SDG 4.a.2 Percentage of students experiencing bullying in the last 12 months

**Metadata**

**Purpose:** This indicator provides information on the extent of self-reported violence and bullying in schools.

**Calculation method:** The indicator would be defined as the percent of students who have experienced any type of bullying in the past year. For assessment $i$, the measure of prevalence of bullying for the assessment’s target population $B_i$ would be defined as

$$B_i = E\left[\mathbf{1}\{b_1 = 1 \text{ or } b_2 = 1 \text{ or } ... \text{ or } b_{ni} = 1\}\right]$$  

(A1.1)

where $\mathbf{1}\{\ldots\}$ denotes the indicator function which takes value 1 if the expression is true and zero if false. Variables $b_1, \ldots, b_{ni}$ denote the various types of bullying included in the question on what types of bullying the student may have experienced; these variables equal 1 if the student has experienced the type of bullying and zero if the student answers no. $E[\ldots]$ denotes the population mean (expected value); the methodology for estimating the expected value of $\mathbf{1}\{b_1 = 1 \text{ or } b_2 = 1 \text{ or } ... \text{ or } b_{ni} = 1\}$ varies by assessment and depends on the assessment’s sampling design. Students who did not answer any of the bullying questions would be omitted from the calculation; for students that omitted some of the questions, the omissions would be treated as zeros. The target population would be that of the surveys/assessment but excluding those unwilling to answer any of the bullying questions. The time period, to match the SDG indicator definition, would be whether bullying was experienced at least once in a year.

**Interpretation:**

A high value indicates a large number of students at the given level of education are experiencing bullying in or near school and thus that the school is not a safe environment in which to promote learning.

**Type of data source:**

School-based surveys/assessments including student assessment data from student background questionnaires.

**Disaggregation:**

By sex, level of education, migrant status and SES. See Annexe on how these were estimated for the international learning assessment datasets.

*Updated September 2020*
Data sources:

Two main sources of data are reported. The first set of sources for this indicator are the Global School-based Student Health Survey (GSHS) developed by the World Health Organization (WHO) and the US Center for Disease Control and Prevention (CDC) in collaboration with UNICEF, UNESCO, and UNAIDS, and the Health Behaviour in School-aged Children study (HBSC) also administered by the World Health Organization. GSHS and HBSC are both school-based surveys conducted primarily among students aged 13-17 years (GSHS) and students aged 11, 13 and 15 years (HBSC). GSHS covers mainly developing countries while the HBSC covers Europe and North America. Table 1 describes the surveys and target populations. The second set of sources currently reported is the international student assessment surveys. Table 2 describes these sources, their target population, and definitions of bullying. Note that data from these sources has been added to the dataset but does not replace existing data points from the GSHS or HBSC.

<table>
<thead>
<tr>
<th>Survey</th>
<th>Target population</th>
<th>Scope of bullying</th>
<th>Asks about bullying in the past...</th>
<th>Years</th>
<th>Number of countries</th>
</tr>
</thead>
<tbody>
<tr>
<td>GSHS</td>
<td>students aged 13-17</td>
<td>Saying or doing bad and unpleasant things, teased a lot in an unpleasant way or when a student is left out of things on purpose. Prompted specifically for types of physical bullying, being made fun of, and being left out</td>
<td>Past 30 days</td>
<td>Rounds: 2003-2008, 2009-2012, 2013-present</td>
<td>101 countries in at least one round (global coverage)</td>
</tr>
<tr>
<td>HBSC</td>
<td>11, 13 and 15 year-old students</td>
<td>As GSHS, except only specific prompts for cyber-bullying</td>
<td>Past 2 months</td>
<td>Latest round: 2013/14*</td>
<td>42 countries in Europe, North America (and Israel)</td>
</tr>
<tr>
<td>Data source</td>
<td>Target population</td>
<td>Bullying question(s)</td>
<td>Responses</td>
<td></td>
<td></td>
</tr>
<tr>
<td>-----------------</td>
<td>-------------------</td>
<td>------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------</td>
<td>-----------</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
| LLECE 2013 (TERCE) | 6th grade students | Do any of these things happen to you when you are at school?  
- I am afraid of one of my schoolmates.  
- I feel threatened by one of my schoolmates.  
- I fear that one of my schoolmates will hit me or hurt me.  
- My schoolmates make fun of me.  
- My schoolmates exclude me.  
- My schoolmates force me to do things that I don’t want to do | bullied: yes  
not bullied: no  
omitted: missing |
| PISA 2018       | 15 year-old secondary students | During the past 12 months, how often have you had the following experiences in school? (Some experiences can happen in social media)  
- Other students left me out of things on purpose  
- Other students made fun of me  
- I was threatened by other students  
- Other students took away or destroyed things that belonged to me  
- I got hit or pushed around by other students  
- Other students spread nasty rumours about me | bullied: “a few times a year”; “a few times a month”; “once a week or more”  
not bullied: “never or almost never” (must be answered to all questions with valid responses to be classified as not bullied)  
omitted: missing for all questions |
| TIMSS 2015      | 8th grade students | During this school year, how often have other students from your school done any of the following things to you (including through texting or the Internet)?  
- Made fun of me or called me names  
- Left me out of their games or activities  
- Spread lies about me  
- Stole something from me  
- Hit or hurt me (e.g., shoving, hitting, kicking)  
- Made me do things I didn’t want to do  
- Shared embarrassing information about me  
- Posted embarrassing things about me online  
- Threatened me | Bullied: “At least once a week”; “Once or twice a month”; “A few times a year”  
Not bullied: “Never” (must be answered to all questions with valid responses to be classified as not bullied)  
omitted: missing for all questions |
| TIMSS 2015      | 4th grade students | During this school year, how often have other students                                                                                                                                                                                                                                                                                                                                                      | Bullied: “At least once a |
from your school done any of the following things to you (including through texting or the Internet)?
• Made fun of me or called me names
• Left me out of their games or activities
• Spread lies about me
• Stole something from me
• Hit or hurt me (e.g., shoving, hitting, kicking)
• Made me do things I didn't want to do
• Shared embarrassing information about me
• Threatened me

Not bullied: “Never” (must be answered to all questions with valid responses to be classified as not bullied)
Limitations and comments:

In addition to the differences in the target populations described above, there are several important notes on the comparability of indicators from the different sources. The GSHS and HBSC datasets report bullying in the past few months rather than one year. For the international student assessment data, comparability is also limited by (1) the additional question in TIMSS on whether the child has experienced the sharing embarrassing information, (2) the explicit question on on-line bullying in TIMSS, (3) the lack of time period specified in LLECE, (4) varying ages within grades which affect the prevalence of bullying, and (5) place of bullying. LLECE (TERCE) does not specify a time period; as a result, the TERCE item may include bullying that occurred more than a year ago. The indicator is based on the self-reporting by students of their experiences of bullying in or near school. There may be instances when some students feel sufficiently intimated that they do not report incidents that have taken place thus resulting in an over-estimate of the safety of the school environment.

Definition of indicator by data source:

<table>
<thead>
<tr>
<th>Survey source (noted in metadata points)</th>
<th>Source of estimates in the dataset</th>
<th>Definition of estimates</th>
</tr>
</thead>
<tbody>
<tr>
<td>HSBC</td>
<td>UNESCO (2019) <em>Behind the numbers: ending school violence and bullying</em>. Paris: UNESCO</td>
<td>“Median prevalence of students who reported being bullied [once or twice]… in the past few months prior to the survey.”</td>
</tr>
<tr>
<td>GSHS</td>
<td>GSHS Fact sheets, various countries, years</td>
<td>“Percentage of students who were bullied on one or more days during the 30 days before the survey”</td>
</tr>
<tr>
<td>LLECE 2013 (TERCE)</td>
<td>UIS estimates</td>
<td>Percent of students reporting being bullied at school</td>
</tr>
<tr>
<td>PISA 2018</td>
<td>UIS estimates</td>
<td>Percent of students reporting being bullied at least a few times per year at school</td>
</tr>
<tr>
<td>TIMSS 2015, average of 4&lt;sup&gt;th&lt;/sup&gt; and 8&lt;sup&gt;th&lt;/sup&gt; grade</td>
<td>UIS estimates, average of estimates for 4&lt;sup&gt;th&lt;/sup&gt; grade and 8&lt;sup&gt;th&lt;/sup&gt; grade</td>
<td>Percent of students reporting being bullied by other students at least a few times per year</td>
</tr>
</tbody>
</table>

Protocol for reporting the indicator

Due to the limitations on comparability of the different data sources, the indicator presents just one data source by country. When more than one data source is available for a given country, the selection criteria prioritized the main data sources (GSHS and HBSC) over learning assessment data. In turn, when there is more than one learning assessment available, the criterion for

*Updated September 2020*
selection was to take the one that best fits the definition of the indicator in terms of the target population and the reference period of the questions referring to bullying, therefore, the order of priority was: PISA, TIMSS, LLECE.
Annexe: metadata for estimates using cross national assessments

Definition of sub-populations

Female and male: The dataset used to estimate the indicators include a question asking whether the student is male or female. For TIMSS, the administrative record of the sex of the student was used following how TIMSS reports learning achievement scores by sex.

High and low SES: All assessments except TIMSS provide a measure of the socioeconomic status of students. This is typically based on the response by students about assets at home as well as education of parents. LLECE 2013 used the responses of the family questionnaire to generate its index. PASEC 2014 and PISA 2018 used student responses; no index was generated for the PASEC 2014 2nd grade students given their young age and reliability of answers. TIMSS reports an index of home learning resources based on household possessions reported by students and this was used as a measure of socioeconomic status. To define high and low SES students, the median was calculate for each country, student above the median were defined as high SES while those below were defined as low SES. See Table 2 for the names of the variables used to define high and low SES in each assessment.

Non-response and small sample sizes: Indicator estimates were not reported for sub-populations if data for the sub-population was available for less than 90 percent of sampled students or if the number of observations for a particular sub-population was less than 100.

Standard errors and confidence intervals methodology

The suggested methodology for estimating standard errors and subsequent confidence intervals varies by assessment and aim to account for clustering at the school-level. All surveys suggest using replicate methods in which the sample variation is obtained from variously defined sub-samples that mimic the sample design; the variation in estimates among the replicates provides an estimate of the sampling variation. The suggested methods were used for all assessments except LLECE 2013. For this survey, replicate weights were provided with each of the learning achievement datasets; however, a large number of students in the background dataset (which included the responses to the bullying and home language questions) were not included in the student achievement dataset. In order to maximize the background data, a linearization method for estimating the standard errors robust to clustering at the school level was used. Table 3 describes the methodology used for each assessment.
Table 1. Variables used to define high and low SES students

<table>
<thead>
<tr>
<th>Assessment</th>
<th>Population</th>
<th>Variable</th>
<th>Respondents</th>
</tr>
</thead>
<tbody>
<tr>
<td>LLECE 2013</td>
<td>Grades 3 and 6</td>
<td>Index of the family’s socioeconomic status (isecf)</td>
<td>Parents</td>
</tr>
<tr>
<td>PASEC 2014</td>
<td>Grade 2</td>
<td>n/a</td>
<td>n/a</td>
</tr>
<tr>
<td>PASEC 2014</td>
<td>Grade 6</td>
<td>Socioeconomic index of the student’s family (ses)</td>
<td>Students</td>
</tr>
<tr>
<td>PISA 2018</td>
<td>15 year-olds</td>
<td>Index of economic, social and cultural status (escs)</td>
<td>Students</td>
</tr>
<tr>
<td>TIMSS 2015</td>
<td>4th grade</td>
<td>Index of home resources for learning (asbghrl)</td>
<td>Students</td>
</tr>
<tr>
<td>TIMSS 2015</td>
<td>8th grade</td>
<td>Index of home educational resources (bsbgher)</td>
<td>Students</td>
</tr>
<tr>
<td>Assessment</td>
<td>Method</td>
<td>Reference for formulas</td>
<td>Software routine</td>
</tr>
<tr>
<td>--------------</td>
<td>---------------------------------</td>
<td>------------------------</td>
<td>-------------------------------------------</td>
</tr>
<tr>
<td>LLECE 2013</td>
<td>Linearized</td>
<td>StataCorp 2013</td>
<td>SVY module for Stata (StataCorp)</td>
</tr>
<tr>
<td>PASEC 2014</td>
<td>Jackknife repeated replication</td>
<td>PASEC 2017</td>
<td>PV module for Stata (Macdonald 2008)</td>
</tr>
<tr>
<td>PISA 2018</td>
<td>Balanced repeated replication</td>
<td>OECD 2009</td>
<td>PV module for Stata (Macdonald 2008)</td>
</tr>
<tr>
<td>TIMSS 2015</td>
<td>Jackknife repeated replication</td>
<td>Foy &amp; LaRoche (2016)</td>
<td>PV module for Stata (Macdonald 2008)</td>
</tr>
</tbody>
</table>
References


