

Report from IEA's TEG

63rd IEA General Assembly, Split, 5 October 2022

Dirk Hastedt

www.iea.nl

Overview

- The IEA TEG met 14–15 November 2021 before the IRC in Dubai, UAE and 17–18 May 2022 in Cork, Ireland
- Both meetings were in hybrid format
- General topics discussed were:
 - Quality issues
 - Delays of the RM assessment system
 - Consequences of the pandemic
- Note: This report will only highlight a few agenda points

TEG members

- Dirk Hastedt (Chair)
- Henry Braun
- Julian Fraillon (New member)
- Marc Joncas
- Michael O. Martin
- Irimi Moustaki
- Ina V.S. Mullis
- Monica Rosén
- Jostein Ryssevik
- Wolfram Schulz
- Heiko Sibberns
- Matthias von Davier



Researching education, improving learning

Project-related Discussions

ICILS 2023



- The study center presented on the impact of RM-related delays and the changes to schedule. Data collection for the field trial was delayed and even suspended in a number of countries. Smaller samples from the field trial were an issue
- For the delivery of process data, several options were considered
- TEG members agreed that it would be safest to focus on the collection of detailed process data for a small number of tasks only
- However, default process data, like timing information, should be collected for all tasks

ICCS 2022



- Due to RM-related issues, data collection was delayed in a number of countries. In addition, some countries experienced faulty players in the first testing days.
- However, these issues were quickly removed

PIRLS 2021



- TEG has discussed a new statistical method for treating non-response data in scaling which shall improve the psychometric properties of the ability estimates
 - TEG appreciated the careful examination of the data and asked for analysis of the effect that this change would have
- Group adaptive design was successfully implemented and PIRLS will continue to move in this direction in future cycles
- The number of Jackknife zones will be increased to 125. The previous number, 75 (established in TIMSS 95), is too small in more than 80% of the countries. However, this will not decrease the error, but increase the precision of the error estimate

TIMSS 2019



- More detailed analyses on the comparatively high amount of not reached and omitted items in the PSI tasks were presented
- Consequences of different treatments, i.e. wrong or not administered, of missing items were discussed
 - TEG appreciated the careful analyses and presentation of the different approaches

TIMSS 2023



- Issues in TIMSS paralleled those in the other IEA studies
- In TIMSS analyses, as in other studies, response data were prioritized over process data
- As a general observation, the nature of the work and the resources needed to accomplish it changed when ILSAs moved toward computer-based assessment
 - TEG acknowledged that some parts of the work process shifted from the international to the national level and the other way around

LaNA and Rosetta Stone



- The results of the Rosetta Stone project were presented, in particular the use of a concordance table which enables reporting students' likely performance on TIMSS and PIRLS based on the performance on one of the regional studies, PASEC or ERCE
- A future goal is to strengthen the concordance tables by including additional countries and also taking additional regions or national assessments into account
- All tools and assessment procedures for LaNA are in place so that energy is now focused on country recruitment

- The study reporting was presented to TEG, in particular what the limitations of the REDS study were, and TEG recommendations were addressed
 - TEG appreciated the clear and transparent handling of the limitations of the study, which are due to the very short time and the circumstances during the pandemic
 - TEG again emphasized that the limitations by omitting essential work steps and QC measures should remain the exception

ICILS Teacher Panel



- The ICILS teacher panel provided evidence that panel studies augmenting IEA studies are feasible
- Since the study report was already launched, the reporting of the study was in focus of the presentation to the TEG
- Different treatments of limitations due to non response were presented
 - TEG appreciated the careful and detailed documentation of all challenges and approaches in the reporting

General Topics



Researching education, improving learning

IEA StudyExpert Online Data Collection Program (Guido Martin)

- The features of the program that are important for IEA were presented in a detailed presentation
- In particular, the integration of open-source components where possible, instant deployment of tests, independence of the instruments, one setup but study-dependent configurability, storage of metadata, support for two screens for adjustments and translations with real previews as typical requirements in IEA studies were highlighted
 - TEG appreciated the thoroughly planned implementation of the new data collection program and acknowledged the integration of the experiences made in the recent years
 - TEG emphasized the need to create interdisciplinary groups in order to combine strengths, knowledge and experience from content experts, software developers, web developers and others who are involved in the development of such a system

Update of IEA Technical Standards

- The current IEA Technical Standards were written in 1999 in a specific context and in order to address specific purposes
- However, the context changed and studies evolved, for example:
 - Stand-alone studies are the exception; long-term study programs with trend reporting the norm
 - There is a change towards more advanced technology and computer-based assessment
 - The complexity of studies increased over time
- Revision of Technical Standards is necessary to address the changes
- Working group to address the changes was established
 - Lead: Michael O. Martin and Heiko Sibberns

Development of Exclusion Rates in IEA Studies (Heiko Sibberns)

- School- and student-level exclusion rates were evaluated by study, cycle, and country
- General finding: Exclusion rates are overall quite stable
- However, there are countries with a steady increase in exclusion rates, in others the rate is steadily declining, and no clear pattern can be identified in the majority of countries
- This situation makes it necessary to monitor exclusion rates continuously at the country level so that countermeasures can be taken if necessary

Moving Beyond Statistical Significance Testing (Matthias von Davier)

- Both the American Statistical Association (A.S.A) and the American Psychological Association (A.P.A) are moving towards de-emphasizing statistical significance testing
- A ritualized zero hypothesis significance testing should also be questioned
- Instead, robust methods should be used in which not every difference, no matter how small, is interpreted as significant due to the size of the sample
- Reporting confidence intervals and levels is a proven alternative. Equivalence testing would also be an option
 - TEG supported the search for alternative methods, but points out that the reports still have to be statistically sound and also have to be understandable for non-statisticians

Thank you!

Dirk Hastedt
d.hastedt@iea.nl



Researching education, improving learning

