United Nations
Educational, Scientific and
Cultural Organization

UNESCO INSTITUTE FOR STATISTICS

TECHNICAL COOPERATION GROUP

SUSTAINABLE DEVELOPMENT GOALS

The quality of international data on teachers to report on SDG target 4.c

TCG6/REF/5

Based on a paper by Patrick Montjourides

Analysis of teachers-related data and metadata in the UNESCO Institute for Statistics database

Contents

1. Introduction

As the global education community strives to achieve the Education 2030-SDG 4 agenda questions on teachers, teacher effectiveness and teacher quality have surfaced and in turn direct the spotlight on the quality of data to monitor progress towards SDG target 4.c. Target 4.c aims, by 2030, to "substantially increase the supply of qualified teachers, including through international cooperation for teacher training in developing countries, especially least developed countries and small island developing States" (United Nations, 2018).

For the first time the international education agenda includes an explicit quantitative target and a set of monitoring indicators on teachers. The global reporting framework uses as global indicator the "Proportion of teachers in: (a) pre-primary education; (b) primary education; (c) lower secondary education; and (d) upper secondary education who have received at least the minimum organized teacher training (e.g. pedagogical training) pre-service or in-service required for teaching at the relevant level in a given country, by sex." (Indicator 4.c.1).

Additionally, six other indicators are recommended in the thematic monitoring framework:

"4.c.2: Pupil-trained teacher ratio by education level

4.c.3: Percentage of teachers qualified according to national standards by education level and type of institution

4.c.4: Pupil-qualified teacher ratio by education level

4.c.5 Average teacher salary relative to other professions requiring a comparable level of qualification

4.c.6: Teacher attrition rate by education level

4.c.7: Percentage of teachers who received in-service training in the last 12 months by type of training". (UNESCO Institute for Statistics, 2018b)

Concerns have however emerged as the status of the international data collection on teacher remains fairly weak. Issues of coverage and definitions substantially undermine the achievement of a sound and sustainable monitoring of the teacher issue. As the custodian agency of SDG 4 data and indicators, the UIS is at the frontline of such discussion and has started to mobilize partners to tackle the issue of more and better teacher data to monitor SDG target 4.c.

Indicators listed above in the global and thematic monitoring frameworks on teachers imply that the global education community has common and understood definitions of trained teachers and of qualified teachers, that most teacher data are disaggregated according to various dimensions (education level, sex, training, type of institutions but also entry into the profession). This represents a consequent jump in comparison to the MDG/EFA era where the main indicators used was the pupil-teacher ratio (PTR).

This paper contributes to the discussion by analyzing UIS data on teachers. The data analyzed are the data collected annually by the UIS through its regular education surveys. The paper includes two parts. The first part is an analysis of teacher training programmes reported by countries in their ISCED

mappings. The second part is an overview of the quality of teacher-related data in the UIS database. These data form the basis for calculating all official SDG 4 indicators and as such play a fundamental role in the international education community.

Currently, SDG target 4.c suffers from consistently low levels of international reporting. Two issues are considered as undermining global reporting: the absence of simplified yet comparable definitions of trained and qualified teachers and the insufficiency of the global teacher data collection process.

The objectives of this paper are thus twofold:

1. To explore how national teacher training programmes can be summarized by a small set of available and comparable variables. Doing so would enable the countries to benchmark their teacher training programmes on a common scale and eventually improve global reporting of trained and qualified teachers.

2. To highlight issues pertaining to the current quality of teachers data in the UIS database and suggest areas of improvement for the global education community to further develop its capability to monitor teachers in the SDG 4 agenda.

It builds on earlier papers commissioned by the UIS, which highlighted the following:

- Data on trained and qualified teachers is critically lagging behind. Only slightly more than a third of low- and middle-income countries and territories are reporting data on trained and qualified teachers to the UIS (UNESCO Institute for Statistics, 2018c). This critically jeopardizes achievement of SDG target 4.c which aims to increase the supply of qualified teachers (United Nations, 2018).

- There are substantial variations in pre-requisites to enter teacher training programmes as well as in actual characteristics of the programmes (content, duration, structure, etc.). Variations are observed between countries (income groups, region) and within countries (federal States, specific teacher training programmes, level of education).

- The global education community would benefit from the development of a typology/taxonomy for teacher training programmes, which would improve the global dialogue on trained and qualified teachers.

- Previous attempts by the UIS to collect data on teacher training programmes have yield unsatisfactory results due to insufficient training and preparation of countries to report data on teacher training programmes (UNESCO Institute for Statistics, 2018c).

- One substantial issue is the high number of dimensions and the variability between teacher training programmes between and within countries. Previous studies have highlighted at least nine dimensions that could be used to characterize a teacher training programme. This in turn has consequences on the feasibility of coming up with an international typology of teacher training programmes (ibid).

While data on teacher training programmes is scarce and incomplete, a number of findings can be highlighted from an analysis of the data currently available. A significant number of countries have reported data on teachers both in their ISCED mapping and in UIS Questionnaire A.

The paper thus seeks to answer the following questions:

- To which extent could existing data and metadata be used to explore the development and operationalization of a teacher training programme typology?

- What additional steps ought to be taken to propose a fully operational project towards an international teacher training programme typology which would reach the same standards of quality and use as other international standards, such as the International Classification of Education (ISCED)?

- What is the status of reporting on teachers and what are some of the key issues hindering the publication of indicators on trained and qualified teachers?

Section 2 provides an analysis of data on the characteristics of teachers training programmes. These data have been reported as part of ISCED regular data collection on all education programmes. This is therefore an exploration of what is currently feasible with existing metadata on teacher training programmes. Section 3 looks at the current state of the UIS data collection on teachers notably in light of the reporting requirements for SDG target 4.c.

2. Analysis of teacher training programmes reported in ISCED questionnaires

This section explores the feasibility of coming up with a simplified classification scheme for teacher training programmes. As highlighted in the introduction, proposals to classify and compare teacher training programmes are currently limited by the large number of characteristics to describe a teacher training programme and their relatively poor reporting in international data collection exercises. The section thus build on existing data reported to the UIS to proposed a simplified baseline to describe and categorize teacher training programmes and eventually better enable the comparison of the share of trained and qualified teachers across countries.

2.1 Dataset

i. ISCED Mappings

Data analyzed in this first part have been extracted by the UNESCO Institute for Statistics from its database and consist of all teacher training programmes found in national ISCED mappings. ISCED mappings are produced by the UIS following UNESCO's Member States reporting through UIS regular data collection. ISCED mapping are an essential tool of international comparability of education systems as they map national education systems onto the International Standard Classification of Education (UNESCO Institute for Statistics 2012). According to the UIS *"ISCED mappings ensure a transparent process of coding national education programmes and related qualifications into comparable categories for use in international statistics by linking the classification criteria to the properties of the education programmes and their related qualifications."* The mappings are produced following a two-stage process. First the UIS gathers detailed information on national education programmes as part of their annual survey (the ISCED questionnaire is available at http://uis.unesco.org/uis-questionnaires). Data collected include for each education programme: entry requirements, entry age, duration and diplomas obtained. The UIS then

proceeds, as second step, to produce an individual ISCED mapping for each country following the guidelines provided by the latest ISCED revision (currently ISCED 2011) and in collaboration with countries. There are also additional exploitable data if one uses the various open fields (entry requirements, programme name, qualification awarded at the end of the programme and note) to characterize each teacher education programme in a structured and consistent way.

Relying on ISCED mappings as the primary source of data for analysing teacher training programmes has however a number of limitations. The first and most obvious one being that a country with no ISCED mapping will not be analysed in this document. Coverage of ISCED mappings is however quite good and 81% of all countries and territories listed on the UIS website have a validated ISCED 2011 mapping of their national education programmes (Table 1). There are, however, variations by region. Coverage ranges from 100% for Arab States to only 43% in Latin America and the Caribbean (LAC). Nevertheless, most other regions are at or around 90% coverage, missing one or two countries' mappings. Only the case of LAC could appear to be problematic regarding the external validity of the results. That is, the possibility to generalize the findings is potentially weakened by under-representation of a specific region. However, whether an ISCED mapping is listed or not on the UIS website does not imply that data do not exist. A number of countries have shared their national education programmes structure with the UIS but are still going through the validation of their ISCED mappings. This can be the case when country focal points did not confirm yet agreement with UIS staff's interpretation of national education programmes or if the UIS faces delays in validating internally the mappings. A number of countries from LAC are in this situation nevertheless the data they submitted is available on the UIS website, on the questionnaire page linked above.

Table 1. Availability of (validated) ISCED mappings

	ISCED 2011 Mapping	*(%)*	No mapping	*(%)*	Total
North America and Western Europe	27	*96*	1	*4*	28
Arab States	20	*100*		*0*	20
Central and Eastern Europe	18	*90*	2	*10*	20
Central Asia	8	*89*	1	*11*	9
East Asia and the Pacific	27	*79*	7	*21*	34
Latin America and the Caribbean	18	*43*	24	*57*	42
South and West Asia	8	*89*	1	*11*	9
Sub-Saharan Africa	42	*91*	4	*9*	46
Total	168	*81*	40	*19*	208

Source: UIS

ii. Teacher training programmes in ISCED mappings

There are other limitations to relying on the ISCED questionnaire as the primary data source for analysing teacher training programmes. First, it requires that teacher training programmes are identified within the list of national education programmes. If teacher training programmes are part of a broader set of education programmes, say a bachelor's or a master's degree, then countries will not list teacher training

programmes as a separate education programme. Teacher training programmes can also be difficult to identify from the list of national education programmes for other reasons. It requires notably that the level at which teachers are trained to teach is specified in one of the text fields of the recorded programme, a requirement that is usually not fulfilled. Eventually, relying on ISCED mapping data also restricts the identification of pathways to the teaching profession, an issue that is worth considering while addressing the challenge of exploring cross-national comparability of teachers. In many countries, it is possible to become a teacher via several pathways besides the initial teacher qualification pathways (UNESCO Institute for Statistics, 2018c). This is also apparent when one looks at the variations in educational attainment of teachers at a given level in many developing countries. While it is possible to deduce initial teacher qualification pathways for a small number of countries in the sample, it is not possible to do it in a systematic way. Consequently, the analysis will mostly focus on the distribution of available characteristics of initial teacher training programmes rather than on the pathways to the teaching profession.

iii. Sample size

The dataset provided by the UIS included eighty-four countries, which reported 171 teacher education programmes. Those are countries which had reported teacher training programmes as part of their ISCED mapping exercise (Table 2). This is less than 40% the total number of 213 countries and territories covered by the UIS. Prospects for deriving meaningful and comprehensive descriptions of teachers training programmes are thus limited by low levels of coverage. These prospects are further limited by variations in regional coverage. Some regions are better covered by teacher training programmes reporting than others. Low-income countries and sub-Saharan African countries in particular are well covered as they have include teacher training programmes in their ISCED mapping. Respectively 65% and 63% of these countries have information on teacher training programmes. Central Asian countries are absent from the analysis in this section as none of them currently have information on their teacher training programmes in UIS database. At the regional level, and with the exception of sub-Saharan Africa and Latin America and the Caribbean (for which more than half of the countries have metadata on their teacher training programmes), the amount of metadata available remain limited. Less than a quarter of countries in Arab States, Central and Eastern Europe, and the Pacific can be analysed and slightly more than a third of countries in East Asia and in North America and Western Europe are included in the analysis based on what is available in the UIS database. The sample is also characterized by relative under-representation of upper-middle- and high-income countries and over-representation of low- and lower-middle-income countries. These points have to be born in mind while discussing the characteristics of teacher training programmes from a global perspective. Eventually it should be considered that the analysis constitutes an analysis of UIS data rather than an analysis of teacher training programmes at the global level.

Table 2. Number of countries reporting teacher training programmes to UIS, by region and by income group

	Total number of countries	**As % of Total**	**Included in the analysis**	***As % of sample***	***(% included)***
Region					
Arab States	20	*9*	4	*5*	*(20)*
Central Asia	21	*10*	0	*0*	*(0)*
Central and Eastern Europe	9	*4*	2	*2*	*(22)*

East Asia and the Pacific					
East Asia	17	*8*	6	*7*	*(35)*
Pacific	17	*8*	4	*5*	*(24)*
Latin America and the Caribbean					
Latin America	19	*9*	10	*12*	*(53)*
Caribbean	24	*11*	14	*17*	*(58)*
North America and Western Europe	31	*15*	11	*13*	*(35)*
South and West Asia	9	*4*	4	*5*	*(44)*
Sub-Saharan Africa	46	*22*	29	*35*	*(63)*
By income group					
High income	78	*37*	26	*31*	*(33)*
Upper middle income	55	*26*	16	*19*	*(29)*
Lower middle income	46	*22*	20	*24*	*(43)*
Low income	34	*16*	22	*26*	*(65)*
Grand Total	**213**		**84**		***(39)***

Source: UIS, author's calculation

The limitations of the sample are compounded by the scarcity of data that can be used to identify the level at which teachers are trained to teach. This information could in theory be derived from the programme name, the certification awarded at the end of the programme (typical categories of ISCED mapping), or the "notes" open text field. But information is often incomplete or imprecise. Additionally, some programmes included in the sample provided by the UIS were not teacher training programmes but pedagogical or orientation counsellors training programmes. These have been excluded from the sample. Finally, some of the data were also not translated even though the database includes a column for English translation, reflecting either an issue in reporting by countries or data processing at the UIS. Manual coding of levels at which teachers are trained to teach has been performed based on the fields *Programme Name* and *Qualification Name in English*, as well as *Notes*. Note that many times it was not possible to deduce the level of education for which teachers are trained.

After recoding, the sample of teacher training programmes exploitable for further analysis became substantially smaller. Eventually, the current state of the UIS ISCED survey only permits analysis of teacher training programmes for 13 countries at ISCED 0, 34 countries for ISCED 1, and respectively 18 and 14 countries for ISCED 2 and ISCED 3. These represent a total of 79 programmes or 46% of all programmes available. An additional 23 programmes (13%) could be identified as training teachers for multiple levels. Furthermore, 45% of the sample is not fully exploitable for further analysis of teacher training programmes as teacher training programmes included in the unspecified group do not provide enough information to infer the specific education level for which teachers are trained (Table 3).

Table 3. Number and distribution of teacher training programmes identifiable in the sample, by region and ISCED level for which teachers are trained

	ISCED level (Single)				Multiple ISCED levels							
	0	**1**	**2**	**3**	**0/1**	**0/1/2**	**0/1/2/3**	**1/2/3**	**1/3**	**2/3**	**Unsp.**	*n*
Arab States	0	30	30	10	0	0	0	0	0	20	10	*10*
Central Asia	0	0	0	0	0	0	0	0	0	0	0	*0*
Central and Eastern Europe	0	0	0	20	0	0	0	0	0	0	80	*5*
East Asia and the Pacific												
East Asia	14	10	5	10	5	5	0	5	0	5	43	*21*
Pacific	0	20	0	0	0	0	0	0	0	0	80	*5*
Latin America and the Caribbean												
Latin America	0	25	0	0	8	0	8	8	0	0	50	*12*
Caribbean	0	14	14	5	5	0	0	0	0	0	64	*22*
North America and Western Europe	19	0	0	0	0	0	0	0	0	10	71	*21*
South and West Asia	0	40	0	40	0	0	0	0	0	0	20	*5*
Sub-Saharan Africa	9	29	16	10	3	0	0	3	1	9	21	*70*
n	*13*	*34*	*18*	*14*	*5*	*1*	*1*	*4*	*1*	*11*	*69*	***171***
% of Total	8	20	11	8	3	1	1	2	1	6	40	

Source: UIS, author's calculations. Note: Unsp. = unspecified

Additional data corrections were made. In some instances, information provided by the UIS differed from what was published on the website. When this was encountered, choice was made to use the officially published data. Several programmes could be entered at various points (for instance after successful completion of ISCED 3 or after completion of ISCED 6 but with a shorter duration). As the objective is to compare teacher education programmes, choice was made to consider the earliest entry point and the longest duration for programmes where this issue arise. A detailed list of data manipulations and additional variables created is provided in Annex 1.

2.2 Characteristics of teacher training programmes

This section aims to detail the distribution of characteristics of teacher training programmes as observed in the UIS database. An initial discussion of patterns in entry requirements, level of training and duration can be found in an earlier paper produced by the UIS (UNESCO Institute for Statistics, 2018c) yet the discussion is based on a less detailed and less structured database, SABER teacher data (World Bank, 2011, 2013). Findings from this section can thus be used to precise the patterns highlighted in earlier studies but also to explore in further detail how these patterns could be used as inputs into the development of an international typology of teacher training programmes.

i. Teacher training programmes by level

a) Early childhood and pre-primary education teachers

In the current database only thirteen ISCED 0 teacher education programmes could be clearly identified (Table 4). Of these, two countries (Germany and Mauritius) have included 3 or more programmes. The analysis is thus limited to eight countries. No clear global pattern could reasonably be discussed from such a sample but a number of points can be highlighted. Teacher education programmes for teaching at the pre-primary education level are usually not at the tertiary education level. Germany (Bachelor level) and Kenya (short-cycle tertiary education) are two exceptions. In the poorest countries of the sample, ISCED 0 teachers only have a diploma or certificate corresponding to completion of upper-secondary education (Ethiopia, Mali, Lao PDR). This represents potentially significant differences in terms of complexity and depth of the skills, knowledge and competencies acquired. The modal level for early childhood and pre-primary teacher education programmes is ISCED 4 (Figure 1) but this pattern remains to be confirmed upon analysis of a better sample.

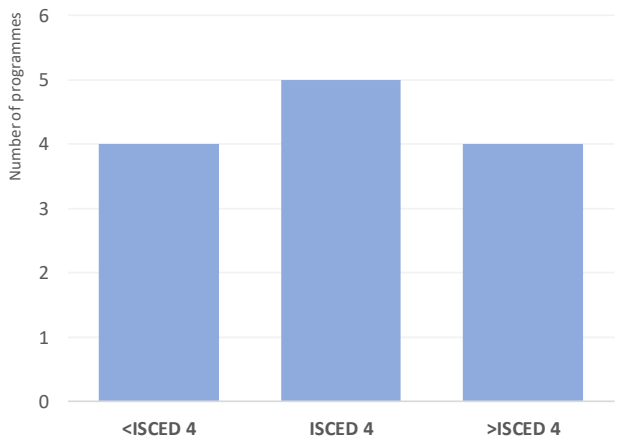
These initial insights are aligned with patterns described in earlier studies where teacher education programmes are of higher ISCED level as we move from the low-income to the higher-income group (UNESCO Institute for Statistics, 2018c).

Table 4. Early-childhood and pre-primary teacher training programmes

Country	Income Group	Region	Programme name	Th. Age	(Ed. cycle point equivalent)	Th. Dur.	ISCED	ISCED level required to teach that programme
Germany	High	NAWE	Practical training year at specialised vocational schools preceding kindergarten teacher training programmes	16	Entr. ISCED 3	1	2	n.a
Germany	High	NAWE	Kindergarten teacher programme at trade and technical schools.	21	ISC6. (2nd year)	3	6	n.a.
Germany	High	NAWE	Kindergarten teacher programme at vocational academies.	19	Entr. ISCED 4,5,6	3	6	n.a.
Germany	High	NAWE	Kindergarten teacher programme at specialised vocational schools.	19	Entr. ISCED 4,5,6	3	6	n.a
Ethiopia	Low	SSA	Pre-school Teaching certificate programme	17	Entr. ISCED 3	<1	3	n.a.
Mali	Low	SSA	Formation des éducateurs du préscolaire	16	Entr. ISCED 3	4	3	6
Kenya	Lower middle	SSA	Pre-Primary Teacher training	18	Entr. ISCED 4,5,6	2	5	6
Cambodia	Lower middle	EAP	Teacher training for Pre school	18	Entr. ISCED 4,5,6	2	4	6
Lao PDR	Lower middle	EAP	Pre-school teacher education program	15	Entr. ISCED 3	3	3	n.a.
Myanmar	Lower middle	EAP	Pre-school Teacher Training Course	17	ISC4. (2nd year)	0.25	4	4
Mauritius	Upper middle	SSA	Teacher's Certificate in Early Childhood Education	19	ISC4. (2nd year)	1	4	5
Mauritius	Upper middle	SSA	Teacher's Diploma in Early Childhood Education	19	ISC4. (2nd year)	1	4	5
Mauritius	Upper middle	SSA	Teacher's Certificate in Pre-vocational Education - School Based Experience	18	Entr. ISCED 4,5,6	1	4	5

Notes: *Th. Age* corresponds to the theoretical entry age of the programme, *Th. Dur.* To the theoretical duration, *Ed. Cycle point equivalent* corresponds to the theoretical point that could be reached at the specified *Th. Age*

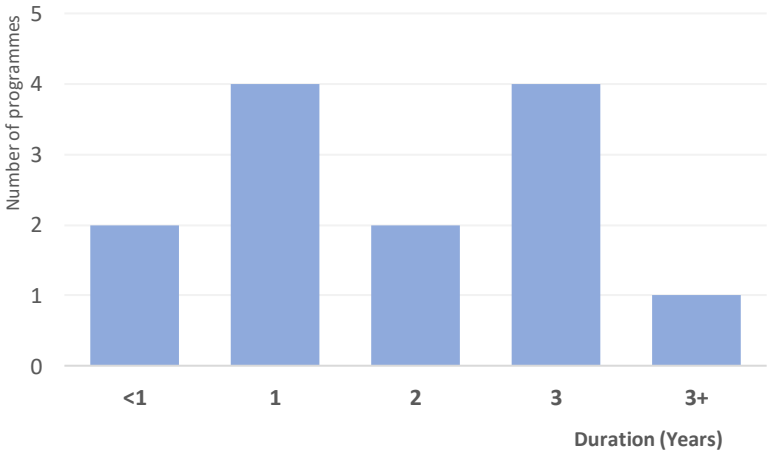
Figure 1. Early-childhood and pre-primary teacher training programmes, by ISCED classification



Source: UIS, author's calculations

Early childhood and pre-primary education teacher programmes are usually entered after successful completion of upper secondary education (an exception being Cambodia which only requires attendance of grade 12 independent of pass or fail) and their duration vary substantially. The modal duration is at 1 year and 3 years although this mostly corresponds to several programmes in two countries (respectively Mauritius and Germany). One country recorded a training programme of 4 years duration (Mali), however this requires further verification as the programme is classified as an ISCED 3 programme and only leads to ISCED 3 attainment. No programme with entry requiring primary education only (as mentioned in an earlier UIS study) was identified.

Figure 2. Early-childhood and pre-primary teacher training programmes, by duration



Source: UIS

Discussion on pre-primary and early childhood education teachers remains substantially limited by the dearth of data available in the UIS database. This should be treated as a particular concern as early childhood and pre-primary education teachers and educators are often neglected in the global discussion despite their critical role in a child's educational journey. Additionally, in many countries, pre-

primary and early childhood education is delivered by private institutions (including community/faith-based institutions) for which it is equally important than for the public sector to ensure that minimum standards of teaching quality are observed.

b) Primary education teachers

Primary education teacher training programmes recorded in the UIS database allow a more detailed analysis with 46 programmes identifiable in total. Programmes which are recorded to train primary teachers as well as secondary teachers are also included (this is the case for El Salvador, Lao PDR, Lesotho, Malaysia, Sierra Leone, Uganda and Uruguay) as no distinction of entrance requirements or duration is made for the level at which teachers are trained to teach. In total, the analysis covers 39 countries as several countries have registered different types of programmes. Lao PDR for instance has two primary teacher training programmes, a mainstream programme and a programme aimed at training teachers in remote areas, with less stringent entrance requirements.

Cross-tabulation of theoretical entry point and duration of the programme is provided in Table 5. Theoretical entry point has been calculated using UIS data on the structure of education systems. Initial data provided by the UIS registered the theoretical entrance age of teacher education programmes, which is not necessarily meaningful from a cross-national perspective.. The theoretical entry point has thus been standardized and recoded for each country. This was done using theoretical entrance age and durations of each ISCED level. For instance, if a country like Seychelles has its upper secondary education cycle starting at 15 and finishing at 18, a teacher training programme accessible at 17 years old would have as theoretical entry point the third year of ISCED 3; ISC3 (3rd year) in Table 5a. In the sample of 46 education programmes for which metadata was available, teacher training programmes are mostly accessed after upper secondary completion (26 programmes in total) although a significant number of programmes can be accessed after completion of lower secondary education (14 programmes). Teacher training programmes last on average 2.3 years, 2 years is the modal value. More than three-quarters (76%) of all primary teacher training programmes have a duration of two years or more.

But these measures of centrality mask a much greater diversity of programmes. Programmes range from short-duration programmes accessed after some years spent in secondary education, such as the *Primary Teaching Certificate* in Pakistan, which can be accessed after the third year of upper secondary and lasts one year, to programmes that are only accessible after successful completion of a Bachelor's degree such as the Post Graduate Education course in Malaysia which trains both primary and secondary school teachers over a year and a half. It has to be noted that Table 5a only includes data for a handful of high-income countries (4 out of the 39 countries). This is a consequence of the lack of detail provided in the ISCED survey. In effect, for high-income countries it is often not possible to confirm the level for which teachers are trained without further information. Most high-income countries have institutions which deliver bachelor's and/or master's degree qualifications for all teachers, independent of the level at which they will teach. With this in mind, a reference table has been provided below (Table 5b). The table provides the distribution of teacher education programmes for level unspecified for high-income countries only. This table can be used as a comparison point under the assumption that most of the level unspecified teacher education programmes also train primary school teachers (together with teacher for other levels).

Table 5a. Primary teacher training programmes, duration and theoretical entry point, all identified programmes

Theoretical entry point \ Duration (years)	1	1.5	2	3	4	Total
Entr. ISCED 3		1	7	4	2	**14**
ISC3 (3rd year)	1			1		**2**
Entr. ISCED 4,5,6	7		8	9	2	**26**
ISC6 (2nd year)			1		1	**2**
ISC6 (3rd year)	1					**1**
Entr. ISCED 7		1				**1**
Total	**9**	**2**	**16**	**14**	**5**	**46**

Note: Identified programmes refer to teacher education programmes which could accurately be identified as training primary education teachers (either through their name, the qualification they deliver or the "note" open field)

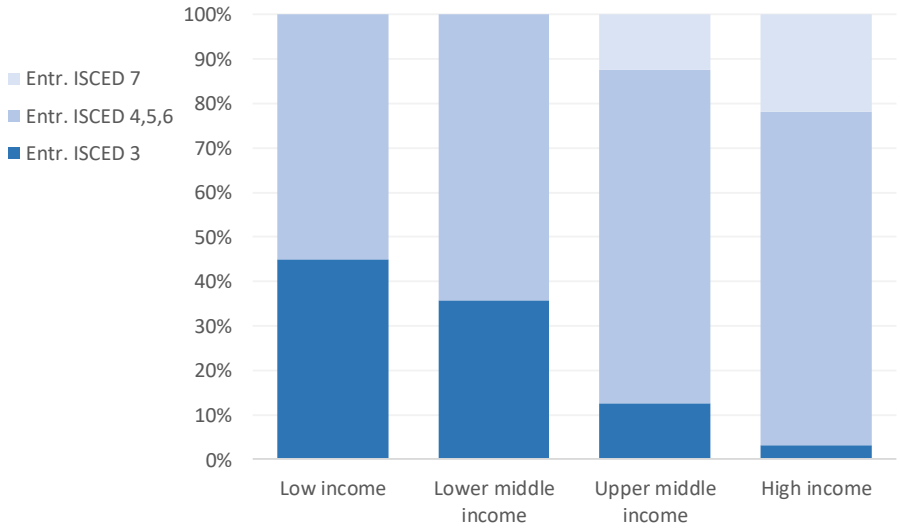
Table 5b. Teacher training programmes, level unspecified, duration and theoretical entry point, high-income countries

Theoretical entry point \ Duration (years)	1	1.5	2	3	4	5	Total
Entr. ISCED 4, 5, 6	1		4	10	5	1	**21**
Entr. ISCED 7	1	2	4				**7**
Total	**2**	**2**	**8**	**10**	**5**	**1**	**28**

Source: UIS, author's calculations

One point apparent from Table 5a is the diversity of entry points and duration that can be found among non-high-income countries. Table 5b, however, potentially_provides insights into how teacher training programmes' entrance requirements evolve as countries become more developed. No teacher training programme is entered before successful completion of upper secondary education and a significant share of programmes are entered after successful completion of an ISCED 6 programme. Figure 3 illustrates at which point in the education cycle teacher training programmes are usually accessed, also taking into consideration the level of economic development of countries. While for low-income countries almost half (45%) of all teacher training programmes can be accessed following successful completion of lower secondary education, this proportion decreases as the level of income increases. Only 13% of programmes offered in upper-middle-income countries can be accessed at upper secondary level, and the earliest access point in high-income countries is at the third year of upper secondary education (Primary Teacher Training Programme in Seychelles). Conversely, high-income countries offer 22% of their programmes only after successful completion of a bachelor's degree, which is not observed in low- and lower-middle-income countries.

Figure 3. Primary teacher training programmes, theoretical entry point by country group



Note: Access points theoretically in a non-entry year (e.g. ISC3-3rd year) have been collapsed to the entrance of the corresponding ISCED level

Source: UIS, author's calculation

Unsurprisingly, a similar pattern is observed with the mapping of teacher training programmes according to ISCED 2011. While overall 61 % of all programmes in the sample are classified as ISCED 5 or above, there are variations by income group. Low-income countries have most (75%) of their teacher training programmes classified as ISCED 3 or 4 (Figure 4). At the other end of the spectrum, high-income countries have 62% of all their programmes classified at ISCED 6 or higher.

Figure 4. Primary teacher training programmes, ISCED level classification by country group

	2	3	4	5	6	7
Low income	5%	40%	35%	15%	5%	0%
Lower middle income	0%	36%	21%	29%	14%	0%
Upper middle income	0%	13%	25%	25%	38%	0%
High income	0%	0%	6%	31%	53%	9%
Grand Total	1%	19%	19%	26%	31%	4%

Note: High income category includes specified (4) and unspecified programmes (28).

Source: UIS, author's calculations

An additional interesting pattern observed is that a number of francophone African countries have recourse to an entrance examination rather than specific achievements (such as a minimum number of O levels) upon completion of lower or upper secondary education, which is a practice observed in anglophone African countries or Caribbean countries. Four of the 16 francophone African countries have a *"concours"* as one of the requirements granting access to primary teacher training programmes.

c) Secondary education teachers

Secondary teacher training programmes are characterized by wider diversity. Fifty programmes, with clear identification of the level for which teachers are trained, have been analyzed. These programmes are offered in 33 countries. These programmes are training teachers for the lower secondary level (18 programmes), upper secondary level (14 programmes), or can include several streams (of same duration), each delivering a qualification to teach a different level (18 programmes). Thus, some countries have programmes where entry requirements, durations and levels of qualification do not change depending on the level teachers will teach. Other countries have two distinct educational programmes for prospective teachers at the lower and upper secondary level. Looking at all programmes together, independent of the target level, the majority of programmes can be entered upon successful completion of upper secondary education. Three out of every five programmes can be accessed after completion of upper secondary education (Table 6). Slightly more than a fifth of the programmes can be accessed following successful completion of a bachelor's degree or a short-cycle tertiary education programme. More strikingly, some programmes can be accessed at the end of lower secondary education. Eight programmes are in this situation, and they usually grant the possibility to teach in vocational lower secondary education.

The pattern for secondary education is closer to what is being observed in high-income countries in Table 5b, with the same modal values for entry point albeit with a shorter modal duration. Five countries (Algeria, Cameroon, Gabon, Madagascar and Turkey) offer upper secondary teacher training programmes of five years duration.

Table 6. Secondary teacher training programmes, duration and theoretical entry point

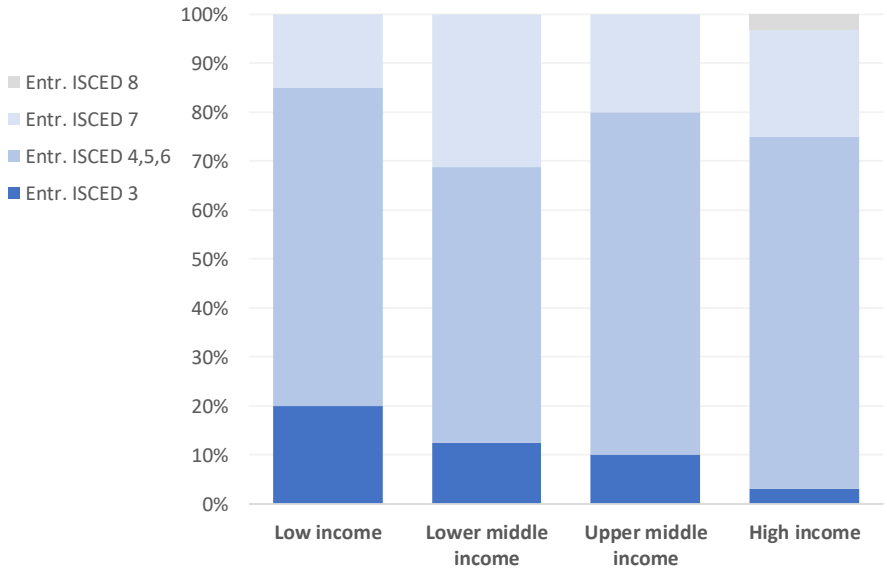
Theoretical entry point \ Duration (years)	1	1.5	2	3	3.5	4	5	Total
Entr. ISCED 3			2	5		1		**8**
Entr. ISCED 4,5,6			9	9	1	5	5	**29**
ISC6 (3rd year)			2					**1**
Entr. ISCED 7	3	1	4					**8**
ISC7 (2nd year)	1		1					**2**
Entr. ISCED 8			1					**1**
Grand Total	**4**	**1**	**19**	**14**	**1**	**6**	**5**	**50**

Note: The table includes all programmes training teachers for secondary education independently of whether they will teach lower or upper secondary.

Source: UIS, author's calculations

The pattern observed for primary school teachers on theoretical entry point is similar albeit less marked. Two key trends are (i) the decreasing importance of programmes accessible after the end of lower secondary education as countries' level of income increases and (ii) the increasing importance of programmes that require at least a bachelor's degree for entry in higher income groups.

Figure 5. Secondary teacher training programmes, theoretical entry point by country group



Note: Access points theoretically in a non-entry year (e.g. ISC6-3rd year) have been collapsed to the entrance of the corresponding ISCED level.

Source: UIS, author's calculations

At the secondary level, teacher training programmes are slightly more homogenous than at the primary level albeit with an increasing share of programmes classified at bachelor's level and master's level the higher the income group. In low-income countries, only 45% of programmes are at the ISCED 6 or above level and there are also 20% of programmes (mostly vocational secondary teachers) which are only classified at ISCED level 3. As a comparison point, secondary teacher training programme are classified as ISCED 6 or higher in 70% of upper-middle-income countries, a situation that does not differ too much from what is being observed among high-income countries (66% of secondary teacher education programmes classified as ISCED 6 or higher).

With regard to other types of entrance requirement, the practice of having an entrance examination is again reserved to francophone African countries (Chad and Madagascar), while selection based on specific grades happens in some anglophone African countries (Mauritius and Sierra Leone). These practices are however a minority of cases as most programmes have neither an entrance examination nor specific grades requirements.

Figure 6. Secondary teacher training programmes, ISCED level classification by country group

	3	4	5	6	7
Low income	20%	5%	30%	20%	25%
Lower middle income	13%	13%	19%	44%	13%
Upper middle income	10%	0%	20%	40%	30%
High income	0%	3%	31%	53%	13%
Grand Total	9%	5%	27%	41%	18%

Source: UIS, author's calculations

2.3 Dimensionality and commonalities of teacher education programmes

One of the main objectives of this paper is to identify commonalities between teacher training programmes to assess the potential for dimensionality reduction and the production of a simplified framework that could enable the comparison of teacher training programmes. As recalled earlier, an earlier UIS study highlighted that at least nine dimensions could be observed for teacher training programmes: pre-requisites for entry, duration, modality, training accent, educational philosophy, length and conditions of probationary or induction periods, additional criteria for admission to teacher education programmes, quality assurance processes/frameworks for programmes, certification or licensing processes.

While this paper is not able to deal with content-related characteristics, or modality of delivery, some reflections can me made about the possibility to classify teacher training programmes according to entry requirements, duration, ISCED classification and qualification awarded at the end of the programme. There is a substantial correlation between ISCED classification of a programme and the qualification awarded at the end, so it is proposed at this stage to drop the latter.

Based on the descriptive statistics presented above, the following reference points can be established:

- There is a higher homogeneity of teacher training programmes among high-income countries. For these countries, teacher training is usually delivered for all levels as part of a common group of education programmes which share entry requirement, duration and ISCED classification. In low- and middle-income countries, there is a significant contrast between primary and secondary education teachers.

- Entrance into teacher training programme happens most often after successful completion of upper secondary education, for both primary and secondary education teachers. But as countries develop, they raise requirements for entry into teacher training programmes. A key difference between primary and secondary teacher training programmes is that even though both can be entered after successful completion of upper secondary, in many cases primary teacher training programmes only classify as ISCED 3 programmes while almost all (nine out of ten) secondary teacher training programmes classify as post-secondary education programmes. The few secondary teacher training programmes which qualify as ISCED 3 are for technical and vocational education teachers.

- Patterns are emerging when income groups are compared. These patterns confirm earlier studies on the characteristics of teacher training programmes. However, this structuration into income group commonalities might render more difficult the development of a global standard of a "trained teacher". Promoting a minimum standard that is too high for education systems in low- and lower-middle-income countries risks stigmatizing these countries even further as they do not have yet the capacity to upgrade their teacher training systems. Thus the typology should probably lean towards a more descriptive rather than prescriptive perspective, or risks to systematically exclude less developed countries.

- All teacher training programmes accessed after successful completion of a bachelor's degree are of short duration (1 or 2 years).

- Primary teacher training programmes are on average shorter than secondary teacher training programmes and mostly qualify as ISCED 4 (post-secondary non-tertiary) or ISCED 5 (short-cycle tertiary) education programmes.

- Entrance examination and grade requirements are the exception rather than the rule and only a handful of countries, mostly in Africa, are using these as entry requirements.

- Programmes which have lower entrance requirements and durations are usually teacher training programmes for specific populations or in vocational education.

- To keep some discriminatory power yet remain simple, the typology would need to consider groupings (within the three dimensions), but the smaller the number of groups the less meaningful the typology will be. Furthermore, there is likely a substantial correlation between the combination of minimum entrance + duration and the ISCED classification of the programme (as they are used to deduce the ISCED classification). So eventually, the ISCED classification of teacher training programmes might be the best summary measure if one were to seek for a single descriptive approach to entry requirements, duration and ISCED classification.

- However, based on the complete list of possibilities identified in the database, a simplified version of possible combinations is proposed in Table 7.

3. Minimum dimensions for categorization and coding

In light of the above, producing an international teacher training programme taxonomy will have to consider the main two sources of variations that are constituted by the level for which teachers are trained and countries' level of development. While the former can be integrated within the taxonomy, the latter cannot but would have to be taken into consideration, should the global education community be willing to develop global minimum standards for teacher training programmes.

Furthermore, teacher training programmes are education programmes in the first place, therefore their ISCED classification already constitutes a differentiating and overarching dimension. It would make sense for a future teacher training programme classification scheme to build on ISCED and further integrate to the ISCED code other dimensions pertaining to teacher training programmes only. A proposal would be to expand the three digits ISCED code to a six digits code; the three last digits potentially coding for the following dimensions:

1. Target level for which teachers are trained to teach

2. Minimum entry requirement as represented by the highest ISCED level completed needed to enter

3. Duration of the programme in years

These three dimensions are proposed on the following basis:

- The first digit is a critical source of variation, as well as an important variable that policy-makers would want to have, to compare teacher training programmes with the same objective.

- The last two digits are recognized among the nine quality differentiating dimensions listed in the literature and are easily measurable and collected, in many cases already available in UIS database.

- Among the nine dimensions listed above, they are ordinal variables and thus enable the discussion on minimum standards.

Eventually, there would be ground to discuss the following additional expansions of the coding scheme:

- Add an initial digit or letter (i.e. 'T') to code for teacher training programmes. This would signal feasibility of expanding classification and comparability to any domain- or profession-specific type of programmes (e.g. specific programmes for nurses, military, police officers, etc.). While not within the scope of the initial discussion, this would position ISCED as a core tool to classify any profession-specific education or training.

- Code for additional criteria for admission (e.g. entrance exams, specific grades, etc.), which are usually available in the ISCED questionnaire. While not ordinal, these would still bring informative background to discussions on teacher training programmes.

- Other dimensions (modality, training accent, educational philosophy, length and conditions of probationary or induction periods, quality assurance processes/frameworks for programmes, certification or licensing processes) could eventually be added, though risking a very long coding scheme.

Table 7 presents the three digits for categorization with their proposed values. Choice was made to keep the same values used in the current ISCED whenever the category corresponds to an existing ISCED level. The first column also proposes a grouping for the initial ISCED classification of teacher education programmes. While there might be some loss of information these grouped categories better reflect the variations observed in the sample. Eventually, one can also revert to the longer three-digit initial ISCED classification of the programme.

Table 7. ISCED-T: Potential categorization of pre-requisite, duration and ISCED classification for teacher training programmes

ISCED classification of the programme (Grouped)	First Digit: Target level	Second Digit: Pre-requisite: Minimum educational level completed	Third Digit: Duration
S: ISCED 2 and 3: Secondary Education	**0: ISCED 0**; Pre-primary	**2:** ISCED 2	**1:** Short (1 year or less)
PS: ISCED 4: Post-secondary non-tertiary	**1: ISCED 1**; Primary	**3:** ISCED 3	**2:** 2 years
TB: ISCED 5,6: Tertiary, bachelor's degree	**6: ISCED 0-1**; Multilevel Pre-primary and primary	**4:** ISCED 4 or some tertiary (ISCED 5 or 6, incomplete)	**3:** 3 years
TM: ISCED 7 and higher: Tertiary, master's degree	**2: ISCED 2**; Lower secondary	**5:** ISCED 5 and 6	**4:** Long (4 years or more)
	3: ISCED 3; Upper secondary	**7:** ISCED 7	
	7: ISCED 2-3; Secondary		
	8: ISCED 1-2-3; Multilevel Primary and Secondary		

Some examples are given below to illustrate how the proposed coding scheme might be used. While quite simple, the scheme enables a partial ordering of teacher training programmes and also provides a rapid understanding of some of the characteristics of the programmes. To take an example at the primary level, the *Diplôme d'Instituteur* in Burundi is a four-year teacher education programme entered at age 16 with the requisite that ISCED 2 is completed. The programme is classified as 354 according to its ISCED mapping. The long coding (including the initial ISCED code) would therefore be 354-124 while the short code would be S-124. In comparison, the *Diplôme d'Instituteur adjoint* in Côte d'Ivoire is also entered after completion of lower secondary but is only 2 years long and is classified as 353 (ISCED 3 completion but without access to tertiary education). The classification would thus be 353-122 or S-122. Both programmes have the same entrance requirements and classification and only the durations differ, one would therefore be able to position both programmes in relation to each other should this be needed.

At the secondary level, Lesotho delivers all its teacher certificates through its teacher training delivered at the College of Education. According to the ISCED mapping data, all programmes have the same entrance requirements (Completion of ISCED 3) and duration (3 years). These programmes are initially classified as ISCED 554. Using ISCED-T the classification would become 554-833 or TB-833.

Examples:

	Burundi	*Diplôme d'Instituteur: Enseignement post fondamental pédagogique, theoretical entrance age 16, trains teacher for the primary level and has a theoretical duration of 4 years. Initial ISCED classification 354.*					
	Initial ISCED mapping				Target level	Min. ISCED level required	Duration
Long	3	5	4	-	1	2	4
Short	S			-	1	2	4

	Côte d'Ivoire	*Diplôme d'Instituteur adjoint: Formation des Instituteurs adjoints, theoretical entrance age 16, trains assistant teachers for the primary level and has a theoretical duration of 2 years. Initial ISCED classification 353.*					
	Initial ISCED mapping				Target level	Min. ISCED level required	Duration
Long	3	5	3	-	1	2	2
Short	S			-	1	2	2

	Lesotho	*Primary Education teacher diploma and Secondary Education teacher certificate: College of education - Teacher training, theoretical entrance age 18, trains teachers for the primary and secondary level and has a theoretical duration of 3 years. Initial ISCED classification 554.*					
	Initial ISCED mapping				Target level	Min. ISCED level required	Duration
Long	5	5	4	-	8	3	3
Short	TB			-	8	3	3

Finally, to support the discussion on minimum benchmarks, an overview of the global distribution of identified teacher training programmes is given in Figure 7. The overview gives the distribution of teacher training programme for each target level. For instance at the primary level, 17% of programmes in the sample are classified as ISCED 2 or 3, accessible after completion of ISCED 2 and of two years duration (ignoring multi-level programmes this would correspond to ISCED-T: S-122). At the other end of the spectrum, another 17% of all primary teacher training programmes are classified as ISCED 5 or 6, accessible after completion of ISCED 3 and of 3 years duration (TB-133).

Figure 7: Distribution of teacher training programmes for each target level by Initial ISCED classification, minimum entrance requirements and duration.

Initial ISCED (short)	Third Digit (duration)	First digit: Target level[1]														
		Pre-primary teachers n= 13					**Primary teachers** n= 46					**Secondary teachers** n= 50				
TM	4+	0	0	0	0	0	0	0	0	0	0	0	10	0	0	0
	3	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0
	2	0	0	0	0	0	0	0	0	0	0	0	0	0	6	4
	1-	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2
TB	4+	0	0	0	0	0	0	4	0	0	0	0	12	0	0	0
	3	0	15	0	8	0	0	17	0	0	0	0	18	0	0	0
	2	0	8	0	0	0	0	9	0	0	0	0	12	0	6	0
	1-	0	0	0	0	0	0	0	0	4	0	0	0	0	6	0
PS	4+	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0
	3	0	0	0	0	0	0	7	0	0	0	2	0	0	0	0
	2	0	8	0	0	0	0	9	0	0	0	0	6	0	0	0
	1-	0	8	23	0	0	0	15	0	0	0	0	0	0	0	0
S	4+	8	0	0	0	0	4	0	0	0	0	2	0	0	0	0
	3	8	0	0	0	0	9	0	0	0	0	8	0	0	0	0
	2	0	0	0	0	0	17	0	0	0	0	4	0	0	0	0
	1-	15	0	0	0	0	0	2	0	0	0	0	0	0	0	0
		ISCED2	**ISCED3**	**ISCED4 or some tertiary**	**ISCED56**	**ISCED7**	**ISCED2**	**ISCED3**	**ISCED4 or some tertiary**	**ISCED56**	**ISCED7**	**ISCED2**	**ISCED3**	**ISCED4 or some tertiary**	**ISCED56**	**ISCED7**
		2	3	4	5	7	2	3	4	5	7	2	3	4	5	7
		Second digit: Minimum education level completed														

Notes:

1. 1. Primary and Secondary teachers training programmes include multi-level programmes, the first digit cannot be fully represented.
2. 2. See Table 7 and coding scheme for legend details.

Source: UIS, author's calculation.

4. Quality of the international data collection on teachers

This second part looks at the current status of the main international data collection on teachers; the annual survey carried out by the UNESCO Institute for Statistics. While there are a number of data collections on teachers, the UIS data collection remains the official global public good for the production of indicators to monitor SDG target 4.c: "by 2030, substantially increase the supply of qualified teachers, including through international cooperation for teacher training in developing countries, especially least developed countries and small island developing States". UIS data collection is done in collaboration with UNESCO Member States and as such includes official endorsement of the final indicators by the countries, as well as a rigorous quality control process at the UIS with countries' focal points and UIS field offices.

4.1 Missingness patterns

i. Data

The data provided by the UIS are analyzed in this section in a way that integrates the UIS data collection protocol. The data collection protocol of the UIS is as follows:

1 -UIS Questionnaires are sent to countries.

2 Countries fill in UIS questionnaires and return completed questionnaires which are immediately entered in a first database, the REP database (see the description below), which contains raw data as reported by countries before any change by the UIS.

3 UIS staff verifies the consistency of data returned by countries. This consistency is verified both internally (consistency between cells within the questionnaire, e.g. male + female = total) and externally (consistency between current questionnaire values and previous data reported by the country). At the end of this process, the UIS gets back to the country with a report detailing issues identified and pending questions. These are subsequently addressed by countries' focal points until final agreement is reached between the UIS and the country.

4 Any change that occurs after the questionnaire is returned and entered in REP is made in a second database, OBS. That is, data from REP are initially transferred to OBS and subsequently modified upon identification and resolution of issues.

5 Once data in OBS are considered cleaned, they are transferred into the EST database. The EST database is the database used to feed UIS' indicators calculation. Additional estimates can be done by the UIS in EST (using time series or observed distribution patterns; see below) and additional data points found in official national documents can also be added at this stage.

6 -In the EST database, the UIS can qualify some data points as unpublishable if the data points generate inconsistencies in indicator values.

In light of this process, looking at the status of a given data point alongside its journey throughout the three databases provides substantial information about:

- When problems occur,

- What could be the reason of problems occurring,
- Which actor of the data collection chain should be targeted to improve identified problems.

The three databased are described below:

REP is the database that includes the original data as reported by the country to the UIS. For the purpose of this paper, the data analyzed are those reported by countries in tables 9 ("Number of classroom teachers by teaching level of education, employment status, type of institution and sex") and 10 ("Number of classroom teachers by employment status, qualified and trained status, teaching level of education, type of institutions and sex") of UIS questionnaire A ("Survey of Formal Education, students and teachers (ISCED 0-4)").

OBS is the database that correspond to clean and validated data by UIS staff and also includes corrections and modifications to the data made by the UIS following exchanges with the country focal point. OBS can also include data added by the UIS after the data collection, on the basis of existing public documents.

EST also includes additional estimations by the UIS to ensure indicator calculations and is the database used to feed the UIS public database. It is also in the EST database that data are marked unpublishable (usually estimates that are based on a heuristic for the sake of calculating averages but cannot be used as data points) should there be some issues regarding indicator values. Issues can include mismatch between population data and education data, out-of-trend variation when compared to previous years, or internal inconsistencies of indicators values.

The three databases were provided by the UIS for all teacher related data points collected for the school year ending in 2013 to the school year ending in 2017. The data come from two tables of UIS questionnaire A; Table 9, which collects data on simple headcount and full-time equivalent teachers by level, sex and institution, and Table 10, which collects the same information but for qualified and trained teachers. These two tables include a total of 330 potential data points (values) on teachers. Table 8 below describes the dimensions that are associated with each data point. Non-discriminating dimensions provided by the UIS were removed to keep only relevant dimensions (see annex). The UIS provided the data in five separate files (one for each year) which were appended in a single data file totalling 351,450 distinct records for which the REP, OBS and EST value is reported or NA (for not available or missing). Each record corresponds to a unique EMC_ID (ID for a data point in the UIS database) for a given country in a given year (330 x 213 countries x 5 years=351,450). Each record contains all the dimensions included in Table 8 below, as well as UIS-specific qualifiers (MQ_ and MG_) and data processing comments added by UIS staff (COMMENTS_). Data processing was done by the author in R and visualization in Excel.

Table 8. Dimensions associated with each data point

VAR	VALUE	RECODE_VAL	DEFINITION
ISCP11_CAT	ISC_CAT5	Voc	Vocational / professional programmes
ISCP11_CAT	_T	All	All educational programmes
ISCP11_CAT	ISC_CAT2	Preprim	Pre-primary education
ISCP11_CAT	ISC_CAT4	General	General / academic programmes
ISCP11_CAT	ISC_CAT1	ECE	Early childhood educational development
STAT_UNIT	NRTEACH	NRTEACH	Newly Recruited teachers
STAT_UNIT	TEACH	TEACH	Classroom Teachers
STAT_UNIT	TRTEACH	TRTEACH	Trained Teachers
STAT_UNIT	TEACH_ISTR_12MTH	TEACH_ISTR_12MTH	Teachers who received training while in-service during the twelve months preceding the end of the reference academic year
STAT_UNIT	QTEACH	QTEACH	Qualified Teachers
SEX	M	M	Male
SEX	F	F	Female
SEX	_T	_T	Total
SECTOR	INST_T	INST_T	Total
SECTOR	INST_PUB	INST_PUB	Public
SECTOR	INST_PRIV	INST_PRIV	Private
ISC11_LEVEL	ISC0	ISC0	ISCED 0 (ECE and Pre-primary)
ISC11_LEVEL	ISC1	ISC1	ISCED 1 (Primary Education)
ISC11_LEVEL	ISC2	ISC2	ISCED 2 (Lower secondary Education)
ISC11_LEVEL	ISC2_3	ISC2_4	ISCED 2 and 3 (Secondary Education - lower + upper)
ISC11_LEVEL	ISC3	ISC3	ISCED 3 (Upper Secondary Education)
ISC11_LEVEL	ISC4	ISC4	ISCED 4 (Post-secondary non-tertiary education)
TABLE_IDENTIFIER	A9	A9	UIS Questionnaire A Table 9
TABLE_IDENTIFIER	A10	A10	UIS Questionnaire A Table 10

ii. Data patterns

This analysis thus examines missing patterns across the three databases. A data point is considered missing if there is no numerical value in EM_FIG_ (the variable containing the actual value for a given data point). A record's missingness type can thus be described by looking at the value of its data point across the three databases (REP, OBS, EST). Missingness patterns are therefore a function of the possible combinations for a given record. Table 9 and Figure 8 describe the missingness patterns found in the UIS database (all years considered) and their distribution. Each EM_FIG_ is described by a binary outcome; 0 if NA (missing) and 1 if there is a numerical value (non-missing). Patterns then corresponds to the combination of these outcomes across all three databases. For instance, pattern 3; "Misreported and corrected", corresponds to: *EM_FIG_REP: Missing + EM_FIG_OBS: Non-missing + EM_FIG_EST: Non-missing* (011), and is found for 4.1% of all records in teachers related data. Missingness patterns have been

labelled according to the author's understanding of the UIS data collection process and might require amendment.

Overall the patterns observed are the followings:

- For the school years ending in 2013 to 2017, two-third of countries' teacher data points have a full missing record (pattern #1; *Unreported and unestimated*). That is, no data was reported, or found in a public document or estimated.

- 22% of all countries' teacher data points have however been through the whole data processing cycle (pattern #2; Ideal scenario). Countries' focal points have reported data in UIS questionnaire A tables 9 and 10 and data have been reviewed, cleaned and validated as inputs into indicator calculations by the UIS.

- An additional 4% of all teacher data points get an OBS value (correction by the UIS or value found in other document) and then get an EST value. This despite either non-reporting or misreporting by the country.

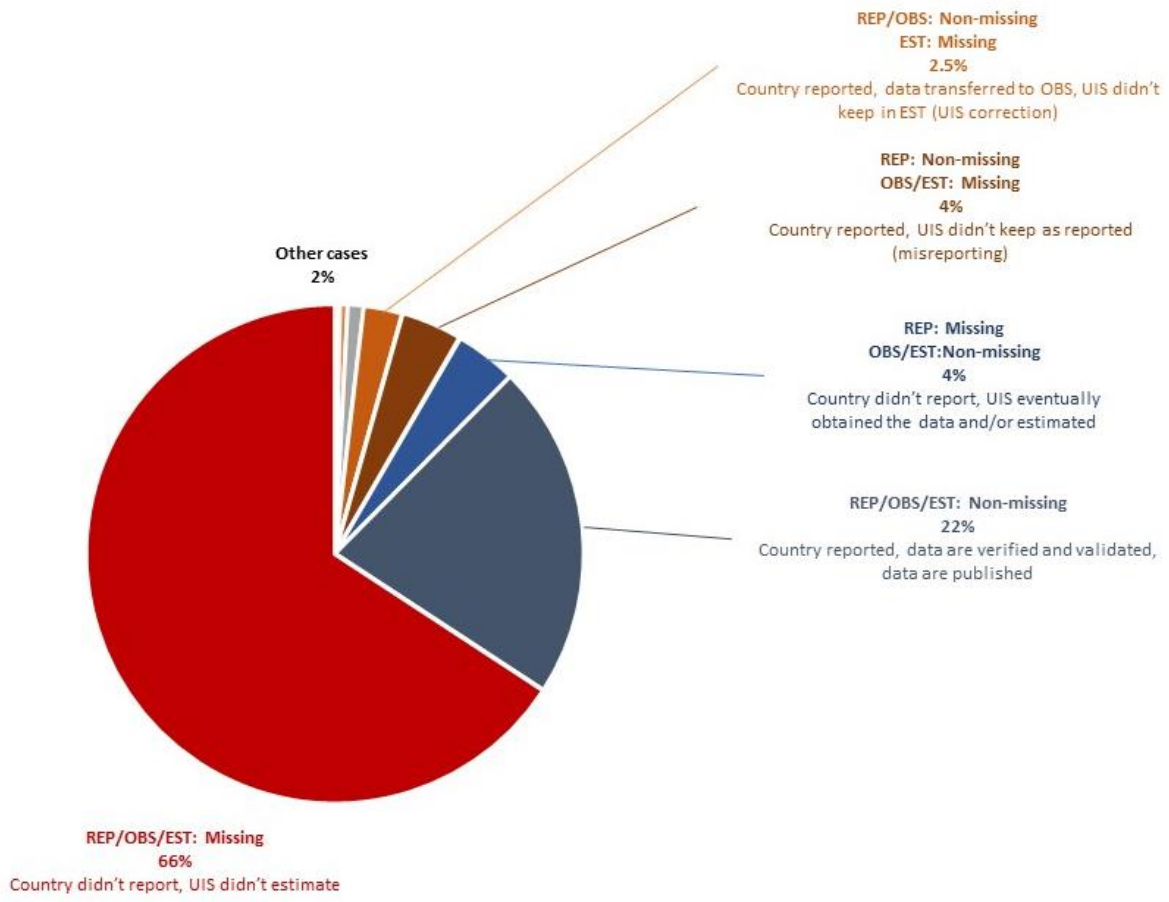
- Lastly, 6.5% of data points have a REP value, that is the country provided data in their first submission of the questionnaire yet, there is no value in the EST series, signalling misreporting by the country (4% of cases where there is a REP value but neither OBS nor EST) or deletion by the UIS in later stages of the process (2.5% of cases where there is a REP and an OBS value but no EST).

Table 9. Missingness patterns in UIS database, all teacher-related data points, 2013-2017

Pattern #	Pattern label	Description	EM_FIG_REP	EM_FIG_OBS	EM_FIG_EST	# missing values	Share of total cases	Number of cases
1	*Unreported and unestimated*	REP/OBS/EST: Missing	0	0	0	3	65.8%	231248
2	*Ideal scenario*	REP/OBS/EST: Non-missing	1	1	1	0	21.8%	76441
3	*Misreported and corrected or found elsewhere*	REP: missing, OBS/EST: Non-missing	0	1	1	1	4.1%	14323
4	*Misreported and deleted*	REP: Non-missing, OBS/EST: missing	1	0	0	2	4.0%	13940
5	*Reported and corrected but deleted in EST*	REP/OBS: Non-missing, EST: missing	1	1	0	1	2.5%	8931
6	*Estimated only*	REP/OBS: missing, EST: Non-missing	0	0	1	2	1.0%	3684
7	*Found elsewhere/Estimated in OBS but deleted in EST*	REP: missing, OBS: Non-missing, EST: missing	0	1	0	2	0.5%	1777
8	*Misreported and deleted but Estimated*	REP: Non-missing, OBS: missing, EST: Non-missing	1	0	1	1	0.3%	1106
	Total		**251032**	**249978**	**255896**	**756906**		**351450**

Source: UIS database, author's calculation

Figure 8. Missingness patterns in UIS database, general overview

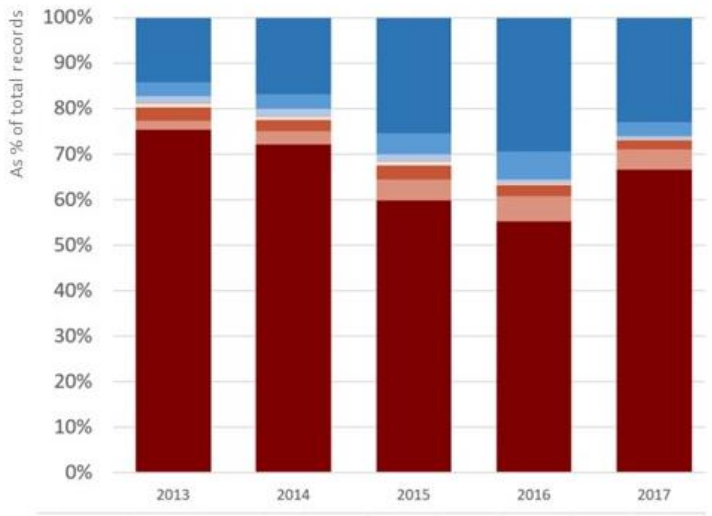


Source: UIS database, author's calculations

These figures represent the average missingness pattern over five years. Looking at data by single year (Figure 9) shows that the general trend is that of an improvement with the proportion of pattern #1 (unreported and unestimated) decreasing from 75% for the school year ending in 2013 (so likely collected in 2014) to 55% for the school year ending in 2016 (data collected in 2017). It is uncertain whether the deterioration observed for the school year ending in 2017 is an actual deterioration or a temporary situation associated with delays in data processing and verification. If this is an observed trend then this will require further analyses of recent changes in UIS data collection but it is more likely to simply reflect the partial status of the database as the data was shared in December 2018.

And while the substantial amount of unreported or misreported records is particularly worrying for the success of monitoring SDG target 4.c, exploring further existing differences in reporting might help the global education community and the UIS in particular to target specific actions in improving the international data collection on teachers.

Figure 9. Missingness patterns, 2013 to 2017



#	Label	2013	2014	2015	2016	2017
1	Unreported and unestimated	75.3	72.2	59.8	55.2	66.6
4	Misreported and deleted	2.1	3.0	4.6	5.6	4.5
5	Reported and corrected but deleted in EST	2.9	2.4	3.1	2.4	1.9
7	Found elsewhere/Estimated in OBS but deleted in EST	0.8	0.5	0.7	0.3	0.2
2	Ideal scenario	14.2	16.8	25.4	29.5	22.9
3	Misreported and corrected or found elsewhere	3.1	3.3	4.7	6.2	3.2
6	Estimated only	1.3	1.6	1.3	0.6	0.4
8	Misreported and deleted but Estimated	0.3	0.2	0.3	0.3	0.4

Note: The total number of records for a single year is 70,290

Source: UIS database, author's calculation

4.2 Reporting issues: disaggregation of teacher-related data

Table 10 shows variations in missing patterns by each of the dimensions available in the database provided by the UIS plus an additional dimension added for income group (World Bank classification as of June 2018). The difference in the proportion of unreported data points is perhaps the most striking between Tables 9 and 10 of UIS questionnaire A. Table A9, which collects data on teachers independently of their trained/qualified status, has 54% of records belonging to pattern #1 (unreported and unestimated) while table A10, which asks for teachers by trained/qualified status, has 73% records for the same pattern. All in all, 32% of data points in table A9 follow the ideal scenario of data reported, cleaned and validated as inputs into indicators, while this is the case for only 16% of data points in table A10.

Table 10. Missingness patterns by type of data points, selected dimensions, all teacher-related data, 2013-2017

	1 Unre-ported and unesti-mated	2 Ideal scenario	3 Misreported and corrected or found elsewhere	4 Misreported and deleted	5 Reported and corrected but deleted in EST	6 Estimated only	7 Found elsewhere/Estimated in OBS but deleted in EST	8 Misreported and deleted but Estimated	Number of records	Total reported	Total unreported	Total available in EST
	(% of Number of records)									(% of Number of records)		
Total	66	22	4	4	3	1	1	0	*351450*	29	71	27
A10	73	16	4	4	2	1	0	0	*223650*	22	78	21
A9	54	32	5	4	3	1	1	0	*127800*	40	60	38
INST_PRIV	64	23	4	4	3	1	0	0	*95850*	30	70	29
INST_PUB	63	24	4	4	3	1	0	0	*95850*	31	69	29
INST_All	68	19	4	4	2	1	1	0	*159750*	26	74	25
Total	66	22	4	4	2	1	1	0	*117150*	29	71	28
F	66	22	4	4	3	1	0	0	*117150*	28	72	27
M	66	22	4	4	3	1	1	0	*117150*	28	72	27
All	66	22	4	4	3	1	1	0	*210870*	29	71	27
General	67	20	3	6	3	1	0	0	*35145*	29	71	24
Voc	71	17	3	5	2	1	0	0	*35145*	24	76	21
ECE	59	24	11	2	1	2	1	0	*35145*	28	72	37
Preprim	65	26	2	3	3	1	0	0	*35145*	32	68	29
ISC0	63	25	6	3	2	1	1	0	*105435*	29	71	32
ISC1	60	33	1	2	3	1	0	0	*35145*	38	62	35
ISC2	69	20	1	6	3	1	0	0	*35145*	29	71	22
ISC2_3	67	18	4	5	3	1	0	0	*105435*	26	74	24
ISC3	68	19	1	5	5	0	0	0	*35145*	30	70	21
ISC4	69	17	7	4	1	1	1	0	*35145*	22	78	25

Table 10 (continued). Missingness patterns by type of data points, selected dimensions, all teacher-related data, 2013-2017

	1 Unre-ported and unesti-mated	2 Ideal scenario	3 Misreported and corrected or found elsewhere	4 Misreported and deleted	5 Reported and corrected but deleted in EST	6 Estimated only	7 Found elsewhere/Estimated in OBS but deleted in EST	8 Misreported and deleted but Estimated	Number of records	Total reported	Total unreported	Total available in EST
	(% of Number of records)									(% of Number of records)		
NRTEACH	87	7	2	3	1	0	0	0	*31950*	11	89	10
QTEACH	69	19	4	4	2	1	0	0	*95850*	26	74	24
TEACH	43	40	6	5	4	1	1	0	*95850*	49	51	48
TEACH_ISTR_12MTH	92	4	2	2	0	0	0	0	*31950*	6	94	6
TRTEACH	70	18	4	4	2	1	0	1	*95850*	24	76	23
Arab States	49	30	6	5	7	1	1	0	*33000*	43	57	38
Caribbean	68	20	4	5	3	0	0	0	*39600*	28	72	24
Central and Eastern Europe	75	18	3	1	2	0	0	0	*34650*	22	78	21
Central Asia	58	25	4	6	6	0	1	0	*14850*	37	63	29
East Asia	56	30	5	4	3	2	1	0	*28050*	36	64	36
Latin America	61	30	3	4	1	1	0	1	*31350*	35	65	35
N. America/W. Europe	83	11	3	0	0	2	0	0	*51150*	12	88	16
Pacific	73	15	4	4	3	1	1	1	*28050*	23	77	20
South and West Asia	49	27	7	8	6	2	1	0	*14850*	41	59	36
Sub-Saharan Africa	64	23	5	5	1	1	0	0	*75900*	30	70	29
High income	76	16	4	1	2	1	0	0	*128700*	19	81	21
Low income	63	23	5	5	2	1	0	1	*56100*	30	70	30
Lower middle income	57	25	5	6	5	1	1	0	*75900*	36	64	31
Upper middle income	61	26	3	6	2	1	0	0	*90750*	35	65	31

Source: UIS database, author's calculation

If one looks in detail at the variations in missing patterns, a few other points can be made. Reporting data by sex and/or type of institution does not seem to carry additional issues. The share of records in pattern #1 or #2 (respectively unreported and ideal scenario) are nearly identical, whether one looks at sex-disaggregated values or total values, and there are more records belonging to the unreported pattern for data points pertaining to "All institutions" types than for data points coding for either public or private (indicating that the missing pattern is associated with non-institution-disaggregated data points). Records on teachers in technical and vocational programmes exhibit slightly higher tendency of non-reporting (71% vs 67% for general programmes) and data at the pre-primary level stand out as having the highest amount of post-reporting transfers and corrections. Eleven per cent of data on teachers in early childhood education belong to missing pattern #3; the country did not report any data yet the OBS and the EST series include a value for the corresponding data point.

Unsurprisingly, countries have more difficulties to report for education levels other than primary education. One third of all records on primary education teachers correspond to pattern #2 while only 20% or less of all records for secondary education teachers and post-secondary education non-tertiary are of the same pattern.

Eventually, the issue that is perhaps of greatest interest in light of SDG 4 reporting is the critical difference in reporting patterns for trained and qualified teachers against simple teacher headcounts. While data points which require simple teacher headcounts (without specifying whether they are trained or qualified) exhibit only 43% of pattern #1, this proportion is 26 percentage points higher as soon as the reporting involves any type of characterization of the trained/qualified status of teachers. If one adds patterns #2, 3, 6 and 8, less than one quarter of all data points on trained and qualified teachers eventually ends up as input into indicator calculation. That is, indicator 4.c.1 (which uses both total number of teachers and the number of trained teachers) can potentially be calculated for only a fifth of all countries.

Table 10 also highlights the critical issue of severe underreporting of data on in-service training and newly recruited teachers. While both types of data can be extremely useful and are needed if one wants to report on the full set of thematic indicators (4.c.6 and 4.c.7), reporting patterns for these data points show worrying trends. Only 6% and 10% respectively of all data points on in-service training and newly recruited teachers end up in the EST series. Prospects for calculating, attrition rates and percentages of teachers who received in-service training in the last 12 months by type of training are jeopardized by this critical gap in the main international database on teachers.

4.3 Reporting issues: regions and income groups, technical capacity or political willingness to report

Given the trends above, one might also wonder whether these missingness patterns are associated with specific country groups or regions. Income groups have been added to the database to explore whether the issue of missingness is associated with variation in countries' wealth (which is potentially associated with stronger national statistical systems and EMIS). It is however interesting to note that high-income countries are the countries where the frequency of pattern #1 is the highest. In fact, 76% of records belong to this pattern for high-income countries while this proportion decreases to 63% in low-income countries, 61% in upper-middle-income countries and 57% in lower-middle-income countries. This trend is confirmed by patterns at the regional level. The highest share of unreported and unestimated records

is found for countries from North America and Western Europe, with 83% of all records belonging to this pattern. Eventually, only 16% of all possible records see a usable value in the EST database, a paradoxical situation when even sub-Saharan African countries manage to report on 30% of all records. Other regions with a high level of non-reporting on teachers include Central and Eastern Europe (75% in pattern #1), the Pacific (73%) and the Caribbean (68%).

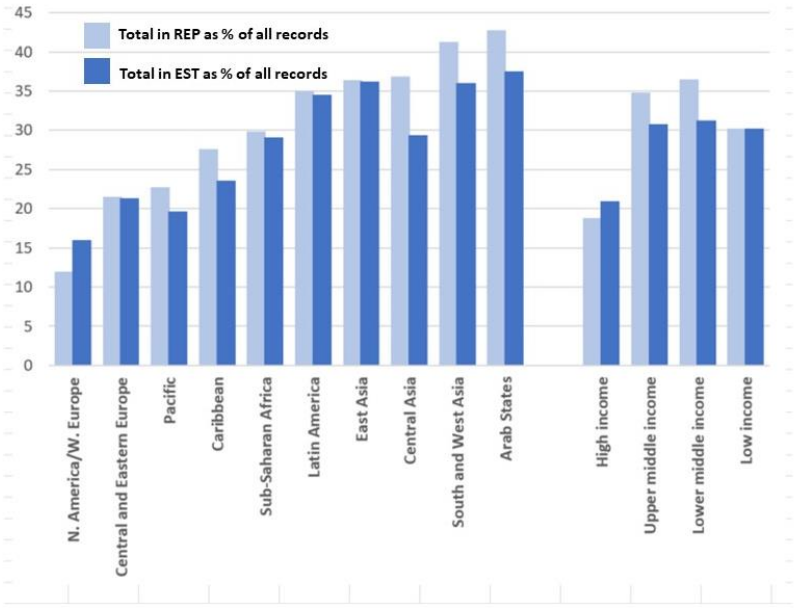
These patterns indicate two issues that can partially explain the non-response rate to the international data collection on teachers. The first one is more related to global governance mechanisms in data collection. High-income countries (and potentially upper-middle-income countries) might not feel compelled to respond to international data collection on teachers because there is an ongoing discussion about definitions of indicators for trained and qualified teachers. The issue and disagreement is so critical for high-income countries that these countries simply refuse to provide these key data. In the case of high-income countries the data collection is not operated by the UIS but is a joint effort between the UIS, the OECD and Eurostat (UNESCO Institute for Statistics, OECD, & Eurostat, 2018a). Data are collected in the *UOE questionnaire on education and training statistics covering educational personnel* (UNESCO Institute for Statistics, OECD, & Eurostat, 2018b) and thus differ from the regular questionnaire sent by the UIS. There is misalignment between the data required to calculate all SDG target 4.c indicators and the data that are collected in the UOE questionnaire. In particular, newly recruited teachers, in-service training and trained/qualified status are not collected at all, making it impossible to calculate any of the indicators listed in either the global or the thematic monitoring framework on teachers. When 49 countries out of 213 listed in the UIS database participate in the UOE data collection, it already means that 23% of all countries will not have a value for any of the SDG target 4.c indicators. The reluctance of developed countries to report these data at international level comes also from the difficulty of having agreed international definitions of the concepts of trained and qualified teachers.

The issue with high-income countries is thus rather an issue of global coordination (between the three international agencies) and potentially of political willingness of developed countries to lead by example and report even data that might seem irrelevant from their perspective. The MDG agenda was criticized for being overly focused on developing countries (Unterhalter, 2013) and its successor, the SDG 4 agenda, was created with the aim to be universally relevant. However, designing data collections that exempt developed countries from appearing in the international monitoring exercise sends the opposite signal and might disserve any future effort to incentivize developing countries to participate in international data collection.

The second issue, associated to high nonresponse rates pertains to national capacity to respond to UIS questionnaires as data collections become more and more complex but national skills or available time are not increasing proportionally to the data collection burden. It is not infrequent in Small Islands Developing States (SIDS) to have a single officer in charge of all education-related data issues. The UIS data collection already includes potentially six questionnaires, adding this to the main task of overseeing and implementing national education data collections it is easy to understand why SIDS have potentially more difficulties to fill in UIS tables on teachers. This capacity issue is also confirmed when the percentage of reported records is compared to the proportion of records which have a value in the EST database. In all developing regions, there is a substantial reduction in coverage when one compares the proportion of data points reported by country focal points and the final amount of data points available in the EST database. The trend is opposite for high-income countries and countries from North America and Western Europe. For North America and Western Europe, there is a 33% gain between what is being reported and what is eventually available in the EST database. However, for regions like the Caribbean

and the Pacific the loss of data corresponds to 15% and 14% respectively once the data have been verified and validated by the UIS. Lower-middle-income countries are the countries with the biggest data loss, with 14% fewer data points in the EST database in comparison to the REP database. This is interesting compared to low-income countries, which see no reduction in their coverage between the two data series, albeit with lower levels of coverage. In comparison, Lower-middle-income countries report 20% more data points than low-income countries but their reported values are more often corrected (or deleted) than low-income countries, leading to very similar coverage rate eventually.

Figure 10. Reported data points versus final coverage in the EST database, by region



Source: UIS database, author's calculations

5. Analysis of reported data, consequences for reporting on SDG target 4.c

In this section, further analysis of reported data is made to highlight some issues beyond missing data. It will look in particular into variations of countries reporting on trained and qualified teachers. While the UIS requests countries to provide both the number of trained teachers and the number of qualified teachers, the difference might not be obvious for respondents unfamiliar with the topic. Consequently, correctly reporting according to international definitions set by the UIS remains a difficult exercise for national staff in Ministries of Education. It is also important to bear in mind that key definitions for the UIS data collection are not included in the questionnaire itself. Definitions are provided in a 40-page manual (UNESCO Institute for Statistics, 2018a) which countries' focal points theoretically need to master before filling in the UIS questionnaire. Such practice is likely to be more error prone than including either built-in help or a glossary at the end of the questionnaire, especially in cases where questionnaires are sent around by the focal point to several offices in charge of various aspects of education.

The definitions provided by the UIS are as follows:

"A **qualified** teacher is one who has the minimum academic qualifications necessary to teach at a specific level of education in a given country. This is usually related to the subject(s) they teach."

"A **trained** teacher is one who has fulfilled at least the minimum organized teacher-training requirements (pre-service or in-service) to teach a specific level of education according to the relevant national policy or law. These requirements usually include pedagogical knowledge (broad principles and strategies of classroom management and organization that transcend the subject matter being taught - typically approaches, methods and techniques of teaching), and professional knowledge (knowledge of statutory instruments and other legal frameworks that govern the teaching profession). Some programmes may also cover content knowledge (knowledge of the curriculum and the subject matter to be taught and the use of relevant materials)." (UNESCO Institute for Statistics, 2018a, p. 26)

While the distinction appears important from a policy planning perspective, it is possible that both the training and the academic qualification happen in the same programme, potentially making it difficult to distinguish for national statisticians in Ministries of Education. It is therefore important to explore the extent to which countries manage to distinctly report trained and qualified teachers in UIS questionnaire A (recalling here that UOE countries do not report on these two critical aspects). Figure 11 details the distribution of countries reporting on qualified and trained teachers, by region and for the year 2016, the year with the best coverage in the sample. Only reported data were retained and these distributions have to be understood in light of the high levels of nonresponse (see above and Figure 12). For instance, little can be said for North America and Western Europe, as data are missing both for qualified and trained teachers in over 95% of cases.

One finding from visual exploration of the data in Figure 11 is that the distinction between qualified and trained teachers does seem to make sense for country focal points as in a substantial number of cases, country focal points reported different values for qualified and trained teachers. In Latin America in particular, 70% of the records (where both trained and qualified teachers are asked to be reported)have a different value for trained teachers and qualified teachers (all levels). The distinction seems however less significant at the pre-primary level, as the share of cases where both values are the same increases substantially, representing between 26% (Central and Eastern Europe) and 64% (South and West Asia) of all cases where the distinction between qualified and trained is made (Figure 11).

Figure 11a. Countries reporting to the UIS, qualified vs trained teachers, by level, 2016, reported data only, all levels

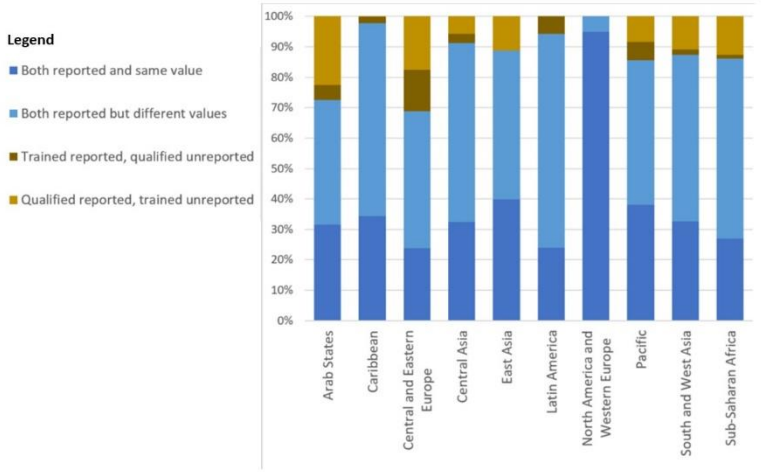
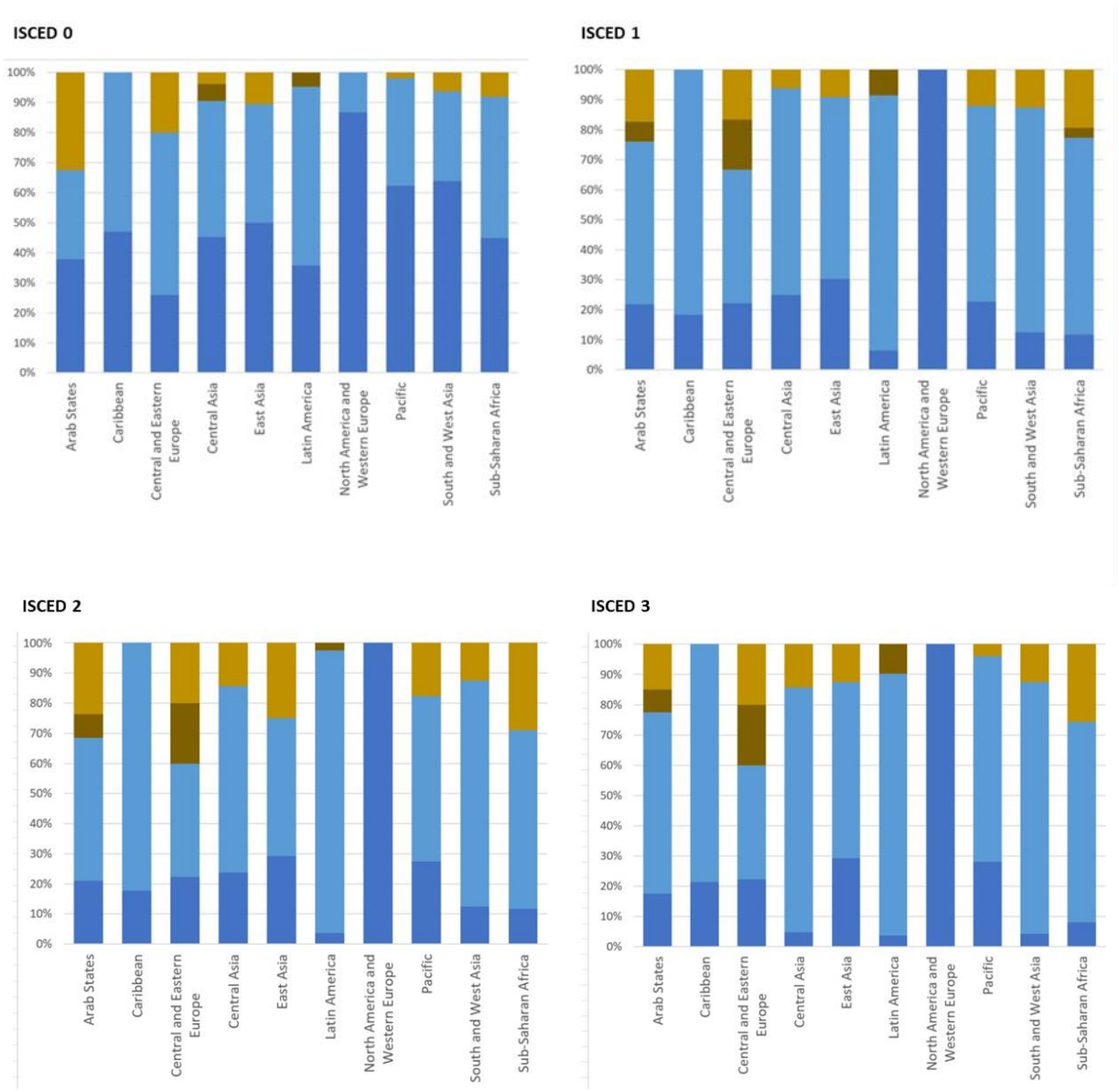
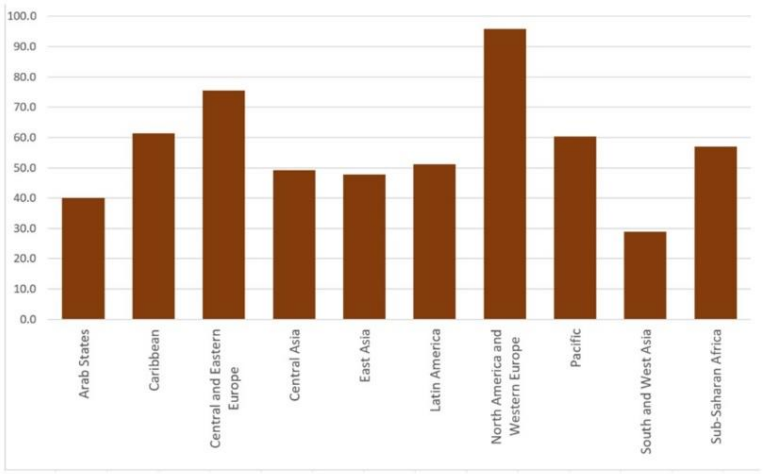


Figure 11b. Countries reporting to the UIS, qualified vs trained teachers, by level, 2016, reported data only, by level



Source: UIS database author's calculation

Figure 12. Proportion of simultaneously missing trained and qualified teachers data points, by region



Source: UIS database, author's calculation

It is difficult to draw conclusions about the added value of the distinction between trained and qualified teachers. Under such low levels of reporting, table 11 illustrates the fact that the share of total reporting which corresponds to either a similar value for trained and qualified teachers, or reporting on only one of the two categories oscillates between 70% (ISCED 0, Arab States and South and West Asia) and 6% in the extreme case of ISCED 2 teachers in Latin America but with values often exceeding 40 or 50% (e.g. all regions except Latin America and the Caribbean and Central and Eastern Europe at ISCED 0). This indicates either that teacher training programmes provide both academic and pedagogic credentials or that countries are not able to make the distinction in these cases and thus choose either to report the same value for both trained and qualified teachers or choose to report only on one of the two categories.

Table 11. Share of total reporting where value for qualified and trained teachers is the same or where only one of the two categories is reported, 2016, by level of education and by region

	ISCED 0	ISCED 1	ISCED 2	ISCED 3	ISCED 2 and 3
Arab States	70	46	53	40	57
Caribbean	47	18	18	21	46
Central and Eastern Europe	46	56	62	62	54
Central Asia	55	31	38	19	20
East Asia	60	39	54	42	40
Latin America	41	15	6	13	15
N America/W Europe	87	100	100	100	100
Pacific	64	35	45	32	49
South and West Asia	70	25	25	17	39
Sub-Saharan Africa	53	34	41	34	28

Source: UIS database, author's calculations

6. Unpublished data, processing comments

While the REP database is the database which enables looking into countries' focal points reporting, the EST database allows a more detailed look at some of the processing and validation choices made by the UIS. Two variables provided by the UIS are of interest here:

- MQ_EST: which codes for the qualifiers added by UIS staff. These qualifiers are taken into account during the indicators calculation process. They indicate to the UIS software whether the data point is of good quality for calculation. Example codes are:
    - 2: National estimate. Estimate provided by a country to the UIS.
    - 3: UIS estimate. Manual estimation produced by the UIS.
    - 5: Unpublishable. Data cannot be used or published.
- COMMENT_EST: These are the comments added by UIS staff to justify the choice of qualifier, if relevant, or to explain how the data were estimated.

While it is not possible to review in detail all the comments included in the dataset, a rapid review of these comments can give the basis for further discussion regarding the type of issues faced by the UIS and countries in reporting teacher data.

6.1 Estimated data points and unpublishable data

Table 12 shows the distribution of UIS qualifiers found in the EST database for each of the missingness patterns observed. The table also helps to confirm some of the labels as it confirms for instance that most records in pattern #8 are either estimated by the UIS or marked as unpublishable (which can be done by the UIS when regional averages are calculated for instance). Taking only records which have a value in the EST database at the end of the data processing chain (Patterns #2, 3, 5 and 8) the amount of data marked as unpublishable by UIS staff stands at 8.4% of all records or 8,055 data points of the 95,554 that end up in the EST database. In total, almost 10% of records with a value are either lost at the end of the process or simply introduced for the sake of estimating regional averages (and thus automatically marked as unpublishable data). Of notable importance as well, the UIS currently estimates only a small fraction of its published data; coverage could be improved by making more use of estimations.

Table 12. Missingness patterns and UIS qualifiers

		2 (national estimate)	3 (UIS estimate)	5 (unpub-lishable)	NULL (no qualifier)	O (observed)	Z (unpub-lishable)	Total	Total records
1	Unreported and unestimated	0.0	0.0	0.0	74.1	25.8	0.1	**100**	**231248**
2	Ideal scenario	0.1	1.0	0.5	0.0	89.6	8.8	**100**	**76441**
3	Misreported and corrected or found elsewhere	0.2	0.7	0.3	0.0	92.3	6.4	**100**	**14323**
4	Misreported and deleted	0.0	0.0	0.0	17.1	82.1	0.8	**100**	**13940**
5	Reported and corrected but deleted in EST	0.0	0.0	0.0	74.3	25.7	0.0	**100**	**8931**
6	Estimated only	1.2	14.1	1.9	0.0	77.0	5.8	**100**	**3684**
7	Found elsewhere/Estimated in OBS but deleted in EST	0.0	0.0	0.0	80.3	19.7	0.0	**100**	**1777**
8	Misreported and deleted but Estimated	0.0	6.8	1.2	0.0	77.8	14.3	**100**	**1106**
		0.0	0.4	0.1	51.7	45.3	2.4	**100**	**351450**

	2 (national estimate)	3 (UIS estimate)	5 (unpub-lishable)	NULL (no qualifier)	O (observed)	Z (unpub-lishable)	Total
Total EST (2+3+5+8)	158	1473	473	0	85395	8055	**95554**
(%)	*0.2*	*1.5*	*0.5*	*0.0*	*89.4*	*8.4*	**100**

Source: UIS database, author's calculations

Following observations made from Table 12, explorations for improvement of coverage need to look into some of the issues encountered by UIS staff. In particular, there are a number of judgement calls that the statisticians have to make. This include for instance decisions to publish particular data points or to proceed with estimation of data points. A cursory look at comments associated with unpublished data gives at least six broad categories of issues to act upon to improve the coverage of the UIS database on teachers. The most frequent issue is related to data processing, as many unpublishable data are marked as such while awaiting a response from the country. The UIS validates all of its data with country focal points and the correction and publication of data relies on agreement between the UIS and country focal points on potential issues. Lack of reactivity of country focal points can result in delays in publishing data on teachers. Another frequent issue encountered is that of unresolved out-of-trend data, which means that a country confirmed the data yet is not able to produce a valid explanation for a sudden change of trend. In many of these cases and in the absence of additional information, the UIS takes the decision to mark the data point unpublishable. Other issues include partial reporting, notably due to absence of data for private institutions; unresolved internal consistency issues, for instance when there are more trained teachers than teachers headcount; or definitional issues, either misunderstanding of definitions or incompatibility of global definitions and processes with national definitions.

Finally, there might be grounds for the UIS to improve its coverage by looking into better estimation procedures. At the moment, it seems that estimation procedures are rather ad-hoc and on a case-by-case basis. Four types of estimations produced by the UIS can be identified. These are (1) estimations based on informal knowledge obtained during workshops, (2) based on other international data sources, such as the OECD Teaching and Learning International Survey (TALIS), (3) estimations based on additional data provided by countries, and (4) face value estimates produced using past trends. To our knowledge, the UIS does not use additional estimation methods for data points while there would likely exist some fit-for-purpose machine learning approaches that could at least partially solve the issue of low coverage.

6.2 Overall assessment of the quality of teacher data in the UIS database

This rapid review of missingness patterns and quality aspects of UIS data on teacher has highlighted a number of points:

- Gender issues are well covered. Data on teachers disaggregated by sex are equally well-reported as data on teachers for both sexes combined. Similarly, data for public and private institutions seem reasonably well reported although there might be an association between absence of reported data on teachers and difficulties to collect data in the private sector.

- Data on teachers in technical and vocational programmes seem to be more difficult to track for countries and are consistently underreported in comparison to data on teachers in general programmes.

- While current trends show an improvement in reporting rates by countries, the present status of the UIS database results in weak prospects for consistent and sustainable reporting on SDG target 4.c. Some dimensions are not reported at all by countries despite the fact that they are needed to report on the thematic agenda (newly recruited teachers, in-service training). Indicators 4.c.6 and 4.c.7 have almost non-existent coverage.

- The main international data collection on teacher indicators applies a different approach to developed countries and to developing countries. UOE countries exhibit in many cases a higher rate of missing data than developing countries, not because the former do not have the capacity to report but because their specific data collection completely excludes all the data necessary to calculate SDG target 4.c indicators. Overall it means that around 23% of all countries will automatically have a missing value for all of their SDG target 4.c indicators.

- Lower-middle-income countries make attempts at responding to more data points but their reporting often seems to be incorrect and has to be corrected in the process by UIS staff. Lower-middle-income countries eventually end up with similar coverage rates as low-income countries despite reporting substantially more data to the UIS.

- Slightly less than 10% of EST data are marked as unpublishable but the reasons behind unpublishable data are not always unsolvable as they often pertain to the lack of responsiveness of countries to UIS queries or to the actual data collection protocol. That is, reinforcing feedback loops and collaboration between the UIS and countries' focal points would likely solve a substantial share of these issues as many cases are actually just pending a response.

- Improving coverage through more estimation procedures should be feasible as the UIS currently only uses a case-by-case approach to estimating data points.

7. Recommendations

This rapid overview of the main international data collection on teacher sheds light on a number of issues which, if addressed, have the potential to yield substantial improvements in the quality of monitoring SDG target 4.c. A number of recommendations can thus be made both with regard to the development of an international teacher training taxonomy and with regard to the improvement of the international data collection.

The main recommendation pertains to improving the current data collection in terms of process, coverage and information collected. The collection of data on teacher training programmes has faced several challenges. At the same time, the ISCED data collection is one of the most successful of the UIS in terms of coverage. The UIS could build on the well-functioning and understood ISCED data collection to collect more on teacher training programmes. One solution would be to include a separate section in each ISCED questionnaire for teacher training programmes. This is less likely to encounter reluctance from countries than a totally new data collection. The new section would for instance collect data on target level (i.e. for which level are teachers trained), as well as more standardized responses for pre-requisites for entry.

Additional data processing standards could be put in place to improve the quality of the data collected. Attention to the actual content of open text fields (and notably fields which require the name of the programmes or diploma award in English, as many of these are not translated) would improve the potential for further analysis of UIS data. From the available data on teacher training programmes, 45% of the sample is not exploitable. The depth and accuracy of the results could be substantially improved upon simple confirmation by countries of the level at which teachers are expected to teach at the end of the training. This is particularly important as a level of uncertainty remains for high-income countries.

Further attention to data collection for developed countries could help send the right signal for developing countries.

The ISCED data collection mostly enables examination of initial teacher training programmes. However, the increasing diversity of pathways to the teaching profession is not reflected (usually 1 or 2 programmes per level only in the current UIS database). Thinking about how to collect data on pathways to the teaching profession needs to be further developed.

There is by definition a high level of correlation between variables collected as part of the ISCED survey and actual ISCED level classification of education programmes. Perhaps a simple combination of number of years of education completed and ISCED classification of the teacher education programme would suffice as a first step before looking into content-related additional variables.

The number of pre-primary and early childhood teacher training programme is especially low and worrying. There is a need to further insist on ECE in data collection to improve the quality of data for indicator 4.c.1(a).

Issues of low reporting (in the UIS REP database) can reflect weak EMIS but also unsuccessful international data collection processes. As the UIS data collections critically rely on the expertise of national staff it becomes paramount to develop solutions that strengthen the commitment and involvement of countries' focal points. Most issues of unpublishable data are linked to lack of responsiveness and/or lack of understanding from the country. This is likely to be associated with confusion between national and international definitions of trained and qualified teachers.

Specific actions need to be targeted based on missingness patterns observed. Understanding why countries do not report teacher data to the UIS is therefore the most important step to improve monitoring of SDG target 4.c. Additionally, better standardization of the data collection process has the potential to improve targeted action. Standard coding of issues (reasons for unpublishable data, etc.) would support a flexible and reactive process.

The UIS should take advantage of current trends in machine learning developments to explore the possibility of coming up with more refined and systematic estimations methods.

References

UNESCO Institute for Statistics. 2012. "International Standard Classification of Education. ISCED 2011."

UNESCO Institute for Statistics. (2018a). *INSTRUCTION MANUAL. Survey of Formal Education*. Retrieved from http://uis.unesco.org/sites/default/files/documents/instruction-manual-survey-formal-education-2017-en.pdf

UNESCO Institute for Statistics. (2018b). *Quick Guide to Education Indicators for SDG 4*. Retrieved from http://uis.unesco.org/sites/default/files/documents/quick-guide-education-indicators-sdg4-2018-en.pdf

UNESCO Institute for Statistics. (2018c). *Teacher preparation (training) and Education 2030. Trends, challenges, typology and taxonomy*. Montréal, Canada: UNESCO.

UNESCO Institute for Statistics, OECD, & Eurostat. (2018a). *UOE data collection on formal education. Manual on concepts, definitions and classifications.* Retrieved from http://uis.unesco.org/sites/default/files/documents/uoe2016manual_11072016_0.pdf

UNESCO Institute for Statistics, OECD, & Eurostat. (2018b). *uoe-qre-education-training-statistics-educational-personnel-2018-en*. Retrieved from http://uis.unesco.org/sites/default/files/documents/uoe-qre-education-training-statistics-educational-personnel-2018-en.xlsx

United Nations. Global indicator framework for the Sustainable Development Goals and targets of the 2030 Agenda for Sustainable Development. , A/RES/71/313, E/CN.3/2018/2 § (2018).

Unterhalter, E. (2013). *Education targets, indicators and a post-2015 development agenda: Education for All, the MDGs, and human development* [Working Paper Series]. Institute of Education, Univeristy of London.

World Bank. (2011). *SABER - TEACHERS. Objectives, Rationale, Methodological Approach, and Products*. Retrieved from http://siteresources.worldbank.org/EDUCATION/Resources/278200-1290520949227/7575842-1290520985992/SABER-Teachers_Framework.pdf

World Bank. (2013). *What Matters Most for Teacher Policies: A Framework Paper*. Retrieved from http://siteresources.worldbank.org/EDUCATION/Resources/278200-1290520949227/7575842-1365797649219/Framework_SABER-Teachers_Apr.13.pdf

Annex 1: Author's transformation of UIS original database

Variables added:

- ID: Programme ID number
- EXAM: Existence of an entrance examination, binary variable. If entry requirements state that an individual must pass an examination to enter the teacher training programme then EXAM takes the value "yes".
- MARKS: Specific marks requirements, binary variable. If entry requirements state that an individual must meet specific educational achievement (specific marks subjects mastered) then MARKS takes the value "yes".
- PROF: Professional experience requirement, binary variable. If entry requirements state that an individual must meet specific professional experience requirements then PROF takes the value "yes".
- PROFY: Continuous variable. Code for the actual length of professional experience required if PROF="yes".
- MULTI: Multi-level teacher education programme, binary variable. If a teacher education programme delivers qualification to teach at several levels then MULTI takes the value "yes".
- Teachlvl: Multinomial variable. Codes for the level at which teachers are taught to teach. Values: ISCED levels: 0,1,2,3, Unspecified: 9, Multilevel: M+ ISCED levels (e.g. M01, M123…).
- ISC0-ISC3: Single level parsing, binary. Coded 1 if yes.
- INCG: World Bank income group, July 2018.
- ThAgerecode: Age recoded based on the data manipulation described below.
- ThDurrecode: Duration recoded based on the data manipulation described below.
- Multientry: A programme can be entered at several points, binary variable. Takes the value 1 if yes.
- EntrPt: Education cycle point equivalent to entry age.

Data manipulation:

Teacher training programmes

- Duration was limited to full-time only. If a programme has two durations (one full time, one part time) only full-time duration is accounted for (example: Mauritius teacher diploma for special education needs).
- Duration intervals (e.g. Sweden postgraduate) are reduce to their middle point (e.g. 1-1.5=1.25).
- If a given programme has several entry points only the longest duration and the earliest entry point are accounted for.
- Programmes for pedagogical and orientation counselling were removed.

UIS Databases REP, OBS and EST

- Non-discriminating dimensions were removed: "FREQ", "EDU_TYPE", "ISCP11_SUB", "GRADE", "FIELD", "COUNTRY_ORIGIN", "COUNTRY_CITIZENSHIP", "AGE", "UNIT", "INFRASTR", "QUAL_LEVEL".
- NULL values transformed to NA.