2019

CHAPTER 4

eAssessment System for TIMSS 2019

Mark Cockle
Heiko Sibberns

Introduction

As described in Chapter 1 of this volume, TIMSS 2019 marked the beginning of the transition to eTIMSS—the digital version of TIMSS designed for computer- and tablet-based administration. eTIMSS offered an engaging, interactive, and visually attractive assessment that enabled TIMSS 2019 to better assess complex areas of the mathematics and science frameworks and increase operational efficiency in translation, assessment delivery, data entry, and scoring. Although the aim is to switch completely to the new digital mode in future assessment cycles, in recognition of the different levels of preparation and infrastructure, countries had the option in 2019 of choosing either eTIMSS or paperTIMSS.

In addition to the overarching requirements for a computer-based system that could produce attractive and engaging assessment items while being reliable, flexible, and easy to use, there were a number of other conditions that had to be taken into account in choosing the system:

- The assessment should be capable of operating on tablets as well as on personal computers.
- Assessment delivery should be via USB memory sticks or through a local server approach whereby the assessment software is installed on a local server that can be accessed by a small number of clients (no more than 30). Full internet-based administration was not a requirement for the TIMSS 2019 assessment cycle.
- Because about half the countries were administering the paperTIMSS version, it was important that items developed in the eAssessment system be as similar as possible to their corresponding paper versions, while capitalizing on interactive computer-based features such as drag-and-drop, multi-select, and drop-down menus.
- Beyond the utilization of features just mentioned in developing individual items, the system should also accommodate more extended Problem Solving and Inquiry Tasks (PSIs) designed to simulate real world or laboratory situations in which students could integrate and apply process skills and content knowledge to solve mathematics problems or conduct virtual

The Designer included various features that could be used by item developers in creating or customizing items, which was particularly important when dealing with trend items where a close match between the electronic item and the paper version was required. For example, the application of labels to images could be made above, below, or to the left or right of an image, or tables could be inserted with invisible lines in order to place objects within columns to obtain a more precise layout.

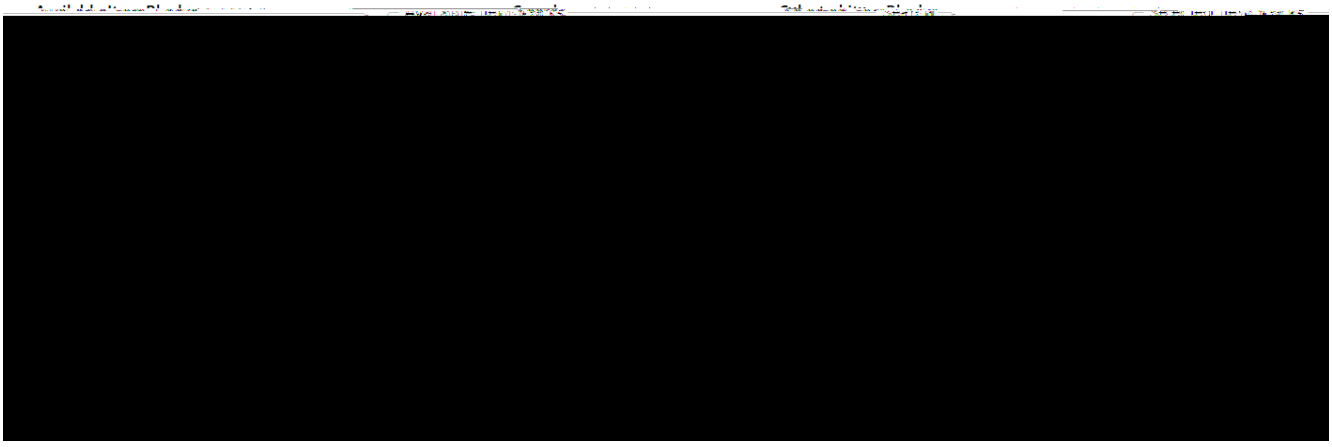
To accommodate items that included images with overlaid text, which are very common in TIMSS, an SVG feature was introduced. This provided great versatility in working with images such as line or bar charts that had overlaid text (e.g., axes labels) that later had to be translated.

It should be noted that the extended Problem Solving and Inquiry Tasks (PSIs) were substantially more complex and interactive than the standard eTIMSS items, and so were constructed independently of the Designer and subsequently combined into item block combinations (or eAssessment "booklets") by the Assembler.

Assembler

The Assembler module was used by IEA Hamburg and the TIMSS & PIRLS International Study Center to combine assessment items into blocks, and then item blocks or PSI tasks into item block combinations along with the assessment directions and eTIMSS questionnaire, in accordance with the TIMSS 2019 matrix-sampling booklet design (see TIMSS 2019 Assessment Design). It was also within this module that the allocation of the booklet number to the instrument was made. Exhibit 4.2 shows part of the Assembler window, and illustrates how a completed booklet has been constructed with six elements; on the left are available blocks that were not included in the construction of this booklet.

Exhibit 4.2: Combining Item Blocks and Directions in an Assembler Window



Once items were assembled into item blocks they could be released country-by-country into the Translation System, where countries and benchmarking participants could begin translation. Item blocks could be “released” for translation one by one as they were completed in the Designer. However, all items

- **Instrument Finalized**: The final status, assigned by the NRC, indicated that the materials had completed all verification steps and were ready for assessment Player production.

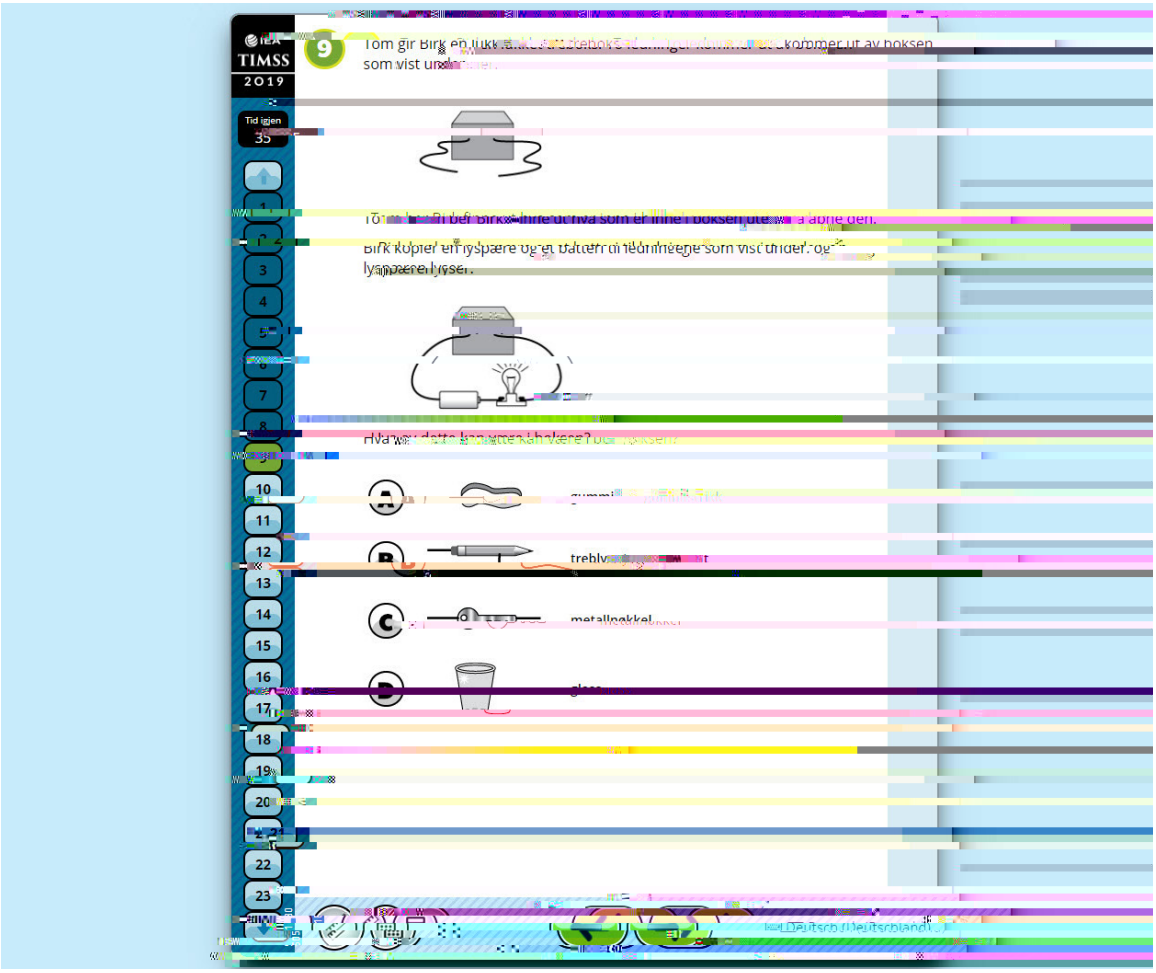
At each step along the way, comments could be left to document the process. The idea was to mimic the workflow and fields used in the National Adaptation Forms for the paper version. In the particular case of translation verification, additional labels were available to indicate specific errors or deviations found/corrected, including a "severity code" assigned by verifiers to each deviation to assist the NRC in deciding whether to accept or reject suggestions made by the verifiers (see Instrument Translation and Layout Verification). In general, comments were labeled so as to indicate for whom they were intended. For example, a layout verifier could leave comments in the system for the attention of the NRC, in which case the label would have been "Layout."

A preview feature was available for all users to display items exactly as they would appear within the final Player. This was especially useful during layout verification, allowing as it did comparison with a preview of the original (untranslated) source version.

Player

The assessment Player is the software that the student interacts with while taking the eTIMSS assessment. The Player presents the assessment items to the student and uploads the student response data to IEA's data servers. After translation and layout verification were successfully completed, a customized version of the Player was produced for each language of instruction in each country. This sometimes required

Exhibit 4.5: eTIMSS Assessment Player User Interface



The software development criteria for the Player encompassed multiple, sometimes competing elements. The key elements were speed of operation, security, and consistency of user experience.

Security

To ensure the security of the student data and test items:

- The Player was configured to run within a ”sandbox”—a virtual space in which software can be run securely–to isolate the Player in a restricted memory range
- Contents of the sandbox were automatically deleted after the testing session.

Consistency of User Experience

To ensure the Player operated the same way for all countries and languages:

- The Player exhibited a close to identical display on Firefox or Chrome browsers or with either the Android or USB Players
- Country- and language-specific CSS files were available to make final layout and font adjustments
- Right-to-left languages had automatically reversed layout, with the ability to revert individual elements back to left-to-right format.

Local Server Method

The local server method was a feature of the USB Player. With the Player program stored on a local drive of a PC, it was possible to start the program as with the standard USB from the main menu. An option then enabled the test administrator to set up the PC as a server and enable computers connected to the local network to connect to this, displaying the assessments in a browser (the Chrome browser was the preferred option).

The minimum specifications for the server computer were above those for standard USB delivery, and no system check was available to test suitability. Therefore the following minimum requirements for the server PCs were defined in order to determine if a machine was able to run the Player successfully:

- OS: Windows 8 or higher
- Processor speed: 2.2 GHz
- Memory: 8GB
- Available storage space: 10GB on SSD drive
- Administrator rights.

The upload procedure was similar to the USB method, with the additional step of stopping the server-client service. Once this was done, the results from all students could be uploaded at once.

Description of eAssessment Data

The assessment Player recorded student item responses as well as other actions taken by the student an the data were stored in a SQLite database. Student actions were broken down into time-4.7 (d i 10GB 1

- A screen ID number: This number indicated the specific screen (or item) on which the response was saved or event occurred.
-