
# Implementation Readiness Package version 2.0

## User Guide


Prepared by UCLA/CRESST

February 28, 2017

# Table of Contents

## Table of Contents

# Introduction

States are responsible for deploying and operating the Smarter Balanced Assessment Delivery System, and in many cases, they will contract for system services. Any platform used to deliver Smarter Balanced assessments should be verified to deliver assessment items with authenticity, score items and tests properly, and deliver assessment results to the Smarter Balanced Data Warehouse according to specifications. To aid with this effort, an Implementation Readiness Package (IRP) was developed. IRP is a web application hosted at http://smarterapp.cresst.net consisting of a blend of software test harnesses and written specifications that States and their vendors can use to guide and inform their Assessment Delivery System compliance testing.

This document is a user guide to the IRP version 2.0. Tables 1 below lists the features contained in IRP version 2.0.

*Table 1: IRP Version 2.0 Features*

| IRP Feature | IRP Interface | Functional Requirements Tested |
|---|---|---|
| Provides IRP data for Assessment Delivery System under test (test harness data). | *TDS Report Analysis/CAT Engine Analysis*<br>• Links and buttons provided to download necessary data to help vendors generated data for IRP analysis | a. The given test harness data allows IRP to compare against the Assessment Delivery System's output |
| Check compliance of Test Item delivery to the student. | *TDS Report Analysis*<br>• Manual upload or Automated upload of TDS Report XML documents. | a. Correctness of item rendering (display) and associated interaction on screen (NOTE: this is a human visual test)<br>b. Capability to input and capture student responses to assessment items |
| Check compliance of delivery of scored TDS Report XML documents to the Data Warehouse | *TDS Report Analysis*<br>Manual Mode<br>• Manual upload of TDS Report files to IRP server for analysis.<br>Automation Mode<br>• Automation Adapter sends TDS Report URLs to server for analysis. | a. Compliance with TDS Report file XML structure<br>b. Capability to score student responses to individual assessment items<br>c. Capability to score assessment<br>d. Correctness of data:<br>○ Student information<br>○ Item Scoring<br>○ Test Scoring |
| Analysis of Adaptive Engine | *CAT Mode*<br>• Web Service to interface with vendor system, involving student responses to assessment items and resulting adaptations. | a. Efficiency and accuracy of vendor adaptive engine analyzed and compared against "gold standard" adaptive engine using known psychometric properties. |

# IRP Usage

IRP is located http://smarterapp.cresst.net. IRP performs two type of analysis: *TDS Report Analysis* and *CAT Engine Analysis*. TDS Report Analysis analyzes TDS Report XML documents for Smarter Balanced specification compliance. CAT Engine Analysis analyzes CAT engine performance. TDS Report Analysis can be achieved in two ways: Manually and Automated. CAT Engine Analysis is performed manually. Each of these types of IRP Analysis can be chosen by selecting the desired IRP Analysis Mode as explained in the following sections.

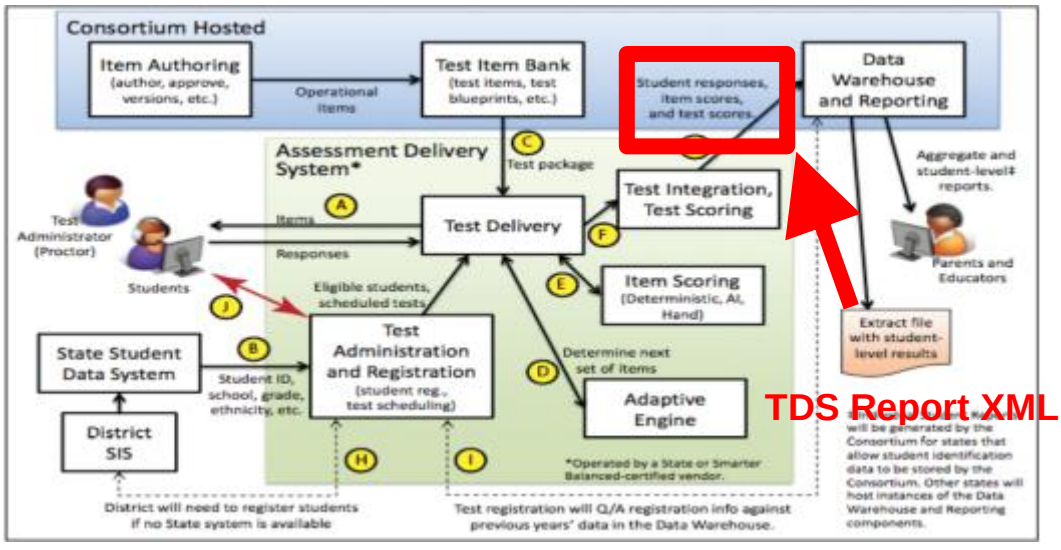
## TDS Report Analysis

Smarter Balanced specifications mandates the format and data requirements that are output from a State's Assessment Delivery System. All specifications can be found at http://www.smarterapp.org/specifications.html. According to the specifications described below, the output is represented as a ***TDS Report XML document*** (also known as *Test Results Transmission Format* document or *Data Warehouse Score Batching* document). Thus, IRP will analyze TDS Report XML documents for compliance with the following Smarter Balanced specifications:

1. **Test Results Transmission Format plus the associated Schema XSD** (http://www.smarterapp.org/specs/TestResultsTransmissionFormat.html) - Defines the TDS Report XML document format and provides the Schema XSD.
2. **Data Dictionary and Logical Data Model** (http://www.smarterapp.org/specs/TestResults-DataModel.html) - Further defines the fields, data types, and acceptable field values contained within the TDS Report XML.
3. **Smarter Balanced Reporting/SmarterApp Interface Specification/Test Score Batching** (http://www.smarterapp.org/documents/DataWarehouse-Spec-ScoreBatching.pdf) - This specification provides Scoring specifications for the TDS Report XML document.

The specifications above define most of the same fields. In case of conflicts between the documents, the definitions in the Data Dictionary and Logical Data Model takes precedence.

*Figure 1: TDS Report XML within a Smarter Balanced Assessment Delivery System*



## IRP Resources

For IRP to analyze TDS Report XML documents that are generated by a State's Smarter Balanced Assessment Delivery System, IRP mandates specific data to be used. IRP provides a data package (resources) containing Test Packages and Content, Students, Student Accommodations, and Student to Test mappings. These resources are to be input into the State's Assessment Delivery System prior to generating TDS Report XML documents for IRP analysis.

The IRP resources are referred to as the *IRP Package* on the IRP website. They are all contained in a single ZIP file located at http://smarterapp.cresst.net/irp-package.zip. When selecting the Manual Mode (described in the next section), a link to download the IRP Package can be found.

The IRP Package contains the following files:

| File | Description |
|---|---|
| IRP V1 Module A Form - Sample.xlsx | See the Item Rendering & Interactivity section. This is a completed sample form for visually verifying item rendering |
| IRP V1 Module A Form.xlsx | See the Item Rendering & Interactivity section. This form will help a State determine item rendering and interaction compliance. |
| IrpContentPackage | This directory contains IRP Items. These Items are referenced by the IRP Test Packages (see TestPackages description below). |
| IRPExplicitEligibility_template.csv | This file contains the IRP Student to IRP Test mappings. It is a template file with placeholder values that need to be replaced with implementation specific data. Once filled out, it can be used to upload into a registration system. |
| IRPStudentGroup_template.csv | This file contains a Student Group definition to group all IRP Students. Using this file is optional and doesn't influence IRP Analysis. |
| IRPStudents_template.csv | This file defines the IRP Students that must be populated into the State's Assessment Delivery System prior to generating TDS Report XML documents. It is a template file with placeholder values that need to be replaced with implementation specific data. Once filled out, it can be used to upload into a registration system. |
| IRPStudents.xlsx | This file also defines the same IRP Students as the templated file above. |
| IRPStudentsDesignatedSupportsAndAccommodations_template.csv | This file defines the Accommodations for the IRP Students. It is a template file with placeholder values that need to be replaced with implementation specific data. Once filled out, it can be used to upload into a registration system. |
| IRPStudentsDesignatedSupportsAndAccommodations.xlsx | This file also defines the same Accommodations as the templated file above. |
| IRPTestStudentMapping.csv | This file is a CSV version of a complete IRP Student to IRP |

| | |
|---|---|
| | Test mapping. It includes mappings for the summative (COMBO) tests along with individual fixed form and CAT tests. |
| IRPTestStudentMapping.xlsx | This file is an Excel version of the same data found in the CSV file. |
| TestPackages | This directory contains IRP Test Packages XML documents. These are separated into three parts:<br>1. Test Packages that need to be loaded into the Administration and Registration component<br>2. Test Packages that need to be loaded into the Test Delivery component<br>3. Test Packages for scoring that need to be loaded into the Test Scoring or Test Integration component.<br>These Test Packages refer to IRP Test Items that are contained in the IrpContentPackage directory described above.<br><br>**Note:** Test Scoring is only defined for Summative tests (also known as Combo tests) which means both a completed Fixed Form test and a CAT test must be taken for a Test Score to be computed. For the Smarter Balanced Open Source Test Integration System, additional mappings need to be inserted into the OSS_TIS database. See https://github.com/SmarterApp/TDS_TestIntegrationSystem/blob/master/TDSQAService/OSS.TIS/SQL/TISDB/4_IRP_DataLoad.sql. |

## Generating TDS Report XML Documents

This section generally describes how to generate IRP analyzable TDS Report XML documents on a State's Assessment Delivery System. IRP does NOT dictate how the State generates TDS Report XML documents, as each Assessment Delivery System will be different, with different methods and dependencies for integrating and processing students and their test results. IRP data must be used, however, because IRP's Analysis Engine checks for the existence of specific information in TDS Report XML documents.

These steps can be performed either entirely manually, partially automated, or fully automated. The steps generally work the same whichever method is used.

### *1. Download the IRP Package*

From the IRP webpage (http://smarterapp.cresst.net), download the IRP Package (described above).

### *2. Process IRP Package Data and Generate TDS Report XML Documents*

Load the State's Assessment Delivery System with the Test Packages and Student Data found in the IRP Package.

The Student Data file, **IRPStudents_template.csv** or **IRPStudents.xlsx**, contains a list of IRP Students to be used. Configure the IRP Students to be able to take the IRP Tests as described in the **IRPExplicitEligibility_template.csv** and **IRPTestStudentMapping.csv** documents. The requirement for States is that these virtual students must be integrated into their Assessment Delivery System. States or their vendor must then complete the following steps:

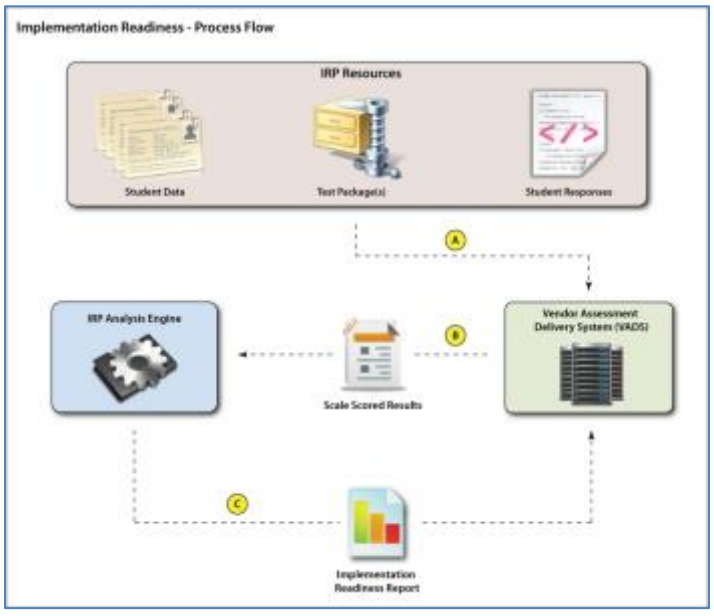
1. Each IRP Student takes their assigned IRP Test.
2. Student Responses and Item Rendering/Interactivity
    a. *Student Responses*. For each student, his or her assessment item responses are entered into the system in exactly the same fashion as if this were an actual student completing the assessment at a remote client site. **The student responses can be random.**
    b. *Item Rendering & Interactivity*. For each of the assessment items contained in the test package, the State must manually verify its proper rendering on the device screen, and its proper interactivity, as appropriate (i.e. clickable features, draggable features, text/equation entry, etc.). This is a manual, human-visual task, that verifies compliance based on the requirements described in the Item Rendering & Interactivity section.
3. The State's Assessment Delivery System must process each student's item responses through its internal scoring engine, and provide Item Scoring for each Test. Similarly, the system must process the Test Scores for the Summative (Combo) tests. The IRP Scoring Test Package is defined to score only Summative (Combo) tests. This means, that an IRP Student must complete a Fix Form and CAT test. The IRP Student to IRP Test mapping files specify mappings for both type of tests for each IRP Student.
4. These results must be captured in a *TDS Report XML document*, suitable for transmission to the Smarter Balanced Consortium Data Warehouse (Figure 1: TDS Report XML within a Smarter Balanced Assessment Delivery System). The TDS Report XML document will include all of the information required for storage in the Data Warehouse, as defined by the specifications listed at the start of the TDS Report Analysis section.

### *3. Upload TDS Report XML Documents to IRP for Analysis*

Once ready, the generated TDS Report XML document(s) will be submitted to IRP for analysis. The following sections describe how the Manual Mode and IRP Automation Adapter Mode can be used to submit the TDS Report XML documents. IRP accepts TDS Report XML documents from individual scored/unscored Fixed Form Tests, individual scored/unscored CAT Tests, or summative (combo) Tests.

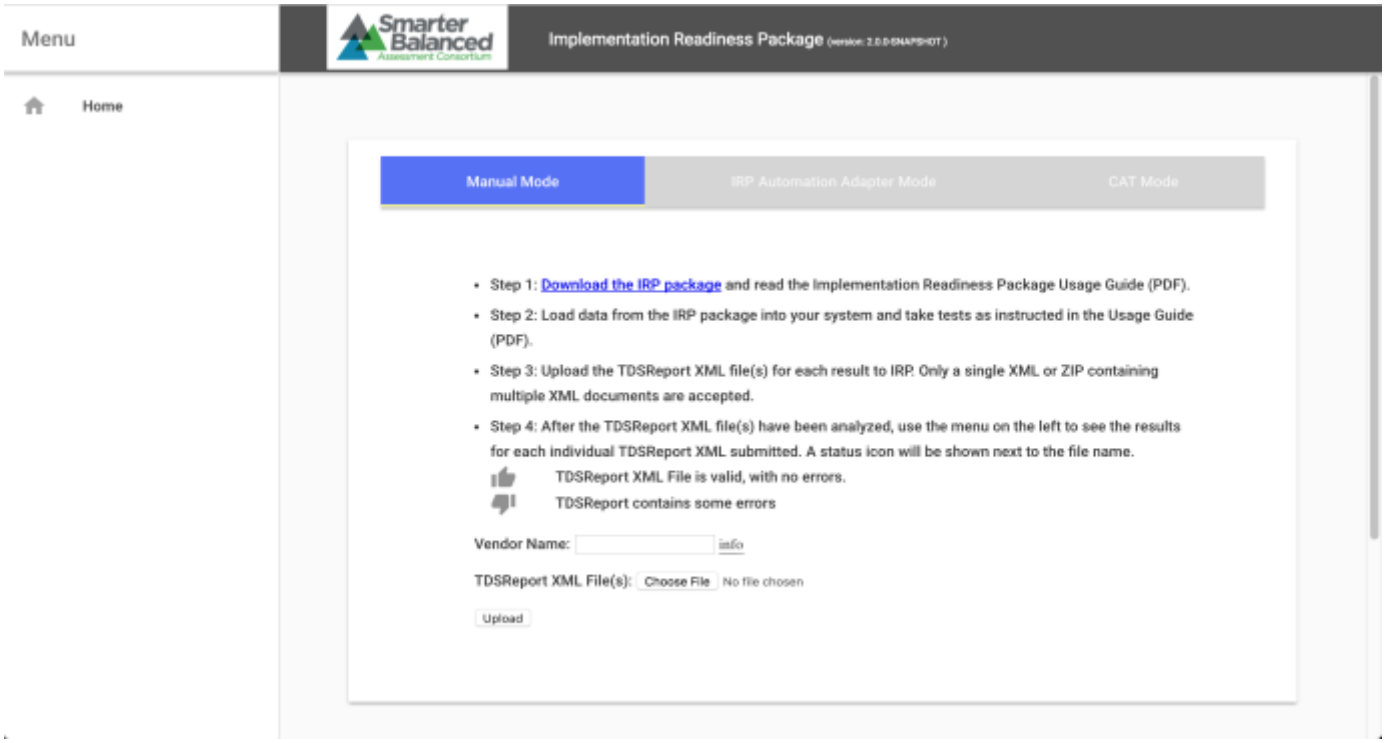
The following figure illustrates the process flow for using IRP to analyze generated TDS Report XML documents.

*Figure 2: Implementation Readiness – TDS Report XML Document Generation and Analysis Process Steps*



## Manual Mode

*Figure 3: Manual Mode*



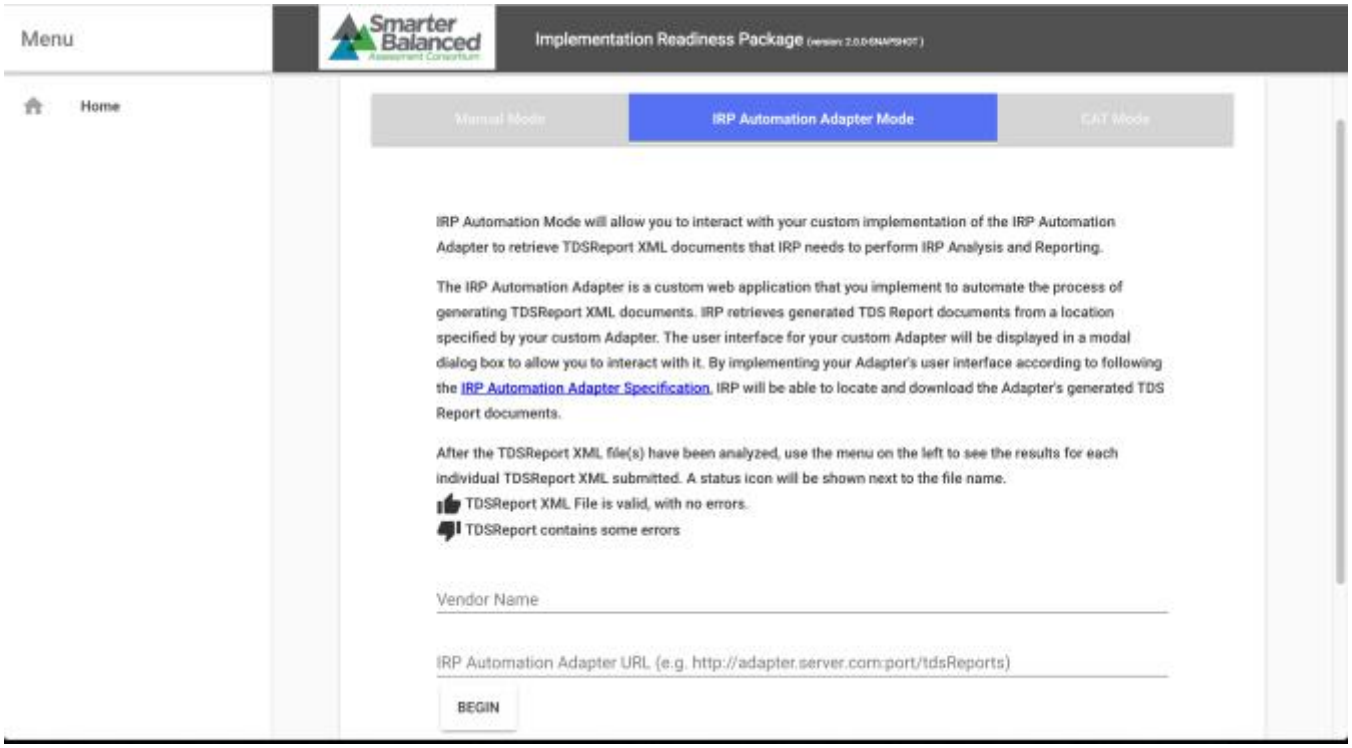

Manual Mode is the default IRP Analysis Mode when navigating to IRP's webpage. Use the Manual Mode to manually upload TDS Report XML documents to IRP for analysis.

To use the Manual Mode, enter a name to identify your organization into the **Vendor Name** textbox, as shown in Figure 3: Manual Mode. The information supplied in this textbox will be used in the final PDF version of the Analysis Report. Next, click the **Choose file** button and then select either a single TDS Report XML document or a ZIP file containing multiple TDS Report XML documents. Finally, click the **Upload** button to submit the file(s) to IRP for analysis.

See the TDS Report Analysis Results section to learn how to navigate the Analysis Report.

## IRP Automation Adapter Mode

*Figure 4: Automation Mode*

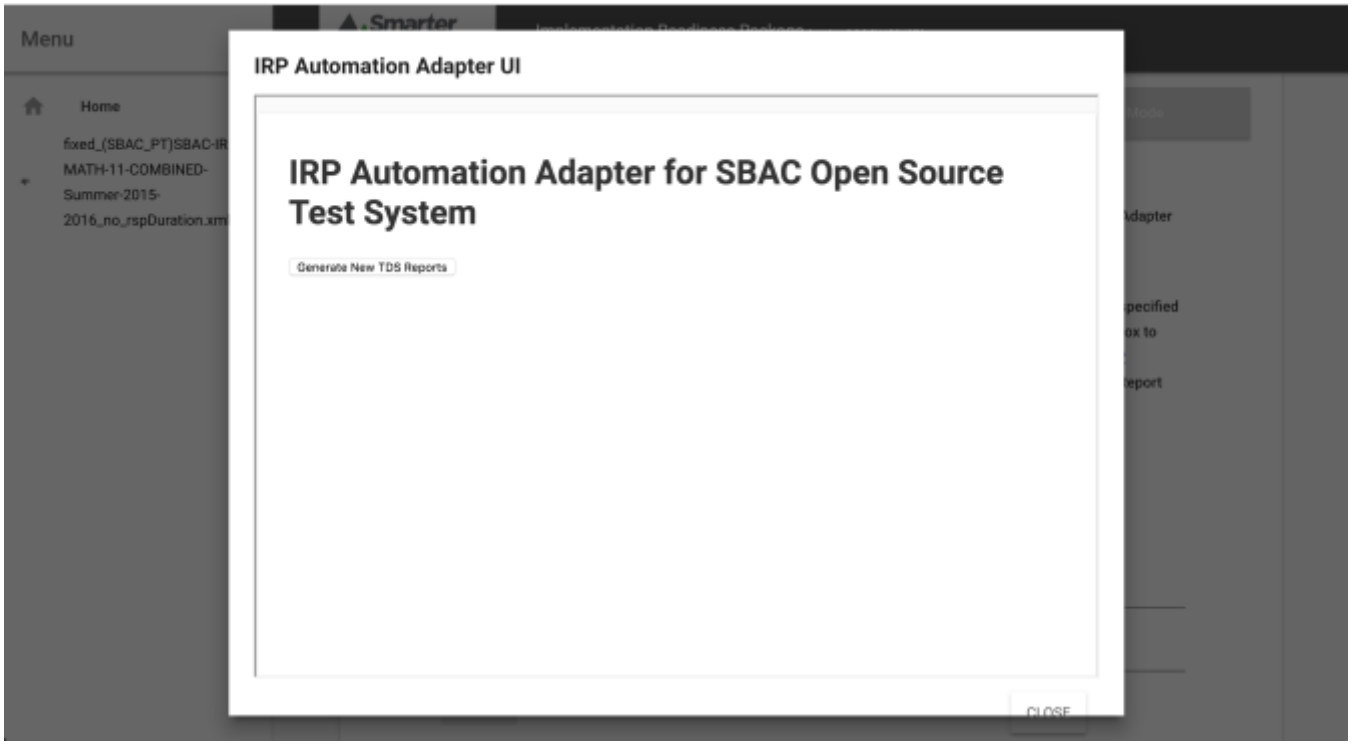



Use the IRP Automation Adapter Mode to interact with a State's Assessment Delivery System implementation of the IRP Automation Adapter (http://smarterapp.cresst.net/irp-automation-adapter-spec.pdf).

An IRP Automation Adapter automates the process of generating TDS Report XML documents. Since each State's Assessment Delivery System will be implemented differently, it is the State's responsibility to deploy an Automation Adapter that can work with their system. A reference implementation of the IRP Automation Adapter has been developed for the Smarter Balanced Open Source Assessment Delivery System. It can be used to help bootstrap custom implementations of an IRP Automation Adapter (https://bitbucket.org/sbcresst/irp-automation-adapter-sboss).

To use the IRP Automation Adapter Mode, select the **IRP Automation Adapter Mode** tab as shown in Figure 4: Automation Mode. Enter a name to identify your organization in the **Vendor Name** text field. Next, enter the URL of the IRP Automation Adapter into the **IRP Automation Adapter URL** text field (ensure it begins with the protocol; for example, http:// or https://). Next, click the **Begin** button to display a modal dialog box containing an iframe that hosts the IRP Automation Adapter as shown in Figure 5: IRP Automation Adapter Modal Dialog.

*Figure 5: IRP Automation Adapter Modal Dialog*




The modal dialog will display the custom IRP Automation Adapter hosted at the URL that was entered into the **IRP Automation Adapter URL** text field. The final step: interact with the hosted IRP Automation Adapter to generate TDS Report XML documents. If the IRP Automation Adapter complies with the specification (http://smarterapp.cresst.net/irp-automation-adapter-spec.pdf), then the modal dialog box will automatically close upon receiving a list of URLs pointing to TDS Report XML documents and submit the list to IRP for analysis.

See the next section TDS Report Analysis Results to learn how to navigate the Analysis Results.

## TDS Report Analysis Results

The TDS Report Analysis component of IRP focuses on an Assessment Delivery System's ability to synthesize: a) student data; b) students' responses to assessment items (item scoring); c) test scoring; and d) results packaging (for the Data Warehouse).

The overall idea is that a set of *sample students* each complete *two tests* that are processed through the Assessment Delivery System to yield a summative (combo) *TDS Report XML* document suitable for transmission to the Data Warehouse, as specified in the Smarter Balanced document: *Reporting Data Specification Format Test Score Batching*, located at: http://www.smarterapp.org/documents/DataWarehouse-Spec-ScoreBatching.pdf.

In this case, though, rather than sending to the Data Warehouse, the TDS Report XML document will be submitted to IRP where it will be analyzed for compliance with the following:

I. File structure (XML schema and data) for integration with the Data Warehouse
II. Completeness and accuracy of student data

III. Item scoring accuracy
IV. Test scoring accuracy

The TDS Report XML document will be analyzed by the IRP system for compliance with the following:

1. File format. Is the XML document properly formatted to meet the requirements for integration with the Data Warehouse?
2. Consistency & accuracy of student information. Has the student data been properly preserved, and is all appropriate student data provided (i.e. identifiable vs. de-identifiable information), according to the Data Warehouse requirements specifications?
3. Item Scoring. Was each student's response to each assessment item scored correctly?
4. Test Scoring. Has the Assessment Delivery System properly scored the tests as a whole, according to existing rubrics and weightings?

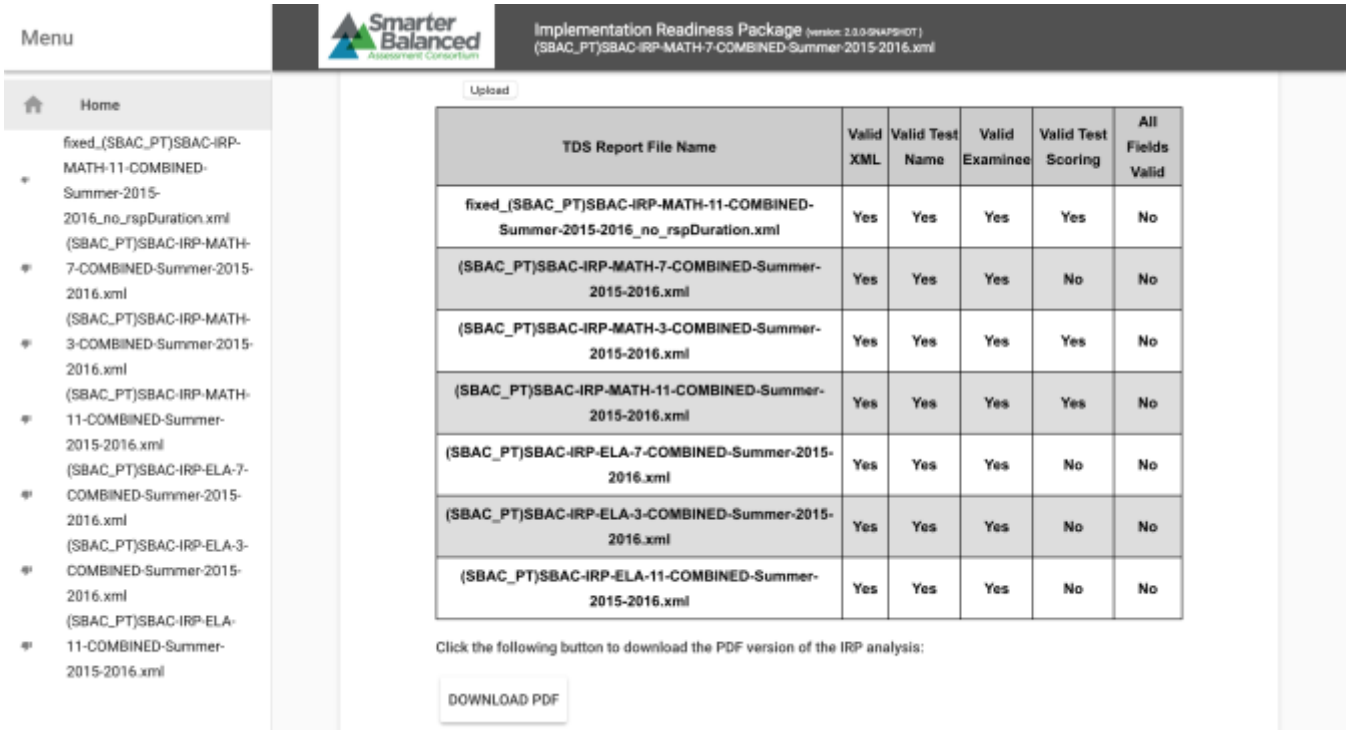
## Navigating Analysis Results

Navigating the analysis results will be the same for both Manual Mode and IRP Automation Adapter Mode.

A summary of the overall analysis results will be displayed in a *Summary Table* as shown below in the Figure 6: Summary Table and PDF Download below. The Summary Table provides a high-level summary of the Analysis Results.

Along with the Summary Table, a button to download a PDF version of the Analysis Results is available. When clicking the *Download PDF* button, IRP will take some time to construct the PDF document. The length of time depends on how many TDS Report XML documents were submitted for analysis.

*Figure 6: Summary Table and PDF Download*

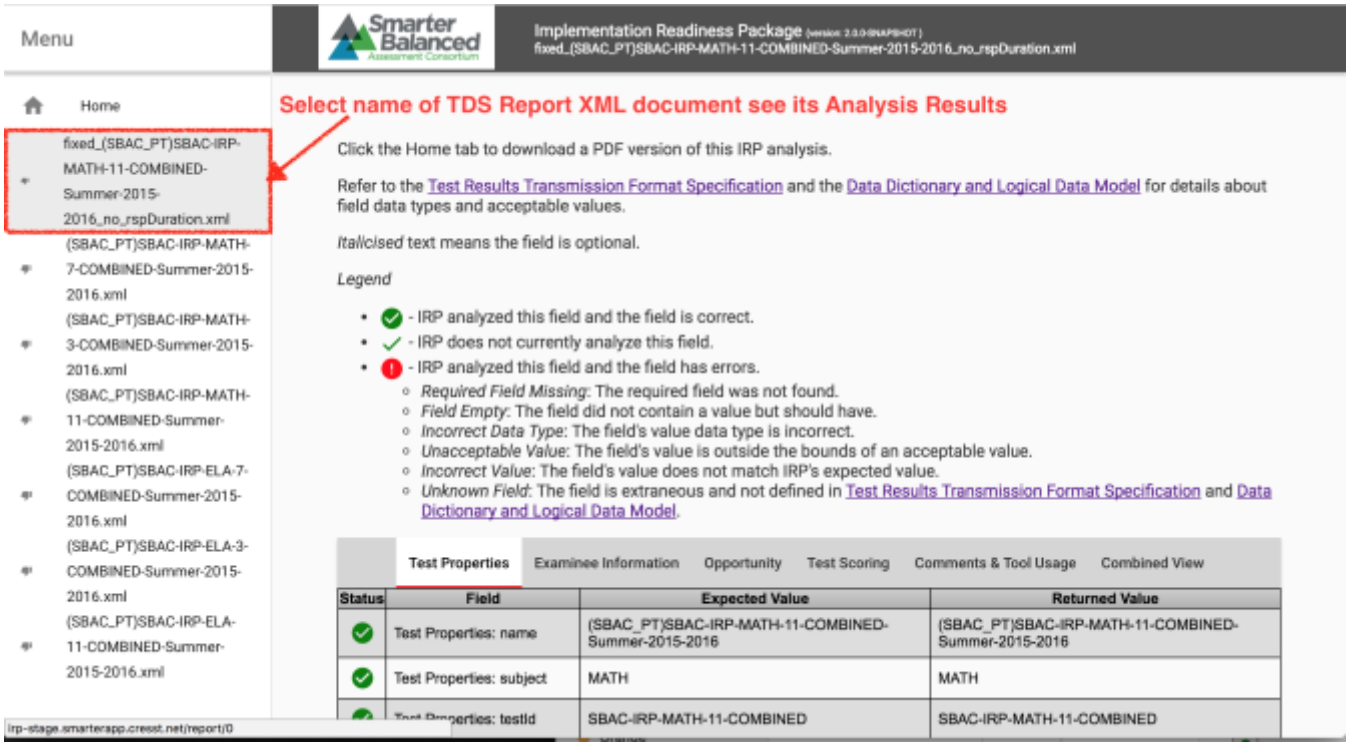


The Summary Table consists of six columns. Each row of the Summary Table represents the summary status of each TDS Report XML document submitted to IRP for analysis. The columns are described below:

1. **TDS Report File Name**: This column represents the filename if using Manual Mode or the test name if using Automation Mode
2. **Valid XML:** Indicates whether or not the XML passed Schema Validation
3. **Valid Test Name**: IRP expects the Test name to be one provided in the IRP Package (resources). If a different Test name is found, then value for this column will be 'No'.
4. **Valid Examinee:** IRP expects the Examinee (Student) to match one the IRP Student that is supposed to take associated Test. IRP provides a mapping of IRP Students to IRP Tests.
5. **Valid Test Scoring:** This column represents the overall validation of Test scoring.
6. **All Fields Valid:** If any of the fields analyzed contain an error the value of this column will be 'No'. Drill down into the TDS Report XML document Analysis Results to see more details.

To drill down and get a more detailed view of the results, click on the TDS Report XML document file name on the left of the screen beneath the Home link as shown in Figure 7: Selecting Results for a TDS Report.

*Figure 7: Selecting Results for a TDS Report*

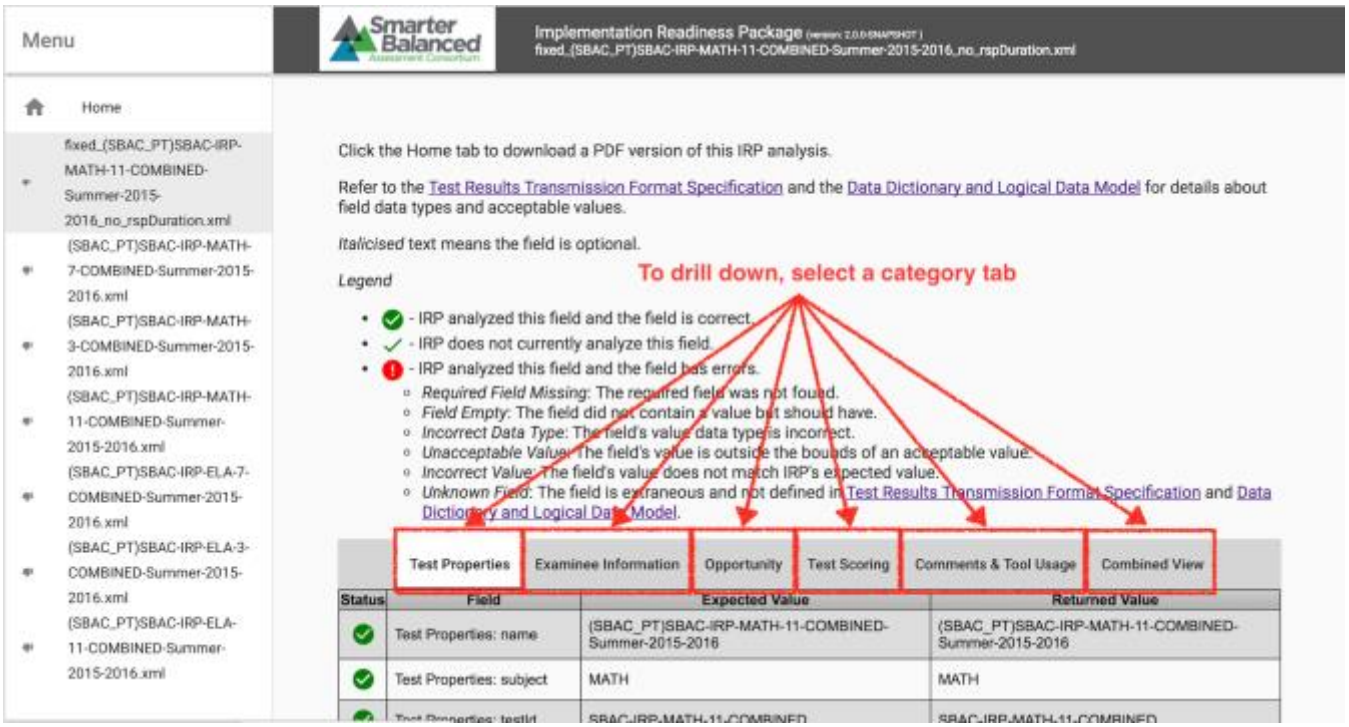



Once the name of a TDS Report XML document has been selected, further drill-down into specific sections of TDS Report XML Analysis can be performed. There are five *Analysis Category* views plus a *Combined View* as shown in Figure 9: Analysis Categories. The Analysis Categories represent sections of a TDS Report XML document as shown in Figure 8: TDS Report XML Document Example that IRP groups together for analysis. The Combined View contains results from all five Analysis Categories in one view.

*Figure 8: TDS Report XML Document Example*

```
<TDSReport>
  <Test grade="03" mode="online" contract="SBAC_PT" bankKey="187" testId="SBAC-IRP-MATH-3-COMBINED" subject="MATH" name="(SBAC_PT)SBAC-IRP-MATH-3-C
  <Examinee isDemo="0" key="22"> </Examinee>
  <Opportunity reportingVersion="0" clientName="SBAC_PT" qaLevel="" windowId="ANNUAL" sessionId="Pro-13" taName="Proctor Proctor" taId="proctor001@
  <Comment context="TestItem" itemPosition="1" date="2014-04-14T10:52:02.440">Walk 1&amp;amp;#92;2 miles&amp;amp;#10;BLAH BLAH</Comment>
  <Comment context="GlobalNotes" itemPosition="" date="2014-04-09T13:38:31.327">Some comments about the test in general</Comment>
  <ToolUsage type="Highlight" code="TDS_Highlight1">
    <ToolPage page="2" groupId="I-200-22489" count="1" />
  </ToolUsage>
</TDSReport>
```

*Figure 9: Analysis Categories*




The five categories are:

1. **Test Properties:** Analysis results on all the Test element attributes found in the TDS Report XML document. An example is shown in Figure 10: Test Properties Category.
2. **Examinee Information:** Analysis results on all the Student's information from the Examinee element in the TDS Report XML document. IRP expects information from an IRP Student to be found in this section. An example is shown in Figure 11: Examinee Category.
3. **Opportunity:** Analysis results on the Test as represented by the Opportunity element and its children in the TDS Report XML document. This includes Test name, Items, and Item Scores. IRP expects the Test to be one of the IRP Tests provided in the IRP Package. This Analysis Category excludes Test Scoring. An example is shown in Figure 12: Opportunity Category.
4. **Test Scoring:** Analysis results on Test Scores found in the Score child elements of the Opportunity element. This Analysis Category is displayed differently from other categories as shown in Figure 14: Test Scoring Category and Figure 15: Test Scoring Category Continued.

5. **Comments & Tool Usage:** Analysis results on the optional Comments and ToolUsage elements in the TDS Report XML document. An example is shown in Figure 16: Comments and Tool Usage Category.

*Test Properties Category*

*Figure 10: Test Properties Category*

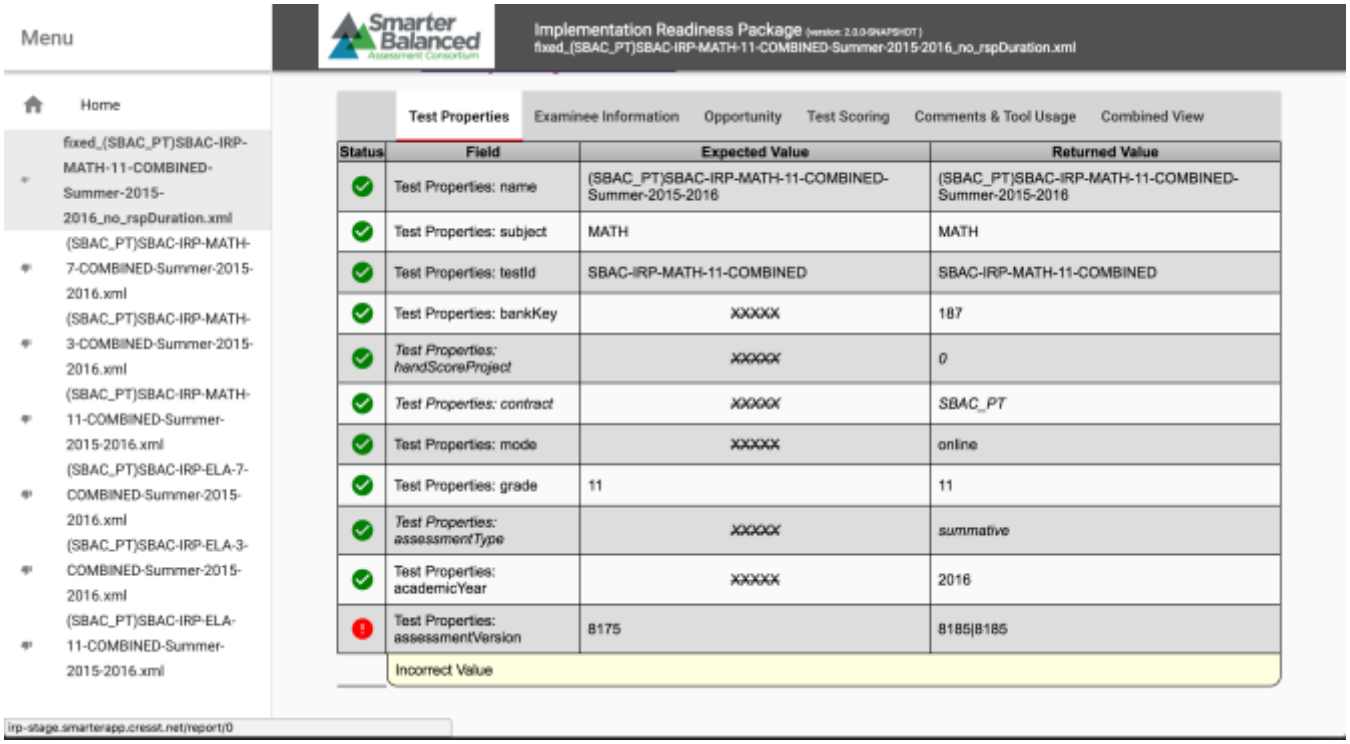


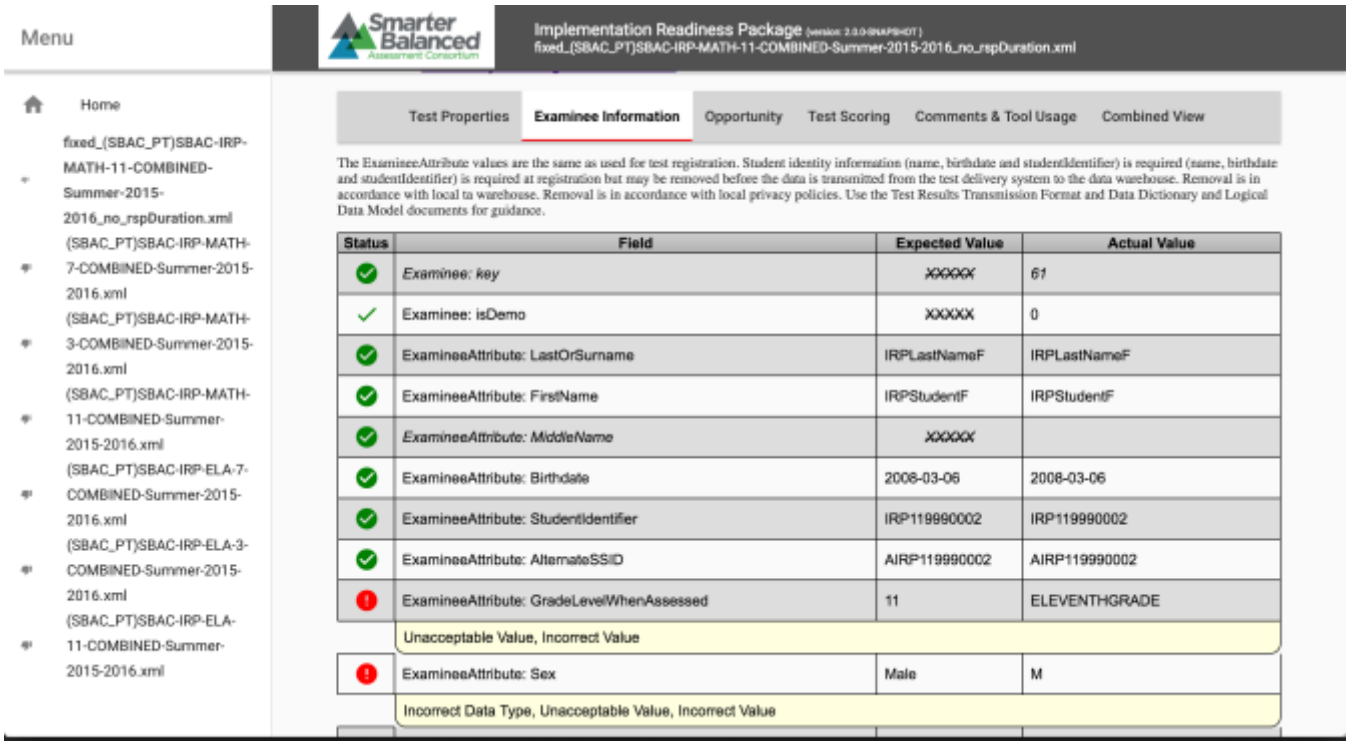

Figure 10: Test Properties Category is an example of how most Analysis Categories are laid out; Test Scoring is the exception (see Test Scoring Category for details). The results shown is displayed as a table with four columns:

1. **Status:** An icon provides a quick visual guide indicating the correctness of the field.
2. **Field:** The name of the field that was analyzed.
3. **Expected Value:** If IRP expects a particular value to be found in the field, it will be displayed in this column. If IRP does not have an expected value, then five crossed out Xs will be displayed: ~~XXXXX~~.
4. **Returned Value:** The value of the field in the TDS Report XML document.

When there are errors, the error will be displayed in the row underneath the row containing the field that was analyzed. The error codes are described in the legend above the Analysis Results table.

*Examinee Category*

*Figure 11: Examinee Category*




Examinee Category contains analysis results related to the Student section of the TDS Report XML document. Analysis of *ExamineeAttribute* and *ExamineeRelationship* are considered "Complex Field Analysis". Those elements consist of similarly named attributes whose values have to be inspected for meaning. For example, *ExamineeAttribute* contains four attributes: *name*, *value*, *context*, *contextDate.* Multiple *ExamineeAttribute* elements can exist as a child element of the *Examinee* element of the TDS Report XML document. To discern the meaning of each *ExamineeAttribute* element, the data in the *name* and *value* attributes have to be analyzed.

IRP performs "Complex Field Analysis" on *ExamineeAttribute* and *ExamineeRelationship* to perform actions such as finding an IRP Student's information.

*Opportunity Category*

*Figure 12: Opportunity Category*

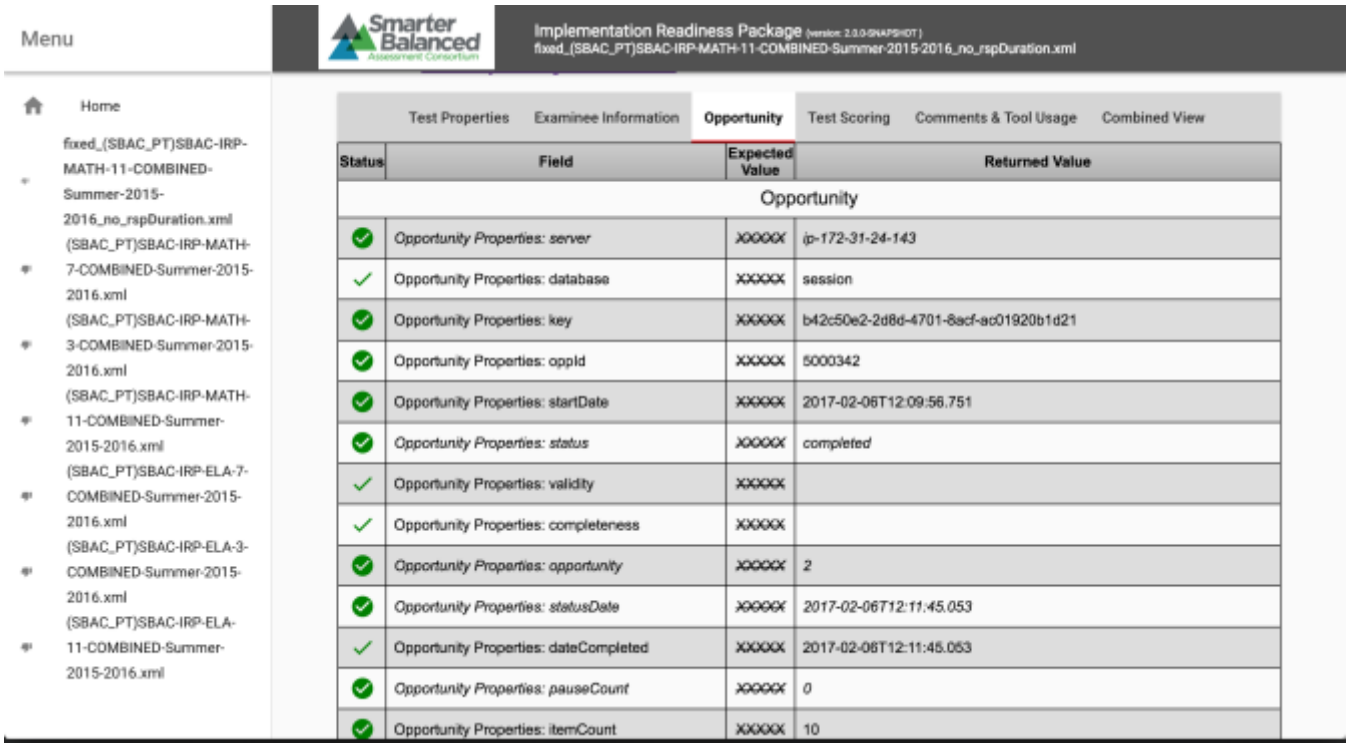
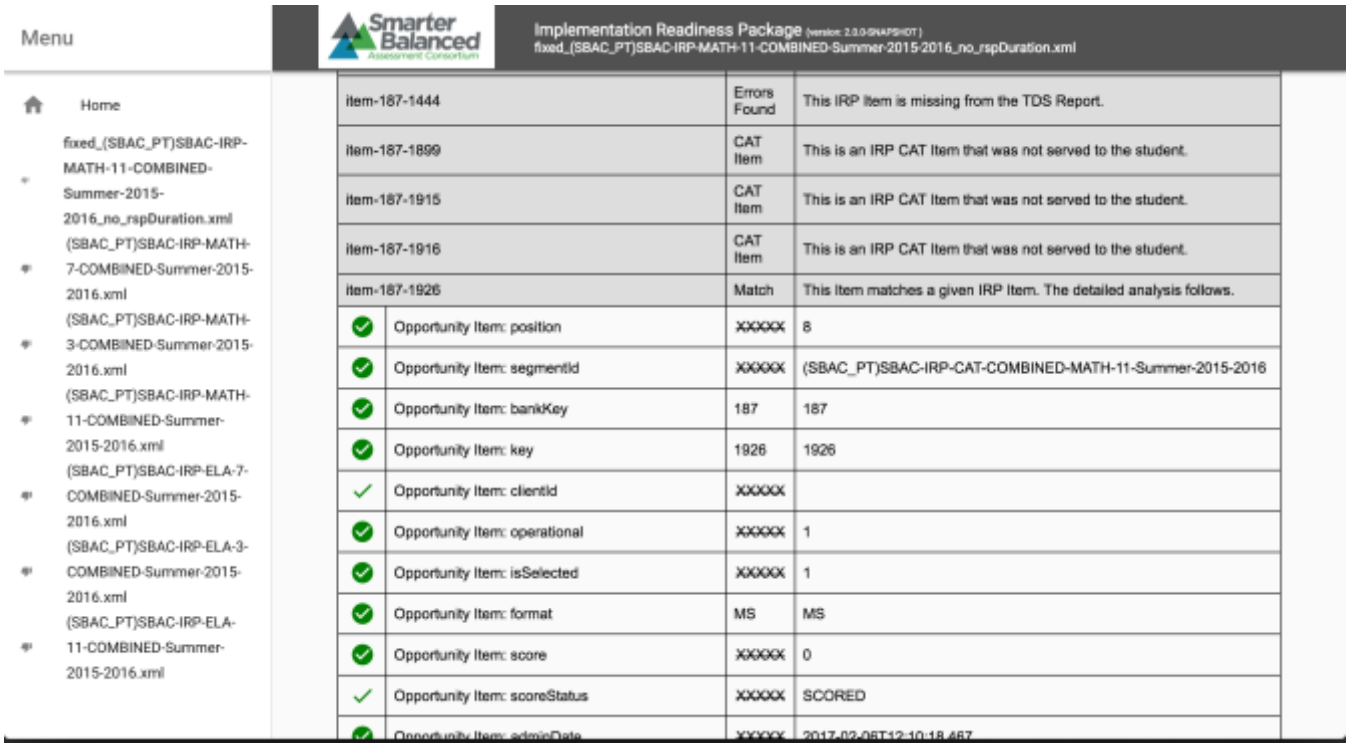



*Figure 13: Opportunity Item Scoring*

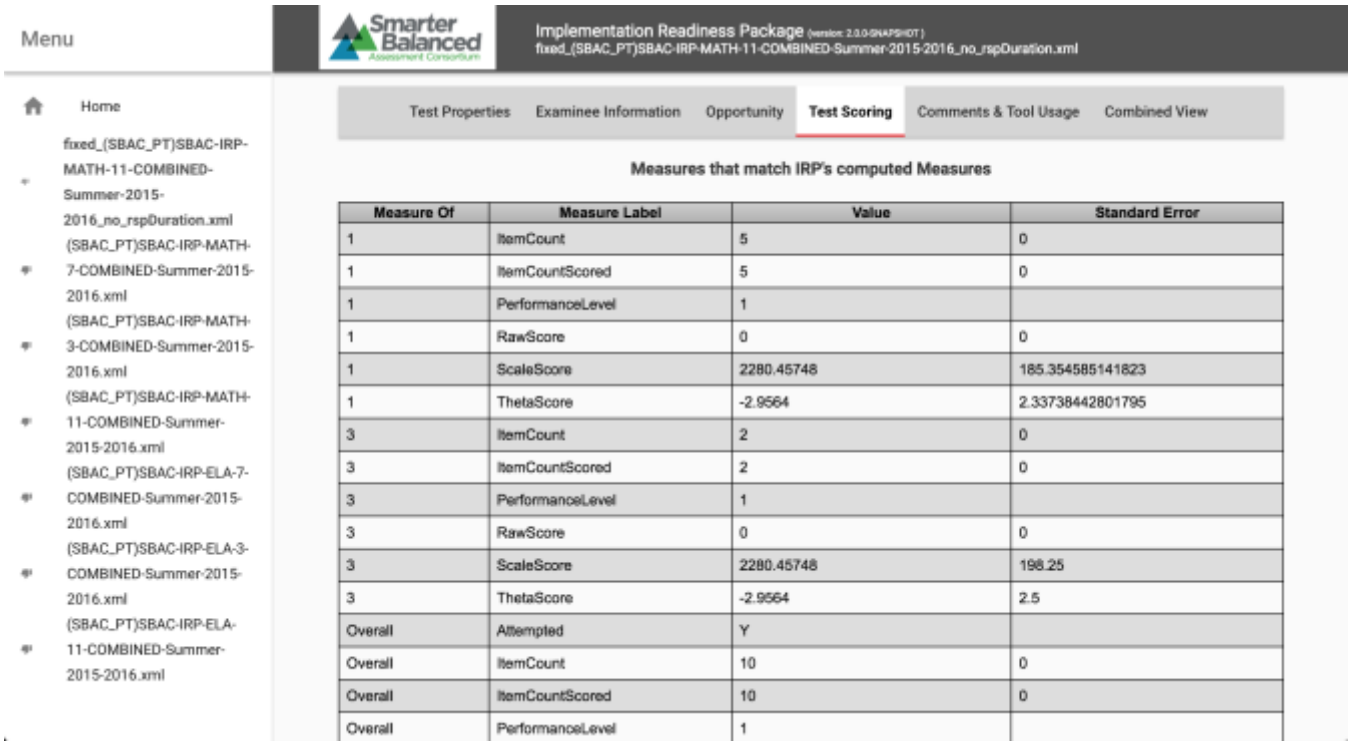


The Opportunity Category contains Analysis Results for the Test and Items as shown in both Figure 12: Opportunity Category and Figure 13: Opportunity Item Scoring.

The Item Scores are analyzed by comparing Item Scores found in the TDS Report XML document to scores calculated by IRP's own Item Scoring Engine. IRP submits the Item response to its Item Scoring Engine, then compares the results found in the TDS Report XML document. The Item Scoring Engine used by IRP is from the Smarter Balanced Open Source Assessment Delivery System.

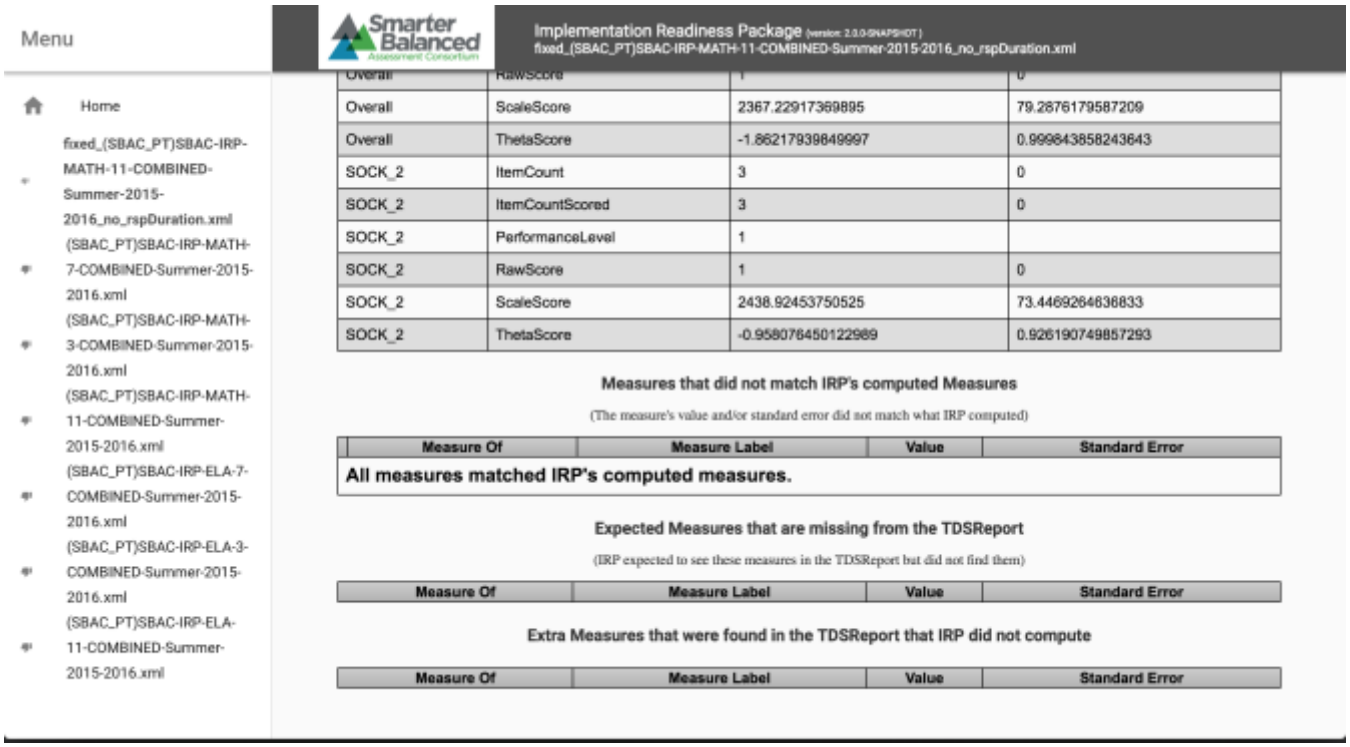
### *Test Scoring Category*

*Figure 14: Test Scoring Category*




| Measure Of | Measure Label | Value | Standard Error |
|---|---|---|---|
| 1 | ItemCount | 5 | 0 |
| 1 | ItemCountScored | 5 | 0 |
| 1 | PerformanceLevel | 1 | |
| 1 | RawScore | 0 | 0 |
| 1 | ScaleScore | 2280.45748 | 185.354585141823 |
| 1 | ThetaScore | -2.9564 | 2.33738442801795 |
| 3 | ItemCount | 2 | 0 |
| 3 | ItemCountScored | 2 | 0 |
| 3 | PerformanceLevel | 1 | |
| 3 | RawScore | 0 | 0 |
| 3 | ScaleScore | 2280.45748 | 198.25 |
| 3 | ThetaScore | -2.9564 | 2.5 |
| Overall | Attempted | Y | |
| Overall | ItemCount | 10 | 0 |
| Overall | ItemCountScored | 10 | 0 |
| Overall | PerformanceLevel | 1 | |

*Figure 15: Test Scoring Category Continued*



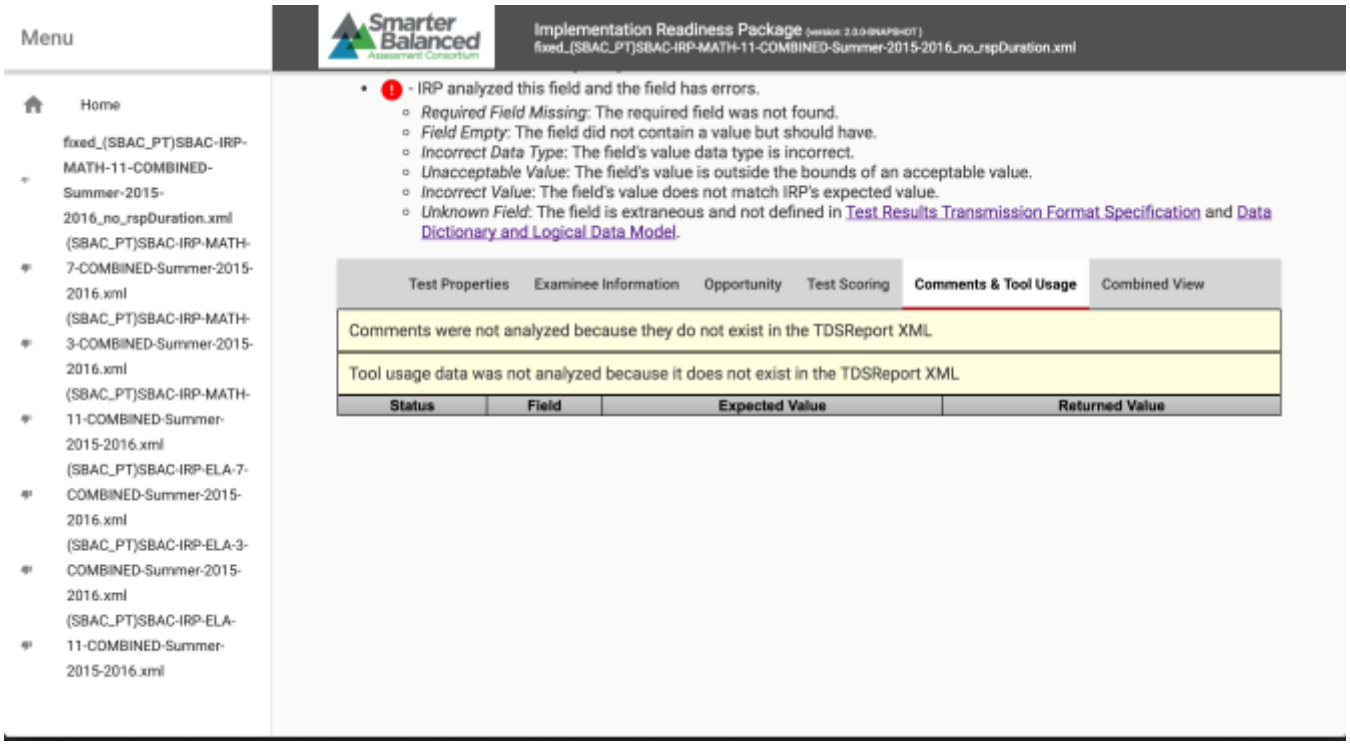

IRP analyzes Test Scoring by removing scoring information from the TDS Report XML document and then submits the scoreless document to its own Test Scoring Engine. The results from IRP's Test Scoring Engine are compared to the results found in the TDS Report XML document. IRP uses Smarter Balanced Open Source Test Integration System as its Test Scoring Engine.

IRP's Test Scoring Category is split into four sections:

1. **Measures that match IRP's computed Measures:** This section lists all the Scoring measures that match those computed by IRP.
2. **Measures that did not match IRP's computed Measures:** This section lists all the Scoring measures that did not match those computed by IRP.
3. **Expected Measures that are missing from the TDSReport:** This section lists measures that IRP did not find in the TDS Report XML document that were computed by IRP's Test Scoring Engine.
4. **Extra Measures that were found in the TDSReport that IRP did not compute:** This section lists measures that IRP's Test Scoring Engine did not compute but were found in the TDS Report XML document.

### *Comments and Tool Usage Category*

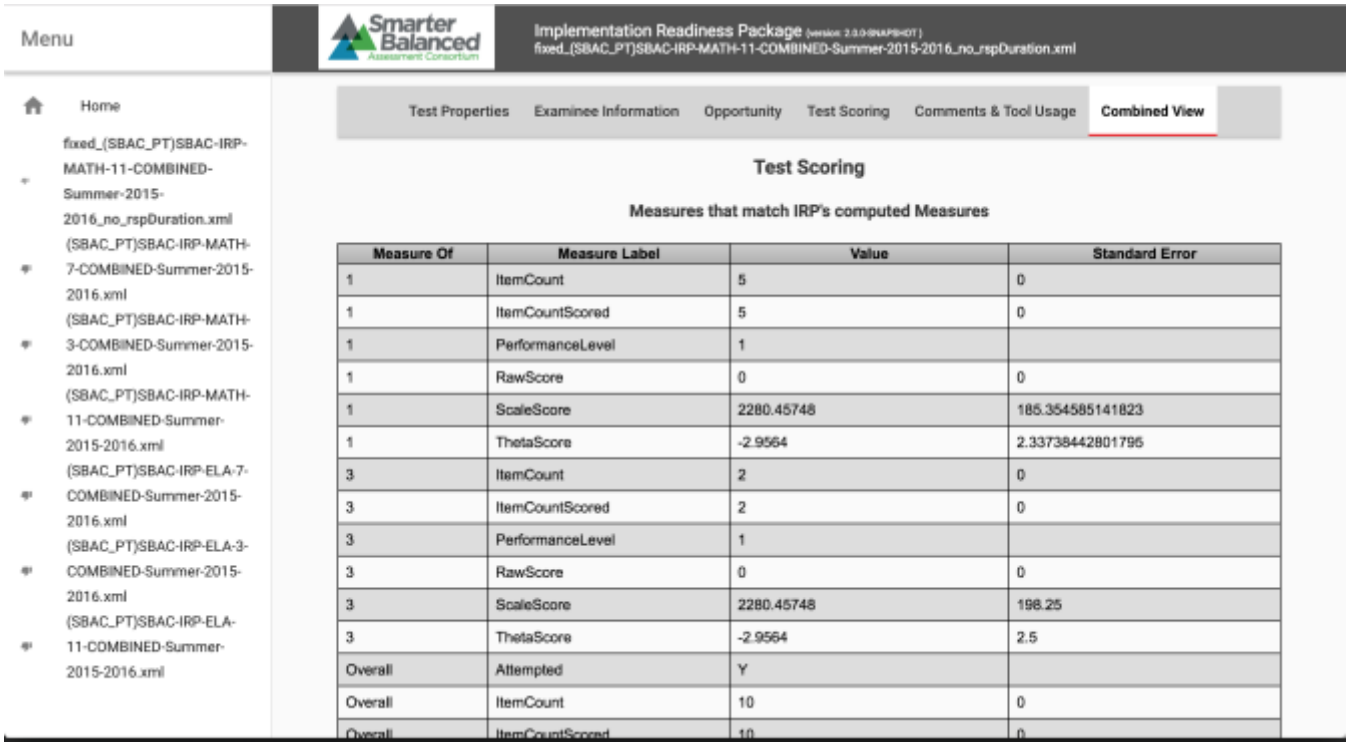
*Figure 16: Comments and Tool Usage Category*




Comments and Tool Usage are optional elements in the TDS Report XML document. If these are missing from an TDS Report XML document, then the message shown in Figure 16: Comments and Tool Usage Category will be shown. Otherwise, an analysis breakdown will be displayed similarly to other Analysis Categories.

*Combined View*

*Figure 17: Combined View*




The Combined View as shown in Figure 17: Combined View combines all five Analysis Categories into one view. This view starts with Test Scoring analysis results followed by the Test Properties Category, Examinee Information Category, Opportunity Category, and Comments & Tool Usage Category.

## CAT Engine Analysis

The IRP (Implementation Readiness Package) Computer Adaptive Testing (CAT) Analysis module is a service offered to vendors that will evaluate the readiness of their test delivery system. Using data and parameters supplied by IRP (which includes: the item pool, item parameters, the test blueprints, and student information), vendors can use the internal capabilities of their own CAT engine to simulate data from computer-adaptive test instances. The results extracted from the test instances contain detailed and summary statistics of how the CAT engine performed. Vendors can export this information into a comma separated file format (CSV), and then upload the data to an IRP module for analysis. The IRP module evaluates the vendor's CAT engine against a minimum and/or suggested Smarter Balanced standards across several dimensions:

- Theta score recovery
- Classification accuracy of achievement level descriptors.
- Item exposure rate
- Bias of estimated proficiencies
- Blueprint fulfillment/violations

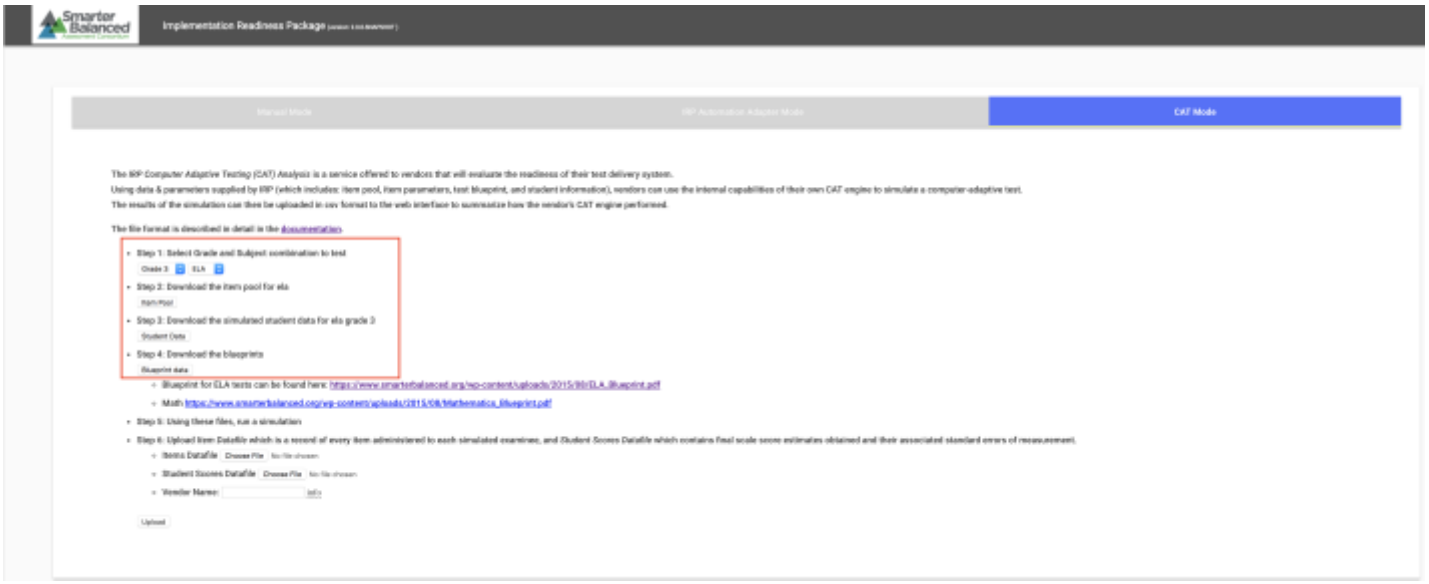
After this analysis, IRP will generate a response back to the vendor in the form of a standardized report that contains the results of the evaluation. Vendors are free to adjust their CAT engine and repeat the process as many times as they like. Both an HTML and PDF report are available to the vendor.

## Analysis Steps

The CAT Analysis module of IRP runs with the following steps:

1. The vendor downloads a *Simulation Package*, which includes all the information needed to run the simulation. This package includes: the item pool, simulated responses, and blueprint specifications. Figure 18: CAT Simulation Package Download shows where to download the Simulation Package.

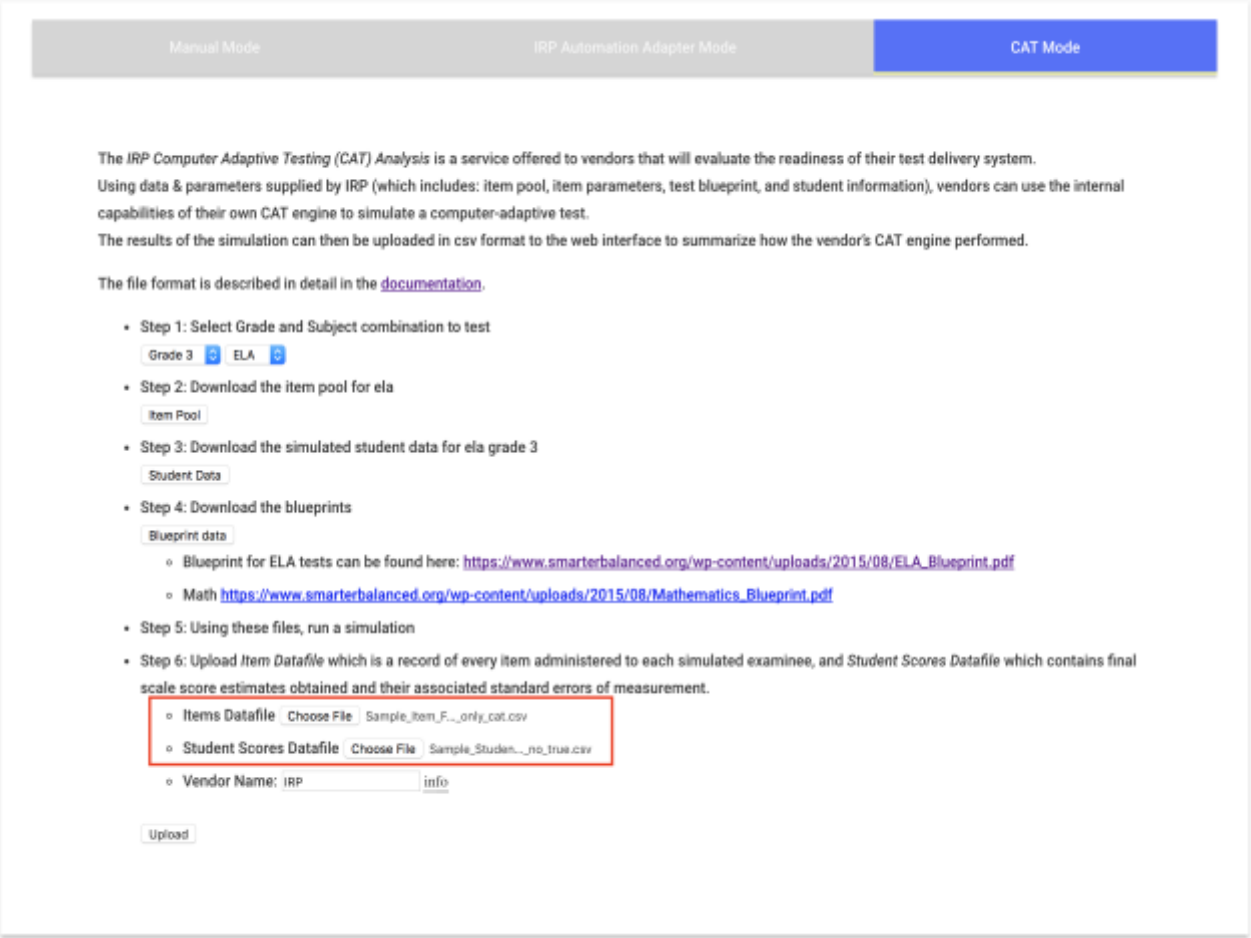
*Figure 18: CAT Simulation Package Download*




2. The vendor uses this information to run a simulation using their own CAT engine.
3. Once the vendor's CAT simulation is complete, the results need to be uploaded to IRP for evaluation. The uploaded data must include two components:
    a. **Items Datafile**: a record of every item administered to each simulees (as selected by the Test Delivery System's adaptive algorithm)
    b. **Student Scores Datafile**: final scale score estimates obtained and their associated standard errors of measurement.

    Data files corresponding to these two components must be delivered in a comma-separated format as described in the Data File Format section below.

*Figure 19: CAT Data File Upload*




4. Once the uploaded data file is received from the vendor, IRP will examine the file to verify that its format conforms to the fields and data types described in Data File Specifications found below. If there are any errors detected, IRP will flag them and issue a report, and the Data Evaluation process will end. If no errors are found, IRP will proceed with the Data Evaluation process.

## Data File Format

The data files are expected to be in CSV format with a header row. The columns are described here.

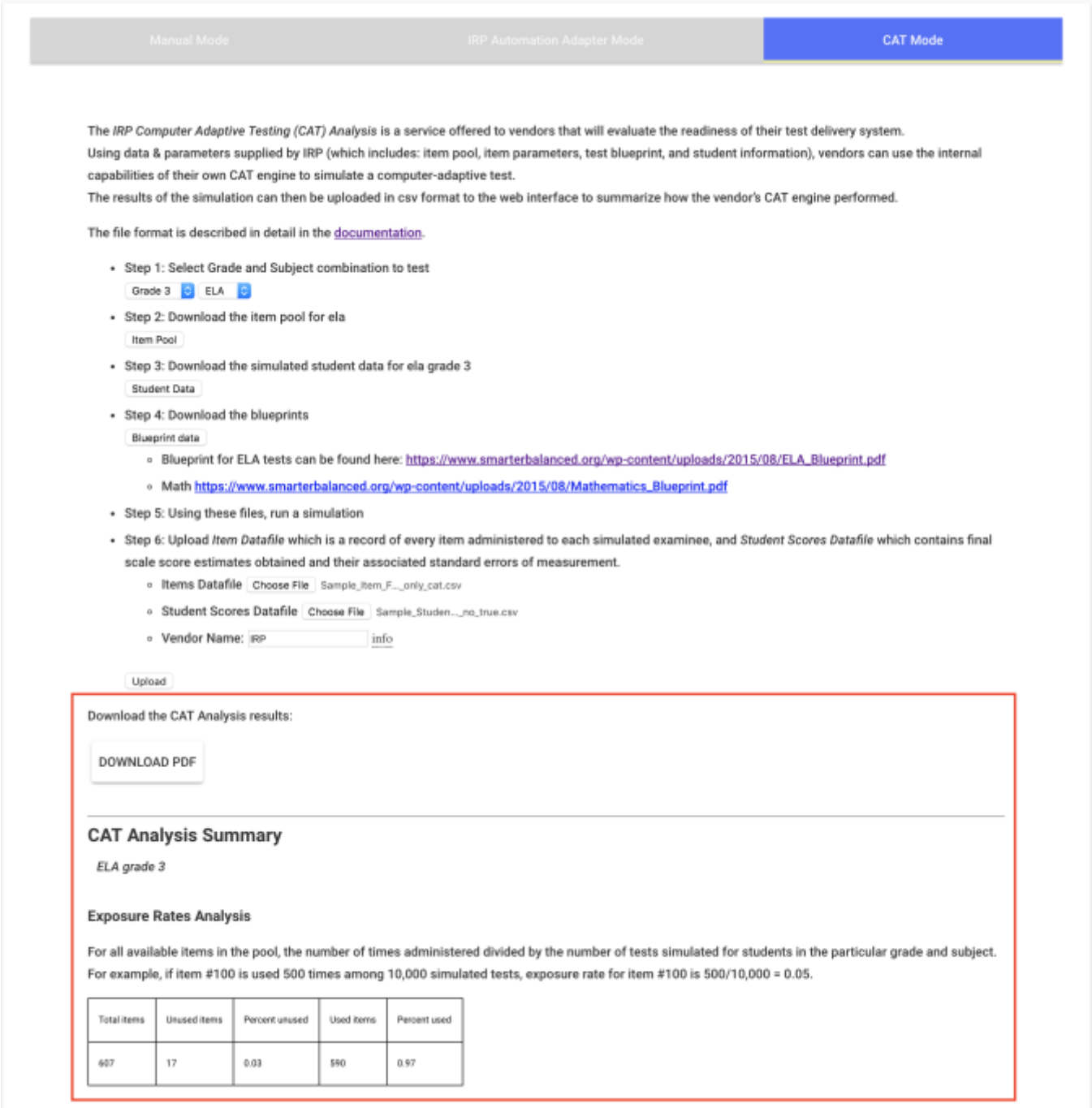
- **Student Score Data [Vendor Upload] 11 Columns:**
  - Field 1, unique student identifier (string/alphanumeric)
  - Field 2, overall score on logit/theta scale (numeric/decimal)
  - Field 3, standard error of measurement for the overall score on logit/theta scale (numeric/decimal)
  - Field 4, claim 1 score on logit/theta scale (numeric/decimal)
  - Field 5, standard error of measurement for the claim 1 score on logit/theta scale (numeric/decimal)
  - Field 6, claim 2 score on logit/theta scale (numeric/decimal)

  - Field 7, standard error of measurement for the claim 2 score on logit/theta scale (numeric/decimal)
  - Field 8, claim 3 score on logit/theta scale (numeric/decimal)
  - Field 9, standard error of measurement for the claim 3 score on logit/theta scale (numeric/decimal)
  - Field 10, claim 4 score on logit/theta scale (numeric/decimal)
  - Field 11, standard error of measurement for the claim 4 score on logit/theta scale (numeric/decimal)
- **Items Administered [Vendor Upload] 3 Columns:**
  - Field 1, unique student identifier (string/alphanumeric)
  - Field 2, unique item identifier (string/alphanumeric)
  - Field 3, item score (numeric/integer)

### Analysis Results

Once the data is uploaded and validated then the results will be shown as given by example in the following figures starting with Figure 20: CAT Analysis Results.

*Figure 20: CAT Analysis Results*



### Distribution of exposure rates

The count of the number exposure rates falling into each percentage range with a bin size of 0.1

| Bin | 0.00 to 0.10 | 0.10 to 0.20 | 0.20 to 0.30 | 0.30 to 0.40 | 0.40 to 0.50 | 0.50 to 0.60 | 0.60 to 0.70 | 0.70 to 0.80 | 0.80 to 0.90 | 0.90 to 1.00 |
|---|---|---|---|---|---|---|---|---|---|---|
| Count | 470 | 114 | 5 | 0 | 0 | 0 | 0 | 0 | 1 | 0 |

### Score Summary

Score estimates are compared against the true/generating scores. Summaries are available as Average Bias and Root Mean Squared Error (RMSE) These summaries are also computed against the scores partitioned into deciles.

| | |
|---|---|
| Average Bias | 0.012 |
| Root Mean Squared Error (RMSE) | 0.311 |

*Average Bias by Decile*

| **Decile** | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|---|
| **Mean Bias** | 0.285 | 0.051 | 0.013 | 0.006 | 0.019 | -0.040 | -0.021 | -0.055 | -0.084 | -0.051 |

*RMSE by Decile*

| **Decile** | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|---|
| **RMSE** | 0.576 | 0.335 | 0.277 | 0.262 | 0.266 | 0.279 | 0.221 | 0.240 | 0.246 | 0.253 |

### Precision Summary

Precision is scored by the overall mean standard error of measurement and by claim.

| **Claim** | Overall | 1 | 2 | 3 | 4 |
|---|---|---|---|---|---|
| Mean Standard Error of Measurement | 0.294 | 0.482 | 0.523 | 0.808 | 0.686 |

- For every true score and estimated scale score, there are corresponding levels. For the overall score, students are assigned to one of four possible levels.
- Cut scores by grade and subject can be found here (pp. 8-9): http://www.smarterapp.org/documents/TestScoringSpecs2014-2015.pdf

### Classification Levels

For every true score and estimated scale score, there are corresponding levels. For the overall score, students are assigned to one of four possible levels. *Classification accuracy* is the number of correctly labeled students compared to the generating student levels divided by the total number of students. This is a sum of the diagonal along the table and dividing by the total number of students.

| Theta Cut between Levels 1 and 2 | Theta Cut between Levels 2 and 3 | Theta Cut between Levels 3 and 4 |
|---|---|---|
| -1.646 | -0.888 | -0.212 |

Score Levels

| | True Level | | | |
|---|---|---|---|---|
| Estimated Level | 1 | 2 | 3 | 4 |
| 1 | 334 | 44 | 0 | 0 |
| 2 | 42 | 207 | 30 | 0 |
| 3 | 0 | 36 | 132 | 13 |
| 4 | 0 | 0 | 23 | 139 |

Classification accuracy: 0.81

### Blueprint Violations

Blueprint violations are summarized from here

- ELA https://www.smarterbalanced.org/wp-content/uploads/2015/08/ELA_Blueprint.pdf
- Math https://www.smarterbalanced.org/wp-content/uploads/2015/08/Mathematics_Blueprint.pdf
  Summaries are given based on the number of tests that fall under, over or match the blueprint specification.

*Violation Counts*

| Specification | Min | Max | Under | Match | Over |
|---|---|---|---|---|---|
| Claim 1 | 14 | 16 | 0 | 1000 | 0 |
| Claim 1 , Literary, Targets: 1, 2, 3, 4, 5, 6, 7 | 7 | 8 | 0 | 1000 | 0 |
| Claim 1 , Informational, Targets: 8, 9, 10, 11, 12, 13, 14 | 7 | 8 | 0 | 1000 | 0 |
| Claim 1 , DOK = 2 | 7 | NA | 2 | 998 | 0 |
| Claim 1 , DOK >= 3 | 2 | NA | 0 | 1000 | 0 |
| Claim 1 , Targets: 2 | 1 | 2 | 8 | 992 | 0 |
| Claim 1 , Targets: 4 | 1 | 2 | 42 | 958 | 0 |
| Claim 1 , Targets: 1, 3, 5, 6, 7 | 3 | 6 | 0 | 1000 | 0 |
| Claim 1 , Targets: 9 | 1 | 2 | 0 | 1000 | 0 |
| Claim 1 , Targets: 11 | 1 | 2 | 13 | 987 | 0 |
| Claim 1 , Targets: 8, 10, 12, 13, 14 | 3 | 6 | 0 | 1000 | 0 |
| Claim 1 , Lit Long passage, Targets: 1, 2, 3, 4, 5, 6, 7 | 1 | NA | 401 | 599 | 0 |
| Claim 1 , Lit Short passage, Targets: 1, 2, 3, 4, 5, 6, 7 | 0 | 2 | 0 | 73 | 927 |
| Claim 1 , Info Long passage, Targets: 8, 9, 10, 11, 12, 13, 14 | 1 | NA | 226 | 774 | 0 |
| Claim 1 , Info Short passage, Targets: 8, 9, 10, 11, 12, 13, 14 | 0 | 2 | 0 | 178 | 822 |
| Claim 1 , 0-1 short text, Targets: 2, 4 | 0 | 1 | 0 | 1000 | 0 |
| Claim 1 , 0-1 short text, Targets: 1, 3, 5, 6, 7 | 0 | 0 | 0 | 1000 | 0 |

There is an option for downloading the analysis as a PDF or viewing as html. Each section is documented to describe the various statistics.

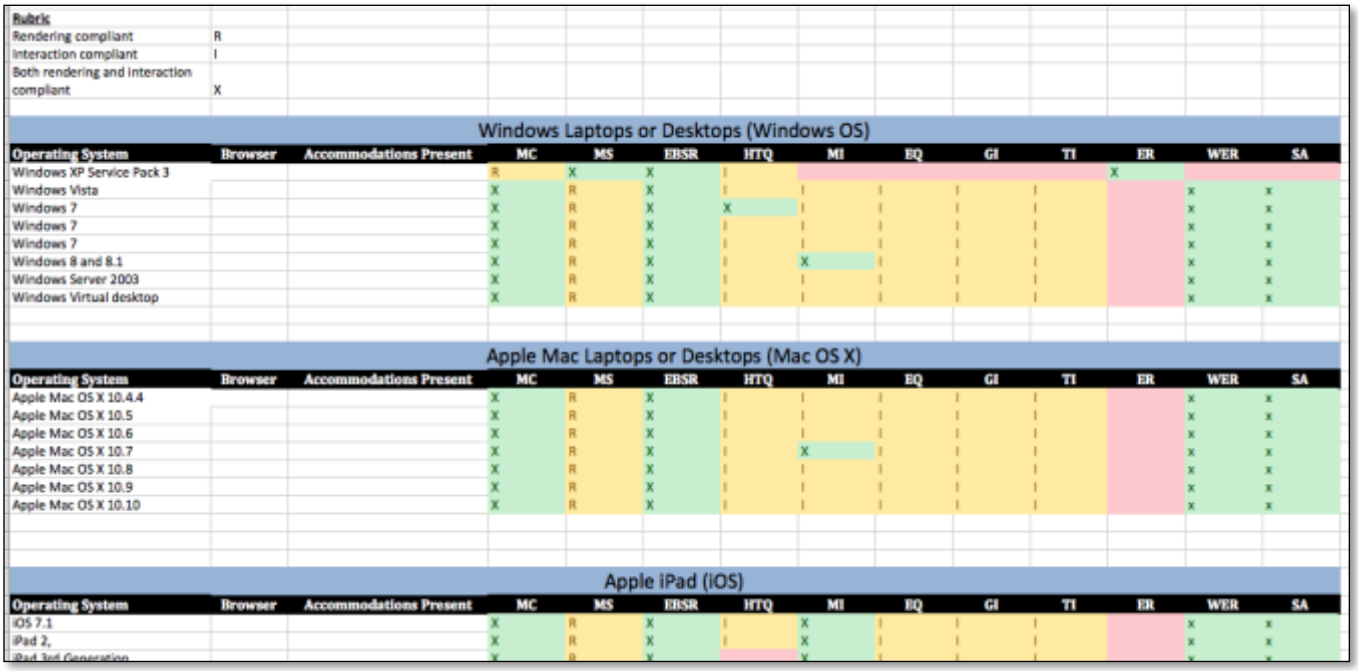
# Item Rendering & Interactivity

As part of the IRP download package, there is an Excel document entitled *IRP v1 Module A Form*, which is a worksheet that can be used to verify compliance with item Rendering and Interaction requirements across different devices, operating systems and item types (see requirements spelled out in the Implementation Readiness Package document hosted on the SmarterApp.org website located here: http://www.smarterapp.org/specs/ImplementationReadinessPackage.html.

Below is a sample screenshot of this worksheet. Under the appropriate device type category, fill in the OS tested, the Smarter Balanced Secure Browser version tested, any Accommodations (i.e. American Sign Language, text color/background, text-to-speech, etc.) in place, and the results of the rendering and interaction for each of the item types, according to the rubric listed above the table (reproduced below). In actual use, the vendor must add as many rows as needed under each device type for each of the system combinations tested. This document should be saved as part of your Vendor Assessment Delivery System IRP Compliance results.

**Rubric (enter the appropriate letter, as follows):**

| Observed Compliance | Representation in Worksheet | Color of Worksheet Cell |
|---|---|---|
| Rendering Compliant | R | Yellow |
| Interaction Compliant | I | Yellow |
| Both Rendering and Interaction Compliant | X | Green |
| Neither Rendering nor Interaction Compliant | [blank] | Red |
| Not Tested | NT | White |
| | | |



Rubric
Rendering compliant R
Interaction compliant I
Both rendering and interaction compliant X

Windows Laptops or Desktops (Windows OS)

| Operating System | Browser | Accommodations Present | MC | MS | EBSR | HTQ | MI | BQ | GI | TI | ER | WER | SA |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Windows XP Service Pack 3 | | | R | X | X | I | | | | | X | | |
| Windows Vista | | | X | R | X | I | I | I | I | I | | x | x |
| Windows 7 | | | X | R | X | X | I | I | I | I | | x | x |
| Windows 7 | | | X | R | X | I | I | I | I | I | | x | x |
| Windows 7 | | | X | R | X | I | I | I | I | I | | x | x |
| Windows 8 and 8.1 | | | X | R | X | I | X | I | I | I | | x | x |
| Windows Server 2003 | | | X | R | X | I | I | I | I | I | | x | x |
| Windows Virtual desktop | | | X | R | X | I | I | I | I | I | | x | x |

Apple Mac Laptops or Desktops (Mac OS X)

| Operating System | Browser | Accommodations Present | MC | MS | EBSR | HTQ | MI | BQ | GI | TI | ER | WER | SA |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Apple Mac OS X 10.4.4 | | | X | R | X | I | I | I | I | I | | x | x |
| Apple Mac OS X 10.5 | | | X | R | X | I | I | I | I | I | | x | x |
| Apple Mac OS X 10.6 | | | X | R | X | I | I | I | I | I | | x | x |
| Apple Mac OS X 10.7 | | | X | R | X | I | X | I | I | I | | x | x |
| Apple Mac OS X 10.8 | | | X | R | X | I | I | I | I | I | | x | x |
| Apple Mac OS X 10.9 | | | X | R | X | I | I | I | I | I | | x | x |
| Apple Mac OS X 10.10 | | | X | R | X | I | I | I | I | I | | x | x |

Apple iPad (iOS)

| Operating System | Browser | Accommodations Present | MC | MS | EBSR | HTQ | MI | BQ | GI | TI | ER | WER | SA |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| iOS 7.1 | | | X | R | X | I | X | I | I | I | | x | x |
| iPad 2, | | | X | R | X | I | X | I | I | I | | x | x |
| iPad 3rd Generation | | | X | R | X | | X | I | I | I | | x | x |

Figure 2. Sample Screenshot of Module A Form Worksheet

## Revision History

| Date | Version | Description | Author |
|---|---|---|---|
| 2017-02-28 | 2.0.0 | Initial document | Paul Espinosa |
| | | | |
| | | | |
| | | | |
| | | | |
| | | | |
| | | | |