## 1. Introduction

At DigitalDesk we believe that the university testing center collects some of the most valuable and least utilized data on campus. Our mission at DigitalDesk is to improve education by providing institutions and faculty with tools to unlock the insights hidden in their testing center data. We do this through a system which leverages cutting-edge information technology to provide cost savings, improve efficiency, and deliver a suite of exam tools to anywhere faculty can access the internet.

Our exam tools allow faculty to know within seconds of their forms being scanned the:

- Overall exam reliability
- High and Low scores
- Mean and Median scores
- Standard Deviation of scores

And allow faculty to delve deeper into individual test items with measures of

- Effectiveness of test questions to differentiate between students
- Question difficulty
- The effectiveness of question distractors

In addition to allowing faculty to access these reports remotely we also allow teachers to remotely:

- Update answer keys
- Rescore Exams
- Release grades to campus LMS and SIS
- Generate emails to students

## 2. Understanding Exam Analysis

The heart of the DigitalDesk vision is effective exam analysis, which allows teachers to create better exams. After we have installed our software and connected your current scanning machines to the network, your faculty will be able to access their exam analysis within seconds of the completion of a scan job, or instantly if you decide to utilize our online testing solution. We present this Exam Analysis in a concise and easy-to-understand report which lays out:

Exam Reliability: We use the data gathered from faculty exams to test the likelihood of exams to produce consistent results. Our reliability coefficient is based on a Chronbach's Alpha analysis. Chronbach's Alpha consists of three factors:

1. **Intercorrelation Between Items:** We measure the relationship between items on the test. The greater the interconnection between items, the stronger there relationships are. Exam reliability increases as the correlations between items increase.
2. **Test Length:** Longer tests are more reliable.
3. **Test Content:** Using a wider variety of content and techniques increases exam reliability.

The exam 'reliability coefficient' we generate can range in value from 0 (no reliability) to 1 (perfect reliability). In practice, exams DigitalDesk scores range in value from 0.5 to 0.9. On an exam with a high reliability coefficient, students who answer one question correctly will tend to answer other questions correctly. This helps faculty understand how well the set of questions on a test effectively measures student knowledge.

Our exam reliability coefficient provides an effective way for teachers to know how good their exams are at a glance:

| RELIABILITY COEFFICENT | INTERPRETATION |
|---|---|
| **.90 and above** | Excellent reliability (at the level of the best standardized tests) |
| **.80 to .90** | Very good for a classroom test. |
| **.70 to .80** | Good for a classroom test. Several items could be improved. |
| **.60 to .70** | Mediocre score. This test should be revisited and the class should be supplemented by other measure to determine student grades accurately. A number of items could be improved. |
| **.50 to .60** | Test must be revised. |
| **.50 and below** | Unreliable test. This test should not contribute to course grade. |

High/Low Scores: The high and low scores help frame test results within the range of student performance.

Mean/Medium Scores: We represent the average score in both points and percentage. Comparing the average score of all students with the medium—or middle—score helps faculty get a feel for how their classes respond to exams.

Standard Deviation: We measure how clustered student scores are around the average to provide faculty with a measure of how effectively faculty instruction is meeting students learning needs. A highly dispersed set of scores indicates that students are responding differently to classroom instructions and assignments. A tightly grouped cluster of scores around the mean indicates all student learning styles are being adequately met by class material.

After gaining a sense of how the overall exam results relate to materials and classroom environment our analysis allows faculty to delve deeper into the individual test items and identify which items need to be revised to improve overall exam reliability. Our item analysis assists faculty in identifying:

Item Difficulty: Our report helps faculty judge question difficulty by analyzing the P-value of a question. Ideal P-values fall a little above halfway between perfect score and chance. Through feedback from this component of our report faculty can reword and perfect questions to create finely honed questions.

Item Discrimination: We help faculty determine question reliability with a Point-Biserial correlation (PBS). A PBS analysis renders a 'discrimination index' which we color code, so faculty can fine-tune their tests much more efficiently by quickly seeing which questions are good, and which should be revised:

| | |
|---|---|
| Best | if the index is above .40 |
| Better | if it is between .30 and .39 |
| Marginal | if it is between .20 and .29 and |
| Review | if it is below .20 |

Distractor Analysis: We further help faculty fine tune their exams by providing them with an analysis of how many students picked each answer option and their combined average scores. By comparing the average scores of test respondents to specific distractors a professor can see in an instant how effectively their answer choices are at discriminating between students based on their level of knowledge. We distill this into the color coating mentioned above for quick test improvement.

## 3. Remote Capabilities

After faculty have used our reports to identify exam items they want to fine tune, our software allows them to remotely update answer keys, and rescore exams. Once they are satisfied test results can be pushed to campus SIS[1] and LMS[2] systems, and reports can be emailed to individual students with the push of a button. With our New WebAsses application faculty can use the information contained in our reports to update future digital versions of their exams remotely as well.

[1] Student Information System

[2] Learning Management System