SDG 4 Ensure inclusive and equitable quality education and promote lifelong learning opportunities for all

METADATA

Target 4.a Build and upgrade education facilities that are child, disability and gender sensitive and provide safe, non-violent, inclusive and effective learning environments for all

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

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

$$B_i = E\big[\mathbf{1}\{b_1 = 1 \ or \ b_2 = 1 \ or \ \dots \ or \ b_{n_i} = 1\}\big] \qquad \text{(A1.1)}$$

where $\mathbf{1}\{\dots\}$ denotes the indicator function which takes value 1 if the expression is true and zero if false. Variables $b_1,\dots, b_{n_i}$ 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. $E[\dots]$ denotes the population mean (expected value); the methodology for estimating the expected value of $\mathbf{1}\{b_1 = 1 \ or \ b_2 = 1 \ or \ \dots \ or \ b_{n_i} = 1\}$ 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.

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

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.

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.

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

Table 1. Summary of bullying data collected by the HSBC and GSHS

SURVEY	TARGET POPULATION	SCOPE OF BULLYING	ASKS ABOUT BULLYING IN THE PAST	YEARS	NUMBER OF COUNTRIES
GSHS	Students aged between 13 and 17 years old	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	Past 30 days	Rounds: 2003-2008, 2009-2012, 2013-present	101 countries in at least one round (global coverage)
HBSC	Students aged 11, 13 and 15 year old	As GSHS, except only specific prompts for cyber-bullying	Past 2 months	Latest round: 2013/14*	42 countries in Europe, North America (and Israel)

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

Table A1. Data sources and questions on bullying

DATA SOURCE	TARGET POPULATION	BULLYING QUESTION(S)	RESPONSES
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

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

DATA SOURCE	TARGET POPULATION	BULLYING QUESTION(S)	RESPONSES
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 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	- 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

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 3. Definition and origin of estimates reported in UIS STAT by source

SURVEY SOURCE (NOTED IN METADATA POINTS)	SOURCE OF ESTIMATES IN THE DATASET	DEFINITION OF ESTIMATES
HSBC	UNESCO (2019) *Behind the numbers: ending school violence and bullying*. Paris: UNESCO	Median prevalence of students who reported being bullied [once or twice]… in the past few months prior to the survey.
GSHS	GSHS Fact sheets, various countries, years	Percentage of students who were bullied on one or more days during the 30 days before the survey
LLECE 2013 (TERCE)	UIS estimates	Percent of students reporting being bullied at school
PISA 2018	UIS estimates	Percent of students reporting being bullied at least a few times per year at school
TIMSS 2015 (Average of 4th and 8th grade)	UIS estimates, average of estimates for 4th grade and 8th grade	Percent of students reporting being bullied by other students at least a few times per year

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 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.

Updated September 2020

Annex - 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.

Urban and rural: All assessments ask about the type of location in which the school is located to the school director; however, only LLECE 2013 asks explicitly whether the school is located in an urban or rural area. The other surveys ask the question in various ways included the number of inhabitants or by description. See Table 1 for the questions from each assessment and how they were mapped to urban or rural.

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 calculated 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. 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.

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

Table 1. Variables used to define high and low SES students

ASSESSMENT	POPULATION	VARIABLE	RESPONDENTS
LLECE 2013	Grades 3 and 6	Index of the family's socioeconomic status (isecf)	Parents
PASEC 2014	Grade 2	n/a	n/a
PASEC 2014	Grade 6	Socioeconomic index of the student's family (ses)	Students
PISA 2018	15 year old	Index of economic, social and cultural status (escs)	Students
TIMSS 2015	4th grade	Index of home resources for learning (asbghrl)	Students
TIMSS 2015	8th grade	Index of home educational resources (bsbgher)	Students

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

Table 2. Methodology for calculating standard errors by assessment

ASSESSMENT	METHOD	REFERENCE FOR FORMULAS	SOFTWARE ROUTINE
LLECE 2013	Linearized	StataCorp 2013	SVY module for Stata (StataCorp)
PASEC 2014	Jackknife repeated replication	PASEC 2017	PV module for Stata (Macdonald 2008)
PISA 2018	Balanced repeated replication	OECD 2009	PV module for Stata (Macdonald 2008)
TIMSS 2015	Jackknife repeated replication	Foy & LaRoche (2016)	PV module for Stata (Macdonald 2008)

References

Foy, P. and S. LaRoche (2016). Chapter 4. Estimating Standard Errors in the TIMSS 2015 Results. In. Michael O. Martin, Ina V.S. Mullis, and Martin Hooper, Editors, *Methods and Procedures in TIMSS 2015*. Boston College, TIMSS & PIRLS International Study Center
available at: https://timssandpirls.bc.edu/publications/timss/2015-methods/chapter-4.html

OECD (2009). *PISA Data Analysis Manual SPSS, Second Edition*. Paris: OECD. Available at: https://www.oecd-ilibrary.org/docserver/9789264056275-en.pdf?expires=1593772556&id=id&accname=guest&checksum=256201EAFD784D9168955CFDB6068789

PASEC (2017). *Manuel d'exploitation des données : Évaluation internationale PASEC2014*. PASEC, CONFEMEN, Dakar. Available at: https://www.pasec.confemen.org/wp-content/uploads/2017/06/Manuel_exploitation_donn%C3%A9es_%C3%A9valuation_internationale_PASEC2014.pdf

StataCorp (2013). *Stata Survey Data Reference Manual Release 13*. College Station, Texas, USA: Stata Press. Available at: https://www.stata.com/manuals13/svy.pdf