed in Kazakhstan to lead the changes in teaching and learning that the potential School Review envisions.

#### **Improve initial training for school leaders and select them based upon their capacity to be instructional leaders**

The availability of masters' programmes in education leadership should be expanded and school leaders' participation in those courses should be encouraged before they take up their duties. The education leadership programme at the Graduate School of Education in Nazarbayev University could act as a model to be emulated (with adaptations) in selected universities that have the capacity to offer this course. Nazarbayev University could also develop an online version of its master's programme in order to help future principals in remote areas to prepare for to take up their posts. Graduates from these programs should be given priority when applying to principal posts. Furthermore, the qualifications that new principals are expected to have should look beyond categorical criteria that might preclude talented graduates from these new programs, such as years of experience in the education system.

#### **Introduce systematic professional development for school leaders**

Principals need access to coaching and mentoring along with well-facilitated opportunities to visit and work in different schools (Greany, 2018<sup>[45]</sup>). One method of developing principals is through co-ordinating capacity building through a central agency. These agencies can have several responsibilities, from training current principals to helping form and drive policy concerning principals. Such agencies have been successfully formed in Singapore and Scotland (Scottish College for Educational Leadership, n.d.<sup>[46]</sup>; Singapore National Institute of Education, n.d.<sup>[47]</sup>).

In Kazakhstan, professional development for principals, like professional development for teachers, has been fragmented. With the potential introduction of School Review, the expectations of principals will be raised and they will need strong, systematic support to fulfil their responsibilities. Kazakhstan should consider approaching professional development for principals in a more central, deliberate manner to ensure that all principals, especially those in rural schools with fewer resources, can receive support.

The OECD suggests that one organisation be identified to lead principal professional development. Because Orleu and the NIS Centre of Pedagogical Excellence have already provided principal training, one of them could serve as this organisation. After being identified, the organisation will need to work with regional methodological cabinets so programmes that are centrally developed can be regionally delivered. Since principals in rural areas might have difficulty attending training, the identified organisation should consider developing online professional development courses that will be shared on their online platforms (both the NIS Centre of Excellence and Orleu already deliver some trainings online). It will also be important for the organisation to work closely with providers of principal training (currently only Nazarbayev University) so the education that aspiring principals' receive is co-ordinated with their in-service training.

### **Provide principals with the resources to carry out self-evaluation**

Since many schools are likely not to be externally evaluated regularly if School Review is introduced, the potential standards positively require self-evaluation so schools can engage in school-led improvement in between external visits. However, self-evaluation has not been emphasised since school attestation was eliminated, and even when it was regularly practiced it was not always a meaningful exercise. Principals will need support if they are to engage in effective self-evaluation.

If Kazakhstan decides to further develop and introduce School Review, it should give principals the School Review standards, the evaluation manual and the classroom observation protocol. Principals can use these materials to understand the new quality expectations for Kazakhstani schools and evaluate for themselves the quality of teaching and learning at their own schools. The policy perspective on the national assessment also recommends the development of diagnostic assessments, which principals can administer in their schools to understand how their students perform. These procedures will not only help principals perform more meaningful self-evaluation, but also help them prepare for external evaluation visits.



## References

- Alaska Department of Education and Early Development (2019), *2019 Educator Guide to Assessment Reports for PEAKS English Language Arts and Mathematics and Alaska Science Assessment*, [https://education.alaska.gov/tls/Assessments/Peaks/EducatorGuide\\_Assessments\\_Reports.pdf](https://education.alaska.gov/tls/Assessments/Peaks/EducatorGuide_Assessments_Reports.pdf) (accessed on 13 March 2020). [73]
- American Educational Research Association (2014), *Standards for Educational and Psychological Testing*. [87]
- Anderson, P. and G. Morgan (2008), *Developing Tests and Questionnaires for a National Assessment of Educational Achievement*, The World Bank, Washington, D.C. [69]
- AQA (n.d.), *How a question paper is created*, <https://www.aqa.org.uk/about-us/what-we-do/getting-the-right-result/how-exams-work/making-an-exam-a-guide-to-creating-a-question-paper/making-an-exam-a-guide-to-creating-a-question-paper-video-transcript> (accessed on 3 April 2020). [113]
- Asian Development Bank (2017), “Innovative Strategies for Accelerated Human Resource Development in South Asia: Student Assessment and Examination—Special Focus on Bangladesh, Nepal, and Sri Lanka”, <http://dx.doi.org/10.22617/TCS179079>. [105]
- Baartman, L. et al. (2007), “Evaluating assessment quality in competence-based education: A qualitative comparison of two frameworks”, *Educational Research Review*, Vol. 2/2, pp. 114-129, <http://dx.doi.org/10.1016/j.edurev.2007.06.001>. [82]
- Baird, J. et al. (2018), *Examination standards: How measures and meanings differ around the world*. [111]
- Bishop, J. (1999), *Are National Exit Examinations Important for Educational Efficiency?*, Cornell University ILR School, <http://digitalcommons.ilr.cornell.edu/articles> (accessed on 10 January 2019). [80]
- Boyle, A. et al. (2000), *State Support for School Improvement: School-level Perceptions of Quality Evaluating the Quality of State Support for School Improvement*, <http://www.air.org> (accessed on 15 April 2020). [33]
- Bray, M. and M. Kobakhidze (2014), “The Global Spread of Shadow Education”, [http://dx.doi.org/10.1007/978-94-6209-650-9\\_11](http://dx.doi.org/10.1007/978-94-6209-650-9_11). [93]
- Bray, M. and O. Kwo (2013), “Behind the façade of fee-free education: shadow education and its implications for social justice”, *Oxford Review of Education*, Vol. 39/4, pp. 480-497, <http://dx.doi.org/10.1080/03054985.2013.821852>. [92]
- Bridges, D. (2014), *Education Reform and Internationalisation: The Case of School Reform in Kazakhstan*, Cambridge University Press, Cambridge. [53]
- Bruns, B., D. Evans and J. Luque (2012), *Achieving World-Class Education in Brazil The Next Agenda Human Development*, World Bank, <https://openknowledge.worldbank.org/handle/10986/2383> (accessed on 20 April 2020). [43]
- Bürger, S., U. Kroehne and F. Goldhammer (2016), “The transition to computer-based testing in large-scale assessments: Investigating (partial) measurement invariance between modes”, *Psychological Test and Assessment Modeling*, Vol. 58/4, pp. 597-616, [64]



- <https://www.researchgate.net/publication/322330675>.
- Chamberlain, S. (2013), “Communication strategies for enhancing qualification users’ understanding of educational assessment: Recommendations from other public interest fields”, *Oxford Review of Education*, Vol. 39/1, pp. 114-127, <http://dx.doi.org/10.1080/03054985.2013.764757>. [109]
- Department Basic Education Republic of South Africa (2018), *Annual performance plan 2018/2019*, <https://www.education.gov.za/Portals/0/Documents/Reports/Annual%20Performance%20Plan%20201819.pdf?ver=2018-03-14-121624-263> (accessed on 3 April 2020). [107]
- Department of Education and Skills (2016), *Action Plan for Education 2016-2019*, Government of Ireland, Dublin, <https://www.education.ie/en/Publications/Corporate-Reports/Strategy-Statement/Department-of-Education-and-Skills-Strategy-Statement-2016-2019.pdf>. [52]
- DuFour, R. (2004), “What Is a Professional Learning Community?”, *Educational Leadership*, Vol. 61/8, pp. 6-11, <http://www.ascd.org/publications/educational-leadership/may04/vol61/num08/What-Is-a-Professional-Learning-Community%C2%A2.aspx> (accessed on 27 February 2020). [13]
- Education Review Office (2020), *Manurewa West School*, <https://www.ero.govt.nz/review-reports/manurewa-west-school-19-03-2020/> (accessed on 29 April 2020). [31]
- Education Review Office (2019), *Longford Intermediate School Evaluation*, <https://www.ero.govt.nz/review-reports/longford-intermediate-28-11-2019/> (accessed on 18 January 2020). [18]
- Education Review Office (2016), *School Evaluation Indicators*, <https://www.ero.govt.nz/assets/Uploads/ERO-15968-School-Evaluation-Indicators-2016-v10lowres.pdf>. [29]
- Education Review Office (2016), *School evaluation indicators: Effective Practice for Improvement and Learner Success*. [19]
- Education Scotland (2015), *How good is our school? (4th edition)*, <http://www.educationscotland.gov.uk/resources/h/hgios4/> (accessed on 27 January 2020). [20]
- Education Scotland (n.d.), *National Improvement Hub*, <https://education.gov.scot/improvement> (accessed on 20 April 2020). [42]
- Ehren, M. and N. Shackleton (2016), “Risk-based school inspections: impact of targeted inspection approaches on Dutch secondary schools”, *Educational Assessment, Evaluation and Accountability*, Vol. 28/4, pp. 299-321, <http://dx.doi.org/10.1007/s11092-016-9242-0>. [23]
- Enseignement Belgique (n.d.), *Non-certification external assessments [Évaluations externes non certificatives]*, The Teaching Portal in the Wallonia-Brussels Federation, <http://www.enseignement.be/index.php?page=25162&navi=2024> (accessed on 5 April 2020). [74]
- ERC (n.d.), *National Assessments of Mathematics and English Reading (NAMER)*, Educational Research Centre, Ireland, <http://www.erc.ie/studies/namer/overview/> (accessed on 20 February 2020). [65]
- ETS (2019), *Content and Construct Validity*, [https://www.ets.org/gre/institutions/about/fairness/content\\_construct\\_validity/](https://www.ets.org/gre/institutions/about/fairness/content_construct_validity/) (accessed on [88]

26 March 2020).

- European Commission (2017), *Networks for Learning and Development across School Education*, Directorate-General Education, Youth, Sport and Culture, Schools and Multilingualism, [https://www.schooleducationgateway.eu/downloads/Governance/2018-wgs5-networks-learning\\_en.pdf](https://www.schooleducationgateway.eu/downloads/Governance/2018-wgs5-networks-learning_en.pdf). [4  
1]
- European Commission/EACEA/Eurydice (n.d.), *Assuring Quality in Education: Policies and Approaches to School Evaluation in Europe*, Publications Office of the European Union., Luxembourg; <http://dx.doi.org/10.2797/678>. [7  
6]
- Faubert, V. (2009), "School Evaluation: Current Practices in OECD Countries and a Literature Review", *OECD Education Working Papers*, No. 42, OECD Publishing, Paris, <https://dx.doi.org/10.1787/218816547156>. [2  
7]
- Government Communication and Information Systems Republic of South Africa (2018), *Official guide to South African 2017/2018 Education*, <https://www.gcis.gov.za/sites/default/files/docs/resourcecentre/pocketguide/08-Education-1718.pdf> (accessed on 3 April 2020). [1  
0  
6]
- Government of Republic of Kazakhstan (2017), *About the Program – Digital Kazakhstan state program*, Government decree, <https://digitalkz.kz/en/about-the-program/> (accessed on 22 November 2019). [6  
3]
- Greaney, V. and T. Kellaghan (2012), *National Assessments of Educational Achievement, Volume 3: Implementing a National Assessment of Educational Achievement*, The World Bank, Washington DC. [5  
7]
- Greaney, V. and T. Kellaghan (2008), *Assessing National Achievement Levels in Education*, The World Bank, Washington, D.C. [6  
0]
- Greaney, V. and T. Kellaghan (2007), *National Assessments of Educational Achievement Volume 1*, The World Bank, <http://dx.doi.org/10.1596/978-0-8213-7258-6>. [1  
1  
5]
- Greany, T. (2018), "Balancing the needs of policy and practice, while remaining authentic: an analysis of leadership and governance in three national school leadership colleges", *Wales Journal of Education*, Vol. 20/2, pp. 63-98, <https://doi.org/10.16922/wje.20.2.5>. [4  
5]
- Hamilton, L. and D. Koretz (2002), "Tests and their use in test-based accountability systems", in L. Hamilton, B. (ed.), *Making Sense of Test-Based Accountability in Education*, RAND Publishing, , Santa Monica, California. [6  
8]
- Hargreaves, A. et al. (n.d.), *Leading from the Middle: Spreading Learning, Well-being, and Identity Across Ontario Council of Ontario Directors of Education Report*. [3  
6]
- Hayward, L. (2015), "Assessment is learning: the preposition vanishes", *Assessment in Education: Principles, Policy and Practice*, Vol. 22/1, pp. 27-43, <http://dx.doi.org/10.1080/0969594X.2014.984656>. [7  
1]
- Hipkins, R. and S. Robertson (2011), *Moderation and Teacher Learning What can research tell us about their interrelationships?*, <https://www.nzcer.org.nz/system/files/moderation-teacher-learning.pdf> (accessed on 1 April 2020). [1  
1  
4]
- HKEAA (2020), *Hong Kong Examinations and Assessment Authority - Governance*, [1  
1]

- [http://www.hkeaa.edu.hk/en/about\\_hkeaa/governance/](http://www.hkeaa.edu.hk/en/about_hkeaa/governance/) (accessed on 8 April 2020). 7]
- IAC (2019), *National Report on the State and Development of the Education System of the Republic of Kazakhstan 2018*, Information Analytic Centre, Nur-Sultan. [4 ]
- IAC (n.d.), *System for education data analysis (SEDA)*, Information Analytic Centre (IAC), <http://seda.iac.kz/about.php> (accessed on 13 December 2019). [7 8]
- INEE (n.d.), *National Institute of Education Evaluation [Instituto Nacional para la Evaluación de la Educación]*, Mexico, <https://www.mejoredu.gob.mx/sobre-el-inee/historia-inee/> (accessed on 20 February 2020). [5 8]
- INVALSI (2019), *National Institute for the Evaluation of the Education and Training System*, <https://www.invalsi.it/invalsi/index.php> (accessed on 25 February 2020). [1 1 8]
- Japan Times (2019), *Open-ended questions for Japan’s new university entrance exams scrapped*, [https://www.japantimes.co.jp/news/2019/12/17/national/open-ended-questions-scrapped-university-entry-exams-japan/#.Xim\\_QMhKiHs](https://www.japantimes.co.jp/news/2019/12/17/national/open-ended-questions-scrapped-university-entry-exams-japan/#.Xim_QMhKiHs) (accessed on 23 January 2020). [9 8]
- Jumabayeva, Z. (2016), “The Key Drivers of the Unified National Test in Kazakhstan: A Critical Analysis of its Impact on School Leavers”, *NUGSE Research in Education*, Vol. 1, [https://www.researchgate.net/publication/311665473\\_The\\_Key\\_Drivers\\_of\\_the\\_Unified\\_National\\_Test\\_in\\_Kazakhstan\\_A\\_Critical\\_Analysis\\_of\\_its\\_Impact\\_on\\_School\\_Leavers](https://www.researchgate.net/publication/311665473_The_Key_Drivers_of_the_Unified_National_Test_in_Kazakhstan_A_Critical_Analysis_of_its_Impact_on_School_Leavers) (accessed on 27 January 2020). [1 0 4]
- Kellaghan, T. and V. Greaney (2019), *Public Examinations Examined*, The World Bank, <http://dx.doi.org/10.1596/978-1-4648-1418-1>. [8 6]
- Kellaghan, T., V. Greaney and T. Murray (2009), *Using the Results of a National Assessment of Educational Achievement National Assessments of Educational Achievement*. [5 6]
- Kirkpatrick, R. (2011), *The Negative Influences of Exam-Oriented Education on Chinese High School Students: Backwash from Classroom to Child*, <https://link.springer.com/content/pdf/10.1186/2229-0443-1-3-36.pdf> (accessed on 30 March 2020). [9 1]
- Kitchen, H. et al. (2019), *OECD Reviews of Evaluation and Assessment in Education: Student Assessment in Turkey*, OECD Reviews of Evaluation and Assessment in Education, OECD Publishing, Paris, <https://dx.doi.org/10.1787/5edc0abe-en>. [8 1]
- Kitchen, H. et al. (2017), *Romania 2017*, OECD Reviews of Evaluation and Assessment in Education, OECD Publishing, Paris, <https://dx.doi.org/10.1787/9789264274051-en>. [1 7]
- Kitchen, H. et al. (2019), *OECD Reviews of Evaluation and Assessment in Education: North Macedonia*, OECD Publishing, Paris, <https://dx.doi.org/10.1787/079fe34c-en>. [9 5]
- Leiva, F., F. Ríos and T. Martínez (2006), “Assessment of interjudge reliability in the open-ended questions coding process”, *Quality and Quantity*, Vol. 40/4, pp. 519-537, <http://dx.doi.org/10.1007/s11135-005-1093-6>. [9 7]
- Li, R. et al. (2019), *OECD Reviews of Evaluation and Assessment in Education: Georgia*, OECD Reviews of Evaluation and Assessment in Education, OECD Publishing, Paris, <https://dx.doi.org/10.1787/94dc370e-en>. [2 4]

- Madaus, G. (1988), "The Distortion of Teaching and Testing: High-Stakes Testing and Instruction", [9  
0]  
*Peabody Journal of Education*, Vol. 65/3, pp. 29-46,  
<http://dx.doi.org/10.1080/01619568809538611>.
- Maghnouj, S. et al. (2019), *OECD Reviews of Evaluation and Assessment in Education: Serbia*, [1  
1]  
OECD Reviews of Evaluation and Assessment in Education, OECD Publishing, Paris,  
<https://dx.doi.org/10.1787/225350d9-en>.
- MCEECDYA (2009), *Principles and Protocols for Reporting on Schooling in Australia*, Ministerial [7  
2]  
Council for Education, Early Childhood Development and Youth Affairs (MCEECDYA),  
<http://www.educationcouncil.edu.au/site/DefaultSite/filesystem/documents/Reports%20and%20publications/Publications/Measuring%20and%20reporting%20student%20performance/Principles%20and%20protocols%20for%20reporting%20on%20schooling%20in%20Australia.pdf>.
- Mcnamara, G. and J. O'hara (2006), "Workable Compromise or Pointless Exercise? School-based [1  
5]  
Evaluation in the Irish Context", <http://dx.doi.org/10.1177/1741143206068218>.
- Meadows, M. and B. Black (2018), "Teachers' experience of and attitudes toward activities to [1  
1  
2]  
maximise qualification results in England", *Oxford Review of Education*, Vol. 44/5, pp. 563-580,  
<http://dx.doi.org/10.1080/03054985.2018.1500355>.
- Michels, B. and H. Eijkelhof (2019), "The Development and Implementation of an Interdisciplinary [1  
1  
6]  
STEM Course for Upper Secondary Schools", *Educational Designer*, Vol. 2/3,  
<https://www.educationaldesigner.org/ed/volume3/issue12/article48/> (accessed on 6 May 2020).
- MINEDUC (n.d.), "Sistema de Medición de Calidad de la Educación (SIMCE)", *Agencia de Calidad [7  
9]  
de la Educación, Ministry of Education, Chile*,  
<https://www.agenciaeducacion.cl/evaluaciones/que-es-el-simce/>.
- Minister of Education and Science of the Republic of Kazakhstan (2016), *On the approval of [1  
0]  
evaluation criteria for educational institutions. Order of the Minister of Education and Science of  
the Republic of Kazakhstan dated February 2, 2016 No. 124*.
- MoES (2019), *Country Background Report: Kazakhstan*, Ministry of Education and Science, Nur- [8  
1]  
Sultan.
- MoES (2016), "Rules for conducting an external assessment of educational achievements", Ministry [7  
5]  
of Education and Science, Nur-Sultan, <http://adilet.zan.kz/rus/docs/V1600013287>.
- Mourshed, M., C. Chijioke and M. Barber (2010), *Education ow the world's most improved school [3  
7]  
systems keep getting better*, McKinsey & Company.
- Mullis, I. et al. (2016), *TIMSS 2015 International Results in Mathematics and Science*, Retrieved [6  
7]  
from Boston College, TIMSS & PIRLS International Study Center,  
<http://timss2015.org/download-center/> (accessed on 6 April 2020).
- Munby, S. and M. Fullan (2016), *Inside-out and downside-up*. [3  
5]
- New Zealand Qualifications Authority (n.d.), *External Advisory Groups*, [1  
0  
0]  
<https://www.nzqa.govt.nz/about-us/our-structure/advisory-groups/> (accessed on 6 April 2020).
- Newton, P. (2007), "Clarifying the purposes of educational assessment", *Assessment in Education: [5  
9]  
Principles, Policy & Practice*, Vol. 14/2, pp. 149-170,  
<http://dx.doi.org/10.1080/09695940701478321>.

- Norwegian Directorate for Education and Training (n.d.), *School gate - Primary school - Overview*, <https://skoleporten.udir.no/> (accessed on 1 May 2020). [7  
7]
- OECD (2019), *PISA 2018 Database*, <https://www.oecd.org/pisa/data/2018database/>. [7  
]
- OECD (2019), *PISA 2018 Results (Volume I): What Students Know and Can Do*, PISA, OECD Publishing, Paris, <https://dx.doi.org/10.1787/5f07c754-en>. [3  
]
- OECD (2019), “PISA 2018 Background questionnaires”, in *PISA 2018 Assessment and Analytical Framework*, OECD Publishing, Paris, <https://dx.doi.org/10.1787/67e1518f-en>. [7  
0]
- OECD (2019), *TALIS 2018 Results (Volume I): Teachers and School Leaders as Lifelong Learners*, TALIS, OECD Publishing, Paris, <https://dx.doi.org/10.1787/1d0bc92a-en>. [1  
2]
- OECD (2018), *Developing Schools as Learning Organisations in Wales*, Implementing Education Policies, OECD Publishing, Paris, <https://dx.doi.org/10.1787/9789264307193-en>. [3  
8]
- OECD (2018), “Education Policy Outlook Kazakhstan”, <https://www.oecd.org/education/Education-Policy-Outlook-Country-Profile-Kazakhstan-2018.pdf>. [8  
3]
- OECD (2017), *Education in Lithuania*, Reviews of National Policies for Education, OECD Publishing, Paris, <https://dx.doi.org/10.1787/9789264281486-en>. [8  
5]
- OECD (2017), *Higher Education in Kazakhstan 2017*, Reviews of National Policies for Education, OECD Publishing, Paris, <https://dx.doi.org/10.1787/9789264268531-en>. [8  
4]
- OECD (2016), *Education in Colombia*. [3  
4]
- OECD (2015), *Education at a Glance 2015: OECD Indicators*, OECD Publishing, Paris, <http://dx.doi.org/10.1787/eag-2015-en>. [4  
9]
- OECD (2015), *Education at a Glance 2015: OECD Indicators*, OECD Publishing, Paris, <https://dx.doi.org/10.1787/eag-2015-en>. [9  
]
- OECD (2014), *Reviews of National Policies for Education: Secondary Education in Kazakhstan*, Reviews of National Policies for Education, OECD Publishing, Paris, <https://dx.doi.org/10.1787/9789264205208-en>. [6  
2]
- OECD (2013), *Synergies for Better Learning: An International Perspective on Evaluation and Assessment*, OECD Reviews of Evaluation and Assessment in Education, OECD Publishing, Paris, <https://dx.doi.org/10.1787/9789264190658-en>. [6  
]
- OECD (2009), *Measuring Government Activity*, OECD Publishing, Paris, <https://doi.org/10.1787/9789264060784-en>. [5  
1]
- OECD (2008), *Measuring improvements in Learning Outcomes: Best Practices to Assess the Value-Added of Schools*, <http://www.sourceoecd.org/educationt/9789264050228www.sourceoecd.org/9789264050228> (accessed on 25 February 2020). [2  
8]
- OECD (2007), *Reviews of National Policies for Education: Higher Education in Kazakhstan 2007*, Reviews of National Policies for Education, OECD Publishing, Paris, <https://dx.doi.org/10.1787/9789264033177-en>. [1  
0  
3]
- OECD/The World Bank (2015), *OECD Reviews of School Resources: Kazakhstan 2015*, OECD [5



- Reviews of School Resources, OECD Publishing, Paris, [1]  
<https://dx.doi.org/10.1787/9789264245891-en>.
- OECD/The World Bank (2007), *Reviews of National Policies for Education: Higher Education in Kazakhstan 2007*, Reviews of National Policies for Education, OECD Publishing, Paris, [1  
0  
2]  
<https://dx.doi.org/10.1787/9789264033177-en>.
- Ofsted (2019), *School inspection handbook*, [3  
0]  
[https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment\\_data/file/843108/School\\_inspection\\_handbook\\_-\\_section\\_5.pdf](https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/843108/School_inspection_handbook_-_section_5.pdf) (accessed on 8 April 2020).
- Ofsted (2018), *Ofsted inspections: myths*, <https://www.gov.uk/government/publications/school-inspection-handbook-from-september-2015/ofsted-inspections-mythbusting> (accessed on [1  
6]  
19 February 2020).
- Opposs, D. et al. (2020), “Governance structure and standard setting in educational assessment”, [9  
9]  
*Assessment in Education: Principles, Policy & Practice*, pp. 1-23,  
<http://dx.doi.org/10.1080/0969594X.2020.1730766>.
- Paulo Santiago, A. et al. (2017), *OECD Reviews of School Resources: Chile*, OECD Publishing, [6  
1]  
Paris.
- Pearson (2018), *PARCC Final Technical Report for 2018 Administration*, [1  
0  
8]  
<https://files.eric.ed.gov/fulltext/ED599198.pdf> (accessed on 31 March 2020).
- Pont, B., D. Nusche and H. Moorman (2008), *Improving School Leadership - Volume 1: Policy and Practice*, OECD Publishing, Paris, [3  
9]  
<https://www.oecd.org/education/school/improvingschoolleadership-volume1policyandpracticevolume2casestudiesonsystemleadership.htm> (accessed on 20 January 2020).
- Republic of Kazakhstan (2019), *State Program for the Development of Education 2020-2025*. [3  
2]
- Republic of Kazakhstan (2016), “State Program for the Development of Education 2016-19”, 205. [1  
]
- Rostiashvili, K. (2004), *Corruption in the Higher Education System of Georgia*, [8  
9]  
[https://tracc.gmu.edu/sites/default/files/Rostiashvili\\_Corruption\\_in\\_HES\\_Eng.pdf](https://tracc.gmu.edu/sites/default/files/Rostiashvili_Corruption_in_HES_Eng.pdf) (accessed on 30 March 2020).
- Schick, A. (2003), “The Performing State: Reflection on an Idea Whose Time Has Come but Whose Implementation Has Not”, *OECD Journal on Budgeting*, Vol. 3/2, [4  
8]  
<http://dx.doi.org/10.1787/budget-v3-art10-en>.
- Schleicher, A. (2015), “Schools for 21st-Century Learners: Strong Leaders, Confident Teachers, Innovative Approaches. International Summit on the Teaching Profession.”, *OECD Publishing*. [4  
4]
- Scottish College for Educational Leadership (n.d.), *Education Scotland and SCEL*, [4  
6]  
<https://professionallearning.education.gov.scot/> (accessed on 28 February 2020).
- Scottish Government (n.d.), *Scottish National Standardised Assessments (SNSA)*, Smarter Scotland, Scottish Government, <https://standardisedassessment.gov.scot/> (accessed on [6  
6]  
20 February 2020).
- SICI (2018), *Country Profile The Netherlands*. [2  
2]

- SICI (2009), *The Inspectorate of England*, <https://www.sici-inspectorates.eu/getattachment/a8732d7f-cb1e-489a-b501-f5a85ad66743/87094.pdf;.jpg;.aspx>. [2  
1]
- Simpson, L. and J. Baird (2013), “Perceptions of trust in public examinations”, *Oxford Review of Education*, Vol. 39/1, pp. 17-35, <http://dx.doi.org/10.1080/03054985.2012.760264>. [1  
1  
0]
- Singapore National Institute of Education (n.d.), *Leaders in Education Programme*, <https://www.nie.edu.sg/professional-and-leadership-development/leadership-programmes/leaders-in-education-programme> (accessed on 28 February 2020). [4  
7]
- Tanujaya, B., J. Mumu and G. Margono (2017), “The Relationship between Higher Order Thinking Skills and Academic Performance of Student in Mathematics Instruction, International Education Studies, 2017”, *International Education Studies*, Vol. 10/11, <https://eric.ed.gov/?id=EJ1159551> (accessed on 25 February 2020). [9  
6]
- Tashibaeva, D. (n.d.), *What is necessary to do so that the school principal becomes a strong link in the education of Kazakhstan*, IAC, <http://iac.kz/en/publishing/what-necessary-do-so-school-principal-becomes-strong-link-education-kazakhstan-0> (accessed on 6 February 2020). [1  
4]
- Tong, C., C. Lee and G. Luo (2020), “Assessment reform in Hong Kong: developing the HKDSE to align with the new academic structure”, *Assessment in Education: Principles, Policy & Practice*, pp. 1-17, <http://dx.doi.org/10.1080/0969594X.2020.1732866>. [1  
0  
1]
- U.S. Department of Education (2018), *Alaska Consolidated State Plan*. [2  
6]
- U.S. Department of Education (2018), *Elementary and Secondary Education Act of 1965 as amended through P.L. 115-224, Enacted July 31, 2018*. [2  
5]
- UNICEF (n.d.), *Dare to Share: Empowering Teachers to be the Change in the Classroom*, UNICEF. [4  
0]
- World Bank (2018), *Population ages 0-14 (% of total population) - Kazakhstan, OECD members | Data*, <https://data.worldbank.org/indicator/SP.POP.0014.TO.ZS?locations=KZ-OE> (accessed on 3 March 2020). [2  
1]
- World Bank (2017), *Kazakhstan - Education Modernization Project*, <http://documents.worldbank.org/curated/en/340301488682834835/pdf/Kazakhstan-PAD-02092017.pdf> (accessed on 8 August 2019). [5  
5]
- World Bank (2012), *Kazakhstan: Student Assessment*, SABER Country Report, Systems Approach for Better Education Results. [5  
4]
- World Bank (2004), *Ten Steps Ten Steps to a Results-Based Monitoring and Evaluation System to a Results-Based Monitoring and Evaluation System*, <https://www.oecd.org/dac/peerreviews/World%20bank%202004%2010%20Steps%20to%20a%20Results%20Based%20ME%20System.pdf> (accessed on 4 December 2019). [5  
0]
- Zhang, Y. (2013), “Does private tutoring improve students’ National College Entrance Exam performance?-A case study from Jinan, China”, *Economics of Education Review*, Vol. 32/1, pp. 1-28, <http://dx.doi.org/10.1016/j.econedurev.2012.09.008>. [9  
4]



## Annex A. Key indicators

#	List of key indicators	Kazakhstan	OECD
Background information			
Economy			
1	GDP per capita, PPP (constant 2011 international \$), 2018*	27 738	40 537
2	GDP per capita growth (annual %), 2018*	4.1	2.3
Society			
3	Population growth (annual %), 2018*	1.3	0.6
4	Population aged 14 years or less (%), 2018*	28.5	17.8
Education indicators			
System			
5	Starting age of compulsory education, 2018***	7	5.7
6	Duration of compulsory education (years), 2017***	9	10.9
Students – net enrolment rates			
7	Pre-primary education (ISCED 0), 2017***	54.9	84.4
	Primary education (ISCED 1), 2018***	87.6	95.6
	Secondary education (ISCED 2 and 3), 2018***	89.4	89.4
8	Tertiary education attainment rate (25 to 34 years old) (ISCED levels 5 to 8), 2015***	50.3	40.9
9	Share of students enrolled in vocational programmes for upper secondary education (15 to 19 year olds), 2017***	39.7	43.1
Teachers			
10	Mean age of teachers (TALIS 2018)	40.9	44.1
11	Share of female teachers in secondary education	75.5	58.6
12	Ratio of students to teaching staff (2018) Primary education (ISCED 1)***	19.6	15.3
13	Ratio of students to teaching staff (2018) Secondary education (ISCED 2 and 3)***	7.0	13.7
Finance			
14	Total expenditure on education as a percentage of GDP, all levels 2016***	3.0	5.4
15	Total public expenditure on primary education as a percentage of total government expenditure, 2017 for Kazakhstan, 2016 for OECD average***	1.0	3.5
Learning outcomes			
16	Mean students' performance in reading, PISA 2018****	387	487
17	Mean students' performance in science, PISA 2018****	397	489
18	Mean students' performance in mathematics, PISA 2018****	423	489

Source: \* The World Bank, World Bank Indicators: Education, <https://data.worldbank.org/topic/education> (accessed on 17 January 2020)

\*\* UIS, UNESCO Institute for Statistics, <http://data.uis.unesco.org> (accessed on 17 January 2020)

\*\*\* OECD (2019), TALIS 2018 Results (Volume I): Teachers and School Leaders as Lifelong Learners, TALIS, , <https://doi.org/10.1787/1d0bc92a-en>

\*\*\*\* OECD (2019), PISA 2018 Results (Volume I): What Students Know and Can Do, <https://doi.org/10.1787/5f07c754-en>. \*

\*\*\*\* OECD (2019), Skills Matter: Additional Results from the Survey of Adult Skills, OECD Skills Studies, , <https://doi.org/10.1787/1f029d8f-en>

This work is published under the responsibility of the Secretary-General of the OECD. The opinions expressed and arguments employed herein do not necessarily reflect the official views of OECD member countries.

This document, as well as any data 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.

The statistical data for Israel are supplied by and are under the responsibility of the relevant Israeli authorities. The use of such data by the OECD is without prejudice to the status of the Golan Heights, East Jerusalem and Israeli settlements in the West Bank under the terms of international law.

The use of this work, whether digital or print, is governed by the Terms and Conditions to be found at <http://www.oecd.org/termsandconditions>.