Biographical data about verifiable accomplishments, collected in a structured questionnaire format, could contribute to the selection of school principals or early identification of persons to be developed for eventual succession into the principalship. Three criterion groups, selected to represent increasing levels of average performance, received surveys that generated response rates of 27 percent for the general principal group, 33 percent for the superintendent or research nominees group, and 33 percent for the nationally distinguished principal group. The survey included a seven-item composite (developed in a prior construction sample) that proved moderately efficient in discriminating high performing principals from principals-in-general in two cross-validation samples; the composite also identified modest positive correlations with principals' self-ratings of job performance. No significant sex or race differences in composite scores were observed. The results suggest the feasibility of using simple, inexpensive biographical inventories to supplement existing methods for selecting principals. (18 references and 3 tables) (KM)
SOME BIOGRAPHICAL CORRELATES OF OUTSTANDING PERFORMANCE AMONG SCHOOL PRINCIPALS

Gary D. Gottfredson and Lois G. Hybl
Center Staff

Edward L. McDill, Co-Director
James M. McPartland, Co-Director

Karl L. Alexander
Henry J. Becker
Barbara A. Bennett
Jomills H. Braddock II
Renee B. Castaneda
Barbara S. Colton
Russell L. Dawkins
Doris R. Entwisle
Joyce L. Epstein
Anna Marie Farnish
Denise C. Gottfredson
Gary D. Gottfredson

Edward J. Harsch
John H. Hollifield
Lois G. Hybl
Nancy L. Karweit
Melvin L. Kohn
Nancy A. Madden
Alejandro Portes
Robert E. Slavin
Carleton W. Sterling
Robert J. Stevens
Tammi J. Sweeney
Shi Chang Wu

Center Liaison

Rene Gonzalez, Office of Educational Research and Improvement

National Advisory Board

Patricia A. Bauch, Catholic University of America
Jere Brophy, Michigan State University
Jeanne S. Chall, Harvard University
James S Coleman, University of Chicago
Edgar G. Epps, University of Chicago
Barbara Heyns, New York University
David W. Hornbeck, Maryland State Department of Education
Michael W. Kirst, Chair, Stanford University
Rebecca McAndrew, West Baltimore Middle School
Sharon P. Robinson, National Education Association
The Center

The mission of the Center for Research on Elementary and Middle Schools is to produce useful knowledge about how elementary and middle schools can foster growth in students' learning and development, to develop and evaluate practical methods for improving the effectiveness of elementary and middle schools based on existing and new research findings, and to develop and evaluate specific strategies to help schools implement effective research-based school and classroom practices.

The Center conducts its research in three program areas: (1) Elementary Schools; (2) Middle Schools, and (3) School Improvement.

The Elementary School Program

This program works from a strong existing research base to develop, evaluate, and disseminate effective elementary school and classroom practices; synthesizes current knowledge; and analyzes survey and descriptive data to expand the knowledge base in effective elementary education.

The Middle School Program

This program's research links current knowledge about early adolescence as a stage of human development to school organization and classroom policies and practices for effective middle schools. The major task is to establish a research base to identify specific problem areas and promising practices in middle schools that will contribute to effective policy decisions and the development of effective school and classroom practices.

School Improvement Program

This program focuses on improving the organizational performance of schools in adopting and adapting innovations and developing school capacity for change.

This report, prepared by the School Improvement Program, examines the potential of using a simple biographical checklist to supplement existing methods for selecting principals.
Abstract

The potential of a simple biographical checklist to identify outstanding school principals was explored. A seven-item composite developed in a construction sample was moderately efficient in discriminating high performing principals from principals-in-general in two cross-validation samples, and it showed modest positive correlations with principals' self-ratings of job performance. No significant sex or race differences in composite scores were observed. Results suggest the feasibility of using simple, inexpensive biographical inventories to supplement existing methods for selecting principals.
America's school systems recruit and select a large number of new school principals each year. Despite the importance of these selection decisions, the procedures typically used in selection are widely recognized as inefficient and inadequate (Baltzell & Dentler, 1983; National Commission on Excellence in Educational Administration, 1987; National Governor's Association, 1986). The former U.S. Secretary of Education described a selection system with "patronage, politics, favoritism, or familiarity edging out merit" (Bennett, 1987, p. iii).

Recent research on the validity of assessment center ratings for predicting subsequent job performance for school principals (Gomez, 1985; Schmitt, Noe, Meritt, & Fitzgerald, 1984) provides one promising approach to improving the validity of selection decisions. Such selection research is rare, however, in a field marked by limited attempts to develop candidates for succession, the dominance of the interview as a selection method, and little use of actuarial combination of information in personnel decisions.

Biographical data have proven useful in the prediction of job performance for managers (Laurent, 1970), researchers (Smith, Albright, Glennon, & Owens, 1961), and other occupations (Schmitt, Gooding, Noe, & Kirsch, 1984) as well as in the prediction of occupational persistence (Cascio, 1976). In addition, biographical data about nonacademic accomplishments in college are useful predictors of a variety of adult accomplishments that are not well predicted by measures of academic potential (Munday & Davis, 1974). Childs and Klimoski (1986) recently showed that dimension scores from a biographical inventory are useful in predicting career success in a sample of persons working in diverse occupations.

Biographical data keys have traditionally been developed by empirically selecting items related to a criterion (England, 1971), and some research implies that such empirical keying can
produce more efficient prediction than weighted composites of homogeneous scales (Mitchell &
Klimoski, 1982). But research by Childs and Klimoski (1986), Eberhardt and Muchinsky
(1982), Munday and Davis (1974), and Owens and Schoenfeldt (1979) have illustrated the theo-
retical interpretability of dimension scores from biographical inventories. A theoretical under-
standing of the structure of biographical data could offer promise for the efficient generation of
biographical items likely to be useful in prediction. Little research has explored this possibility,
but the usefulness of Holland's (1985) classification and theory is suggested by the Eberhardt and
Muchinsky (1982) and Munday and Davis (1974) results.

Method

Instrument

Eighteen biographical items were assembled in a longer survey booklet that informed partic-
ipants of the purposes of the research, requested their voluntary participation, sought demo-
graphic information, requested critical incident data (not examined in this report), and called for
self-ratings of job performance. The eleven self-ratings were Likert-type items developed by
translating the most important and time-consuming job elements identified in an earlier job anal-
ysis (Gottfredson & Hybl, 1987) into items. (For example, "Observing teachers' instruction and
classroom management practices.")

Biographical data. The occupation of school principal is classified as Social-Enterprising-
Investigative. (SEI) according to Holland's occupational classification (Gottfredson & Holland,
in press), so 13 items that appeared related to these dimensions of activity or accomplishment
were written. For instance, serving as a volunteer for a social or health service agency is a Social
activity, collecting and analyzing data for an empirical evaluation is an Investigative activity, and
being elected an officer in high school or college government is an Enterprising accomplishment.
Three items to reflect prior recognition (e.g., been appointed by a local or state superintendent to
serve on a committee or task force) were included. Finally, two items that seemed related to
energy or initiative (e.g., worked at least 15 hours per week while attending college) were added.
Each item was verifiable in principle and called for a yes or no response.

Samples

Three criterion groups, selected to represent increasing levels of average performance, were
developed in the following manner.

Principals in general. From a pool of principals who had responded to an earlier survey
(Gottfredson & Hybl, 1987) we randomly selected a sample of 387 principals composed of
approximately equal groups of persons in a three-by-two classification of (a) urban, suburban,
rural and (b) elementary, middle/junior high school. The initial source of this pool was a com-
mmercial mailing list. (A sample of respondents to an earlier survey was used to allow merging
responses from the two surveys for a research purpose unrelated to the present report.)

Nominees. Letters were sent to 665 public school superintendents and research and evalua-
tion personnel asking for nominations of two elementary and two middle or junior high school
principals whom they judged to be outstanding performers. We attempted to limit this mailing to
districts with 20 or more schools to ensure the opportunity for respondents to be reasonably
selective. In all, 306 responses were received producing 884 nominations -- 277 by superinten-
dents and 607 by research and evaluation personnel. All superintendent nominees were included
in the sample, and enough research personnel nominees were added by random selection within
cells of the three-by-two (location-by-level) classification to match as closely as possible the
composition of the sample of principals-in-general. The constraint that districts have at least 20
schools to be included in this sample resulted in fewer nominees from rural districts than for the
in-general sample. The nominee sample numbered 386.

National distinguished principals. The population of 198 individuals who were honored in
the National Association of Elementary School Principals (NAESP) National Distinguished Prin-
cipal Program for the years 1984-1987 composed a third sample. NAESP state affiliates each
year select one principal whom they judge to be leading schools "in which a commitment to
excellence is evident; in which the programs are designed to meet the academic and social needs
of all students; and in which community ties with parents and local business organizations have
been firmly established" (NAESP, 1987, p. 36)

Mailings and Response Rates

Nine-page survey booklets were sent with a personalized cover letter by first class mail. A
follow-up mailing was sent a month later to principals who had not yet responded. This proce-
dure produced six booklets returned as undeliverable and 103 usable responses from the in-gen-
eral sample, for a 27% response rate; five returned as undeliverable and 126 completed by super-
intendent or research personnel nominees, for a 33% response rate; and six booklets returned as
undeliverable and 63 completed by national distinguished principals, for a 33% response rate.

Analyses

The in-general and nominated respondents were each divided into two subsamples using
computer-generated pseudo-random numbers. Two-thirds of each of these groups served as a
construction sample and one-third served as a validation sample to assess the efficiency of pre-
dictions in an unbiased sample. The small number of national distinguished principals was
treated as a second cross-validation sample.

The following seven biographical items which were empirically associated with group mem-
bership in the construction sample were selected to form an accomplishment composite: (a) con-
ducted a formal training workshop for other principals, (b) been elected an officer in a local, state
or national education organization, (c) presented an address on an educational, social, or scien-
tific topic before a community group other than at your school (e.g., Kiwanis, Rotary, church, or
business group), (d) published a paper in an educational journal or magazine or authored a book
that was commercially published, (e) received an award or honor for your performance as a prin-
cipal from a school system for which you worked, (f) served as a paid consultant on educational
problems outside your own school system, and (g) been appointed by a local or state school superintendent to serve on a committee or task force involving educators from diverse locations. Items were combined by simple unit weighting.

Cutting points on the composite were selected to produce assignments to higher and lower performing criterion groups in approximate proportion to their base rates, and the composite scores and these cutting points were applied in construction and cross-validation samples to assess the efficiency of classification. Cohen's (1960) *kappa* and standard errors for *kappa* (Fleiss, Cohen, & Everitt, 1969) were used to assess the magnitude and statistical significance of greater-than-chance agreement between concurrent predictions and actual group membership. Mean scores for criterion groups were also compared.

Internal consistency item analysis in the pooled respondent samples was used to assess the homogeneity of the accomplishment composite (*alpha* = .58) and a scale composed of the 11 self-ratings of job performance (*alpha* = .88).

Finally, the correlations of the accomplishment composite with a number of demographic and personal history variables were examined. These include self-reports of (a) number of years employed as a principal, (b) location of the principal's school (urban = 1, suburban = 2, rural = 3), (c) district size (20 or more schools = 1, fewer = 0), (d) school enrollment, (e) percentage of students receiving free lunch (an indicator of students' poverty), (f) sex (female = 2, male = 1), (g) race (minority status = 1, white = 0), (h) highest degree earned (less than bachelor's degree = 1, doctoral degree = 7), (i) college grade-point average, and (j) self-ratings of job performance.

Subsidiary analyses were conducted to determine if scores on the accomplishment composite differed by sex or race and if interactions with sex or race in the prediction of criterion group membership occurred. An *F* test for the increment to $R^2$ due to the addition of an interaction term after the appropriate main effects were already in the model was used to test for heterogeneity of prediction by race and sex.
Results

The most important results are displayed in Table 1. This table shows the percentage of correct classifications ("hits") in the construction sample and the two validation samples. Results shown in the lower (validation sample) portion of the table involve a comparison of each named sample with the random hold-out sample of principals-in-general. Percentage hits are shown alongside the base-rate percentages. Coefficient kappa may be interpreted as the proportion of greater-than-chance possible agreement that was observed. The z statistics show that kappas were significant at the $p < .001$ level except for the validation sample involving nominees where the $05$ level was not reached ($p < .08$).

The main results are summarized in another way in Table 2, which shows means and standard deviations for each group. As expected the composite has a significantly higher mean for nominees than principals in general in the construction sample ($p < .001$). The overall $F$ for group differences in the validation sample comparisons is also significant ($p < .001$), and Scheffe's test for post hoc comparisons shows the national distinguished principals to have higher scores than the other two criterion groups ($p < .001$). The nominees do not significantly differ from the principals-in-general in the validation sample. The national distinguished principals mean score exceeds the mean for the in-general sample by 1.3 standard deviations for the in-general group. This large between-group difference corresponds to the large (.57) value of kappa shown in Table 1 for the classification of principals as members of the in-general group or the distinguished group.

Correlations between a variety of personal history and background variables are displayed in Table 3. For this table, data from construction and validation samples are pooled, and correlations are shown separately within each group as well as for the total sample. The most important results in this table are the negligible correlations between race or sex and the biographical accomplishment composite, implying that the use of this composite in screening would not produce adverse impact.
Also of importance in Table 3 are the small and usually nonsignificant correlations between highest degree earned and accomplishments (-.14 to .19) and between college grades and accomplishments (-.03 to .18) implying that the accomplishments are relatively independent of these two credentials which are sometimes used in making selection decisions.

Table 3 also shows that for principals-in-general and the national distinguished principals the scores on the accomplishment checklist correlate modestly (.22 and .30) with self-ratings of job performance. Other analyses showed, however, that the self-rating scores did not differ significantly among criterion groups (the means were nearly identical), casting doubt on the utility of the self-ratings as a criterion measure. Overall, years as a principal is modestly (.17 to .24) correlated with the number of accomplishments claimed and these correlations are sometimes significant. In the pooled sample, there is a small (-.14) significant correlation between percentage of students receiving a free lunch and accomplishments. Finally, among principals-in-general, urban principals from larger districts claim significantly more accomplishments.

Tests for heterogeneity of prediction (using the pooled construction and validation samples for greater statistical power) found no significant interaction for either sex or race in discriminating national distinguished principals from principals-in-general or for sex in discriminating nominees from the in-general group. A significant race by composite interaction was found in discriminating nominees from principals-in-general ($F_{2.211} = 4.9, p < .03$). For whites, the correlation between the composite and membership in the nominee group was .36 ($n = 192, p < .001$) and for minorities the correlation was a nonsignificant .09 ($n = 23$). The small minority group had slightly and nonsignificantly higher mean scores on the composite than did the white group (3.74 versus 3.68) and nearly the same proportions of both groups were members of the nominee group (.54 versus .55 for minorities and whites, respectively). Although it is difficult to know what to make of this differential prediction for the small, mixed (68% black) minority group, the equal mean scores imply that adverse impact would not result if this composite were used in making decisions.
Explorations of the data implied that national distinguished principals came from more affluent schools according to the percent-free-lunch indicator than did principals-in-general (\(M = 20.2\) versus 31.4, \(p < .05\) by the Scheffe test, overall \(F\) for group differences \(p < .02\)). Distinguished principals also had a nonsignificantly lower percentage of students receiving free lunch than did nominees (\(M = 29.0\)). This observation, together with the Table 3 results showing a small (-.14) but significant negative correlation between the biographical composite and percentage of students receiving free lunch raises the possibility of criterion contamination: Perhaps school affluence makes principals look good and the good fortune of leading an affluent school leads to nomination as outstanding or to the national honor. In addition, other differences in the schools or districts (location, level, size of school) or personal history or demographic characteristics (years of experience as a principal, highest degree earned, sex, race, or college grades) could be correlated with both the accomplishment composite and with the criteria and so render the association of composite scores with criterion-group membership spurious.

To explore this possibility, two regression analyses using the pooled construction and validation subsamples were performed to determine if the addition of the accomplishment composite significantly improved prediction after all other significant predictors from the sets listed above had been allowed to enter the regression equation. One analysis used the contrast between principals-in-general and nominees as the criterion (excluding national distinguished principals from the analysis), and the second analysis used the contrast between the in-general group and national distinguished principals as the criterion (excluding nominees from the analysis).

For the nominee criterion, number of schools in the district and percentage of students receiving free lunch entered the equation with significant positive and negative regression coefficients, respectively (\(R^2 = .28, p < .001\)). The addition of the accomplishment composite to the equation significantly improved prediction (increment to \(R^2 = .04, p < .001\)). For the distinguished principal criterion, school level, percentage of students receiving free lunch, and level of highest degree entered the equation with significant coefficients ... first two negative, the third
positive, $R^2 = .16, p < .001$). The addition of the accomplishment composite to the equation significantly improved prediction (increment to $R^2 = .25, p < .001$).

**Discussion**

The results imply that biographical data about verifiable accomplishments collected in a structured questionnaire format could contribute to the selection of school principals or the early identification of persons who might be developed for eventual succession into the principalship. Because this type of data appears to have been overlooked in educational organizations, more systematic research and development using biographical data as predictors would likely lead to the development of moderately valid yet inexpensive assessment tools to complement methods now used.

The seven-item composite developed in this research was relatively independent of academic accomplishments and showed no race or sex differences, suggesting that inventoried biographical accomplishments may add to the predictive efficiency of assessments when combined with other forms of data and will not contribute to adverse impact.

No item written to represent Holland's (1985) Investigative dimension was selected in the criterion-related item analysis. All Investigative items had nonsignificant negative correlations with the criterion. Empirically keyed items appear to be related to Holland's Social and Enterprising dimensions and to reflect a recognition by others of a person's expertise or helpfulness.

Limitations of the present research--low response rates, the use of only a brief checklist of biographical accomplishments, the use of two judgmental methods of identifying criterion groups, the small sample sizes for the cross-validation samples, the small numbers of minority group members, and the use of extreme groups--imply that more systematic research using alternative criteria, lengthier biographical check lists from which to glean items, larger samples, and samples containing larger numbers of minority group members is required to clarify the extent of the potential of biographical data in selection and identification of principals. Despite the
limitations, the present results imply that further development is warranted and suggest some directions for that development.
References


Table 1

Efficiency of Accomplishment Composite in Discriminating Nominees and National Distinguished Principals from Principals in General

<table>
<thead>
<tr>
<th>Sample and contrast group</th>
<th>Base rate (%)</th>
<th>Percent hits</th>
<th>kappa</th>
<th>z</th>
<th>n</th>
</tr>
</thead>
<tbody>
<tr>
<td>Construction sample</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nominees</td>
<td>57</td>
<td>72</td>
<td>.44</td>
<td>5.20</td>
<td>142</td>
</tr>
<tr>
<td>Validation samples</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nominees</td>
<td>53</td>
<td>61</td>
<td>.21</td>
<td>1.82</td>
<td>74</td>
</tr>
<tr>
<td>Distinguished principals</td>
<td>62</td>
<td>79</td>
<td>.57</td>
<td>5.47</td>
<td>92</td>
</tr>
</tbody>
</table>

Note. A cutting score of 4 was used in predicting membership in the principals-in-general group versus nominees, and a cutting score of 5 was used in predicting membership in the group of distinguished principals. The percentage of individuals correctly classified in this way is shown in the column headed percent hits. The z statistic for kappa was calculated according to the method of Fleiss, Cohen, and Everitt (1969).
Table 2

Mean Accomplishment Composite Scores for Construction and Validation Sample Principals-in-General and Nominees, and for National Distinguished Principals

<table>
<thead>
<tr>
<th>Group</th>
<th>M</th>
<th>SD</th>
<th>n</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Construction sample</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Principals in general</td>
<td>2.98</td>
<td>1.53</td>
<td>61</td>
</tr>
<tr>
<td>Nominees</td>
<td>4.26</td>
<td>1.42</td>
<td>81</td>
</tr>
<tr>
<td><strong>Validation sample</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Principals in general</td>
<td>3.31</td>
<td>1.64</td>
<td>35</td>
</tr>
<tr>
<td>Nominees</td>
<td>3.97</td>
<td>1.65</td>
<td>39</td>
</tr>
<tr>
<td>Distinguished principals</td>
<td>5.44</td>
<td>1.34</td>
<td>57</td>
</tr>
</tbody>
</table>

Note. F(1,140) for the construction sample = 26.20, p < .001. Overall F(2,128) = 23.82, p < .001 for the validation sample groups. The distinguished principals differ from the other two validation sample groups at the p < .001 level by the Scheffe test.
Table 3

Correlations of Accomplishment Composite with Other School and Personal Variables

<table>
<thead>
<tr>
<th>Variable</th>
<th>Principals in general (n=90-96)</th>
<th>Nominees (n=113-120)</th>
<th>Distinguished principals (n=54-57)</th>
<th>Total sample (n=259-273)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Years as a principal</td>
<td>.17</td>
<td>.24**</td>
<td>.18</td>
<td>.24**</td>
</tr>
<tr>
<td>Rural location (vs. urban)</td>
<td>-.27**</td>
<td>.03</td>
<td>-.21</td>
<td>.08</td>
</tr>
<tr>
<td>District size</td>
<td>.24*</td>
<td>-.15</td>
<td>.07</td>
<td>.10</td>
</tr>
<tr>
<td>School enrollment</td>
<td>.19</td>
<td>.06</td>
<td>-.02</td>
<td>.08</td>
</tr>
<tr>
<td>Percentage of students receiving free lunch</td>
<td>-.08</td>
<td>-.04</td>
<td>-.11</td>
<td>.14*</td>
</tr>
<tr>
<td>Female</td>
<td>.04</td>
<td>-.02</td>
<td>-.05</td>
<td>.04</td>
</tr>
<tr>
<td>Minority</td>
<td>.10</td>
<td>-.05</td>
<td>.11</td>
<td>.04</td>
</tr>
<tr>
<td>Highest degree</td>
<td>.19</td>
<td>.08</td>
<td>-.14</td>
<td>.14*</td>
</tr>
<tr>
<td>College GPA</td>
<td>.18</td>
<td>.07</td>
<td>-.03</td>
<td>.13*</td>
</tr>
<tr>
<td>Self-rating of performance</td>
<td>.22*</td>
<td>.05</td>
<td>.30*</td>
<td>.18**</td>
</tr>
</tbody>
</table>