Validity Study

Cognitive Ability, Personality, and Low Fidelity Simulation Measures in Predicting Training Performance among Customer Service Representatives

Saul Fine and Shlomo Dover
CareerHarmony, Ltd.

A predictor battery of cognitive ability, personality, and low fidelity simulation measures was completed by a sample of 193 U.S.-based customer service representatives. The results showed cognitive ability and simulation scores to be significantly correlated with final training course scores, whereas personality scales were not. Following simultaneous regression analysis, a unit-weighted composite was suggested for the test battery composed of all scales. This composite was significantly correlated (r = .30, uncorrected) with training performance.

Sample

N	193 customer service representatives
Gender	63% were women, and 37% were men
Race	65% were White, 18% were African American, 14.5% Hispanic/Latino, 2% were Asian, and .5% were Native American Indian
Age	Mean = 31.98, standard deviation = 10.89

Location

The data for this study were collected from a large call center organization in the United States, in 2003.

Predictor Information

The predictors in this study consisted of four web-based assessment tools. Two cognitive tests (Numerical Reasoning and Verbal Reasoning), one low fidelity simulation test (Systems Management), and one personality-based inventory (Interpersonal Coping Style) were administered in a proctored environment. Numerical Reasoning is primarily a power test consisting of 30 items. Each item presents a series of numbers that are organized according to a specific rule, and the test taker must enter the number that best continues the series. Verbal Reasoning is primarily a power test consisting of 25 items. Each item consists of a pair of verbal analogies, and the test taker must choose the closest matching analogy from a set of multiple-choice answers. The System Management test presents a simulated operating system where the test taker is required to respond to a series of 20 display messages in accordance with the operating system's user manual. The Interpersonal Coping Style Inventory includes 54 items. The items consist of two multiple choice

statements, each describing an interpersonal behavior, and each representing one of the three ipsative scales measured: Avoiding, Competing, and Compromising. Table 1 contains the reliabilities and intercorrelations of the four tests' scales.

Criterion Information

The criterion utilized was overall training course scores. Upon commencement of employment, participants went through a 2 to 3 week training course, preparing them for work in their respective project assignments. While training course tracts differed slightly per project assignment, all training courses included classroom training followed by supervised on-the-job experience. In each case, a single score, ranging from 0 to 100, was given for overall training performance upon completion of the course. To avoid inter-course scoring differences, training scores were standardized within project tracts (from a total of 3 tracts), and then aggregated.

Validity Information

Table 2 presents the correlation of the predictor measures with the criterion for this concurrent validity study. The aptitude and simulation tests were all moderately correlated with the criterion, although none of the personality scales were significantly correlated with the criterion. It is likely that this is related to the possibility that the training course, as opposed to actual performance, is more focused on learning procedures than on interpersonal interactions.

Simultaneous multiple regression analysis was run on the predictor measures against the criterion, yielding an R^2 of .098 (.067 adjusted R^2), $F(6,177) = 3.19$, $p < .005$. None of the beta weights were significant. With the non-significant bivariate correlations removed (i.e., the personality scales), the R was reduced from .31 to .28, $F(3,180) = 5.17$, $p < .002$. This difference was not significant ($p < .05$), indicating that the personality scales did not significantly increment the explained variance above that by the simulation and aptitude tests alone.

However, since in larger sample sizes small differences in validity may become significant and have an influence on the utility of the selection system, it was proposed to include all scales with unit weights (with personality scales as negative weights) into the final composite for the prediction of training performance. This produced a final predictor-criterion correlation of .30 ($p < .001$).

Table 1
Predictor reliabilities and intercorrelations

Predictor	(1)	(2)	(3)	(4)	(5)	(6)
1. Numerical reasoning	(.88)	.43**	.29**	- .08	.00	- .11
2. Verbal reasoning		(.83)	.38**	- .11	.13	.24**
3. Systems management			(.82)	- .08	.13	- .16
4. Competing				(.67)	- .22**	- .20**
5. Compromising					(.67)	- .33**
6. Avoiding						(.59)

Note: Reliability estimates collected from a larger sample (N=382) are in parentheses.

*p < .05; **p < .01

Adverse Impact

A slight positive effect size of .31 was found for the composite on non-Whites, $t(184) = 4.60$, p < .001, .27 on women, $t(184) = 4.0$, p < .001, and .02 for employees over the age of 45, $t(183) = .21$, p > .83. This indicates that Whites and men scored marginally higher than non-Whites and women, and that age had no effect. While statistically significant, the magnitudes of these differences are not likely to have substantial operational effects. Nevertheless, operational usage of this selection composite should monitor compliance with the Uniform Guidelines on Employee Selection Procedures (1978) (e.g., the 80% rule). Job applicants' demographic data were not available to analyze EEO compliance within the present incumbent sample.

Adverse impact may also be measured by the differential predictability of the criterion between groups. No such significant differences were found for race (White: .26, non-White: .31), age (<45: .31, ≥45: .20), or gender (males: .33, females: .29). All validity coefficients for these subgroups were significant ($p < .05$), with the exception of the 'older than 45' group.

Table 2
Correlations between test scores and training performance

Predictor	Correlation
1. Numerical reasoning	.22**
2. Verbal reasoning	.21**
3. Systems management	.19**
4. Competing	- .02
5. Compromising	- .06
6. Avoiding	- .10

* $p < .05$; ** $p < .01$

References

Uniform Guidelines on Employee Selection Procedures (1978); 43 FR 38295.

Author Notes

Please address correspondence to:

Saul Fine, PhD
CareerHarmony
P.O. Box 7616, Petach Tikva 49170, Israel.
Phone: ++972-3-924-6569
Fax: ++972-3-924-6567
Email: saulf@careerharmony.com