

Schmidt, F. & Hunter, J. (1998). The validity and utility of selection methods in personnel psychology: Practical and Theoretical Implications of 85 years of research findings. *Psychological Bulletin*, 124(2), 262-274.

Researcher Notes

Categories addressed: Recruitment, selection.

Participants: Literature Review.

Method: This study uses a meta-analysis method, and shows the validity of 19 selection procedures for predicting job performance and training performance, as well as the validity of paired combinations of general mental ability (GMA) and the 18 other selection procedures.

Independent Variables: Selection procedures (e.g. general mental ability, work sample tests, integrity test, structured and unstructured employment interviews, job knowledge tests, tryout procedure, peer ratings, behavioral consistency method, reference checks, job experience, biographical data, point method, interests, graphology, age of applicants).

Dependent Variables: Job performance and training performance.

Purpose of the study: Researchers look at the implications of these selection procedures in making decisions about hiring, training, and developmental assessments, as well as the development of theories of job performance. The use of effective hiring methods increases employee performance, monetary value of output, and job-related-skills (Hunter, Schmidt, & Judiesch, 1990).

Findings:

- On the basis of meta-analytic findings, GMA (also called general cognitive ability and general intelligence) occupies a special place, for several reasons. First, of all procedures that can be used for all jobs, whether entry level or advanced, it has the highest validity and lowest application cost. Second, GMA has been shown to be the best available predictor of job-related learning. Third, the theoretical foundation for GMA is stronger than any other personnel measure (theories of intelligence have been developed and tested by psychologists for over 90 years).
- The value of .51 for the validity of GMA is from a very large meta-analytic study conducted for the U.S Department of Labor (Hunter, 1980; Hunter & Hunter, 1984). The database for this unique meta-analysis included over 32,000 employees in 515 widely diverse civilian jobs. This meta-analysis examined both performance on the job and performance in job training programs. Because of the special status of GMA, it can be considered the primary personnel measure for hiring decisions, and the remaining 18 personnel measures as supplements to GMA measures.
- Work sample tests increase job performance validity by 24% over GMA alone. However, work sample tests can be used only with applicants who already know the job. Such workers do not

need to be trained, and so the ability of work sample tests to predict training performance has not been studied.

- Integrity tests (used to hire employees with reduced probability of counterproductive job behaviors, such as drinking or drugs on the job, fighting on the job, stealing from the employer, sabotaging equipment and other undesired behaviors) increase the validity and utility of job performance by 20% over GMA alone. In the prediction of training performance, integrity tests appear to produce higher incremental validity than any other measure studied to date.
- The average validity of structured interviews is .51, versus .38 for unstructured interviews. An equally weighted combination of structured interview and GMA measure yields a validity of .63.
- Job knowledge tests increase validity by .07 over the GMA measures alone, yielding a 14% increase in validity and utility. However, like work sample measures, job knowledge tests cannot be used to evaluate and hire inexperienced workers.
- Job tryout procedure increases the validity and utility by the same amount as job knowledge tests (14%). Yet, tryout procedures are very expensive to implement, and low job performance by minimally screened probationary workers can lead to serious economic losses.
- Peer ratings also increase the validity and utility by 14%. But, like the job tryout procedures, peer ratings have some limitations. First, they cannot be used for evaluating and hiring applicants from outside the organization; they can be used only for internal job assignment, promotion, or training assessment. Peer ratings have been found to predict performance in training programs with a mean validity of .36.
- The behavioral consistency method (based on the principle that the best predictor of future performance is past performance. Applicants are asked to describe their past achievements that best illustrate their ability to perform functions at a high level, and supervisors score the answers with the aid of specific scales) increases the validity and utility by 14%. Use of this method is not limited to applicants with previous experience on the job in question. Especially for higher level jobs, this method may be well worth the cost and effort.
- No information is available on the validity of the job tryout or the behavioral consistency procedures for predicting performance in training programs.
- Reference checks increments validity by 12%, only two percentage points less than the increments for the preceding methods. However, special attention should be paid regarding to the legal climate in the United States and its laws on legal liability for employers.
- Regarding to job experience, studies have found that when experience on the job does not exceed 5 years, the correlation between amount of job experience and job performance is considerably larger .33 when job performance is measured by supervisory ratings and .47 when job performance is measured using a work sample test.
- Biographical data increase the validity over GMA for only .01 on average (a 2% increase). This is because biographical data correlates substantially with GMA.
- Point method of evaluating previous training and experience (T&E) is used mostly in government hiring. All point methods are credentialistic; typically an applicant receives a fixed number of points for (a) each year or month of experience on the same of similar job, (b) each year of relevant schooling or each course taken, (c) each relevant training program completed,

and so on. This method has low validity and produces only a 2% increase in validity over GMA alone. This T&E method has not been used to predict performance in training programs.

- Amount of education has even lower validity for predicting job performance than the T&E point method (.10). However, amount of education predicts learning in job training programs better than it predicts performance on the job. Hunter and Hunter (1984) found a mean validity of .20 for performance in training programs.
- Interest of applicants is believed to have higher job prediction validity. This has a validity of .10 which shows that this is true only to a very limited extent.
- Graphology has no empirical evidence to demonstrate validity in job performance and training programs.
- Likewise, age of applicants shows no validity for predicting job performance.

Practice Implications:

- GMA can be considered the primary personnel measure for hiring decisions. Previous research also suggests that other central determining variables in job performance may be job experience (e.g. opportunity to learn), and the personality trait of conscientiousness.
- This is consistent with this study's conclusion that a combination of GMA tests and integrity test (which measures mostly conscientiousness) has the highest high validity (.65) for predicting job performance.
- Another combination of high validity (.63) is GMA plus a structured interview, which may be in part measure conscientiousness and related personality traits (such as agreeableness and emotional stability, which are also measured in part by integrity tests).