Reading Technical Reports of International Large-Scale Assessments

Simon

1. What is the full name of the assessment?

   ICCS 2016 Technical Report stands for International Civic and Citizenship Education Study 2016 Technical Report.

2. Who coordinates and administers the assessment? What partner organizations does the sponsoring organization work with? Describe the roles of each.

   ICCS 2016 was organized and conducted by a consortium of three partner institutions which closed cooperated with the national research coordinators (NRCs) from the participating countries:

   a. The Australian Council for Educational Research (ACER)
      It acted as the international study center, for the general development of study design, assessment framework, and international student instruments.
   b. The International Association for the Evaluation of Educational Achievement (IEA) acted as the project coordination center, with responsibilities for data processing, sampling, scaling, data analysis, and general coordination activities.
   c. Laboratorio di Pedagogia Sperimentale(LPS) at Roma Tre University, with responsibilities for the development of the teacher and school surveys and a European regional module.
   d. And there are many dedicated groups to help manage the study including:
      i. Joint management commitment (JMC)
      ii. Project advisory committee (PAC)
      iii. NRCs played a crucial role in this study as they coordinated the work of the national research centers, oversaw the local implementation of survey procedures, and contributed to the development of the assessment framework, instruments and reporting through a series of face-to-face meetings, and through a regular communication network.

3. What are the aims/goals of the assessment?

   The study will show the students' knowledge and understanding of civics and citizenship, as well as their attitude, perceptions, and activities related to civics and citizenship. Furthermore, it examined differences among counties in relation to these outcomes of civic and citizenship education and explored how cross-national differences relate to student characteristics, school and community contacts, and national characteristics.

4. How long has the assessment been in existence? How often is the assessment administered? How many cycles have been administered to date? Have there been any changes in the administration of the assessment over time?

   ICCS 2016 is a continuation and an extension of ICCS 2009.
   I think the cycle time is uncertain but it depends on the enduring and emerging challenges of educating young people in a world where contexts of democracy and civic participation continue to change at the national, regional, and global levels. And it will be conducted when the world needs it.
   ICCS2016 is the second cycle of this assessment.
   ICCS2016 is designed to maintain continuity with ICCS2009, enabling enduring aspects of civic and citizenship educational contexts, processes, and outcomes to be measured, while supporting comparison of outcomes and contexts between 2009 and 2016, so there are some key materials and variables that are statistically linked to enable changes.

5. Who participates? What are the requirements for countries to participate? Which individuals take the assessment? (e.g., students, parents, teachers, principals)

   ICCS 2016 gathered data from more than 94,000 students in their eighth year of schooling in about 3800 schools from 24 countries. Also most of these countries had participated in ICCS 2009.
   So, participation in ICCS is open to all IEA member countries and affiliated countries and education systems. Each county decides whether or not it will participate in an IEA study.
   The main targeted group is students, and the tests/questionnaires were given by more than 37,000 teachers in those schools and by contextual data collected from school principals and national research centers.
   And there are two additional questionnaires for European and Latin American students.

6. What is material is covered in the assessment? Which domains? Which content?

   ICCS2016 mainly focused on the following research questions:

a. The way civic and citizenship education is implemented in participating countries, including the aim and principles for this learning area, the curricular approaches are chosen to provide it, and changes and/or developments since 2009.
b. The extent of students' knowledge and understanding of civics and citizenship, and the factors associated with its variation across and within countries.
c. Students' current and expected future involvement in civic-related activities, their perceptions of their capacity to engage in these activities, and their perceptions of the value of civic engagement.
d. Students' beliefs about contemporary civil and civic issues in society, including those concerned with civic institutions, rules, and social principles (democracy, citizenship, and diversity), as well as their perceptions of their communities and threats to the world's future.
e. The ways in which schools organize civically and citizenship education, with a particular focus on general approaches, the processes used to facilitate civic engagement, interaction with their communities, and schools' and teachers' perceptions of the role of this learning area.

7. How was the research instrument developed? Who was involved in the process?

The ICCS2016 comprised the following instruments administered to students, teachers, school principals, and national centers:

a. 88 items measuring civic and citizenship knowledge, analysis, and reasoning contained in the student cognitive test were assigned to 8 booklets. Each student completed one of the 45-minute booklets.
b. Student questionnaire - for students.
c. Teacher questionnaire - for teachers.
d. School questionnaire - for school principals
e. National research coordinators compiled and synthesized the information from national experts in response to an online national contexts survey.

8. What is the sampling method? How are schools sampled? How are students sampled? Explain the sampling process in some detail. How is the sampling method implemented?

The school samples were designed as stratified two-stage cluster samples;

Schools were randomly selected at the first stage with probability proportional to size, and intact classrooms were sampled at the second stage. Typically, each country aimed for a sample size of 150 schools, the total number of sampled students ranging between 3000 and 4500. Furthermore, ICCS aimed to sample 15 teachers from all teachers teaching the target grade at each sampled school.
Firstly, it has defined the target population to involve in this analysis, for most countries, the targeted grade is 8 grade, and targeted teachers are school staff members who provide student instruction through the delivery of lessons to students.
Then it specified the intended sample sizes and achieved sample sizes. The overall goal of the student sample design was to achieve an effective sample size of at least 400 students for the main variable of interest.
The civic-knowledge score and questionnaire scales reflecting civic-related perceptions were regarded as the main variables of interest.
To be more specific, In the ICCS student survey, the ICCS sampling team asked each participating country to have a minimum intended school sample size of 150 selected schools. This meant selecting at least one intact class from each school. Once non-participation of schools and students had been taken into account, these requirements were expected to result in an achieved student sample size of roughly 3000 tested students.

Country	Originally sampled schools (*n*)	Student survey		Teacher survey	
		Participating schools (*n*)	Participating students (*n*)	Participating schools (*n*)	Participating teachers (*n*)
Belgium (Flemish)	165	162	2931	157	2021
Bulgaria	150	147	2966	140	1549
Chile	180	178	5081	169	1452
Chinese Taipei	150	141	3953	144	2239
Colombia	150	150	5609	136	1580
Croatia	178	175	3896	176	2723
Denmark	240	184	6254	59	489
Dominican Republic	150	141	3937	128	754
Estonia	175	164	2857	49	403
Finland	185	179	3173	170	2097

The IEA used as its general approach a stratified two-stage probability sampling design, in which the schools were selected systematically with probability proportional to size (PPS) within each stratum. IEA sampling team asked national centers to provide a list of schools with students enrolled in the target grade. The team carefully double-checked the ICCS school-sampling frames in order to ensure that they provided complete coverage of the target population and did not include incorrect entries, duplicate entries, or entries that referred to elements that were not part of the target population. The team then verified the plausibility of the information against official statistics.

9. How are test items sampled?

They used the item response theory (IRT) scaling methodology to scale the test items. Use of the one-parameter (Rasch) model for dichotomous items means that the probability of selecting Category 1 instead of 0 is modeled as

$$P_i(\theta) = \frac{\exp(\theta_n - \delta_i)}{1+\exp(\theta_n - \delta_i)}$$

In the case of items with more than two categories (as, for example, with constructed open-ended test items with partial and full scores), we can generalize this model to the partial credit model (Masters & Wright, 1997), which takes the form:

$$P_{X_i}(\theta_n) = \frac{\exp\sum_{j=0}^{x}(\theta_n - \delta_i + \tau_{ij})}{\sum_{h=0}^{m_i}\exp\sum_{j=0}^{h}(\theta_n - \delta_i + \tau_{ij})} \quad x_i = 0,1,\ldots,m_i$$

Then the test coverage and item dimensionality, when measuring cognitive abilities, it is important to use test items that cover the different levels of achievement found in the target population. As a start point, we estimated the distribution of cognitive abilities among ICCS 2016 students and the location of item thresholds

10. What quality assurance measures are taken so that the assessment is reliable?

IEA developed the quality assurance program, which was an integral part of the study both nationally and internationally. Two independent quality assurance programs were implemented in each participating country; these were designed to offer evidence of the procedures employed in data collection in order to advocate for data comparability and included an international quality assurance program, conducted by the IEA, and a national quality assurance program, managed by each national center and based on the guidelines and manuals provided internationally.
The major task of the independent international quality observers (IQOs)
IQOs was to conduct on-site visits during the main survey data collection. In each country, the IQO and the IQO's assistant(s) visited a sample of 10% of the participating schools during the ICCS administration. When there was a benchmarking participant from the same country, and only one centrally organized national center was responsible for all aspects of data collection, five additional school visits were required for the benchmarking entity.

11. Reflect on what you have learned about your assessment. What are your thoughts about how the assessment is designed and delivered?

To conduct an international large-scale assessment is incredibly difficult.
By looking through the ICCS2016 assessment, I am surprised by the scale of this assessment and the resources to be used for this.
Firstly, you have to think about the objectives of conducting this large-scale assessment. You have to define profound research questions to conduct a meaningful assessment. Then you have to gather the resources and sponsorships to organize and coordinate different parties to implement the assessment. Luckily, then you have to think up useful models and sampling strategies to do the analysis. Then you have to design data collecting formats, like questionnaires, etc. After collecting those data then you have to validate the data by using quality assurance measures techniques. Then you start to do data analysis to get a conclusion from the data you collect.
During these processes, each step needs to be scientifically implemented with appropriate methods. With this level of data, the noise of data will cost a huge amount of effort to deal with. And to make sure of the accuracy of the assessment, the data also needs to be as detailed and objective as possible.
I googled online and found 3 principles to develop and design LSLAs:

1. Technically sound. They are assessment methodologies, analysis, and interpretation of data that follow scientific principles;
2. Following standardized field operations;
3. Designed to be ethical, fair, and inclusive of the target population.

Even though LSA could give people information, it still has some limitations such as LSLAs generally focus on a limited range of learning dimensions and address a defined number of the multiple purposes of education. It may not measure other variables such as classroom and school settings.
LSA contains a lot of subjects to learn and I am excited to know more about it and those practical and useful methods!