6 Monitoring and evaluating equity and inclusion

This chapter analyses the fundamental role that monitoring and evaluation can play in ensuring the continuous improvement and the effectiveness of policies targeted at addressing diversity in education and improving equity and inclusion in education systems. Specifically, this chapter focuses on monitoring equity and inclusion in education, evaluating policies and practices to address diversity, and promoting equity and inclusion at the system, local and school levels. The chapter concludes by proposing policy pointers for approaches to monitoring and evaluation that promote equity and inclusion in education.

Introduction

This chapter discusses the fundamental role that monitoring and evaluation can play in ensuring that an education system is not only introducing policies to improve equity and inclusion, but is also implementing them and achieving its objectives.

Without relevant information on the current state of equity and inclusion and progress towards these, policy makers might judge the system according to the imperfect data they have available. This might misdirect them or, in the case of absence of data, may mean that they are unaware of challenges that need action. Monitoring systems are therefore important to assess progress in improving equity and inclusion in education. They are crucial in providing feedback to inform improvements across the education system, as well as in identifying necessary support measures for schools.

Evaluation is important in determining whether policies are having the intended effects and in informing necessary adjustments. The evaluation process can help policy makers decide, for instance, whether policies are having inadvertent effects and should be discontinued, or whether they warrant further support. In the case of smaller programmes, this can entail their upscaling or institutionalisation. While evaluation can pose challenges, it is particularly important to seek co-operation and consensus across all stakeholder levels in evaluating and subsequently implementing changes that lead to equity and inclusion.

Programmes that support equity and inclusion in education are also designed and implemented at the local and school levels. Hence, the monitoring of progress towards more equitable and inclusive education systems is important at these levels, too. Individual actors need to assess their circumstances and thus receive feedback on their performance in the areas of equity and inclusion. This might also help them identify how to improve their interventions to support equity and inclusion in individual schools. External, as well as internal, school evaluation can be highly informative in improving equity and inclusion in education.

This chapter is organised in five sections. After the introduction, the second section elaborates on monitoring progress in improving equity and inclusion. The third section examines evaluations of policies, programmes and processes to improve equity and inclusion. The fourth section explores supports for schools in improving equity and inclusion practices through evaluation processes. The final section provides pointers for future policy development.

Monitoring progress in improving equity and inclusion in education

Education systems differ in how they monitor progress in improving equity and inclusion in education. The following sections summarise what is monitored, what instruments are used to measure progress and the use of monitoring results.

Operationalisation of equity and inclusion in monitoring systems

Monitoring and evaluation frameworks can be used for accountability and improvement purposes (OECD, 2013[1]). A major accountability objective is to inform the public of the quality of the education system, including the quality of education for diverse groups. Another objective is to provide feedback on reforms in the education system that can be used to improve educational processes and outcomes of all students. In general, six major aims can be distinguished (OECD, 2013[1]):

- Monitoring of student academic and broader well-being outcomes, including the disaggregation by dimensions of diversity, socio-economic background and geographic location;
- Monitoring of student outcomes over time;
- Monitoring of the impacts of a policy;

- Monitoring demographic, administrative and contextual data which can explain the outcomes of the education system;
- Generating feedback and information for stakeholders in the education system; and
- Using the generated information for development and implementation of policies.

These aims can be subsequently tailored for the purposes of monitoring progress in improving equity and inclusion in education. The extent to which these aims are present in a given education system varies. Equity and inclusion are intrinsically linked to the context of each education system – so their monitoring may differ considerably from country to country. Indeed, in academic literature, international definitions or national practices, there is no consensus on either the definitions of equitable and inclusive education systems, or the difference between the two. For the purposes of this report, readers are invited to refer to the explanations provided in Chapter 1.

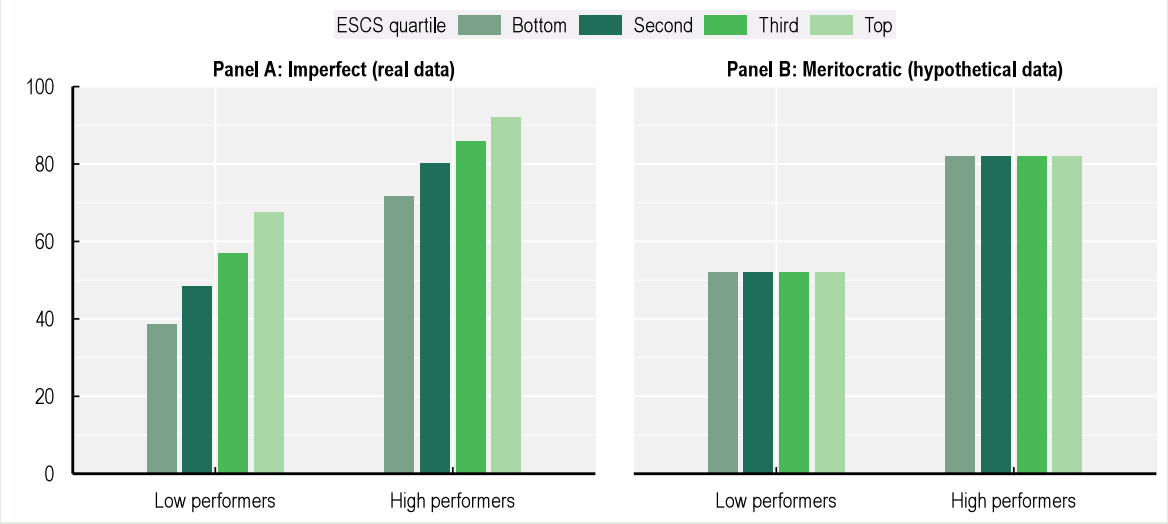
Equitable education systems are those that ensure the achievement of educational potential regardless of personal and social circumstances, including factors such as gender, ethnic origin, Indigenous background, immigrant status, sexual orientation, gender identity, special education needs (SEN) and giftedness (Cerna et al., 2021[2]; OECD, 2017[3]). A closely related term is that of "equal educational opportunities". In a system that offers equal educational opportunities, educational outcomes are the result of actions within the individual student's control and their ability to reach their full potential is not hindered by circumstances beyond their control. Under this notion, educational outcomes should be a result of actions in individuals' control and not of circumstances beyond their control so that they can reach their full potential. One of the roles of monitoring progress in improving equity in education is to look at the disparities in educational outcomes as well as (in)equalities in opportunities. In order to simplify and classify the concept of equity in education, the UNESCO Institute for Statistics (2018[4]) published the *Handbook on Measuring Equity in Education* that outlines a classification of equity in education into five concepts: meritocracy, minimum standards, impartiality, equality condition and redistribution (Box 6.1).

Box 6.1. Concepts of equity in education

Meritocracy

Under this concept, educational outcomes are redistributed based on merit. For instance, the OECD reports the percentage of students who expect to complete tertiary education by the level of their reading performance and socio-economic status (Panel A in Figure 6.1). In 2018, on average across OECD and even within the same reading proficiency level (merit), the expectations depended heavily on the socio-economic background of students. Under a meritocratic distribution, approximately equal levels of expectations within proficiency levels would be expected (Panel B in Figure 6.1). One of the biggest challenges when evaluating equity based on the concept of meritocracy is to find a suitable measurement of merit. Indeed, reading performance might not be the best merit based on which to evaluate equity in expectations to complete tertiary education.

Figure 6.1. Percentage of students who expect to complete tertiary education



Note: ESCS quartile relates to OECD Programme for International Student Assessment (PISA) index of economic, social and cultural status (ESCS) (OECD, 2019[5]).
Source: OECD (2019[5]), PISA 2018 Results (Volume II): Where All Students Can Succeed, Table II.B1.6.6 and Table II.B1.6.7, 10.1787/b5fd1b8f-en (Panel A).

StatLink https://stat.link/huze8y

Minimum standards

The measure of minimum standards focuses on whether some minimum educational outputs are achieved by everyone. For instance, Target 4.1 of the Sustainable Development Goals aims to ensure that all girls and boys complete primary and secondary education by 2030 (United Nations, n.d.[6]). The minimum standard is, in this case, full primary and secondary education completion.

Impartiality

Impartiality quantifies the relationship between an outcome variable and a measure of circumstance. The measure of circumstance can be gender, socio-economic background, immigrant status, etc. For example, the OECD regularly reports PISA reading, mathematics and science scores by gender, socio-economic and immigrant background (OECD, 2019[5]). Education at a Glance 2021, which specifically focused on equity in education, reported a wide range of outcomes related to participation and progression through education disaggregated by gender, socio-economic status, immigrant background (country of origin) and geographic location (sub-national regions) (OECD, 2021[7]). There is no single way of quantifying the relationship between an outcome variable and a measure of circumstance. Thus, the practice varies from simple disaggregation of student outcomes by groups to more complex estimations of relationships (correlation coefficients or proportion of variance explained by circumstances).

Equality condition

While impartiality focuses on absolute levels of (in)equity, the equality condition is concerned with distribution of educational variables (often resources) across the population. For instance, the OECD reports how various resources hinder instruction (based on school leaders' reports) across performance or socio-economic groups (OECD, 2020[8]).

Redistribution

Redistribution concerns whether educational inputs are distributed equally or, for instance, unequally to compensate for existing disadvantages. Redistribution is often applied in education finance. For example, in some countries, such as in England (United Kingdom), pay scales for teachers in high-income areas are higher to compensate for the fact that living expenses there are higher (Department for Education, 2022[9]).

Source: UNESCO-UIS (2018[4]), Handbook on Measuring Equity in Education, http://uis.unesco.org/sites/default/files/documents/handbook-measuring-equity-education-2018-en.pdf (accessed 10 July 2022).

In practice, however, the perspectives taken towards the operationalisation of equity are often narrow and simplified to a disaggregation of student outcomes or an observation of inequalities in students' performance (e.g., through the standard deviations of students' scores, percentile deviations of students' scores) (Appels et al., 2022[10]). Studies rarely explain what equity involves beyond the disaggregation of student performance, such as the role of schools in counterbalancing inequities, or the interactions between student performance, background characteristics, teachers and the broader environment (ibid.).

The importance of monitoring progress in improving equity in education is closely related to the rationale for equitable education in general. As mentioned in Chapter 1, disparities in learning outcomes are related to a range of negative outcomes later in life (UNESCO-UIS, 2018[4]). Furthermore, high inequities in education resulting from a lack opportunities can lead to the misallocation of skills and talent, and thus hinder economic growth (Hsieh et al., 2013[11]; OECD, 2018[12]).

While there is no agreed distinction between the monitoring of progress in improving equity and in improving inclusion, it is possible to differentiate the two by focusing on the conceptualisation in Chapter 1. Based on these definitions, the monitoring of progress in improving inclusion should be broader in focus than for equity. In addition to examining whether all students have equal opportunities to reach their potential, the monitoring of progress in improving inclusion should also focus on how students feel at school, their well-being outcomes and socio-emotional development (Mezzanotte and Calvel, Forthcoming[13]). Inclusion indicators can also examine whether students are truly included in the school setting (e.g., their sense of belonging) or just integrated; and explore the potential barriers students may face with regard to inclusion. One approach to monitoring inclusion is to focus on processes in the education system, in addition to inputs and outcomes (Box 6.2).

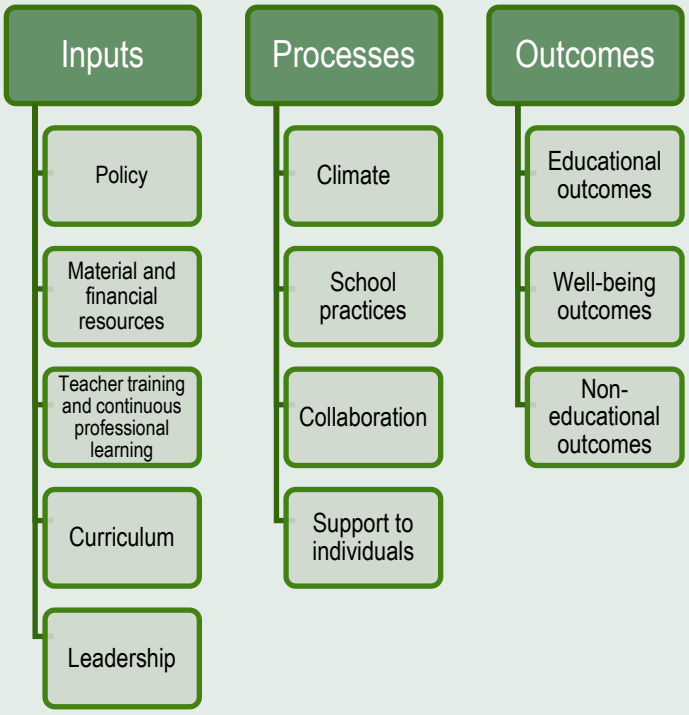
Box 6.2. Inputs-processes-outcomes model

The inputs-processes-outcomes model combines three dimensions to create a comprehensive framework for identifying areas requiring more intervention. Inputs include all sources provided to the system. These include not only financial resources but also policies, teacher training, curriculum and leadership. Processes are practices in schools such as the development of school climate, collaboration or support to individuals. Processes ultimately transform inputs into outcomes. Outcomes include educational outcomes (e.g., participation, dropout, grade repetition rates and achievement), well-being outcomes (such as sense of belonging, mental health and school climate) and non-educational outcomes (for example economic and labour market outcomes, and health outcomes) (Figure 6.2). By focusing on inputs and processes in addition to outcomes, the model can shed more light on the potential causes for regressed outputs.

The inputs-processes-outcomes model allows for the monitoring of progress in relation to both equity and inclusion. By focusing on outcomes, it can measure, for instance, how educational well-being or non-educational factors differ by student groups. The process dimension of the model goes beyond

educational outcomes and examines factors such as school climate, collaboration and the support individuals – and can thus, to some extent, show whether the system has the capacity to adjust to the needs of the student groups under analysis. The inputs part of the model can shed light on both equity and inclusion: it can show, for instance, the extent to which financial resources are distributed equally to compensate for existing disadvantages but also, for example, whether curricula are truly inclusive of all students. The model has been adopted in education, such as in Education at a Glance 2018 (OECD, 2018[14]).

Figure 6.2. Inputs-processes-outcomes model



Source: Adapted from Mezzanotte and Calvel (Forthcoming[13]), Indicators of inclusion in education: a framework for analysis.

The monitoring of progress in improving inclusion also involves a greater emphasis on individual experiences rather than those of groups or categories (UNESCO, 2020[15]). Indeed, the more inclusive a school is, the less it needs to use categorical data given that fewer students require identification or support (ibid.). School-level approaches towards monitoring the progress in improving inclusion might thus be more appropriate. These are summarised in the section on Supporting schools in improving equity and inclusion practices through evaluation processes.

Data collection practices are diverse

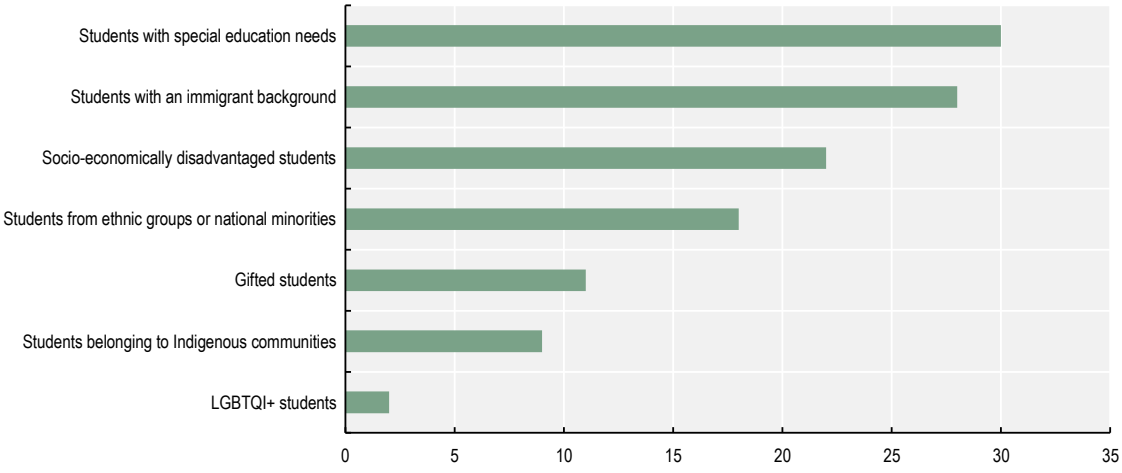
The evaluation of the progress towards reaching inclusion and equity goals cannot happen without robust data collections that monitor the access, participation and achievement of all learners. This can include monitoring across specific groups (by gender, immigrant background, SEN, socio-economic or ethnic/Indigenous background, giftedness, and sexual orientation and gender identity) as well as various student outcomes.

The Strength through Diversity Policy Survey 2022[1] indicates that a range of student groups are included in the national data collections of education systems across the OECD. Thirty education systems reported

collecting data on students with SEN, 28 systems on students with an immigrant background, 22 on socio-economically disadvantaged students, 18 on students from certain ethnic groups or national minorities, 11 on gifted students and nine on students belonging to Indigenous communities. Only Canada and Chile collected data on LGBTQI+ students (Figure 6.3).[2]

Figure 6.3. Data collections on diversity (2022)

Number of education systems that collect data on the following groups (ISCED 2)



Note: This figure is based on answers to the question "Does a national (or sub-national) authority collect data on these groups of students at ISCED 2 level?". Thirty-one education systems responded to this question. Response options were not mutually exclusive.
Options selected have been ranked in descending order of the number of education systems.
Source: OECD (2022[16]), Strength through Diversity Policy Survey 2022.

StatLink https://stat.link/4le65c

Some dimensions of diversity (namely, giftedness, sexual orientation and gender identity) are underrepresented in data collections, as is acknowledged in international research (McBrien, Rutigliano and Sticca, 2022[17]; Rutigliano and Quarshie, 2021[18]). There are a range of reasons why data for particular dimensions of diversity may not be collected at the national (or sub-national) level. Legislative frameworks in some countries may not allow for the collection of some characteristics (e.g., sexual orientation) due to the private and sensitive nature of such data. Some education systems, such as Portugal (Box 6.3), do not categorise students based on their characteristics but instead focus on the support measures they require. Other education systems adopt colour-blind policies whereby data on certain characteristics, such as ethnic background, are prohibited to be collected by law (see the next section).

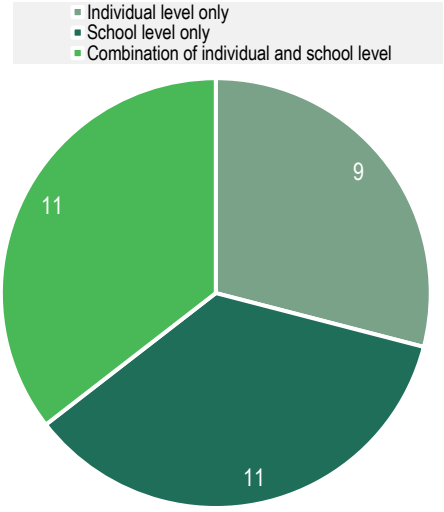
The methods for collecting data also differ across education systems. Some national (or sub-national) authorities collect data at the individual student level while others at an aggregated (e.g., school) level. Individual-level data collections mean that data is collected about each student and then sent to a national (or sub-national) authority. Aggregated level data means that data is sent to a national (or sub-national) authority in aggregates. For instance, each school sends the total number of students with an immigrant background, rather than data on each student's immigration status (which would be collected in an individual-level approach).

Results from the Strength through Diversity Policy Survey 2022 show that aggregated collections were more common as a means for gathering data compared to individual-level data collections (Figure 6.4).

Data was reported as being collected solely at an aggregate level in 11 education systems and solely at the individual level in nine systems. In 11 systems, the level of data collection differed depending on the student characteristic.

Figure 6.4. Type of data collection (2022)

Number of education systems that collect data about students at the following levels (ISCED 2)



Note: This figure is based on answers to the question "Does a national (or sub-national) authority collect data on these groups of students at ISCED 2 level? If so, are data points collected on an individual student level or aggregated on e.g., school level?". Thirty-one education systems responded to this question.
Source: OECD (2022[16]), Strength through Diversity Policy Survey 2022.

StatLink https://stat.link/9mgvjt

Individual-level data collections enable policy makers to consider intersectionality in their analyses and policy implications. Aggregated data are only as applicable for an intersectional analysis as they are set up during the collection. For example, a country with aggregated data collections would only be able to know the number of female students with an immigrant background if it specifically asked for that particular number and dimensional intersection from each school. On the other hand, by collecting individual-level data with attributes on both gender and an immigrant background, it is straightforward to create aggregate statistics for the intersection.

Practices around labelling diverse students differ

Labelling students with a particular need, ethnicity or other type of background can have both positive and negative impacts. Labelling is viewed as advantageous by some teachers: classification can help to identify and explain the limitations and potential negative consequences of current practice. For example, some teachers feel that student Attention Deficit Hyperactivity Disorder (ADHD) classifications can help explain why regular practice does not suffice and legitimates a different approach (Mezzanotte, 2020[19]; Wienen et al., 2019[20]). Classifications can also bring consistency to research and communication, and can be useful in the assessment and placement of students in special programmes or settings (OECD, 2022[21]; Thomson, 2012[22]). For some typologies of SEN, the label can also bring some “relief” to students or parents, who can then better understand the driver of certain difficulties or struggles at school. Diagnosis or clarification can thus explain the causes of student behaviour, bring empathy and offer resolution (Mezzanotte, 2020[19]; Wienen et al., 2019[20]).

The absence of student labels can also make some students invisible to policy makers and can silence the experiences of diverse groups (Öhberg and Medeiros, 2017[23]; Simon, 2017[24]). After all, if data is not collected, no gaps can be seen. This may cause stakeholders in the education system to remain or become ignorant of the needs of some students. In the Netherlands, students with SEN are not labelled as such and individual schools are meant to monitor the progress of students with a "progress and development plan" (Inspectorate of Education, 2022[25]). However, while individual schools might have a good overview of these students, national data are of poor quality (ibid.). An accurate picture of the trends in the number of students with SEN (with a progress and development plan) and their outcomes in and through education is therefore limited (ibid.). In the Slovak Republic, it was estimated that due to limits in data collections on "disadvantaged socio-economic background", only 39% of students at risk of poverty or social exclusion have been targeted by financial contributions for disadvantaged students (Hellebrandt et al., 2020[26]). In the area of intersectionality, research points out that the often reported gender gap in learning outcomes varies significantly by socio-economic status or the ethnicity of students (OECD, 2019[5]; Strand, 2014[27]). However, these kinds of intersections of students' identities and heterogeneities are often absent in monitoring systems and thus are not considered in policy responses (Varsik and Gorochovskij, Forthcoming[28]).

Opponents to labelling argue that equality and national cohesion of ethnic groups and national minorities is achieved through invisibility, i.e., all are equal before the law without a distinction between ethnic groups (Balestra and Fleischer, 2018[29]; Simon, 2017[24]). Some European countries have on this basis adopted a "colour-blind" approach to data collection on ethnic groups. Under this policy, data on ethnic groups are not collected or, if they are collected, they are not considered in policy making.

Some researchers have also argued that labelling may result in teachers having lower expectations of the performance of certain students due to their preconceptions regarding the abilities of students belonging to diverse groups (Hart, Drummond and McIntyre, 2007[30]). Labelling may in this way result in students not being viewed as individuals but judged based on stereotypical preconceptions (OECD, 2022[21]; Osterholm, Nash and Kritsonis, 2007[31]). According to some teachers, labelling also has no value for educational practice without a further analysis of the student needs (Mezzanotte, 2020[19]; Wienen et al., 2019[20]). Furthermore, by labelling students, teachers and other stakeholders in the system might focus on deficits of students, rather than directing their attention as to how the system can help underperforming students (Ainscow and Messiou, 2017[32]). Finally, in the context of inclusion of students, inclusion cannot happen one group at a time; the process must encompass all students (UNESCO, 2020[15]). From this standpoint, data disaggregated by characteristics might seem irrelevant. As a result, some education systems are changing their approaches in order to limit the potentially negative consequences of labelling (Ebersold et al., 2020[33]). In Portugal, for example, students are no longer categorised by their characteristics (ethnic or immigrant background, SEN etc.), but by the type of educational support measure(s) they need (Box 6.3).

Box 6.3. Reform and monitoring of inclusive education in Portugal

As part of this legislative changes since 2018, Portugal shifted its emphasis from identification of student characteristics to identification of student support measure(s). Portugal no longer collects information on student characteristics (except for gender and nationality) on the rationale that it is not necessary to categorise in order to intervene. The identification of students' needs happens at the school level and is conducted as early as possible in co-operation with a range of stakeholders including parents/guardians, social services or relevant teaching as well as non-teaching staff. The initial identification is followed by an approval for assessment and mobilisation of a multidisciplinary team. The team is also responsible for implementation and monitoring of the support measures.

As a result, the system collects information on students falling into one of three categories of support

measures: universal measures, selective measures and additional measures. A student in each category can benefit from a wide spectrum of interventions ranging from curriculum accommodation/enrichment, tutoring, pedagogical-psychological support to redesigning of the pedagogical strategy, including significant curricular adjustments.

Portugal has co-operated with the European Agency for Special Needs and Inclusive Education to develop a monitoring system that will enable stakeholders to assess the effectiveness of the inclusion law. This new monitoring system was developed in 2022 and consists of six standards, 11 indicators and 19 related questions that monitor the level of implementation of the identified standards (European Agency for Special Needs and Inclusive Education, 2022[34]).

Source: OECD (2022[21]), Review of Inclusive Education in Portugal, https://doi.org/10.1787/a9c95902-en.

Finally, identifying groups at an aggregate level might be required for other purposes than targeted support, such as maintaining statistical databases or allocating resources. As elaborated in Chapter 3, almost all education systems that provided answers to the Strength through Diversity Policy Survey 2022 reported accounting for various student backgrounds in their funding formulas. These labels may not impact how teachers view their students. In fact, assigning a label based on administrative data can occur without assigning the same label in the classroom (UNESCO, 2020[15]). In Finland, for instance, diagnostic labels are not used in the classroom, because the system focuses on the needs of individual students, regardless of their characteristics. For administrative purposes, however, certain statistics can be disaggregated by student groups, such as those of students with physical impairments (Jahnukainen and Itkonen, 2010[35]).

Monitoring mostly focuses on academic outcomes

The lives and experiences of students are shaped by a range of factors. Apart from learning, students also spend a considerable time at school socialising with their peers and interacting with school staff. Academic outcomes are only one aspect of the overall school experience, and it is important to understand how happy and satisfied students are with different aspects of their life, how connected they are to others and whether they enjoy good physical and mental health (Cerna et al., 2021[2]). This understanding can be developed through collecting data on a range of student well-being outcomes, including academic, psychological, physical, social and material (ibid.). These dimensions are key ingredients of the concurrent well-being of individuals and contribute to their personal development in the short-, medium- and long-term (ibid.).

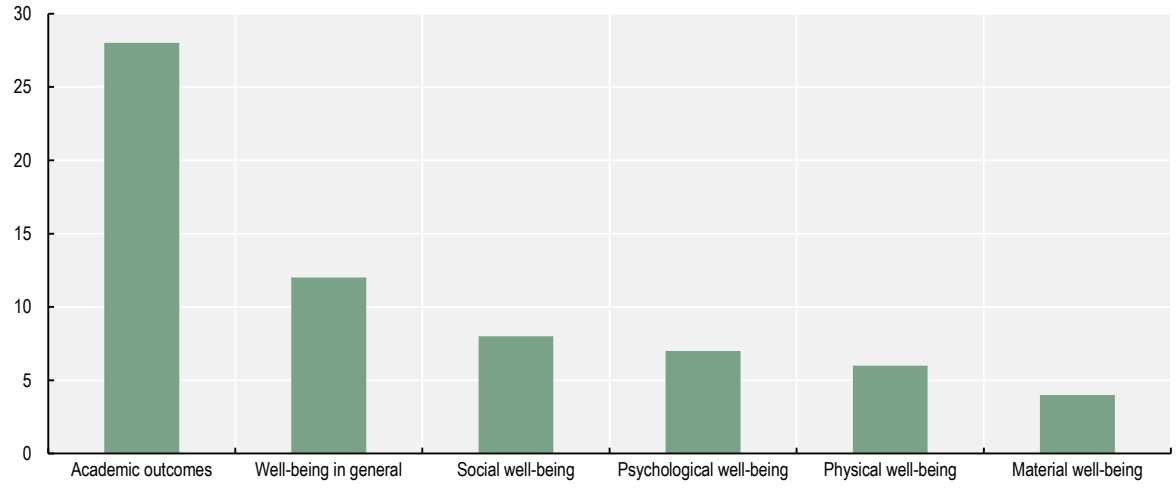
The Strength through Diversity Project considers multiple aspects of student well-being: academic, psychological, physical, social and material (Cerna et al., 2021[2]). The psychological dimension of students' well-being includes students' views about life, their engagement with school, the extent to which they have a sense of agency, identity and empowerment, and their opportunities to develop goals and ambitions for their future. Physical well-being relates to students' health status, safety and security, the ability to engage with others without physical barriers in access and mobility. Social well-being refers to the quality of students' social lives. This includes relationships with peers, family and school staff. Finally, material well-being considers the material resources available that enable families and schools to cater to students' needs.

However, despite growing research on the positive associations between high levels of student well-being and positive and fulfilling life-experiences (Pollard and Lee, 2003[36]), performance (Gutman and Vorhaus, 2012[37]) and negative associations with risky behaviours (such as drinking and smoking) (Currie et al., 2012[38]), schooling is in many education systems organised with the aim of maximising learning outcomes, sometimes at the expense of overall well-being. This is reflected in data collections, which tend to focus on learning outcomes, with comparatively little attention given to indicators of other aspects of well-being.

Most education systems (28) who participated in the Strength through Diversity Policy Survey 2022 reported focusing their official data collections on academic outcomes for students irrespective of specific groups (Figure 6.5). While 12 systems collected data on general well-being outcomes, collecting data on specific types of well-being was less common for students irrespective of specific groups. Eight education systems considered social well-being, seven systems psychological well-being and six physical well-being in their data collections. Material well-being outcomes were rarely included in national or sub-national data collections (with only four education systems reporting that they collected data on this aspect of well-being).

Figure 6.5. Data collections on students irrespective of specific groups (2022)

Number of education systems that collect data at least once during ISCED 2 on students irrespective of specific groups



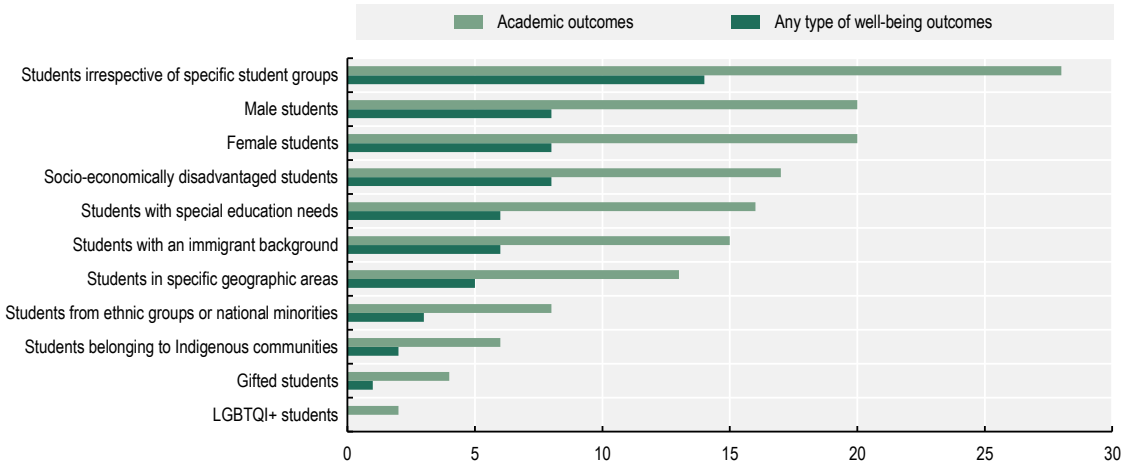
Note: This figure is based on answers to the question "Which dimensions of student outcomes are nationally (or sub-nationally) collected at least once during ISCED 2 level? [Students irrespective of specific student groups]". Thirty-one education systems responded to this question. Response options were not mutually exclusive.
Options selected have been ranked in descending order of the number of education systems.
Source: OECD (2022[16]), Strength through Diversity Policy Survey 2022.

StatLink https://stat.link/d9ptlb

Education systems most commonly reported collecting data on the academic and well-being outcomes of students irrespective of specific groups (Figure 6.6). The dimensions of diversity that were most common in group-focused data collections were gender, socio-economic and immigrant background, SEN, and location in specific geographic areas. Only eight and three education systems collected academic and well-being data respectively for students from ethnic groups or national minorities. Six and two systems collected academic and well-being data respectively on students belonging to Indigenous communities and data collections on gifted and LGBTQI+ students were even rarer. For all student groups, data on any type of well-being outcomes were collected considerably less often than data on academic outcomes.

Figure 6.6. Data collections on academic and well-being outcomes (2022)

Number of education systems that collect data at least once during ISCED 2 on the following groups



Note: This figure is based on answers to the question "Which dimensions of student outcomes are nationally (or sub-nationally) collected at least once during ISCED 2 level?". Thirty-one education systems responded to this question. Response options were not mutually exclusive. Any type of well-being outcomes can include one or more of the following: psychological well-being outcomes, social well-being outcomes, material well-being outcomes, physical well-being outcomes, well-being outcomes in general.
Options selected have been ranked in descending order of the number of education systems that selected any type of well-being outcomes.
Source: OECD (2022[16]), Strength through Diversity Policy Survey 2022.

StatLink https://stat.link/zuyf6h

This trend is also visible in international research. In many education systems, student experiences at school remain unmonitored. For instance, systematic data collections on anti-LGBTQI+ bullying in schools are in place in only four countries (Finland, France, the Netherlands and Sweden), despite the fact that LGBTQI+ students are consistently reporting higher rates of bullying compared to their peers (IGLYO, 2022[39]; McBrien, Rutigliano and Sticca, 2022[17]). Furthermore, those collections often differ in the extent of detail they cover. In Sweden, for instance, the school inspectorate monitors bias-motivated bullying in schools that can be based on the sexual orientation or gender identity of the victim (IGLYO, 2022[39]). In other countries, however, the data is less detailed and more generic (McBrien, Rutigliano and Sticca, 2022[17]).

Instruments used to measure progress towards equity and inclusion

Data can be used by policy makers and other stakeholders to monitor progress, evaluate outcomes and ultimately improve students' learning and other well-being outcomes. One common approach to summarising data collections comes in the form of indicator frameworks. Other instruments include the use of national assessments with background data, national or international longitudinal and key stakeholder surveys, and specific reviews on equity and inclusion.

Data collections happen within an educational context under different policy regimes. As a result, the methods for collecting data differ across education systems and no country uses a single instrument to measure progress towards equity and inclusion. Most often the approaches are combined. For instance, results from national assessments and surveys are often used in indicators frameworks and specific reviews on equity and inclusion.

Indicator frameworks

The joint international standardised data collection by UNESCO, OECD and EUROSTAT has been a major driver of the collection of international information on equity in education. The data are summarised in indicators and published in various reports, including Education at a Glance, which focuses on equity and inclusion in education every three years (most recently in 2021) (OECD, 2021[7]). The OECD has published a number of reports that focus on equity in and beyond education, although these mainly focus on disparities in terms of socio-economic background, gender, immigrant background and geographic location (e.g., in terms of urban/rural differences) (OECD, 2012[40]; OECD, 2017[3]; OECD, 2018[14]). Performance scores are often disaggregated by one or more dimensions of diversity. The effect of socio-economic background on performance is often examined: as score-point difference in performance associated with one-unit increase in the index of economic, social and cultural status, or percentage of the variation in performance explained by the index (OECD, 2019[41]). Beyond education, indicators focus on labour market outcomes, such as earnings or labour market participation by educational attainment (OECD, 2018[14]). Data on adult skills or educational attainment by socio-economic background are also often reported (OECD, 2017[3]). More recently, the OECD engaged in developing a dashboard of indicators on equity in and through education (Box 6.4).

A key challenge is designing indicators in a way that adequately represents the value they are measuring. While national education goals and objectives may be comprehensive and broad, monitoring systems may be rather limited in the information they can offer. For instance, as was elaborated in the previous sections, measures of inclusion need to collect a wide range of data outside of the domain of learning outcomes. These may not be available, challenging and costly to obtain if new data sources and data infrastructures need to be created. In some areas, such as quality of the teaching force, data might not even be possible to obtain. In this context, it is difficult to create indicator frameworks that are feasible, reliable, with high coverage and validity (OECD, 2013[1]). It is therefore important to consider the purpose of each indicator so that it designed in a considered manner with the goal to measure what is valued, rather than value what is measured (Ainscow, 2005[42]). That way, policy makers and other stakeholders can limit the possibility that the measure itself becomes the target.

Box 6.4. Dashboard of indicators on equity in and through education

In 2022, the OECD engaged in developing a dashboard, whose objective was to position OECD countries across a range of indicators on equity in and through education. The dashboard has two overarching aims and five policy aims. The overarching aims are: (1) enabling all learners to developing the knowledge, skills, attitudes and values to thrive in equitable and inclusive societies; and (2) ensuring that education contributes to equitable economic and social outcomes. The five policy aims include: (a) raising educational outcomes through more equitable education opportunities; (b) investing in the early years; (c) empowering teachers and school leaders to support equity in and through education; (d) aligning resources with the needs of learners; and (e) enabling an inclusive school environment. As a result, equity in and through education is measured using 35 key comparative indicators often disaggregated by age, gender and socio-economic status.

In the process of the dashboard development, the OECD also identified several gaps and limitations in the existing data. These included a lack of information on the early years of education, particularly in terms of country coverage, teacher quality and teaching practices. Data on the quality of initial education and continuous professional learning for teachers and school leaders was also generally hard to find. Challenges were also visible in measuring equity throughout the student's education, especially when focusing on enrolment in, access to and graduation from upper secondary and tertiary education by students' characteristics.

Some important areas could not be covered due to lack of information or data based on subjective students', teachers' or school leaders' reports. These areas include attitudes and values of students and young adults, skills to thrive in a digital world and on socio-emotional skills, social outcome indicators (civic engagement, for instance), and engagement with parents and communities. Other highlighted challenges included the timeliness of data (e.g., the Survey of Adult Skills, a product of the Programme for the International Assessment of Adult Competencies (PIAAC), is relatively old), lack of data focusing on inclusion as a process, and insufficient data on ethnic groups, refugees, LGBTQI+ students, gifted students and students with SEN.

Source: OECD (2022[43]), Dashboard of indicators on equity in and through education, OECD document for official use, EDU/EDPC(2022)20/ANN.

Reference points for measuring progress in advancing equity and inclusion

At the national (or sub-national) level, many education systems embed indicators into strategies, action plans or national improvement frameworks that set out goals for equity and inclusion. This can be done to monitor progress, and clarify the vision and objectives of the administration, while reducing and aggregating the abundance of available information to several key elements (Gouëdard, 2021[44]). Furthermore, by including an equity component in the strategy (e.g., disaggregation of outcomes by dimensions of diversity), the success of the policy can also be measured in relation to specific student groups.

New Zealand's Child and Youth Well-being Strategy, for instance, measures progress towards six defined well-being outcomes (and equity in relation to those outcomes) (Department of the Prime Minister and Cabinet, 2020[45]). In education, these cover participation, attendance, literacy, numeracy, science skills, socio-emotional skills and self-management skills. In Japan, the Third Basic Plan for the Promotion of Education (2018-22) focuses on well-being outcomes (OECD, 2019[46]). These include improvement in "the percentage of students in [primary and secondary] schools who do not eat breakfast" or improvement in "the percentage of students in [primary and secondary] schools who go to bed at around the same time every day and who wake up at around the same time every day" (Ministry of Education, Culture, Sports, Science and Technology, 2018[47]). In many cases, the strategies also include more specific instruments that measure the progress towards equity and inclusion.

Goals related to equity and inclusion are also central to the Scottish (United Kingdom) National Improvement Framework, with the specified key priorities including closing the attainment gap between advantaged and disadvantaged students, improving students' health and well-being, and placing the human rights and needs of every child and young person at the centre of education. These key priorities reflect the vision for education specified in the Framework, which is to deliver both excellence and equity, ensuring that every child and young person is able to thrive and has "the best opportunity to succeed, regardless of their social circumstances or additional needs" (Scottish Government, 2021[48]). Similarly, in Latvia, ensuring equal opportunities is one of the strategic objectives of Latvia's National Development Plan for 2021-2027 (Cross-Sectoral Coordination Center, 2020[49]).

National curricula can serve as reference points, given that countries often use national assessments to monitor the implementation of the national curriculum and/or progress against specific student learning objectives or educational standards (OECD, 2013[1]). National curricula can also articulate equity and inclusion as overarching goals for education to ensure that the value of equity is both taught and modelled in schools (OECD, 2021[50]). In Estonia, for instance, the values of "honesty, compassion, respect of life, justice and human dignity" are embedded in the national curriculum across various themes (OECD, 2021, p. 95[50]).

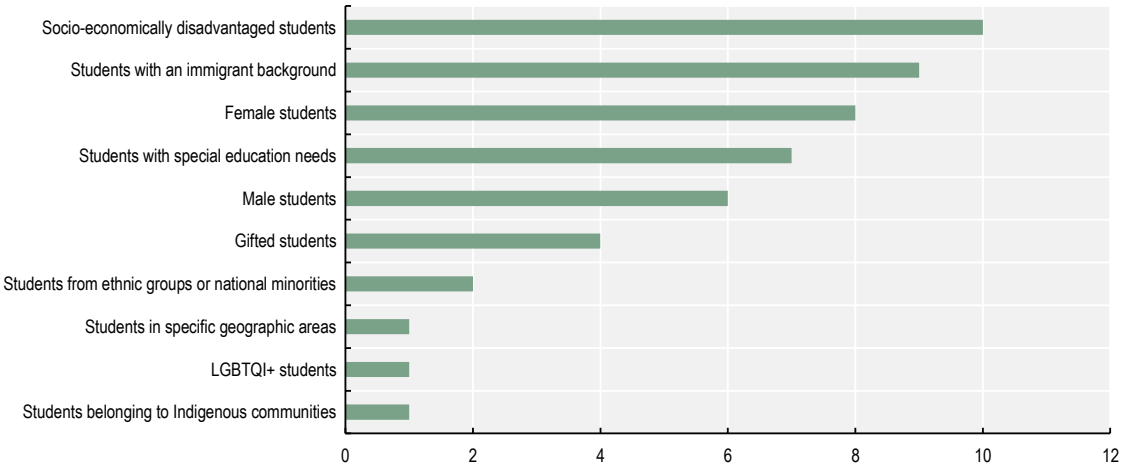
Education systems define a wide range of indicators

Indicators often include targets that can help instigate action. Targets can help policy makers identify the biggest gaps, quantify how much action is needed in different priority areas and monitor progress over time. If a strategy contains a specific target value for a policy or a set of policies, policy makers and other stakeholders can use it to evaluate whether the goals were met.

The results of the Strength through Diversity Policy Survey 2022 show that most OECD education systems did not define indicators specific to student groups (Figure 6.7). The indicators that were reported focused on equity and mostly measured impartiality (Box 6.1), i.e., whether educational outcomes differ for various student groups. Most education systems defined indicators with a focus on socio-economically disadvantaged students (ten systems), with an immigrant background (nine systems), female or male students (eight and six systems respectively), and students with SEN (seven systems). Four education systems reported having indicators specific to gifted students, and two systems specific to ethnic groups or national minorities. The other dimensions of diversity were represented even less often.

Figure 6.7. Indicators (2022)

Number of education systems that define indicators specific to the following groups (ISCED 2)



Note: This figure is based on answers to the question "Does your education jurisdiction define nationally (or sub-nationally) indicator(s) specific to any of the groups of students at ISCED 2 level?". Twenty-five education systems responded to this question. Response options were not mutually exclusive.
Options selected have been ranked in descending order of the number of education systems.
Source: OECD (2022[16]), Strength through Diversity Policy Survey 2022.

StatLink https://stat.link/c2qm9j

Where education systems had indicators related to students with an immigrant background, these mostly examined their school participation rates and results (OECD, 2022[16]). Iceland and Spain reported monitoring immigrants' results based on international large-scale surveys, such as the PISA, with Iceland having set a target for students with an immigrant background to reach the OECD PISA average in reading, mathematics and science (OECD, 2022[16]). Targets in Latvia focused on full participation of minors who were granted asylum status and children of returning migrants, and in Lithuania on full participation of citizens of foreign countries.

Education systems reported monitoring students with SEN and gifted students largely in terms of whether they were labelled as such, which educational setting they were placed in (e.g., dedicated or mainstream schools) and whether they had received specific educational support (OECD, 2022[16]). In Spain, for instance, the indicators focus on the number of students with SEN and gifted students in education disaggregated by the typology of SEN, gender and the type of school. In Lithuania, targets are set for the proportion of students whose needs for additional support are met (85% in 2025 and 97% in 2030) as well as proportion of students with disabilities receiving inclusive education in mainstream education settings (85% in 2025 and 90% in 2030).

Socio-economically disadvantaged students were also explicitly targeted with indicators in some instances (OECD, 2022[16]). In England (United Kingdom), the education system monitors the disadvantaged gap index. It summarises the relative attainment gap between disadvantaged students and all other students. In Scotland (United Kingdom), achievements in the expected level in literacy and numeracy are reported by socio-economic background of students. Moreover, broader well-being outcomes are monitored, particularly as a potential consequence of the COVID-19 pandemic: the gap in total difficulties score between students (aged 13 and 15) in the most deprived and least deprived areas.[3]

Other education systems reported monitoring disadvantaged students indirectly (OECD, 2022[16]). One of the indicators in the Flemish Community of Belgium focuses on early leavers from education and training developed by EUROSTAT.[4] In Latvia, the provision of portable computer equipment for disadvantaged students is monitored as part the 2021-3 Action Plan of the Education Development Guidelines 2021-7.

Indicators targeting gender gaps often focus on academic outcomes (OECD, 2022[16]). In Iceland, the percentage of female and male students reaching level 2 or higher in PISA in reading, mathematics and science are monitored, with targets being set for the year 2025 for each subject and gender. In Northern Ireland (United Kingdom), mid-upper secondary examination results are reported as being disaggregated by gender. In Estonia, drop-out rates from lower secondary education are monitored by gender, while the participation of girls in science, technology, engineering and mathematics (STEM) programmes are monitored in the Flemish Community of Belgium and Ireland.

Two education systems (Chile and the Slovak Republic) reported having indicators related to students from ethnic groups and national minorities and one education system had indicators related to students belonging to Indigenous communities. In the Slovak Republic, the “Roma Strategy of Equity, Inclusion and Participation” sets the target of decreasing the number of Roma early school leavers by half (to 36%) and decreasing the number of Roma students attending ethnically homogenous classes by half (to 30%) by 2030 (OECD, 2022[16]).

Finally, some education systems reported having developed indicator frameworks related to Indigenous peoples. Such indicators are, for instance, set out in Australia’s “Closing the Gap” framework, which was developed to close the gap between Indigenous peoples and the majority population in areas such as child mortality, early childhood education, school attendance, literacy and numeracy, upper secondary education attainment, employment, and life expectancy (Commonwealth of Australia, 2020[51]). For each of these areas, there is a clearly stated target provided with a timeframe (e.g., “95 per cent of all Indigenous four-year-olds enrolled in early childhood education (by 2025)”) and progress with relevant statistics documented. Data is also provided at the provincial level.

National assessments with background data

In 2015, 27 and 30 out of 38 surveyed OECD systems administered national assessments or examinations during primary and upper secondary levels (general programmes), respectively (OECD, 2015[52]). National assessments and examinations are standardised student achievement tests. While national assessments do not affect students’ progression through school or certification, national examinations have a formal consequence for students, such as an impact on a student’s eligibility to progress to a higher level of education or to complete an officially-recognised degree (ibid.). The uses as well as methods for

administering the national assessments and/or examinations varied across education systems. The main purpose of national assessments at lower secondary level was to provide teachers with student diagnostic information and the main purpose of national examinations at the upper secondary level was to determine student entry to tertiary education (ibid.). National assessments/examinations also varied greatly in terms of the subjects covered and whether they were administered to all students or just a sample (ibid.).

While sample-based surveys generally suffer from shrinking samples and larger estimation errors as the focus shifts to individuals with multiple specific characteristics (UNESCO, 2020[15]), national assessments/examinations are often administered to the full cohort of students, giving way for analyses that explore the outcomes of groups of students that have very low populations. However, various factors have an impact on student results and countries exercise caution when publishing disaggregated outputs. Analyses with an intersectional focus that consider multiple dimensions of diversity are thus rarely applied. When conducted, these revealed, for instance, that SEN identification is heterogeneous across various dimensions of diversity. Disadvantaged students aged 5-16 were approximately twice as likely to be identified with SEN in England (United Kingdom), controlling for gender and ethnic background (Strand and Lindorff, 2018[53]; Strand and Lindsay, 2008[54]). Boys were more likely to be identified with SEN compared to girls, controlling for ethnic and socio-economic background (Strand and Lindorff, 2018[53]; Strand and Lindsay, 2008[54]).

Comprehensive information on the extent to which assessments include background data of students is not available. The results from the Strength through Diversity Policy Survey 2022 on data collections on diversity (Figure 6.3) indicated that some dimensions of diversity, such as sexual orientation or gender identity, were collected to a much lower extent (if at all) than others. Indeed, gender and geographic locations are two characteristics that are often used to present results from student assessments. For instance, Italy and Poland publish results from their student assessments disaggregated by gender and geographic location (Central Examination Commission, 2019[55]; INVALSI, 2022[56]). Some education systems publish information from student assessments by a range of other categories. In England (United Kingdom) results of student assessments at the upper secondary level are published by SEN, ethnicity and socio-economic background (determined on the basis of free school meals eligibility) in addition to gender and geographic location (Department for Education, 2022[57]). Estonia publishes results of the state examination by students' language of instruction, including sign language, gender and geographic location (Examination Information System, 2022[58]). The Slovak Republic publishes selected results from the lower secondary student assessment disaggregated by gender, socio-economic background, type of school and geographic location (NUCEM, 2022[59]).

Some students, such as students who arrived in the country during the school year in which they would normally be tested or less than one year beforehand, or students with SEN (particularly if they are enrolled in dedicated schools), might be excluded from participation in nation-wide assessments (European Education and Culture Executive Agency, Eurydice, 2011[60]). Exclusion of students from assessments may result in limited data on the progress of these student groups at the national (or sub-national) level. Sometimes, education systems make accommodations to the tests so that participation increases. For instance, several education systems provide tests in Braille or enlarged letters for students with visual impairments, or adapted material for pupils with physical disabilities (ibid.). For those with more severe typologies of SEN, students can also have an assistant, interpreter or special teacher at disposal during testing (European Education and Culture Executive Agency, Eurydice, 2011[60]; Hellebrandt et al., 2020[26]). Other accommodations include more time to take the test, more frequent breaks during testing, and the use of various forms of support (including magnifying glasses, special reminders or information leaflets, etc.) (European Education and Culture Executive Agency, Eurydice, 2011[60]). Finally, computer-based assessments can, in some instances, improve the accessibility of the test for some students with SEN, given that computerised accessibility tools can be incorporated into the assessment platform, as was recently done in the Pan-Canadian Assessment Programme (Council of Ministers of Education Canada, 2019[61]).

Publication of national assessment results: impacts on equity

National assessments and examinations are sometimes also used with the additional purpose of holding schools accountable. A common approach in this regard is to make aggregated school results publicly available. The provision of information could, in theory, achieve two outcomes: (1) lower the information asymmetry about overall student performance between school insiders and outsiders (e.g., parents, national (or sub-national) authorities); and (2) provide incentives to align the actions taken by school leaders with those set by governing bodies or national (or sub-national) authorities by focusing on pre-defined standards, subjects or contents that are measured (Torres, 2021[62]). While the impact of publication of school results has been shown to be positive on the average academic achievement, the evidence is mixed in regards to the impacts on equity in education (Hanushek and Raymond, 2005[63]; Torres, 2021[62]). On the one hand, school accountability systems were correlated with lower overall results for low-performing students in secondary education and increased gaps between majority and minority ethnic groups in England (United Kingdom) and the United States (Burgess et al., 2005[64]; Hanushek and Raymond, 2005[63]). On the other hand, student assessments can have positive consequences for equity in education if they are used to improve awareness of the main challenges in the education system of for particular students (OECD, 2020[8]). Positive correlations were observed between equity in education and the use of student assessments to inform parents about their child's progress; the use of student assessments to identify aspects of instruction or the curriculum that could be improved; and the use of written specifications for student performance on the school's initiative. Higher equity in education was also observed among countries that used assessments to seek feedback from students and to have regular consultations on school improvement (ibid.). Finally, studies showed no substantial impact of test-based accountability practices on educational inequalities (Torres, 2021[62]). For instance, the publishing of school results in Japan did not reveal adverse distributional effects on student performance (Morozumi and Tanaka, 2020[65]).

Depending on how they are designed, the publishing of school results can have unintended consequences with negative implications for equity and inclusion in education (Torres, 2021[62]). These can include increased social segregation among schools (due to parents choosing to enrol their children in better-performing schools) (Davis, Bhatt and Schwarz, 2015[66]), focusing teachers' attention only on the measured subjects, or only on the students who have a realistic chance of passing a given proficiency threshold (Neal and Schanzenbach, 2010[67]). They may also result in teachers focusing on preparing students to "sit the test" rather than the entire curriculum (Jennings and Rentner, 2006[68]). Creating league tables can also cause difficulties in recruiting and retaining teachers in low-performing schools (Clotfelter et al., 2004[69]).

Previous studies have suggested that accountability should not only apply to learning outcomes, but also to school resources, professional capacity and other school processes (Darling-Hammond and Snyder, 2015[70]). The rationale is that schools' outcomes need to be compared only if other aspects of the schooling process are considered, including resourcing and the (quality of) the teaching force (Torres, 2021[62]).

Longitudinal surveys

Longitudinal surveys use administrative data or specialised sample-based surveys to track the same people over time. The biggest advantage of longitudinal population surveys is that they can better determine patterns over time because changes can be observed for the same individuals. In contrast, cross-sectional surveys only provide snapshots of a given situation in time.

Education authorities in several countries co-operate with academic institutions or non-governmental organisations to administer longitudinal surveys. For example, the Millennium Cohort Study, Next Steps and Our Future longitudinal surveys have been conducted by higher education institutions and social research organisations with funding and in co-operation with several governmental departments in the United Kingdom (Centre for Longitudinal Studies, n.d.[71]; Centre for Longitudinal Studies, 2022[72];

CLOSER, n.d.[73]). The surveys offer a wealth of data that can be used for various group comparisons and intersectional analyses related to education and other areas (e.g., health, labour market outcomes). Most of these sources offer information on gender, SEN, ethnic and socio-economic background and some also on immigrant background and sexual orientation (ibid.). Similarly, the Longitudinal Study of Australian Children and the High School Longitudinal Study in the United States contain information on all dimensions of diversity except giftedness (Australian Institute of Family Studies, n.d.[74]; US Bureau of Labor Statistics, n.d.[75]).

Longitudinal surveys can also be created from administrative datasets that contain a unique student identifier that does not change over time. Since 1996, New Zealand assigns a unique student identifier (the National Student Number, NSN) to each student (NZQA, 2022[76]). This allows for monitoring of student enrolment, attendance and tracking of students' educational paths. It is also helpful for various statistical purposes, research purposes and ensuring that student educational records are accurately maintained. The availability of NSNs was more recently used to calculate the Equity Index, consisting of 37 variables that measure socio-economic background and educational achievement in upper secondary education examinations (Ministry of Education, 2022[77]). The 37 variables include parental socio-economic indicators (e.g., education level, income, prison custody, mother's age at her first child), child socio-economic indicators related to poverty, abuse or neglect (e.g., care and protection placement/notification/ investigation of child), national background variables (e.g., country of birth) and transience variables related to moving home or school (e.g., number of home and school changes) (ibid.). Similarly, in the Slovak Republic, the use of unique student identifiers has enabled the monitoring of Roma students' participation in education (from the primary to tertiary levels), and allowed for the analysis of early school leaving rates disaggregated by Roma/non-Roma as well as socio-economically advantaged/disadvantaged populations (Hellebrandt et al., 2020[26]). The use of unique student identifiers has also enabled researchers to compare the educational outcomes of both Roma students and students from socio-economically disadvantaged groups (excluding Roma) with those of the general population and to analyse issues such as within- and between-school segregation and the unequal distribution of SEN identification (ibid.). In Chile, the Provisional School Identifier allows for monitoring and continuity in the educational trajectories of students including those of foreign nationality, and also enables the certification of their studies (Ministry of Education, 2022[78]).

Key stakeholder surveys

Information regarding students' psychological, physical, social and material well-being outcomes are sometimes collected using surveys based on a sample. Such key stakeholder surveys can be used to broaden the evidence on student outcomes to domains that are outside of academic outcomes. Representative sample-based surveys can provide an overall picture of well-being outcomes, while limiting the administrative costs of running a full-population study. The disadvantage is, however, that analyses of student sub-groups is limited by the sample size. This is a particularly pressing issue in quantitative research focusing on intersectionality, where, by definition, researchers need data on students belonging to several sub-groups. Such two-, three- or more dimensional intersections, however, have often extremely small number of observations.

Statisticians have developed several techniques that can ensure representativeness while avoiding low cell counts for certain intersections: stratified random sampling, and purposive, quota or snowball sampling (Else-Quest and Hyde, 2016[79]). Under stratified random sampling, the population is first divided into sub-groups (e.g., intersections) from which individuals are randomly sampled. Purposive, quota or snowball sampling, while no longer representative, are more suitable for qualitative methods, whereby participants are recruited to be typical of the population of interest through networks or by asking research participants to refer eligible peers (Scottish Government, 2022[80]).

The OECD Survey on Social and Emotional Skills uses stratified random sampling and provides reliable and comparable insights at an international level for policy makers and educators on how social and emotional skills relate to key life outcomes (OECD, 2021[81]). The survey collects information on environments at home, in school, among peers and other background characteristics (including gender, immigrant and socio-economic backgrounds) to analyse the complex interactions between these factors and student well-being outcomes. Similarly, the New Brunswick Student Wellness Survey in Canada allows for comparisons of some aspects of psychological, physical, social and material well-being disaggregated by several dimensions of diversity (with some limitations[5]) considered by the Strength through Diversity Project (New Brunswick Health Council, 2022[82]).

Since 2016, the *Institut für Arbeitsmarkt- und Berufsforschung (IAB), Bundesamt für Migration und Flüchtlinge (BAMF), Sozio-oekonomische Panel (SOEP)* (the Institute for Employment Research, Federal Office for Migration and Refugees, Socio-Economic Panel) conduct the IAB-BAMF-SOEP Survey of Refugees in Germany to obtain reliable information on the circumstances faced by people who have sought protection in Germany in recent years. For this purpose, information on refugees' schooling and vocational training and on their current work situation is collected, among other things (Federal Office for Migration and Refugees, 2022[83]). In many countries, however, refugees are absent from systematic and comprehensive data collections in schools (Siarova and van der Graaf, 2022[84]). Researchers have also noted that data collections focusing on refugees are often not comparable between countries, their methodologies vary and available data is often limited and fragmented (European Union and the United Nations, 2018[85]; Wiseman and Bell, 2021[86]).

Absence of data is a particularly challenging problem for researchers focusing on education outcomes of LGBTQI+ students (McBrien, Rutigliano and Sticca, 2022[17]). Most countries do not systematically collect information on sexual orientation and gender identity, nor include it in regular censuses (OECD, 2019[87]).[6] However, some practices around the collection of data on LGBTQI+ students exist. In the United States, the National School Climate Survey provides information on some aspects of student psychological, social and material well-being outcomes with particular focus on LGBTQI+ students. The survey reports on experiences of discrimination, harassment, school climate and resources of lesbian, gay, bisexual, transgender and queer youth (Kosciw et al., 2020[88]).

Survey non-response presents a major challenge

High non-response rates make analyses of some student groups particularly challenging in sample-based surveys (UNESCO, 2020[15]). Some student groups might be hesitant to disclose sensitive information of their identity for fears of discrimination or persecution. This is particularly challenging for ethnic minorities and Indigenous populations. In the United States, Black or African American alone or in combination, American Indian or Alaska Native alone or in combination, and Hispanic or Latino populations were the most undercounted in the 2020 census (Jensen and Kennel, 2022[89]). Roma populations in Europe and Indigenous populations are also hesitant to self-identify resulting in their underrepresentation in censuses (Csata, Hlatky and Liu, 2020[90]; Jamieson et al., 2021[91]).

Improving response rates is not an easy task, and a range of factors and strategies can be considered. In terms of survey design, it is important to carefully tailor questions to the intended audience and to ensure that is "user-friendly" with an appealing appearance (Smith and Bost, 2007[92]). Survey length is also crucial. While longer surveys generally yield lower response rates, some research shows that the quality of provided answers is not affected (Deutskens et al., 2004[93]; the iConnect consortium, 2011[94]). The way in which answers are collected (by mail, electronically, by phone, in person, etc.) should also be considered in light of the target audience (Smith and Bost, 2007[92]). In the survey administration phase, it is important to clearly explain who will be able to view individual responses (in order to alleviate participants' potential concerns relating to confidentiality) and how the data will be used and how it can help communities (ibid.). For instance, as a strategy to improve response rates for Indigenous students, some schools in Nova Scotia (Canada) focused on building awareness among parents and students regarding how the

information collected in the particular surveys/data collections would be used to determine how to more effectively support Indigenous students such as through the allocation of Indigenous student support workers and recommendations targeted scholarship opportunities (OECD, 2017[95]).

Alter-identification is also a strategy that can be used where, due to the risk of stigma or discrimination (among other reasons), individuals may be reluctant to self-identify as having a particular background or belonging to a particular group. Alter-identification is where a third party estimates the size of a particular group and/or their participation in specific projects and programmes (European Social Fund Learning Network, 2018[96]). For instance, the Slovak Republic collects data on Roma communities by asking municipality representatives to estimate the share of Roma population under their jurisdiction (Ministry of Interior of the Slovak Republic, 2022[97]).

International surveys and international benchmarking

Surveys administered by international or non-governmental organisations can be useful in filling some of the gaps that exist in national data collections. For instance, OECD countries take part in PISA, where results are disaggregated by gender, immigrant and socio-economic background, to the extent sample sizes allow for it. The individual national versions of international questionnaires accompanying international large-scale assessments can also be complemented with specific items. For example, some countries complemented the PISA 2018 student questionnaire with questions about students' ethnic/Indigenous background or gender identity (OECD, n.d.[98]). To protect individuals' privacy rights, these data are neither made publicly available nor are reported in PISA international reports. They can be useful, however, in national analyses and monitoring frameworks.

Obtaining internationally comparable data is an issue for the concept of special education needs. While data collections on students with SEN are common (Figure 6.3), definitions among countries vary (Chapter 1). One of the consequences is that there is a large variance in the shares of students with SEN, from 1% in Sweden to 21% in Scotland (United Kingdom) (European Agency for Special Needs and Inclusive Education, 2022[99]). Definitional issues are also present in the concept of giftedness. Chile, the Flemish Community of Belgium, Iceland, Ireland, Scotland (United Kingdom) and Spain use the terms "ability", "high cognitive skills", "exceptionally able" and combinations or variations of these to describe giftedness (OECD, 2022[16]). Korea, Türkiye and the United States used the word "talent" or its variations (ibid.). The French Community of Belgium, Italy and Norway defined gifted students in terms of their high potential (ibid.).

One of the few sources of international (though not necessarily internationally comparable) data on SEN is the European Agency for Special Needs and Inclusive Education, which publishes information on the number of students with an "official decision of SEN" by the International Standard Classification of Education (ISCED) level as well as gender in 31 European member countries (European Agency for Special Needs and Inclusive Education, 2022[99]). In addition, the agency provides contextual country background notes. These include the basic definitions of "special education needs", educational assessment procedures, legal background and other information (European Agency for Special Needs and Inclusive Education, 2022[99]).

Information on LGBTQI+ individuals can be compiled by international non-governmental organisations. For example, at the European level, the International Lesbian, Gay, Bisexual, Transgender, Queer and Intersex Youth & Student Organisation compiles a regular report on LGBTQI+ inclusive education systems based on data collected from national civil society organisations (Box 6.5).

Box 6.5. The LGBTQI Inclusive Education Index

Based on questionnaires submitted by civil society organisations that focus on the rights of sexual minorities and gender identity issues, The International Lesbian, Gay, Bisexual, Transgender, Queer and Intersex Youth & Student Organisation created ten indicators that aim to capture the extent to which governments have implemented laws and policies that foster the inclusion of all learners in the education system.

The indicators are developed around ten areas. Within each area, several factors are "graded" usually on a scale 0 to 10 (with some going to negative values indicating a high degree of discrimination). The areas are as follows:

- **Anti-discrimination law applicable to education**: this focuses on anti-propaganda laws that might ban the display of sexual orientation, gender identity and gender expression, or variations in sex characteristics within educational settings. It also covers anti-discrimination legislation and the extent to which it covers sexual orientation, gender identity and expression and variations in sex characteristics;
- **Policies and action plans**: this assesses whether anti-bullying policy or national action plans are in place;
- **Inclusive national curricula**: this determines whether compulsory national curricula are inclusive of LGBTQI+ people;
- **Mandatory teacher training**: this indicates whether sexual orientation, gender identity and gender expression, or variations in sex characteristics are mentioned in compulsory teacher training programmes;
- **Legal gender recognition**: this assesses whether gender recognition provisions are in place;
- **National or regional data collection on bullying and harassment**: this determines whether evidence on bullying and harassment based on sexual orientation, gender identity and gender expression, or variations in sex characteristics is collected;
- **Support systems for young people**: this focuses on the existence of LGBTQI+ youth services and groups, as well as support and guidance from school staff;
- **Information and guidelines**: this asks whether policies prohibit the presence of LGBTQI+-related information;
- **School environment and inclusion**: this focuses on hostility or inclusiveness of extra-curricular activities available to LGBTQI+ students. It assesses whether LGBTQI+ students were excluded from extra-curricular activities or whether the establishment of LGBTQI+ student groups was prohibited;
- **International commitments**: this assesses whether the country is a member of the European Governmental LGBTI Focal Points Network and has signed the UNESCO Call for Action by Ministers on Inclusive and Equitable Education for All learners.

Source: IGLYO (2022[39]), LGBTQI inclusive education report, https://www.education-index.org/wp-content/uploads/2022/05/IGLYO-LGBTQI-Inclusive-Education-Report-2022-v3.pdf (accessed 23 May 2022).

The lack of recognition of some ethnic groups (e.g., Roma) as official minorities, combined with colour-blind approaches, results in scarce international data collections on ethnicity (Rutigliano, 2020[100]). Moreover, administrative categorisation of "ethnicity" often differs. Some countries refer to "ethnic minority groups" or "minority ethnic groups" others base their definitions on nationality. In the latter case, ethnicity as a concept is then often unrecognised in official statistics and policy making (Rutigliano, 2020[100]). For some ethnic

groups, such as Roma, there is a substantial heterogeneity in the definition and terminology of "Roma population", making international data collections inherently more challenging (ibid.). In regard to Roma populations, evidence is often based on international surveys, such as those conducted by the European Agency for Fundamental Rights. The 2016 European Union Minorities and Discrimination Survey collected a wide range of information on education, health, housing, labour market participation, discrimination and living conditions of the Roma population. Within education, they specifically focused on participation, segregation and educational attainment (European Union Agency for Fundamental Rights, 2018[101]).

Specific reviews on equity and inclusion

External evaluations in the areas of equity and inclusion can also come from school evaluation bodies, such as school inspectorates and independent institutions under the ministries of education. These reports often use evidence gathered from school inspection visits, data from administrative sources (e.g., national assessments) as well as international large-scale assessments. In Denmark, the *Danmarks Evalueringsinstitut* (Danish Evaluation Institute) is an independent institution the aim of which is to undertake systematic and mandatory evaluations of teaching and learning at all levels of the education system from pre-school to postgraduate level (The Standing International Conference of Inspectorates, 2021[102]). At the primary level, one of the themes they focus on is "inclusive learning environments" and several reports were published on this topic. One project includes a collection of experiences on the development of inclusive learning environments. More detailed reports by the Institute focused on the co-operation between teachers and special education teachers, and the process of collaborative teaching methods[7] (Danish Evaluation Institute, 2020[103]). Other reviews covered the inclusion of newly arrived students in the country into mainstream classes along with a detailed description of six selected municipalities and a specific review on teaching students with dyslexia (Danish Evaluation Institute, 2019[104]; Danish Evaluation Institute, 2020[105]).

In 2022, the Lithuanian *Nacionalinė švietimo agentūra* (National Education Agency) undertook a review of inclusion in and through education, analysing factors such as the representation of students with SEN in education above lower secondary level, the extent to which teachers are prepared to meet the diverse needs of students through continuous professional learning, and the participation of students with an immigrant and refugee background at various levels of education (National Education Agency, 2022[106]). The review also covered psychological well-being and labour market outcomes of young Lithuanians, as well as geographic disparities in education quality.

In the Netherlands, the *Inspectie van het Onderwijs* (Dutch Inspectorate of Education) covers the topic of inclusion in their annual education review. Progression through education of until the end of the secondary level is covered for students with SEN, with an immigrant background and for students whose parents have a lower level of education. The report also covers school changes, school segregation, school repetition and drop-out. Beyond student outcomes, the review also covers quality assurance and its contributions to inclusive education, the geographic dispersion of teacher shortages, and the general working conditions of school staff (Dutch Inspectorate of Education, 2022[107]).

Use of monitoring results

Education systems use several approaches to utilise and publish the results of their data collections. As elaborated in the previous sections, these include the production of an annual statistical report, a governmental report on the state of education, reports by specific evaluation agencies, summary reports on results from national assessments, national audit reports on the education sector or part of it, and various reports by sub-national educational jurisdictions (OECD, 2013[1]). Countries often use a combination of the formats that vary in design and content significantly (ibid.).

Apart from products delivered in written formats, education systems heighten the accessibility and use of education evaluation results by publishing them in the form of databases and dashboards for a wide range

of stakeholders, including the general public. In the United States, for instance, data from the National Assessment of Educational Progress are summarised in dashboards that disaggregate data, including student learning outcomes, by geographic location, gender, ethnic background and socio-economic background (e.g., parental education level and National School Lunch programme eligibility) (The Nation's Report Card, 2022[108]). The Czech School Inspectorate designed a dashboard that portrays various indicators in a map, such as the number of students by various typologies of SEN (Czech School Inspectorate, 2022[109]). Data can also be visualised by non-public organisations. Prokop Analysis and Quantitative Research (PAQ Research) in the Czech Republic, for example, uses mapping tools to illustrate a wide range of indicators that can cause educational inequities at the regional level, including the number of children in housing needs, the number of children in inadequate of overcrowded dwellings, resourcing per student, early school leaving indicators, grade repetition and many others (PAQ Research, 2022[110]). Data can also be downloaded in user-editable formats for further analyses.

Strengthening the channels between monitoring systems and practices

A considerable challenge for policy makers is to use the full potential of the instruments used to measure progress towards equity and inclusion to inform policy development. However, a systematic review of Roma-inclusion measures in the fields of education, employment, healthcare, housing and non-discrimination/fighting antigypsyism[8] revealed that the results of monitoring and evaluation reports are often not considered due to the lack of co-operation between stakeholders (Fresno et al., 2019[111]). This includes a lack of co-ordination between administrators of the intervention and local authorities, a lack of close co-operation with key stakeholders and the lack of a common understanding of the main goals in the case of several institutions coordinating one project. Furthermore, there is often a disconnect between research and the policy-making process. An obvious, yet often overlooked point is that researchers need to communicate their findings to policy makers in a clear language that can be easily understood by non-specialists (Oliver and Cairney, 2019[112]). Thus, ensuring co-operation between different actors is key in ensuring that the information obtained through monitoring is used to inform policy development.

Some organisations have reflected on these issues and are attempting to make research more accessible to schools. The Education Endowment Foundation/Sutton Trust, for instance, set up the “Research Schools Network”. Schools that are members of the network can benefit from continuous professional learning for teachers and school leaders on how to improve classroom practices based on evidence. Some initiatives that were explored include the support of students with SEN during testing times, the reintegration of disengaged learners, and strategies to reduce the impact of school closures on disadvantaged students during the COVID-19 pandemic (Research Schools Network, 2020[113]; Research Schools Network, 2020[114]; Research Schools Network, 2022[115]).

Co-operation is also important within the public sector, where responsibilities are often strictly divided between ministries and departments. Many issues will necessarily concern multiple agencies and departments and co-operation is important to ensure the development and implementation of an effective monitoring and evaluation framework. Refugee integration in education, for example, often requires co-ordination with the departments responsible for housing, employment and health. In France, all relevant ministries meet every six months to review the implementation of the National Strategy for Refugee Reception and Integration. Moreover, all relevant stakeholders in refugee integration, including local actors, are involved in the interpretation and monitoring of the evidence, and review of the implementation of the strategy at an annual conference organised by the inter-ministerial delegation for the reception and integration of refugees (Siarova and van der Graaf, 2022[84]). The French Council for School Evaluation is another example where stakeholders from different sectors meet to discuss evidence focused around school evaluation while also focusing on equity in education (Box 6.6).

Box 6.6. The French Council for School Evaluation

The main missions of the independent French Council for School Evaluation are to ensure consistency in the evaluation of public education policies; comment on methodologies, tools and results of school assessments; define and comment on tools for school self-assessments, and synthesise research on students' educational achievement.

There are 13 members in the Council, four of whom are members of the parliament, three are senior officials from the Ministry of Education and six are various education experts working under the supervision of the Council's president.

To provide advice on removing inequalities in academic success (whether based on socio-economic status, geographic location or gender), the Council formed a committee that builds an inventory of studies about evaluation techniques on equity and equality in student assessments. The aim of the committee is to develop a targeted evaluation plan in this area, and ultimately implement student assessments that respect the principles of equity and that are adapted to inclusive education (The French Council for School Evaluation, 2021[116]).

Source: Ministry of National Education and Youth (2022[117]), Conseil d'évaluation de l'École [The French Council for School Evaluation], https://www.education.gouv.fr/conseil-d-evaluation-de-l-ecole-305080 (accessed 23 May 2022).

Monitoring systems are also used to provide useful information for the management of local authorities and schools. However, while schools are the primary units that send data to national (or sub-national) authorities, they do not always receive a statistical analysis of their profile to support them in their internal analysis and further planning. One way in which national (or sub-national) authorities can support schools is to compare their data with central-level indicators or with other relevant benchmark groupings (e.g., comparing the outputs with schools with similar inputs), or by providing them with frameworks and guidelines for internal school evaluation. These practices are further elaborated in the section on Supporting schools in improving equity and inclusion practices through evaluation processes.

Evaluating policies, programmes and processes to improve equity and inclusion in education

In line with other OECD definitions, evaluation refers to the structured assessment of policies or programmes to reliably determine their merit and value according to the specific criteria (Golden, 2020[118]). Evaluation can inform policy development, curriculum, planning, reporting, resource allocation decisions and performance management and, in the context of limited resources, can be crucial in ensuring the highest value for money (Cerna et al., 2021[2]).

Evaluation frameworks of policies and programmes in the area of equity and inclusion are shaped by the broader context of education policies as well as existing traditions, values and cultures. While the need for evaluation in the education sector is widely acknowledged, traditions and cultures in education shape the nature and significance of evaluation and assessment activities (OECD, 2013[1]; OECD, 2022[119]). As such, there are no common definitions and concepts of policy evaluation and practices across OECD countries vary (Golden, 2020[118]).

Education system evaluation can be implemented with the aims of accountability, development or diagnosis (OECD, 2013[1]). Accountability relates to measuring student or school outcomes in order to identify underperforming actors in the system. The consequences can then take the form of rewards or sanctions, such as career advancement or salary progression of school staff. Development relates to

identifying the strengths and weaknesses of the system or individual schools. Finally, diagnosis aims to measure the extent to which school outcomes are aligned with objectives.

This section mostly focuses on the development aim of evaluation. Furthermore, while evaluation can also take the form of assessments (such as examinations; see Chapter 5) and appraisals (e.g., judging the performance of teachers), this chapter focuses on judgements of the effectiveness of schools, programmes and policies, with a particular emphasis on equity and inclusion.

Challenges in evaluating equity and inclusion

Despite the fact that the evaluation of policies was among the identified priority areas of many education systems between 2008 and 2019, OECD reviews have continued to identify a possible absence or underdevelopment of system evaluation components, insufficient clarity in evaluation processes, possible gaps in data collections that could inform improvement and issues with quality (OECD, 2019[46]). Across OECD countries, only about one in ten education reforms is followed by any attempt to evaluate its impact (OECD, 2015[120]). Evaluations in the area of equity and inclusion are also rare. Researchers have identified a lack of programme evaluations for LGBTQI+ students, gifted students, and ethnic minorities and Indigenous populations (McBrien, Rutigliano and Sticca, 2022[17]; Rutigliano, 2020[100]; Rutigliano, 2020[100]). This section discusses the different challenges in developing and implementing effective evaluation systems for equity and inclusion in education.

Evaluations for equity and inclusion are often not effectively resourced. The process of evaluation can be resource intensive and often requires specialised tools, skill sets and the involvement of third-party institutions (e.g., from the academic sector). Despite growing investment, funding for education research remains lower than in other comparable sectors (OECD, 2022[119]), also affecting evaluations in the areas of equity and inclusion. Choosing the appropriate methodologies and metrics for evaluation can also be a resource-intensive task. This is particularly true if causal impacts are to be estimated; and incorrect causal conclusions can have significant costs for education systems (Cook, 2002[121]). Indeed, the identification of the quality of research is a recurring theme in education research. For instance, definitions and standards of what merits a high-quality research are often left undefined (OECD, 2022[119]). Consequently, evaluations are often left without causal interpretations or are limited to descriptive analyses (to the extent such analyses can still be considered “evaluations”).

In the area of inclusion of European Roma minorities, Fresno et al. (2019[111]) analysed more than 60 reports in monitoring, assessment and evaluation of programmes in the areas of education, employment, discrimination and others. The authors concluded that evaluations using robust evaluation methods are scarce. While many studies described the socio-economic situation of Roma groups, only a few looked at the effectiveness of measures attempting to alleviate the effect of socio-economic background on outcomes (Rutigliano, 2020[100]).

Similarly, the general lack of data as well as resources results in a low number of policy and programme evaluations related to Indigenous populations. Even when evaluations are conducted, they are often qualitative in nature and small in scale. Furthermore, they often cover communities in specific educational contexts, making generalisations challenging (OECD, 2017[95]).

Evaluations require a commitment to continue in the evaluation process over a longer period of time (Golden, 2020[118]). However, evaluations are often limited in terms in their time scale. For instance, in the area of giftedness, assessment within the field is often related to identification challenges. However, once students are identified as gifted, their progress in education is rarely evaluated over the long-run (Parekh, Brown and Robson, 2018[122]). This prevents a proper evaluation of programmes and policies addressing the needs of gifted students. Thus, in many countries, an additional challenge to adequately identify students’ needs lies in a continuous assessment of their progress.

Socio-political considerations can undermine the evaluation process or even discourage evaluations from taking place (Golden, 2020[118]). Some evaluations can be hampered by a lack of political will, while others might be published during an inconvenient time in the political cycle (Bamber and Anderson, 2012[123]). Fears of policy failure can result in evaluations not taking place at all. Furthermore, given that evaluations can span across political cycles, an incumbent government might not always feel enthusiastic in acknowledging positive evaluation results for a policy introduced and undertaken by a different administration (Golden, 2020[118]).

Moreover, evaluations are often absent for dimensions of sexual orientation and gender identity (McBrien, Rutigliano and Sticca, 2022[17]). A possible reason might be that many initiatives are relatively new and thus cannot be evaluated yet. However, the topic of LGBTQI+ is also considered controversial in public debates in some regions, and initiatives that support LGBTQI+ students are at times contended (McBrien, Rutigliano and Sticca, 2022[17]).

Finally, evaluation results do not always permeate to local and classroom level. Research has identified that a lack of co-operation can result in a barrier between researchers and school staff hindering the process of monitoring and evaluation (Fresno et al., 2019[111]; OECD, 2022[119]). This disconnect between research and schools can have many negative consequences. In regards to students with ADHD, for instance, evidence shows that giving additional time during exams is a widely adopted practice, yet with little support in the academic literature. In contrast, peer-to-peer learning, which has been found effective in terms of improving student well-being outcomes, is less frequently applied (Mezzanotte, 2020[19]).

Enabling factors for evaluation in the area of equity and inclusion

Each education system is unique and shaped by the broader context of education policies as well as existing traditions, values and cultures. This means that there is no one-size-fits-all approach in terms of ensuring the implementation of effective evaluation policies, with the complexities inherent in different education systems requiring consideration of a variety of methods and strategies (Golden, 2020[118]). Despite this, several factors that can help to facilitate the implementation of evaluation policies can be identified. These are discussed below.

Evaluations cannot happen without developing a shared mind-set of evaluative thinking. This means emphasising and valuing a deep and critical enquiry process, question assumptions and the status quo and view mistakes as necessary in the learning process (Golden, 2020[118]). In the area of inclusion this can also entail addressing concerns regarding inclusive practices at all levels of the education system (AuCoin, Porter and Baker-Korotkov, 2020[124]). To address these concerns, stakeholders might need to reflect on public priorities, renew a shared vision, bring diverse perspectives and experiences to the process, and ultimately fundamentally change the attitudes about diversity, equity and inclusion at all levels of society (ibid.). This was the approach taken in several evaluation rounds of the inclusive system in New Brunswick (Canada) (Box 6.7).

Box 6.7. Inclusive reforms in New Brunswick (Canada)

Bill 85, enacted in 1986, gave public schools a mandate to include all students within the education system of New Brunswick (Canada). After two decades, a review of the policy highlighted several challenges, such as the need for clarity in the policies and particularly the need to develop a consensus on the definition of inclusion. As a result, the government adopted a clear definition of inclusive education along with several related terms. Among others, the common learning environment concept was defined which broadened the definition of inclusive education to all students, not just selected groups, e.g., those with SEN.

In 2012, the government mandated another review that focused on approaches at the district, school

and classroom level to better support learners. As a result, Policy 322 was developed to provide guidance and clarity on programmes and procedures to every public school in the province. Given New Brunswick's dedication to inclusive reforms, the province was awarded the UNESCO/Emir Jaber al-Ahmad al-Jaber al-Sabah Prize for "long-standing commitment to a systemic approach to inclusive education for learners with disabilities and for its vision of education as a catalyst for social justice and equity" (UNESCO, 2014[125]).

The evaluations were framed by bringing diverse perspectives and experiences to the process. Apart from deputy ministers, school district officials and parent advocacy groups, they also included school leaders and teachers, and eventually also university faculty and teacher professional associations. The opportunities to engage in the process resulted in a greater commitment on the side of school leaders and teachers, and their subsequent investments into continuous professional learning in the area of inclusion.

Source: AuCoin, Porter and Baker-Korotkov (2020[124]), New Brunswick's journey to inclusive education, 10.1007/s11125-020-09508-8.

Evaluations will not happen without consensus-building through a consultative and co-operative process. Significant divergences and views and interests exist within various stakeholder groups and policy makers need to reconcile these to avoid the perception that evaluation and assessment policy is imposed to other groups in a top-down fashion (OECD, 2013[1]). These collaborative efforts can include expert groups, committees, taskforces or communal efforts by different evaluators (Golden, 2020[118]). Collaborative approaches to evaluation can also promote learning and knowledge transfer between national (or sub-national) authorities and local and regional authorities. For instance, in 2011, Canada upgraded its gender-based analysis and created the Gender-Based Analysis Plus approach (GBA+) (Government of Canada, 2022[126]). This approach is an *ex ante* evaluation budgeting process based on the premise that policy making should consider gender as well as other dimensions of diversity and the interactions between them. To spread the use of the GBA+ approach to all provinces and territories of Canada, the Government engaged sub-national entities in a co-operative process by developing GBA+ awareness weeks. During these events, the participants explore the changing realities and inequalities of diverse groups, and examine how the GBA+ approach can be used to create effective policies, programs and services (Government of Canada, 2022[126]). As a result, various Canadian provinces have already embraced this process (Hankivsky and Mussell, 2018[127]).

The process of evaluation can be streamlined by involving internal administrative and analytical capacities. Many central authorities already have internal monitoring and/or evaluation units, although they often operate in silos (e.g., education management and information systems operate independently of policy analysis units) (Golden, 2020[118]). In the Czech Republic, the inspectorate was involved in two evaluations of the inclusive school reform (Box 6.8).

Box 6.8 Evaluation of inclusive reforms in the Czech Republic

As of 2016/17, the Czech school system assumes that all students are educated in the mainstream school setting (including students with SEN and gifted students). Education in special school/class settings is allowed only in exceptional circumstances when supporting mechanisms are not sufficient. Students are entitled to appropriate supporting mechanisms that address their special needs (understood broadly and include education needs as well as needs resulting from disadvantaged socio-economic, cultural or immigrant background). Supporting mechanisms have different stages (and costs attached to them) and are assigned at the discretion of teachers, schools or special education centres. The Czech School Inspectorate evaluated this policy change at the end of 2016/17 and 2017/18 in the

form of several hundreds of school visits.

The 2016/17 evaluation found that schools were often using individualised teaching methods, co-operation between teachers and teaching assistants had improved, school activities were accessible to all students and after-school activities had also benefited from supporting mechanisms. On the other hand, some schools observed an increased administrative burden. Furthermore, many schools were using supporting mechanisms inefficiently (e.g., buying new school resources even though old ones could have been reused), although the 2017/18 evaluation no longer identified this as a challenge. Few teachers attended continuous professional learning related to the policy change and schools often struggled to find qualified non-instructional staff. In 2017/18, an additional challenge presented the involvement of teacher assistants in the process of teaching. For instance, some school visits revealed that the teaching assistant focused on one particular student and the teacher on the rest of the class, with the effect that the student in question was therefore not fully and equally included in the teaching process.

Source: Czech School Inspectorate (2017[128]), Společné vzdělávání ve školním roce 2016/2017 [Common education in 2016/2017 school year], https://www.csicr.cz/getattachment/7734c437-a133-4411-b8b6-ed11776ad4fe/TZ-Spolecne-vzdelavani-16-10-2017.pdf (accessed 20 May 2022) and Czech School Inspectorate (2018[129]), Vybrané aspekty implementace společného vzdělávání v období 1. pololetí školního roku 2017/2018 [Selected aspects of the implementation of common education in the first half of 2017/2018], https://www.csicr.cz/CSICR/media/Prilohy/2018_p%c5%99%c3%adlohy/Dokumenty/TZ-Vybrane-aspekty-implementace-spolecneho-vzdelavani-v-1-pololeti-2017-2018.pdf (accessed 20 May 2022).

The clarity of purpose of the evaluation and the clear communication of objectives are also key factors that researchers have identified for the development and implementation of effective evaluations (OECD, 2013[1]). A comprehensive evaluation framework and clear guidelines can play a valuable role in this respect (Golden, 2020[118]). Gender policy is an area with many guidelines and frameworks. In fact, the research in the area has permeated into policy making and *ex ante* gender impact evaluations are now sometimes conducted. These aim to ensure that during the policy-making process, disadvantages specific to women, and more recently also non-binary individuals in some jurisdictions (such as in Victoria (Australia)), are not overlooked (European Institute for Gender Equality, 2016[130]; Federal Minister for Women and Civil Service at the Federal Chancellery, 2012[131]; Victoria State Government, n.d.[132]). The development of toolkits for policy makers to conduct gender impact evaluations can help to clarify the purpose and the development of evaluations. The European Institute for Gender Equality, for instance, provides a framework for gender impact assessments (Box 6.9). As a result, gender impact evaluations are not uncommon and interventions have been evaluated in terms of boys' engagement with school (Glass, 2013[133]), impact on students' literacy skills by male as well as female role models (Marx and Roman, 2002[134]; Wood et al., 2016[135]) and others. In some countries, such as Austria, impact assessment on gender equality are even required by law (Federal Minister for Women and Civil Service at the Federal Chancellery, 2012[131]).

Box 6.9. Gender Impact Evaluations: Gender Mainstreaming Toolkit

The European Institute for Gender Equality structures gender impact assessments into five steps grouped in three stages:

- **Gender relevance assessment**: In the first two steps (definition of the policy purpose, checking gender relevance), the aim is to describe the purpose of a policy, law or programme and assess its impact, both direct and indirect, on gender equality;
- **Gender impact assessment**: In the subsequent two steps (gender-sensitive analysis, weighing gender impact), the gender impact analysis is carried out. Such analysis is necessary for the

stakeholders to understand the current situation and how it could evolve without public intervention. Moreover, the analysis should be able to evaluate how the intervention might change the existing situation and how it will impact gender equality and gender relations; and

- **Gender quality assessments**: The last step (findings and proposals for improvement) consists of the formulation of conclusions in terms of the impacts on targeted women and men. Proposals promoting gender equality should be put forward.

Finally, to provide quality assurance to the whole process, it is also recommended to follow up on the gender impact assessment and thus identify potential adjustments needed in policy or law. The Gender Mainstreaming Toolkit is available on the website of the European Institute for Gender Equality in an interactive form so that various stakeholders can integrate it in their frameworks.

Source: European Institute for Gender Equality (2016[130]), Gender Impact Assessment: Gender Mainstreaming Toolkit, https://eige.europa.eu/publications/gender-impact-assessment-gender-mainstreaming-toolkit (accessed 10 May 2022).

Gender impact assessments are also conducted at the regional level. The Swedish Association of Local Authorities and Regions created a set of checklists and guidelines to systematise the decision-making process at the local level and thus clarify the purpose of gender impact assessments (European Institute for Gender Equality, 2016[130]). The aim of these resources is to highlight any possible data gaps, analyse the consequences of the proposed policy for men and women, and describe how the proposed changes lead towards gender equality goals.

Upscaling and institutionalisation following evaluations that improve equity and inclusion

Upscaling or institutionalisation can be viewed as results of an evaluation. Upscaling can be defined as the process of expanding the effects of a practice not only to a larger group of beneficiaries, but also to achieve longer-term changes in practice and belief (depth), continuation of intervention effects after initial implementation (sustainability), and strong ownership of the reform (Coburn, 2003[136]). Upscaling can result in institutionalisation, i.e., the process through which new practices or innovations are integrated into the context of focus and become prevailing practices in an organisation, system or society (Nworie, 2015[137]).

The challenges that prevent the upscaling and institutionalisation of programmes or policies are similar to those of evaluations itself. There is often a lack of resources for the process of upscaling and institutionalisation, such as money, tools, materials and time commitment (Ely, 1999[138]; Koehler, Palaiologou and Brussino, 2022[139]). The broader communities are often not prepared for the scaled implementation of the practice, or lack a shared understanding of the benefits of the programme (de Hoop et al., 2019[140]; Seelos and Mair, 2016[141]). The process often lacks leaders who are equipped with the skills needed to support the upscaling process, and who understand both the local context of the programme as well as broader societal factors that can hinder upscaling (Seelos and Mair, 2016[141]). The process often lacks iteration, i.e., it does not foster a continuous re-examination and evaluation of the interventions before they scaled up (Glennan et al., 2004[142]). Finally, political support is often lacking in the process of institutionalisation (Koehler, Palaiologou and Brussino, 2022[139]).

Despite these challenges, smaller programmes have become more widespread following evaluations, and have led to improved learning and well-being outcomes for diverse students. The Tutoring Online Programme in the city of Milan (Italy), for instance, provided online tutoring by trained volunteers from higher education institutions for lower secondary students (Carlana and La Ferrara, 2022[143]). Students participating in the programme were lagging behind their peers during distance learning as a result of school closures during the COVID-19 pandemic. A large proportion of participating students were

socio-economically disadvantaged, with language barriers, an immigrant background or SEN. The results of the evaluation showed positive effects on learning outcomes as well as student effort, attendance and behaviour. In terms of non-academic outcomes, the evaluation revealed positive effects on students' beliefs in their ability to control the outcome of events in their lives and psychological well-being. Some of these effects were even more pronounced for disadvantaged students, students with an immigrant background or students with learning disabilities. Given the generally positive results, the programme was scaled up to 3 000 students in 2021/22 as part of a wider co-operation with the Italian Ministry of Education.

Successful scaled-up practices can also be found in the area of teacher co-operation and knowledge-sharing. In New Zealand, the Learning and Change Networks initially started as voluntary local networks between schools to boost student achievement, including that of Indigenous students. Teachers came together to analyse data, share classroom practices and students' views on their learning. The successful feedback on this initiative led the government to scale up the practice to a country-wide initiative, Communities of Learning, to incentivise more schools for such knowledge sharing (OECD, 2017[95]).

The *Schlauschule* in Bavaria (Germany) is an example of a programme that started locally and eventually became institutionalised at the federal level. The programme targets refugee students between the ages of 16 and 21, including unaccompanied minors (Koehler, Palaiologou and Brussino, 2022[139]). Each school year, about 300 students are taught in 20 classes. An individualised approach is promoted with the ultimate goal of students becoming more familiar with the Bavarian schooling system and transitioning to the mainstream track. Depending on prior education and learning needs, students spend one to four years in a *Schlauschule*. Besides academic competencies, the programme aims to develop general knowledge and key competences, such as solidarity and equal treatment independent of gender, origin, age and socio-economic status. It started as a small-scale programme responding to the need for education of asylum-seekers in Bavaria. Initially, the programme educated only 20 unaccompanied minors. In 2004, it was approved as a vocational preparation school for refugees by the Bavarian Ministry for Education. This fostered partnerships with new schools and a wider participation of refugees in the programme. Once Germany ratified the UN Declaration of the Rights of the Child in 2010, the programme inspired the Federal State of Bavaria to introduce the *Berufsvorbereitungsklassen* (vocational preparation classes) that follow the example of *Schlauschule*.

Supporting schools in improving equity and inclusion practices through evaluation processes

Interventions to support equity and the inclusion of students are not only managed at the national level, but also at a local and school level. Individual school evaluations can be an important tool that assists with decision making, resource allocation and school improvement. The effective monitoring and evaluation of schools, including the aspects of equity and inclusion, are central to their improvement. Schools need feedback on their practices to identify areas for improvement.

The way in which school evaluations are designed and implemented can have important impacts on efforts to improve equity and inclusion (Cerna et al., 2021[2]). On the one hand, effective systems need to be in place to ensure that local actions are being taken and that they are in line with national requirements (OECD, 2022[21]). On the other hand, disproportionate blame for systemic educational problems on any actor can have serious negative side effects, widening inequality and damaging learning (UNESCO, 2017[144]).

The evaluation at the school level can happen externally (e.g., by school inspectorates) and internally (self-evaluations). In 2018, internal as well as external evaluations were common across OECD countries. More than 94% and 76% of 15-year-old students attended schools whose leaders reported the existence of internal and external evaluations respectively in their schools (OECD, 2020[8]). Individual country

practices, however, vary and there is no “one-size-fits-all” approach that would be applicable to all education systems.

Equity and inclusion in external evaluation processes

The nature of feedback delivered by the evaluation bodies differs across education systems. In some education systems, development processes are encouraged through indirect feedback. In other systems, the evaluation focuses more on whether schools have met pre-defined criteria (Ehren, Perryman and Shackleton, 2014[145]). Education systems can also focus specifically on the topics of equity and inclusion in their policy frameworks for school evaluations. In 2022, 11 out of 34 OECD education systems provided criteria for an assessment of equity and inclusion in their policy frameworks for school evaluation (the Flemish Community of Belgium, France, Iceland, Ireland, Italy, Latvia, Lithuania, New Zealand, Portugal, Spain and Türkiye) (OECD, 2022[16]).

A number of factors can support school evaluation processes in promoting equity and inclusion. First, clear reference standards can be defined with a particular focus on equity and inclusion (Cerna et al., 2021[2]; OECD, 2015[120]). These can be developed in national quality standards that define the characteristics of quality in terms of equity and inclusion in addition to other domains (OECD, 2015[120]). In North Rhine-Westphalia (Germany), for example, addressing student diversity is incorporated within the four domains of the region’s quality standards (teaching and learning, school culture, continuous professional learning, and leadership and management) (Ministry for Schools and Education of the North Rhine-Westphalia, 2022[146]). For instance, one indicator focuses on whether diversity is respected and taken into account in schools. Reference standards can also draw from equity and inclusion policies that are currently in place. In Scotland (United Kingdom), for instance, school evaluation guidelines draw on principles and values of the “Getting it right for every child” policy (Education Scotland, 2015[147]). These principles and values place each student at the heart of the learning process, appreciate differences among students, and consider and address inequalities (Scottish Government, n.d.[148]).

Second, benchmark information on key indicators related to equity and inclusion can be offered so that schools can adequately compare themselves to broader averages (Cerna et al., 2021[2]; OECD, 2015[120]). In Portugal, for example, the Equity Indicator measures disadvantaged students’ completion of each school cycle in the expected time and with positive scores in national examinations. Each school cluster and municipality can compare their index score with the national average. The values are also used during visits by the school inspectorate (OECD, 2022[21]).

The third important aspect is the training of external evaluators in the area of equity and inclusion so that they can adequately assess a school’s performance and potentially offer assistance for improvement (Cerna et al., 2021[2]; OECD, 2015[120]). Through this process and their experience, evaluators can share good school-level practices that are in line with national educational priorities from other schools they have evaluated. Ensuring that the external school evaluators have in-depth expertise in education and teaching so as to be able to guide and support others in the process of school development is also highly important from a legitimacy perspective, which is crucial for the school evaluation process to be effective (OECD, 2013[1]). In Austria, following the 2017 education reform, school quality managers are responsible for external school evaluation (Federal Ministry of Education, Science and Research, 2022[149]). Their roles vary from the implementation of reforms and development targets, through control of regional educational planning, to provision of pedagogical expertise in areas relevant to equity and inclusion (e.g., diversity-oriented pedagogies, co-operation between teachers and support staff) (Federal Ministry of Education, Science and Research, 2022[150]). The Federal Ministry of Education, Science and Research organises training events, specialised courses for school quality managers and an annual school supervisory congress to deepen and expand school quality managers’ professional skills (Federal Ministry of Education, Science and Research, 2022[149]).

Lastly, school evaluation needs a clear purpose and to provide clear guidance on how school processes can be improved in terms of equity and inclusion (Cerna et al., 2021[2]; OECD, 2015[120]). To this end, some education systems have created clear guidelines on what is evaluated in terms of equity and inclusion. The aim of these guidelines is to provide a common base for the evaluation and draw attention to the aspects that are important. In Scotland (United Kingdom), these are summarised in the form of several quality indicators and complemented with features of highly effective practices (Box 6.10).

Box 6.10. Evaluation framework in Scotland (United Kingdom)

In Scotland (United Kingdom), the framework for school evaluation is published by Education Scotland, the institution responsible for external school evaluation. The guidelines, questions and indicators outlined in this framework are also used by school inspectors during school visits. The framework is structured around 15 quality indicators within three categories. For each indicator, the framework outlines features of highly effective practices and questions that can guide schools towards improvement.

Leadership and management

This area covers the school leadership and schools' approach to improvement. The indicators include the topics of self-evaluation for self-improvement, leadership of learning, leadership of change, leadership and management of staff, and management of resources to promote equity.

Learning provision

This area covers the quality of care and education. The indicators include the topics of safe-guarding and child protection; curriculum; learning, teaching and assessment; personalised support; family learning; transition; and partnerships.

Successes and improvements

This area covers whether schools are best placed to ensure the best possible outcomes for all learners. The indicators include the topics of well-being, equality and inclusion; raising attainment and achievement; and increasing creativity and employability.

Source: Education Scotland (2015[147]), How good is our school?, https://education.gov.scot/improvement/Documents/Frameworks_SelfEvaluation/FRWK2_NIHeditHGIOS/FRWK2_HGIOS4.pdf (accessed 11 August 2022).

Equity and inclusion in internal evaluation processes

School self-evaluation is a long-established process in OECD education systems. In some, the practice is required by law, while in other countries it is recommended or required only indirectly (e.g., by developing school guidelines) (European Commission/EACEA/Eurydice, 2015[151]; OECD, 2015[120]). Some education systems focus specifically on the topics of equity and inclusion aspects in their policy frameworks for internal evaluations. In 2022, 19 out of 34 OECD education systems provided guidelines for an assessment of equity and inclusion in frameworks for school self-evaluation (Australia, Canada, Colombia, the Flemish Community of Belgium, France, Greece, Iceland, Ireland, Italy, Latvia, Lithuania, New Zealand, Northern Ireland (United Kingdom), Norway, Portugal, Scotland (United Kingdom), Slovenia, Spain and Sweden) (OECD, 2022[16]).

By encouraging school self-evaluations, governments highlight that schools are best placed to analyse their own strengths and weaknesses and identify the main areas for improvement (OECD, 2015[120]). The

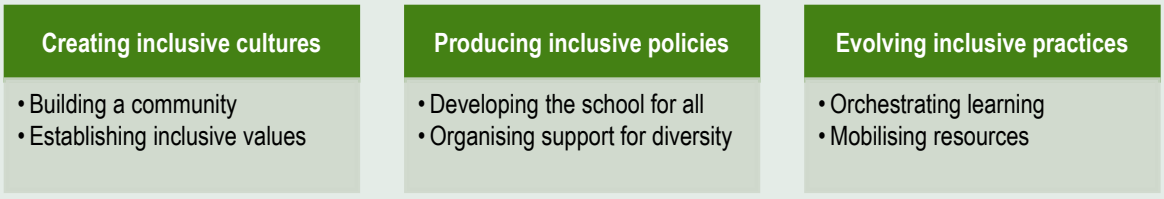
objective of a school self-evaluation is in essence similar to that of an external evaluation: to assess the effectiveness of structures and processes, including teaching and community engagement, and ultimately improve the quality of instruction and thereby student achievement (OECD, 2015[120]; OECD, 2015[52]). However, internal evaluation can lead to increased reflection on school quality and intentions to improve, a greater ownership of change, and a greater sensitivity to areas in need of improvement (Godfrey, 2020[152]). It can also help with the identification of continuous professional learning needs for teachers and the subsequent increased use of professional learning.

In the context of equity and inclusion, internal evaluation can lead to revisions in the content or organisation of the curriculum, and the provision of targeted support for groups of students. Self-evaluation can also serve to identify barriers that hinder inclusive education for all students. This can be done by analysing school climate, relationships, learning support, barriers to continuous professional learning and other aspects of the school environment. A useful framework that can guide schools in self-evaluating in this area is the Index for Inclusion (Box 6.11). The Index can also help school leaders to conduct self-evaluations that have a real and lasting impact on school processes – a common challenge for self-evaluations identified in the literature (Nelson, Ehren and Godfrey, 2015[153]; OECD, 2013[1]).

Box 6.11. Index for Inclusion

The Index for Inclusion is a self-evaluation tool for schools to help them identify barriers to and subsequently develop an inclusive school setting. It has been adopted and modified in several countries, and translated into 35 languages (Centre for Studies on Inclusive Education, 2020[154]). The resource is comprised of a set of questions and a comprehensive document that defines the terminology (e.g., inclusion in education) and explains the main concepts. One key characteristic of the Index is that it does not focus on a particular group of students (e.g., those with SEN), but focuses on all students. The Index is structured around three main dimensions of school improvement: creating inclusive cultures, producing inclusive policies and evolving inclusive practices (Figure 6.8). All dimensions are necessary to the development of inclusion at schools. Each dimension is divided into two sections to focus attention on what needs to be done to increase learning and participation in school. Each section contains several statements that schools can self-assess.

Figure 6.8. Dimensions of the Index for Inclusion



Creating inclusive cultures

Building a community indicators focus on co-operation between stakeholders, i.e. between students, between teachers, between students and teachers, between staff and parents, between schools and wider communities, and between schools and other governing structures. Indicators on **establishing inclusive values** then evaluate whether staff and students feel valued, how they treat each other and how the challenges of participation and discrimination are addressed. The principles and values of inclusive school settings should guide day-to-day decisions and should be continuously developed.

Producing inclusive policies

Developing the school for all assesses whether inclusive policies are in place for staff recruitment and student admission. This domain also includes indicators on how newcomers are helped and welcomed in a school and whether all new members are valued. **Organising support for diversity** is concerned with whether there is co-ordination between various stakeholders in school to support the needs of all students, and whether policies designed to promote inclusion take a broad approach or rather target only specific student groups, such as those with SEN. The domain also addresses barriers to attendance and bullying.

Evolving inclusive practices

The **orchestrating learning** domain is concerned with classroom practices, whether they encourage the participation of all students, embrace diversity, co-operation and mutual respect. It also focuses on participation in out-of-school activities, homework and co-operation between teachers to plan lessons. **Mobilising resources** addresses whether students support each other, whether staff share knowledge, advice and practices for the benefit of all students, and whether all school resources are utilised effectively. It also includes indicators on community engagement and distribution of school resources.

Source: Booth and Ainscow (2002[155]), The Index for Inclusion: developing learning and participation in schools, https://www.eenet.org.uk/resources/docs/Index%20English.pdf (accessed 11 April 2022).

Clear reference standards approved by the national authority, which focus on equity and inclusion, are often developed to help schools self-evaluate (Cerna et al., 2021[2]; OECD, 2015[120]). For example, the "Every School a Good School" policy developed by the Department of Education in Northern Ireland (United Kingdom) sets out a series of principles and values that should guide school improvement processes, in line with the overarching vision of ensuring "that every learner fulfils his or her full potential at each stage of his or her development" (Department of Education, 2009[156]). The policy also outlines the roles and responsibilities of different actors in school improvement and specific targets for monitoring and evaluating progress (some of which focus on equity and inclusion).

Some education systems provide schools with self-evaluation guidelines. Guidelines for effective self-evaluation can make the process easier for school staff, who often lack the capacity and time to engage in self-evaluation (Nelson, Ehren and Godfrey, 2015[153]). Time is also needed for collaborative discussions that are important in articulating the goals and questions (ibid.). In Norway, the Ministry of Education provides tools to support schools in conducting voluntary self-evaluation. The goal is to promote common reflections on school practice and help schools prioritise areas for further development, including in the area of inclusion (Directorate of Education, 2022[157]). Similarly, the Irish Inspectorate publishes school self-evaluation guidelines (Department of Education and Skills, 2016[158]). The guidelines outline a step-by-step approach for initiating the self-evaluation process, provide a set of standards that can assist schools in making judgements about the quality of teaching and learning, describe a range of evaluation approaches, provide guidance about recording the self-evaluation process, and outline a framework with domains and standards to enable schools to evaluate teaching and learning. Standards in the "leading learning and teaching" domain include whether school leaders foster a commitment to inclusion, equality of opportunity and the holistic development of each student.

Tools to support schools' evaluation can also be provided by other public bodies. Statistics Norway, for instance, is responsible for publishing value-added indicators for schools (Statistics Norway, 2017[159]). Compared to more traditional measures of equity that disaggregate outcomes by student characteristics, value-added indicators attempt to consider the composition and level of skills of students when they started their study at a particular school. By comparing each school's actual results with expected results based on the level of skills and characteristics of the student intake, they can determine whether the school exceeded the expectations, performed *at par* with expectations or underperformed. This provides schools with an indication on the extent to which they are supporting students' learning.

Improving equity and inclusion as a result of school evaluation

Whether school evaluation leads to improved progress towards equity and inclusion depends on numerous factors. It is important to consider that school evaluation impacts each school differently and various conditions determine whether and to what extent schools accept and act on the results from the analyses (OECD, 2013[1]).

Finding a common language can help translate evaluations into improved processes

The extent to which feedback obtained through the process of evaluation is actually implemented is a widely debated question. Due to resistance from schools, lack of capacity or lack of communication on the part of the evaluators, school evaluation often does not motivate schools to improve (Ehren et al., 2013[160]). Internal evaluations often suffer from poor quality and sometimes fail to answer the questions that were set at the beginning of the evaluation process (Nelson, Ehren and Godfrey, 2015[153]).

One important enabling factor to help ensure that evaluation leads to school improvement is the alignment of views and beliefs on what constitutes a "good education", and finding a common language between the evaluators and schools (Ehren et al., 2013[160]). Within schools, it is important that any differences in the meanings of inclusion are resolved and school staff arrive at a shared understanding that draws them around a common purpose (Azorín and Ainscow, 2018[161]). Shared understanding of the goals of an evaluation can foster evaluative thinking, reasoning and decision making within schools, and focus attention on the aspects relevant to equity and inclusion, such as outcomes for diverse students, school conditions that promote equity, and relationships with diverse communities. Building a common language between evaluators and schools is not an easy task. One approach is to work with schools in a cyclical process of evaluations and conversations that emphasise schools' priorities and the specific needs of individual school communities (Box 6.12). This process fosters an environment where evaluators work closely with schools. Research shows that evaluations that were followed by schools' improvement often included a close co-operation between evaluators and schools (Ehren and Visscher, 2008[162]).

Box 6.12. School evaluation in New Zealand

In New Zealand, schools are expected to take part in an "ongoing, cyclical process of evaluation and inquiry for improvement" (Education Review Office, 2016, p. 6[163]). This process is intended to be participatory and collaborative, and includes schools' self-reviews and the specific needs of individual school communities.

The Educational Review Office uses two types of indicators: outcome and process indicators. Outcome indicators concern student achievement, progress, and related goals, but broaden the focus from learning outcomes to well-being, social and emotional learning, and resilience. For instance, one of the indicators asks whether students "feel included, cared for, and safe and secure" or whether Māori students "enjoy education success as Māori".

Process indicators are divided into six domains each of which contains indicators related to equity and/or inclusion. For instance, the second domain focuses on whether school leadership pursues not only excellence, but also equity (e.g., in allocation of resources or alignment of teachers' continuous professional learning and student learning needs). The third domain also explores whether schools involve diverse communities in the school life (with a particular focus on Māori communities), and whether students have tools that support learning at home. The fourth domain focuses on culturally responsive curriculum, effective teaching and opportunity to learn.

The process of evaluation starts with noticing (e.g., examination of the outcome indicators) and investigating (e.g., examination of the process indicators). The process then continues with

collaborative sense making through conversations on effective and meaningful practices to meet the school community's specific aspirations. The conversations continue in selecting the most appropriate actions to take. The evaluation concludes with monitoring and evaluating the impact. Outcome indicators are applied to see the effect and the process restarts.

The co-operative and investigative nature of the school evaluation process is illustrated by the fact that there are many more process indicators compared to outcome indicators. The indicators are provided to schools as a source for dialogue and planning rather than to impose policy or make summative judgements. The inquiry thus focuses on identifying which school conditions are contributing to poor performance, and, conversely, which school processes and activities are contributing to excellence.

Source: Education Review Office (2016[163]), School Evaluation Indicators, https://ero.govt.nz/how-ero-reviews/schoolskura-english-medium/school-evaluation-indicators (accessed 18 November 2022) and Gergen and Gill (2020[164]), *Beyond the Tyranny of Testing*.

Co-operation can foster inclusive environments

Co-operation in school evaluations can be fostered between researchers and schools. Co-operation with academia can foster the development of evaluation literacy among teachers, by, e.g., providing training and facilitating collaborative discussions (Nelson, Ehren and Godfrey, 2015[153]). Co-operation between "outsiders" (such as academics) and practitioners can also contribute to the development of better understandings of educational processes that transform the values of equity into practice (Ainscow, 2010[165]). External support can also produce honest feedback that is not biased by internal school structures (ibid.). However, a reorientation of values and goals might be required for a fruitful co-operation (Hiebert, Gallimore and Stigler, 2002[166]). Teachers, for instance, may need to be encouraged to be more open to sharing their successful practices across the board and to view teaching as an activity that can be continuously improved (ibid.). Teachers are in the unique position of having tested the programmes and can thus describe their experiences to researchers first-hand (Link Engineering Educator Exchange, 2017[167]). Furthermore, if teachers share ownership of a research, they might find it more relevant and applicable for their practice (OECD, 2022[119]). Teachers can also help researchers to refine hypotheses so that they are relevant, while researchers can ensure their suitability for research (ibid.). School leaders, in turn, are crucial in setting the vision of the evaluation process, managing collaborative discussions, allocating resources (including time) and developing a culture that is inquiry and improvement oriented (Nelson, Ehren and Godfrey, 2015[153]).

In England (United Kingdom), a co-operation initiative between schools and a university focused on ways to support students from disadvantaged socio-economic backgrounds. The programme used a collaborative inquiry to identify the issues schools are facing. Subsequently and in co-operation with university researchers, evidence on how schools' practices could be developed was collected. The research was then applied to stimulate new thinking, continuous professional learning, and to identify appropriate response strategies to diverse student needs. The initiative developed close links between schools and the academia in order to improve student learning opportunities and outcomes of students, and enabled schools to develop more inclusive practices tailored to the strengths and weaknesses of each school (Ainscow et al., 2016[168]).

In Murcia (Spain), researchers applied a co-operative and evaluative research methodology to facilitate a self-assessment process (Arnaiz, De Haro and Guirao, 2015[169]). The purpose of the study was to help schools identify the barriers to participation and learning of all students, and to evaluate how inclusive the processes in the schools were. The researchers helped with the co-ordination and facilitation of the discussions among school leaders, psychologists, counsellors, teachers and special education teachers (from the schools as well as nearby psycho-pedagogical centres). These discussions led to suggestions

on continuous professional learning for teachers, collaborative teaching practices and school self-evaluation methods.

"School to school" approaches that invest into capacity-building and continuous professional learning of teachers and school leaders by sharing knowledge and best practices have been promoted as strategies that can foster a productive evaluation process and the development of policies and practices to advance student outcomes, including in schools with large proportions of diverse students and in schools facing challenging circumstances (Ainscow, 2012[170]; Ehren et al., 2017[171]; OECD, 2015[120]; West, 2010[172]). These collaborative approaches can take the form of schools getting together to learn from each other and provide support in the process of self-evaluation (Nelson, Ehren and Godfrey, 2015[153]). A shared leadership is a crucial factor in developing these networks. Shared leadership challenges the notion that schools compete with each other, and instead builds trust and fosters collaborations that can improve inclusive practices in schools (Ainscow, 2010[165]). For instance, Ainscow et al. (2016[168]) set up an equity research network in which each school team gathered evidence about students identified as being "most vulnerable to underachievement, marginalisation or exclusion". Subsequently, the teams shared their findings with colleagues in the partner schools as well as researchers to deepen understandings of practices, beliefs, assumptions, and organisational processes (Ainscow et al., 2016[168]). As a result, school staff had a clearer focus and plan on what they wanted to achieve. Causal inferences are difficult to make, but teachers also felt that overall school performance increased as a result of this approach (ibid.).

An evaluation of an initiative that focused on collaboration of schools in Wales (United Kingdom) also showed improved results for all students, including socio-economically disadvantaged students (Box 6.13).

Box 6.13. Central South Wales Challenge

The Central South Wales (CSW) Challenge was an initiative in Wales (United Kingdom) launched in 2014. The government invested large amounts of additional resources in schools with particularly high proportions of disadvantaged students. The schools chose one of four strands of collaborative activities:

- **Pathfinders** focused partnerships on building partnerships between schools identified as being high performing and those identified as being low-performing. The intent was to accelerate improvement in schools facing particularly challenging circumstances.
- **Hubs** provided continuous professional learning activities for teachers and school leaders based on regional needs. After two years, the hubs also focused on bringing schools and practitioners into collaborative networks and groups.
- **School Improvement Groups** were designed to break down barriers among schools and among local authorities.
- **Peer enquiry** was developed by school leaders in high performing schools who felt that they lack support. Senior leaders supported one another in a process of mutual enquiry. Eventually, peer enquiry also included a wider range of lower performing schools.

While causal implications are difficult to make, the evaluation revealed that student scores in all five local authorities that were part of the CSW Challenge improved at the fastest rate of all Welsh regions. The two most underperforming local authorities improved to the greatest extent. Moreover, average progress among socio-economically disadvantaged students (those eligible for free school meals) also revealed encouraging trends by exceeding government targets.

Source: Hadfield and Ainscow (2018[173]), Inside a self-improving school system: Collaboration, competition and transition, 10.1007/s10833-018-9330-7.

Some education systems' external evaluators already focus specifically on whether schools foster co-operations. In Northern Ireland (United Kingdom), every post-primary school is a member of an area learning community (ALC). The aim of the ALC is for schools to come together (along with further education colleges and other training providers) and plan a broad and balanced area-relevant curriculum (Department of Education, n.d.[174]). Inspections also focus on whether the ALCs meet these goals. According to some inspectors, the ALCs improve collaboration and facilitate better transition between primary and lower secondary schools, among other positive impacts. Inspections have also shifted their focus from accountability to the use of self-evaluation and help to facilitate strategic planning and "joined-up thinking" (Ehren et al., 2017[171]).

In the Netherlands, mainstream and special schools are required to work together to provide inclusive education for all children (Ehren et al., 2017[171]). As such, several dozens of school networks have been formed and are governed by new education authorities to facilitate the provision of high-quality education to all students. The inspectorate evaluates not only individual schools, but also the networks. In particular, the inspections focus on how each school, in collaboration with other schools in the network, provides adequate support to all students and suggests areas for improvement (ibid.).

Pointers for policy development

This chapter has reviewed the available evidence and country approaches relating to the monitoring of progress towards equity and inclusion and to the evaluation of policies and practices that seek to promote equity and inclusion. The policy pointers that follow are drawn from analytical work, country reviews and the available research literature. It should be stressed that there is no single model or global best practice of a monitoring and evaluation framework. The development of practices always needs to consider country-specific contexts and features of the respective education systems. Consequently, the following pointers for future policy development will not be of equal relevance for different countries. In some systems, the policy pointers might already be implemented while in others, they might not apply due to different social, economic and educational structures and traditions.

Design monitoring systems that assess progress towards equity and inclusion

Education systems differ in how they collect data on diversity in schools. In some education systems, data are collected on a small set of dimensions of diversity. In others, a wide range of student characteristics is considered in policy making. In some education systems a limited set of data is available at the national (or sub-national) level, while a broader selection of student characteristics is available at the school level. Some education systems are moving away from collecting data on dimensions of diversity and, instead, focus on support measures.

Monitoring systems should be designed with the aim of informing stakeholders whether progress towards equity and inclusion is being achieved and, if not, what improvements can be made in this respect. Progress in improving equity and inclusion in education can be measured using a wide range of instruments (indicator frameworks, national assessments with background data, longitudinal/ key stakeholder/international surveys, or specific reviews on equity and inclusion). In combination, these have the potential to provide a complex picture on the state of progress in improving equity and inclusion education. As such, the education system should not be focused solely on learning outcomes of students. After all, academic outcomes are only one aspect of the overall school experience. It is equally important to understand how happy and satisfied students are with different aspects of their life, how connected they are to others, and whether they enjoy a good physical and mental health. This understanding can be developed through collecting data on a range of student well-being outcomes, including academic, psychological, physical, social and material.

At the same time, monitoring systems should not be entirely focused on student outcomes. In considering equity and inclusion, it is important to take into account inputs and processes that also create equitable and inclusive school environments. Inputs can consider resources, access to continuous professional learning and others. Processes can include school climate, teaching practices etc. A comprehensive data collection supports decision-makers in informing improvements across the whole system.

Consider monitoring equity and inclusion in education in a comprehensive strategy

The monitoring of progress in improving equity and inclusion in an education system can be key to a better understanding of the needs of students, teachers, school leaders and other stakeholders involved in creating an equitable and inclusive school environment. Without relevant information on the diversity within the system or the progress in improving equity and inclusion, policy makers might judge the system according to the often-imperfect data they have available. This might misdirect them or, in the case of absence of data, may mean that they are unaware of challenges that need action.

Instruments that are used to measure progress towards equity and inclusion can be embedded in a strategy with clearly stated long-term goals for the education system. This can help to align the views of various stakeholders. Short- and medium-term priorities, in turn, can help identify the intermediary steps. Progress of the education system to reaching these can be monitored using indicators that can contain specific target values that a policy or a set of policies is aimed to achieve. When developing these indicators, it is important to consider the purpose of each of them so that it is carefully designed with the goal to measure what is valued, rather than value what is measured. That way, policy makers and other stakeholders can limit the possibility that the measure itself becomes the target. In ensuring that the information obtained through monitoring is used to inform policy development, it is important to form co-operative relationships between different actors and within the public sector.

Leverage evaluations to identify policies, programmes and processes that best address the needs of students

Apart from monitoring, education systems should evaluate the effectiveness of implemented policies, programmes and processes. Evaluation is a necessary pre-requisite for informing evidence-based policy development, curriculum, planning, reporting, resource allocation decisions and performance management. In the context of equity and inclusion, evaluations help with the identification of policies, programmes and processes that address the needs of students most effectively. Subsequently, evaluation can be instructive in identifying initiatives that can be scaled up and potentially institutionalised. Despite many challenges, some education systems engaged in evaluations that have led to improvements in equity and inclusion.

The process of evaluation should be approached with a mind-set that addresses concerns regarding inclusive practices at all levels of the education system. This entails bringing diverse perspectives and experiences to the process, and ultimately changing the attitudes about diversity, equity and inclusion at all levels of society. Indeed, the process of evaluation needs to be based on a consultative and co-operative process with a wide range of stakeholders. Clarity of purpose and the use of available resources can help to streamline the process of evaluation.

Ensure school evaluations can be used for improvements in equity and inclusion

School external evaluation (e.g., by school inspectorates) and internal evaluation (self-evaluation) have the potential to identify barriers in the processes that hinder equitable or inclusive education for all students by analysing school climate, relationships, learning support, barriers to continuous professional learning and other aspects of the school environment. The effective monitoring and evaluation of schools, including

around the aspects of equity and inclusion, are central to their improvement. Schools need feedback on their practices to identify, and address, areas for improvement.

Several practices can improve equity and inclusion as a result of school evaluation. First, it is important to find a common language on what constitutes a good education, clarify the meaning of inclusion and reach a shared understanding that draws school staff around a common purpose. Second, co-operative approaches can foster better understandings of educational processes that transform the values of equity and inclusion into practice. They can lead to the exchange of good practices, peer reviews and other environments in which schools are motivated to learn from each other. These approaches are increasingly common not only among schools, but also between external evaluators and schools, as well as researchers and schools.

References

Ainscow, M. (2012), "Moving knowledge around: Strategies for fostering equity within educational systems", *Journal of Educational Change*, Vol. 13/3, pp. 289-310, https://doi.org/10.1007/s10833-012-9182-5. [170]

Ainscow, M. (2010), "Achieving excellence and equity: reflections on the development of practices in one local district over 10 years", *School Effectiveness and School Improvement*, Vol. 21/1, pp. 75-92, https://doi.org/10.1080/09243450903569759. [165]

Ainscow, M. (2005), "Developing inclusive education systems: what are the levers for change?", *Journal of Educational Change*, Vol. 6/2, pp. 109-124, https://doi.org/10.1007/s10833-005-1298-4. [42]

Ainscow, M. et al. (2016), "Using collaborative inquiry to foster equity within school systems: opportunities and barriers", *School Effectiveness and School Improvement*, Vol. 27/1, pp. 7-23, https://doi.org/10.1080/09243453.2014.939591. [168]

Ainscow, M. and K. Messiou (2017), "Engaging with the views of students to promote inclusion in education", *Journal of Educational Change*, Vol. 19/1, pp. 1-17, https://doi.org/10.1007/s10833-017-9312-1. [32]

Appels, L. et al. (2022), "Unpacking equity. Educational equity in secondary analyses of international large-scale assessments: A systematic review", *Educational Research Review*, p. 100494, https://doi.org/10.1016/j.edurev.2022.100494. [10]

Arnaiz, P., R. De Haro and J. Guirao (2015), "La evaluación en educación primaria como punto de partida para el desarrollo de planes de mejora inclusivos en la Región de Murcia", *Revista Electrónica Interuniversitaria de Formación del Profesorado*, Vol. 18/1, p. 103, https://doi.org/10.6018/reifop.18.1.214351. [169]

AuCoin, A., G. Porter and K. Baker-Korotkov (2020), "New Brunswick's journey to inclusive education", *PROSPECTS*, Vol. 49/3-4, pp. 313-328, https://doi.org/10.1007/s11125-020-09508-8. [124]

Australian Institute of Family Studies (n.d.), *Growing Up in Australia: The Longitudinal Study of Australian Children - Data Dictionary*, https://growingupinaustralia.gov.au/sites/default/files/release_9c1_data_dictionary.pdf (accessed on 22 February 2022). [74]

Azorín, C. and M. Ainscow (2018), "Guiding schools on their journey towards inclusion", *International Journal of Inclusive Education*, Vol. 24/1, pp. 58-76, https://doi.org/10.1080/13603116.2018.1450900. [161]

Balestra, C. and L. Fleischer (2018), "Diversity statistics in the OECD: How do OECD countries collect data on ethnic, racial and indigenous identity?", *OECD Statistics Working Papers*, No. 2018/09, OECD Publishing, Paris, https://doi.org/10.1787/89bae654-en. [29]

Bamber, V. and S. Anderson (2012), "Evaluating learning and teaching: institutional needs and individual practices", *International Journal for Academic Development*, Vol. 17/1, pp. 5-18, https://doi.org/10.1080/1360144x.2011.586459. [123]

Booth, T. and M. Ainscow (2002), *The Index for inclusion: developing learning and participation in schools*, Centre for Studies on Inclusive Education, https://www.eenet.org.uk/resources/docs/Index%20English.pdf (accessed on 11 April 2022). [155]

Burgess, S. et al. (2005), *Who Wins and Who Loses from School Accountability? The Distribution of Educational Gain in English Secondary Schools*, https://ssrn.com/abstract=837284 (accessed on 18 November 2022). [64]

Carlana, M. and E. La Ferrara (2022), *Improving Educational Outcomes through Online Tutoring during Schools Closures in Italy during the Covid-19 Outbreak*, https://www.povertyactionlab.org/evaluation/improving-educational-outcomes-through-online-tutoring-during-schools-closures-italy (accessed on 19 May 2022). [143]

Central Examination Commission (2019), *Osiągnięcia uczniów kończących gimnazjum w roku 2019 [Achievement of Students Graduating from Upper Secondary Schools in 2019]*, Central Examination Commission, https://cke.gov.pl/images/_KOMUNIKATY/Sprawozdanie_egzamin%20gimnazjalny%202019.pdf.pdf (accessed on 8 November 2022). [55]

Centre for Longitudinal Studies (2022), *Next Steps*, https://nextstepsstudy.org.uk/ (accessed on 22 February 2022). [72]

Centre for Longitudinal Studies (n.d.), *Millennium Cohort Study*, https://cls.ucl.ac.uk/cls-studies/millennium-cohort-study/ (accessed on 22 February 2022). [71]

Centre for Studies on Inclusive Education (2020), *Index for Inclusion: developing learning and participation in schools*, http://www.csie.org.uk/resources/inclusion-index-explained.shtml (accessed on 12 December 2022). [154]

Cerna, L. et al. (2021), "Promoting inclusive education for diverse societies: A conceptual framework", *OECD Education Working Papers*, No. 260, OECD Publishing, Paris, https://doi.org/10.1787/94ab68c6-en. [2]

CLOSER (n.d.), *Longitudinal Survey of Young People in England: Cohort 2*, https://www.closer.ac.uk/study/lsype-2/ (accessed on 22 February 2022). [73]

Clotfelter, C. et al. (2004), "Do school accountability systems make it more difficult for low-performing schools to attract and retain high-quality teachers?", *Journal of Policy Analysis and Management*, Vol. 23/2, pp. 251-271, https://doi.org/10.1002/pam.20003. [69]

Coburn, C. (2003), "Rethinking Scale: Moving Beyond Numbers to Deep and Lasting Change", *Educational Researcher*, Vol. 32/6, pp. 3-12, https://doi.org/10.3102/0013189x032006003. [136]

Commonwealth of Australia (2020), *Closing the Gap Report 2020*, https://ctgreport.niaa.gov.au/sites/default/files/pdf/closing-the-gap-report-2020.pdf (accessed on 28 April 2022). [51]

Cook, T. (2002), "Randomized Experiments in Educational Policy Research: A Critical Examination of the Reasons the Educational Evaluation Community has Offered for not Doing Them", *Educational Evaluation and Policy Analysis*, Vol. 24/3, pp. 175-199, https://doi.org/10.3102/01623737024003175. [121]

Council of Ministers of Education Canada (2019), *PCAP 2019*, https://www.cmec.ca/697/PCAP_2019.html (accessed on 8 November 2022). [61]

Cross-Sectoral Coordination Center (2020), *National Development Plan of Latvia for 2021-2027*, Cross-Sectoral Coordination Center, https://pkc.gov.lv/sites/default/files/inline-files/NAP2027__ENG.pdf (accessed on 7 November 2022). [49]

Csata, Z., R. Hlatky and A. Liu (2020), "How to head count ethnic minorities: validity of census surveys versus other identification strategies", *East European Politics*, Vol. 37/3, pp. 572-592, https://doi.org/10.1080/21599165.2020.1843439. [90]

Currie, C. et al. (eds.) (2012), *Social determinants of health and well-being among young people. Health Behaviour in School-aged Children (HBSC) study: international report from the 2009/2010 survey*, WHO Regional Office for Europe, https://www.euro.who.int/__data/assets/pdf_file/0003/163857/Social-determinants-of-health-and-well-being-among-young-people.pdf (accessed on 23 May 2022). [38]

Czech School Inspectorate (2022), *České školství v datech [Czech Education in Data]*, https://www.vzdelavanivdatech.cz/ (accessed on 8 November 2022). [109]

Czech School Inspectorate (2018), *Vybrané aspekty implementace společného vzdělávání v období 1. pololetí školního roku 2017/2018 [Selected aspects of the implementation of common education in the first half of 2017/2018]*, Czech School Inspectorate, https://www.csicr.cz/CSICR/media/Prilohy/2018_p%c5%99%c3%adlohy/Dokumenty/TZ-Vybrane-aspekty-implementace-spolecneho-vzdelavani-v-1-pololeti-2017-2018.pdf (accessed on 20 May 2022). [129]

Czech School Inspectorate (2017), *Společné vzdělávání ve školním roce 2016/2017 [Common education in 2016/2017 school year]*, Czech School Inspectorate, https://www.csicr.cz/getattachment/7734c437-a133-4411-b8b6-ed11776ad4fe/TZ-Spolecne-vzdelavani-16-10-2017.pdf (accessed on 20 May 2022). [128]

Danish Evaluation Institute (2020), *Samarbejde om inkluderende læringsmiljøer [Collaboration on inclusive learning environments]*, Danish Evaluation Institute, https://www.eva.dk/grundskole/samarbejde-om-inkluderende-laeringsmiljoeer (accessed on 18 November 2022). [103]

Danish Evaluation Institute (2020), *Undervisning af ordblinde elever i folkeskolen [Teaching dyslexic pupils in primary school]*, Danish Evaluation Institute, https://www.eva.dk/grundskole/undervisning-ordblinde-elever-folkeskolen (accessed on 18 November 2022). [105]

Danish Evaluation Institute (2019), *Direkte indskrivning af nyankomne elever i almenklasser [Direct enrollment of newly arrived students in general classes]*, Danish Evaluation Institute, https://www.eva.dk/grundskole/direkte-indskrivning-nyankomne-elever-almenklasser (accessed on 18 November 2022). [104]

Darling-Hammond, L. and J. Snyder (2015), "Meaningful Learning in a New Paradigm for Educational Accountability: An Introduction", *Education Policy Analysis Archives*, Vol. 23, p. 7, https://doi.org/10.14507/epaa.v23.1982. [70]

Davis, T., R. Bhatt and K. Schwarz (2015), "School Segregation in the Era of Accountability", *Social Currents*, Vol. 2/3, pp. 239-259, https://doi.org/10.1177/2329496515589852. [66]

de Hoop, T. et al. (2019), *Scaling Education Innovations in Complex Emergencies*, American Institutes for Research, https://www.unhcr.org/hea/wp-content/uploads/sites/125/2021/02/HEA_MetaEvaluation_Final-2.pdf (accessed on 3 June 2022). [140]

Department for Education (2022), *Key stage 4 performance: academic year 2020/21*, https://www.ethnicity-facts-figures.service.gov.uk/education-skills-and-training/11-to-16-years-old/gcse-results-attainment-8-for-children-aged-14-to-16-key-stage-4/latest (accessed on 8 November 2022). [57]

Department for Education (2022), *Occupations paid by national pay scales*, https://www.gov.uk/government/publications/national-pay-scales-for-eligible-teaching-and-education-jobs/national-pay-scales-for-eligible-teaching-and-education-leadership-occupation-codes (accessed on 10 July 2022). [9]

Department of Education (2009), *Every School A Good School*, Department of Education, https://www.education-ni.gov.uk/sites/default/files/publications/de/ESAGS%20Policy%20for%20School%20Improvement%20-%20Final%20Version%2005-05-2009.pdf (accessed on 23 May 2022). [156]

Department of Education (n.d.), *Entitlement Framework*, https://www.education-ni.gov.uk/articles/entitlement-framework (accessed on 23 May 2022). [174]

Department of Education and Skills (2016), *School self-evaluation guidelines 2016-2020*, Department of Education and Skills, https://assets.gov.ie/25262/f4a6f2a21e1c4c26a55234511085d5a3.pdf (accessed on 18 November 2022). [158]

Department of the Prime Minister and Cabinet (2020), *Child and youth wellbeing strategy*, https://dpmc.govt.nz/our-programmes/child-and-youth-wellbeing-strategy#:~:text=The%20Child%20and%20Youth%20Wellbeing,and%20how%20others%20can%20help (accessed on 31 May 2022). [45]

Deutskens, E. et al. (2004), "Response Rate and Response Quality of Internet-Based Surveys: An Experimental Study", *Marketing Letters*, Vol. 15/1, pp. 21-36, https://doi.org/10.1023/b:mark.0000021968.86465.00. [93]

Directorate of Education (2022), *Ståstedsanalysen for skole [Situational analysis for schools]*, https://www.udir.no/kvalitet-og-kompetanse/stastedsanalyse/om-stastedaanalysen-for-skoler/ (accessed on 8 November 2022). [157]

Dutch Inspectorate of Education (2022), *The State of Education 2022*, Dutch Inspectorate of Education, https://english.onderwijsinspectie.nl/documents/annual-reports/2022/04/28/state-of-education-2022 (accessed on 18 November 2022). [107]

Ebersold, S. et al. (eds.) (2020), *Changing Role of Specialist Provision in Supporting Inclusive Education: Mapping Specialist Provision Approaches in European Countries*, European Agency for Special Needs and Inclusive Education. [33]

Education Review Office (2016), *School Evaluation Indicators*, Education Review Office, https://ero.govt.nz/how-ero-reviews/schoolskura-english-medium/school-evaluation-indicators (accessed on 18 November 2022). [163]

Education Scotland (2015), *How good is our school?*, Education Scotland, https://education.gov.scot/improvement/Documents/Frameworks_SelfEvaluation/FRWK2_NIHeditHGIOS/FRWK2_HGIOS4.pdf (accessed on 11 August 2022). [147]

Ehren, M. et al. (2013), "Impact of school inspections on improvement of schools—describing assumptions on causal mechanisms in six European countries", *Educational Assessment, Evaluation and Accountability*, Vol. 25/1, pp. 3-43, https://doi.org/10.1007/s11092-012-9156-4. [160]

Ehren, M. et al. (2017), "Evaluation and decentralised governance: Examples of inspections in polycentric education systems", *Journal of Educational Change*, Vol. 18/3, pp. 365-383, https://doi.org/10.1007/s10833-017-9297-9. [171]

Ehren, M., J. Perryman and N. Shackleton (2014), "Setting expectations for good education: how Dutch school inspections drive improvement", *School Effectiveness and School Improvement*, Vol. 26/2, pp. 296-327, https://doi.org/10.1080/09243453.2014.936472. [145]

Ehren, M. and A. Visscher (2008), "The relationship between school inspections, school characteristics and school improvement", *British Journal of Educational Studies*, Vol. 56/2, pp. 205-227, https://doi.org/10.1111/j.1467-8527.2008.00400.x. [162]

Else-Quest, N. and J. Hyde (2016), "Intersectionality in Quantitative Psychological Research", *Psychology of Women Quarterly*, Vol. 40/3, pp. 319-336, https://doi.org/10.1177/0361684316647953. [79]

Ely, D. (1999), *New Perspectives on the Implementation of Educational Technology Innovations*, https://files.eric.ed.gov/fulltext/ED427775.pdf (accessed on 3 June 2022). [138]

European Agency for Special Needs and Inclusive Education (2022), *Desenho de um sistema de monitorização da implementação do regime jurídico da educação inclusiva em Portugal [Design a System to Monitor the Implementation of the Law on Inclusive Education in Portugal]*, European Agency for Special Needs and Inclusive Education, https://www.dge.mec.pt/sites/default/files/Noticias_documentos/sistema_de_monitorizacao_para_a_educacao_inclusiva_em_portugal.pdf (accessed on 18 November 2022). [34]

European Agency for Special Needs and Inclusive Education (2022), *Home*, https://www.european-agency.org/ (accessed on 23 May 2022). [99]

European Commission/EACEA/Eurydice (2015), *Assuring Quality in Education: Policies and Approaches to School Evaluation in Europe*, Publications Office of the European Union, https://data.europa.eu/doi/10.2797/678 (accessed on 23 May 2022). [151]

European Education and Culture Executive Agency, Eurydice (2011), *National testing of pupils in Europe : objectives, organisation and use of results*, European Commission, https://data.europa.eu/doi/10.2797/18294 (accessed on 8 November 2022). [60]

European Institute for Gender Equality (2016), *Gender Impact Assessment: Gender Mainstreaming Toolkit*, Publications Office of the European Union, https://doi.org/10.2839/907321 (accessed on 10 May 2022). [130]

European Network Against Racism (2022), *Antigypsyism*, https://www.enar-eu.org/about/antigypsyism/ (accessed on 18 November 2022). [178]

European Social Fund Learning Network (2018), *How to monitor and evaluate Roma-related initiatives under Structural and Investment Funds*, https://www.euromanet.eu/wp-content/uploads/2018/02/ESF-Roma-Inclusion-LN_Handbook-ME.pdf (accessed on 31 May 2022). [96]

European Union Agency for Fundamental Rights (2018), *Second European Union minorities and discrimination survey: Roma: selected findings*, Publications Office, https://data.europa.eu/doi/10.2811/469 (accessed on 23 May 2022). [101]

European Union and the United Nations (2018), *Expert Group on Refugee and Internally Displaced Persons Statistics*, Publications Office of the European Union, https://unstats.un.org/unsd/demographic-social/Standards-and-Methods/files/Principles_and_Recommendations/International-Migration/2018_1746_EN_08-E.pdf (accessed on 23 May 2022). [85]

EUROSTAT (2019), *Glossary:Early leaver from education and training*, https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Glossary:Early_leaver_from_education_and_training (accessed on 16 June 2022). [179]

Examination Information System (2022), *Testide tulemuste statistika [Statistics of Test Results]*, https://eis.ekk.edu.ee/eis/eksamistatistika (accessed on 8 November 2022). [58]

Federal Minister for Women and Civil Service at the Federal Chancellery (2012), *Regulatory Impact Assessment and the Impact Dimension of Gender Equality*, Federal Minister for Women and Civil Service at the Federal Chancellery, http://regulatoryreform.com/wp-content/uploads/2015/02/Austria-Regulatory-Impact-Assessment-and-the-Impact-Dimension-of-Gender-Equality-2012.pdf (accessed on 10 May 2022). [131]

Federal Ministry of Education, Science and Research (2022), *Der Qualitätsrahmen für Schulen [The Quality Framework for Schools]*, https://www.qms.at/qualitaetsrahmen/der-qualitaetsrahmen-fuer-schulen (accessed on 18 November 2022). [150]

Federal Ministry of Education, Science and Research (2022), *Schulaufsicht - Schulqualitätsmanager/innen [School supervision - school quality managers]*, https://www.bmbwf.gv.at/Themen/schule/bef/qum/schulaufsicht.html (accessed on 18 November 2022). [149]

Federal Office for Migration and Refugees (2022), *The IAB-BAMF-SOEP Survey of Refugees*, https://www.bamf.de/SharedDocs/ProjekteReportagen/EN/Forschung/Integration/iab-bamf-soep-befragung-gefluechtete.html?nn=447028 (accessed on 4 28 2022). [83]

Fresno, J. et al. (2019), *A meta-evaluation of interventions for Roma inclusion*, Publications Office of the European Union, https://doi.org/10.2760/641471 (accessed on 23 May 2022). [111]

Gergen, K. and S. Gill (2020), "Relational Approaches to School Evaluation", in *Beyond the Tyranny of Testing*, Oxford University Press, https://doi.org/10.1093/oso/9780190872762.003.0007. [164]

Glass, T. (2013), *Creating learning environments for disengaged boys: Bridging the gender gap with universal design for learning*, https://mspace.lib.umanitoba.ca/jspui_org/bitstream/1993/17596/1/Thesis%20March%202013.pdf. [133]

Glennan, T. et al. (2004), *Summary: Toward a More Systematic Approach to Expanding the Reach of Educational Interventions*, RAND Corporation, https://www.jstor.org/stable/10.7249/mg248ff.25. [142]

Godfrey, D. (2020), "From External Evaluation, to School Self-evaluation, to Peer Review", in *School Peer Review for Educational Improvement and Accountability, Accountability and Educational Improvement*, Springer International Publishing, Cham, https://doi.org/10.1007/978-3-030-48130-8_1. [152]

Golden, G. (2020), "Education policy evaluation: Surveying the OECD landscape", *OECD Education Working Papers*, No. 236, OECD Publishing, Paris, https://doi.org/10.1787/9f127490-en. [118]

Gouëdard, P. (2021), "Developing indicators to support the implementation of education policies", *OECD Education Working Papers*, No. 255, OECD Publishing, Paris, https://doi.org/10.1787/b9f04dd0-en. [44]

Government of Canada (2022), *Gender-based Analysis Plus (GBA+)*, https://women-gender-equality.canada.ca/en/gender-based-analysis-plus.html (accessed on 7 June 2022). [126]

Gutman, L. and J. Vorhaus (2012), *The Impact of Pupil Behaviour and Wellbeing on Educational Outcomes*, Department for Education, https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/219638/DFE-RR253.pdf (accessed on 23 May 2022). [37]

Hadfield, M. and M. Ainscow (2018), "Inside a self-improving school system: Collaboration, competition and transition", *Journal of Educational Change*, Vol. 19/4, pp. 441-462, https://doi.org/10.1007/s10833-018-9330-7. [173]

Hankivsky, O. and L. Mussell (2018), "Gender-Based Analysis Plus in Canada: Problems and Possibilities of Integrating Intersectionality", *Canadian Public Policy*, Vol. 44/4, pp. 303-316, https://doi.org/10.3138/cpp.2017-058. [127]

Hanushek, E. and M. Raymond (2005), "Does school accountability lead to improved student performance?", *Journal of Policy Analysis and Management*, Vol. 24/2, pp. 297-327, https://doi.org/10.1002/pam.20091. [63]

Hart, S., M. Drummond and D. McIntyre (2007), "Learning without Limits: Constructing a Pedagogy Free from Determinist Beliefs about Ability", in *The SAGE Handbook of Special Education*, SAGE Publications, https://doi.org/10.4135/9781848607989.n39. [30]

Hellebrandt, T. et al. (2020), *Groups at Risk of Poverty or Social Exclusion Spending Review: Final Report*, Ministry of Finance of the Slovak Republic, https://www.mfsr.sk/files/archiv/56/ReviziavydavkovnaohrozeneskupinyZSverziaFINALENG.pdf (accessed on 14 April 2022). [26]

Hiebert, J., R. Gallimore and J. Stigler (2002), "A Knowledge Base for the Teaching Profession: What Would It Look Like and How Can We Get One?", *Educational Researcher*, Vol. 31/5, pp. 3-15, https://doi.org/10.3102/0013189x031005003. [166]

Hsieh, C. et al. (2013), *The Allocation of Talent and U.S. Economic Growth*, National Bureau of Economic Research, Cambridge, MA, https://doi.org/10.3386/w18693. [11]

IGLYO (2022), *LGBTQI inclusive education report*, IGLYO — The International Lesbian, Gay, Bisexual, Transgender, Queer and Intersex Youth & Student Organisation, https://www.education-index.org/wp-content/uploads/2022/05/IGLYO-LGBTQI-Inclusive-Education-Report-2022-v3.pdf (accessed on 23 May 2022). [39]

Inspectorate of Education (2022), *State of Education*, https://english.onderwijsinspectie.nl/documents/annual-reports/2022/04/28/state-of-education-2022 (accessed on 18 November 2022). [25]

INVALSI (2022), *Dati [Data]*, https://invalsi-serviziostatistico.cineca.it/ (accessed on 8 November 2022). [56]

Jahnukainen, M. and T. Itkonen (2010), "Disability or Learning Difficulty? Politicians or Educators? Constructing Special Education in Finland and the United States", *Comparative Sociology*, Vol. 9/2, pp. 182-201, https://doi.org/10.1163/156913210x12536181351033. [35]

Jamieson, L. et al. (2021), "Challenges in identifying indigenous peoples in population oral health surveys: a commentary", *BMC Oral Health*, Vol. 21/1, https://doi.org/10.1186/s12903-021-01455-w. [91]

Jennings, J. and D. Rentner (2006), "Ten Big Effects of the No Child Left behind Act on Public Schools", *Phi Delta Kappan*, Vol. 88/2, pp. 110-113, https://doi.org/10.1177/003172170608800206. [68]

Jensen, E. and T. Kennel (2022), *Who Was Undercounted, Overcounted in the 2020 Census?*, https://www.census.gov/library/stories/2022/03/who-was-undercounted-overcounted-in-2020-census.html (accessed on 18 November 2022). [89]

Koehler, C., N. Palaiologou and O. Brussino (2022), "Holistic refugee and newcomer education in Europe : Mapping, upscaling and institutionalising promising practices from Germany, Greece and the Netherlands", *OECD Education Working Papers*, No. 264, OECD Publishing, Paris, https://doi.org/10.1787/9ea58c54-en. [139]

Kosciw, J. et al. (2020), *The 2019 National School Climate Survey: The experiences of lesbian, gay, bisexual, transgender, and queer youth in our nation's schools*, GLSEN, https://www.glsen.org/sites/default/files/2021-04/NSCS19-FullReport-032421-Web_0.pdf (accessed on 22 February 2022). [88]

Link Engineering Educator Exchange (2017), *Teachers as policy makers*, https://www.linkengineering.org/Explore/LE_Blog/45832.aspx (accessed on 31 January 2022). [167]

Marx, D. and J. Roman (2002), "Female role models: Protecting women's math test performance", *Personality and Social Psychology Bulletin*, Vol. 28/9, pp. 1183-1193, https://doi.org/10.1177/01461672022812004. [134]

McBrien, J., A. Rutigliano and A. Sticca (2022), "The Inclusion of LGBTQI+ students across education systems: An overview", *OECD Education Working Papers*, No. 273, OECD Publishing, Paris, https://doi.org/10.1787/91775206-en. [17]

Mezzanotte, C. (2020), "Policy approaches and practices for the inclusion of students with attention-deficit hyperactivity disorder (ADHD)", *OECD Education Working Papers*, No. 238, OECD Publishing, Paris, https://doi.org/10.1787/49af95e0-en. [19]

Mezzanotte, C. and C. Calvel (Forthcoming), *Indicators of inclusion in education: a framework for analysis*, OECD. [13]

Ministry for Schools and Education of the North Rhine-Westphalia (2022), *Qualitätstableau NRW Kompaktversion [Quality tables NRW compact version]*, https://www.schulministerium.nrw/dokument/qualitaetstableau-nrw-kompaktversion (accessed on 18 November 2022). [146]

Ministry of Education (2022), *An Introduction to the new Equity Funding system for schools and kura*, Ministry of Education, https://assets.education.govt.nz/public/Documents/our-work/changes-in-education/An-Introduction-to-the-new-Equity-Funding-system-for-schools-and-kura.pdf (accessed on 20 May 2022). [77]

Ministry of Education (2022), *Identificador Provisorio Escolar (IPE) e Identificador Provisorio del Apoderado (IPA) para personas extranjeras [Provisional School Identifier (IPE) and Provisional Representative Identifier (IPA) for foreigners]*, https://www.chileatiende.gob.cl/fichas/49443-identificador-provisorio-escolar-ipe-e-identificador-provisorio-del-apoderado-ipa-para-personas-extranjeras (accessed on 18 November 2022). [78]

Ministry of Education, Culture, Sports, Science and Technology (2018), *The Third Basic Plan for the Promotion of Education (Provisional Translation)*, https://www.mext.go.jp/en/policy/education/lawandplan/title01/detail01/1373799.html (accessed on 5 July 2022). [47]

Ministry of Interior of the Slovak Republic (2022), *Atlas rómskych komunít 2019 [Atlas of the Roma Communities 2019]*, https://www.minv.sk/?atlas-romskych-komunit-2019 (accessed on 31 May 2022). [97]

Ministry of National Education and Youth (2022), *Conseil d'évaluation de l'École [The French Council for School Evaluation]*, https://www.education.gouv.fr/conseil-d-evaluation-de-l-ecole-305080 (accessed on 23 May 2022). [117]

Morozumi, A. and R. Tanaka (2020), *Should School-Level Results of National Assessments Be Made Public?*, https://www.iza.org/publications/dp/13450/should-school-level-results-of-national-assessments-be-made-public (accessed on 18 November 2022). [65]

National Education Agency (2022), *Lietuva. Švietimas šalyje ir regionuose [Lithuania. Education in the country and regions 2022]*, National Education Agency, https://www.nsa.smm.lt/wp-content/uploads/2022/08/Svietimas-Lietuvoje-2022-web.pdf (accessed on 18 November 2022). [106]

Neal, D. and D. Schanzenbach (2010), "Left Behind by Design: Proficiency Counts and Test-Based Accountability", *Review of Economics and Statistics*, Vol. 92/2, pp. 263-283, https://doi.org/10.1162/rest.2010.12318. [67]

Nelson, R., M. Ehren and D. Godfrey (2015), *Literature Review on Internal Evaluation*, http://schoolinspections.eu/wp-content/uploads/downloads/2015/09/Literature-review-internal-evaluation.pdf (accessed on 8 November 2022). [153]

New Brunswick Health Council (2022), *New Brunswick Student Wellness Survey*, https://nbhc.ca/surveys/new-brunswick-student-wellness-survey (accessed on 13 April 2022). [82]

NUCEM (2022), *Testovanie 9 2022 [Testing 9 2022]*, NUCEM, https://www2.nucem.sk/dl/5302/Vysledky%20Testovania%209%202022%20-%20prezentacia.pdf (accessed on 9 November 2022). [59]

Nworie, J. (2015), "Institutionalization of Teaching and Learning Gains in Higher Education", *Educational Technology*, Vol. 55/5, pp. 21-28, https://www.jstor.org/stable/44430404 (accessed on 7 December 2022). [137]

NZQA (2022), *National Student Number (NSN)*, https://www.nzqa.govt.nz/login/national-student-number-nsn/ (accessed on 8 November 2022). [76]

OECD (2022), *Dashboard of indicators on equity in and through education, OECD document for official use, EDU/EDPC(2022)20/ANN*. [43]

OECD (2022), *Review of Inclusive Education in Portugal*, Reviews of National Policies for Education, OECD Publishing, Paris, https://doi.org/10.1787/a9c95902-en. [21]

OECD (2022), *Strength through Diversity Policy Survey*, OECD. [16]

OECD (2022), *Who Cares about Using Education Research in Policy and Practice?: Strengthening Research Engagement*, Educational Research and Innovation, OECD Publishing, Paris, https://doi.org/10.1787/d7ff793d-en. [119]

OECD (2021), *Adapting Curriculum to Bridge Equity Gaps: Towards an Inclusive Curriculum*, OECD Publishing, Paris, https://doi.org/10.1787/6b49e118-en. [50]

OECD (2021), *Beyond Academic Learning: First Results from the Survey of Social and Emotional Skills*, OECD Publishing, Paris, https://doi.org/10.1787/92a11084-en. [81]

OECD (2021), *Education at a Glance 2021: OECD Indicators*, OECD Publishing, Paris, https://doi.org/10.1787/b35a14e5-en. [7]

OECD (2020), *PISA 2018 Results (Volume V): Effective Policies, Successful Schools*, PISA, OECD Publishing, Paris, https://doi.org/10.1787/ca768d40-en. [8]

OECD (2019), *Education Policy Outlook 2019: Working Together to Help Students Achieve their Potential*, OECD Publishing, Paris, https://doi.org/10.1787/2b8ad56e-en. [46]

OECD (2019), *PISA 2018 Results (Volume I): What Students Know and Can Do*, PISA, OECD Publishing, Paris, https://doi.org/10.1787/5f07c754-en. [41]

OECD (2019), *PISA 2018 Results (Volume II): Where All Students Can Succeed*, PISA, OECD Publishing, Paris, https://doi.org/10.1787/b5fd1b8f-en. [5]

OECD (2019), *Society at a Glance 2019: OECD Social Indicators*, OECD Publishing, https://doi.org/10.1787/soc_glance-2019-en. [87]

OECD (2018), *Education at a Glance 2018: OECD Indicators*, OECD Publishing, Paris, https://doi.org/10.1787/eag-2018-en. [14]

OECD (2018), *Equity in Education: Breaking Down Barriers to Social Mobility*, PISA, OECD Publishing, Paris, https://doi.org/10.1787/9789264073234-en. [12]

OECD (2017), *Educational Opportunity for All: Overcoming Inequality throughout the Life Course*, Educational Research and Innovation, OECD Publishing, Paris, https://doi.org/10.1787/9789264287457-en. [3]

OECD (2017), *Promising Practices in Supporting Success for Indigenous Students*, OECD Publishing, Paris, https://doi.org/10.1787/9789264279421-en. [95]

OECD (2015), *Education at a Glance 2015: OECD Indicators*, OECD Publishing, Paris, https://doi.org/10.1787/eag-2015-en. [52]

OECD (2015), *Education Policy Outlook 2015: Making Reforms Happen*, OECD Publishing, Paris, https://doi.org/10.1787/9789264225442-en. [120]

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://doi.org/10.1787/9789264190658-en. [1]

OECD (2012), *Equity and Quality in Education: Supporting Disadvantaged Students and Schools*, OECD Publishing, Paris, https://doi.org/10.1787/9789264130852-en. [40]

OECD (n.d.), *PISA 2018 national questionnaires*, https://www.oecd.org/pisa/publications/pisa-2018-national-questionnaires.htm (accessed on 20 April 2022). [98]

Öhberg, P. and M. Medeiros (2017), "A sensitive question? The effect of an ethnic background question in surveys", *Ethnicities*, Vol. 19/2, pp. 370-389, https://doi.org/10.1177/1468796817740379. [23]

Oliver, K. and P. Cairney (2019), "The dos and don'ts of influencing policy: A systematic review of advice to academics", *Palgrave Communications*, Vol. 5/21, https://doi.org/10.1057/s41599-019-0232-y. [112]

Osterholm, K., W. Nash and W. Kritsonis (2007), "Effects of Labeling Students "Learning Disabled": Emergent Themes in the Research Literature 1970 Through", *Focus on Colleges, Universities, and Schools*, Vol. 1/1, pp. 3-9, http://www.nationalforum.com/Electronic%20Journal%20Volumes/Osterholm,%20Karen%20Effects%20of%20Labeling%20Students%20Learning%20Disabled.pdf (accessed on 7 December 2022). [31]

PAQ Research (2022), *Mapa vzdělávacího ne/úspěchu*, https://www.mapavzdelavani.cz/ (accessed on 8 November 2022). [110]

Parekh, G., R. Brown and K. Robson (2018), "The Social Construction of Giftedness", *Canadian Journal of Disability Studies*, Vol. 7/2, pp. 1-32, https://doi.org/10.15353/cjds.v7i2.421. [122]

Pollard, E. and P. Lee (2003), "Child Well-being: A Systematic Review of the Literature", *Social Indicators Research*, Vol. 61/1, pp. 59-78, https://doi.org/10.1023/a:1021284215801. [36]

Research Schools Network (2022), *Supporting pupils with SEND during testing times*, https://researchschool.org.uk/news/supporting-pupils-with-send-during-testing-times (accessed on 18 November 2022). [115]

Research Schools Network (2020), *Building a successful foundation for reintegration*, https://researchschool.org.uk/news/building-a-successful-foundation-for-reintegration (accessed on 18 November 2022). [113]

Research Schools Network (2020), *Sustaining the learning of disadvantaged pupils during school lockdown*, https://researchschool.org.uk/news/sustaining-the-learning-of-disadvantaged-pupils-during-school-lockdown (accessed on 18 November 2022). [114]

Rutigliano, A. (2020), "Inclusion of Roma students in Europe: A literature review and examples of policy initiatives", *OECD Education Working Papers*, No. 228, OECD Publishing, Paris, https://doi.org/10.1787/8ce7d6eb-en. [100]

Rutigliano, A. and N. Quarshie (2021), "Policy approaches and initiatives for the inclusion of gifted students in OECD countries", *OECD Education Working Papers*, No. 262, OECD Publishing, Paris, https://doi.org/10.1787/c3f9ed87-en. [18]

Scottish Government (2022), *Using intersectionality to understand structural inequality in Scotland: evidence synthesis*, Scottish Government, https://www.gov.scot/publications/using-intersectionality-understand-structural-inequality-scotland-evidence-synthesis/ (accessed on 4 October 2022). [80]

Scottish Government (2021), *Achieving Excellence and Equity: 2022 National Improvement Framework and Improvement Plan*, Scottish Government, https://www.gov.scot/publications/achieving-excellence-equity-2022-national-improvement-framework-improvement-plan/documents/ (accessed on 7 November 2022). [48]

Scottish Government (n.d.), *Getting it right for every child (GIRFEC)*, https://www.gov.scot/policies/girfec/ (accessed on 18 November 2022). [148]

Seelos, C. and J. Mair (2016), "When Innovation Goes Wrong", *Stanford Social Innovation Review*, Vol. 14/4, pp. 27-33, https://doi.org/10.48558/wtsa-3b80. [141]

Siarova, H. and L. van der Graaf (2022), "Multi-stakeholder approach for better integration of refugee students: Stakeholder engagement in the practice-research-policy transfer in refugee education policy", *OECD Education Working Papers*, No. 265, OECD Publishing, Paris, https://doi.org/10.1787/82b390fb-en. [84]

Simon, P. (2017), "The failure of the importation of ethno-racial statistics in Europe: debates and controversies", *Ethnic and Racial Studies*, Vol. 40/13, pp. 2326-2332, https://doi.org/10.1080/01419870.2017.1344278. [24]

Smith, S. and L. Bost (2007), *Strategies for Increasing Response Rates*, National Post-School Outcomes Center, https://files.eric.ed.gov/fulltext/ED575738.pdf (accessed on 12 July 2022). [92]

Statistics Canada (2020), *Updated content for the 2021 Census of Population: Family and demographic concepts, and activities of daily living*, https://www12.statcan.gc.ca/census-recensement/2021/ref/98-20-0001/982000012020001-eng.cfm (accessed on 12 April 2022). [175]

Statistics Norway (2017), *Differences in the value added of primary and lower secondary schools in Norway*, https://www.ssb.no/en/utdanning/artikler-og-publikasjoner/differences-in-the-value-added-of-primary-and-lower-secondary-schools-in-norway (accessed on 8 November 2022). [159]

Strand, S. (2014), "Ethnicity, gender, social class and achievement gaps at age 16: intersectionality and 'getting it' for the white working class", *Research Papers in Education*, Vol. 29/2, pp. 131-171, https://doi.org/10.1080/02671522.2013.767370. [27]

Strand, S. and A. Lindorff (2018), *Ethnic disproportionality in the identification of Special Educational Needs (SEN) in England: Extent, causes and consequences*, http://www.education.ox.ac.uk/research/the-unequal-representation-of-ethnic-minorities-in-special-education-in-england-extent-causes-and-consequences (accessed on 22 February 2022). [53]

Strand, S. and G. Lindsay (2008), "Evidence of Ethnic Disproportionality in Special Education in an English Population", *The Journal of Special Education*, Vol. 43/3, pp. 174-190, https://doi.org/10.1177/0022466908320461. [54]

The French Council for School Evaluation (2021), *Pour une compréhension éclairée de notre École [For an enlightened understanding of our school]*, https://www.education.gouv.fr/media/117745/download (accessed on 18 November 2022). [116]

the iConnect consortium (2011), "Effect of questionnaire length, personalisation and reminder type on response rate to a complex postal survey: randomised controlled trial", *BMC Medical Research Methodology*, Vol. 11/1, https://doi.org/10.1186/1471-2288-11-62. [94]

The Nation's Report Card (2022), *Data Tools*, https://www.nationsreportcard.gov/data_tools.aspx (accessed on 8 November 2022). [108]

The Scottish Government (2020), *Scottish Index of Multiple Deprivation 2020*, https://www.gov.scot/news/scottish-index-of-multiple-deprivation-2020/ (accessed on 4 July 2022). [177]

The Scottish Government (2016), *National Improvement Framework for Scottish Education*, The Scottish Government, https://www.gov.scot/publications/national-improvement-framework-scottish-education-2016-evidence-report/documents/ (accessed on 4 July 2022). [176]

The Standing International Conference of Inspectorates (2021), *The Inspectorate of Education of Denmark*, https://www.sici-inspectorates.eu/members/inspection-profiles (accessed on 18 November 2022). [102]

Thomson, M. (2012), "Labelling and self-esteem: does labelling exceptional students impact their self-esteem?", *Support for Learning*, Vol. 27/4, pp. 158-165, https://doi.org/10.1111/1467-9604.12004. [22]

Torres, R. (2021), "Does test-based school accountability have an impact on student achievement and equity in education?: A panel approach using PISA", *OECD Education Working Papers*, No. 250, OECD Publishing, Paris, https://doi.org/10.1787/0798600f-en. [62]

UNESCO (2020), *Global Education Monitoring Report 2020: Inclusion and education: All means all*, Paris, UNESCO. [15]

UNESCO (2017), *Global Education Monitoring Report 2017/8: Accountability in education*, UNESCO, https://en.unesco.org/gem-report/report/2017/accountability-education (accessed on 8 November 2022). [144]

UNESCO (2014), *Celebrating excellence in education for people with disabilities*, https://webarchive.unesco.org/20161026074335/http://www.unesco.org/new/en/education/resources/online-materials/single-view/news/celebrating_excellence_in_education_for_people_with_disabilities/ (accessed on 18 November 2022). [125]

UNESCO-UIS (2018), *Handbook on Measuring Equity in Education*, UNESCO Institute for Statistics, http://uis.unesco.org/sites/default/files/documents/handbook-measuring-equity-education-2018-en.pdf (accessed on 10 July 2022). [4]

United Nations (n.d.), *Ensure inclusive and equitable quality education and promote lifelong learning opportunities for all*, https://sdgs.un.org/goals/goal4 (accessed on 10 July 2022). [6]

US Bureau of Labor Statistics (n.d.), *National Longitudinal Surveys*, Center for Human Resource Research (CHRR), Ohio State University, https://www.bls.gov/nls/ (accessed on 22 February 2022). [75]

Varsik, S. and J. Gorochovskij (Forthcoming), *Intersectionality in Education*, OECD. [28]

Victoria State Government (n.d.), *A Guide to Conducting Gender Impact Analysis*, https://www.parliament.vic.gov.au/images/stories/committees/paec/Inquiry_into_Gender_Responsive_Budgeting/Submissions/20a_Department_of_Premier_and_Cabinet_-_Attachment_A_-_A_Guide_to_Conducting_Gender_Impact_Analysis.pdf (accessed on 7 December 2022). [132]

West, M. (2010), "School-to-school cooperation as a strategy for improving student outcomes in challenging contexts", *School Effectiveness and School Improvement*, Vol. 21/1, pp. 93-112, https://doi.org/10.1080/09243450903569767. [172]

Wienen, A. et al. (2019), "The advantages of an ADHD classification from the perspective of teachers", *European Journal of Special Needs Education*, Vol. 34/5, pp. 649-662, https://doi.org/10.1080/08856257.2019.1580838. [20]

Wiseman, A. and J. Bell (2021), "Education without evidence: Gaps in data availability for refugee, asylee, and humanitarian migrant students in US schools", *Research in Education*, Vol. 112/1, pp. 95-108, https://doi.org/10.1177/00345237211034885. [86]

Wood, C. et al. (2016), *Premier League Reading Stars: Evaluation Report 2015-2016*, https://cdn.literacytrust.org.uk/media/documents/Premier_League_Reading_Stars_evaluation_report_2015-2016.pdf (accessed on 7 December 2022). [135]

Notes

[1] See Annex in Chapter 1 for more information about survey methodology.

[2] Gender was not considered in the Survey question.

[3] Total difficulties score is a measure of five aspects of children's development: emotional symptoms; conduct problems; hyperactivity/inattention; peer relationship problems; and pro-social behaviour. The results are collected in the Strengths and Difficulties Questionnaire in the Scottish Health Survey. The questions cover themes such as consideration, hyperactivity, malaise, mood, sociability, obedience, anxiety and unhappiness (The Scottish Government, 2016[176]).

The most deprived areas are defined as those in the bottom quintile of the Scottish Index of Multiple Deprivation. The Index ranks several thousands of areas in Scotland based on income, employment, health, education, access to services, crime and housing (The Scottish Government, 2016[176]; The Scottish Government, 2020[177]).

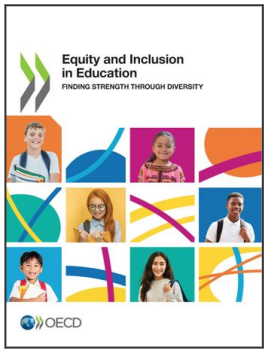
[4] "Early leaver from education and training, previously named early school leaver, refers to a person aged 18 to 24 who has completed at most lower secondary education and is not involved in further education or training; the indicator 'early leavers from education and training' is expressed as a percentage of the people aged 18 to 24 with such criteria out of the total population aged 18 to 24" (EUROSTAT, 2019[179]).

[5] In the 2021/22 edition for lower secondary students, immigrant status can be inferred based on the question whether the student was born in Canada or not, and the socio-economic background can be proxied based on students reporting "Always" or "Often" going to school or to bed hungry because there is not enough food at home.

[6] Canada being one of the few exceptions. The Canadian 2021 census included questions on both sex at birth and gender identity. While only two options were provided for sex at birth, the gender question also included an "other" text box (Statistics Canada, 2020[175]).

[7] See Chapter 4 for more information on collaborative teaching.

[8] Structural and institutional racism and discrimination against Roma people (European Network Against Racism, 2022[178]).



From:

Equity and Inclusion in Education

Finding Strength through Diversity

Access the complete publication at:
https://doi.org/10.1787/e9072e21-en

Please cite this chapter as:

OECD (2023), "Monitoring and evaluating equity and inclusion", in *Equity and Inclusion in Education: Finding Strength through Diversity*, OECD Publishing, Paris.

DOI: https://doi.org/10.1787/34da1f82-en