



Selection testing of prospective employees is well accepted in many industries as an aid in hiring the most productive workers. Most farm workers, however, are hired on a first-come, first-hired basis, and there are few reports of testing in agriculture.

In a limited study in 1986, a test was used successfully to predict tomato workers' performance on the job (*California Agriculture*, May-June 1987), but the validity of employment tests in other areas of agriculture remains largely unstudied.

On the hypothesis that job-sample tests of tasks such as pruning, budding, and picking could be used to predict the performance of vineyard workers, I undertook a study in three San Joaquin Valley vineyards in 1986-87. The specific purpose of the study was to determine how well a job-sample test (when workers know they are being tested) can predict performance on the job (when workers do not know their work is being measured).

Methods

The study involved four separate groups in three vineyards. The farms were selected because they paid workers on a piece-rate basis, employed 40 or more

Predicting vineyard pruner performance

Gregory Encina Billikopf

A job-sample test showed high correlation with work performance



Workers marked the beginning and ending of test trials with colored clothespins. Performance was measured by the number of vines pruned in a 46-minute period.

workers, and had cooperated on other research projects. There were approximately 115 workers on farm 1, 45 on farm 2, 116 on farm 3A, and 67 on farm 3B. To keep working conditions comparable, the study included only cordon-pruned grape varieties. Most vines on the same farm that were the same variety, age, and spacing were considered sufficiently similar to constitute uniform conditions.

The study, which took place during the winter dormant season when pruning is normally done, used both "predictive" and "concurrent" analyses. In a predictive study, job applicants' test results are compared with actual on-the-job performance later. In a concurrent study, the work of present employees during a trial period is compared with their regular on-the-job performance. In both cases, workers would be aware they were being tested, but would not know when their on-the-job performance was being measured.

Workers were asked to prune vines during two identical 46-minute periods (here labeled tests 1 and 2). Workers understood they needed to prune as fast as possible and still maintain quality. Performance was measured in numbers of vines pruned per 46-minute period for each worker.

In the concurrent studies, worker participation was voluntary, but all workers participated. Test results from the predictive studies (farms 2 and 3A) were used to make hiring decisions.

Work performance information was obtained from each farm's payroll records from two randomly selected days and two randomly selected grape varieties, after the pruning season was over.

Results were evaluated statistically by correlation analysis to determine the association or closeness of the relationships (as opposed to attempting to show a cause-and-effect relationship). Four validity coefficients were established for each study by relating each test 1 and test 2 to each performance 1 and performance 2.

Test consistency

Consistency between test 1 and test 2 results was high (average correlation $r=.83$). One might expect that workers would be more motivated to perform well in predictive than in concurrent test conditions, since performance on the test could mean obtaining or losing a job opportunity. This did not seem to be the case.

Informal observation shed some light on motivation among the various worker groups. Workers on all three farms (in both predictive and concurrent tests) seemed very competitive, and several tried but were not permitted to get a head start. Some pruners told others to slow down, often using the fear of a speed-up

(where farmers might reduce the piece rate) as a reason for their colleagues to slow down. Such comments as *No me dejes atrás; no te apures tanto* ("Don't leave me behind; don't hurry so much"), and *¿Quieres podar por menos?* ("Do you want to prune for less?") were common.

Notwithstanding the calls to slow down, most workers seemed motivated to do their best. It was anticipated that workers might be concerned about speed-ups in the concurrent studies, but it was not anticipated that this concern would be just as prevalent in the predictive ones.

Even more surprising was that, in the predictive study on farm 2, one worker stopped working to help another. This lack of concern over his own results could perhaps be explained by the fact that some groups of workers prefer to work together, or need to because of transportation. In such circumstances, getting the job when a friend or spouse did not would be useless.

There was a large range of scores within each study (table 1). For example, in the study on farm 3A, test 2, workers pruned from 3 to 24 vines in the test period. On farm 1, test 1, the range was 12 to 28 vines.

Work performance consistency

Consistency of work performance was high except on farm 2 (average correlation, excluding farm 2, $r=.61$). Possible explanations for farm 2 results are that (1) the sample size was small; (2) workers might not have been motivated to do their best even if they were being paid on a piece-rate basis; and (3) the farmer, like some who pay on a piece rate rather than by the hour, might not have been careful about documenting exact working hours.

A review of the results with the manager of farm 2 suggested that the third explanation was the case for them. The other farms, on the other hand, were very care-

ful to document exact starting and finishing times for workers. In addition to careless time measurements, some supervisors credit workers, for a given day, for partially finished vines, so that the total number of vines pruned in a day might actually have been more or less than that recorded. Sloppy recordkeeping could account for major differences in scores from day to day.

Validities

Validities or correlations between test and on-the-job performance measurements ranged from significant to highly significant for most farmwide results. Only farm 2 showed no significance between the two measurements. This finding does not seem to contradict the notion that very inconsistent work performance measurements would make a test invalid, no matter how reliable it was. Worker productivity consistency for farm 2 was so low that lack of validity for this group was expected.

The only other farmwide results in which correlation between test and on-the-job performance was extremely low were found in the low test 1, performance 2, results of farm 3B (concurrent study). Individual crew validities for this farm were not as low as the farmwide results. They ranged from 0.46 (ns) to 0.63 ($p<.01$). It is possible that performance 2 involved vineyard blocks of different difficulty levels. This could explain why validity coefficients were higher (although not necessarily significant) for each crew than for the whole farm. Other than the exceptions described, the validity coefficients in this study are among the highest reported in the employment testing field.

Conclusions

There seems to be great potential for the use of work-sample employment tests in agriculture. Nevertheless, the need to establish the validity of the relationship between test and on-the-job performance at the individual farm remains.

There seems little doubt that both concurrent and predictive tests can predict performance, but employers cannot use the test and assume that it will always work. The test on farm 2 turned out to be totally invalid. High consistency of test and of on-the-job performance is essential for obtaining a valid test.

Gregory Encina Billikopf is Personnel Management Farm Advisor, University of California Cooperative Extension, Stanislaus County. A more detailed report is available from the author at 733 County Center III Court, Modesto, CA 95355. The author is grateful to Dr. James Wakefield, California State University, Stanislaus, for guidance in this study.

TABLE 1. Work scores in vineyard pruning tests

Farm & test*	No. of workers	No. of vines pruned		Average correlation (validity)†
		Test sample	On the job	
1 (Con)	115			.68 ***
Test 1		12 - 28	10 - 37	
Test 2		14 - 30	20 - 50	
2 (Pre)	45			.14 (ns)
Test 1		6 - 24	23 - 47	
Test 2		7 - 26	13 - 36	
3A (Pre)	116			.57 **
Test 1		5 - 26	20 - 46	
Test 2		3 - 24	16 - 42	
3B (Con)	67			.41 **
Test 1		12 - 38	20 - 49	
Test 2		10 - 40	16 - 42	

* Con = concurrent. Pre = predictive.

† Average correlation (validity) between test and job.

** $f<.01$. *** $f<.001$.