

Empirical Item Keying Versus a Rational Approach to Analyzing a Psychological Climate Questionnaire

Christopher W. Hornick and Lawrence R. James
Institute of Behavioral Research, Texas Christian University

Allan P. Jones
Naval Health Research Center, San Diego, California

The present study compared two approaches to scoring a *Psychological Climate Questionnaire*—an empirical keying of items using item analysis and a rational approach which focused on identifying the underlying constructs measured by the questionnaire. The approaches were compared with respect to prediction of performance criteria and the ability to ascertain the theoretical underpinnings of the instrument. Questionnaires were completed by 398 male firemen in a large metropolitan area. Results demonstrated that a summation of item scores to reflect carefully designed a priori constructs did not necessarily result in weaker prediction of performance criteria. Similarities of the present study with previous work are discussed.

Three popular approaches to scoring a set of questionnaire items (e.g., personality inventories, biographical inventories, attitude surveys) for which there is no single correct response have been evident in the psychological literature (cf. Baehr & Williams, 1967; Goldberg, 1972; Levine & Zachert, 1951; Matteson, Osburn & Sparks, 1969; Nunnally, 1967). These three approaches have been labeled the *external* (or *empirical*) *approach*, the *internal approach*, and the *intuitive approach* (see Goldberg, 1972, for an extensive review of these approaches).

The *empirical approach* places initial emphasis on the prediction of an external criterion, typically using an empirical keying procedure based on item-criterion relationships. Understanding of underlying relationships or constructs accounting for the prediction has received secondary or no emphasis. This approach has been thought to maximize prediction of an external criterion (Guion, 1965). It has been criticized, however, because a capitalization on unique item variance often results in criterion-specific scoring keys with low internal consistency if the criterion is heterogeneous (Goldberg, 1972; Guion, 1965; Nunnally, 1967). The resulting lack of internal consistency often hinders determination of the theoretical composition or construct validity of the scoring keys (Baehr & Williams, 1967; Matteson et al., 1969; Nunnally, 1967). Thus, little understanding of conceptual relationships between the items and the criterion may be conveyed by this approach.

The *internal approach* uses the internal structure of the questionnaire items as the sole source of information to determine where and how items are to be used in the construction of item composites. Typically, this has been accomplished through the use of factor-analytic procedures or a strategy directed toward maximization of internal consistency (e.g., Loevinger, Gleser, & Dubois, 1953).

APPLIED PSYCHOLOGICAL MEASUREMENT
Vol. 1, No. 4 Fall 1977 pp. 489-500
© Copyright 1977 West Publishing Co.

The *intuitive approach* has typically utilized expert judgments to determine the suitability of an item for inclusion in a particular composite. These judgments are often based solely on theoretical considerations. However, recent investigators (e.g., Jones, James, Bruni, Hornick, & Sells, 1975; Matteson et al., 1969) have utilized both theoretical considerations and internal consistency analyses to construct item composites from a pool of questionnaire items. This latter approach, involving a mixture of intuitive and internal considerations, has been labeled a *rational approach* by Goldberg (1972) and others.

Proponents of the *rational approach* usually place initial emphasis on developing items to measure a theoretically meaningful set of constructs. Analyses of the item data focus on producing composites for each construct with high internal consistency. Prediction of one or more criteria then proceeds by regressing each criterion on the composites. Criticisms of a rational approach have been twofold: (1) the approach focuses on the communality of items and thus deletes or suppresses unique item variance which might have overlapped with the criterion and thereby improved prediction; and (2) construct validity, unless it is based upon more than inter-item inter-test relationships, is circular (Bechtoldt, 1959; Cronbach & Meehl, 1955; James, 1973; Loevinger, 1957).

The present study was designed to provide a comparison of two of the above approaches; that is, the rational approach and the empirical approach. Because the rational approach combined both internal and intuitive strategies for constructing item composites, a separate internal approach was not attempted.

A previous study by Goldberg (1972) compared the empirical and the rational approaches with many other strategies for constructing scales from a personality inventory to predict a variety of psychological criteria. The present study can be viewed as an extension and generalization of Goldberg's work to a different criterion, setting, sample, and type of questionnaire.

Comparisons of the results of these two studies are made later in this paper.

The questionnaire used in the present study was the *Psychological Climate Questionnaire* (Jones et al., 1975). This questionnaire was chosen for two reasons. First, previous effort devoted to defining and understanding the nature of psychological climate (cf. Forehand & Gilmer, 1964; Guion, 1973; James & Jones, 1974a) provided a basis for comparing the empirical and rational approaches in terms of their ability to represent underlying theoretical constructs. The second reason concerned the underlying theoretical relationship between psychological climate and behavior, which is developed below.

As presented by James, Hartman, Stebbins, and Jones (in press), *psychological climate* was defined as "the individual's internalized representations of organizational conditions and interrelationships among organizational conditions, [which] reflects a cognitive structuring of perceived situational influences in the situation." This definition was predicated on the assumption that psychological climate (1) is primarily descriptive of organizational conditions rather than evaluative or affective; (2) involves psychological processing, abstracting, and structuring of situational perceptions and the development of a cognitive map; (3) is multidimensional, where each dimension describes perceptions/cognitions of situations in terms of their perceived situational influence (e.g., warm, ambiguous, challenging, supporting, etc.); (4) is determined primarily by those characteristics of situations that have direct and immediate ties to individual experiences (e.g., leadership, role requirements, rewards); and (5) occupies an intervening role in a model of individual cognitive processing, serving to mediate between situational stimuli and individual attitudes and behavior.

Three points are of interest with respect to the above definition and assumptions. First, psychological climate operates at the individual level of explanation (i.e., a perceptual/cognitive process)

and is not to be confused with "organizational climate," which describes situations. Although conceptual and demonstrated empirical relationships exist between psychological and organizational climates (cf. James & Jones, 1974a), psychological climate is nonetheless an individual attribute and at least some individual differences are expected in perceptions/cognitions of situations. Second, although psychological climate is an individual attribute, it is not tautological with evaluative, affective variables such as attitudes. As noted in the assumptions above, psychological climate describes the perceptions/cognitions of situational stimuli which are then employed, in a model of cognitive information processing, to formulate affective reactions by comparing the perceptions/cognitions to individual needs, wants, and desires. Thus, psychological climate and affective reactions are presumed to be related, which could also include reciprocal interactions of affect upon perceptions/cognitions, but not a tautology (cf. James & Jones, 1974a for a review of this literature). Third, many current models of behavior (e.g., social-system models) suggest that behavior is a function of *both* individual and situational attributes. Furthermore, the major impact of situations on individual behaviors is presumed to be based upon individuals' perceptions/cognitions of the situation (cf. Endler & Magnusson, 1976). Thus, to the extent that perceived situational attributes were related to individual behaviors in the sample studied, a significant relationship was expected between psychological climate and job performance. Therefore, an investigation of the relationship of psychological climate to individual job performance provided a second basis for comparison of the two approaches.

In summary, the two scoring procedures for the questionnaire items were compared in terms of the prediction of individual performance criteria (using a double cross-validation design) and the ability to ascertain dimensions or constructs underlying the prediction instrument. This provided a means for examining the relative merits of the two approaches.

Method

Subjects

A 72% sampling of male firemen working in two municipal fire departments in a large metropolitan area in the southwestern United States yielded 398 respondents. The remaining 28% were either not available for group administration sessions (e.g., off-duty, illness) or were deleted in edits of the data (e.g., missing data). The average age of this sample was 36 years, and average tenure with the fire department was 11 years. A total of 78% were high school graduates; 43% had completed one or more years of college.

Criteria

A standard civil-service performance rating completed at regular 6-month intervals by an individual's immediate supervisor served as the performance criterion for this study. Ten major areas of job performance were represented in each rating: *devotion to duty, reliability, personal behavior and appearance, ability to make judgments, physical fitness, initiative, public relations, fire fighting and prevention knowledge, fire-fighting skills, and respect for authority*. Within each of these major areas, individuals were rated on a scale from 70 to 100. On the basis of results of criterion analyses, which are reported below, the average rating across the above ten areas was used to indicate each individual's overall or general performance level. The two most recent performance evaluations for each subject served as criteria. These ratings were obtained 3 months and 9 months prior to administration of the *Psychological Climate Questionnaire*, and thus relationships between the criteria and scores derived from the questionnaire technically would be designated post-dictive validation.

Psychological Climate Questionnaire

The *Psychological Climate Questionnaire* consisted of 170 items developed to measure 42 a

priori item composites representing perceived situational influences in the organizational environment. The composites were subgrouped into four areas: characteristics of the job; leadership at the level of the immediate supervisor; the workgroup; and major subsystems as well as the total organization. A more extensive discussion of the development of the questionnaire is reported elsewhere (James & Jones, 1974b; James et al., in press; Jones et al., 1975), and thus only a brief description is presented here.

Items and composites were based upon reviews of the literature appropriate for each respective area, previous climate research (cf. James & Jones, 1974a; Jones et al., 1975), pilot studies (James & Hornick, 1973; Jones, 1973), and interviews with firemen. Items for each composite were chosen on the basis of previously demonstrated construct validity whenever possible. For example, the leadership measures of *support*, *goal emphasis*, *work facilitation*, and *interaction facilitation* were based on the four-factor theory of leadership presented by Bowers and Seashore (1966) and the empirical operationalization by Taylor (1971). A separate continuum was constructed for each item, employing either descriptive attributes on a continuous scale or a Likert format.

Procedure for Rational Approach

The initial step in the rational approach involved the assessment of item intercorrelations within each of the 42 a priori composites. On the basis of low and inconsistent item intercorrelations in this sample as well as in two other samples (Jones et al., 1975), seven composites were dropped from all subsequent analyses. This resulted in a subset of 144 items representing 35 composites (see Table 1). Note that the composites are organized in the table according to the four conceptual areas enumerated earlier. Scores were computed for each of the 35 composites by summing the items (item variances were similar). Because the composite intercorrelations indicated that a more parsimonious set

of dimensions could be employed to describe the data, the composite scores were entered into a principal-components analysis. A varimax rotation was conducted for the components with eigenvalues greater than or equal to 1.0. Interpretation of the components was based on composites with loadings greater than or equal to $|\pm .40|$. The results of principal components analyses of the same 35 composites in two other samples also provided a basis for interpretation (cf. James & Jones, 1974b; Jones et al., 1975).¹ Finally, for prediction purposes, component scores were computed directly (Harman, 1967).

For the prediction analyses, the total sample was randomly split into two equal subsamples ($N = 199$). Within each subsample, the two composite performance criteria were regressed separately on all component scores by means of stepwise multiple-linear regression. The prediction equations generated within the subsamples were then cross-validated in a double cross-validation design.

Procedure for Empirical Approach

In order to facilitate comparison, the above cross-validation subsamples were also employed in the empirical keying approach. While the rational approach employed the 144 items that defined the retained 35 composites, the present approach included all the original 170 items, because this would have been the procedure used if

¹The other two samples included a sample of 4,315 Navy enlisted personnel and one of 504 lower, middle, and top managers in a private health-care foundation. The same 35 composites were submitted to principal component analyses with six components resulting in both analyses. Invariance of the six components across the three samples was examined by computing Tucker's coefficients of congruence (Mulaik, 1972). Five of the components were considered to generalize across the three samples (coefficients of invariance ranged from .74 to .97; Jones et al., 1975). Although differences existed in the component structures produced by the three analyses, these were a function of capitalization on unique characteristics within each sample to a large extent. Therefore, the names given the components differed somewhat depending on the sample being discussed. For a more detailed discussion of these analyses, see James and Jones (1974b) and Jones et al. (1975).

a rational approach had not been attempted. The items were entered into separate item analyses to predict each of the two composite ratings for each subsample. In this way, two empirical keys were developed for each subsample (one for each criterion) by selecting item alternatives which received significant ($p < .05$) biserial correlations with the criteria (Guion, 1965; Nunnally, 1967). A unit-weighting procedure was employed to construct the scoring keys which were also cross-validated in a double cross-validation design. Finally, the items comprising each key were content analyzed in an attempt to ascertain what was being measured by the scoring keys.

Results

Psychological Climate Components

Table 1 presents the rotated component structure of the 35 composites, communalities, and the number of items comprising each composite. The communalities, which provided lower bound estimates of reliability, were generally moderate (mean communality = .63). This was largely a function of the fact that the composites were designed to measure different, but not mutually exclusive, aspects of the organizational environment, and each composite encompassed a relatively small number of items. However, as reported below, the estimates of reliability for the component scores, based upon internal consistencies, were within traditionally accepted ranges (internal consistency estimates for the composites were not examined because of the small number of items per composite).

The principal components analysis provided six components with eigenvalues greater than or equal to 1.0. These components accounted for 63% of the trace of the correlation matrix.

The first component was defined by high positive loadings for emphasis on efficiency, professional spirit, and opportunities for growth and advancement. High negative loadings were obtained for role conflict and job pressure. This component reflected an emphasis on performing

tasks efficiently according to professional standards with an accompanying lack of role conflict and pressure. Moreover, such emphasis was perceived as associated with opportunities for growth and advancement and an awareness of employee needs and problems. In general, this component reflected an organizationally supported orientation toward professional expertise and appropriate rewards and was thus designated *Professional Orientation*.

The second component focused on characteristics of the workgroup, with high loadings by all the workgroup composites. This component indicated a cooperative and warm group environment with an emphasis on workgroup and organizational effectiveness and was designated *Workgroup Effectiveness and Cooperation*.

The third component was defined by composites measuring perceptions of leader behavior, especially the dimensions of the four-factor theory of leadership (Bowers & Seashore, 1966; Taylor, 1971). The name given this component was *Leader Facilitation and Support*.

All loadings above .40 for the fourth component were provided by selected composites from the job domain and focused on challenging and important tasks with accompanying opportunities to interact with others. The job-standards composite also loaded highly on this component. In general, this component reflected challenging and interesting aspects of tasks which were also associated with high quality performance (i.e., job standards). Thus, this component was designated *Job Challenge, Importance, and Quality*.

The highest loadings for the fifth component were generally provided by climate composites included in the larger subsystems and total organization domain. Composites with the highest positive loadings included organizational conflict, organizational ambiguity, and role ambiguity (from the job domain). Interdepartmental cooperation and fairness and objectivity of the reward process had high negative loadings on this component. This component indicated that organizational conflict and ambiguity were associated with a lack of interdepartmental cooperation.

Table 1
Principal Components of Psychological Climate Composites

Composites	Components						h ²	No. of Items
	I (λ=4,5)	II (λ=4,0)	III (λ=4,4)	IV (λ=3,0)	V (λ=3,6)	VI (λ=2,5)		
<u>Job Characteristics</u>								
Role Ambiguity	-.36	-.14	-.23	-.28	<u>.56</u>	-.19	.64	6
Role Conflict	<u>-.72</u>	-.17	-.02	-.04	.29	-.11	.66	6
Job Autonomy	<u>.56</u>	.12	.26	.13	-.12	.41	.60	4
Job Variety	<u>.37</u>	-.02	.13	<u>.55</u>	.04	.38	.60	4
Job Importance	.33	.33	-.01	<u>.55</u>	-.26	.01	.58	4
Feedback on Job	.23	.11	.40	.28	-.34	.34	.54	3
Job Challenge	.14	.21	<u>.27</u>	<u>.66</u>	-.24	.15	.65	4
Job Pressure	<u>-.63</u>	-.10	.08	.26	.31	.04	.58	6
Emphasis on Efficiency	<u>-.73</u>	.10	.18	.26	-.18	.05	.67	5
Job Standards	.07	.29	.29	.64	-.10	-.10	.62	4
Dealing with Others	-.03	.18	.09	<u>.62</u>	-.04	.16	.46	2
<u>Leadership</u>								
Leadership Support	.16	.30	<u>.61</u>	.07	-.28	.36	.70	5
Leadership Goal Emphasis	.14	.34	<u>.73</u>	.18	-.22	.12	.77	4
Leadership Work Facilitation	.12	.26	<u>.77</u>	.19	-.23	.27	.83	5
Leadership Interaction Facilitation	.09	.24	<u>.73</u>	.16	-.23	.28	.76	4
Leadership Planning and Coordination	.29	.33	.49	.18	-.31	.25	.63	3
Upward Interaction	.12	.04	<u>.70</u>	.10	-.10	-.08	.53	2
Confidence and Trust - Up	.32	.13	.20	.17	-.19	-.50	.48	2
Confidence and Trust - Down	.30	.20	.19	.03	-.25	<u>.65</u>	.66	5

(Continued on Next Page)

Table 1 (cont'd)
Principal Components of Psychological Climate Composites

Composites	Components						No. of Items
	I ($\lambda=4.5$)	II ($\lambda=4.0$)	III ($\lambda=4.4$)	IV ($\lambda=3.0$)	V ($\lambda=3.6$)	VI ($\lambda=2.5$)	
<u>Work Group</u>							
Work Group Cooperation	.02	<u>.70</u>	.14	.18	-.25	<u>.40</u>	.76
Work Group Friendliness and Warmth	.07	<u>.60</u>	.18	.15	-.24	<u>.46</u>	.69
Work Group Efficiency	-.02	<u>.78</u>	.14	.07	.17	<u>-.14</u>	.69
Work Group Pride	.09	<u>.80</u>	.27	.20	-.02	<u>.11</u>	.78
<u>Total Organizational Characteristics</u>							
Openness of Expression	<u>.48</u>	.20	.34	.22	-.30	.32	.62
Organizational Communication - Down	.35	.06	.35	.26	-.48	.34	.66
Interdepartmental Cooperation	.08	-.04	.17	.16	<u>-.73</u>	.16	.62
Organizational Conflict	-.37	-.28	-.12	-.13	.58	-.14	.60
Organizational Ambiguity	-.18	-.28	-.25	-.24	<u>.58</u>	-.06	.57
Organizational Consistency	.30	.23	.36	.07	-.40	.06	.44
Organizational Espirit	.28	<u>.50</u>	.20	.31	<u>-.16</u>	.26	.55
Professional Espirit	<u>.71</u>	-.05	.19	.08	-.08	.24	.61
Organizational Effectiveness	.26	<u>.70</u>	.22	.28	-.13	.11	.72
Reward Process (Fair & Objec.)	.15	<u>-.10</u>	.17	-.11	<u>-.60</u>	.06	.44
Opportunities for Growth and Advancement	<u>.65</u>	.03	.37	.32	-.07	.15	.70
Awareness of Employee Needs and Problems	<u>.40</u>	.11	.39	.19	-.41	.24	.58

tion and equitable rewards. Moreover, role ambiguity covaried positively with organizational conflict and ambiguity, which supports recent organizational theory (cf. House & Rizzo, 1972). This component was designated *Organizational Conflict and Ambiguity*.

The sixth component was more specific than the others and reflected primarily a mutual trust between subordinates and superiors. This component was named *Leadership Confidence and Trust*.

Internal Consistency of Climate Components

The internal consistency of each psychological climate component was assessed by computing the alpha coefficient (α) for component scores, developed by summing across the item composites shown in Table 1 with salient loadings ($\geq |\pm .40|$) of each component. For example, the internal consistency for the second component, *Workgroup Effectiveness and Cooperation*, was based on the intercorrelations of the scores from six item composites; namely, the four workgroup composites plus organizational spirit and organizational effectiveness. Reasonably high estimates of reliability were obtained for *Professional Orientation* (.87), *Workgroup Effectiveness and Cooperation* (.87), *Leader Facilitation and Support* (.90), and *Organizational Conflict and Ambiguity* (.86). Estimates of moderate internal consistency were obtained for *Job Challenge, Importance, and Quality* (.75) and *Leadership Confidence and Trust* (.78).

Criterion Analyses

An average across the ten major areas of job performance described earlier was used to represent overall or general performance in this study. The decision to construct a composite of the ten criteria was based on both theoretical and empirical considerations.

Theoretically, a lengthy debate has concerned whether criteria should be combined to represent an ultimate criterion or whether multiple

criteria, usually representing relatively independent dimensions, should be retained (cf. James, 1973). A combination of criteria has been viewed by some as being justified only if the criteria are at least significantly related (Guion, 1965). However, others (e.g., Ghiselli, 1956) have suggested that a condition approximating unidimensionality should exist before criteria are combined.

Empirical considerations thus centered on the relationships among the supervisory ratings. Inspection of these criterion ratings suggested that an overall combination of criteria would be a meaningful procedure. In order to confirm the suggested unidimensionality of the performance ratings, two analyses of the correlations among the ten ratings for each evaluation were examined for a random sample of 120 firemen.

The first analysis consisted of calculating an internal consistency reliability (coefficient α) for the sum of the ten ratings from each evaluation. The resulting internal-consistency estimates were .92 and .96.² The item-total correlations ranged from .57 to .88 for the earlier ratings and .73 to .93 for the later ratings. The rating intercorrelations were all significant and positive, ranging from .21 to .91 and from .50 to .91 for the two respective evaluations. All ratings were made on a scale from 70 to 100; however, supervisors nearly always restricted their ratings to between 83 and 97. This produced quite similar standard deviations for all ratings.

The results of a principal components analysis of each set of ten ratings also strongly supported the unidimensionality of the criteria. Two components (with eigenvalues ≥ 1.0) resulted from the analysis of the first set of ratings, with the first unrotated component accounting for 63% of the total variance and the second accounting for only an additional 12%. Analysis of the sec-

²The correlation between the two composite performance criteria was calculated to be .92 for the total sample of 398 firemen. Because the two criteria were identically measured except for the six-month interval separating them, this coefficient could be interpreted as test-retest reliability for the composite criterion.

Table 2
 Cross Validities for Psychological Climate Questionnaire Against Job Performance Criteria (N = 199 for each sample)

Sample	Method of Analysis			
	Item Analysis of Items		Multiple Regression of Component Scores	
	Criterion 1	Criterion 2	Criterion 1	Criterion 2
A	.49	.45	.55	.53
B	.47	.44	.46	.42

ond set of ratings resulted in a single component (with eigenvalue ≥ 1.0) accounting for 76% of the total variance.

In summary, the criterion analyses generally indicated unidimensionality in the ratings. Very high item-total correlations, high internal-consistency reliability, and strong indications of a single factor underlying the criteria appeared to support the use of an average rating for each evaluation to represent overall performance.

Validation Analysis for Rational Approach

The cross-validity estimates for the rational approach, obtained by the multiple regression of the six component scores against the two composite performance criteria, are presented in Table 2. The regression equations constructed on Sample A and cross-validated on Sample B provided cross-validity coefficients of .46 and .42 for the two performance criteria, respectively. The regression equations constructed on Sample B and cross-validated on Sample A provided cross-validities of .55 and .53 for the same respective performance criteria.

Validation Analysis for Empirical Approach

With respect to the empirical keying approach, results of the double cross-validation of the scoring keys are also presented in Table 2. The number of items included in the scoring keys constructed on Sample A were 59 and 63 items for the two respective composite criteria. For the scoring keys developed on Sample B, 65 and 55 items, respectively, were included to predict the two criteria. Table 2 shows cross-validities of .47 and .44 to predict the two respective criteria in Sample B from the scoring keys developed in Sample A. The scoring keys constructed on Sample B and cross-validated on Sample A provided cross-validities for the respective criteria of .49 and .45.

Content analyses of the four scoring keys generated in the item analysis provided no readily apparent content domains. The items compris-

ing each of the keys cut across most of the composites, including items from 26 to 28 of the 35 composites.

Comparisons of Rational and Empirical Approaches

Comparisons between cross-validities obtained using similar techniques required four significance tests (two criteria compared for each of the two subsamples) using the *t*-test for correlated variables (McNemar, 1969). For example, the cross-validity (.49) for the first criterion in Sample A using the empirical keying procedure was compared to the equivalent cross-validity (.55) using the rational approach, $t(196) = -1.20$, $p > .20$. None of the differences reached statistical significance, although the cross-validities tended to favor the rational approach in one sample and the empirical approach in the other.

Discussion

The major argument for the use of an empirical keying approach has been that it maximizes prediction of a complex criterion. In the present study, however, this contention was not supported because the cross-validities derived from an empirical approach were not significantly higher than those yielded by a rational approach. Such results are further strengthened by the fact that the empirical keying procedure included all 170 original items, while the rational approach used only 144 items. These findings did not appear to support the criticism that an empirical keying procedure yields higher cross-validities than a rational approach due to the deletion of unique item variance. In fact, the cross-validities resulting from the rational approach tended to be slightly, although not significantly, higher than the cross-validities provided by the empirical approach in Sample A.

These results support those of Goldberg (1972), who used similar approaches to study the relationships between a personality inventory

and a set of individual, psychological criteria. Goldberg's results indicated that an intuitive-rational approach for scoring a personality inventory produced higher cross-validities than an empirical procedure. In fact, the rational approach proved superior for all seven of his most predictable criteria. As pointed out by Goldberg, his results may have been a function of the fact that a rational approach capitalized less on the idiosyncrasies of the validation sample than did an empirical procedure. This same rationale appeared to be applicable in the present study as well.

It is important to note the differences between the present study and the Goldberg study in order to delineate the extent to which the current study extends and generalizes Goldberg's findings. First, the settings and samples were quite different. Goldberg's sample was comprised of female college freshmen whereas the sample for the present study, which was much larger, consisted of male firemen in a world-of-work situation. Second, although more numerous, Goldberg's criteria were psychologically oriented whereas ours consisted of performance or job-success ratings. Third, the questionnaire examined in the present study employed Likert-type item response formats in contrast to Goldberg's use of true-false item response formats. Finally, the content domains of the two questionnaires were different. Thus, considering the many differences in the studies, the results of the present study appeared to generalize as well as provide strong support for Goldberg's (1972) findings.

The present study also provided additional information by attempting to develop an understanding of what was being measured by the *Psychological Climate Questionnaire*. Whereas the construct validity of the personality inventory employed in Goldberg's study was not under investigation, it was a major concern in the present study. With respect to this concern, our results indicated that the rational approach appeared to provide a better understanding of what the questionnaire measured. That is, by

constructing items specifically designed to measure selected content domains, a clearer framework for analysis and a better understanding of what was being measured by the questionnaire were provided. In contrast, content analysis of items comprising the scoring keys produced by the empirical keying procedure did not provide for a clear understanding of what each represented. This was further indicated by the fact that each scoring key included items from 26 to 28 of the original 35 composites. Thus, prediction provided by use of the empirical approach did not aid in understanding the relationships between predictors and criteria.

On the other hand, the rational approach was better suited for understanding predictor-criterion relationships. Meaningful interpretation of these relationships was enhanced by selecting measures with at least the beginning of the establishment of a nomological net and thus some, but not complete, understanding of construct validity (cf. Royce, 1963). For example, prediction equations generated in both samples against both criteria yielded significant weights for only the first and fifth components (*Professional Orientation* and *Organizational Conflict and Ambiguity*, respectively). Both of these components received high negative weights. The negative relationship between *Organizational Conflict and Ambiguity* and the performance criteria appeared logical and consistent with the literature (House & Rizzo, 1972). Firemen who perceived low levels of organizational conflict and a lack of organizational ambiguity tended to be the better performers. Thus, low levels of conflict and ambiguity were related to high levels of performance.

The negative relationship between *Professional Orientation* and individual performance, however, is not as easy to explain. In fact, it appears counterintuitive; that is, it is not clear why low performance should be related to a high *Professional Orientation* (i.e., an organizationally supported orientation toward professional expertise and appropriate rewards).

Although the data did not permit a direct test,

recent literature suggests a possible explanation. Studies examining turnover (Farris, 1971) and relationships between turnover and psychological climate (Bruni, Jones, & James, 1975) have indicated the importance of equally attractive, readily available employment alternatives for the more qualified employees. That is, the characteristics which are related to higher performance might tend to give individuals a greater probability of finding other positions. The poorer performers, on the other hand, may have fewer opportunities for other jobs, and thus tend to be more oriented to their present positions. For example, Bruni et al. (1975) found that the less trained, less qualified sailors reported higher professional spirit and were more likely to reenlist than more qualified, better trained sailors. Farris (1971) similarly reported that a personal orientation to career growth was negatively related to retention. Such explanations for the findings of the present study must be regarded as tentative, however, because of their post hoc nature and require further investigation.

In summary, further comparisons of empirical and rational approaches are needed before definite conclusions may be drawn regarding the superiority of either. However, the results of the present study, as well as Goldberg's (1972) work, indicate that it is possible to place major emphasis on the rational approach without a necessary loss of predictive effectiveness. Moreover, in the present study, this approach provided a better understanding of the theoretical underpinnings of the questionnaire items than did the empirical keying procedure by suggesting possible implications and interpretations which were not evident in the empirical keying approach.

References

- Baehr, M. E., & Williams, G. B. Underlying dimensions of personal background data and their relationship to occupational classification. *Journal of Applied Psychology*, 1967, 51, 481-490.
- Bechtoldt, H. P. Construct validity: A critique. *American Psychologist*, 1959, 14, 619-629.

- Bowers, D., & Seashore, S. Predicting organizational effectiveness with a four-factor theory of leadership. *Administrative Science Quarterly*, 1966, 2 (2), 238-263.
- Bruni, J. R., Jones, A. P., & James, L. R. *Correlates of first-term reenlistment behavior aboard Navy ships*. (Tech. Report 75-8). Fort Worth, TX: Texas Christian University, Institute of Behavioral Research, 1975.
- Cronbach, L. J., & Meehl, P. E. Construct validity in psychological tests. *Psychological Bulletin*, 1955, 52, 281-302.
- Endler, N. S., & Magnusson, D. Toward an interactional psychology of personality. *Psychological Bulletin*, 1976, 83, 956-974.
- Farris, G. F. A predictive study of turnover. *Personnel Psychology*, 1971, 24, 311-328.
- Forehand, G. A., & Gilmer, B. V. H. Environmental variation in studies of organizational behavior. *Psychological Bulletin*, 1964, 62, 361-382.
- Ghiselli, E. E. Dimensional problems of criteria. *Journal of Applied Psychology*, 1956, 40, 1-4.
- Goldberg, L. R. Parameters of personality inventory construction and utilization: A comparison of prediction strategies and tactics. *Multivariate Behavioral Research Monograph*, 1972, (No. 72-2).
- Guion, R. M. *Personnel testing*. New York: McGraw-Hill, 1965.
- Guion, R. M. A note on organizational climate. *Organizational Behavior and Human Performance*, 1973, 9, 120-125.
- Harman, H. *Modern factor analysis* (2d ed.). Chicago: University of Chicago Press, 1967.
- House, R. J., & Rizzo, J. R. Role conflict and ambiguity as critical variables in a model of organizational behavior. *Organizational Behavior and Human Performance*, 1972, 7, 467-505.
- James, L. R. Criterion models and construct validity for criteria. *Psychological Bulletin*, 1973, 80, 75-83.
- James, L. R., & Hornick, C. W. An analysis of criteria, situational data, organizational climate, and individual characteristics: Results of a pilot study. In E. K. Gunderson (Chair.), *Organizational analysis: Models, methods, and criteria*. Symposium presented at the meeting of the American Psychological Association, Montreal, 1973.
- James, L. R., & Jones, A. P. Organizational climate: A review of theory and research. *Psychological Bulletin*, 1974, 81 (12), 1096-1112. (a)
- James, L. R., & Jones, A. P. An organizational model: Components and measurement. In B. T. King (Chair.), *Men in social systems: Results of a three-year multiorganizational study*. Symposium presented at the meeting of the American Psychological Association, New Orleans, September 1974. (b)
- James, L. R., Hartman, E. A., Stebbins, M. W., & Jones, A. P. Relationships between psychological climate and a VIE model for work motivation. *Personnel Psychology*, in press.
- Jones, A. P. *Functioning of organizational units related to differences in perceived climate and habitability*. Paper presented at the meeting of the American Psychological Association, Montreal, August 1973.
- Jones, A. P., James, L. R., Bruni, J. R., Hornick, C. W., & Sells, S. B. *Psychological climate: Dimensions and relationships*. (Tech. Report No. 75-3). Fort Worth, TX: Texas Christian University, Institute of Behavioral Research, 1975.
- Levine, A. S., & Zachert, V. Use of biographical inventory in Air Force classification program. *Journal of Applied Psychology*, 1951, 35, 241-244.
- Loevinger, J. Objective tests as instruments of psychological theory. *Psychological Reports*, 1957, Monograph Supplement 9.
- Loevinger, J., Gleser, G. C., & DuBois, P. H. Maximizing the discriminating power of a multiple-score set. *Psychometrika*, 1953, 18, 309-317.
- Matteson, M. G., Osburn, H. G., & Sparks, C. P. *A computer-based methodology for constructing homogeneous keys with applications to biographical data*. (Report 1). Houston TX: University of Houston, Personnel Psychology Services Center, 1969.
- McNemar, Q. *Psychological statistics*. New York: John Wiley & Sons, 1969.
- Mulaik, S. A. *The foundations of factor analysis*. New York: McGraw-Hill, 1972.
- Nunnally, J. C. *Psychometric theory*. New York: McGraw-Hill, 1967.
- Royce, J. R. Factors as theoretical constructs. *American Psychologist*, 1963, 18, 522-528.
- Taylor, J. C. An empirical examination of a four-factor theory of leadership using smallest space analysis. *Organizational Behavior and Human Performance*, 1971, 6, 249-266.

Acknowledgements

This paper is based on an earlier version presented at the 82nd Annual Convention of the American Psychological Association, New Orleans, Louisiana. The study was supported by the Office of Naval Research Contract Numbers N00014-A-0001, RR-42-08-01, NR170-743 and the Bureau of Medicine and Surgery, Department of the Navy, Research Work Unit NFS1524002-50tDX5f. The authors gratefully acknowledge the invaluable assistance of Kurt Helm in the collection of data for this study.

Author's Address

Chris W. Hornick, Institute of Behavioral Research, Texas Christian University, Fort Worth, TX 76129