BUROS
Center for Testing

Phase I Report: Document Preparation, Scanning and Editing

Matthew W. Grady, Ph.D. and Kurt F. Geisinger, Ph.D.

Buros Center for Testing

Consultants to the Florida Department of Education

May, 2011

Overview

During the week of May 1, 2011, representatives from the Buros Center for Testing observed the document preparation, scanning, and editing of test documents in accordance with Phase I of Buros' FDOE proposal. Dr. Matthew Grady observed the process at Pearson's Austin, Texas facility, and Dr. Kurt Geisinger observed the process in Pearson's Iowa City, Iowa facility. This report outlines the document preparation, scanning, and editing processes, as implemented by Pearson, and provides some recommendations for process improvements where appropriate.

Document Receipt and Check-in

All boxes of to-be-scored documents are received in secure, dry receiving areas. Boxes are checked for obvious damage, including water damage and tearing. If obvious damage to the boxes is observed, Pearson personnel take pictures of the damaged boxes and notify the Florida Department of Education (FDOE) to discuss the resolution of issues that might affect scanning. Inbound box labels are scanned into a receiving system, the boxes are manually counted, and the box counts are compared to the counts on the box labels. Counts obtained from the scanned inbound labels are then checked against each district's material return list and the inbound label receiving system, and discrepancies are reported to the Program Team, who then contact district assessment coordinators to resolve the discrepancies. It was our observation that the receiving areas were organized and clean, and the staff appeared to know exactly what they were doing.

Data Preparation

Boxes of to-be-scored documents are moved from the secure receiving area to a data preparation area on pallets. The boxes are opened and the to-be-scored documents are removed by Pearson personnel. Pearson personnel then check and sign off on each of the items on the list of contents for the box. To ensure that all to-be-scored documents have been removed from a box, Pearson personnel use a "buddy check" system in which two people examine the bottom of each box for any remaining contents. "Batches" of to-be-scored documents are then created and logged into a Work Flow Management system with a count for each batch. Batches of to-be-scored documents are then placed on carts and moved to a holding area where they await transfer to a "batching" area.

Batching. Batch carts of to-be-scored documents are moved to the batching area. Pearson personnel ("batchers") remove the paper bands from the to-be-scored documents and remove any extraneous material from the to-be-scored documents, (e.g. Post-it notes, etc.). It was our observation that the adhesive used on the paper bands is quite aggressive and in some cases sticks directly to the stack of to-be-scored documents. In very limited cases this causes tearing of the to-be-scored document when the paper band is removed during batching. Tearing of the to-be-scored documents makes scanning of the documents more difficult and time consuming. It is our recommendation that FDOE consider using paper bands with less aggressive adhesive as this

would likely eliminate this problem. It is also possible that another type of fastener could be used.

After de-banding, batches of to-be-stored documents are counted and the counts are recorded on a sheet that stays with the batch cart throughout its scanning and editing cycle. The batches of to-be-scored documents are then stacked to a standardized height and a standardized red "X" is drawn across the bottom of the stacked documents. This red "X" is designed to aid in the reassembly of a stack of to-be-scored documents in the rare event of a "dropped stack." A dropped stack is said to have occurred when documents fall from the batch cart due to the cart tipping, etc. The batchers then sign off on the batches before they are placed on carts and moved to a quality control area where they are independently checked by Pearson personnel. Throughout most of this process, the batches remain on the carts. After they are slit and scanned, they are placed back on the original batch carts.

Quality Control (QC) Check. Within the quality control area (a separate staging area and staff) Pearson personnel ensure that each stack of to-be-scored documents in the batch has been counted correctly and that a standardized red "X" has been drawn across the bottom of each stack. The quality control personnel sign off on the batch, which is then moved to a logging station. At the logging station, Pearson personnel log an estimated batch and stack count in the Work Flow Management system and sign off on the batch, which is then moved to a holding area in preparation for scanning.

It was our observation that the data preparation areas were organized and clean. Batch carts traveled only a limited distance between receiving, batching, and QC areas, which decreased the likelihood that to-be-scored documents would be misplaced or dropped. Each Pearson employee who handled the to-be-scored documents at any stage of the data preparation process signed off on a Control Sheet that traveled with the batch cart throughout its remaining cycle. This procedure allows floor managers to track and target document-processing problems to a specific stage at any point during the data preparation process. It was also our observation that Pearson data preparation personnel had adequate access to senior employees and supervisors. In some cases, space appeared limited and workspaces crowded, but not in a manner that impeded effective work.

Humidity Control, Document Acclimation, and Moisture Checking

The goal of humidity control, document acclimation, and moisture checking is to ensure that the moisture level in to-be-scored documents is not high enough to interfere with document scanning. The process differs slightly between the Austin and Iowa City locations.

In both the Austin and Iowa City facilities, the humidity level in the data preparation area is constantly monitored with a running average humidity level displayed on a clearly-visible digital display overhead. When batches are created in the data preparation area, each batch is assigned an acclimation time that indicates the earliest time that the batch can be scanned. The

purpose of assigning the acclimation time is to ensure that to-be-scored documents have spent at least eight hours in a humidity-controlled environment before they are moved to the document scanning area. After acclimating in the humidity-controlled environment for at least eight hours, a moisture probe is used to test the moisture level of a stack of documents from each batch. If the moisture level observed for a given batch is below a criterion level established by Pearson, the batch can be moved to the scanning area, which is also humidity-controlled. If the moisture level exceeds the criterion level, the batch of to-be-scored documents remains in a humidity-controlled holding area until the moisture level decreases below the criterion level. Most of the carts have four shelves with three stacks placed side-by-side on each shelf. Typically, the middle stack on each shelf is probed to determine the moisture level. Once the moisture level for the batch has decreased below the criterion level, as indicated by the moisture probe, the to-be-scored documents are then moved to the scanning area. In the Iowa City facility, if the moisture level observed for a given batch is below a criterion level established by Pearson, the batch can be moved to the scanning area which is humidity-controlled. If the moisture level exceeds the criterion level after the designated time, the batch of to-be-scored documents may be moved to a special holding area termed the "Sahara Room" until the moisture level decreases below the criterion level. This room is primarily utilized for batches that probe abnormally high upon delivery from Data Prep. This room has an even lower humidity threshold compared to the larger acclimation room. In addition, fans are mounted on the walls in order to provide even better air circulation in that room.

Once the moisture level for the batch has decreased below the criterion level, as indicated by the moisture probe, the to-be-scored documents are moved to the humidity-controlled scanning area.

It was our observation that considerable care was exercised in ensuring that the moisture level of to-be-scored documents was sufficiently low so as to not interfere with document scanning. As there is no "Sahara Room" at the Austin facility, if the moisture level exceeds the criterion level after the designated time, the batch of to-be-scored documents stays in the acclimation room until the moisture level decreases below the criterion level. It is our understanding that Pearson developed the moisture criterion and the humidity-level criterion in collaboration with paper science specialists at the University of Georgia.

Document Slitting and Scanning

Slitting. Upon arrival in the scanning area, Pearson personnel check the contents of batch carts. After the initial check, batches are moved to the "slitting" area. During slitting, a batch of to-be-scored documents is fed through an industrial-designed slitting machine that removes the bound edge of each document. The stacks of to-be-scored documents (formerly bound into test booklets) are now in single-sheet format, (i.e. "loose" sheet format). After slitting, the stacks of loose-sheet booklets in the batch are placed back on the batch cart and reviewed by Pearson personnel, who check their contents and sign off again on the Control

Sheet. Once the batch cart of slit sheets has been signed off on in the slitting area, the documents are moved to a nearby holding area where they await scanning, using the red "X" from Data Prep as another stack integrity check.

Before a stack of to-be-scored documents is guided through the slitting machine, a series of test sheets are fed through the slitter to ensure that the machine is correctly calibrated and will slit the correct side of the documents. Failure on a test sheet triggers an automatic shutdown of the slitter, and the problem is resolved according to established guidelines. Although unusual, it is sometimes the case that the slitting machine blade will cut too close and slit across the "timing tracks" of a document. These mis-slit documents must be repaired during the scanning process or "flatbed" scanned by hand during the hand-editing process.

Scanning. Once batch carts of slit, to-be-scored documents are signed off on by the document staging, data prep, QC, humidity-control, and slitting personnel, they are ready to be scanned. Before a batch of to-be-scored documents is fed into a scanner, a series of test sheets are fed through the scanner to ensure that the scanner is properly aligned and reading at the desired level of accuracy (scan value intensities). Failure on any of the test sheets will trigger an automatic shutdown of the machine that must be resolved before scanning of any live documents/sheets can continue.

The scan application is designed to only allow the doc id codes assigned to that project. If any other material is intermixed with invalid doc codes for that application, those documents will be rejected and will not scan. These documents are pulled and given to the data prep team to be batched with their appropriate scan application.

There are a number of additional error checks built into the scanning process. During scanning of actual to-be-scored documents, the scanner automatically shuts down:

- if an answer document does not match the document key, which is "hidden" in the "skunk" marks on each page of a document. This mark identifies the document number and page number of each sheet scanned.
- if the documents are fed in an order that is inconsistent with the document key (e.g., items out of order, hence pages out of order). If the document pages are out of order, the scanner operator will put them back in order.
- if a page of a document is skipped or if multiple pages of a document are fed through simultaneously (i.e., a "multiscan"/stuck together).
- if any extraneous material is detected on a page of a document, (e.g., Post-it note, etc.).

The scanner operators made frequent use of compressed air to clear the machines of dust and debris and make use of clear or colored tape embedded with tracking marks on it to restore the tracking marks if they had been damaged. While such damage is rare, it did occur

occasionally when test takers removed their security label from the test documents or doodled inside the skunk marks.

During scanning, document images are captured and stored electronically and a unique identifying number is laser printed onto each sheet of the physical documents. If a document cannot be successfully scanned for any reason, (e.g., the document is ripped, or the document was damaged during slitting, or the document cannot be read accurately due to damage or ink, etc.) a scanning alert is logged in the Workflow Management System. Documents with minor damage can often be repaired by the scanner operator using the special tape mentioned above that allows the documents to be scanned successfully. Documents that cannot be successfully scanned for other reasons must be flatbed scanned, and unscannable data must be manually entered during the hand-editing stage of document processing.

After scanning, the stacks of documents are placed on their original batch cart and again signed off on by the scan operators on the Control Sheet. The batch cart is then moved to a nearby holding area where it awaits transfer to the hand-editing area.

It was our observation that the period between the slitting and scanning of answer documents is the period when a dropped stack would likely have the most negative consequences. During this period, the to-be-scored documents are transformed into single-sheet format, (i.e., loose sheet format). Although Pearson maintains a detailed "dropped stack procedure," the reassembly of a dropped stack during this stage of the document scoring process would be difficult if it were to occur, as the to-be-scored documents are no longer bound into test booklets. Currently, FDOE addresses this problem by utilizing a serialization system for its Reading and Math retake test and answer books, as well as its Grade 3 Math test and answer book. By this system, a unique serial number is printed on each page of a test document during document printing. These serial numbers effectively ensure that dropped stacks of serialized documents can be accurately reassembled. A dropped stack of loose document pages from grades without serial numbers, however, would be more difficult to reassemble. We were happy to learn that FDOE included the requirement of lithocodes on all student answer documents in its future production specifications. Like serialization systems, lithocoding systems print a unique identifying number on each page of each answer document during document printing. We believe that the utilization of the lithocoding system will reduce the severity of problems resulting from dropped stacks, should any occur.

Scanner Maintenance and Calibration. Scanners undergo routine maintenance with every 2.5 to 3 million sheets scanned. Each scanner has a programmatic maintenance countdown clock that indicates to the scanner operator the number of documents that can be scanned before the scanner must undergo routine maintenance. If the scanner reaches the "countdown" number, it shuts down automatically until routine maintenance has been performed. Attached to each scanner is a maintenance log that shows the maintenance history and includes a brief description of the type of maintenance performed. Pearson maintenance personnel are on-site during daytime

hours and on-call during nighttime shifts. Scanner calibration frequency reports are also maintained by scanning managers. Scanner calibration frequency reports contain information on scanner calibrations that include checks on the mark intensity thresholds and double mark/multi-mark identification threshold settings. It was our observation that Pearson scanner maintenance was performed with adequate regularity and that scanner calibration frequency reports were adequately maintained.

An additional "alert log" of slitting and scanning anomalies is maintained by the Pearson program team. The alert log records and provides a description of abnormal occurrences, (e.g., poorly slit document pages, document pages rendered unscannable by damage caused by adhesive on document paper bands, a dropped stack). The log is maintained as a Microsoft Excel spreadsheet by a scanning manager and can be summarized by scanning machine, scanning operator, or scanning section to identify trends. It was our observation that these logs were adequately maintained.

Hand Editing

Batch carts of scanned physical documents are moved to the hand-editing area where they are scanned into the Workflow Management System to ensure there is a record of their physical arrival in the hand-editing area. The scanning program sends electronic scan data from the scanning area to an image-editing program used in the hand-editing area. Documents that trigger a scan alert during scanning are identified by the image-editing program and Pearson editors apply the necessary edits according to guidelines agreed on with the Department. If Pearson editors are unable to resolve the alerts, the Pearson program team is contacted. It was our observation that Pearson personnel within the hand-editing area had adequate access to senior employees and supervisors.

Documents that could not be scanned via the machines are considered "unscannable" and/or "unscorable", and they must be flatbed scanned by hand and scannable marks present in the document must be entered manually by hand-editors. If documents are too damaged to be flatbed scanned the program team is notified. Manual entry of item data not scanned is "double key-entered" meaning that two independent Pearson editors electronically enter student marks so that the entries can be checked for agreement. If there is disagreement between hand entries, the disagreement is resolved by a senior Pearson editor.

As an extra QC measure, Pearson personnel in the hand-editing area also check a sample of all non-alerted scanned documents. All student marks on every 250th physical document are checked against the scanned file for agreement. This serves as an additional check of the scanner accuracy. Lack of agreement between the physical documents and electronic scan files could indicate a potential problem in scanning which would be addressed immediately by a senior Pearson editor.

For some answer documents for some grade levels, examinees "write in" their answers to questions in boxes located above the answer bubbles. These questions in the Mathematics tests are known as "gridded response" questions. Examinees then "bubble in" the corresponding bubbles according to written instructions. It was our observation that there were occasional discrepancies between the "written in" and "bubbled in" answers on answer documents. For example, the examinee may write in the correct answer but bubble in the answer document incorrectly. In such an instance, the response would be considered incorrect. The fact that these types of discrepancies occur does not indicate that the answer documents are being inaccurately scanned, scored, or hand-edited by Pearson. However, we include this information in our report so that FDOE is aware that these types of discrepancies can occur. Our recommendation is that FDOE consider this problem and how it might be ameliorated – possibly by ensuring that students have adequate practice bubbling in answer documents when necessary. It might also be possible that software could be developed to read the written-in responses directly. However, at present, we are unaware of any software that is capable of performing this function.

Transcription

All documents requiring transcription for Spring 2011 were routed to Iowa City. Documents requiring transcription include a mix of circled, bubbled (gridded), and handwritten responses. Data from the approximately 1,500 large print documents are manually "double key-entered into Pearson's FALCON system meaning that two independent, trained senior transcribers enter student marks so that the entries can be checked for agreement systematically and resolved by a third individual. Transcription of data from a large print answer document to a regular answer document is completed by a senior transcriber and the transcription is checked by a second employee. The transcribed data is checked against the Falcon system data and an error report is generated for any mis-matches. All mis-matches are investigated and corrected as necessary. Those responses that involve Braille are sent to a specialized company (Florida Instructional Materials Center) that transcribes the answers into a scannable document.

Site Security

It was our observation that the Austin and Iowa City sites were highly secure sites. No unauthorized persons were allowed into the receiving, data preparation, scanning, or hand-editing areas of either facility. It was our impression that students' personally identifiable information was safeguarded in strict compliance with FERPA.

Overall Impressions

Our overall impression of the document preparation, scanning, and editing processes performed by Pearson was quite positive. Pearson took adequate care to ensure that documents were handled appropriately and accounted for. To-be-scored documents were checked for moisture. The Pearson facilities were clean and organized and detailed process documents were available to Pearson personnel. Scanning machines were adequately maintained, as were scanner

calibration frequency reports. Unscannable data were double key-entered and checked for agreement, as were data from large print documents. Scanned documents were always stored in secure, dry locations. Overall, our process improvement recommendations are minor and we feel that Pearson's document preparation, scanning, and editing processes are of high quality.