Update on U.S. Navy Recruiter Selection Research

Walter C. Borman
Lisa M. Penney
Personnel Decisions Research Institutes, Inc.

Ronald M. Bearden
Navy Personnel Research Studies and Technology

The purpose of the present research was to develop and evaluate the validity of a selection test for the job of Navy recruiter. To accomplish this, we conducted a job analysis of this target job, developed criterion measures for the validation, prepared a personality test tailored to military recruiting, and conducted concurrent and predictive validity studies.

Briefly, U.S. Navy recruiters face an increasingly difficult recruiting environment characterized by lower youth enlistment propensity and, during the 1990s, a very tight labor market. Over the last several years, these conditions have led the Navy to increase its recruiter force by approximately 500 recruiters to more than 5,000. In order to staff the recruiter force, young enlisted personnel must be moved from some critically undermanned jobs into recruiting duty. Because recruiters serve as representatives of the Navy to the community and provide the critical manpower supply, it is important for the Navy to make good decisions about which of these sailors will be effective as recruiters. This was the motivation for the research reported here.

The validation approach chosen for this project was a criterion-related strategy. This is typically accomplished by obtaining the test scores of job applicants and collecting measures of these same individuals' job performance. Test scores are then related to how well individuals perform on the job. Successful validation of this type provides confirmation that use of the selection measures will, in fact, identify the most qualified candidates for the job. A concurrent strategy was used during the first phase of the project, whereas a predictive strategy was employed during the second phase. Specifics of the research, including details of criterion development, predictor preparation, and validation of the predictors are provided in the following sections.

Development of Criterion Measures of Recruiter Job Performance

In a validation study, accurate performance measurement is critical for several reasons. First, in order to demonstrate the utility of the tests for selecting qualified candidates, it is necessary to show that they are, in fact, related to job performance. This can only be achieved if comprehensive, reliable, and valid measures of job performance are available.

Our goal was to develop criteria that would fully capture the Navy recruiter job performance domain. Two measures of recruiter job performance were used in the validation study: recruiter rating scales designed to measure typical performance over time, i.e., what each recruiter actually does to perform effectively or ineffectively; and recruiter production, measuring a critical outcome of job performance.

The recruiter rating scales were developed in a Navy recruiter job analysis (Penney, Borman, Hedge, Abrahams, & Drenth, 2001). The eight behavioral dimensions that emerged were: (1) Locating and Contacting Qualified Prospects; (2) Gaining and Maintaining Rapport; (3) Obtaining Information from Prospects and Making Good Person-Navy Fits; (4) Salesmanship Skills; (5) Establishing and Maintaining Good Relationships in the Community; (6) Providing Knowledgeable and Accurate Information about the Navy; (7) Administrative Skills; and (8) Supporting Other Recruiters and the Command. Subsequent analyses in one of the validation studies resulted in a three factor summary of these eight dimensions: Selling Skills, Human Relations Skills, and Organizing Skills. Validation analyses were done using this three factor system.

Regarding the other criterion measure, typically, production measures focus on the number of recruits or contracts signed in a specified time period, for example, number of recruits signed per month. Production measures are inherently appealing as indicators of recruiter performance. Not only are they "bottom-line" measures, but also they are apparently objective, and they are certainly quantifiable and readily available. Unfortunately, their usefulness as a recruiter performance criterion is limited in several ways.

Among those, production numbers depend to some extent on territorial influences beyond the recruiters' control. Second, the time of the calendar year production data are available has an effect that is, again, beyond the control of recruiter efforts. Finally, the number of months data are available for a recruiter is an issue, because small numbers of months' data may be quire unreliable.

Each of the potential problems was addressed in our validation studies. First, production data were corrected for the region's mean. Thus, recruiters working in a difficult region (with low average production rates across recruiters in the region) get more credit for their numbers of recruits compared to a recruiter from an easier region. Second, production data were standardized within month to overcome the problem of some months being easier than others (e.g., the summer when youths are just out of school is relatively easy). Finally, we computed the reliabilities for data from two through 12 months and determined that four months was a reasonable minimum ($r = .72$ for both the concurrent and predictive samples).

Correlations between the criterion measures were also computed to investigate the construct validity of the measures. These correlations for the predictive study appears in Table 1. They were very similar for the concurrent study.

Table 1. Correlations Between Criterion Measures

Criterion Measure	Production	Selling Skills	Human Relations Skills	Organizing Skills	Overall Performance
Selling Skills	.61**				
Human Relations Skills	.33**	.59**			
Organizing Skills	.23*	.40**	.49**		
Overall Performance	.44	.95	.86	.62	
N=123					

Production correlated highest with the Selling Skills rating factor (.61), the factor that most closely represents the skill, effort, and ability of recruiters to enlist prospects. Production correlated less highly with the Human Relations Skills (.33) and Organizing Skills (.23) rating factors. As in the predictive validation sample, this pattern of relations provides construct validity evidence for both the ratings and production index.

Development of the Predictor Measure

Based on the job analysis and experience we had studying Navy recruiting in the 1980s, we targeted five personality constructs: (1) Ambition; (2) Conscientiousness; (3) Emotional Stability; (4) Extroversion; and (5) Leadership. The Recruiter Assessment Battery (RAB) contained 109 items and is scored on five-point Likert scale ranging from "Definitely false" to "Definitely true." Example items are: "I often set higher standards for myself than others set for themselves" (Ambition) and "I obey a rule even if everybody around me is not obeying it" (Conscientiousness).

Validation Results

Again, we conducted a concurrent and a predictive validity study. For the concurrent study, we administered the RAB to approximately 200 on-production recruiters in three different Recruiting Districts, and also collected performance ratings from supervisors and peers at the same locations on the eight scales described previously, and finally, captured production data from the Recruiting Commands' files. Interrater reliabilities ranged from .60 to .81 for the three composite factors, and as mentioned, reliability for the production data was .72.

Table 2 presents the validities for the predictor scales against the performance criteria.

Table 2. Correlations Between RAB Factors and Criteria (Concurrent)					
	Selling Skills	Human Relations Skills	Organizing Skills	Overall Performance	Production
Ambition	.22**	.14*	-.01	.20**	.13
Conscientiousness	.03	.04	.18*	.06	.01
Extroversion	.21**	.20**	-.08	.20**	.16*
Leadership	.09	.10	-.02	.09	.12
Emotional Stability	.13	.12	-.02	.12	.19**

*p<.05, **p<.01
N = 199 for Rating Criteria
N = 200 for Production

For the predictive study, 623 students entering the Recruiter School were administered the RAB during their first week of school. Then, more than a year later, we were able to follow up on approximately 120 of this group to obtain supervisory ratings and 535 of them to get production data. Because there was one rater per ratee, no interrater reliability estimates were possible; reliability of the production data was again .72. Validities for the predictor scales are depicted in Table 3.

Table 3. Correlations Between RAB Factors and Criteria (Predictive)					
	Selling Skills	Human Relations Skills	Organizing Skills	Overall Performance	Production
Ambition	.05	.02	-.09	.07	.10
Conscientiousness	-.02	-.09	.02	-.04	.07
Extroversion	.18*	.21*	-.07	.20*	.13**
Leadership	.02	.02	-.13	-.01	.05
Emotional Stability	.27**	.26**	.14	.27**	.10*

N = 119 for Rating Criteria
N = 535 for Production

Overall, validities for the two studies are promising. Extroversion is the most consistent predictor of performance and productivity across studies. Ambition was a good predictor in the concurrent study, and Emotional Stability did very well in the predictive study.

Development of a Scoring Key

To develop a scoring key based on the validation research accomplished here *and* to estimate the likely validity of that key, we completed the following additional steps: (1) used the item-level concurrent validities to select items for the scoring key; (2) formed a composite of those items; (3) cross-validated the composite in the predictive validation sample; and (4) developed an interim final key by including primarily items with good validities in the predictive sample. The rationale for this approach was, first, that the best estimates of the likely operational validities come from the predictive study. In that study, the predictors were administered before the recruiter trainees had any recruiting experience, similar to what would be the case in their operational use.

The second part of the rationale for this approach is that although the Extroversion and Emotional Stability scales had reasonably consistent validities across the two studies, arguing perhaps for using a composite of those two scales as the key, there were several of those scales' items that showed poor validities *and* several Ambition and Leadership items that had very good validities. Thus, a combination of the reasonably large sample sizes providing good estimates of item level validities and the fact that several items from two of the less valid scales had promising validities argued for this strategy.

Validation results for the item-level analyses appear in Table 4. As expected, validities for the concurrent sample, where the items were selected based especially on their correlations with production, Selling Skills, and Overall Performance (51 items), are as high as .28. However, cross-validation results in the predictive sample are also promising, reaching .21 for production and .26 and .27, respectively, against the Sales Skills and Overall Performance rating criteria.

Table 4. Correlations Between RAB Score and Criteria

	Selling Skills	Human Skills	Relations	Organizing Skills	Overall Performance	Production
RAB Score (in CV)	.28**	.21**	-.02	.25**	.28**
RAB Score (in PV)	.26**	.26**	-.03	.27**	.21**

for CV N = 200 for production; 199 for ratings
for PV N = 537 for production; 127 for ratings

Conclusions

Although we have demonstrated reasonably good validity for the RAB, there is certainly a possibility that sailors taking the test in an operational setting where the scores will be used to select or screen out individuals for recruiting duty could slant their responses (i.e., fake good or fake bad). Warnings about faking and persuasive messages encouraging honest responses may help, but test takers determined to slant their responses may still be able to do so.

One strategy that would likely address this problem is to administer the RAB routinely to all Petty Officers soon after their first reenlistment. Test scores could then be placed in a file and later used to encourage or discourage coming into recruiting. This strategy might also be expanded to include other special assignment jobs such as drill instructor. The RAB could then be used as more of a classification or career counseling tool than a selection tool.

References

Penney, L. M., Borman, W. C., Hedge, J. W., Abrahams, N. T., & Drenth, D. J. (2001). *Development of recruiter and recruiter-in-charge (RINC) performance criteria*. (Technical Report 381). Tampa, FL: Personnel Decisions Research Institutes.