Part I

Selection

1

Select on Intelligence

Frank L. Schmidt

Other things equal, higher intelligence leads to better job performance on all jobs. Intelligence is the major determinant of job performance, and therefore hiring people based on intelligence leads to marked improvements in job performance – improvements that have high economic value to the firm. This principle is the subject of this chapter.

This principle is very broad: it applies to all types of jobs at all job levels. Until a couple of decades ago, most people believed that general principles of this sort were impossible in personnel selection and other social science areas. It was believed that each organization, work setting, and job was unique and that it was not possible to know which selection methods would work on any job without conducting a validation study on that job in that organization. This belief, called the theory of situational specificity, was based on the fact that different validity studies of the same selection procedure(s) in different jobs in the same organization and/or different organizations appeared to give different results. However, we now know that these "conflicting findings" were mostly due to statistical and measurement artifacts and that some selection procedures have high validity for predicting performance on all jobs (e.g. intelligence) and others do a poor job of predicting performance on *any* job (e.g. graphology) (Schmidt and Hunter, 1981, 1998). This discovery was made possible by new methods, called meta-analysis or validity generalization methods, that allow researchers to statistically combine results across many studies.

Meta-analysis has also made possible the development of general principles in many other areas beyond personnel selection (Hunter and Schmidt, 2004; Schmidt, 1992). For example, it has been used to calibrate the relationships between job satisfaction and job performance with precision (Judge, Thoresen, Bono, and Patton, 2001) and between organizational commitment and work-related outcomes including job performance (Cooper-Hakim and Viswesvaran, 2005).

What is Intelligence?

Intelligence is not the ability to adapt to one's environment; insects, mosses, and bacteria are well adapted to their environments, but they are not intelligent. There are many ways

in which organisms can adapt well to their environments; use of intelligence is only one possible way. Intelligence is the ability to grasp and reason correctly with abstractions (concepts) and solve problems. However, perhaps a more useful definition is that intelligence is the ability to learn. Higher intelligence leads to more rapid learning, and the more complex the material to be learned, the more this is true. Intelligence is often referred to as general mental ability (GMA) and general cognitive ability, and we use all these terms interchangeably in this chapter.

Intelligence is the broadest of all human mental abilities. Narrower abilities include verbal ability, quantitative ability, and spatial ability. These narrower abilities are often referred to as special aptitudes. These special aptitudes do predict job performance (although less well than GMA), but only because special aptitude tests measure general intelligence as well as specific aptitudes (Brown, Le, and Schmidt, 2006; Schmidt, Ones, and Hunter, 1992). It is the GMA component in these specific aptitude tests that predicts job performance. For example, when a test of verbal ability predicts job or training performance, it is the GMA part of that test – not the specifically verbal part – that does the predicting (Brown et al., 2006).

Intelligence predicts many important life outcomes in addition to job performance: performance in school, amount of education obtained, rate of promotion on the job, ultimate job level attained, income, and many other things (Brody, 1992; Herrnstein and Murray, 1994; Gottfredson, 1996; Jensen, 1998). It is even involved in everyday activities such as shopping, driving, and paying bills (Gottfredson, 1996). No other trait – not even conscientiousness – predicts so many important real world outcomes so well. In this sense, intelligence is the most important trait or construct in all of psychology, and the most "successful" trait in applied psychology.

The thousands of studies showing the link between intelligence (GMA) and job performance have been combined into many different meta-analyses. Ree and co-workers have shown this for military jobs (Olea and Ree, 1994; Ree and Earles, 1991, 1992; Ree, Earles, and Teachout, 1994), as have McHenry, Hough, Toquam, Hanson, and Ashworth (1990) in the famous Project A military study. (With a budget of 24 million dollars, Project A is the largest test validity study ever conducted.) Hunter and Hunter (1984) have shown this link for a wide variety of civilian jobs, using the US Employment Service database of studies. Schmidt, Hunter, and Pearlman (1980) have shown it for both civilian and military jobs. Other large meta-analytic studies are described in Hunter and Schmidt (1996), Schmidt (2002), and Schmidt and Hunter (2004). Salgado and his colleagues (Salgado, Anderson, Moscoso, Bertua, and de Fruyt, 2003a, 2003b) demonstrated the link between GMA and job performance across settings in the European countries. The amount of empirical evidence supporting this principle is today so massive that it is hard to find anyone who questions the principle.

There has been an important development since the first edition of this book appeared in 2000: a new and more accurate method for correcting for the biases created by range restriction has been developed and applied (Hunter, Schmidt, and Le, 2006; Schmidt, Oh, and Le, 2006; Schmidt, Shaffer, and Oh, 2008). (Range restriction is the condition in which variability of the predictor (here intelligence) in one's sample of people (job incumbents) is artificially lower than in the population of people (job applicants) one wants to get estimates for.) Application of this procedure to existing data shows that previous estimates of the validity of GMA – including those in the 2000 version of this chapter – were underestimated by 25% to 30%. In this chapter, I present the updated, more accurate validity estimates. When performance is measured objectively using carefully constructed work

sample tests (samples of actual job tasks), the correlation (validity) with intelligence measures is about .84–84% as large as the maximum possible value of 1.00, which represents *perfect* prediction. When performance is measured using ratings of job performance by supervisors, the correlation with intelligence measures is .66 for medium complexity jobs (over 60% of all jobs). For more complex jobs, this value is larger (e.g. .74 for professional and managerial jobs), and for simpler jobs this value is not as high (e.g. .56 for semi-skilled jobs). Another performance measure that is important is amount learned in job training programs (Hunter et al., 2006). Regardless of job level, intelligence measures predict amount learned in training with validity of about .74 (Schmidt, Shaffer, and Oh, 2008).

Why Does Intelligence Predict Job Performance?

It is one thing to have overwhelming empirical evidence showing a principle is true and quite another to explain *why* the principle is true. *Why* does GMA predict job performance? The primary reason is that people who are more intelligent learn more job knowledge and learn it faster. The major direct determinant of job performance is not GMA but job knowledge. People who do not know how to do a job cannot perform that job well. Research has shown that considerable job knowledge is required to perform even jobs most college students would think of as "simple jobs," such as truck driver or machine operator. More complex jobs require even more job knowledge. The simplest model of job performance is this: GMA causes job knowledge, which in turn causes job performance. But this model is a little too simple: there is also a causal path directly from GMA to job performance, independent of job knowledge. That is, even when workers have equal job knowledge, the more intelligent workers have higher job performance. This is because there are problems that come up on the job that are not covered by previous job knowledge, and GMA is used directly on the job to solve these problems. Many studies have tested and supported this causal model (Hunter, 1986; Ree, Earles, and Teachout, 1994; Schmidt, Hunter, and Outerbridge, 1986). This research is reviewed by Schmidt and Hunter (1992), Hunter and Schmidt (1996), and Schmidt and Hunter (2004). It has also been shown that over their careers people gradually move into jobs that are consistent with their level of GMA (Wilk, Desmariais, and Sackett, 1995; Wilk and Sackett, 1996). That is, a process that sorts people on GMA takes place gradually over time in everyday life. People whose GMA exceeds their job level tend to move up to more complex jobs; and people whose GMA is below their job level tend to move down.

There is a broader theory that explains these research results: the traditional psychological theory of human learning (Hunter and Schmidt, 1996; Schmidt and Hunter, 2004). This theory correctly predicted that the effect of GMA would be on the learning of job knowledge. The false theory of situational specificity became widely accepted during the first eight decades of the 20th century in considerable part because personnel psychologists mistakenly ignored the research on human learning.

Many lay people find it hard to believe that GMA is the dominant determinant of job performance. Often they have known people who were very intelligent but who were dismal failures on the job because of "bad behaviors" such as repeated absences from work, carelessness at work, hostility toward the supervisor, unwillingness to work overtime to meet a deadline, or stealing from the company. These are examples of so-called "counterproductive work behaviors" (CWBs). Integrity tests predict CWBs with a validity of about

.35 (Ones, Viswesvaran and Schmidt, 1993). People with lower scores on integrity tests show more CWBs. The personality trait of conscientiousness also predicts CWBs (again, negatively). However, a recent large-scale study ($N > 800$) found that GMA predicted CWBs with a validity of .47; when the more accurate correction for range restriction is applied, this figure becomes .57. So it is possible that the best predictor of CWBs is GMA. People who are more intelligent show fewer CWBs.

There is also a facet of job performance called "contextual performance" (CP). CP is just good citizenship behaviors, while CWB is bad citizenship behaviors as discussed above. CP behaviors include willingness to help train new employees, willingness to work late in an emergency or on a holiday, supporting the community relations and reputation of the company, and many other such behaviors. CP behaviors and CWBs are different from core job performance but are often confused with core job performance by lay observers. CP and CWB behaviors are predicted by measures of the personality traits of conscientiousness and to a lesser extent agreeableness (Dalal, 2005). We do not yet know whether GMA predicts CP behaviors; these studies have yet to be done. Low ability leads to an inability to perform well; low conscientiousness and low agreeableness lead, not primarily to low performance on core job tasks but to lack of CP and/or more displays of organizationally disruptive behaviors (CWBs). These disruptive behaviors are more visible to lay observers (and to many supervisors) than differences between employees in core job performance, probably because they appear so willful. On the other hand, a low ability employee has difficulty learning how to perform the job, but if he/she has a "good attitude," this employee often seems like less of a problem than one showing CWBs. This makes it difficult for some to clearly see the GMA–performance link in the real world (Hunter and Schmidt, 1996).

Of course, low conscientiousness can lead to less effective performance if it results in reduced effort (see Chapter 2, this volume). For objective measures of job performance, empirical evidence indicates that on typical jobs this effect is limited, probably because most jobs are fairly structured, reducing the scope for individual differences in effort to operate (Hunter, Schmidt, Rauschenberger and Jayne, 2000; Hunter and Schmidt, 1996). However, it is important to remember that when supervisors rate job performance, they incorporate into their ratings both CP behaviors and CWBs, in addition to core job performance (Orr, Sackett, and Mercer, 1989; Rotundo and Sackett, 2002). Hence supervisory ratings reflect a combination of core job performance and citizenship behaviors, both good and bad. In the case of ratings, low conscientiousness and low agreeableness lead to poorer citizenship behaviors, which lead to lower ratings of overall performance. For the typical job, the weight on conscientiousness in predicting objectively measured core job performance is only 20% as large as the weight on GMA. In predicting supervisory ratings of job performance, it is 40% as large (Schmidt, Shaffer, and Oh, 2008).

What is Required to Make This Principle Work?

There are three conditions that are required to make this principle work. That is, there are three conditions that are required for companies to improve job performance levels by using GMA in hiring and to reap the resulting economic benefits.

Selectivity

First, the company must be able to be selective in who it hires. For example, if the labor market is so tight that all who apply for jobs must be hired, then there can be no selection and hence no gain. The gain in job performance per person hired is greatest with low selection ratios. For example, if one company can afford to hire only the top scoring 10%, while another must hire the top scoring 90% of all applicants, then with other things equal the first company will have a much larger gain in job performance.

There is another way to look at this: companies must provide conditions of employment that are good enough to attract more applicants than they have jobs to fill. It is even better when they can go beyond that and attract not only a lot of applicants, but the higher ability ones that are in that applicant pool. In addition, to realize *maximum* value from GMA-based selection, employers must be able to *retain* the high performing employees they hire.

Measuring general mental ability

Second, the company must have some way of measuring GMA. The usual and best procedure is a standardized employment test of general intelligence, such as the Wonderlic Personnel Test. Such tests are readily available at modest cost. Less valid are proxy measures such as grade point average (GPA) or class rank. Such proxy measures are partial measures of intelligence. Also, intelligence can be assessed to some extent during the employment interview (Huffcutt, Roth, and McDaniel, 1996), although this is a much less valid measure of GMA than a standardized written test.

Variability in job performance

Third, the variability in job performance must be greater than zero. That is, if all applicants after being hired would have the same level of job performance anyway, then nothing can be gained by hiring "the best." This condition is always met. That is, on all jobs studied there have been large differences between different workers in quality and quantity of output. Hunter, Schmidt, and Judiesch (1990) meta-analyzed all available studies and found large differences between employees. In unskilled and semi-skilled jobs, they found workers in the top 1% of performance produced over three times as much output as those in the bottom 1%. In skilled jobs, top workers produced 15 times as much as bottom workers. In professional and managerial jobs, the differences were even larger. These are very large differences, and they are the reason it pays off so handsomely to hire the best workers.

There is another advantage to hiring the best workers: the pool of talent available for future promotion is greatly increased. This is of great value to employers, because it helps ensure high performance all the way up through the ranks of managers. When the right people are promoted, their value to the firm in their new jobs is even greater than in their original jobs. Thus selection of high ability people has implications not only for the job they are hired onto, but for other jobs in the organization, too.

Are There Exceptions to This Principle?

As long as the three conditions described above are met, there are no known exceptions to this principle. That is, there are no known cases or situations in which it is inadvisable to select employees for general intelligence.

However, there are some people, particularly labor leaders, who believe there is an exception. These people believe that companies should not select on mental ability if they can select on job experience instead. That is, they believe that job experience is a better predictor of job performance than general intelligence. What does research show? For applicants with job experience of between none and five years, experience *is* a good predictor of job performance. But in the range of higher levels of experience, say from five to 30 years of job experience, job experience does not predict performance very well (Schmidt, Hunter, Outerbridge, and Goff, 1988; Hunter and Schmidt, 1996). On most jobs, once people have about five years of experience, further experience does not contribute much to higher performance. This is probably because experience beyond five years does not lead to further increases in job knowledge. This, in turn, may be due to the fact that after five years of on-the-job learning, people in the typical job are forgetting job knowledge about as fast as they are learning new job knowledge.

Another important fact is this: even for new hires in the one to five year range of job experience, where experience is a valid predictor of job performance, the validity declines over time. That is, experience predicts performance quite well for the first three years or so on the job and then starts to decline. By 12 years on the job, experience has low validity. But GMA continues to predict job performance quite well even after people have been on the job 12 years or more.

What this means is that job experience is *not* a *substitute* for GMA. In the long run, hiring on intelligence pays off much more than hiring on job experience (Hunter and Schmidt, 1996). So if you had to choose, you should choose GMA. However, typically, you do not have to choose; more than one procedure can be used. It may be desirable to use *both* experience and GMA in hiring; as discussed later, it is usually best to use multiple hiring methods. But in this case, the *weighting* given to GMA should be higher than the weighting given to job experience.

Issues in Implementing an Ability-based Hiring System

Can intelligence be too high?

One issue is whether an applicant can have too much intelligence for a job. Recently, an applicant was rejected for a job as a police officer in a New Jersey city on grounds that his intelligence test score was too high! This city believed something that many people believe: that intelligence leads to better job performance *but only up to a point.* After that, more intelligence leads to *lower* job performance. Hundreds of studies have shown that this is false. Higher intelligence leads to better job performance up to the highest levels of intelligence (Coward and Sackett, 1990). There is a straight line (linear) relationship between intelligence and job performance. Why do so many people believe otherwise? Probably because they imagine a university professor or a medical doctor working as a janitor, and they think "This person would be so bored with this job that he would do a poor job." They forget that the university professor or doctor would never apply for the janitor's job to begin with.

Among people who actually apply to get real jobs, there is a straight line relationship between intelligence and performance; the higher the intelligence, the better the job performance. Hence, we do not have to worry about hiring people who are too intelligent for the job.

Does only intelligence matter in jobs?

A second issue is the one alluded to earlier: Although intelligence is the best predictor of job performance, it does not follow that use of intelligence *alone* in hiring is the best way to select people. In fact, it is well known that other predictors can be used along with intelligence to produce better predictions of job performance than intelligence alone. For example, for most jobs an intelligence test combined with an integrity test (a composite personality of conscientiousness, emotional stability, and agreeableness) is 20% more valid than an intelligence test alone. Adding a structured employment interview to an intelligence test increases validity by 14% (Oh, Schmidt, and Shaffer, 2008). It is almost always possible to add supplementary measures that increase validity. Some of these measures are discussed in other chapters in this book (e.g. Barrick and Mount's chapter on selection of conscientiousness and emotional stability).

Are there legal risks in selecting for intelligence?

A third issue is the potential for legal risks. Members of some minority groups, particularly blacks and Hispanics, typically have lower average scores on GMA tests, leading to lower hiring rates. Government agencies such as the Equal Employment Opportunity Commission refer to these lower hiring rates as "adverse impact." The term adverse impact is deceptive, because it implies that the GMA tests *create* the difference in test scores, when in fact the tests only measure real pre-existing differences in mental skills. This is shown by the fact that minorities and non-minorities with the same test scores have the same level of later job performance. That is, the test scores predict equally accurately for all groups; they are predictively fair or unbiased (Schmidt, 1988; Wigdor and Garner, 1982).

Despite this fact, a lower hiring rate for minorities does sometimes lead to lawsuits. Employers can win these suits by demonstrating that the tests are valid predictors of job performance. Today, such demonstrations rely increasingly on summaries of the kinds of research findings discussed in this chapter, rather than on studies conducted by the employer. (This is part of the move away from the theory of situational specificity, discussed earlier.) Since around the mid 1980s, employers have been winning more and more such suits, and today they prevail in 80% or more of such suits. Research shows that the value of the increases in job performance from good selection overshadows any potential legal costs stemming from defending against such suits. But a key fact is that today there are far fewer such suits to begin with. Currently, less than 1% of employment-related lawsuits are challenges to selection tests or other hiring procedures. This is almost certainly due to the greatly reduced chances of winning such suits.

Political risks

However, this does not mean that all employers are willing to use intelligence tests in hiring. Although the percentage of employers using GMA tests has been increasing, some

firms view even the possibility of a lawsuit as a public relations disaster. They feel that even if they win, they still lose on the public relations front. And they believe that public relations problems can reduce sales and profits. These firms – mostly larger companies that sell directly to consumers – are willing to tolerate lower levels of job performance to avoid even the possibility of such a problem. Unfortunately for such firms, not using GMA tests does not remove the possibility of lawsuits. Other selection procedures also produce "adverse impact." Employers have tried to reduce adverse impact by introducing various forms of minority preferences in hiring, but courts have recently begun to strike down many forms of minority preferences. For example, under the 1991 Civil Rights Act, it is illegal to adjust test scores or other scores to equalize minority and non-minority hiring rates. This issue is one that will probably remain unsettled for some time.

Many firms that rarely use written GMA tests build oral GMA tests into the interview process. For example, in many employment interviews at Microsoft, job applicants are asked to solve complex mental puzzles that require high GMA to answer correctly. In fact, even ordinary job interviews have been found to be correlated with GMA scores (Huffcutt et al., 1996). And, as would be expected from this fact, it has recently been found that even ordinary job interviews show larger minority–majority differences (and thus "adverse impact") than was previously believed to be the case (Roth, Bobko, Switzer, and Dean, 2001).

The effect of testing for intelligence on employee attitudes

A fifth issue is whether the use of mental ability tests turns off applicants. Some have argued that applicants do not like to take ability tests. However, surveys of applicant attitudes reveal that they view mental ability and GMA tests as generally relevant to job performance (more so than they do personality, bio-data, and integrity tests, for example), and that they do not have a negative attitude toward such tests (Hausknecht, Day, and Thomas, 2004). It also appears to be the case that when GMA or other ability tests are used, applicants view the selection requirements as being higher and this increases the status of the job and hence its attractiveness. That is, something that is harder to attain is viewed as being more valuable.

The economic value of hiring on intelligence

A final issue is whether the economic value of the job performance gains from GMA-based hiring is cancelled out by higher wages and salaries. The argument is that if a firm hires more intelligent people, they will have to pay them more and this will cancel out the gains from the increased job performance. However, in most cases it appears that there is no increase in compensation costs, at least initially. This is especially likely to be the case when few of the firm's competitors use GMA measures in their hiring. Typically, there is a pool of available applicants in the area for a particular type of job, and the higher GMA applicants have no immediate effective way to command higher initial wages.

However, after some time on the job, when higher GMA employees have developed high levels of performance, the employer can afford to share some of these gains with such employees in the form of higher wages or salaries. In some cases, this might be necessary to retain high performing employees. In any event, the payoff to the employer in terms of enhanced job performance is much greater than any increase in compensation cost.

Although most employers, for most jobs, do not pay different people in the same job at different rates, they do typically promote the top workers to higher level jobs, and this does result in higher pay. But at promotion the value of the worker's performance to the firm increases much more than the worker's pay, creating another large net benefit to the firm of good selection. On the other hand, employers that hire only mediocre or poor workers at entry level find that their higher level jobs also become filled with mediocre or poor performers. Again, as noted earlier, selection based on GMA improves performance not only in the job in question, but also later in higher level jobs in the firm.

Case Examples

We will first look at two negative examples and then examine two positive examples of real world applications of GMA-based hiring.

US Steel plant at Fairless Hill, PA

Up until 1978, the US Steel plant at Fairless Hills, PA, selected applicants into their skilled trades apprentice programs based on the applicants' total scores of a battery of ability tests. These total scores were a good measure of GMA, and selection was from the top down. The plant maintained apprentice programs in the wide variety of skilled trades needed to run a steel mill: machinists, tool and die makers, electricians, sheet metal workers, etc. The local unit of the United Steelworkers Union, however, did not like this selection method. In negotiations with the union, the company agreed to modify the selection system. In the new system, all applicants who scored above a low cut-off on each test, set at about the 7th grade level, were considered equally qualified and eligible for hire. Only a few applicants were screened out by this procedure. Applicants in the passing group were selected based on plant seniority only. Hence, this plant went from a GMA-based hiring system to one in which GMA played only a very minor role.

The apprentice training center at Fairless Hills was a well-run facility that kept excellent records of apprentice performance from both before and after the change in the selection system. These records showed that after the new selection system was introduced, performance plummeted. Scores on the mastery tests of amount learned in training declined markedly. The flunk-out and drop-out rates increased dramatically. The training time and training costs of those who did make it through the program increased substantially – because many apprentices had to retake multiple units in the training. And finally, the ratings of later performance on the job out in the plant declined.

This was a well-controlled natural quasi-experiment. The only change made was the lowering of mental ability standards in selection. The training program and the tests given in the program remained the same. The decline in performance was clearly due to the lower intelligence of the new apprentices.

The Washington, DC police force

Up until the mid 1980s the Washington, DC police force was one of the best in the USA. Applicants were selected for Police Academy training based on a general intelligence test

constructed for the District of Columbia by the US Office of Personnel Management (OPM), as required by then existing Congressional regulations. This test had been challenged legally and the case had gone all the way to the US Supreme Court, where it had been upheld. A background investigation was also part of the selection process. The mayor of Washington, Marion Barry, repeatedly voiced opposition to both the test and the background check on grounds that the failure rate on both was higher for blacks. In 1987, when Congress relinquished control over the selection process to the Mayor's office, Barry took responsibility for the selection process out of OPM's hands. He then eliminated both the GMA test and the background test. The replacement selection process was somewhat unclear, but reputedly involved fairly perfunctory interviews.

The first consequence was that the flunk-out rate in the Police Academy soared, with over 80% of the new hires being incapable of completing the required training. Failure rates that high were viewed as unacceptable, and so the content of academy training was "dumbed down." When this reduced the failure rate only slightly, the content was further dumbed down, and then dumbed down again. This process of successive adjustments ultimately "solved" the flunk-out problem.

However, the police officers being produced were incompetent. Large numbers of murder indictments had to be dismissed because the reports written by the officers on the scene were unintelligible, due to the low literacy levels. The solution rate for murder cases, formerly one of the highest in the USA, declined precipitously to one of the lowest. Firearms accidents soared because officers did not know how to use their sidearms properly. Complaints of police abuse and incompetence from citizens soared. In addition, crime on the police force became quite common. For example, a group of police officers was found to be selling handguns previously confiscated from criminals *back to criminals*! These changes and others are described by Carlson (1993a, 1993b).

In this example, unlike the US Steel example, *two* things are happening. First, people low in intelligence are being hired, resulting in plummeting job performance. Second, criminals are being hired because there was no background investigation to ensure that they were not, and the result was crime on the police force.

Employment in the federal government

We now turn to a more positive example – or at least a less negative one. For many jobs in the federal government, people can either be hired from the outside using a GMA test or they can be promoted from within. When they are promoted from within, GMA tests are usually not used – although they sometimes are. Instead, people are evaluated based on records of their education and training and on appraisals by their supervisors of their performance in their present jobs. These procedures do have some validity but would not be expected to be as valid as GMA-based hiring.

So we can ask the following question. After people have been on the job some time, is the job performance higher for those initially selected using a GMA test? Government researchers at OPM addressed this question in a detailed study of three representative mid-level government jobs: IRS auditor, social security claims examiner, and customs inspector. In each of these jobs, people hired both ways had been on the job from five to eight years.

The measure of job performance was unusually good: it was the sum of a hands-on work sample test, a job knowledge test, and supervisory ratings of job performance.

In all three jobs, those selected years earlier using GMA tests had higher job performance. The average job performance of the non-GMA-selected employees was at the 50th percentile, while that of the GMA-selected employees was at the 70th percentile. This is a large difference. If this difference is projected over the federal workforce as a whole, it amounts to *billions* of dollars per year in increased output (Schmidt, Hunter, Outerbridge, and Trattner, 1986). We can also look at this another way. Americans expect their federal government to perform a wide variety of socially important tasks (e.g. administer the social security program, protect homeland security, run the federal tax system fairly and accurately, catch people who commit federal crimes, etc.). To the extent that the federal government hires less competent people, these jobs are done less well. As shown in this research, failure to select on GMA results in the hiring of less competent people and produces lower job performance.

This study was a reasonably controlled quasi-experiment. During the study, the researchers did not know which employees had initially been selected using a GMA measure and which had not. The only relevant difference between the two groups of workers was the method by which they had been hired. This study provides strong evidence that GMA-based hiring pays off in higher job performance.

The Philip Morris plant in Cabarrus County, North Carolina

The US Employment Service began a new nationwide program of employment testing, operated through state employment offices, in the early 1980s. Like its earlier program, it was based on the General Aptitude Test Battery (GATB). One of the three abilities measured in that program was GMA (the other two were general perceptual ability and general psychomotor ability). This new program was based on the methods of meta-analysis or validity generalization that were mentioned at the beginning of this chapter.

The large Philip Morris plant in Cabarrus County, North Carolina, was one of the first employers to subscribe to this testing program. They signed an agreement under which the state employment service tested and referred the higher scoring applicants to Philip Morris for possible hire. For the jobs at Philip Morris, most of the weight was placed on GMA in determining who was hired.

The human resources department at Philip Morris decided to conduct a study to compare the performance of GATB–GMA-selected workers and workers hired without use of the test. They found that the GMA-selected workers were superior across a variety of performance measures. For example, there was a 35% gain in output. The GMA-selected workers learned 8% more skills during job training, had 25% fewer operator failures and 58% fewer disciplinary actions. The incidence of unsafe job behaviors was 35% less and the reduction in work days lost to accidents was *82%*.

These are large differences. The Philip Morris personnel researchers, Dennis Warmke and William Van Arnam, noted the employment interview used might have contributed somewhat to the performance superiority of these workers. However, they stated that because it was the GMA test that screened out most of the applicants who were not hired,

the GMA test was the dominant influence producing the performance improvements. This research is described in McKinney (1984).

Conclusion

Higher intelligence leads to better job performance on all jobs, and the increases in job performance resulting from hiring on GMA have high economic value for organizations. Higher intelligence causes higher job performance primarily because it causes people to learn job knowledge faster and to learn more of it. However, intelligence is also used directly on the job to solve performance-related problems, independent of prior job knowledge. The primary requirement that an organization must meet to make GMA-based hiring work well is the ability to attract job applicants and to retain them once they are hired. Despite beliefs to the contrary, hiring on job experience is inferior to hiring on GMA. Although GMA is the most important determinant of job performance, it is not the only determinant. Therefore, firms should use other valid procedures along with GMA. Finally, we have seen four concrete, graphic, real world examples of the impact of GMA on job performance.

References

Brody, N. (1992). *Intelligence* (2nd edition). San Diego, CA: Academic Press Inc.

Brown, K., Le, H., and Schmidt, F. L. (2006). Specific aptitude theory revisited: Is there incremental validity for training performance? *International Journal of Selection and Assessment,* 14, 87–100.

Carlson, T. (1993a). D.C. blues: The rap sheet on the Washington police. *Policy Review*, Winter, 27–33.

Carlson, T. (1993b). Washington's inept police force. *The Wall Street Journal*, November 3.

Coward, W. M., and Sackett, P. R. (1990). Linearity of ability–performance relationships: A re-confirmation. *Journal of Applied Psychology*, 75, 295–300.

Cooper-Hakim, A., and Viswesvaran, C. (2005). The construct of work commitment: Testing an integrative framework. *Psychological Bulletin*, 131, 241–259.

Dalal, R. S. (2005). A meta-analysis of the relationship between organizational citizenship behavior and counterproductive work behavior. *Journal of Applied Psychology,* 90, 1241–1235.

Gottfredson, L. S. (1996). Why g matters: The complexity of everyday life. *Intelligence*, 24, 79–132.

Hausknecht, J. P., Day, D. V., and Thomas, S. C. (2004). Applicant reactions to selection procedures: An updated model and meta-analysis. *Personnel Psychology*, 57, 639–683.

Herrnstein, R. J., and Murray, C. (1994). *The Bell Curve: Intelligence and Class Structure in American Life.* New York: The Free Press.

Huffcutt, A. I., Roth, P. L., and McDaniel, M. A. (1996). A meta-analytic investigation of cognitive ability in employment interview evaluations: Moderating characteristics and implications for incremental validity. *Journal of Applied Psychology*, 81, 459–473.

Hunter, J. E. (1986). Cognitive ability, cognitive aptitudes, job knowledge, and job performance. *Journal of Vocational Behavior*, 29, 340–362.

Hunter, J. E., and Hunter, R. F. (1984). Validity and utility of alternate predictors of job performance. *Psychological Bulletin*, 96, 72–98.

Hunter, J. E., and Schmidt, F. L. (1996). Intelligence and job performance: Economic and social implications. *Psychology, Public Policy, and Law*, 2, 447–472.

Hunter, J. E., and Schmidt, F. L. (2004). *Methods of Meta-analysis: Correcting Error and Bias in Research Findings*. Thousand Oaks, CA: Sage.

Hunter, J. E., Schmidt, F. L., and Judiesch, M. K. (1990). Individual differences in output variability as a function of job complexity. *Journal of Applied Psychology*, 75, 28–42.

Hunter, J. E., Schmidt, F. L., and Le, H. (2006). Implications of direct and indirect range restriction for meta-analysis methods and findings. *Journal of Applied Psychology*, 91, 594–612.

Hunter, J. E., Schmidt, F. L., Rauschenberger, J. M., and Jayne, M. (2000). Intelligence, motivation, and job performance. Chapter in C. L. Cooper and E. A. Locke (eds), *I/O Psychology: What We Know about Theory and Practice*. Blackwell Publishers, pp. 278–303.

Jensen, A. R. (1998). *The g Factor: The Science of Mental Ability*. Westport, CT: Praeger.

Judge, T. A., Thoresen, C. J., Bono, J. E., and Patton, G. K. (2001). Another look at the relationship between job satisfaction and job performance. *Psychological Bulletin*, 127, 376–407.

McHenry, J. J., Hough, L. M., Toquam, J. L., Hanson, M. L., and Ashworth, S. (1990). Project A validity results: The relationship between predictor and criterion domains. *Personnel Psychology*, 43, 335–354.

McKinney, M. W. (1984). Final report: Validity generalization pilot study. Southern Test Development Field Center, US Employment Service, Raleigh, North Carolina.

Oh, I., Schmidt, F. L., and Shaffer, J. (2008). Comprehensive updating of the validity and incremental validity of selection procedures based on 95 years of research. Working paper, Department of Management, Tippie College of Business, University of Iowa, Iowa City, Iowa.

Olea, M. M., and Ree, M. J. (1994). Predicting pilot and navigator criteria: Not much more than g. *Journal of Applied Psychology*, 79, 845–851.

Ones, D. S., Viswesvaran, C., and Schmidt, F. L. (1993). Comprehensive meta-analysis of integrity test validities: Findings and implications for personnel selection and theories of job performance. *Journal of Applied Psychology*, 78, 679–703.

Orr, J. M., Sackett, P. R., and Mercer, M. (1989). The role of prescribed and nonprescribed behaviors in estimating the dollar value of performance. *Journal of Applied Psychology*, 74, 34–40.

Ree, M. J., and Earles, J. A. (1991). Predicting training success: Not much more than g. *Personnel Psychology*, 44, 321–332.

Ree, M. J., and Earles, J. A. (1992). Intelligence is the best predictor of job performance. *Current Directions in Psychological Science*, 1, 86–89.

Ree, M. J., Earles, J. A., and Teachout, M. (1994). Predicting job performance: Not much more than g. *Journal of Applied Psychology*, 79, 518–524.

Roth, P. I., Bobko, P., Switzer, F. S., and Dean, M. A. (2001). Prior selection causes biased estimates of standardized ethnic group differences: Simulation and analysis. *Personnel Psychology*, 54, 591–617.

Rotundo, M., and Sackett, P. R. (2002). The relative importance of task, citizenship, and counterproductive performance to global ratings of job performance: A policy-capturing approach. *Journal of Applied Psychology*, 87, 66–80.

Salgado, J. F., Anderson, N., Moscoso, S., Bertua, C., and de Fruyt, F. (2003b). International validity generalization of GMA and cognitive abilities: A European community meta-analysis. *Personnel Psychology*, 56, 573–605.

Salgado, J. F., Anderson, N., Moscoso, S., Bertua, C., de Fruyt, F., and Rolland, J. P. (2003a). A meta-analytic study of general mental ability validity for different occupations in the European community. *Journal of Applied Psychology*, 88, 1068–1081.

Schmidt, F. L. (1988). The problem of group differences in ability scores in employment selection. *Journal of Vocational Behavior*, 33, 272–292.

Schmidt, F. L. (1992). What do data really mean? Research findings, meta analysis, and cumulative knowledge in psychology. *American Psychologist*, 47, 1173–1181.

Schmidt, F. L. (2002). The role of general cognitive ability in job performance: Why there cannot be a debate. *Human Performance*, 15, 187–210.

Schmidt, F. L., and Hunter, J. E. (1981). Employment testing: Old theories and new research findings. *American Psychologist*, 36, 1128–1137.

Schmidt, F. L., and Hunter, J. E. (1992). Development of causal models of processes determining job performance. *Current Directions in Psychological Science*, 1, 89–92.

Schmidt, F. L., and Hunter, J. E. (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.

Schmidt, F. L., and Hunter, J. E. (2004). General mental ability in the world of work: Occupational attainment and job performance. *Journal of Personality and Social Psychology*, 86, 162–174.

Schmidt, F. L., Hunter, J. E., and Outerbridge, A. N. (1986). The impact of job experience and ability on job knowledge, work sample performance, and supervisory ratings of job performance. *Journal of Applied Psychology*, 71, 432–439.

Schmidt, F. L., Hunter, J. E., Outerbridge, A. N., and Goff, S. (1988). The joint relation of experience and ability with job performance: A test of three hypotheses. *Journal of Applied Psychology*, 73, 46–57.

Schmidt, F. L., Hunter, J. E., Outerbridge, A. N., and Trattner, M. H. (1986). The economic impact of job selection methods on the size, productivity, and payroll costs of the Federal work-force: An empirical demonstration. *Personnel Psychology*, 39, 1–29.

Schmidt, F. L., Hunter, J. E., and Pearlman, K. (1980). Task difference and validity of aptitude tests in selection: A red herring. *Journal of Applied Psychology*, 66, 166–185.

Schmidt, F. L., Oh, I., and Le, H. (2006). Increasing the accuracy of corrections for range restriction: Implications for selection procedure validity and other research findings. *Personnel Psychology*, 59, 281–305.

Schmidt, F. L., Shaffer, J. A., and Oh, I. (2008). Increased accuracy for range restriction corrections: Implications for the role of personality and general mental ability in job and training performance. *Personnel Psychology*, 61, 827–868.

Schmidt, F. L., Ones, D. S., and Hunter, J. E. (1992). Personnel selection. *Annual Review of Psychology*, 43, 627–670.

Wigdor, A. K., and Garner W. R. (eds) (1982). *Ability Testing: Uses, Consequences, and Controversies* (Report of the National Research Council Committee on Ability Testing). Washington, DC: National Academy of Sciences Press.

Wilk, S. L., Desmarais, L. B., and Sackett, P. R. (1995). Gravitation to jobs commensurate with ability: Longitudinal and cross-sectional tests. *Journal of Applied Psychology*, 80, 79–85.

Wilk, S. L., and Sackett, P. R. (1996). Longitudinal analysis of ability–job complexity fit and job change. *Personnel Psychology*, 49, 937–967.

Exercises

Hiring office workers

You are the human resources director at a large firm and you are faced with designing a system for hiring office workers. An office manager comes to you and says the firm should not use written GMA tests because of the danger of law suits. He says he knows GMA is important to job performance but maintains that you can use "GMA-loaded" interviews to measure GMA and thus get the benefit of using GMA without leaving a "paper trail" of test scores that could stimulate a law suit. Respond to this manager based on what you learned from this chapter. What would you tell him? What is the foundation for your response?

Educating the CEO

You are the human resources director in your organization. The CEO calls you to her office for a meeting and tells that she knows from 35 years of experience in dealing with people that the key determinant of high job performance is personal values and sense of responsibility. She says she would like to have all hiring in the company done using measures of values and sense of responsibility. Based on what you learned in this chapter, what would you tell her? What is the basis for the position you are taking?