Advanced Ergonomics
Physical Ability Testing
Program Review

by
Charles K. Anderson, Ph.D., CPE
President
Advanced Ergonomics, Inc.

ADVANCED ERGONOMICS, INC.
5550 LBJ Freeway
LB40, Suite 350
Dallas, Texas 75240
Phone: (800) 682-0169
Fax: (972) 239-3757

Website: www.advancedergonomics.com
E-Mail: aei@advancedergonomics.com

Table Of Contents

Section 1: Introduction

In 1989, Advanced Ergonomics (AEI) began widespread implementation of physical ability testing programs. These programs were built upon a substantial base of statistical validation research that AEI had amassed, and has further expanded over time. Today, there are more than 1400 client locations using the AEI physical ability testing service. The program is primarily used for assessing the ability of applicants to meet the physical demands of the jobs for which they are applying, but it has also been used for assessing individuals returning from extended leaves or injuries, transferring to jobs with higher demands, and other applications.

The intent of this report is to review the design of the program and the results that have been achieved. Included are sections regarding:

- testing implementation process;
- ongoing program management;
- program effectiveness; and
- legal reviews.

One consequence of a mismatch between worker ability and job demand is increased risk of injury. Hence, it would be anticipated that an effective testing program would also lead to a reduction in injuries. In fact, companies have found a 57% reduction in back injury experience, and 41% reduction in overall injury experience, on average, with implementation of the testing program. Data from the various validation studies suggest that new-hires who fail the battery are more than seven times as likely to have a back injury as those who pass.

The testing program has been designed to meet the requirements of the Americans With Disability Act (ADA), Title VII of the Civil Rights Act, the Age Discrimination in Employment Act (ADEA), and the various state counterparts to these pieces of legislation. The program has also been validated in accordance with the Uniform Guidelines on Employee Selection Procedures (29 CFR Part 1607), Section 15.

As will be discussed in more depth later, time has afforded several opportunities for program reviews by the Equal Employment Opportunity Commission (EEOC) and their state equivalents, the Office of Federal Contract Compliance (OFCCP) and National Labor Relations Board (NLRB) arbitrators. In every instance, the testing program has "passed muster".

Section 2: Implementation Process

The first two steps in the physical ability testing implementation process are ergonomic job analysis and test battery design. The third step is preparation of the documentation required by EEOC, which validates that the battery is an objective indicator of ability to perform the job.

Job Analysis

In order for a testing program to be effective, let alone withstand a legal challenge, it is absolutely critical that it is based on a thorough job analysis. Advanced Ergonomics has spent more time studying the physical demands of warehouse workers than any other company in the world. Our method has been used in more than 450 distribution centers to collect information about the strength, endurance and postural demands encountered when handling cases.[1] As part of these analyses, Advanced Ergonomics has measured the energy expenditures of more than 3,000 workers in a wide range of industries.

Analysis Of Strength Requirements

The strength requirements of a job need to be characterized in terms of the load that is handled, the starting location, and the location to which it is moved. With most jobs, the postural requirements are intertwined with strength requirements since the stressfulness of handling a load is, in part, a function of the posture used when lifting. As most of us have experienced, it is much more difficult to lift a heavy object from floor level, far out in front of us or overhead, as compared to handling that object close to the body at waist level. The frequency of handling is very important as well, in that it is critical to focus on heavy weights that are routinely encountered, rather than the isolated instance, say, where a 100 lb. bag is handled once per month.

As an example, food warehouse workers typically have to be able to occasionally lift very heavy loads (50 to 60 lb. at least, and in some situations as much as 100 lb.) over the course of a shift in which they frequently lift cases with an average weight of 20 to 30 lb. These cases are being

[1]The extensive experience in food distribution warehouses was one of the reasons that the Food Distributors International (FDI) trade association asked Advanced Ergonomics to write the Voluntary Ergonomics Program Management Guidelines For Food Distribution Centers. Also because of this experience, Dr. Anderson was invited to speak in London to the members of CIES, which is the European equivalent of FDI.

lifted from heights ranging from floor level to the full extent of overhead reach, and from locations ranging from in front of the body to the back of slots. Heavier cases tend to be slotted on the bottom. The wide range of case weights and handling locations makes it difficult to provide a simple profile or description of the lifting requirements for warehouse workers. The difficulty of providing a simple characterization of the handling requirement is further compounded by the fact that after lifting the case from the slot, the warehouse worker then carries the case and places it on the outbound pallet at a location that could range from near floor level to the full extent of overhead reach. The placement of any given case on the outbound pallet will vary depending on the cases selected before and the cases yet to be selected for the order. In other words, there is little that is routine about the weight of the next case, where it will be lifted from, or where it will be set. Complex handling requirements are a common situation in virtually all industries. Rarely are the loads handled of the same weight, handled from and to the same locations, or handled at a constant pace throughout the day.

A major ramification of this complexity is that it is very difficult to adequately translate that requirement into the design of a physical ability test that will accurately assess an individual's capacity to perform the demand in its full complexity. The difficulties inherent in developing a test of repetitive lifting capability are discussed in further detail later in this section under the subheading "Alternative Tests".

It is possible, though, to define the extremes of the strength requirements. These would be the most-demanding handling situations that routinely occur. If the worker is unable to perform these most demanding tasks, then he or she would be unable to perform the job, even if he or she has the strength to handle the majority of the cases. For example, 13% or less of the cases handled in food distribution warehouses are more than 50 lb., so it is not typical that a warehouse worker would be handling multiple heavy cases in close sequence. Hence, it cannot be said that a warehouse worker has to *repetitively* handle the heaviest of the cases. On the other hand, warehouse workers do repetitively handle cases, though the case weight ranges widely.

Analysis Of Endurance Requirements

The strength and endurance demands are interrelated in that the maximum weight that individuals can handle at a given point in time is a function of their peak strength when they are

rested or "fresh", and the degree to which they have become fatigued by prior physical exertion. Probably all of us have experienced that the amount we can lift at the end of a strenuous day is much less than what we can lift when we are well-rested. The degree of fatigue we will experience when performing prolonged strenuous activity, such as selecting cases for an entire shift, is a function of our cardiovascular fitness level, which is sometimes referred to as our aerobic capacity.

One way to capture the complex repetitive handling requirement is to quantify the amount of energy expended to perform the work. The energy expenditure requirement can then be directly compared to an individual's aerobic capacity in order to objectively determine if they have the physical capability to perform the complex repetitive lifting, carrying, walking and stepping involved in a full shift (or more) of activity. To reemphasize, this approach has the added advantage of taking into account all of the physical activity involved in a shift's work -- not just the repetitive handling.

Conclusions

Jobs like "warehouse worker" can be very demanding from both a strength and endurance standpoint. Furthermore, these jobs often require that the worker be able to adopt a wide range of postures. Because of the issues related to developing meaningful physical ability assessment tests, Advanced Ergonomics has chosen to characterize the demands in terms of the maximum handling requirements that are encountered routinely, and the overall energy expenditure demand of the job.

Typically, comparison of the strength and endurance requirements to the capabilities of the workforce as a whole reveals that most males can meet the strength requirements of the job. In other words, they are able to demonstrate the ability to lift the heaviest cases from the most demanding positions at least one time. *Because of the high level of energy expenditure associated with jobs such as "warehouse worker", it is less likely that the individual will have the aerobic capacity to work an entire shift at the required energy expenditure.* Only about 70% of the male applicants to distribution centers have the aerobic capacity to do the job, whereas more than 90% have the strength to lift the heaviest cases. The percents of the female population are even smaller than those for the male population, but the same basic trend holds. Aerobic capacity is as much or more of a limiting factor than strength in many jobs. *Therefore, the most critical element in evaluating a person's ability to do jobs like "warehouse worker" is to accurately assess their aerobic capacity*

relative to the energy expenditure requirement of the job. The next section discusses how tests were developed to assess an individual's ability to meet those demands.

Test Battery Design

The exact design of the test battery for a given job will be a function of the requirements of that job. Since requirements vary tremendously, so will test battery designs. The commonalties that are seen relate to the basic categories of physical demand that may need to be assessed. The general categories are strength, cardiovascular endurance, flexibility, agility and posture tolerance.

Strength Assessment

As was discussed in the last section, the approach of Advanced Ergonomics to profiling the manual handling requirement of a job is to capture the peak demands, the postures that are involved, and the overall demand. For instance, the peak demand might be characterized by the heaviest case weight that would be routinely encountered. The overall demand is best captured in terms of the average energy expenditure and the typical shift length.

Consideration of the heaviest case routinely handled has to account also for the postural element. This can be illustrated with an example from grocery warehousing. Generally, cases can be pulled from the back of a slot to the front with a "rake" or "hook", so it would not be representative to ask participants to lift far from the body, even though some cases may be stored at the full extent of reach. Our measurements of handling locations for the heaviest cases routinely encountered indicated that the heaviest cases are typically stored in the lower slot, and are generally placed below mid-chest height on the outbound pallet. The highest case on a full pallet of heavy product was also typically close to mid-chest height. Cases that were higher than mid-chest level generally could be slid down the side of the stack, which would mean that the warehouse worker would not have to handle the full weight of the case until it was taken away from contact with the stack, which typically happened between mid-chest and waist height. Therefore, the strength tests for a grocery warehouse worker were designed to evaluate the individual's ability to lift from floor level and in the mid-chest region.

The tests are performed in the following manner. The participant is asked to place weight in a box and demonstrate the lifts in the two regions described above. The individual starts with a low

amount of weight and adds to the box in preset increments until reaching the maximum weight they can safely handle. The cutoff is based on the ergonomic job analysis for that distribution center.

The same basic approach would be taken for developing strength tests for any other job. In brief, the essential handling tasks are identified and then simulated. A key factor in choosing a particular manner of simulation is to select a method that bears the greatest resemblance to the way the most demanding routine handling requirement is performed. Additional considerations include safety of test administration, reliability and objectivity of results.

Aerobic Capacity Assessment

An individual's aerobic capacity can be assessed in a number of ways. The most accurate assessment is to perform a maximal stress test, which involves having the individual walk on a treadmill or pedal a bicycle at increasing paces until their heart rate reaches its maximum. The energy expenditure achieved at maximum heart rate is their aerobic capacity. This procedure is quite uncomfortable for the participant, is costly to perform and carries a greater risk of untoward health consequence than would be desired in a routine test for employment purposes. The alternative is a sub-maximal stress test, which requires the participant to work at a level that represents 65% to 85% of their maximum heart rate. The results of the submaximal test can be used to calculate aerobic capacity. As with the maximal stress test, any of a number of modes can be used, such as treadmills, bicycles, or a step bench. Advanced Ergonomics elected to use a test that involves a step bench so as to keep the cost to a minimum.

Great care was taken by Advanced Ergonomics in selecting the specific step test protocol. Research on the protocol used indicates that it has a correlation coefficient of .92 with aerobic capacity as determined with a maximal test; perfect correlation would be 1.0. Therefore, the sub-maximal step test used by Advanced Ergonomics is very predictive of the results that would be obtained from a much more expensive maximal treadmill test.

The step test can be performed in less than fifteen minutes, yet it is able to be used to accurately assess an individual's aerobic capacity, which indicates the ability to maintain the productivity standard for an entire shift. The participant wears a heart rate monitor during the test. The heart rate is recorded at various points during the test, and entered into a computer program along with other data about the individual. This program is used to predict aerobic capacity based

on equations generated from extensive research on this particular protocol. The pass/fail cutoff score is based on the metabolic expenditure requirements found on the job at the particular facility.

Flexibility, Agility and Postural Tolerance Tests

By virtue of the fact that the strength test is designed to reflect the manner in which strength must be exerted on the job, it will also assess whether the individual has the level of flexibility needed. For instance, a strength test involving a lift from floor level also evaluates whether the individual has the ability to reach to floor level. Independent assessments of flexibility have been set up where there are no significant strength requirements.

The same can be said of agility and postural tolerance tests. It is possible to set up very specific tests for jobs that require significant agility, such as climbing ladders, moving through spaces with a low overhead, or climbing through openings of various size. The test is usually a mockup of the exact task required on the job. A posture tolerance test would involve having the individual work in the required posture for a period representative of what would be required on the job.

One factor influencing whether such tests would be included in the battery is the anticipated pass rate. If it is expected that virtually all individuals would pass a given test, then that test may be dropped from the battery. We have found that many of the flexibility, agility and posture tolerance tests that might be considered actually have very high pass rates for the typical new-hire populations, so few of the batteries include these tests.

Alternative Tests

There are a variety of strength testing protocols available other than the one used by AEI. AEI is constantly evaluating alternatives to determine if any new advances would improve the effectiveness of identifying individuals who can safely perform the job and/or reduce adverse impact for females, individuals over the age of forty, or individuals considered to have a disability as defined by the Americans With Disabilities Act.

Isokinetic Strength Tests

For example, some vendors use isokinetic strength testing equipment. This equipment is designed such that the participant exerts force against a bar or handles that are constrained to move at a constant velocity. This design creates a problem for assessing the applicant's ability to lift, since lifting is not typically performed at a constant velocity. Generally a person performing a lift starts the movement slowly while overcoming inertia, builds speed, and then slows the movement near the termination of the lift. A more accurate assessment of capability is obtained when the natural pace of lifting is allowed. Furthermore, some isokinetic tests involve strapping the participant into a fixture that only allows movement in the low back, knee or shoulder. The strength of the specific muscles are then tested in isolation. The problem with this approach is that lifting is performed with the whole body, so the assessment does not directly evaluate the ability to lift. When performing an actual lift, an individual can compensate for weakness in low-back muscles by utilizing leg muscles, thereby allowing them to perform a lift that they could not accomplish with their low-back muscles alone.

Another set of issues to be considered in evaluating alternatives relates to whether the results of the test can be directly linked to the requirements of the job. Isokinetic equipment typically measures the force exerted in units of torque or moment, which is difficult to compare to the actual weight that has to be handled on the job. Some vendors who advocate the isokinetic approach try to get around this problem by using the strength rating table from the Dictionary of Occupational Titles (DOT). In addition, they use this same DOT strength rating table to attempt to ascertain the weight an applicant could handle repeatedly over a shift from the torque generated in one or a few trials. Unfortunately, the strength rating table is not well-suited for either of these uses, and may lead to denying employment to applicants who can actually perform the job. Clearly, this would be an ineffective approach that could also create legal liability for unfair discrimination in hiring practices.

Progressive Isoinertial Lifting Evaluation (PILE) Test

As another example, tests such as the progressive isoinertial lifting evaluation (PILE test) or derivatives thereof are commonly used in clinics to evaluate patients' readiness to return to work. After reviewing the PILE test, we concluded that it also does not improve the ability of the battery to assess an individual's repetitive lifting capability compared to the method currently used. The

lumbar PILE test involves lifting a box from zero to 30 inches at a rate of four lifts every 20 seconds. This lifting pace translates to a rate of 12 lifts per minute or 720 lifts per hour. The weight is increased by five to 10 lb. every 20 seconds until the individual reaches his or her maximum capacity. As a safety factor, the individual's heart rate is monitored while he or she is performing the test. If the heart rate exceeds 85% of the maximum heart rate, the test is stopped. It is important to note that the heart rate is not used to estimate aerobic capacity. Said another way, the PILE protocol is not a test of endurance.

In general, the problems found with most alternative tests are that the test is not highly similar to what is performed on the job, and/or the results of the test are difficult to compare to the ability level the job requires. While the tests may be useful for generic medical evaluations, the results may not be directly comparable to the physical job requirements. Hence, the medical test would not be appropriate for employment decision-making purposes.

Conclusions

The Advanced Ergonomics test battery design consists of tests that reflect the demands of the specific job for which an individual is being evaluated. The components may include strength tests, flexibility tests, agility tests, posture tolerance tests and a test of aerobic capacity. The strength test involves having the participant handle a box, just as would occur on the job. The results of the test indicate the maximum weight the participant can handle occasionally, which is then compared to the maximum weight a worker could expect to routinely encounter as part of the job. Any other test, such as ones where there are multiple repetitions of the lift at given weights, do not appear to improve the ability to predict the individual's single or repetitive handling capacity or risk of injury.

Repetitive handling capacity is ascertained in the Advanced Ergonomics battery by assessing the participant's aerobic capacity. Knowledge of the individual's aerobic capacity can be translated into the energy expenditure he or she could handle without excessive fatigue for the given shift length. That energy expenditure can then be compared to the energy expenditure required by the job. With warehouse work, the test of aerobic capacity is often the most critical type of test. Roughly 70% of the participants have the aerobic capacity to meet the energy expenditure requirements of the job, whereas more than 90% of the participants have the strength and flexibility

to meet the lifting requirements. Hence, it is absolutely critical that aerobic capacity be taken into account when evaluating the individual's ability to safely perform many of the physically-demanding jobs in industry today.

Section 3: Comprehensive Service

Up to this point, the discussion has focused on details of the Advanced Ergonomics program, as it relates to job analysis, test battery design and validation. In closing the discussion, it is important to reemphasize that an *effective* physical ability testing program is much more than the administration of a test and generation of a report. This is particularly true when the testing is being performed for a company with multiple locations around the country. A comprehensive, well-managed service requires that the following elements also be addressed:

- rapid pass/fail response;
- quality control;
- data management;
- preparedness for legal review;
- test administration management; and
- ongoing physical demand review.

Rapid Pass/Fail Response

The human resource manager needs a rapid response as to whether the individual has the physical ability to safely perform the job. Advanced Ergonomics guarantees delivery of a pass/fail decision within 24 hours of receipt of the test data. However, most tests are processed and results delivered within one to two hours of receipt of the test data.

Quality Control

Managers depend heavily on the accuracy of the test results when making employment decisions. Hence, Advanced Ergonomics constantly monitors the data submitted for any indication of poor test administration, or any other problem that could influence the accuracy of the test. In fact, we follow-up on approximately 12% of the tests that are submitted in order to verify information and resolve unusual findings. One of the advantages of having Advanced Ergonomics provide this service is that our client is assured of an unbiased, consistent quality check across all of the clinics performing tests.

Data Management

It is vitally important to monitor program effectiveness and integrity across all of the sites where testing is being performed. For this reason, Advanced Ergonomics tracks the pass rate trends

by location, and other trends in the data within one location or across locations that would be indicative of problems needing to be addressed. AEI provides quarterly summaries of test results by location and across locations that address these needs. We monitor the pass rates on a quarterly basis to detect problems in the quality of test administration or changes in the applicant pool. Furthermore, EEOC requires periodic reports regarding adverse impact on females when implementing this sort of testing.

Preparedness For Legal Review

There will be discrimination complaints filed in relation to the physical ability test battery. Our experience has been that when the program addresses the requirements EEOC lays out in the Uniform Guidelines on Employee Selection Procedures, complaints are typically resolved by sending the validation document to the agency making the request. This means that the complaint goes no further than an initial review, thereby avoiding the drain of resources associated with prolonged interactions with governmental agencies. Advanced Ergonomics stands behind its work and will defend its decisions.

Test Administration Management

Clinics have turnover of personnel, which necessitates ongoing training and refresher courses. Likewise, companies may choose to change health care providers. Advanced Ergonomics serves as the stable element for the ongoing certification of test administrators and assurance of test administration quality.

Ongoing Physical Demand Review

EEOC requires a periodic review of the job requirements to assure that the test protocol remains appropriate. Advanced Ergonomics reviews current handling tasks in order to identify shifts in strength requirements, and circulates surveys to gather information regarding any changes in work practices that could influence the energy expenditure requirements of the job.

Summary

When evaluating the testing services of various providers, it is important to consider not only the manner in which they perform job analysis, design the test battery and validate the program, but also the manner in which they provide ongoing support. A checklist for comparing testing services of alternative providers is provided in Appendix A to assist in performing a comprehensive review.

Section 4: Program Effectiveness

The intent of the physical ability test program is to identify individuals who have the physical ability to safely perform a given job. Two consequences of a mismatch between worker ability and job demand are decreased tenure and increased risk of injury. The effectiveness of the program can be measured in a number of ways. One measure is the pass rate. It reflects the additional information gained by including physical ability testing in the decision-making process. A second measure is the length of time new workers stay on the job. A third measure is reduction in injury rates. These last two measures reflect different dimensions of the impact of better matching applicants and jobs.

Pass Rates

Due to the high job-relatedness of the test battery, the pass rate is a direct reflection of the difficulty of the particular job for which it was designed. For that reason, the pass rate varies from job to job, though it generally falls in the range of 50% to 95%.

People who fail the battery are generally ones that were about to be hired if testing were not in place. Hence, many managers see the pass rate as an indicator of program effectiveness. Lower pass rates suggest greater difficulty for managers to determine an individual's ability to perform the job in the absence of the information provided by physical ability testing.

Tenure

In a recent analysis by AEI for one of our clients, we were able to look at new-hire tenure over a period of more than four years. Tenure patterns for 350 applicants hired before implementation of physical ability testing were compared to tenure patterns of 453 applicants hired after implementation of testing. It was projected that new-hires who failed the battery had an average tenure of approximately 100 hours, whereas new-hires who passed had an average tenure greater than 2,000 hours. One ramification of this is that it would be expected that those who fail the test would need to be replaced within 100 hours on-the-job, thus wasting the time, effort and expense that went into making those hires. By screening applicants on the basis of the physical ability test battery, significant expense could be saved in the hiring process. The expense reduction would be a product of the pass rate and the expense to advertise, interview, hire and train a new worker.

Reduction in Injury Rates

One method to assess the impact on injury rate involves a criterion-related predictive validation design. This design provides the strongest support for the effectiveness of a test battery in identifying applicants who can safely perform the job at issue. In this type of design, applicants for employment are tested prior to hire. Their test results are **not** used in determining whether or not they should be hired. Instead, results of the testing are kept confidential from all Company personnel. Performance of new-hires who fail the battery can be directly compared to the performance of those who pass.

The drawbacks of a criterion-related predictive validation design are that new-hires who fail the battery must be hired, and that all new-hires need to be tracked for periods up to two years. Once the effectiveness of the battery has been demonstrated with this type of study, it is undesirable to place future applicants who fail at risk by hiring them for a job for which they have not demonstrated the physical ability to safely perform. Furthermore, there is reduced incentive for a company to wait for up to two years to begin screening with the test battery in order to have enough applicants tested and hired so as to meet the requirements of the prospective validation design.

An alternative method for assessing the effectiveness of a testing battery is to compare the performance of new-hires who began work before the test battery was implemented to new-hires who began after implementation. This design is referred to hereafter as a "pre/post-implementation analysis." The major benefit of this study design is that the testing program can be immediately used for making screening decisions rather than waiting until a sufficient sample of new-hires who fail the battery are brought on the job in order to meet the prospective study sample-size requirements. The drawback is that the study design involves comparing performance from two different time periods. Any other changes between those two periods may impact the ability to detect the effectiveness of the screening program. This issue can be addressed by comparing the groups in relatively tight time periods, such as one year, pre- and post-implementation.

Although the pre/post-implementation analysis design is not as strong relative to a predictive validation study design in providing evidence to make causal inferences concerning the implementation of the program, obtaining a consistent effect size across different populations, industries, settings, and time **does** provide a strong inference that the observed effects are due to the

program and not to artifacts such as regression to the mean and selection bias. Taken together, the results of this type of study can provide a strong inference that the observed effects are due to the implementation of the testing process.

Prospective Validation Studies

Three prospective validation studies have been performed by AEI (Anderson, 1992a; Anderson, 1992b; Anderson, 1999). One was based on data from 16 food distribution warehouses in which comparisons were made among performance for new-hires who failed the AEI test battery vs. new-hires who passed (Anderson, 1999). The second was based on data from nine soft drink distribution warehouses (Anderson, 1992a). As with the food distribution warehouse validation study, comparisons were made between performance for new-hires who failed the AEI test battery vs. new-hires who passed. The third was based on data from three retail distribution warehouses (Anderson, 1992b). The same comparisons of performance were made between new-hires who failed and passed the AEI test battery.

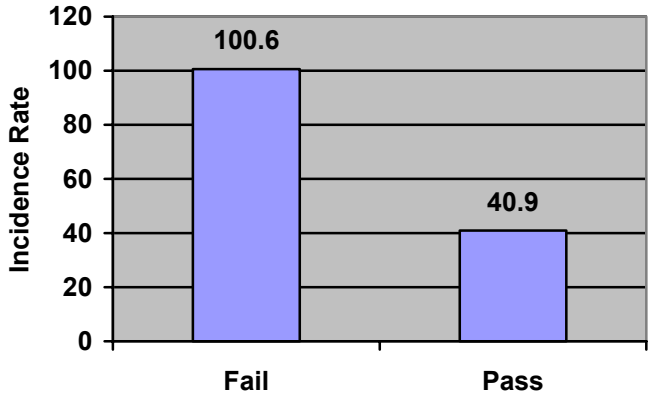
These three studies form the foundation of evidence indicating the effectiveness of the AEI screening program. As was discussed above, this validation study design is the most powerful way to demonstrate program effectiveness. This study design basically involves comparing the performance of new-hires who fail the test battery to the performance of new-hires who pass. Normally, applicants who fail the battery would not be hired, but their test results are not used in the employment decision-making process during the data-gathering phase of this type of study so they are, in fact, hired.

Figure 4.1 provides an example of the process of comparing performance between new-hires who fail the battery vs. those who pass. This figure shows the incidence rate for food distribution workers who had one or more test scores below the cutoffs (failed the battery) vs. those who had test scores that exceeded each cutoff. The incidence rate for a group is calculated by dividing the total number of injuries for the group by the total number of hours worked, and then normalizing the ratio by multiplying by 200,000 hours. This method is in keeping with the approach used by OSHA. The value of normalizing the rates to a common base of 200,000 hours is that it takes into account differences in the number of employees, hiring volume, hours worked, and puts the information reported into a single metric.

It can be seen in Figure 4.1 that the incidence rate for new-hires who failed the battery was 100.6 injuries per 200,000 hours worked, compared to an incidence rate of 40.9 injuries per 200,000 hours worked for new-hires who passed. Said another way, the injury rate for new-hires who pass the battery is 59% lower than the rate for those who fail or less than half as high.

Physical job requirements were quite similar across the locations included in these three studies, so the results can be combined so as to see the overall effectiveness of physical ability testing for these sorts of jobs. The similarity in job requirements was primarily due to the fact that all of the jobs were associated with distribution of merchandise of similar weight and configuration. The test cutoffs were directly based on the physical job requirements. Hence, the cutoffs, in turn, were quite similar across locations and industries.

Figure 4.1: Comparison of Incidence Rates for Grocery Distribution Study



Data was available on injury experience and retention for all locations in the three studies. Retention was measured by calculating the percent of new-hires still employed as of the eighth week. Similar trends were seen at later time intervals. The results are summarized in Table 4.1. Across the three studies, new-hires who pass the battery have 55% fewer injuries and are 35% more likely to still be employed as of the eighth week. These results are statistically significant at $\alpha<0.05$. This indicates that the AEI physical ability test battery is effective in identifying individuals who are more likely to have longer tenure and safer performance.

Table 4.1: Summary of Prospective Validation Results
(New-Hires Who Pass Relative To New-Hires Who Fail)

Study	Decrease In Incidence Rate	Increase In Retention
Food Distribution	59%†	37%†
Soft-Drink Distribution	77%†	166%†
Retail Distribution	38%†	21%
Overall	**55%†**	**35%†**

† indicates statistical significance at $\alpha<0.05$

Pre/Post-Implementation Studies

As was mentioned earlier, pre/post-implementation studies differ from the prospective studies just discussed in that performance is compared between new-hires who started in the year prior to implementation and new-hires who start in the year after implementation. In a pre/post-implementation comparison, applicants who fail the battery post-implementation are not hired, whereas in a predictive validation study, applicants who fail the battery are hired during the study period. Said another way, applicants hired in the pre-implementation phase include ones who would have failed the test battery if it were in use, whereas only applicants who pass the battery are hired in the post-implementation phase. Hence comparisons between the pre- and post-implementation periods are effectively comparisons between the pre-implementation group that has both new-hires who pass and new-hires who fail the battery to the post-implementation group which has only new-hires who pass.

AEI has performed pre/post-implementation studies at 175 locations in different industries including food distribution, soft drink distribution, and retail distribution. These studies cover a total of more than 21 million worker-hours pre- and post-implementation. Samples of results from these studies and other analyses performed by clients themselves are shown in Appendix B. Table 4.2 summarizes the results. The table shows reductions in injury experience for four different injury criteria: 1) Back injuries, 2) other strain/sprain injuries, 3) contact injuries, and 4) overall injuries.

Back injuries are strain or sprain injuries which occur specifically to the back. *Other strains/sprains* refers to strains or sprains of various joints and muscles other than the back, including the shoulder, elbow, wrist, and knee. *Contact injuries* are those injuries which occur due to contact with an object. Examples include lacerations, contusions and fractures. *Overall injuries* are defined as any type of injury that is reported. Negative numbers in the table indicate instances in which the incidence rate of injuries increased.

Overall, there was a 37% to 54% decrease in incidence rate, depending on the type of injury at issue. Decreases were found in all categories of injuries – sprains/strains as well as contact injuries.

Table 4.2: Summary of Pre/Post-Implementation Injury Rate Comparisons:
Percent Reduction in Injury Rate Pre- vs. Post-Implementation.

Industry	Number Locations	Back Sprain/ Strains	Other Sprain/ Strains	Contact Injuries	All Injuries
Food Distribution	69	46%	40%	38%	42%
Soft Drink Distribution	83	78%	40%	55%	48%
Retail Distribution	9	48%	32%	-24%	18%
Other	4	67%	39%	-17%	37%
Overall	**175**	**54%**	**39%**	**37%**	**41%**

The reduction in contact injuries is hypothesized to arise from there being less physical fatigue for a worker who has the strength and endurance in excess of what the job requires. Physical fatigue can lead to loss of coordination, which in turn can lead to injuries related to dropping merchandise, slipping, and falling, among other types. The increases in contact injury rates found in the retail distribution and in "other" industries were thought to be related to process changes, such as introducing a different form of box cutter, as opposed to being related to implementation of the testing battery. There is no reason that implementation of testing would increase an injury rate – at

most there would be no change. It is interesting to note that back injuries and other sprain/strains dramatically decreased at the locations in the retail distribution and "other" industries, which is where the primary impact of physical ability testing would be expected to manifest.

The injury experience for new-hires who fail can be estimated by extracting the anticipated number of hours and injuries for the new-hires who passed from the overall injury rate for the pre-implementation group, based on the experience of the post-implementation group. An injury rate for the new-hires who fail can then be estimated by dividing the remaining injuries by the remaining hours and normalizing by 200,000 hours. Table 4.3 summarizes a comparison of extrapolated injury rates for pre-implementation new-hires who fail the battery with the actual injury rates for post-implementation new-hires, all of whom passed the battery, based on the data from the 175 locations covered in the meta-analysis of pre/post-implementation studies. The projected decrease in incidence rate for those who pass relative to those who fail the battery ranges from 73% to 85% depending on the type of injury. These projections are comparable to the results found in the prospective validation studies. The relative risk of injury (ratio of injury rate for new-hires who fail to injury rate for new-hires who pass) ranges from 4.01 to 6.62. This means, for instance, that new-hires who fail the battery are expected to have more than a six-fold higher risk of back injury than new-hires who pass.

Table 4.3: Injury Rates By Test Result Status

	Back Injuries	Other Sprain-Strain	Contact Injuries	All Injuries
Fail (projected)	49.22	51.90	40.25	137.78
Pass	7.43	12.94	10.74	31.52
Percent Decrease	85%	75%	73%	77%
Relative Risk	6.62	4.01	3.75	4.37

In summary, the results of the three prospective validation studies and the 175 post-implementation studies indicate a substantial reduction in physical injuries as a direct function of the

use of physical ability tests. The fact that the studies were performed across a wide range of industries and geographic locations attests to the stability and robustness of the results achieved. These studies proved that physical ability testing programs can be designed to meet the objective of determining whether a given individual is able to safely perform the job.

Furthermore, these studies provided the foundation for implementing physical ability testing programs in a wide range of industries that have significant physical demands. It would be expected that similar results would be achieved for other jobs with similar demands. Examples include bottled water distribution, beer distribution, auto parts distribution, ambulance service, barge operations, steel manufacturing, shipbuilding, construction and aluminum product production.

Section 5: Legal Reviews

Physical ability test batteries implemented for companies by AEI have been subject to legal review by Federal and State agencies whose responsibility is to monitor employment discrimination. Reviews have come about through a variety of circumstances, ranging from routine audits performed by OFCCP as part of contract compliance reviews to complaints filed by job applicants who felt they were discriminated against on the basis of their age, race, ethnic group, gender or disability. When a job applicant files a complaint, the governmental agency follows up with the employer to determine if there has been a cause for the complaint. If the agency determines that there may be substantive grounds to believe there has been cause for complaint, a more extensive inquiry is initiated. If the complaint cannot be resolved through the conciliation process, it is taken to litigation. A few program reviews have arisen in the process of EEOC auditing a company's compliance with a conciliation agreement that related to other hiring practices in place.

AEI is aware of over 80 reviews that have been completed over the course of time of the program being available. There have probably been additional OFCCP audits, but AEI only becomes aware of them if OFCCP asks the company for further information about the physical ability testing program. The outcome of all of these reviews has been that agencies have allowed the testing program to continue to be used. All indications are that this successful history with legal reviews will continue into the future.

References

Anderson, C.K. "The Impact Of Physical Ability Testing On Injuries And Retention For The Coca Cola Bottlers Association". Technical Report, Advanced Ergonomics, Inc., 1992a.

Anderson, C.K. "Impact Of Physical Ability Testing On Workers' Compensation Injury Rate And Severity - Target Distribution Centers". Technical Report, Advanced Ergonomics, Inc. 1992b.

Anderson, CK. *Physical Ability Testing Prospective Validation Study for Food Distribution Warehousing*. Advanced Ergonomics Technical Report, 1999.

Appendix A: Testing Services Comparison

Attached is a worksheet that you can use to compare the testing services offered by Advanced Ergonomics to other potential service providers. You'll want to know how they compare not only on price, but also on the issues we've listed. This will ensure that you have a comprehensive service that will not only reduce your injury experience but also avoid potential costs associated with implementing a testing program that cannot be defended against complaints of discrimination on the basis of sex, age, race or disability. Those costs can include back wages for all applicants who filed the class action, lawyer fees and potentially even punitive damages.

Testing Services Comparison Worksheet

	Other Provider	AEI
Principle business	?	Physical ability testing
Able to service all locations	?	Yes
Provides pass/fail decision	?	Yes
Uses job-related tests and cutoffs	?	Yes
Overall pass rate	?	82%
Evaluates endurance capability relative to job requirement	?	Yes
Third party review of test quality and accuracy	?	Yes
Documented reduction in injuries for clients	?	Yes
Has performed prospective validation studies	?	Yes
PhD-level expert provided to respond to discrimination complaints who has experience dealing with complaints on the basis of sex, age, race and disability	?	Yes
Successful resolution in all closed discrimination complaints	?	Yes
Provides validation documentation meeting EEOC requirements	?	Yes
Annual review of job requirements	?	Yes

Appendix B: Sample Analyses Performed By Advanced Ergonomics

Advanced Ergonomics compares incidence rates in the year before and the year after implementation of testing as part of the process of evaluating the effectiveness of the program. The pre-implementation, or baseline, incidence rate is typically based on the experience of the new-hires in the year prior to implementation. If the sample size of pre-implementation new-hires is too small, the rates are unreliable. In that case, the incidence rate for all incumbents in the year prior to implementation is used as the baseline.

Listed below are results of a representative sample of pre-/post-implementation analyses performed by AEI. Also included are results of studies performed by clients.

Kroger - Nashville, TN

Overall injury rates for Order Assemblers dropped by 55%.
Contact injury rate for Order Selectors dropped 69%.
(incumbent baseline)

Kroger - Houston, TX

Selector back injuries down 35%.
(new-hire baseline)

Kroger - Indianapolis, IN

Overall injury rate for selectors dropped 39%.
(new-hire baseline)

Kroger - Fort Wayne, IN

61% reduction in overall injury rate for selectors.
69% reduction in selector back injuries.
Sprain/strain injury rate down 63% for selectors.
(new-hire baseline)

SUPERVALU, Inc. - Belle Vernon, PA

Overall injury rate dropped 69.5% for selectors
(new-hire baseline)

SUPERVALU, Inc. - Hazelwood, MO

Selector back injury rate dropped 31%
Selector contact injury rate dropped 51%
Overall injury rate dropped 26% for selectors
(new-hire baseline)

SUPERVALU, Inc. - Pleasant Prairie, WI

Selector back injury rate is 15% lower
Overall injury rate is 8% lower for selectors
(new-hire baseline)

SUPERVALU, Inc. - Fort Wayne, IN

Selector back injury rate dropped 28%
Selector strain/sprain injury rate dropped 7%
(new-hire baseline)

SUPERVALU, Inc. - Fargo, ND

No injuries to new-hire selectors tested

SUPERVALU, Inc. - New Stanton, PA

Selector back injury rate dropped 53%
Selector contact injury rate dropped 59%
Overall injury rate dropped 65% for selectors
(new-hire baseline)

SUPERVALU, Inc. - Bismarck, ND

Overall injury rate dropped 77% for selectors
(incumbent baseline)

SUPERVALU, Inc. - Tacoma, WA

Selector back injury rate dropped 35%
(incumbent baseline)

The SYGMA Network - Mechanicsburg, PA

Driver overall injury rate dropped 35%.
Warehouse overall injury rate dropped 24%.
(new-hire baseline)

The SYGMA Network - Grand Prairie, TX

No injuries among new-hires tested.

The SYGMA Network - San Antonio, TX

No back injuries to workers hired since testing.
Overall injury rate dropped 30% for selectors and 15% for drivers.
(incumbent baseline)

The Coca-Cola Bottling Group (Southwest) - Amarillo, TX

Injury rates for new-hires - zero.

The Coca-Cola Bottling Group (Southwest) - Lubbock, TX

Overall injury rates down 59%.
Musculoskeletal injury rate down 49%.
(new-hire baseline)

The Coca-Cola Bottling Group (Southwest) - Abilene, TX

No injuries for new-hires.

Sample Analyses Performed By Clients

The Great Atlantic & Pacific Tea Co., Inc. (A&P)
Carl Frey, Corporate Safety Director
The AEI program is the primary reason our company's accident rate plummeted 68%. The payback was very high. We calculate that we saved $5 million in 1992.

Delta Beverage Group
Jimmy St. Pierre, Area Loss Control Manager
We were able to reduce our workers compensation injury frequency rate by 40.9% through the use of the Advanced Ergonomics testing program.

FirstFleet Trucking
Debra Parsons, Director of Risk Management
The Advanced Ergonomics program serves a valid purpose in the process of new employee selection. We currently use this program at all our locations and have seen very good results throughout.

Pepsi
Matt Hirsheimer, Corporate Safety Director
The Advanced Ergonomics process insures we hire people who can perform our jobs more safely. The two pilots we ran showed a 40% and 57% reduction in injuries to new hires, which is significant in creating and maintaining a safe work environment.

Rhodes Furniture
Bob Weir, Corporate Safety Director
Rhodes Furniture began using the Advanced Ergonomics physical ability testing program in 1995 for the screening of new-hire warehouse, home delivery and furniture repair candidates. During this seven-year time span, Rhodes Furniture realized 3.8 million dollars in bottom line savings in workers' compensation costs for an investment of just over $200,000 dollars in implementation and testing fees. The total workers' compensation savings for Rhodes since 1995 jumps to over 20 million dollars if conservative indirect costs are also included.

Given these impressive results, the Advanced Ergonomics physical ability testing battery continues to be a major program in our continuing efforts to make the workplace safer and more cost effective.