

Explaining predictive validity

Wim Bloemers, assessmentanarchy.com

A company wants to select 50 new employees, and attracts 100 candidates. Assume, after pre-selection, a base rate of qualified applicants of .50, so 50 of the 100 candidates are indeed qualified (valid positives). How good is your test with a predictive validity of .40 in this situation?

	situation 1 predict. val. = 1		situation 2 predict. val. = 0		situation 3 predict. val.= .40	
	qualified		qualified		qualified	
	yes	no	yes	no	yes	no
hired	50		25	25	35	15
rejected		50	25	25	15	35

With a predictive validity of 1.0 all 50 valid positives are selected, 50 unqualified candidates are rejected by the test.

With no predictive validity, random distribution occurs, so all 4 cells contain 25 candidates. Notice that thus with a test with a zero validity, still 25 qualified candidates are selected, due to a base rate of .50. That's why preselection/attracting the right people is so important. Your test validity playground focuses on the 25 'extra' qualified candidates above chance level that are among the group of 100 candidates.

The third situation is the realistic one. With a validity of .40, 10 extra qualified candidates are selected above chance (.40*(50-25)). You can also see that predictive validity is direct proportional to the number of hired valid positives. E.g with a validity of .80, you would attract .80*(50-25)= 20 extra qualified candidates above chance, thus 45 in total. That's twice as much compared to a validity of .40.

Source: W.K.B. Hofstee; Rationele personeelselectie; er wordt te weinig getest. (Rational personnel selection: too little tests are being used) In: *Harvard Holland Review*, 7, pp. 35-41