# OSCAR REPORTS

OSCAR provides an array of reporting functionality intended to provide full transparency into any scoring project in real-time. A full list and description of reports is provided below.

**Individual Dashboard** – Individual scorers are provided a scoring dashboard that allows them to track their progress and accuracy throughout a scoring project with visibility into their QC results, scoring activity, mail, scoring assignment and recent activity.

**Supervisor Dashboard** – Scoring supervisors can identify any scoring or scorer issues from their Supervisor Dashboard that reports for scoring assignment, IRR & Validity, scoring activity and user activity.
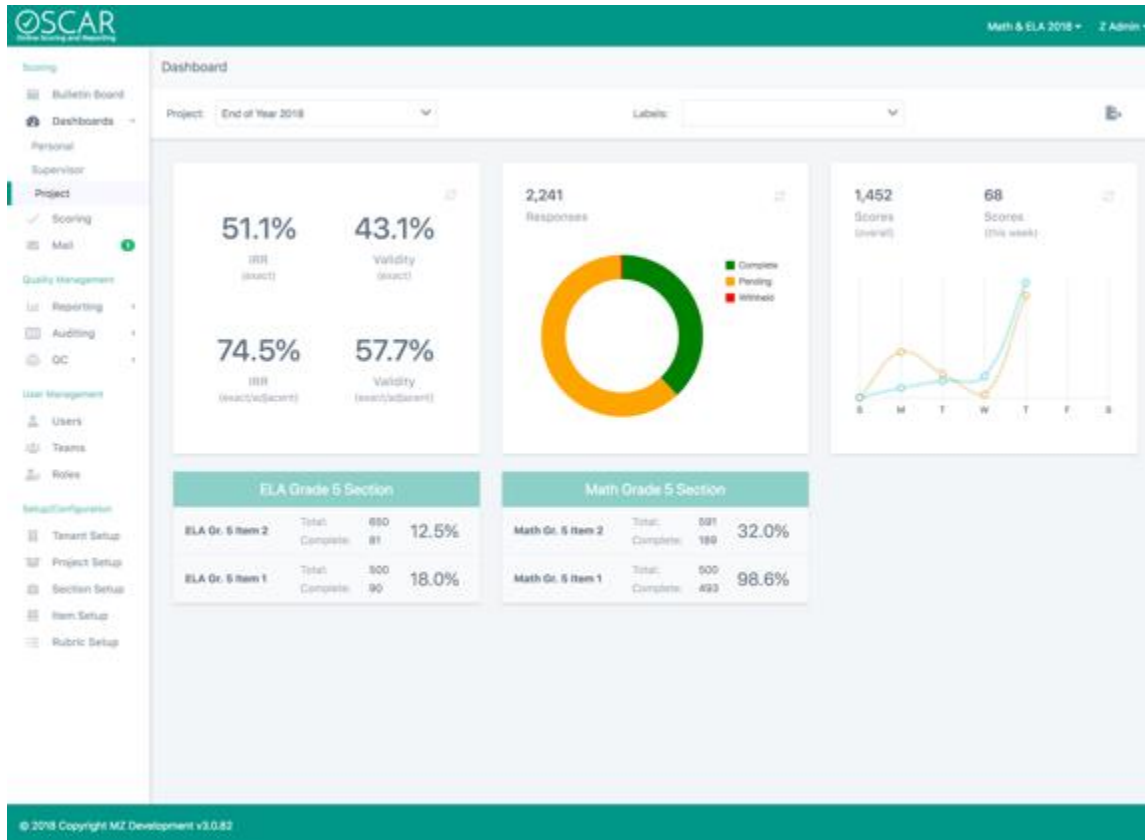
**Project Dashboard** – The project dashboard provides an overall snapshot of project activity and progress that includes IRR & Validity, project completion, scoring activity and more granular reports on each section.



**Completion Report** – The Completion Report allows scoring supervisors to monitor the status of every response loaded into the scoring platform. A scoring supervisor or administrator can monitor the state of the responses and determine how close an item/project is to completion. The Completion Report also presents a view of the number of responses requiring resolution reads as well as the number of back read scores that have been applied. This report allows the administrator and supervisor to know at what point to move scorers from one item to another as scoring is completed, and it ensures that back reads are not overlooked. If scoring a field test or research items, the report can be used to show a summary of responses that have been flagged because of difficulties in scoring.
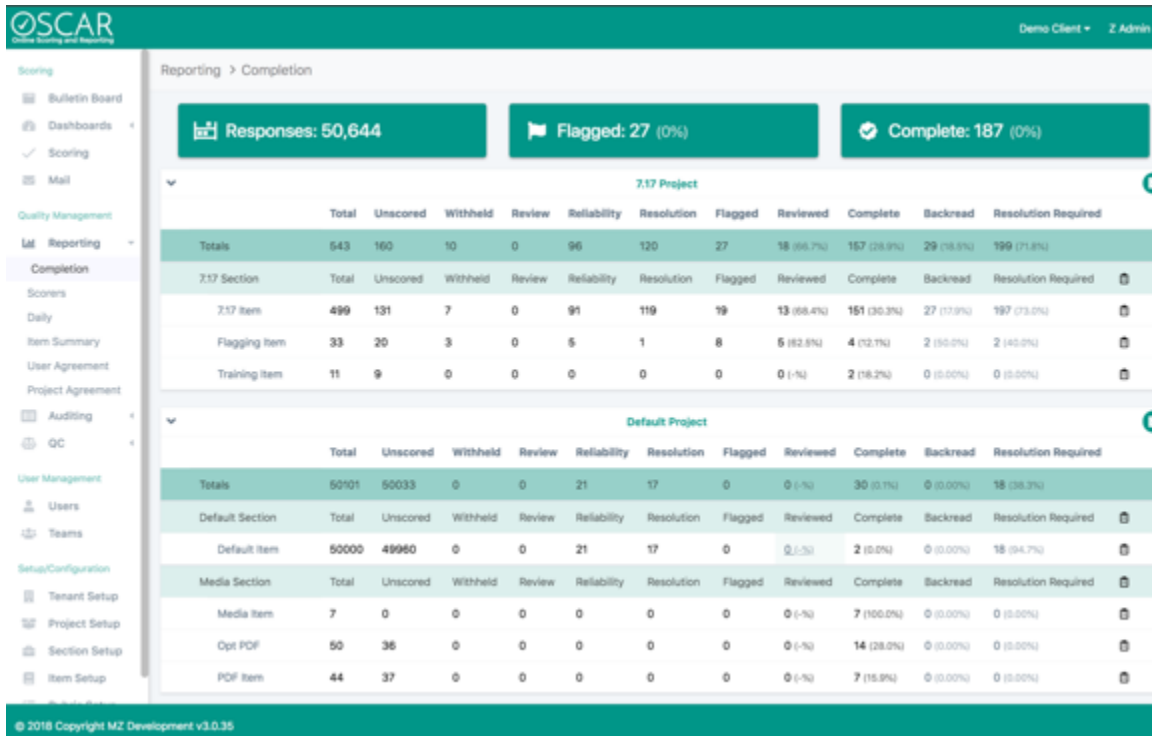
The Completion Report allows scoring supervisors to look at:

| | |
|---|---|
| Total | The total number of responses per item |
| Unscored | The number of responses that have not received a first or second score |
| Withheld | The number of responses that have been flagged as non-scorable or have been flagged due to a scorer's question regarding the scoring of that response. Responses remain in the withheld stage until they are released by to the pool by a scoring supervisor. |
| Review | The number of responses that have been routed to a "review" pool through a rubric condition code. |
| Reliability | The number of responses awaiting a second score |
| Resolution | The number of responses awaiting a rescore by a scoring supervisor as required by a project's resolution scoring rules |
| Flagged | The number of responses that have a flag associated with it. A flag will remain with the response through all states unless deleted by a supervisor. These flags have not been reviewed by a supervisor. |
| Reviewed | The number of responses with flags that have the flag reviewed. |
| Complete | The number of scorable responses that have completed the scoring process and have received either one or two scores depending on the requirements of the project |
| Backread | The number of responses that have received a manual rescore by a scoring supervisor |
| Resolution Required | The total number of responses that were rescored as required by a project's resolution scoring rules |

**Burndown Report** – The Burndown Report is a subset of the Completion Report and provides a forecast for project completion with the capability to see how reallocation of scorer resources might impact future scoring activity and project completion.

⊘SCAR
Online Scoring and Reporting

Math & ELA 2018 ▼    Z Admin ▼

Scoring
- Bulletin Board
- Dashboards
- Scoring
- Mail

Quality Management
- Reporting
- Auditing
- QC

User Management
- Users
- Teams
- Roles

Setup/Configuration
- Tenant Setup
- Project Setup
- Section Setup
- Item Setup
- Rubric Setup

Reporting > Completion > Burndown

ELA Grade 5 Section    ▼      ELA Gr. 5 Item 1    ▼                    Run Report

| Day | Scores | Time/Score | Completed/Scorer | Scorers | Completed | Remaining |
|---|---|---|---|---|---|---|
| History (1 week) | | | | | | |
| Sep 4 | 1 | 00:10 | 1.0 | 1 | 1 | 411 (82.20%) |
| Sep 5 | 9 | 00:01 | 1.0 | 1 | 1 | 410 (82.00%) |
| Sep 6 | 28 | 01:11 | 0.0 | 1 | 0 | 410 (82.00%) |
| Average | 12.7/d | 00:52 | 0.7/d | 1.0/d | 0.7/d | |
| Projected (1 week) | | | | | | |
| Sep 7 | 26 | 00:52 | 1 | 10 | 10 | 390 (78.00%) |
| Sep 10 | 26 | 00:52 | 1 | 10 | 10 | 380 (76.00%) |
| Sep 11 | 26 | 00:52 | 1 | 10 | 10 | 370 (74.00%) |
| Sep 12 | 26 | 00:52 | 1 | 10 | 10 | 360 (72.00%) |
| Sep 13 | 26 | 00:52 | 1 | 10 | 10 | 350 (70.00%) |

Complete/Scorer/Day    Scorers    Complete/Day    Responses
1                      10 ⇕       10              500

☑ M  ☑ T  ☑ W  ☑ T  ☑ F
☐ S  ☐ S

Scores left: 1036
Score-time left: 15.2h
Finish date: Nov 1 (100%)

History ▢ Projected

© 2018 Copyright MZ Development v3.0.82

**Scorer Report** – The Scorer Report is designed to allow scoring supervisors access to scorer progress for a project, item, or for all scoring results. By checking this report, the scoring supervisor can see how frequently a scorer has had quality checks using validity and calibration responses. To fully analyze the quality performance of the scorers, the supervisor can move from this report to the Daily Report (Frequency Distribution and Inter-Rater Reliability) and to the Quality Control Reports. For projects that require a quick turnaround, the Scorer Report is used to help the supervisor monitor the project and to use resources and time more efficiently.
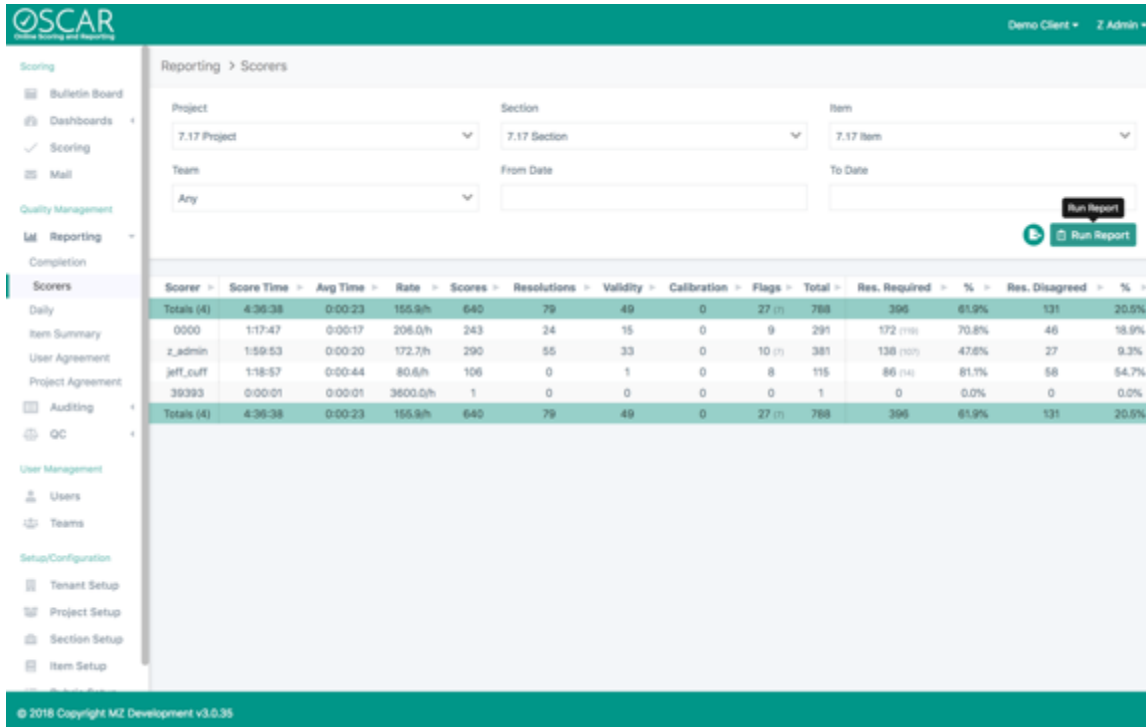


The Scorer Report allows scoring supervisors to look at:

| | |
|---|---|
| Last Login | The day and time a scorer last logged in to the scoring platform |
| Total System Time | The total time a scorer has logged in to the scoring platform during the project |
| Total Score Time | The total time spent accessing and scoring responses (from the time a response is accessed to the time it receives a score) |
| Average Score Time | The average time spent accessing and scoring a response (time spent reading a response and assigning a score) |
| Rate | The average number of responses scored per hour |
| Scores | The number of responses scored |
| Validity | The number of validity responses scored |
| Calibration | the number of calibration responses a scorer has been assigned to score for retraining purposes |
| Flags | The number of responses flagged as being non-scorable |
| Res. Required | The total number of responses scored by that scorer that required a resolution read |
| % | The % of responses scored by that individual scorer that required a resolution (third) read |
| Res. Disagreed | The number of responses scored by a scorer that required a third read and their score disagreed with the resolution read. |
| % | The % of responses scored by a scorer that required a third read and their score disagreed with the resolution read. |

**Daily Report** – The Daily Report is a tool that scoring supervisors use to identify drift, scorer issues and overall trait level scoring performance by comparing an individual scorer's score distribution to the score distribution of the entire scorer pool for an item. The Daily Report provides real-time scoring frequency distribution and inter-rater reliability (IRR) results by trait. Data can be filtered by day or by any defined time over the course of a project.

The totals line of the report shows the scoring distribution of the entire scoring pool for an item.

The following example is based on an item that is scored using a 3-point rubric (1-3). Based on the scoring of the entire group, the totals line shows the percentage of responses were assigned a score of one (1), two (2) and three (3). Underneath the totals line is the individual scorer's score distribution and the comparison to the aggregate scoring for the group (+/-). Scoring supervisors can identify if a scorer is struggling and allows them to identify an individual scorer's scoring tendencies or biases. By monitoring frequency distribution, the scoring supervisor can provide feedback directed at correcting any scoring tendencies or biases and by analyzing the data, the scoring supervisor can provide more meaningful scoring feedback and recalibration training for the scorers.

The Daily Report also shows how often a scorer is in exact agreement with another scorer when scoring a response that requires two scores. The report also shows the percentage of responses that a scorer assigned a score that was either adjacent high or adjacent low to another scorer, and it shows the percentage of responses that received a non-adjacent score. This information provides additional data to identify an individual scorer's scoring tendencies or biases.

**User Summary Report** – The User Summary report provides a comprehensive "snapshot" view of all user activity associated with a section/user/item. It includes time on task, IRR/Validity/Resolution overview, QC overview, item summary overview and recent activity.
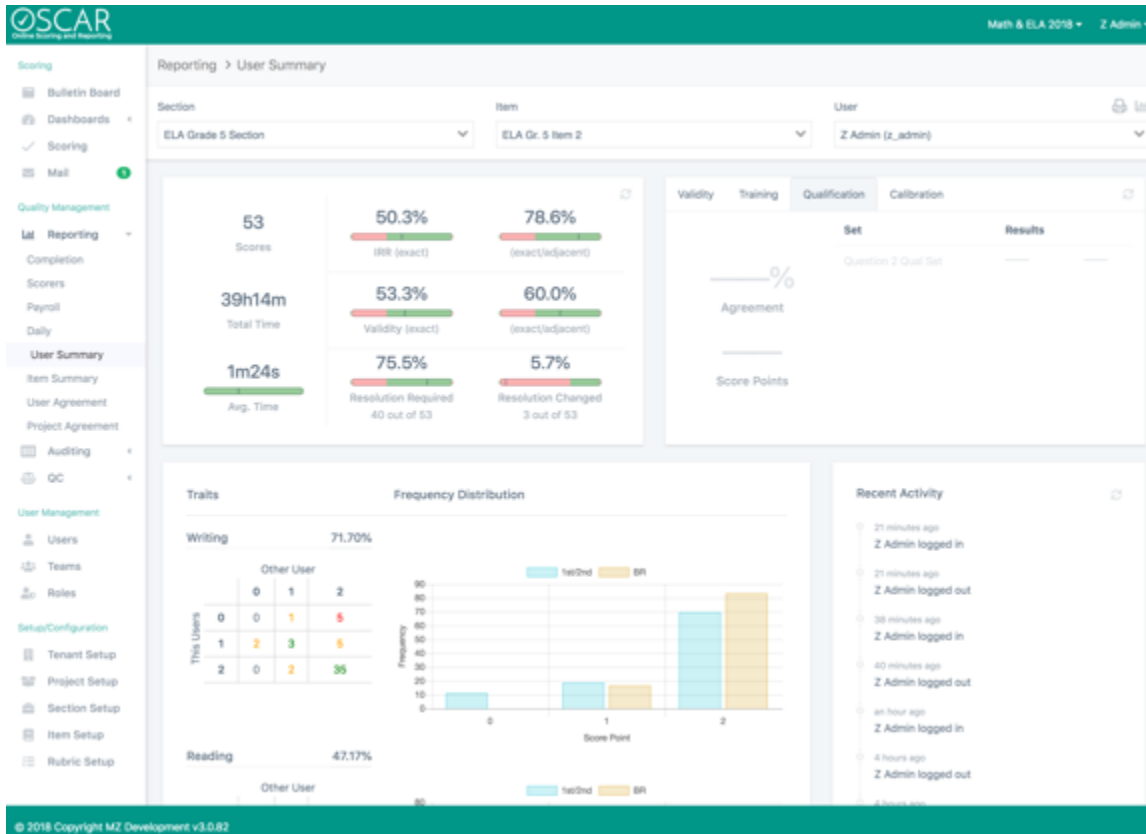
**Item Summary Report** – The Item Summary report provides a matrix view of IRR and the ability to see a specific scorer's trends by trait. This allows for targeted recalibration where needed. If a specific scorer is selected, the report will compare that scorer on the Y axis vs. the scorer of the second read on the X axis. The report shows the number of responses when the individual scorer agreed with the same exact score assigned by another scorer, and the number of responses when the scorer was adjacent high or low and non-adjacent high or low.

The matrix provides the scoring supervisor with the ability to identify traits that have high numbers of non-adjacent scores (if the item is setup to have first and second reads). The frequency distribution chart shows the percentage of scores at each rubric score point by trait. This granular reporting allows for more streamlined recalibration.
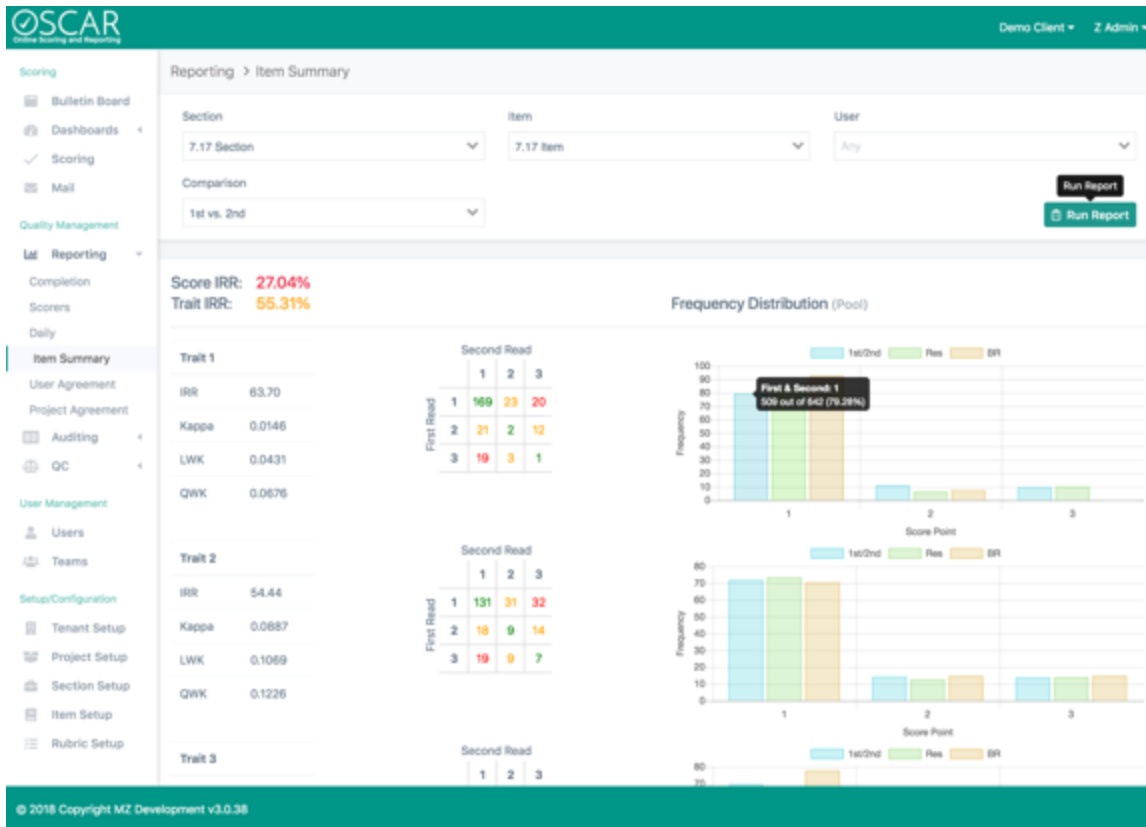
**User Agreement Report** – The User Agreement Report provides IRR and Validity exact or exact/adjacent percentage reporting for an item in aggregate and by trait for each scorer as well as Resolution required and resolution agreement percentages.



| Scorer | | IRR (exact) | | IRR (exact/adjacent) | | Validity (exact) | | Validity (exact/adjacent) | | Resolution Required | | Resolution Disagreed | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Totals (6) | | 45.4% | | 66.5% | | 42.4% | | 54.9% | | 67.7% | | 10.2% | |
| 0000 | − | 45.9% | | 59.3% | | 46.7% | | 53.3% | | | | | |
| + Development | | + 46.7% | | + 55.6% | | + 40.0% | | + 50.0% | | 64.4% | | 8.9% | |
| + Comprehension | | + 46.7% | | + 64.4% | | + 50.0% | | + 50.0% | | 29 out of 45 | | 4 out of 45 | |
| + Organization | | + 44.4% | | + 57.8% | | + 50.0% | | + 60.0% | | | | | |
| scorer_2 | + | 60.6% | | 82.8% | | 48.5% | | 57.6% | | 51.5% | | 0.0% | |
| 100015 | + | 45.6% | | 66.7% | | 46.3% | | 61.1% | | 77.9% | | 15.8% | |
| 100013 | + | 32.1% | | 59.6% | | 43.3% | | 56.7% | | 82.7% | | 25.0% | |
| z_admin | + | 48.3% | | 74.5% | | 33.3% | | 48.1% | | 58.2% | | 0.0% | |
| 100012 | + | 42.8% | | 53.6% | | 48.1% | | 59.3% | | 67.4% | | 15.2% | |
| Totals (6) | | 45.4% | | 66.5% | | 42.4% | | 54.9% | | 67.7% | | 10.2% | |

**Project Agreement Report** – The Project Agreement Report provides IRR and Validity exact or exact/adjacent percentage reporting for a project in aggregate and by trait for each scorer.



| Section | Item | | IRR (exact) | IRR (exact/adjacent) | | Validity (exact) | Validity (exact/adjacent) | |
|---|---|---|---|---|---|---|---|---|
| Totals (4) | | | 51.1% | 74.5% | | 43.1% | 57.7% | |
| ELA Grade 5 Section | ELA Gr. 5 Item 1 | + | 45.4% | 66.5% | | 42.4% | 54.9% | |
| ELA Grade 5 Section | ELA Gr. 5 Item 2 | − | 57.7% | 83.7% | | 48.0% | 65.3% | |
| | • Writing | | • 68.2% | • 88.5% | | • 36.0% | • 68.0% | |
| | • Reading | | • 54.1% | • 86.0% | | • 52.0% | • 64.0% | |
| | • Spelling | | • 51.0% | • 76.4% | | • 56.0% | • 64.0% | |
| Math Grade 5 Section | Math Gr. 5 Item 2 | + | | | | 56.3% | 71.9% | |
| Math Grade 5 Section | Math Gr. 5 Item 1 | + | | | | 27.7% | 46.2% | |
| Totals (4) | | | 51.1% | 74.5% | | 43.1% | 57.7% | |

# QUALITY CONTROL (QC) REPORTING

The QC Reports show how the scoring team is doing on Training, Qualification, Validity and Calibration sets. The QC Reports allows a view of the entire team or an individual scorer.

**Qualification/Training/Validity/Calibration Report** – These reports allow a scoring supervisor to see how the scoring team is performing on any QC scoring set.

In the example Qualification Report below, the green check mark shows the responses that each scorer scored correctly, and the red circle shows the responses that an individual scored incorrectly. The report allows the supervisor to assess the strengths and weaknesses of the scoring pool. It can help to identify what type of responses may be presenting difficulties to many scorers within the group. The scoring supervisor can then provide additional training to address the issue before live scoring begins.

By clicking the green button located at the end of each row, the supervisor is taken to the QC Report for that individual scorer.



The Validity Report (similar to the Qualification Report format) allows a view of overall performance on all validity responses (also known as validity papers, or anchors) that have been scored by a team or by an individual. Validity responses are responses that have been assigned scores by the scoring supervisors. (That score is considered the 'true score' and typically agreed upon by the customer.) At some point in the project, each validity response will be assigned scores by all or most of the scorers in the group, depending on the requirements of the project.

Typically, these validity responses will be served blind to the scorer, meaning the scorer will not know that he or she is scoring a validity response. After the scorer submits a score, the scorer will immediately see his or her own scores compared to the true scores assigned to that response by the scoring supervisors. By using this report, along with the Daily Report, a supervisor should be able to provide necessary targeted calibration to correct scoring issues that may arise.

The Individual QC Report is also available to individual scorers through the Dashboard tab; however, the scorer can only access his or her QC results. On the individual QC Report, scorers can review their performance on training, qualification, validity, and calibration sets. Every response from each set can be accessed by the scorer along with the score assigned by that scorer and the true score assigned by the scoring supervisor. This information is available to the scorer for review over the entire course of the project.

# OSCAR
Online Scoring and Reporting

Demo Client ▾    Z Admin ▾

**Scoring**
- Bulletin Board
- Dashboards ‹
- Scoring
- Mail

**Quality Management**
- Reporting ‹
- Auditing ‹
- QC ⌄
  - QC Setup
  - Reporting

**User Management**
- Users
- Teams

**Setup/Configuration**
- Tenant Setup
- Project Setup
- Section Setup
- Item Setup
- Rubric Setup

## QC Reporting

| Training | Qualification | Validity | Calibration |

| Trait | NAL | AL | E | AH | NAH |
|---|---|---|---|---|---|
| Trait 1 | 2 | 3 | 21 | 1 | 1 |
| Trait 2 | 2 | 1 | 22 | 2 | 1 |
| Trait 3 | 2 | 1 | 22 | 1 | 2 |

Section: 7.17 Section ⌄

Item: 7.17 Item ⌄

QC Set: Training Set 3 ⌄

| Scorer | 1 | 2 | 3 | 4 | 5 | 6 | 7 | Actions |
|---|---|---|---|---|---|---|---|---|
| 0000 | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ↺ |
| z_admin | ✓ | ⊘ | ⊘ | ✓ | ⊘ | ⊘ | ✓ | ↺ |
| basic_scorer | ✓ | ⊘ | ⊘ | ⊘ | ⊘ | ⊘ | ✓ | ↺ |
| demo_scorer | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ↺ |
| 100011 | - | - | - | - | - | - | - | ↺ |
| 39393 | - | - | - | - | - | - | - | ↺ |