PeopleHawk

THE SCIENCE BEHIND PEOPLEHAWK:

Valid assessments can more than double the chances **of success in candidate selection**

Predictive efficiency of cognitive assessments

Predictive efficiency

Predictive efficiency refers to the degree of success at reducing the error in predicting an individual's scores on a criterion, such as income level. The index of forecasting efficiency is one such measure, as is the percentage of criterion variance 'explained' by (shared in common with) the predictor. The latter, which is a squared correlation, is used in evaluating how well theoretical models "fit" the data. It also seems to be the measure of predictive importance favoured by detractors of *g* (e.g., Gould, 1994), perhaps because it yields the smallest-appearing estimates of importance.

Educators, selection psychologists, policymakers and others who work in the realm of practical affairs are not concerned with fully explaining particular outcomes. Instead, they usually want to know how much 'return on investment' for a set criterion they will get by changing some input by a certain amount. However, even predictors with low predictive efficiency can yield huge effects over extended periods of time. Similarly, to take a hypothetical example, a vocational counsellor would find it more useful to know that a particular counselee's SAT score gives them only a 60% chance of achieving a C average or better at the college he or she wishes to attend, than to know that SAT scores account for 30% of the variation in grades in that institution.

Prediction of an individual's odds of success

Life has unknown ups and downs so, like gamblers, we all have to play the odds, as far as we understand them. Predictive validities can be used to show how our odds of success or failure (being admitted to university, successfully graduating) rise or fall depending on our traits (intelligence, grades) or circumstances (parents' income or education).

In particular, correlations can be used to calculate the expected rates of criterion success for any particular range of scores on the predictor. Personnel professionals and college admissions officers, for example, use such tables to set minimum cut-off scores in hiring workers and admitting students. They cannot know for sure which particular applicants will succeed, but they can be fairly certain about the proportions who will and how those proportions rise or fall with predictive validity.

For example, consider a situation where 60% of individuals succeed. To be specific, about 60% of the working population is above the intelligence level (IQ 100) required for adequate performance as a bank teller, for example. If a bank hired randomly, the odds of successful performance would be 1.5: 1 (60:40, in Jensen, 1980, Taylor-Russell tables on p. 307). If the bank selected applicants based on a test with a predictive validity of only .3 (and hired half the applicants), the odds of success would rise to over 2: 1 (69:31). Switching to a cogitive test with a validity of .45 would raise the odds to 3:l (74:26), thus doubling the original odds.

The higher the correlation between predictor and criterion, the more sharply the odds of success will diverge for the same two scores (say, IQ 85 vs. 115). Odds are especially useful in assessing the life chances of individuals at different levels of the IQ continuum. They often differ by multiples far greater than 3:l.

Prediction of groups’ average performance levels

Odds deal with success versus failure, that is, with either meeting or falling below some minimum performance level. But many institutions are more interested in predicting average performance levels above some minimum and how much they might change under different conditions (new selection procedures). This is the realm of “effect sizes” and “utility” analysis (respectively, Lubinski & Humphreys, 1997; Boudreau, 1991).

Such calculations are of particular importance for social policy because it usually concerns itself with gradual shifts in population outcomes, favourable or unfavourable. Predictive validity need not be large for effect sizes to be substantial in both human and economic terms (Lubinski & Humphreys, 1997). Utility analyses of alternative selection procedures make the same point. The percentage gain in aggregate worker performance that is to be expected when switching from one selection procedure to another is a direct function of their respective predictive validities. For example, a predictive validity of .4 (or .5) means that using a battery of cognitive assessments can achieve up to 40% (or 50%) of the gains that would be possible by using a perfectly valid test (predictive validity of 1 .0) compared with random selection (predictive validity of zero). Even much smaller increases in validity (say, from .2 to .4) often translate into thousands of dollars per hire per year (see Boudreau, 1991, for an extended discussion of utility analyses in personnel selection).

Relative importance

The importance of a predictor is often also judged by comparing its validity with that of other predictors. The predictor with the tighter link to important outcomes is the more important. More formally, a set of such correlations can be mathematically modelled to estimate the independent effects of each predictor. Path analyses that do so have evidenced that cognitive tests are the single best predictor of job performance, better even than job trials, biographical data, reference checks, education, interviews and even grades. See Table 1. Note that personality profiling was not one of the predictors evaluated by Hunter and Hunter, 1984. It was not until the decades that followed that personality profiling was studied for such purposes.

TABLE 1

Mean Validities and Standard Deviations of Various Predictors for Entry-Level Jobs for Which Training Will Occur After Hiring

Predictor	Mean validity	SD	No. correlation	No. subjects
Ability Composite*	.53	.15	425	32,124
Job tryout	.44		20	
Biograhpical inventory	.37	.10	12	4,429
Reference check	.26	.09	10	5,389
Experience	.18		425	32,124
Interview	.14	.05	10	2,694
Training and experience ratings	.13		65	
Academic achievement	.11	.00	11	1,089
Education	.10		425	32,124
Interest	.10	.11	3	1,789
Age	.01		425	32,124

Source. Hunter and Hunter, 1984
*Cognitive Assessments
Note: Personality profiling was not condisered within this study.

Conclusion

Finally, it is important to realise that the same validity coefficient may be of little use for one purpose (accurately predicting an individual's behaviour) and yet quite powerful for another (predicting rates of behaviour in different groups). Also, a number of even small factors that might be inconsequential individually can cumulate for enormous impact over an individual's lifetime. An analogy is the small but inexorable odds favouring the house in casino gambling (Gordon et al., 1988).

REFERENCES

Boudreau, J.W. (1991). Utility analysis for decisions in human resource management. In M.D. Dunnette & L.M. Hough (Eds.), *Handbook of industrial and organizational psychology* (Vol. 2). Palo Alto: Consulting Psychologists Press.

Gordon, R.A., Lewis, M.A., & Quigley, A.M. (1988). Can we count on muddling through the *g* crisis in employment? *Journal of Vocational Behavior*, *33*, 424-451.

Gould, S.J. (1994, November 28). Curveball [*Review of The bell curve: Intelligence and class structure in American life*]. *New Yorker*. 139-149.

Jensen. A.R. (1980). *Bias in mental testing*. New York: Free Press.

Lubinski, D., & Humphreys, L.G. (1997). Incorporating general intelligence into epidemiology and the social sciences. *Intelligence*, *24*, 159-201.