

APPLICATIONS

Evaluation of Two New Assessment Methods for Selecting Telephone Customer Service Representatives

The Phonemini Company is a distributor of men's and women's casual clothing. It sells exclusively through its merchandise catalog, which is published four times per year to coincide with seasonal changes in customers' apparel tastes. Customers may order merchandise from the catalog via mail or over the phone. Currently, 70% of orders are phone orders, and the organization expects this to increase to 85% within the next few years.

The success of the organization is obviously very dependent on the success of the telephone ordering system and the customer service representatives (CSRs) who staff the system. There are currently 185 CSRs; that number should increase to about 225 CSRs to handle the anticipated growth in phone order sales. Though the CSRs are trained to use standardized methods and procedures for handling phone orders, there are still seemingly large differences among them in their job performance. The CSRs' performance is routinely measured in terms of error rate, speed of order taking, and customer complaints. The top 25% and lowest 25% of performers on each of these measures differ by a factor of at least three (e.g., the error rate of the bottom group is three times as high as that of the top group). Strategically, the organization knows that it could substantially enhance CSR performance (and ultimately sales) if it could improve its staffing "batting average" by more accurately identifying and hiring new CSRs who are likely to be top performers.

The current staffing system for CSRs is straightforward. Applicants are recruited through a combination of employee referrals and newspaper ads. Because turnover among CSRs is so high (50% annually), recruitment is a continuous process at the organization. Applicants complete a standard application blank, which asks for information about education and previous work experience. The information is reviewed by the staffing specialist in the HR department. Only obvious misfits are rejected at this point; the others (95%) are asked to have an interview with the specialist. The interview lasts 20–30 minutes, and at the conclusion the applicant is either rejected or offered a job. Due to the tightness of the labor market and the constant presence of vacancies to be filled, 90% of the interviewees receive job offers. Most of those offers (95%) are accepted, and the new hires attend a one-week training program before being placed on the job.

The organization has decided to investigate fully the possibilities of increasing CSR effectiveness through sounder staffing practices. In particular, it is not pleased with its current methods of assessing job applicants; it feels that neither the

application blank nor the interview provides the accurate and in-depth assessment of the KSAOs that are truly needed to be an effective CSR. Consequently, it engaged the services of a consulting firm that offers various methods of KSAO assessment, along with validation and installation services. In cooperation with the HR staffing specialist, the consulting firm conducted the following study for the organization.

A special job analysis led to the identification of several specific KSAOs likely to be necessary for successful performance as a CSR. Three of these (clerical speed, clerical accuracy, and interpersonal skills) were singled out for further consideration because of their seemingly high impact on job performance. Two new methods of assessment, provided by the consulting firm, were chosen for experimentation. The first was a paper-and-pencil clerical test assessing clerical speed and accuracy. It was a 50-item test with a 30-minute time limit. The second was a brief work sample that could be administered as part of the interview process. In the work sample, the applicant must respond to four different phone calls: from a customer irate about an out-of-stock item, from a customer wanting more product information about an item than was provided in the catalog, from a customer who wants to change an order placed yesterday, and from a customer with a routine order to place. Using a 1-5 rating scale, the interviewer rates the applicant on tactfulness (T) and concern for customers (C). The interviewer is provided with a rating manual containing examples of exceptional (5), average (3), and unacceptable (1) responses by the applicant.

A random sample of 50 current CSRs were chosen to participate in the study. At Time 1 they were administered the clerical test and the work sample; performance data were also gathered from company records for error rate (number of errors per 100 orders), speed (number of orders filled per hour), and customer complaints (number of complaints per week). At Time 2, one week later, the clerical test and the work sample were readministered to the CSRs. A member of the consulting firm sat in on all the interviews and served as a second rater of CSRs' performance.

Results for Clerical Test

	Time 1	Time 2
Mean score	31.61	31.22
Standard deviation	4.70	5.11
Coefficient alpha	.85	.86
Test-retest r	-.31**	-.37**
r with error rate	.41**	.39**
r with speed	-.11	-.08
r with complaints	.21	.17
r with work sample (T)	.07	.15
r with work sample (C)		

Results for Work Sample (T)

	Time 1	Time 2
Mean score	3.15	3.11
Standard deviation	.93	1.01
% agreement (raters)	88%	79%
r with work sample (C)	.81**	.77**
r with error rate	-.13	-.12
r with speed	.11	.15
r with complaints	-.37**	-.35**

Results for Work Sample (C)

	Time 1	Time 2
Mean score	2.91	3.07
Standard deviation	.99	1.10
% agreement (raters)	80%	82%
r with work sample (T)	.81**	.77**
r with error rate	-.04	-.11
r with speed	.15	.14
r with complaints	-.40**	-.31**

(Note: ** means that r was significant at $p < .05$)

on the work sample at Time 1 and Time 2. It is expected that the clerical test and work sample will have positive correlations with speed and negative correlations with error rate and customer complaints.

Based on the description of the study and the results above,

1. How do you interpret the reliability results for the clerical test and work sample? Are they favorable enough for Phonemin to consider using them "for keeps" in selecting new job applicants?
2. How do you interpret the validity results for the clerical test and work sample? Are they favorable enough for Phonemin to consider using them "for keeps" in selecting new job applicants?
3. What limitations in the above study should be kept in mind when interpreting the results and deciding whether to use the clerical test and work sample?

Conducting Empirical Validation and Adverse Impact Analysis

Yellow Blaze Candle Shops provides a full line of various types of candles and accessories such as candleholders. Yellow Blaze has 150 shops in shopping malls