Hiring Method Evidence Chart

A one-page reference on predictive validity by selection method, from the 2022 Sackett re-analysis, with how to read it and what it does not say.

One page to put in front of anyone who asks why you are changing how interviews are run.

Predictive validity by selection method

METHOD	VALIDITY (R)
Structured interviews	.42
Job knowledge tests	.40
Empirically keyed biodata	.38
Work sample tests	.33
Cognitive ability tests	.31
Integrity tests	.31
Assessment centres	.29
Interests	.24
Conscientiousness measures	.19
Unstructured interviews	.19
Years of job experience	.07

Source: Sackett, Zhang, Berry and Lievens (2022), *Journal of Applied Psychology*. These are revised mean operational validities, corrected for a long-standing statistical error that had inflated earlier figures by roughly .10 to .20.

How to read a validity coefficient

It is a correlation between what a method measures and later job performance, on a scale from 0 to 1. It is not an accuracy percentage. A validity of .42 does not mean the method is right 42% of the time; it means scores and performance move together reasonably strongly.

Rough reference points: above .30 is useful, above .40 is strong, and nothing in personnel selection reaches .70.

Three things this chart does not say

That the average is your number. These are means across many studies. Structured interviews have an 80% credibility interval of roughly .18 to .66, which is an enormous range. A structured interview built from a real analysis of the job, with trained interviewers and written criteria, lives at the top of it. One that is structured in name only lives at the bottom. The method matters less than the execution.

That you should pick one. The methods are not alternatives. A role profile, an assessment and a structured interview cover different ground, and combining them predicts better than any single row here.

That validity is the only criterion. Cognitive ability tests have respectable validity and the largest subgroup differences of any common method. Any serious selection decision weighs predictive power against adverse impact, candidate experience and legal defensibility together.

Why the ranking changed

For two decades the reference point was Schmidt and Hunter's 1998 summary, which put cognitive ability at the top and was quoted in almost every textbook. Sackett and colleagues showed that the standard correction for range restriction had been applied in a way that systematically overstated validities. Correcting it lowered most estimates and reordered the top: structured interviews, a method most organizations already run, moved above cognitive ability tests.

The practical reading is encouraging. The strongest tool available to most teams is one they already use every day. It just has to be built properly.

A note on vendor figures

You will see higher numbers quoted for commercial processes. GoodJob, for instance, reports .65 to .75 for its full PATH behavioral hiring process, which combines a role profile, an assessment, structured interviews and onboarding.

That is not the same kind of number as the table above, and it should not be read as a row in it. The table reports independent meta-analyses of single methods. A figure like .65 to .75 comes from a practitioner's own data and describes a whole multi-step process. Both can be informative. Only one has been through independent replication, and anyone quoting either should say which they are holding.