Assessment in Education: Principles, Policy & Practice

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/caie20

PISA 2018: did Sweden exclude students according to the rules?

Christian Andersson & Sofia Sandgren Massih

To cite this article: Christian Andersson & Sofia Sandgren Massih (2023) PISA 2018: did Sweden exclude students according to the rules?, Assessment in Education: Principles, Policy & Practice, 30:1, 33-52, DOI: 10.1080/0969594X.2023.2189566

To link to this article: https://doi.org/10.1080/0969594X.2023.2189566

ASSESSMENT IN EDUCATION: PRINCIPLES, POLICY & PRACTICE
2023, VOL. 30, NO. 1, 33–52
https://doi.org/10.1080/0969594X.2023.2189566

Routledge
Taylor & Francis Group

PISA 2018: did Sweden exclude students according to the rules?

Christian Andersson [a,b] and Sofia Sandgren Massih [a]

[a]Department of Performance Audit, The Swedish National Audit Office, Stockholm, Sweden; [b]Department of Industrial Economics, Blekinge Institute of Technology, Karlskrona, Sweden

ABSTRACT

This study assesses whether student exclusions from PISA 2018 in Sweden followed the criteria set by the OECD. We do this using both qualitative and quantitative methods. Our conclusion is that the exclusions made in PISA 2018 in Sweden did not follow OECD criteria and were much too high. Furthermore, interviews with school coordinators indicate that many of them misunderstood the OECD criteria. We also conclude that the National Agency for Education did not sufficiently follow up on exclusions. A review of the Swedish exclusion rate made by the OECD did not present credible results but accepted the results. A recalculation of PISA 2018 scores for Sweden where we assume non-participating students to be low performers show that results are significantly affected.

ARTICLE HISTORY
Received 20 January 2022
Accepted 27 February 2023

KEYWORDS
PISA; Sweden; OECD; participation rates; exclusion rate

The Programme for International Student Assessment (PISA) survey is an international assessment that measures 15-year-old students' reading, mathematics, and science literacy every three years. In many countries, including Sweden, the survey serves as a basis for policy decisions regarding future reforms and the allocation of resources to the school system. The survey is especially prominent in Sweden, considering the country's lack of national knowledge measurements of educational development over time.

Since results from international studies are used to make comparisons over time and between countries, strict criteria must be followed to ensure consistency. Furthermore, for results to be comparable for the same country over time, the criteria must also be applied in the same way for each test cycle. According to the OECD's guidelines, countries participating in PISA may exclude a maximum of 5% of students to ensure the results are consistent and comparable. However, in Sweden, more than 11% of students were excluded from PISA 2018.

Sweden's results from PISA 2015 indicated that a downward curve for Sweden had been broken, as results in all three subject areas had improved. The results of the 2018 survey confirmed this trend reversal (National Agency for Education, 2019). However, immediately after the results were published in December 2019, there was public

CONTACT Christian Andersson christian.andersson@riksrevisionen.se Department of Performance Audit, The Swedish National Audit Office, Stockholm 102 33, Sweden

questioning of the number of students excluded from the tests. The numbers of excluded and absent students increased significantly between 2015 and 2018. In 2018 the National Agency for Education (NAE) for the first time had full responsibility for the implementation of PISA. Between 2002 and 2015 the Mid-Sweden University were responsible for the implementation, and before 2002 Kalmar University. As the debate escalated, the OECD was asked by representatives of the Swedish government, in consultation with the National Agency for Education (NAE), to review the Swedish implementation of PISA 2018. However, the results of this review are questionable since the report is based on assumptions about the Swedish school system and preparatory classes that are not correct. Thus, it seems worthwhile to evaluate whether the Swedish PISA 2018 survey was carried out reliably and whether the Swedish government responsibly dealt with the doubts, especially considering the political relevance of PISA.[1] Exclusions and participation rates in PISA has also led to a public debate in both Canada and England following the work by Anders et al. (2021) and Jerrim (2021).

Therefore, this paper focuses on exclusion and non-response related to the implementation of the PISA survey in Sweden 2018. The main research question is if Sweden followed the OECD criteria when it comes to exclusions in PISA 2018. In this work we combine both qualitative and quantitative methods. We interview school coordinators to explore how the survey was implemented at schools. We also calculate how many students might have been excluded from the survey if the criteria had been adhered to. Finally, in an appendix we re-estimate the PISA results for Sweden using different and (in our opinion) more realistic assumptions about the exclusion rate than those used by the OECD in the above-mentioned report.

The remainder of this paper is structured as follows. The next section provides the background to the design of PISA. After that, we review the existent literature before discussing the methods used and the data we collected. The results are then provided and discussed, with our conclusions and final remarks offered in the final section.

Background

Sweden has participated in the PISA survey since the first survey in 2000. In 2018, a sample of approximately 5,500 15-year-old students from 207 compulsory schools and 16 upper secondary schools in Sweden took the test. The government (see ordinance 2015, 1047) has delegated responsibility for Sweden's participation in the PISA survey to the NAE. Since 2015, the NAE has also served as the national project centre for PISA, which means that they are responsible for carrying out the survey according to the rules put forth by the OECD. Previously, the NAE procured this function from Mid Sweden University.

Sweden's results in the PISA survey from 2000–2018

Sweden achieved its best PISA results in 2000 and thereafter the results gradually declined until 2012. The 2015 results showed a statistically significant improvement compared to the 2012 results. The PISA 2018 results showed that Sweden's reading, mathematics, and science scores had returned to roughly the same levels as in 2006, which is when the sharp decline in results began (Figure 1).

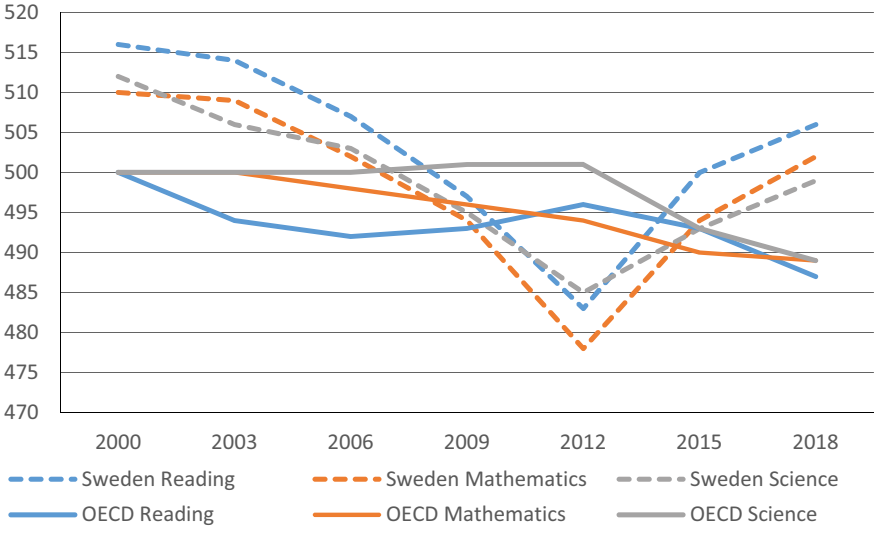

Figure 1. Sweden's performance in the PISA survey compared with the OECD average, 2000–2018. Source: OECD's performance reports.

In PISA 2018, Swedish students performed better than the OECD average in all three subject areas. Compared to PISA 2015, scores statistically insignificantly increased in all subject areas.

Previous research

Extensive research concerning the PISA survey is available in the literature and has increased over time – an overview of PISA research has been presented by Hopfenbeck et al. (2018). Here, we focus on research investigating how student sampling, student exclusions, and non-response might affect different countries' results and, thus, the comparability between countries and over time.

Anders et al. (2021) analysed Canada's results in PISA 2015. Although Canada has performed very well in PISA, it has a comparatively high student exclusion rate, low levels of school participation, and high rates of student absence. Therefore, the researchers concluded that serious problems arise when comparing Canada's PISA 2015 data to those of other countries. Sensitivity analyses show that Canada's PISA results would be lower than other high-performing countries – such as Estonia, Finland, Japan, and South Korea – if the country had not had such a high proportion of excluded students and non-response.

The authors suggest that the comparability of PISA results between countries can be improved by raising the minimum permitted response rate for each country from 80% to 90%. They also argue that the OECD's 5% limit for student exclusions should be applied more strictly. Also, the non-response bias analysis performed by the OECD should be more thorough and transparent. For example, the OECD could provide a security rating to indicate the reliability of each country's statistics.

Jerrim (2021) analysed PISA data from the UK, concluding that the data contain many anomalies. The combination of non-response, exclusions from the test, and technical details surrounding eligibility criteria led to a total non-participation rate of around 40%. Such a high exclusion rate means that the results' accuracy is highly uncertain. The researcher concluded that the OECD must be more transparent when defining the technical details underlying PISA.

Durrant and Schnepf (2018) analysed England's non-response in PISA 2000 and 2003 to see if non-response could distort the results. Using detailed administrative student data, the authors examined which student characteristics correlated with non-response both at the school and student level. Of particular interest is whether students' abilities are related to non-response. The results show that low-performing students are much more likely than high-performing students to belong to the non-response group. The researchers recommend implementing regular non-response analyses for the PISA survey and collecting additional statistics to improve future non-response analyses.

Schools impact the non-response rate; therefore, close cooperation with schools is recommended to encourage student participation. Freitas et al. (2016) also analysed the problem of representativity, finding a sizeable bias between the effective student distribution and PISA samples. Furthermore, they showed that the representativeness problems associated with PISA samples in Portugal can be corrected by applying post-stratified weights.

All major international knowledge surveys, such as the PISA, exclude students with disabilities. According to Schuelka (2013), this can make these surveys unrepresentative. Furthermore, by excluding students with disabilities, countries risk marginalising this group. As a result, this group might receive less attention in school debates and from politicians.

Brzyska (2018) analysed trends in the proportion of students with special educational needs (SEN) excluded from taking the PISA. Changes in definitions and legislation in participating countries regarding SEN students could affect the PISA response rate and results. Given the importance of PISA results in many countries, the author acknowledges that countries might use exclusion criteria to improve their positions compared to other countries. The OECD could deal with this problem by publishing the percentage of students with special needs excluded from the national selection of students in National PISA Reports.

LeRoy et al. (2018) also analysed students with special needs in PISA 2003–2012 and identified three main problems. Firstly, reports on students with special needs are not published regularly in connection with PISA. Secondly, the OECD controls access to statistics about students with special needs, making it difficult for researchers to analyse this issue. Thirdly, there is no standard format for reporting on SEN students within the PISA survey. Additionally, the author pointed out some major differences in how disability is classified in different countries, especially when it comes to cognitive and emotional disabilities.

Aursand and Rutkowski (2021) studied possible explanations for the increasing and relatively high PISA exclusion rates in Norway between 2000 and 2018. Based on interviews with six school leaders, the authors found a high degree of subjectivity among school leaders regarding exclusions. In some cases, school leaders want to minimise the feeling of failure among students by excluding them from the test. Their results show that

some school leaders confused OECD's exclusion criterion regarding time in the Norwegian school system with the time a student could attend preparatory classes. Also, subtle differences in language usage may impact whether some students are excluded. The researchers concluded that it may be important to provide extra guidance to school management staff at schools that offer preparatory classes.

Rutkowski and Rutkowski (2016) discussed methodological areas within PISA and recommend exercising caution when using and interpreting PISA results. For example, a high proportion of excluded students can lead to greater-than-expected distortions and, thus, possibly incorrect conclusions. Therefore, they recommend including a dedicated limitations chapter or section in every PISA report. Furthermore, in all tables and results, it should be clearly marked which countries did not meet the sampling standards.

Implementation of PISA

To ensure compliance with the OECD's PISA rules, each participating school must choose a school coordinator. School coordinators act as liaisons between the school and the national project manager. The principal of each selected school decides who should be the school coordinator, and an assistant principal or teacher is usually appointed. School coordinators play an important role in providing the NAE with lists of students who attend the school. School coordinators also identify which students meet the OECD's criteria for exclusion and, therefore, should be exempt from participation.

There are detailed OECD rules for much of the work to be done by school coordinators, and the NAE plays an important role in ensuring appropriate participation. The NAE's role includes training and providing support to school coordinators regarding the rules for selection, exclusion, and non-response. The NAE must ensure that the rules are applied and provide school coordinators with relevant manuals explaining how the tests should be carried out. The manual developed by the OECD (2017) for national project managers also states that the NAE must ensure that the student lists collected from each participating school are accurate.

The selection of students who participate in PISA surveys takes place in several steps.[2] Initially, all schools that have students of the right age are listed. Then, special needs schools and schools for pupils with learning disabilities are excluded (but included when calculating the exclusion rate) before schools are officially selected. The OECD then performs a stratified random selection process to determine which schools will participate. School coordinators provide lists of all students at the chosen school born in the appropriate year and specify which students require some form of extra educational support. Different codes are used for this purpose, depending on the form of support the student receives – for example, physical and mental disabilities and the need for additional language training correlate with different codes. However, in Sweden, all students who require extra support are assigned the same code.

Once the student lists have been submitted to the NAE, students are randomly selected to participate. In Sweden, 37 students from each school are chosen. If a school does not have 37 students who qualify, all qualifying students are selected. When school coordinators know which students have been selected, they must fill in a student protocol and specify which students should be excluded according to the OECD criteria. Not all students who receive special support are to be excluded during this process.

To maintain comparability between all participating countries, any permitted exclusions must be centrally defined and limited. If too many students are excluded, the survey results cannot be considered representative of the country's school system. The criteria for exclusion have remained unchanged since the start of the PISA survey in 2000 (OECD, 2015b).

The total exclusion rate may not exceed 5% of the target population (which includes excluded schools). The exclusion rate is calculated as the number of excluded students divided by the total number of students who would have participated if no schools or students had been excluded. The OECD (2005, 2015a) provides the following criteria to minimise exclusions:

- Students who have a moderate to severe permanent physical disability (such as a visual impairment) that prevents them from being able to take the PISA test.
- Students who have a cognitive, behavioural, or emotional disability confirmed by qualified staff, meaning they cannot take the PISA test. The functional impairment must be so severe that the student cannot follow even the general instructions of the assessment.
- Students who meet the following three criteria:
  - not a native speaker of the language in which the assessment is given
  - have limited proficiency in the assessment language
  - have received less than one year of instruction in the assessment language.

All 37 students from each school (minus excluded students) are expected to take the test. However, some students are not allowed to participate by their custodians. Additional students fail to show up to the test for different reasons. All such students are encoded as non-response. In addition, not all students who participate complete the test. A small proportion of students answer only a few questions – these students also receive the non-response code.

Exclusions, non-response, and coverage over time

The NAE's legal department re-interpreted the now revoked Personal Data Act in 2016. After that, individual exclusion categories were not allowed to be reported separately in Sweden – now, only the total number of excluded students are reported. In 2018, 681 (over 11%) of the selected students were excluded from the PISA survey. Of these, 9.8% points were student exclusions, and the rest were due to school exclusions.

In PISA 2015, 275 students were excluded, 154 due to functional disabilities and 121 due to limited language proficiency issues. Just under 6% of selected students were excluded in 2015, of which 4.5% points were student exclusions. Sweden has reported a gradual increase in the proportion of exclusions between 2003 and 2015. However, between 2015 and 2018, the exclusion rate rose by 94% (Figure 2). In 2018, 25 schools excluded eight or more students (i.e. more than 20% of the selected students).

Since Sweden do not classify the reason for exclusion it is difficult to know within which group the increase appeared. Data from the Swedish National Board of Health and Welfare regarding adolescents with neuropsychiatric diagnoses indicate that diagnoses

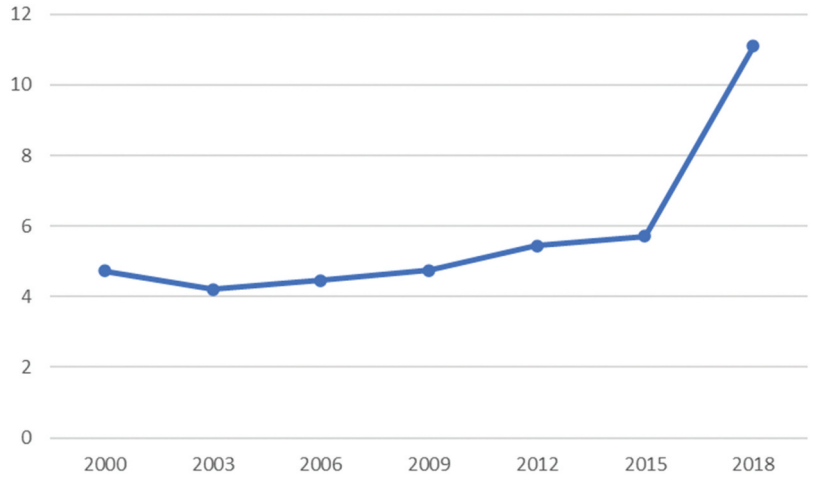

Figure 2. Exclusion rate (in percentages) for Sweden in PISA 2000–2018. Source: OECD.

have increased by around 20% since 2015. However, not all adolescents with diagnoses should be excluded from PISA. Only those who cannot follow the general instructions of the tests are to be excluded. Still, the increase indicates that the proportion of exclusions due to neuropsychiatric disabilities may have increased but only with approximately 0,5% points.

Assuming that exclusions due to physical or cognitive disabilities have not changed (but that behavioural and emotional disabilities increased by about 20%) yields an exclusion rate of just under 3% due to disabilities in 2018.[3] However, Sweden had a student exclusion rate of almost 10%, meaning that almost seven percentage points are unexplained.

Furthermore, even though several countries had total exclusion rates above the 5% limit set by the OECD in 2018, Sweden had the highest total exclusion rate of all participating countries. Furthermore, Sweden also has the highest student exclusion rate of all participating countries (the OECD limit for student exclusions is 2.5%; Figure 3).

The OECD adjudication board reviews the data for all cases in which the exclusion rate exceeds 5%. This board consists of representatives from the OECD, the contractor, the technical advisory group, and the sampling referee. Usually, the board approves the data. No explanations are given for countries just over the 5% limit, and the explanations offered for countries whose exclusion rates are much higher than the threshold are often incomplete and sometimes even inaccurate.

Methods and Data

We employ both quantitative and qualitative methods in this study. The quantitative parts help us answering if Sweden followed the OECD criteria when it comes to student exclusions in PISA 2018 and how the Swedish PISA results might have been affected by the high exclusions. The qualitative part gives us better insight in how the work with

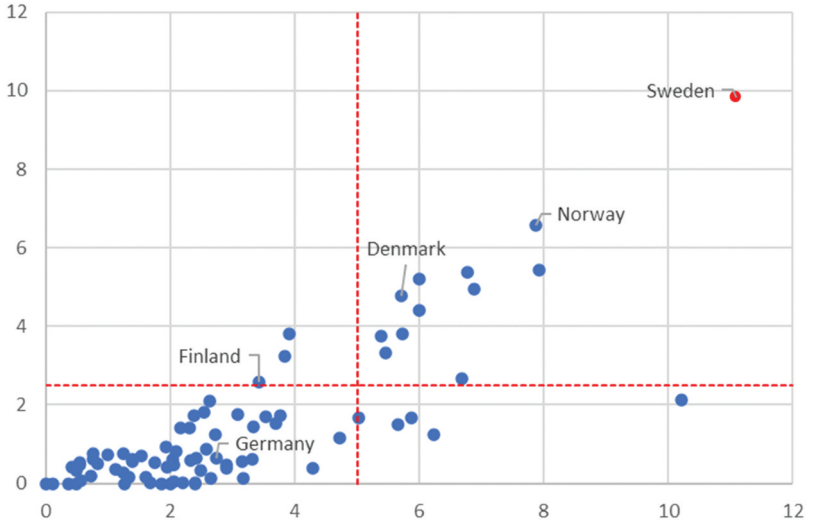

Figure 3. Total exclusions (x-axis) and student exclusions (y-axis) for different countries in the 2018 PISA survey, given as percentages. OECD limits for total and student exclusions given as dashed red lines. *Source: Statistics from the OECD processed by the Swedish NAO.*

exclusions was carried out at schools. Both parts complement each other and makes our overall conclusions stronger.

Quantitative methods

We use multiple data sources in determining whether the exclusions from PISA 2018 in Sweden were valid. From NAE we collected information from the Student register on each immigrant student in participating schools regarding the year and month they received residence permit in Sweden. We then use data from the Swedish Migration Agency on processing times for residence permits. Our main purpose is to determine whether Sweden followed the rules for exclusions using data from schools that participated in PISA 2018. We base our analyses on information regarding the number of students at each school and how many students participated in PISA 2018, the exclusion and non-response rates of each school. Data from the Student register is available up until October 2017.

Several data limitations need to be mentioned. For one, although we know how many students from each school were excluded from PISA 2018, we do not know precisely which students were excluded. Also, while we know how many newly arrived students attended each participating school, the data do not express for how long they had attended a Swedish school. How long a student has resided in Sweden before they are included in the population register differs between individuals and between grounds for residence permits. For example, immigrants seeking asylum have the right to participate in school almost immediately after arriving in Sweden, which they do long before receiving official permission to stay. In contrast, those who wish to move to Sweden because they have family living there cannot arrive before receiving their residence permit.

Furthermore, we can observe the flow of students with foreign backgrounds from year to year, but we do not know exactly which month each student started school. We also know how many students applied for (and were granted) a residence permit before the tests, but we cannot connect these data.

Due to the limitation mentioned above, we make several different calculations and assumptions to ensure that we ultimately provide a reliable answer to the research question. We also add some additional time to our assumptions regarding how long the newly arrived are assumed to have been in Sweden – we would rather overestimate the number of exclusions allowed than underestimate it. For example, even those who were granted asylum in the early 2018 may have been in Sweden for longer than 15 months, but since we do not have reliable information on these individuals, we consider them eligible exclusions.

Our analysis is then carried out as follows. Using the Student register, we map out the student bases from the schools that participated in PISA 2018 in terms of the number of newly arrived students. According to the NAE's statistical definition a newly arrived student is a student who has immigrated within the last four years. The crucial exclusion criterion related to limited language proficiency is that the student has had less than one year of education in the Swedish language. Thus, it is not sufficient to be a newly arrived student according to the statistical definition in order to be excluded.

The receiving municipality must enrol newly arrived students in school within one month of arrival and can then use up to two months to screen them. Therefore, we assume that all students who had been living in Sweden for at least 15 months should have been included in the PISA survey since they must have been enrolled in a class for more than a year. All children living in Sweden have the right to attend a Swedish school, and, as a rule, immigrant children are registered for school as soon as they arrive in Sweden. We know that for immigrants seeking asylum born in 2002 (i.e. the PISA-age in 2018) it took on average 19 months to receive a Swedish residence permit. Thus, we assume that students granted asylum in 2017 could not have been excluded due to the language criterion since they have been in school for more than one year. However, we make an exception for students who have received a residence permit as quota refugees because they receive their permits before arriving in Sweden. We also assume that students who have received a residence permit as a relative of a Swedish citizen in 2017 (or December 2016) can be excluded since these individuals receive their permit before entering the country.

There were also over 300 students in the sampled schools who had no personal identity number. Most of these students had not yet been granted a residence permit but they may also be undocumented migrants or students whose residence permit applications were rejected. Since we do not know these details, we assume that most students without a personal identity number can be excluded; however, we assume that some of them were granted asylum in the final months of 2017. We also assume that some students were granted residence permits as relative immigrants after October 2017 (when our last data point was recorded). Thus, we add them as potential exclusions.

In addition, in Appendix A we present another quantitative analysis. We show how Sweden's PISA 2018 results would have differed had all selected students participated. In this section we follow the methodology described by Anders et al. (2021) and extend the

analyses by also running simulations for other countries. For the simulations, we use PISA microdata made available online by the OECD.

Qualitative method

We obtained the qualitative data by conducting interviews with 32 school coordinators at Swedish schools that participated in PISA 2018. Most of the interviewed school coordinators had clear memories from their work with PISA. Only 3 school coordinators that we contacted declined participation since they did not believe they remembered enough. 7 school coordinators that were sampled were not possible to reach. The interviews were semi-structured using a prepared set of questions.[4] Stratified sampling was used to select school coordinators to interview. The different strata used were type of school, location, exclusion rate and non-response. Most of the interviews were conducted online, due to the Corona virus. All interviews were done by two persons who both took notes. We also interviewed former national project managers at Mid Sweden University to understand how they worked with previous PISA cycles and what factors led to success in that work.

Analysis and results

First, we present our quantitative analysis regarding exclusions from PISA 2018 in Sweden and whether the exclusions are valid per the OECD's rules. Also, we show how exclusions and non-response might have affected Sweden's 2018 PISA results. We then present our qualitative analysis based on interviews with school coordinators. We also discuss how the NAE informed schools and followed up on exclusions. In the following sub-section, we examine the OECD's review of Sweden's PISA 2018 exclusion rate.

Results from the quantitative analysis

Sweden's PISA 2018 exclusion rate

Based on the assumptions given above, we conclude that 2.5% of the students at the selected schools could be excluded per the language criterion (see Table 1 column 1 below). This represents an increase of 0.5% points compared to the language criterion exclusion rate for PISA 2015. The calculation considers all students at the schools selected for PISA 2018, not only the sample of 37 students from each selected school. Since the 37

Table 1. Calculations of potential exclusion rates per the OECD's language criterion based on different assumptions.

	(1) The Swedish NAO's preferred calculation	(2) All students in the school sample with an immigration year of 2017 can be excluded	(3) Same assumption as (2) but for all students throughout Sweden	(4) Same assumption as (3) but for students in PISA 2015	(5) Actual exclusions from PISA 2015 based on the language criterion
Exclusion rate	2.5%	3.0%	3.0%	2.3%	2.0%

Source: Calculations are based on statistics provided by the Swedish Migration Agency and the NAE.

students from each school are selected randomly, there is no reason to suggest that our inclusion of all students would distort the results of the calculations.

We have also made calculations where we allow all students in the selected schools whose immigration year is 2017 to be excluded, as well as students without a personal identity number, making special adjustments for the period from October to December. We derive a possible exclusion rate of 3% (Table 1 column 2). The corresponding calculation for the whole country yields an exclusion rate of just under 3% (Table 1 column 3). The result of a corresponding calculation based on 2015 data was 2.3% (Table 1 column 4). This is based on students born in 1999 and an immigration year of 2014, plus students without a personal identity number.

Our calculations indicate an increase in students who could have been excluded from PISA 2018. However, according to statistics from the Migration Agency, there were no more asylum seekers before PISA 2018 than there were before PISA 2015; also, there is only a marginal difference in terms of the number of residence permits granted. This indicates that there should be no significant increase in the exclusion rate. It is also possible that a larger proportion of individuals without a personal identity number (who are excluded in our calculations) could have been in Sweden for a longer time than we assumed.

Thus, our preferred estimate is that 2.5% of the student population could have been excluded due to insufficient knowledge of the assessment language. In PISA 2018, 11% of students in Sweden were excluded. Even though another 4% of students could have been excluded because of disabilities, this still leaves 4.5% points unexplained.

The impact of exclusions and non-response on Sweden's PISA 2018 results

To investigate how sensitive the Swedish PISA 2018 results are in terms of the sample and number of students who took the test, we replicate the method used by Anders et al. (2021) for PISA 2015 in Canada. The details of the analysis are presented in Appendix A. The overall conclusion from this exercise is that when non-participating students are assumed to be low performers and included in the results simulations the Swedish results in PISA 2018 are significantly affected. For example, our calculations show that Sweden's results would have been noticeably lower if excluded students had performed at the 25th score percentile on average. Also worth noticing is that the results are much more affected than for neighbouring countries like Denmark, Finland, Germany, and Norway.

Results from the qualitative analysis

School coordinators' interpretation of the OECD's exclusion criteria

Overall, our interviews with school coordinators indicate that they perceive the NAE's information meetings as useful, clear, and informative. Nevertheless, their responses indicate that they have different understandings of the information given at these meetings. Only about half of the interviewed school coordinators understood the information correctly (i.e. in accordance with the OECD's criteria).

Some school coordinators perceived that excluding students, in general, is up to the school coordinator's discretion, possibly together with other teachers. Among the school coordinators who misunderstood the criteria concerning limited language proficiency, two types dominate:

- Those who understood that newly arrived students may be excluded but refer to the NAE's definition of newly arrived students for statistical purposes (i.e. students who have lived in Sweden for up to four years).
- Those who thought that they should assess students' language proficiency levels themselves, regardless of how long the student had been taught in Swedish, and, on that basis, determine whether the student should be excluded.

The exclusion of students due to disabilities is slightly more complicated. For instance, some students who follow the syllabus for students with learning disabilities do so in regular schools. Although these students are to be excluded from the PISA, they might be elected since they are enrolled at regular schools. Special needs educators were often involved in decisions about whether children with behavioural or emotional disabilities should participate – however, the school coordinator sometimes made the decision him/herself. From our interviews, we conclude that some students with behavioural or emotional disabilities were excluded even if their disabilities did not hinder their comprehension of the instructions.

In some cases, it was obvious that school coordinators excluded students incorrectly. For example, some school coordinators excluded school refusals, and some excluded students who might find it difficult to sit through the test. One also excluded students whose custodians did not allow the student to take the test even though these students should have been registered in the non-response category.

Even though most of the coordinators we interviewed believed they had understood the rules correctly, about half of them made mistakes when excluding students. Their errors concerned both immigrant students and students born in Sweden with behavioural or emotional disabilities. Even though it is difficult to generalise from our interviews with approximately 14% of the coordinators who participated in PISA 2018, it is obvious that too many misunderstood the rules and made exclusions not in accordance with the criteria.

An assessment of OECD's review of Sweden's PISA participation 2018

When critiques were raised regarding Sweden's sharply increased exclusion rate for PISA 2018, the government and NAE asked the OECD (2020) to review Sweden's PISA 2018 statistics to confirm whether the results could be trusted. A report was delivered and published on 30 September 2020. Of the report's 55 paragraphs, seven are about Sweden. There is also a fact box where calculations are presented. The remaining text, taken from the OECD's manuals and reports, is very general. The review states that Sweden's national project manager (i.e. the NAE) has understood the rules for exclusions and that they were 'applied correctly to the best of their knowledge' (OECD 2020, §48).

The OECD concluded that Sweden's high exclusion rate was probably due to the high number of newly arrived students in the preceding years. However, the conclusions of the report are based on incorrect assumptions – namely, that all newly arrived students attended preparatory classes and that students who attended preparatory classes did so for 2.5 years. Moreover, the time spent in preparatory classes does not count as tuition in Swedish in the OECD report.

Based on these assumptions, the OECD makes calculations for the whole country using data from the student register. The year of immigration is not considered in these calculations, but students born in 2002 within the school system on October 15 in the respective year are divided into three categories: Swedish, Swedish-born with a foreign background, and foreign-born with a foreign background. It is assumed that the change in the number of students each year represents the number of students in the last category who started school in the current year.

The OECD further assumes that all students born abroad with a foreign background who were added to the student register after October 2014 could be excluded from the PISA per the exclusion criterion regarding limited language proficiency. However, some of these students may have attended Swedish school for up to 3.5 years by the time the PISA survey was conducted in March 2018. The statistics that the OECD used were delivered and calculated by the NAE.

Newly arrived students sometimes take preparatory classes – this depends on whether their school have chosen to organise the introduction of newly arrived students by way of preparatory classes, not on pupils' abilities. Many municipalities and schools do not have preparatory classes and, instead, integrate immigrants into ordinary classes right away. Thus, many newly arrived students do not attend preparatory classes. Preparatory classes are provided in Swedish and follow the same curricula as ordinary classes but with extra assistance. By law (per Chapter 3, Section 12 of the Education Act), students cannot attend preparatory classes for more than two years, and the students shall not be in preparatory class fulltime. Most students attend preparatory classes for a considerably shorter period than two years.

In the OECD's calculations, all foreign-born students with a foreign background who started studying in Swedish school after October 2014, according to the student register, are summarised. Then, 189 students were added to the calculations; these students were assumed to have started after October 2017 and, therefore, were not included in the student register. This adds up to over 6,000 students (1,774 + 2,931 + 1,133 + 189 = 6,027). However, the weighted number of students excluded due to limited language proficiency in PISA 2018 was about 7,500.[5]

Despite making assumptions that (in our opinion) are too inclusive, the OECD did not reach the figure of 7,500 students. However, the OECD declared that 6,027 is close enough to regard Sweden's data as reliable. The Ministry of Education and Research (2020) shortly thereafter made a statement that OECD had audited and accepted Sweden's results in PISA 2018. After the Swedish NAO informed the OECD in March 2021 that students can only participate in preparatory classes for two years, the OECD altered their calculations, arriving at a total of 5,300 students who could potentially be excluded due to the language criterion in the 2018 survey (i.e. 2,200 fewer students than were excluded). The number is weighted to represent the entire target population of students. Despite this, they do not alter their conclusion about the reliability of Sweden's data.

Based on the above discussion, we conclude that the OECD review does not credibly explain the high student exclusion rate for Sweden in PISA 2018.

Conclusion

The PISA survey is a significant international survey, as many countries alter their educational systems based on their PISA results. Thus, the survey is gaining attention among researchers. Several recent studies have explained some difficulties related to the PISA regarding, for example, transparency and representativeness.

In this study, we have shown that too many Swedish students were excluded from PISA 2018. While the OECD allows a maximum exclusion rate of 5%, Sweden's exclusion rate was 11%. According to our calculations, approximately 6.5% of selected students could have been excluded according to the OECD's rules. We also show that the discrepancy between the valid exclusion rate and the actual exclusion rate is mainly due to the mistakes made by school coordinators and the fact that the NAE did not properly follow up on the work with exclusions. Of note, such follow-ups are not required by the OECD.

Furthermore, the OECD did not adequately investigate why Sweden had such a high exclusion rate when they approved the figures provided in the adjudication report. Later, in a review procured by the NAE, the OECD tried to explain the exclusions. However, since the review involves some misunderstandings of the Swedish school system, we conclude that the OECD's assessment is not valid.

According to our simulations, for which we assumed that most of the selected students participated, Sweden's results might have been noticeably lower if the excluded students had performed at, for example, the 25^{th} score percentile on average.

We believe that the OECD should pay more attention to PISA exclusion rates when approving the data. Specifically, the adjudication board should more carefully investigate the causes for high exclusion rates before they approve data. They should also be much clearer when explaining why they have approved unexpected data. We also believe that reports should more often use asterisks to highlight cases where a country's data might not be reliable. Thus, with the release of the PISA 2022 results, it will be advisable to scrutinise the response and exclusion rates of different countries before the results are presented and accepted.

Notes

1. This paper is based on Audit Report 2021:12 by the Swedish National Audit Office (2021).
2. For detailed information on how the sampling in PISA is conducted please see OECD (2018).
3. In comparison, at the age of 15 for the cohort born in 2002, 18% more adolescents had some form of neuropsychiatric disability than in the cohort born in 1999. In 2015, 2.49% of the students were excluded due to a disability. We assume that 0.16% points of these were excluded due to cognitive disabilities (i.e. integrated students at schools for pupils with learning disabilities), which leaves 2.33% for other disabilities. $2.33 \times 1.18 = 2.75$. Added to this figure is the students integrated into schools for pupils with learning disabilities in 2018 (0.17), giving an exclusion rate of students with disabilities of 2.92%. A total of 9.84% of students were excluded in 2018, leaving 6.92% points unexplained. Calculation made with data sources from Statistics from the National Board of Health and Welfare and the National Agency for Education.
4. The interview guide with the prepared set of questions is provided as an online supplementary material.

5. We assume that the exclusions due to disabilities have not changed since 2015.
6. The number of weighted students is taken from the PISA Technical Report, Chapter 11, Table 11.7 and the variable Number of Students Assessed (Weighted). This is the weighted number of students calculated after adjusting for non-response. Since the Swedish NAO does not have access to student weights before adjustments for non-response for countries other than Sweden, the information from Table 11.7 is used for all countries. When student weights before adjusting for non-response are used for Sweden, the results are only marginally affected.

Disclosure statement

No potential conflict of interest was reported by the authors.

Notes on contributors

Christian Andersson holds a PhD in Economics from Uppsala University and his primary research interest is within Economics of Education.

Sofia Sandgren Massih holds a PhD in Economics from Royal Institute of Technology and her primary research interest is within Economics of Education.

ORCID

Christian Andersson http://orcid.org/0000-0002-6710-2920

References

Anders, J., Has, S., Jerrim, J., Shure, N., & Zieger, L. (2021). Is Canada really an education superpower? The impact of non-participation on results from PISA 2015. *Educational Assessment, Evaluation and Accountability*, *33*(1), 229–249. https://doi.org/10.1007/s11092-020-09329-5

Aursand, L., & Rutkowski, D. (2021). Exemption or Exclusion? A study of student exclusion in PISA in Norway. *Nordic Journal of Studies in Educational Policy*, *7*(1), 16–29. https://doi.org/10.1080/20020317.2020.1856314

Brzyska, B. (2018). Trends in exclusion rates for students with special educational needs within PISA. *Oxford Review of Education*, *44*(5), 633–650. https://doi.org/10.1080/03054985.2018.1496907

Durrant, G., & Schnepf, S. (2018). Which schools and pupils respond to educational achievement surveys?: A focus on the English programme for international student assessment sample. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, *181*(4), 1057–1075. https://doi.org/10.1111/rssa.12337

Freitas, P., Catela Nunes, L., Balcão Reis, A., Seabra, C., & Ferro, A. (2016). Correcting for sample problems in PISA and the improvement in Portuguese students' performance. *Assessment in Education: Principles, Policy & Practice*, *23*(4), 456–472. https://doi.org/10.1080/0969594X.2015.1105784

Gottfried, M. (2009). Excused versus unexcused: How student absences in elementary school affect academic achievement. *Educational Evaluation and Policy Analysis*, *31*(4), 392–415. https://doi.org/10.3102/0162373709342467

Government ordinance with instructions for the National Agency for Education. (2015:1047).

Hopfenbeck, T., Lenkeit, J., El Masri, Y., Cantrell, K., Ryan, J., & Baird, J. –. (2018). Lessons learned from PISA: A systematic review of peer-reviewed articles on the programme for international

student assessment. *Scandinavian Journal of Educational Research*, *62*(3), 333–353. https://doi.org/10.1080/00313831.2016.1258726

Jerrim, J. (2021). PISA 2018 in England, Northern Ireland, Scotland and Wales: Is the data really representative of all four corners of the UK? *Review of Education*, *9*(3), 1–41. https://doi.org/10.1002/rev3.3270

Knighton, T., Brochu, P., & Gluszynski, T. (2010). Measuring up. Canadian results of the OECD PISA study. Retrieved May 3, 2022, from https://www.cmec.ca/Publications/Lists/Publications/Attachments/254/PISA2009-can-report.pdf.

LeRoy, B., Samuel, P., Deluca, M., & Evens, P. (2018). Students with special educational needs within PISA. *Assessment in Education: Principles, Policy & Practice*, *26*(4), 386–396. https://doi.org/10.1080/0969594X.2017.1421523

National Agency for Education. (2019). PISA 2018 – 15-åringars kunskaper i läsförståelse, matematik och naturvetenskap. Retrieved May 3, 2022, from https://www.skolverket.se/getFile?file=5347.

OECD. (2005). PISA 2003 technical report. Retrieved May 3, 2022, from https://www.oecd.org/education/school/programmeforinternationalstudentassessmentpisa/35188570.pdf/.

OECD. (2015a). PISA 2003 technical report. Retrieved May 3, 2022, from https://www.oecd.org/education/school/programmeforinternationalstudentassessmentpisa/35188570.pdf .

OECD. (2015b). PISA 2018 technical standards. Retrieved May 3, 2022, from https://www.oecd.org/pisa/pisaproducts/PISA-2018-Technical-Standards.pdf .

OECD. (2017). *PISA national project manager manual*. Retrieved March 13, 2023, from https://www.oecd.org/pisa/pisaproducts/PISA-NATIONAL-PROJECT-MANAGER-MANUAL.pdf .

OECD. (2018). PISA technical report. Retrieved May 23, 2022, from https://www.oecd.org/pisa/data/pisa2018technicalreport/ .

OECD. (2020). A review of the PISA 2018 technical standards in regard to exclusion due to insufficient language experience. Retrieved May 3, 2022, from https://www.skolverket.se/download/18.2f324c2517825909a16357f/1628070064497/Revised_note_on_SWE_PISA_2018_20210310-2.pdf.

Rutkowski, L., & Rutkowski, D. (2016). A call for a more measured approach to reporting and interpreting PISA results. *Educational Researcher*, *45*(4), 252–257. https://doi.org/10.3102/0013189X16649961

Schuelka, M. (2013). Excluding students with disabilities from the culture of achievement: The case of the TIMSS, PIRLS, and PISA. *Journal of Education Policy*, *28*(2), 216–230. https://doi.org/10.1080/02680939.2012.708789

Swedish National Audit Office. (2021). Pisa-undersökningen 2018 – arbetet med att säkerställa ett tillförlitligt elevdeltagande. Retrieved March 13, from https://www.riksrevisionen.se/download/18.4cb651881790d82c1eb2111a/1619525207429/RiR%202021_12%20Anpassad.pdf.

Appendix

Appendix A

To investigate the impact of exclusions and non-response on Sweden's PISA 2018 results we have replicated the method used by Anders et al. (2021) where they examine Canada's PISA results in 2015. Their approach assumes that students who are not attending school, students excluded from the study, and students registered as non-respondents probably have a different result distribution in PISA compared to students who took the test. The proportion of 15-year-olds who do not attend school in Sweden is very low, and the same applies to neighbouring countries. Since we do not know the individual characteristics of the students who did not participate (microdata are not available), we make some assumptions about the PISA results for these individuals. The results can be altered depending on what assumptions are made.

To start, we assume that the PISA results for non-participants (excluded students, students registered as non-respondents, and 15-year-olds who do not attend school) would be lower than those of students who took the test. Gottfried (2009) showed that students who are often absent perform worse at school than those who attend regularly; Knighton et al. (2010) showed that this also applies to PISA in Canada.

Since there is no way to know how much worse non-participating students would have performed than other students on the PISA, our analysis instead examines how Sweden's PISA 2018 results might have changed if different assumptions had been made about how excluded students or non-respondents would have performed on the PISA if they had taken the test. At the individual level, PISA results consist of 10 plausible values. Plausible values are part of the imputation methodology that PISA is based on and are a selection of likely proficiencies for students that attained each score. We simplify the calculations by using the first plausible value in the same way as Anders et al. (2021). We also make comparisons with some of Sweden's neighbouring countries (Norway, Denmark, Finland, and Germany) by performing corresponding analyses for these countries under the same assumptions as for Sweden. We do this because some neighbouring countries have had significantly lower exclusion and non-response rates than Sweden. For instance, for PISA 2018, Germany and Finland had exclusion rates of about 2.7% and 3.4%, respectively.

First, we take the total number of 15-year-olds in Sweden from the technical report for PISA 2018 (108,622 students) and divide these into two groups: the number of participating students, weighted by the final weights (79,604 students), and the weighted number of non-participants (29,018 students).[6] Thus, for the participating group, the PISA results are used as described in the international PISA database but weighted so that the total is 79,604 students. Here, the term *participants* refer to the total weighted number of students who took the PISA test.

For students who did not participate in the survey, we randomly draw 29,018 results from a normal distribution. This is done using different assumptions about the sample's mean average while keeping the same standard deviation applied to the participating group. The value used as the mean for this normal distribution corresponds to different percentiles (the 45th, 40th, 35th, 30th, 25th, 20th, 15th, and 10th percentiles are used) of the observed PISA results for Sweden (and, correspondingly, for the neighbouring countries for which we also perform the analysis). Therefore, the results express the extent to which exclusions and non-response could have impacted Sweden's PISA 2018 results. We also make the corresponding calculations for neighbouring countries, which, in some cases, had significantly lower exclusion and non-response levels than Sweden.

Table A1 presents the simulated results for the reading test included in PISA 2018. Column 1 refers to the percentile in the distribution of reading results in PISA 2018 that we assume the non-participating students would have achieved if they had taken the test.

Sweden's average score on the reading portion of PISA 2018 was 506 points. This is slightly higher than the scores of Germany (498 points), Norway (499 points), and Denmark (501 points),

Table A1. Simulated reading scores for PISA 2018 under different assumptions about the performance of non-participating students.

Non-participants' scores as a percentile of the observed scores	Simulated PISA scores, mean				
	Sweden	Norway	Denmark	Finland	Germany
Original score	506	499	501	520	498
45	503	497	499	519	498
40	499	495	495	517	496
35	495	492	491	516	494
30	491	489	487	514	492
25	486	486	483	512	490
20	481	482	478	510	487
15	474	478	473	508	485
10	466	472	466	505	481

Source: Swedish NAO's processing of OECD microdata for PISA 2018.

but lower than Finland's score (520 points). Of these differences, only that with Finland is statistically significant.

When we recalculate Sweden's PISA 2018 score for reading (while making different assumptions about how well students who did not take the test might have performed), the scores change. When we assume that students who did not take the test performed marginally worse than average scores (i.e., on average, their scores are at the 40th percentile), the reading score falls to 499. This change is greater than that for comparison countries.

If we change our assumption about the scores of students who did not take the test, the simulated average reading score also changes. For example, if we assume that these students would perform very poorly on the reading test (i.e., if their average score was at the 10th percentile), their scores have a much more significant impact on the score. Specifically, Sweden's score would have decreased by 40 points, from 506 to 466 points.

By comparison, Finland's score would have been reduced by only 15 points following the same assumptions. Finland's results are not affected as strongly as Sweden's because a low percentage of Finnish students did not take the PISA in 2018. We find a similar result for Germany's score under the same assumptions (a 17-point decrease).

Furthermore, the differences between Norway's and Denmark's simulated PISA results and their actual results are not as large as the difference between Sweden's simulated and real results. If we were to assume that the students who did not take the test had performed at the 25th score percentile on average, Sweden would no longer have a better score than Norway, as both countries would have achieved an average score of 486 points. Also, Swedish students would have very slightly outperformed Danish students, whose average score would be 483 points.

Figure A1 illustrates the changes in PISA reading scores based on various assumptions about the performances of students who did not participate. As noted above, the differences between Sweden's average official PISA 2018 result and those of other countries are not statistically significantly different, with the exception of Finland, which had a statistically significantly better reading score than Sweden.

The results presented in Figure A1 show that Sweden's score is affected more intensely than the comparison countries when the scores of students who did not take the test are simulated. This is made particularly clear by comparisons with Finland and Germany. The analysis presented above for the different subject areas is based on assumptions that are probably not fulfilled. Nevertheless, but by making the same assumptions for all countries, it is easier to compare how the results in different countries could be affected if all selected 15-year-olds had participated.

Sensitivity analysis

We examine how sensitive the results are to our assumptions by considering that a certain number of students in Sweden did not participate in the PISA test for various reasons. This sensitivity

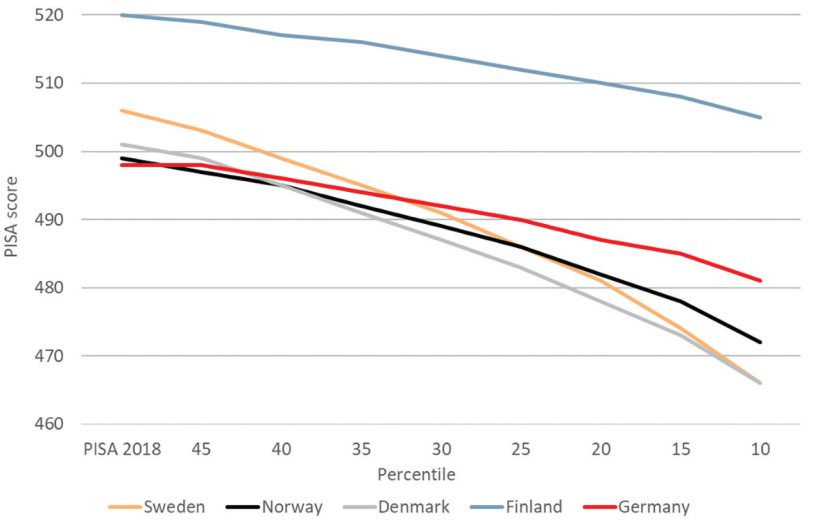

Figure A1. PISA 2018 scores in reading based on different assumptions about the results of non-participating students. *Source: Swedish NAO's processing of OECD microdata for the 2018 PISA survey.*

analysis is performed for results in reading, which was the main subject of PISA 2018. In the first of these analyses (option 1), we allow 2.5% of students to be excluded due to the high immigration of refugees to Sweden in the years before the PISA 2018 survey. However, some of Sweden's neighbours also received relatively large numbers of refugees during the same period. However, there is a notable difference regarding how many such students these countries excluded per the language criterion. For example, even though Germany (like Sweden) received many refugees in the years before the PISA 2018 survey, the percentage of students excluded due to the language criterion was low (0.4%).

In the second analysis (option 2), we also allow for an over-coverage of 3.1% (equivalent to 3,344 students) for Sweden. According to the NAE, the main over-coverage in the PISA 2018 survey is due to three closed compulsory schools and selected upper secondary schools that had no students born in 2002.

In a third analysis (option 3), we also do not simulate the results of 0.7% of excluded students (798 individuals) due to a discrepancy between the total number of individuals born in 2002 and the total number of individuals born in 2002 in Swedish school year 7 or higher.

In the concluding analysis (option 4), we also allow a 1.7% exclusion rate due to cognitive disabilities, which is equivalent to the median value for all participating OECD countries. The results of the different sensitivity analyses are presented in Figure A2.

Overall, the change in Sweden's potential results on the reading portion of PISA 2018 is relatively small when different assumptions are made about how many exclusions are allowed. However, the change increases the worse we assume that non-participating students would have performed. For example, the difference between the original simulation (when no exclusions and non-response are allowed) and option 4 (allowing exclusions due to the language criterion, over-coverage, and cognitive disabilities) is approximately three points when non-participating students are assumed to perform at the 30th percentile. Meanwhile, this difference is approximately 10 points when assuming that non-participating students would have performed at the 10th percentile on average.

Interview guide

- To start, could you please describe and share your thoughts about student exclusions and absences at your school in relation to PISA 2018.

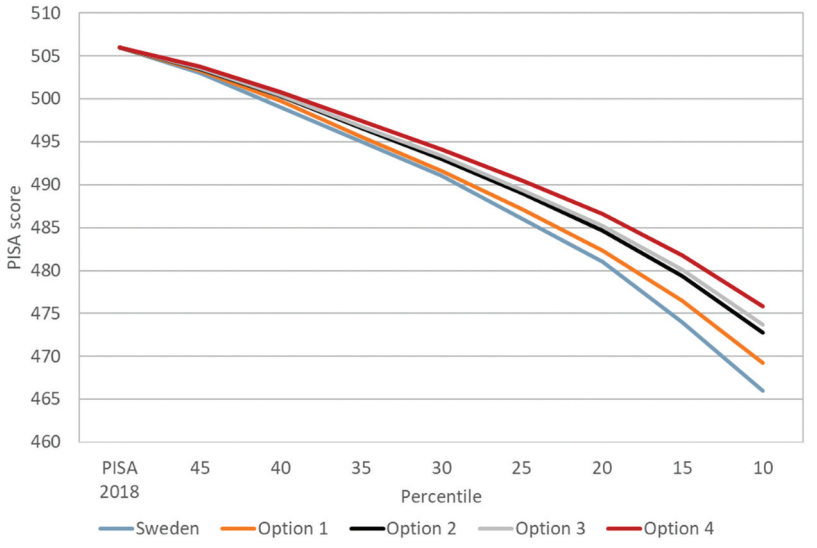

Figure A2. Changed PISA 2018 scores for reading achieved by Sweden under different assumptions about the number of allowed exclusions and the performance of students who did not take the test. *Source: Swedish NAO's processing of OECD microdata for the PISA 2018 survey.*

- How did you understand the OECD rules regarding which students should be excluded from PISA 2018 due to disability or lack of language skills (that is code 3)?
- How did you perceive that the work with exclusions functioned at your school?
- Your school had **X** students in need of support and X students who were excluded because of this reason. Can you please tell us how you reasoned about the students who were excluded from PISA 2018?
- Sometimes guardians do not let students take part in the PISA survey. How did you deal with this at your school? (X students did not participate for this reason at your school)?
- At your school X students were absent from PISA 2018. Do you know if that was because they did not show up, or because they could not answer enough questions? Do you have any reflections regarding this?
- Do you have any comments and thoughts on the OECD exclusion rules regarding functional disabilities and language barriers?
- How was the dialogue with the Swedish National Agency for Education on these issues? How did you experience the information and support given by the Swedish National Agency for Education?
- To your knowledge, how common is it that information on the students' school history is missing?
- Did your school admit a lot of new students during the refugee wave in the autumn of 2015? Were these students included or excluded in PISA 2018?