

TEST VALIDATION AT NORTHWESTERN SUGAR

Mark A. Johnson, Department of Management, College of Business, Idaho State University, Campus Box 8020, Pocatello, ID 83209, 208-282-2155

James P. Jolly, Department of Management, College of Business, Idaho State University, Campus Box 8020, Pocatello, ID 83209, 208-282-3455

ABSTRACT

This case describes a situation where two consultants were hired to establish a valid employee selection system for a sugar processing plant. The case illustrates how a consulting team had to first educate managers about the technical and legal aspects of test validation and how they proceeded to validate the system. At the end of the case, in a meeting with managers and the consultants, the plant manager announced that the firm would begin testing at the plant and at other company facilities in ten days. The consultants were at a loss because the validity study of the testing battery had not been finalized and therefore validity had not been established. This proceedings paper contains only forty-three percent of the text of the full case.

BACKGROUND

In a November 2001 meeting, Brandon Smith, Plant Manager at Northwestern Sugar, announced to the others in the room, "Great, we will begin testing applicants in 10 days. Also, we have been talking with corporate and they are very interested in initiating the testing battery at our other four plants as soon as we get started here. We have really impressed them by getting this program online." Brian Higgins, a university professor, and Dennis Hill, a doctoral student, were in shock. They weren't sure how to respond. "Where the hell did that come from?" Higgins thought to himself.

Higgins reflected back on the events that had occurred in January of that year which had led to the current situation. He and Hill had been working on the development of an employee selection system for Northwestern Sugar for ten months. The project had occurred almost by accident, spurred by a comment Smith had made earlier that year in January during a management-training program they were conducting for managers through the Management Institute of a major Western university.

"Testing is illegal," Smith had exclaimed. "There was this Griggs case that disallows testing because of discrimination against minorities," he went on to say.

Higgins and Hill glanced at one another in a way that basically communicated to the other, "Oh boy, here we go again," and fought back smiles, not wanting to embarrass Smith. But the two had heard this sort of comment many times before. Unfortunately, the manager's assertions were consistent with the perceptions and understandings of many practicing managers, but they were not accurate or reflective of current employment law.

Some information in this case has been changed to preserve the anonymity of the organization and its members.

Higgins and Hill discussed the current legal status of employment testing with Smith and the other trainees. The two trainers also met with Smith after that day's management program to discuss testing and validation processes in more detail.

Northwestern Sugar was a privately held corporation that processed sugar beets. The firm employed more than 2,000 employees at its five packaging plants. The plants were located in three Northwestern states.

THE FEBRUARY MEETING: TO TEST, OR NOT TO TEST?

In February of 2001, just one month after the training program, Higgins and Hill were sitting with Smith in a conference room at Northwestern Sugar. Also in attendance were the Packaging Department Manager and Assistant Manager. The purpose of the meeting was to plan a strategy for implementing a selection-testing program for hiring the plant's operating employees. After further discussions with Higgins and Hill, Smith had done an about-face since his initial assertion that testing was illegal. He now realized that employment testing could be conducted without violating employment laws.

Smith went on to discuss a major reason for Northwestern's interest in testing. He described the company's plan to implement an Enterprise Resource Planning system. The effects of this information system would pervade every aspect of the company's operations, going all the way down to the plant floor. "Traditionally, the work requirements at our plant demanded more physical effort than brains," Smith commented, "but now, well, a lot has changed and a lot more is going to change with the implementation of the ERP system. Our plant workers will have to use the computer systems to record operations, to receive information, and to monitor the processes. This is unlike anything they have had to do in the past." Smith continued:

We need workers who have the aptitudes and skills to effectively run these systems. Moreover, we have had a problem finding capable workers to promote into our higher skills jobs and, eventually, into supervisory and managerial positions. Our current selection processes do not provide us with enough capable plant floor workers and this in turn limits our selection pool for promotions. For years we were bound to adhere to our promotion-from-within system. We had a very formalized bidding system, and seniority was the most important criterion. Over the last few years we have increased the importance of hiring the most qualified internal candidate, but seniority is still given a lot of weight. I sense that we will continue the trend toward increasing the skill requirements for bidding and promotions and rely less and less on seniority.

Smith paused for a moment as if to catch his breath and to give those listening a chance to digest what he had said. He continued:

The bottom line is that we need to not only hire more qualified people into our entry-level positions but also simultaneously improve the level of all our workers in the higher wage classes so we have good people to promote. What we need is a testing program that will improve the average aptitudes and performance of workers, that will improve the quality of our internal ranks for promotion, and that will enable us to objectively screen for higher classification and supervisory positions.

As the meeting between Higgins, Hill, and the three Northwestern Sugar managers continued, Higgins asked the managers to describe the current hiring processes.

Higgins then asked the managers to describe the job paths for new hires. Collectively they described how new hires begin in the classification “Casual Employees” who work part-time, and often on an as needed basis. These workers typically perform some sanitation/laborer functions, cleaning equipment and the plant. Others work on the processing and packaging lines, performing mundane types of work, such as making boxes, stuffing product packages into cases, unjamming raw product from equipment, and the like.

Higgins and Hill then took the opportunity to describe the concept of validation, different validation strategies, and the legal implications of testing, to the three managers. They informed the three that valid selection processes use only job related measures. The managers were told that there are five major types of validation — criterion-related validity (which includes predictive and concurrent), content validity, construct validity, transportability of validity, and validity generalization.

Higgins went on to discuss that most cognitive ability tests have adverse impact. Adverse impact occurs in testing when a selection device selects proportionally more majority group members (i.e., Caucasian) than members of protected groups (i.e., African Americans or Hispanics). Adverse impact does not prove that a firm has illegally discriminated but merely establishes a prima facie case of discrimination. Generally, employers can counter a prima facie case by demonstrating the validity of a selection instrument. However, the validation study must be professionally developed following the guidelines presented in the EEOC’s Uniform Guidelines and the APA’s Principles. Collectively, these references provide some guidelines for using the validation approaches previously discussed. Though it is generally considered desirable from a social and legal perspective to avoid adverse impact, valid selection methods may be used even when adverse impact occurs.

A VALIDATION STUDY PROPOSAL

By the end of the February 2001 meeting, Smith and his two managers expressed their interest in the prospect of using tests for entry level selection and for bidding into higher classification jobs and promotions into the supervisory ranks. Higgins and Hill cautioned them that there were no guarantees that a validation study would support the job-relatedness of a test or test battery. The purpose of a validation study is just that – to “study” prospective tests to find whether the tests are valid in the employer situation for the job or job classes studied. Only then may tests that result in adverse impact be used as evidence of job relatedness. Despite these caveats, Smith commissioned Higgins and Hill to develop a validation study proposal. By mid-March, the two researchers had completed a project proposal that was subsequently accepted in the first week of April by Northwestern Sugar.

JOB ANALYSIS FINDINGS AND TESTS CONSIDERED

The researchers completed the job analysis procedures and written job descriptions for all of the packaging department jobs in September. Qualitative data from the interviews, observation, and the job analysis questionnaire provided a great deal of useful information detailing the job duties and tasks. In addition, the information obtained from the task inventory rating forms provided quantitative data regarding the similarities and differences across the jobs and the relative importance and frequency of tasks performed. Most of the tasks included on the task inventory were rated as having been performed frequently and/or were considered important to the job by the worker raters.

Collectively, the job analysis information provided the researchers with job-related data to assist in their selection of tests to include in the testing battery and which subset of these to include in the concurrent validation study. Based on this information, Higgins and Hill decided to test current workers using a custom-developed math test, a reading test, a mechanical ability test, and a manual dexterity test.

Testing of packaging department employees (for validation study purposes) began in September of 2001 and was completed in the first week of October. The results on the four tests would be analyzed statistically along with performance measurement data (which would serve as the criterion).

COMPLETING THE VALIDATION STUDY

In early November, Higgins and Hill had collected the test and performance data necessary to complete the statistical portion of the project. The data had been entered into a statistical software program and some preliminary statistical analysis had been performed. However, a variety of analyses were needed before the output could be meaningfully interpreted and a lot of work was necessary to document validity if in fact the results supported a conclusion that some or all of the tests were valid. If the results demonstrated sufficient validity for any of the tests they would have to document the results in a technical validation report as required by the EEOC Guidelines. They would also have to examine how the tests worked together. Even if two tests are found to be valid, one may not contribute meaningfully to the overall validity when used in conjunction with another test. This is known as incremental validity. In addition, the consultants would need to determine cut-off scores for the different predictors. Additionally, Higgins and Hill would have to develop a Test User's Manual that addresses issues such as test administration, scoring, security, and retesting policies. Despite all the work that remained, Higgins and Hill were excited that they had completed the data collection phase and some statistical analysis. They knew they were well on their way to finalizing the project.

THE NOVEMBER MEETING: LET'S START TESTING?

Smith had called a meeting for later that week to discuss the "progress" of the validation project. Also in attendance were the two packaging department managers, Steven Daley from Human Resources, Higgins, and Hill. The two researchers proudly communicated that all of the criterion-related testing and performance ratings data had been collected and some initial statistical analysis had been completed. "At this point, all we need to do is finish the statistical analysis, interpret the results, and complete our validation report." It was at this point that Higgins and Hill were thrown the curve when Smith announced his eagerness to begin testing and to expand the testing program to other plants: "Great, we will begin testing applicants in 10 days. Also, we have been talking with corporate and they are very interested in initiating the testing battery at our other four plants as soon as we get started here. We have really impressed them by getting this program online." Higgins was in shock. "How did they jump to this conclusion," he thought. The validation study was performed for jobs at the specific plant. At no time was there any mention of testing at any of the other plants. But Higgins wondered if there was a way to legally use the tests at the other plant facilities if he and Hill could establish the validity at the current location? At that point, Higgins finally managed to get a few words out, "Where should I begin?" Then he thought it best not to respond. Despite his great desire to address Smith's plan during the meeting, Higgins decided to bite his tongue for the time being. "What now, he thought to himself?"