

Updating India's Harmonized Test Scores for the Human Capital Index¹

I. Background

The Human Capital Index (HCI) measures the level of human capital that a child born today can expect to attain by age 18, given the risks of poor health and poor education that prevail in the country where she lives. The health and education components of the index are combined in a way that reflects their contributions to worker productivity. The index units represent productivity relative to a benchmark of complete education and full health, on a scale of 0 to 1.

The education component of the HCI—Learning-Adjusted Years of School—combines information on the quantity of school using Expected Years of School and its quality using Harmonized Test Scores (HTS). HTS are national average scores from major international and regional student achievement testing programs (ISATs and RSATs respectively), harmonized into common units. The HTS uses TIMSS-equivalent units, where 300 is minimal attainment and 625 is advanced attainment. HTS for the 2018 HCI were retrieved from the Global Database on Education Quality (Patrinos and Angrist, 2018).² The 2020 HCI update uses the January 2020 update of this database.

The HTS database harmonizes scores from seven ISATs and RSATs into HTS units by creating an exchange rate between international standardized achievement tests (ISATs) such as PISA, TIMSS, and PIRLS and their regional counterparts (RSATs) such as SACMEQ, LLECE, PASEC, and PILNA, as well as EGRAs. This exchange rate is derived by comparing average scores for countries that participate in both an RSAT and an ISAT in a given time period, schooling level (primary and secondary), and subject. This exchange rate is then applied to country-level average scores for all countries in the testing program.

II. An updated HTS for India

This note describes the process for estimating nationally-representative test scores for India for use in the September 2020 update of the World Bank's Human Capital Index (HCI). India was originally included in the 2018 version of HTS database using scores from the 2009 PISA administered to 15-year-old students in the states of Himachal Pradesh and Tamil Nadu. The 2009 PISA was the only ISAT in which India has participated, and thus was the only way to include India in the HTS database.³ The HTS value of 355 based on the 2009 PISA was used in the 2018 HCI. However, these data are now over 10 years old.

One alternative to calculating an HTS that has become available since the release of the 2018 HCI is the Global Alliance to Monitor Learning (GAML) process for benchmarking minimum proficiency levels across a variety of international, regional, and national assessments. The GAML process identifies the level of

¹ This note was produced by a team led by Aart Kraay (DEC) and Roberta Gatti (HDCE) and including Ritika D'Souza (HDCE) and Paul Corral (HDCE). Substantive inputs were provided by Joao Pedro Azevedo, Marguerite Clarke, Jorge Coarasa, and Shabnam Sinha. The team is grateful for insightful comments from Cristian Aedo, Junaid Ahmad, and Lynne Sherburne-Benz.

² Patrinos, Harry and Noam Angrist (2018). "Global Dataset on Education Quality: A Review and Update (2000-2017). World Bank Policy Research Working Paper No. 8592. <http://documents.worldbank.org/curated/en/390321538076747773/Global-Dataset-on-Education-Quality-A-Review-and-Update-2000-2017>.

³ Although the 2009 PISA assessment covered only two states, calculations by the HTS team suggest that the average score of the two was nevertheless approximately representative of the national average. See Patrinos and Angrist (2018), Annex 3.

attainment on an assessment that is most comparable to a globally agreed definition of minimum proficiency, based on a process of “expert moderation” that includes reviewing the representativeness and difficulty level of questions on the assessment. Usually, this process culminates in the selection of the most appropriate from among an existing set of attainment levels/cut points on the assessment. India’s 2017 National Assessment (NAS) was included in this GAML process.

The GAML process has primarily been used to derive internationally-comparable measures of the proportions of children reaching minimum levels of proficiency in reading and math at three different levels of schooling (Grades 2/3, end of primary, and end of lower secondary) as a part of SDG 4.1.1 reporting. The World Bank has already drawn on the GAML process to inform its global measure of learning poverty (i.e., identifying the proportion of children at the end of primary school who are “below minimum proficiency” in reading based on GAML-identified cut points on national and international assessments).

The minimum proficiency levels on assessments identified by the GAML process can, in principle, also be used as a basis for calculating comparable country-level average scores analogous to the HTS. This is best explained with an example. According to the GAML process, the scale score on the PIRLS fourth-grade reading test that marks the boundary of “minimum proficiency” is 400 (PIRLS scale scores are the numeraire units for the HTS for primary reading).⁴ In the case of India’s NAS, the GAML process identified a scale score of 306 on the NAS Grade 5 reading test as marking the boundary of minimum proficiency. This score marks the cut-point between the “basic” and “proficient” achievement levels on the NAS. Therefore, the GAML process can be thought of as assuming an equivalence between a PIRLS score of 400 (in HTS units) and a NAS score of 306 (in NAS units).⁵

The remaining step in developing an HTS through this approach is to adjust for the fact that PIRLS scores have a mean of 500 and standard deviation of 100 while the NAS scores have different means and standard deviations across subjects and levels. However, it is possible to use the reported NAS mean scores and standard deviations to rescale the NAS proficiency thresholds to the values they would have had if the NAS used the PIRLS mean and standard deviation of 500 and 100 respectively.⁶

⁴ Descriptors of ISAT proficiency thresholds and corresponding boundary scale scores can be found at <http://timssandpirls.bc.edu/timss2015/international-results/timss-2015/mathematics/performance-at-international-benchmarks/>; <http://timssandpirls.bc.edu/timss2015/international-results/timss-2015/science/performance-at-international-benchmarks/>; and <https://www.oecd.org/pisa/publications/pisa-2018-results.htm>, Table 1.5.1.

⁵ Conceptually, there is no particular justification for focusing on the upper boundary of the minimum proficiency level for this linking process. This choice is made only for the pragmatic reason that the boundary scale scores are readily available in the publicly-available NAS documentation. With access to student-level scores from the NAS as well as from PIRLS where they are publicly-available, it would be possible to do more sophisticated linking of the entire distribution of scale scores over a particular range (for example by percentile matching). Unfortunately, however student-level scale scores from the NAS were not publicly available at the time of this analysis.

⁶ Specifically, let μ and σ represent the mean and standard deviation of NAS scale scores, and let Z represent the proficiency threshold in NAS. The proficiency threshold rescaled to 500-100 units is $\frac{Z-\mu}{\sigma} \times 100 + 500$. Expressing this as a ratio to the PIRLS BMP of 400 gives an exchange rate between the NAS and PIRLS of $XR = 400 / \left(\frac{Z-\mu}{\sigma} \times 100 + 500 \right)$. Applying this exchange rate to the rescaled NAS mean gives $HLO = XR \times 500 = 400 / \left(\frac{Z-\mu}{\sigma} \times 100 + 500 \right) \times 500$.

Table 1 presents the results of this exercise for India using NAS 2017 data for grades 5 and 8.⁷ Averaging across subjects at each grade level yields HTS ranging from 393 to 399.

Table 1: HLO using NAS 2017 and GAML Linking Process

India National Achievement Survey (NAS)					Corresponding ISAT				Conversion to HLO		
Grade	Subject	Mean	Standard Deviation	Proficiency Threshold	Name	Subject	Grade	Proficiency Threshold	Rescaled Proficiency Threshold	Implied Exchange Rate	Implied HLO
V	Language	319	59	306	PIRLS	Reading	4	400	478	0.84	418
V	Math	310	57	320	TIMSS	Math	4	400	518	0.77	386
V	Env. Stud.	310	54	315	TIMSS	Science	4	400	509	0.79	393
<i>Average</i>											399
VIII	Language	307	54	320	PISA*	Reading	8	407	524	0.78	388
VIII	Math	269	53	275	TIMSS	Math	8	400	511	0.78	391
VIII	Science	274	58	275	TIMSS	Science	8	400	502	0.80	399
<i>Average</i>											393
Overall Average											396

* Note that PISA is the numeraire test for secondary reading in the HLO database.

This approach allows for the incorporation of more recent, nationally-representative data on learning levels into India's HTS calculation. The GAML linking appears to be most rigorous for Grade 5⁸ and using test score data for this grade produces an HTS for India of **399**.⁹

III. Use in the HCI Update

Following best practices with index updates, the 2020 update to the HCI will include a back-calculated version of the 2018 HCI using the most recent data available now for 2018 for all countries. For India, the HTS estimate of 399 based on the NAS is the most recent measurement available as of 2018 and is used in the back-calculated 2018 HCI as well as in the 2020 HCI. The updated HTS of 399 based on NAS is not immediately comparable with the HTS value of 355 based on PISA that was previously used to calculate the HCI for 2018, and the resulting change between the original 2018 HCI and the back-calculated 2018 HCI should not be interpreted as an improvement in learning as measured by the HTS.

⁷ As published in MHRD/NCERT (*forthcoming*). "National Achievement Survey 2017: Class III, V, VIII: National Report to Inform Policy, Practices, and Teaching Learning", Tables 6.2, F3, F4 and F5.

⁸ Only the Grade 5 results are officially reported on the UIS website for SDG 4.1.1; the Grade 8 data are not shown.

⁹ The methodology of creating an HTS based on equating boundaries of proficiency thresholds is an imperfect and provisional methodology that is being used in the particular case of India for lack of better alternatives. If/when India participates in an ISAT such as PISA on a nationally-representative basis, and/or includes linking items in its NAS, data from these tests can replace provisional linking methodologies.