A Prospective Study of Cognitive and Noncognitive Selection Criteria as Predictors of Resident Performance

PEGGY S. WOOD, BA,* WILBUR L. SMITH, MD,† ELIZABETH M. ALTMAIER, PHD,* VALERIE S. TARICO, PHD,‡ AND E. A. FRANKEN, JR., MD‡

Wood PS, Smith WL, Altmaier EM, Tarico VS, Franken EA. A prospective study of cognitive and noncognitive selection criteria as predictors of resident performance. Invest Radiol 1990;25:855-859.

The primary purpose of this study was to compare objective measures of diagnostic radiology resident applicant performance, such as National Board (NBME) scores, with nonobjective measures, such as conscientiousness and interpersonal skills, in their prediction of resident performance. A second purpose of the study was to examine the predictive usefulness of the Accomplishment Interview (AI), a behavioral selection interview. Thirty resident applicants were interviewed using standard interviews, the AI, and objective data on the applicants were gathered. Four years later, during their diagnostic radiology residency, evaluations were gathered on these applicants, allowing for comparison of current resident performance with past predictive data. Results indicated that noncognitive factors were as important as cognitive factors in the prediction of resident behavior. Further, objective measures, such as NBME scores, failed to adequately predict residents' performance. The implications of these findings for resident selection are discussed.

Key words: resident; prediction; evaluation.

RESEARCH THAT EXAMINES the usefulness of selection criteria to predict resident performance has produced equivocal results. Several studies have reviewed the use of such cognitive predictors as parts I and II of the National Board of Medical Examiners (NBME), membership in Alpha Omega Alpha (AOA), class rank, medical school grades, and the dean's letter of recommendation.[1-5] NBME scores, especially those on part II of the examination, emerge as the strongest predictor of later performance,[1,5] but the actual correlation between these scores and later performance is very low. Membership in AOA is often given great weight in selection equations,[6] yet at least one study has found a complete lack of correlation between AOA membership and residency performance.[1] In addition, other traditional criteria, such as GPA in medical school and scores on the MCAT, also have demonstrated low correlations with resident performance ratings.[7]

Making matters more confusing, most of the cognitive criteria, such as AOA membership and the dean's letter, are inextricably related to, and possibly influenced by, NBME scores. For example, the dean knows the student's NBME scores when she or he writes a letter for the student; AOA membership is in part determined by NBME scores at many institutions.[3] Therefore, it is difficult to know whether residencies are using several different cognitive criteria in their selection process or one cognitive criterion expressed through several methods.

Most residency programs also use personal interviews as part of the selection process in order to assess personality characteristics, communication abilities, and other noncognitive factors.[8] The validity of traditional interviews as a selection criterion has been questioned by several researchers.[9,10] Prior research that has found no relationship between interviews and later resident performance can be criticized for failure to use standard interviews or interviews of sufficient length to adequately assess the applicant's competence.[10]

However, arguments in favor of the interview as a selection tool have pointed to its usefulness as a means of gathering information on noncognitive areas of competence, such as the ability to interact with others, maturity, confidence, and integrity.[5,11-13] These variables are often ignored by traditional selection processes; yet, they appear to be an important aspect of resident performance. Keck et al[5] found that cognitive and

From the *Division of Psychological and Quantitative Foundations, College of Education, University of Iowa, and the †Department of Radiology, College of Medicine, The University of Iowa, Iowa City, Iowa, and the ‡Department of Psychiatry and Behavioral Science, University of Washington, Seattle, Washington.

Reprint requests: Wilbur L. Smith, MD, The University of Iowa, Hospital and Clinics Department of Radiology, Iowa City, IA 52242.

noncognitive variables in combination were better than any individual variable or any class of variable alone in predicting the performance of physicians just after completion of residency. Tarico et al[13] and Altmaier et al[14] further noted that the behaviors most valued among residents by radiology staff dealt with social and attitudinal variables, two types of noncognitive variables that are often ignored in the selection and evaluation of residents.

Since the interview will likely continue to be used in selection, given its apparent usefulness in assessing certain variables, there is a need for the development of a standardized interview format. Such an interview format could be evaluated in a systematic manner by comparing residents' later clinical performance to the performance that would be expected based on the interview. This comparison would provide a validation of the interview as a predictor of later performance and, thus, support its use in selection decisions.

The purpose of this study was to examine the comparative predictive value of several measures: (1) traditional cognitive measures, such as NBME scores and membership in AOA; (2) traditional noncognitive measures obtained by usual faculty and resident interviews; and (3) a nontraditional, noncognitive interview, the Accomplishment Interview (AI), a standardized method of gathering information about the residents' competencies in five noncognitive areas (conscientiousness, curiosity, confidence, recognition of limits, and interpersonal skills).[12] We examined the degree to which each of these measures predicted radiology resident performance in the fourth year of residency.

Materials and Methods

The subjects of the study were 63 applicants who applied for a radiology residency at The University of Iowa in 1983. As a part of the application process, these applicants supplied their NBME scores (parts I and II), class rank, AOA membership status, the number of research projects that they had been involved in during medical school, and their number of publications. From September through December 1983, the applicants participated in a traditional, unstructured, on-campus selection interview conducted by various faculty members. This interview was given a mean score based on the subjective evaluation scores of four faculty interviewers. Applicants were also interviewed by at least one current radiology resident who used the same interview format. The faculty members and the residents were not blind to the applicant's AOA status, NBME scores, or class rank.

The AI was also administered at the time of the interview. Prior to their visit, applicants were sent a questionnaire that asked them to describe past situations in which they had best demonstrated the following qualities: interpersonal skills, conscientiousness, recognition of limits, confidence, and curiosity. Each of these qualities was briefly defined in the questionnaire. During their visit, applicants met with an interviewer to discuss and clarify their questionnaire responses. The interviewer who gathered the AI information was a graduate student in counseling psychology trained in the use of the AI and blind to the applicant's other application data.

The responses to the AI were scored independently by two raters, a graduate student in counseling psychology and a counseling psychology faculty member. The raters used behavioral bench marks as their guide in scoring each applicant for each area of behavior. The bench marks were developed by two experienced radiology faculty members to reflect poor, average, or outstanding performance in each area of behavior.

In 1987, an attempt was made to contact the 63 interviewees at their current diagnostic radiology residency (in which they were in the middle of their R-3 or R-4 year). The residents were asked to sign a consent form to allow the investigators to contact their current medical education director. After the resident agreed, the medical education director was sent two evaluation forms. One evaluation form was of the type currently used for performance assessment by The University of Iowa's Department of Radiology. It assesses skills in four areas: motivation, interaction with others, interpretive skills, and manual dexterity. The evaluator is asked to rate the resident on several components of these areas using a rating scale from 1 to 5, with 5 representing outstanding performance.

The second evaluation form was the Behavioral Observation Scale (BOS), a scale that was created from behaviors that job experts (in this case, radiologists) had generated and endorsed as being vital to competent resident performance.[13,14] On the BOS, the evaluator is asked to rate the resident on a Likert-type scale, ranging from 1 (almost never) to 5 (almost always), for each of 24 specific behaviors. Six behaviors were directly related to each of the areas tapped by the AI (conscientiousness, curiosity, interpersonal skills, and confidence; the fifth category, recognition of limits, was not included in this form of the BOS). An example of a BOS item from the conscientiousness area is "Rushes procedures in order to leave on time." High scores in a given category indicate greater competence in that area.

Of the 63 residents who were contacted, one was dropped from the study because she was currently doing a residency in a specialty other than radiology. Of the 62 remaining residents, 14 residents indicated that they did not wish to participate. We were unable to locate 13 of the residents at the address listed in the Radiology Resident Directory for 1987. Four residents did not return the consent form. Thirty-one residents indicated that they would participate, and of those 31 residents, we were able to obtain completed evaluations on 30 residents, for a participation rate of 49%.

Results

Pearson product-moment correlations were used to examine the relationships among the variables. First, the objective predictor variables were examined (Table 1). Board scores were positively correlated with membership in AOA ($r = .67$; $P < .001$). The correlation between research and publications was .45 ($P < .001$). These results suggest a high degree of interrelationship among the objective criteria.

Next, the faculty interview, the resident interview, and the total AI interview score were examined (Table 2). There was a moderate relationship between how faculty members and residents rated applicants based on the unstructured interview. The AI did not correlate with either the faculty interview or the resident interview,

TABLE 1. Intercorrelations Among Objective Predictor Variables

	AOA	Board Scores	Research	Publications
	r	r	r	r
AOA	—	.67*	.21	.05
Board Scores	—	—	−.02	.04
Research	—	—	—	.45*

*$P < .01$.
AOA: Alpha Omega Alpha, National Medical Honor Society.

TABLE 2. Intercorrelations Among Faculty Interview, Resident Interview, AI, and Objective Measures

	Faculty Interview	Resident Inverview	AI	Combined Objective Measure
Faculty interview	—	.33*	.15	.47*
Resident interview	—	—	−.22	.42†
AI	—	—	—	−.10

*$P < .01$.
†$P < .05$.
AI: accomplishment interview.

suggesting that an applicant's rating based on the AI was unrelated to the rating obtained by the other two interviews.

The faculty interview was significantly correlated with a combined objective measure (AOA + Board Scores + Research + Publications; $r = .47$ and $P < .01$). The interview conducted by the resident was also related to the objective measures ($r = .42$; $P < .05$). This fairly strong relationship between the objective measures and the ranking based on unstructured interviews might indicate that faculty and resident interviewers are basing some of their interview ratings on objective measures. The AI interview was not significantly correlated with the combined objective measure ($r => .10$, NS).

Prior to examining the prediction of performance, the interrelationships among evaluation variables were examined. The BOS total score and the standard radiology evaluation had a positive correlation of .42 ($P < .05$). Several BOS subscales were also significantly correlated with some of the standard radiology evaluation subscales. The interpretive skills subscale of the standard evaluation was correlated with the conscientiousness ($r = .57$, $P < .001$), the confidence ($r = .34$, $P < .04$), and the interpersonal ($r = .59$, $P < .001$) subscales of the BOS. The BOS curiosity subscale was moderately correlated with the motivation subscale of the standard evaluation ($r = .36$, $P < .05$). Finally, there was also a moderate correlation between the BOS interpersonal subscale and the manual dexterity subscale of the standard evaluation ($r = .37$, $P < .05$). These results support the construct validity of the BOS scale, because it should be correlated moderately with other measures of resident performance, but not so strongly that it appears to measure the same construct.

The correlations between the criterion variables and the predictor variables were examined to determine which variables were accurate predictors of later resident performance (Table 3). Of the objective predictors, NBME scores negatively correlated with several criterion variables: manual dexterity as measured by the radiology evaluation, motivation as measured by the radiology evaluation, interpersonal skills as measured by the BOS, and confidence as measured by the BOS. NBME scores did not positively correlate with any of the criterion variables.

AOA membership did not significantly correlate with any of the criterion variables, as was also the case with the objective measure of the amount of research. Number of publications, however, was positively correlated with curiosity in the residency, as measured by the BOS, indicating that resident applicants who publish prior to residency are likely to be rated fairly high during their residency on behaviors that demonstrate scientific and intellectual curiosity.

TABLE 3. Intercorrelations Among Objective Predictors and Evaluation

	Radiology evaluation measure				BOS measure			
	Motivation	Interaction	Manual dexterity	Interpretive skills	Conscien-tiousness	Curiosity	Confidence	Interpersonal skills
	r	r	r	r	r	r	r	r
Board scores	−.36	−.38	−.53*	−.23	−.02	−.44	−.45	−.42
AOA	−.17	−.00	−.08	.12	−.16	−.23	−.20	−.25
Research	−.07	.14	.07	.09	.10	−.19	−.27	.01
Publications	.17	.10	.02	−.02	.21	.40	.03	−.31

*$P < .01$.
†$P < .05$.
BOS: Behavioral Observation Scale; AOA: Alpha Omega Alpha.

TABLE 4. Intercorrelations Among Nonobjective Interview Data and Evaluation Scores†

	Radiology Evaluation Score			
AI score	Motivation	Interaction with others	Manual dexterity	Interpretive skills
Conscientiousness	−.08	−.22	−.11	−.01
Curiosity	.03	.06	−.01	.13
Confidence	−.21	−.11	−.01	.40*
Recognition of limits	−.09	−.39*	−.41*	−.22
Interpersonal skills	.21	.21	.09	.15

*$P < .05$.
†Radiology form.

TABLE 5. Intercorrelations Among Nonobjective Interview Data and BOS Scores†

	BOS Score			
AI score	Conscientiousness	Curiosity	Confidence	Interpersonal skills
Conscientiousness	.15	.00	.17	−.07
Curiosity	.04	.25	.21	.09
Confidence	.40*	−.48†	.03	.38*
Recognition of limits	−.26	−.00	−.04	−.46†
Interpersonal skills	.14	.16	−.02	.02

*$P < .01$.
†$P < .05$.
BOS: Behavioral Observation Scale; AI: accomplishment interview.

Analyses of the nonobjective interview data revealed that neither the faculty nor the resident interview ratings were significantly correlated with later resident behavior (Table 4). The AI interview, however, correlated with numerous resident behaviors. Residents' degree of confidence was positively correlated with their level of interpretive skills in residency, indicating that applicants' higher levels of confidence are likely to be associated with better interpretive skills in the residency. Resident applicants' recognition of their own limits was moderately negatively correlated with interpretive skills and manual dexterity, suggesting that the greater the tendency for the resident to recognize his or her limits, the worse the resident seems to be rated on the skills of manual dexterity and interpretation during the residency.

When results on the BOS were examined, there were numerous significant correlations between the AI and later resident behavior. As Table 5 reveals, applicants' degree of confidence was positively and significantly related to conscientiousness and interpersonal skills in the residency. Degree of confidence was moderately negatively correlated with later curiosity. Finally, applicants' ability to recognize their own limits was significantly negatively correlated with interpersonal skills during the residency.

Discussion

An important finding from this study was that objective measures, such as NBME scores and AOA membership, were found to not correlate with later behavior or to correlate in the opposite direction from what would be expected. AOA membership did not correlate with any performance measures; in this study, it was not a valid predictor of any future resident behavior. NBME scores were correlated with several outcome measures, although in the opposite direction: applicants who had better NBME scores demonstrated poorer resident performance in the areas of manual dexterity, interpersonal skills, and confidence. These results cast doubt on the usefulness of placing too much weight on NBME scores when making resident selections. NBME scores were designed to measure performance on content that is taught in medical school; they were never intended to be used to assess preparation for residency. The inappropriate use of these scores is not a remedy to solve selection woes.[15]

Ironically, the one objective measure that correlated in the direction expected was number of publications, an objective measure that has not emerged as a significant variable in most research on selection. In fact, number of publications is rarely mentioned as a variable that has been investigated. In this study, it was positively related to demonstration of curiosity during the residency.

Related to this issue is the importance of assessing noncognitive variables in the selection process. The best predictors of later performance appeared to be residents' degree of confidence and recognition of their limits. This finding suggests that those interested in the selection of residents should attend to noncognitive as well as cognitive skills. For example, our results demonstrated that level of confidence predicts interpretive skills, conscientiousness, and interpersonal skills. These skills have been rated by radiology faculty as being crucial to effective resident performance.[13] It was surprising that confidence was also negatively correlated with level of curiosity demonstrated in the residency. It would appear that resident applicants who exhibit too much ability to recognize their own limits are later rated lower on interpretive and interpersonal skills in the residency, perhaps because of a lack of assertiveness; this finding warrants further investigation.

Our results suggest that the AI interview may be a useful method for obtaining predictive information on resident applicants. The AI, a structured interview based

on desired job behaviors described by radiologists, appears to be a potential answer to the troubling issue of whether to put much weight on interviews when selecting residents. It would appear that the lack of predictive validity found for unstructured interviews in past studies[10,11] may be the result of retaining faulty interview protocols, using interviews that are unstructured, or allowing faculty members to interview applicants in any way that suits them. Added to the validity of the AI as an interview technique, the finding that neither the faculty nor the resident interview were predictive of later resident performance suggests that it would be beneficial for residencies, at minimum, to adopt a structured interview if they are going to continue interviewing as part of the selection process.

The finding that faculty and resident interviews were positively correlated and that both were correlated with objective measures implies that faculty and residents who are not blind to objective data prior to the interview can exhibit a confirmatory bias about the interviewee. If this is the case, then traditional interviews are not a good assessment of noncognitive variables but rather are an assessment by faculty of cognitive variables that have already been assessed by objective measures. A structured interview, such as the AI, conducted by an interviewer who is blind to the objective measures, would ensure that the interview assesses noncognitive variables in an unbiased manner.

In conclusion, this study provides support for the use of both cognitive and noncognitive variables in the selection of residents. This investigation also points to the use of a structured interview as a valid predictor of later resident performance, while at the same time casts doubt on the exclusive use of objective measures, such as board scores or AOA membership, as predictors of resident behavior.

Acknowledgment

The authors thank Michelle Greiner and Randall Ross for helping to gather data.

Supported by a grant from the Iowa Measurement Research Foundation (Altmaier).

References

1. Yindra KJ, Rosenfeld PS, Donnelly MB. Medical school achievements as predictors of residency performance. J Med Educ 1988;63:356–363.
2. Ross CA, Leichner P. Criteria for selecting residents: A reassessment. Can J Psychiatry 1984;29:681–686.
3. McCollister RJ. The use of part I National Board Scores in the selection of residents in ophthalmology and otolaryngology. JAMA 1988;259:240–242.
4. Wagoner NE, Suriana JR, Stoner JA. Factors used by program directors to select residents. J Med Educ 1986;61:10–21.
5. Keck JW, Arnold L, Willoughby L, Calkins V. Efficacy of cognitive/noncognitive measures in predicting resident physician performance. J Med Educ 1979;54:759–765.
6. Weiss ST, Rosa RM, Jofe T, Munoz B. A prospective evaluation of performance during the first year of the medical residency. J Med Educ 1984;59:967–968.
7. Richards JM, Jr, Taylor CW, Price PG. The prediction of medical intern performance. J Appl Physiol 1961;46:142–146.
8. Leonard A, Harris I. An approach for defining selection criteria of applicants for medical residency training. J Med Educ 1980;55:57–59.
9. Tremonti LP. Intern selection. Ann Intern Med 1973;79:277.
10. Komives E, Weiss ST, Rosa RM. The applicant interview as a predictor of resident performance. J Med Educ 1984;59:425.
11. Dean RE, Dean KB, Nicholas WR, Scholten DJ. The interviewing process as it relates to the selection of candidates for general surgical residency programs. Curr Surg 1987;44:1–6.
12. Tarico VS, Altmaier EM, Smith WL, Franken EA, Jr, Berbaum KS. Development and validation of an accomplishment interview for radiology residents. J Med Educ 1986;61:845–847.
13. Tarico VS, Smith WL, Altmaier E, Franken EA, Jr, Van Velzen D. Critical incident interviewing in evaluation of resident performance. Radiology 1984;152:327–329.
14. Altmaier E, Smith W, Wood P, et al. Cross institutional stability of behavioral criteria desirable for success in radiology residency. Invest Radiol (in press).
15. Volle RL. Using National Board of Medical Examiners scores in selection of residents. JAMA 1988;259:266.