DOCUMENT RESUME

ED 283 867 TM 870 394

AUTHOR	Gomez, Joseph J.
TITLE	The Evaluation of the Management Assessment Center. Final Report.
INSTITUTION	Dade County Public Schools, Miami, FL. Office of Educational Accountability.
PUB DATE	Dec 85
NOTE	96p.
PUB TYPE	Tests/Evaluation Instruments (160) -- Reports - Evaluative/Feasibility (142)
EDRS PRICE	MF01/PC04 Plus Postage.
DESCRIPTORS	Administrator Evaluation; *Administrator Selection; *Assessment Centers (Personnel); Correlation; Elementary Secondary Education; *Evaluation Methods; Interrater Reliability; Job Performance; Personnel Selection; *Predictive Validity; Principals; Program Evaluation; *Rating Scales; *School Administration
IDENTIFIERS	*Dade County Public Schools FL

ABSTRACT

The Management Assessment Center (MAC) of the Dade County (Florida) Public Schools is a unique project, employing multiple techniques to evaluate behavior for school-level administrator personnel selection. This final report of an evaluation deals primarily with issues of the validity and the utility of the MAC. To ascertain validity of the MAC, the performance of candidates was correlated with their subsequent job performance. The data analysis of the results revealed that the validity correlations were positive and statistically significant. The evaluation noted: (1) inter-rater reliability was high; (2) the validity correlations were substantially higher than those generally produced by the interview method; (3) the validity correlations compared favorably with assessment centers; and (4) there is evidence that the validity correlations are still rising. Thus, it was concluded that the MAC does predict job performance. To investigate whether the resources allocated to the MAC are a worthwhile investment in the improvement of the selection process, the evaluation compared the results of the present to the former process. The results revealed that the interview MAC selection process is not superior to the interview-only process. The evaluation recommended that the minimum passing score of the MAC be raised to improve the effectiveness of the existing process. An appendix includes the M/R Rating Instrument. (BAE)

DADE COUNTY PUBLIC SCHOOLS

FINAL REPORT

ON THE EVALUATION OF

THE MANAGEMENT ASSESSMENT CENTER

OFFICE OF EDUCATIONAL ACCOUNTABILITY
PROGRAM EVALUATION DEPARTMENT

DECEMBER 1985

THE SCHOOL BOARD OF DADE COUNTY, FLORIDA

Mr. Robert Renick, Chairman
Dr. Kathleen B. Magrath, Vice-Chairman
Mr. G. Holmes Braddock
Mr. Paul Cejas
Dr. Michael Krop
Ms. Janet R. McAliley
Mr. William H. Turner

Dr. Leonard Britton
Superintendent of Schools

FINAL REPORT ON THE EVALUATION OF THE MANAGEMENT ASSESSMENT CENTER

Principal evaluator/author: Joseph J. Gomez, Ph.D.

Dade County Public Schools
Office of Educational Accountability
1450 Northeast Second Avenue
Miami, Florida 33132

December, 1985

The evaluation of the Management Assessment Center (MAC) spanned three years and generated two reports. The first report, which was titled Preliminary Report on the Evaluation of the Management Assessment Center, was published in March of 1984. It dealt primarily with the practices and procedures that constitute the operation of the MAC. The following document is the second and final report. While the final report deals briefly with the operation of the MAC, it does not replicate the preliminary report. The final report deals primarily with the critical issues of the validity and the utility of the MAC.

TABLE OF CONTENTS

EXECUTIVE SUMMARY. 1

INTRODUCTION 4

Assessment Centers in Education. 4

Management Assessment Center 5

Assessment Cycles of the MAC 6

VALIDITY OF ASSESSMENT CENTERS 9

DESIGN OF THE EVALUATION 12

Interview and Survey of the MAC Assessors: Design 12

Criterion-Related Validity of the MAC: Design 12

Comparison of the Selection Processes: Design 16

RESULTS OF THE EVALUATION. 18

Interview and Survey of the MAC Assessors: Results. 18

Criterion-Related Validity of the MAC: Results. 18

Intercorrelations of the Predictor 18

Intercorrelations of the Criterion 19

Inter-Rater Reliability of the Predictor 19

Consistency of the Criterion 22

Validity Correlations. 23

Longitudinal View of the Validity Correlations 36

Check for Criterion Contamination. 36

Comparison of the Selection Processes: Results. 40

CONCLUSIONS. 44

Interview and Survey of the MAC Assessors: Conclusions. 44

Criterion-Related Validity of the MAC: Conclusions. 44

Comparison of the Selection Processes: Conclusions. 46

RECOMMENDATIONS. 48

BIBLIOGRAPHY 49

APPENDIX 50

LIST OF TABLES

1. Success Rate per Assessment Cycle. 6

2. Success Rate per Candidate Group 7

3. Success Rate by Sex and Ethnicity for Cycle 1 Through Cycle 4: Adverse Impact of the MAC. 10

4. Skill-Exercise Rating Matrix of the MAC. 14

5. Job Dimensions Assessed by the Criterion Instruments 15

6. Inter-Rater Reliability of the MAC Skill Ratings Based on the Pearson r. 20

7. Inter-Rater Reliability of the MAC Skill Ratings Based on Coefficient Alpha. 21

8. Correlations Between the Composite Ratings of the Criterion Instruments. 22

9. Correlations Between the MAC Overall Skill Ratings and the Dimension Ratings Provided by the Supervisors on the Job Performance Scale. 24

10. Correlations Between the MAC Overall Skill Ratings and the Dimension Ratings Provided by the Support Persons on the Job Performance Scale. 25

11. Correlations Between the MAC Overall Skill Ratings and the Dimension Ratings Provided by the Subordinates on the Job Performance Scale. 26

12. Correlations Between the MAC Overall Skill Ratings and the Dimension Ratings Provided by the Subjects on the Job Performance Scale. 27

13. Comparison of the Validity Correlations Based on the Job Performance Scale in the MAC Evaluation and the Schmitt Study. 29

14. Correlations Between the MAC Overall Skill Ratings and the Dimension Ratings Provided by the Supervisors on the Effectiveness Scale. 31

15. Correlations Between the MAC Overall Skill Ratings and the Dimension Ratings Provided by the Support Persons on the Effectiveness Scale. 32

16. Correlations Between the MAC Overall Skill Ratings and the Dimension Ratings Provided by the Subordinates on the Effectiveness Scale. 33

LIST OF TABLES
(continued)

17. Correlations Between the MAC Overall Skill Ratings and the Dimension Ratings Provided by the Subjects on the Effectiveness Scale. 34

18. Correlations Between the MAC Overall Skill Ratings and the Dimension Rating Provided by the Each Rater Group on the Overall Rating Item. 35

19. Correlations Between the MAC Overall Skill Ratings and the 1983 and 1984 Composite Ratings on the Job Performance Scale for the Original 47 Subjects 37

20. Correlations Between the MAC Overall Skill Ratings and the 1983 and 1984 Composite Ratings on the Effectiveness Scale for the Original 47 Subjects 38

21. Correlations Between the MAC Overall Skill Ratings and the 1983 and 1984 Ratings on the Overall Rating Item for the Original 47 Subjects 39

22. Check for Criterion Contamination. 41

23. Comparison of the Job Performance Ratings of the Interview-Only Subjects and the MAC-Interview Subjects 43

EXECUTIVE SUMMARY

An assessment center is an assessment method that employs multiple techniques to evaluate behavior. The techniques can include written tests or interviews, but they are most often limited to job simulation exercises. The subject's behavior is observed by a group of assessors, who pool their observations to form a final evaluation. While industry has utilized the assessment center method for personnel selection since the 1950's, true assessment centers are relatively new in public education. For this reason, the Management Assessment Center (MAC) of the Dade County Public Schools is a unique project.

The MAC was developed in 1982 by Assessment Designs, Inc., a management consulting firm. The funds for the development of the MAC were provided by the state under the provisions of the Management Training Act of 1981. The district, however, underwrites the annual operating budget of the MAC, which excluding assessor time (approximately 520 days) is currently $94,982.

The conceptual framework of the MAC is based on a job analysis of the district's school-level administrators conducted by Assessment Designs. The job analysis identified the following nine skills as necessary for successful job performance: (a) leadership, (b) organizing and planning, (c) perception, (d) decision making, (e) decisiveness, (f) interpersonal, (g) adaptability, (h) oral communication, and (i) written communication. In order to assess these skills, three exercises were developed for the MAC. They include an in-basket exercise, a parent conference simulation and a teacher observation simulation.

The primary function of the MAC is screening candidates for the job of school-level administrator. Before a candidate can interview for a vacant position of principal or assistant principal, he/she must demonstrate through the MAC exercises the ability to successfully perform the job. Successful performance at the MAC means obtaining a minimum score of four on a seven-point rating scale for each of the nine skills. The skill ratings are provided by incumbent administrators (pay grade 43 or higher), who are specially trained to function as MAC assessors. The skill ratings are the composite judgement of three assessors, who observe the candidate's performance on the exercises. (For more detailed information on the MAC procedures, see the subsection titled Management Assessment Center on pages 5 - 6.)

The principal focus of the evaluation of the MAC was the validation of the process. Validation basically involves accumulating sufficient data on the process and its outcome to warrant confidence in decisions based on it. The validation of the MAC process was mandated by both legal and fiscal considerations. In reference to the legal consideration, personnel selection methods have repeatedly been challenged in the federal courts on the grounds of "adverse impact". Adverse impact is a situation where a personnel selection method works to the disadvantage of a legally protected race, sex or ethnic group. While assessment centers have been legally challenged less often than some other personnel selection methods (e.g., paper and pencil tests), many assessment centers do exhibit adverse impact. The MAC is no exception. Although limited in degree, the MAC exhibits adverse impact in the categories of race and ethnicity. And under the circumstances, legal prudence mandates that the validity of the MAC be documented.

In reference to the fiscal consideration, it should be acknowledged that assessment centers in general are more expensive than other personnel selection methods. In the interest of cost efficiency, the district must determine if the resources allocated to the MAC are a worthwhile investment in the improvement of the selection of school-level administrators. The initial step in making this determination is the validation of the MAC. (For more detailed information on the validity issues concerning the MAC, see the section titled Validity of Assessment Centers on pages 9 - 11.)

The evaluation of the MAC spanned three years and generated two reports, a preliminary report and this final report. The preliminary report, which was published in March of 1984, focused primarily on the MAC process. The report noted that during the first year of operation in 1981-82, the MAC had experienced some start up problems. The MAC staff, however, had been very responsive in addressing these problems, and thus had facilitated the subsequent development of the MAC. Consequently, the MAC assessors, who were in a unique position to observe the operation of the center, were very supportive of both the MAC staff and the MAC process. Indeed, the only significant problem in the MAC process identified by the preliminary report was the center's passing rate which was found to be comparably high. (For more detailed information on this phase of the evaluation, contact the Office of Educational Accountability and request a copy of Preliminary Report on the Evaluation of the Management Assessment Center.)

Of greater importance than the MAC process, however, is the intended outcome of the process, which is the prediction of a candidate's subsequent job performance. The degree to which the MAC achieves this objective is a measure of its validity as a personnel selection method. To ascertain the validity of the MAC, the performance of candidates at the MAC was correlated with their subsequent performance on the job. The data analysis of the results revealed that the validity correlations were positive and statistically significant. Moreover, the evaluation noted: (a) the inter-rater reliability, which is considered a prerequisite to validity in an assessment center, was high; (b) the validity correlations were substantially higher than those generally produced by the interview method; (c) the validity correlations compared favorably with those of other assessment centers; and, (d) there is evidence that the validity correlations are still rising. Thus, it was concluded that the MAC does predict job performance. (For more detailed information on the conclusions regarding the validity of the MAC, see the subsection titled Criterion-Related Validity of the MAC: Conclusions on pages 44 - 46.)

Beyond the question of validity is the question of the MAC's utility. In other words, are the resources allocated to the MAC a worthwhile investment in the improvement of the selection process for school-level administrators? In order to answer this question, the evaluation compared the results of the district's present selection process with the former selection process. The former selection process essentially consisted of a series of interviews for the qualified candidates. The present selection process differs in the use of the MAC to screen the qualified candidates prior to the interviews. The results of the comparison revealed that, despite the validity of the MAC, the interview-MAC selection process is not superior to the interview-only selection process. Thus, under the existing operating procedures, the MAC has no utility.

This outcome, nevertheless, is understandable given the minimum passing score of the MAC. The minimum passing score of the MAC is such that the few candidates who are eliminated from consideration would probably have been eliminated anyway by the interviews. Under the circumstances, the interviews in effect become the overriding factor in both selection processes. Thus, it was concluded that there was no advantage in incorporating the MAC into the selection process, not because of a deficiency in its validity but because its validity was essentially not used. (For more detailed information on the conclusions regarding the utility of the MAC, see the subsection titled Comparison of the Selection Processes: Conclusions on pages 46 - 47.)

Consequently, this evaluation recommends that the minimum passing score of the MAC be raised. This upward adjustment in the passing score should be done under the direction of a qualified consultant, since it will likely increase the adverse impact of the MAC. Assuming an appropriate adjustment in the passing score, the evaluation also recommends that the district retain the MAC as part of its selection process of school-level administrators. This recommendation is based on the established validity of the MAC, as well as the demonstrated competence of the MAC staff. The MAC by employing a higher minimum passing score will improve the effectiveness of the existing selection process. Without such an adjustment, however, there is no advantage in retaining the MAC.

INTRODUCTION

An assessment center is a personnel selection method which has been used in industry since the 1950's. Current estimates of the number of assessment centers in operation in the United States range from 300 to 1,000. The exact number is difficult to determine due in part to confusion over the definition of an assessment center. The term has often been incorrectly used to refer to a physical facility or to any program of assessment. A true assessment center, however, conforms to the "Standards and Ethnical Considerations for Assessment Center Operations." This set of standards, which was endorsed by the Third International Congress on the Assessment Center Method in 1975 and revised in 1980, defines an assessment center in the following manner:

> An assessment center consists of a standardized evaluation of behavior based on multiple inputs. Multiple trained observers and techniques are used. Judgements about behavior are made, in part, from specially developed assessment simulations.
>
> These judgements are pooled by the assessors at an evaluation meeting during which all relevant assessment data are reported and discussed, and the assessors agree on the evaluation of the dimensions and any overall evaluation that is made (Task Force on Assessment Center Standards, 1980, p. 35).

Thus, an assessment process is not an assessment center, if: (a) it relies on a single technique, (b) only a single assessor is involved, or (c) there is no pooling of data by the assessors.

Assessment Centers in Education

In the 1970's the assessment center method was introduced into the field of public education. In 1975 the National Association of Secondary School Principals (NASSP) in conjunction with the American Psychological Association developed an assessment center model. The purpose of this pilot project was to "demonstrate another approach to selecting potentially successful administrators that may set a standard for others to emulate" (Hersey, 1977). By 1984, a total of 19 assessment centers had been patterned after the NASSP model. These assessment centers served more than 200 school districts in 15 states; and, according to the NASSP, several additional centers were in the planning stages (Ogawa, 1984). This project consequently represents the most extensive use of assessment centers in public education to date.

In 1979, a three year validity study of the NASSP assessment centers was undertaken by a team of Michigan State University psychologists headed by Dr. Neal Schmitt. The subjects of the study were 167 individuals who had been assessed at seven separate assessment centers and subsequently hired as school-level administrators. When the group had achieved a job tenure of upward to four years, job performance ratings of each subject were obtained from the immediate supervisor, two teachers, two support staff employees, and the subject. The criterion instrument utilized by these raters was a performance scale that assessed 15 job dimensions of a school-level administrator. The scale was developed for the study from a national job analysis of principals and assistant principals conducted by Schmitt and his associates. The results of the study included a number of significant, positive validity correlations,

and Schmitt concluded:

> Examination of the criterion-related validity of the assessment center indicated that there were positive relationships between assessment center skill ratings and ratings of subsequent job performance (Schmitt, Noe, Meritt, Fitzgerald & Jorgensen, 1983, pg. 2).

Management Assessment Center

The Management Assessment Center (MAC) of the Dade County Public Schools is an assessment center that has been used since 1982 by the district to select principals and assistant principals. Although there are similarities between the MAC and the NASSP assessment center model, the MAC was developed independently for the district by Assessment Designs, Inc., a private consulting firm. The funds for the development of the MAC were provided by the state under the provisions of the Management Training Act of 1981. The state also provided the funds for the subsequent development of two additional sets of parallel exercises. The district, however, underwrites the annual operating budget of the MAC, which according to the Supervisor of the MAC is currently $94,982. This sum includes the cost of materials and supplies ($1,037), the cost of substitute services for MAC candidates currently holding a teaching position ($8,150), and the salaries of the MAC staff including fringe benefits ($85,795)[1]. The staff consists of a secretary (pay grade 19), and a supervisor (page grade 44) who reports to the Executive Director of the Bureau of Staff Development.

The conceptual framework of the MAC is based on a job analysis of Dade County school-level administrators conducted by Assessment Designs. The job analysis determined that the following nine skills are necessary for successful job performance: (a) leadership, (b) organizing and planning, (c) perception, (d) decision making, (e) decisiveness, (f) interpersonal, (g) adaptability, (h) oral communication, and (i) written communication. In order to assess these skills, three exercises were developed for the MAC. They include an in-basket exercise, a parent conference simulation, and a teacher observation simulation. During the first day of the two-day assessment process, each candidate's performance on these exercises is observed by three incumbent administrators (pay grade 43 or higher) who are specially trained to function as MAC assessors. On the second day, the three assessors pool their observations and through consensus rate the candidate's performance. In order to be successful, a candidate must obtain an average rating of four or higher on a seven point scale for each of the nine skills. Following the rating of the candidate, the assessors prepare the final report, and subsequently one of them meets with the candidate to review the report.

The primary function of the MAC is the screening of the candidates for the jobs of principal and assistant principal. Essentially, the MAC is one step in the selection process that includes the following five steps: (a) the candidate obtains administration/supervision certification from the state, (b) the candidate

---

[1] Not included in the sum is assessor time ($104,000, based on 520 days at an average daily rate of $200), since it represents no additional cost to the district.

takes the Technical Skills Test for School-Level Administrators[2], (c) the candidate is assessed at the MAC, (d) the candidate who is successful at the MAC is placed in a pool of eligible candidates, and (e) the eligible candidate is free to interview for appropriate open positions for a period of two years. Thus, before a candidate becomes eligible to interview for a job, the candidate must demonstrate in the job simulations of the MAC the ability to successfully perform the job.

Assessment Cycles of the MAC

An assessment cycle refers to the operational period of the MAC during a school year. The first assessment cycle occurred during 1981-82. There has subsequently been an assessment cycle every school year for a total of four cycles. Table 1 displays the number of candidates assessed per cycle, as well as the success rate. A review of the table reveals that the success rate for all candidates peaked during cycle 2 at 72.7%, and it has dropped to its lowest point of 64.0% during the most recent assessment cycle (Bureau of Staff Development, 1985). This trend is encouraging; because, as noted in the preliminary report, the success rate at the MAC appears to be much higher than the success rate of assessment centers described in the literature (Office of Educational Accountability, 1984).

Table 1

Success Rate per Assessment Cycle

		Candidates		
	Job	Assessed	Successful	Success Rate
Cycle 1 (1981-82)	Assistant Principal	205	132	64.3%
	Principal	120	91	75.8%
	Total	325	223	68.6%
Cycle 2 (1982-83)	Assistant Principal	240	170	70.8%
	Principal	90	70	77.7%
	Total	330	240	72.7%
Cycle 3 (1983-84)	Assistant Principal	179	123	68.7%
	Principal	68	48	70.5%
	Total	247	171	69.2%
Cycle 4 (1984-85)	Assistant Principal	208	128	61.5%
	Principal	70	50	71.4%
	Total	278	178	64.0%

[2]This district test is currently used as a diagnostic tool rather than an employment test. Candidates are simply informed of the results and advised on remediation. The results are not used to screen candidates.

Another method of examining the success rate of the MAC is to focus not on a point in time, i.e., an assessment cycle, but rather on a group of candidates. Since the candidates have more than one opportunity to be assessed at the MAC, it is more realistic to compute the success rate based on the eventual success of a designated group of candidates. The selection of a designated group of candidates for this purpose is circumscribed by two factors: (a) the provisions of Board Rule 6Gx13-4A-1.13, Assignment, Transfer and Appointment - Principals and Assistant Principals; and (b) the remediation practices recommended for the

Table 2

Success Rate per Candidate Group

		Candidates		
	Job	Assessed	Successful	Success Rate
Cycle 1 & 2	Assistant Principal	205	171	83.4%
	Principal	120	113	94.1%
	Total	325	284	87.3%
Cycle 2 & 3	Assistant Principal	240	176	73.3%
	Principal	90	74	82.2%
	Total	330	250	75.7%
Cycle 3 & 4	Assistant Principal	179	133	74.3%
	Principal	68	54	79.4%
	Total	247	187	75.7%

unsuccessful candidates. In reference to the first factor, Board Rule 6GX13-4A-1.13 states that the MAC candidates "assessed unacceptable twice... shall not be eligible to reapply for a period of two years following the last assessment." This means that a candidate who was unsuccessful in cycles 1 and 2 cannot be reassessed until cycle 5. Since cycle 5 is scheduled for 1985-86 at the earliest, is not necessary to accommodate a candidate in this situation in identifying a designated group of candidates. In reference to the second factor, the remediation practices recommended for the unsuccessful candidates encourages them to be reassessed in the cycle immediately following the initial assessment. While it is possible for an unsuccessful candidate to be reassessed in some other subsequent cycle, the Supervisor of the MAC has noted that this rarely occurs. Therefore, in keeping with these two factors, the designated group of candidates is defined as those candidates being assessed for the first time in any given cycle; and the method of computing the group's success rate is based on the results of this assessment cycle plus the results of the reassessment during the following cycle.

A comparison of Table 2 with Table 1 reveals that the success rates computed by this method are higher than the success rates based on the assessment cycle. For example, the success rate for all candidates based on the group ranges from 75.7% to 87.3% while the rate based on the assessment cycle ranges from 64.0%

to 72.7% (Bureau of Staff Development, 1985). Both methods, however, reveal the same trend of decreasing success rate. Nevertheless, the success rates computed by either method remain comparably high, since "in a typical assessment center only about half of those assessed are rated acceptable" (Cascio, 1982).

VALIDITY OF ASSESSMENT CENTERS

The basic focus of the present evaluation is the validity of the Management Assessment Center (MAC). While the process of validating a personnel selection method like the MAC is technically complex, the concept of validity is intuitively simple. It begins with the premise that most decisions involving personnel selection are made with something less than complete information. Since a degree of uncertainty is always present, enough must be known about the selection method to reduce the uncertainty to an acceptable level. Thus, a selection method is termed "valid" when sufficient knowledge has been accumulated about its process and outcome to warrant confidence in a decision based on it.

There are two fundamental reasons for establishing the validity of the MAC. The first reason is the legal issue. "Employee selection procedures have been challenged in Federal Courts literally hundreds of times and have not been upheld about half the time they have been challenged" (Boehm, 1982). Most of these challenges have been based on Title VII of the Civil Rights Act, the Office of Federal Contract Compliance Executive Order 11/46, and the 14th Amendment of the U.S. Constitution. These regulations and the numerous court rulings on the matter have led to the establishment of the Uniform Guidelines on Employee Selection Procedures by the Equal Employment Opportunity Commission (EEOC). The current issue of the Uniform Guidelines was published August 25, 1978, and it has been endorsed by the major regulatory agencies.[3]

According to the Uniform Guidelines, it is advisable to document the validity of an assessment center when the results of the center exhibit adverse impact. Adverse impact occurs when there is a personnel selection rate that "works to the disadvantage" of any legally protected race, sex or ethnic group. If a suit is brought against an organization's personnel selection practices, the enforcement agencies will generally resort to the Uniform Guidelines "80% rule of thumb" to detect adverse impact. The rule of thumb is defined as:

> A selection rate for any race, sex or ethnic group which is less than four fifths (or 80%) of the rate for the group with the highest rate will generally be regarded ... as evidence of adverse impact (p. 38297).

It should be noted that this rule is not a legal definition of discrimination. It is rather a "practical device" that is used to detect adverse impact when a personnel selection method is legally challenged.

The application of the 80% rule of thumb to the MAC reveals a degree of adverse impact, as Table 3 illustrates. Table 3 is a compilation of the total number of candidates assessed since cycle 1. The candidates are divided by sex, ethnicity/race and the job sought.[4] The application of the 80% rule of thumb reveals

---

[3] It should be noted that court rulings since the publication of the Uniform Guidelines have altered the applicability of some of the guidelines.

[4] The Uniform Guidelines makes no provisions for checking adverse impact by various sex-ethnic/race groups, e.g., Black females.

adverse impact in four categories: (a) Black principal candidates, (b) Hispanic principal candidates, (c) Black assistant principal candidates, and (d) Hispanic assistant principal candidates.[5] It should be noted that in the second category, i.e., Hispanic principal candidates, there would not have been adverse impact had one more candidate been successful. However, a degree of adverse im-

Table 3

Success Rate by Sex and Ethnicity for Cycles 1 Through Cycle 4: Adverse Impact of the MAC

	Sex		Ethnicity		
Principal Candidates	Male	Female	White	Black	Hispanic
Number of Candidates	201	147	164	120	64
Successful Candidates	140	119	139	77	43
Success Rate	69.6%	80.9%	84.7%	64.1%	67.1%
80% Rule of Thumb	NPC	HSR	HSR,NPC	75.6%	79.2%
Adverse Impact (X)				X	X

	Sex		Ethnicity		
Assistant Principal Candidates	Male	Female	White	Black	Hispanic
Number of Candidates	349	483	426	275	131
Successful Candidates	234	319	327	155	71
Success Rate	67.0%	66.0%	76.7%	56.3%	54.2%
80% Rule of Thumb	HSR,NPC	98.5%	HSR,NPC	73.4%	70.6%
Adverse Impact (X)				X	X

Note. HSR= group with "highest success rate"; NPC= "not a protected class" under Title VII of the Civil Rights Act.

pact is exhibited in the other three categories. Thus, it is legally prudent to document the validity of the MAC.

The MAC, however, is only one step in the personnel selection process. There is a possibility that the adverse impact is not carried through to the actual hiring of the candidates. This "bottom line" check for adverse impact, which was tentatively endorsed by the Uniform Guidelines, is no longer sufficient. In

[5] For example, the calculation of the 80% rule of thumb for the Hispanic assistant principal candidates involves dividing the group's success rate of 54.2% by the highest success rate in this category, i.e., 76.7% for White assistant principal candidates. The answer is multiplied by 100 and the product is 70.6%; since this is less than 80.0%, it is evidence of adverse impact.

Connecticut vs. Teal (1982), the U.S. Supreme Court ruled that an individual step within a personnel selection process could be challenged despite a bottom line that exhibited no adverse impact. Consequently, it is advisable to validate any step in a personnel selection process that exhibits adverse impact regardless of the bottom line outcome.

The second fundamental reason for establishing the validity of the MAC is the fiscal issue. Any personnel selection method involves some expenditure by the organization, but assessment centers are by comparison more expensive. Many organizations employing assessment centers accept the expense, because studies have generally demonstrated the value of assessment centers in personnel selection. Nevertheless, faith in the method is not a sufficient reason for justifying the expense of an assessment center. In the interest of cost efficiency, the district must determine if the resources allocated to the MAC are a worthwhile investment in the improvement of the selection of school-level administrators. The initial step in making this determination is the validation of the MAC.

DESIGN OF THE EVALUATION

The basic objective of the Management Assessment Center (MAC) is the prediction of the candidates' subsequent job performance. The degree to which the MAC achieves this objective is a measure of its validity as a personnel selection method (Division of Industrial-Organization Psychology, 1980). The present evaluation ascertained the criterion-related validity of the MAC by establishing a statistical relationship between performance at the MAC and subsequent performance on the job. In addition, the evaluation compared the results of the district's present selection process for school-level administrators with the former selection process. The essential difference between the two is that the existing process incorporates the MAC. Therefore, the comparison served to reveal the bottom line effect of incorporating the MAC into the selection process. In order to accommodate both the criterion-related validity and the comparison of the selection processes, the MAC evaluation was partitioned into three phases: (a) the interview and the survey of the MAC assessors, (b) the criterion-related validity of the MAC, and (c) the comparison of the selection processes. For the sake of clarity, each phase will be addressed separately in this section and in several subsequent sections of the report.

Interview and Survey of the MAC Assessors: Design

The initial phase of the MAC evaluation involved the interview and survey of the MAC assessors. Both the interview and the survey were conducted after the completion of assessment cycle 1 in 1981-82. The assessors were targeted for this inquiry, because they were in a unique position to provide insights into the operation of the MAC. Structured interviews were conducted with 12 select assessors (approximately 15% of the assessors at that time). The group included representatives from the central office, the area offices, and every school-level. The assessors were asked nine open-ended questions regarding the general operation of the MAC. Their responses served as a guide in the development of the survey instrument. All the MAC assessors (81 at that time) were forwarded the survey instrument, and 77.7% returned usable responses.

Criterion-Related Validity of the MAC: Design

The second phase of the evaluation involved the criterion-related validity of the MAC. Criterion-related validity is a method of statistically verifying a probable relationship between a predictor and a criterion. In this case, the predictor was the MAC skill ratings of a designated group of subjects, and the criterion was the subsequent job performance ratings of these subjects. The relationship between these two variables was calculated by the Pearson product-moment correlation coefficient, i.e., the Pearson r. Thus, the basic elements of the design of the second phase of the evaluation consisted of: (a) the predictor, (b) the criterion, (c) the subjects and (d) the Pearson r. In the subsequent paragrphas, each element is individually examined.

The predictor, as previously noted, consisted of the subjects' skill ratings from the MAC. The skill ratings of each subject include an individual score by exercise by skill for a total of 24 scores. This number is less than the pro-

duct of the three exercises by nine skills, because every skill is not assessed in every exercise. As Table 4 illustrates, adaptability is not assessed in the in-basket exercise; leadership is not assessed in the parent conference exercise; and written communication is not assessed in the teacher observation exercise. In addition to these individual 24 scores, each subject receives nine overall skill scores consisting of the weighced average of the individual scores within a skill. Each of the individual and overall skill scores is based on the seven-point rating scale illustrated in Table 4. Successful performance requires that the subjects obtain a minimum score of four on each of the nine overall skill scores.

The criterion, as previously noted, consisted of the subjects' job performance ratings. The job performance ratings were obtained by means of two instruments, the Job Performance Scale and the Effectiveness Scale (see Appendix). The Job Performance Scale is the instrument developed by the team of Michigan State University psychologists headed by Dr. Neal Schmitt for the validity study of the NASSP assessment centers (Schmitt et al., 1983). This instrument is a behaviorally anchored, six-point rating scale which assesses the 15 job dimensions of a school-level administrator listed in Table 5. The instrument was deemed appropriate for this evaluation because it was developed from a national job analysis of school level administrators. The use of this instrument, furthermore, afforded an opportunity to compare the results of the MAC evaluation and Schmitt et al. (1983) study. The second instrument employed was the Effectiveness Scale. This instrument was developed specifically for this evaluation by Dr. Larry Skurnik, an industrial psychologist formerly employed by the district.[6] The instrument was based on Assessment Design's job analysis as reflected in both the exercises and the skills assessed at the MAC. A five-point rating scale is used in the instrument to assess the 14 job dimensions of a school-level administrator listed in Table 5. In addition, the instrument incorporates an overall rating item designed to evoke a holistic response. This item differs from the other items in the instrument in the use of a stanine scale. Consequently, this single item was treated in the evaluation as a separate instrument under the operational name of Overall Rating Item.

The MAC evaluation spanned the two school years of 1982-83 and 1983-84. The job performance ratings of the subjects were obtained in June of each school year. The ratings for each subject were provided by four individuals who were in a position to observe the subject's job performance. They included: (a) the subject's supervisor (an area director if the subject was a principal; a principal if the subject was an assistant principal), (b) a support person (an assistant principal randomly selected), (c) two subordinates (two teachers randomly selected), and (d) the subject him/herself. To insure objectivity, a coding system was used to keep the ratings confidential. Moreover, each rater had the option of not responding to any item in the instruments that requested information he/she could not provide. These factors probably contributed to the high response rate. The composite number of raters returning usable criterion ratings was 91.8% in 1983 and 92.2% in 1984.

[6] Dr. Skurnik is presently employed by Educational Testing Service.

Table 4

Skill-Exercise Rating Matrix of the MAC

Skill/Exercise	1 In-Basket	2 Parent Conference	3 Teacher Observation	Overall Skill Score
1. Leadership		a		
2. Organizing and Planning				
3. Perception				
4. Decision Making				
5. Decisiveness				
6. Interpersonal				
7. Adaptability	a			
8. Oral Communication				
9. Written Communication			a	

MAC Rating Scale

7 = Outstanding

6 = Well above satisfactory

5 = Above satisfactory

4 = Satisfactory

3 = Below satisfactory

2 = Well below satisfactory

1 = Weak

[a]This skill is not assessed in this exercise.

Table 5

Job Dimensions Assessed by the Criterion Instruments

Job Performance Scale

1. Curriculum and Instructional Leadership: Monitoring Curriculum Objectives
2. Curriculum and Instructional Leadership: Monitoring Individual Progress
3. Coordination of Student Activities: Supervision
4. Student Activities: Participation
5. Direction of Support Services
6. Support Services: Directing the Behavior of Students
7. Staff Evaluation
8. Developmental Activities
9. Community Relations
10. Interpersonal Effectiveness
11. Community Relations: Parents
12. Coordination with District and Other Schools
13. Fiscal or Monetary Management
14. Maintenance of School Plant
15. Structures Communication Which Provide for Cooperation Among Various Groups in School

Effectiveness Scale

1. Adaptability
2. Conferring with Parents
3. Decision Making
4. Decisiveness
5. Interpersonal
6. Management Skill
7. Observing Teachers
8. Oral Communication
9. Paper Work
10. Planning and Organizing
11. Perception
12. Leadership
13. Technical Know-How
14. Written Communication

The subjects in this phase of the evaluation consisted of 121 incumbent school-level administrators, who had been candidates at the MAC during cycles 1, 2 and 3.[7] This group included 101 incumbents who had been successful at the MAC and subsequently hired by March of 1984; and 20 incumbents who, although unsuccessful at the MAC, were incumbent school-level administrators. These 20 incumbents consisted of assistant principals who had been assessed for the job of principal. Their inclusion in the group of subjects provided representation of the unsuccessful candidates, who constitute approximately 30% of the assessed population (see Table 1). In addition, it lessened the restriction in the range of the MAC ratings which would have resulted in a suppression of the correlations between the predictor and the criterion.

The Pearson r was used to compute the correlations between the predictor and the criterion, i.e., the validity correlations. The Pearson r is a formula for calculating the degree of linear relationship between two variables. The formula yields a correlation coefficient which ranges from - 1.00 to +1.00. A positive correlation coefficient indicates that the high values of the predictor (the first variable) are associated with the high values of the criterion (the second variable) and/or the low values of the predictor are associated with the low values of the criterion. A negative correlation coefficient indicates an inverse relationship; in other words, the high values of the predictor are associated with low values of the criterion and/or vice versa. A correlation coefficient near zero indicates that there is no relationship between the predictor and the criterion.

The Pearson r was used to compute two sets of validity correlations. The first set involved the entire group of 121 subjects, and it was computed using the job performance ratings obtained in June of 1984. Since this set of validity correlations is based on the last performance ratings of all the subjects, it represents the most crucial statistical relationship in this phase of the evaluation. The second set of validity correlations involved the 47 subjects hired during 1982-83, and it was computed using first the job performance ratings obtained in June of 1983, then those obtained in June of 1984. A comparison between the 1983 and 1984 validity correlations was done to identify the longitudinal trend in the values of the coefficients. The longitudinal trend is of interest, because several studies have noted a tendency of the values of the validity correlation coefficients to rise over time (Bray & Grant, 1966; Mitchel, 1975; Hinichs, 1978). In addition to the validity correlations, the data analysis of this phase of the evaluation included the following procedures: (a) the intercorrelations of both the predictor and the criterion, (b) the inter-rater reliability of the predictor, (c) the correlations of consistency of the criterion, and (d) a check for criterion contamination.

Comparison of the Selection Processes: Design

The third and final phase of the MAC evaluation involved a comparison of the district's present selection process for school-level administrators with the former selection process. The former selection process essentially consisted of a series of interviews for the candidates. The present selection process differs in the use of the MAC to screen the candidates prior to the interviews. The focus of the comparison was the subsequent job performance of the candidates selected by these two respective processes.

[7] The group originally consisted of 122 subjects, but the transfer of one incumbent hired during 1982-83 to a non-school-level administrative position reduced the number to 121.

The subjects in this phase of the evaluation consisted of 109 incumbent school-level administrators. The interview-only process was used to select 62 of them. These 62 subjects were hired during 1981-82, the last year the interview-only selection process was operational. The remaining 47 subjects were hired the following school year using the MAC-interview selection process. Incidently, these 47 subjects were the same subjects used in the longitudinal view of the validity correlations in the second phase of the evaluation.

Since the interview-only subjects had on an average one more year of job tenure than the MAC-interview subjects, the comparison was staggered in time to equate the two groups on this factor. The June of 1983 job performance ratings of the interview-only subjects were compared to the June of 1984 ratings of the MAC-interview subjects. The method of rating the job performance of each group adhered to the same instruments and procedures used in the second phase of the evaluation.

It should be acknowledged that the design of this phase of the evaluation is not a true experimental design, because the initial distribution of subjects into the two groups was not achieved through randomization. This was, of course, precluded by the operational status of the MAC. While it would have been preferable to equate the two groups through randomization, no reason has been identified that would suggest that the two groups were characteristically different prior to selection. Not only is it unlikely that the pool of potential candidates changed significantly from 1981-82 to 1982-83, but self-nomination was the initial step in both selection processes.

Finally, the statistical comparison of the two groups was done by means of analysis of variance (ANOVA). ANOVA is a method of dividing the variation of quantitative data into different parts in order to determine the source or cause. The relative magnitude of the variation from each source can then be assessed to ascertain whether it is greater than expected. In this phase of the evaluation, ANOVA was used to test the statistical significance of the difference between the means of the job performance ratings of the two groups. If the MAC-interview selection process is indeed superior to the interview-only selection process in the selection of better school-level administrators, the mean of the job performance ratings of the MAC-interview subjects should be significantly higher than the mean of the interview-only subjects.

RESULTS OF THE EVALUATION

The results of the evaluation of the Management Assessment Center (MAC) will be addressed in this section of the report. It should be noted that this section adheres to the same format employed in the Design of the Evaluation, in that the results of each of the three phases of the evaluation are addressed separately.

Interview and Survey of the MAC Assessors: Results

Since the interview and the survey of the MAC assessors were conducted after the completion of assessment cycle 1 in 1981-82, the results of this phase of the evaluation were presented in the preliminary report which was published in March of 1984. These results, therefore, will not be replicated in this report. A reader who wishes to review the results of the first phase and/or to examine the interview and survey instruments should contact the Office of Educational Accountability and request a copy of Preliminary Report on the Evaluation of the Management Assessment Center.

Criterion-Related Validity of the MAC: Results

The second phase of the evaluation involved the criterion-related validity of the MAC. The results of the computation of the validity correlations will be presented in this section. In addition, the results of several data analysis procedures encompassed by the design of this phase will be included. Specifically, the results of the following procedures will be addressed: (a) the intercorrelations of the predictor, (b) the intercorrelations of the criterion, (c) the inter-rater reliability of the predictor, (d) the consistency of the criterion, (e) the validity correlations, (f) the longitudinal view of the validity correlations, and (g) the check for criterion contamination.

Intercorrelations of the Predictor

The intercorrelations of the predictor refers to the correlations between the skill ratings of the MAC. These intercorrelations are important, because they indicate whether the skill ratings are providing unique information on the subjects' performance. Relatively low intercorrelations denote uniqueness, while high intercorrelations denote communality. Communality implies that some of the skill ratings are tapping a common factor in the subjects' performance. Under such circumstances, it may be possible to delete some skills from the set without an appreciable loss of information.

The MAC evaluation revealed that the intercorrelations of the skill ratings are moderate with the majority of the coefficients falling in the 0.50's and 0.60's.[8]

[8] In describing the correlation coefficients within this report, positive coefficients in the 0.40's, 0.50's and 0.60's are designated as "moderate"; coefficients below or above this range are designated accordingly as "low" or "high".

This moderate level of intercorrelations, however, is not unexpected given the similarity of some of the skills. The overriding consideration is whether the similarity of the skills resulted in such uniformly high intercorrelation coefficients as to warrant the deletion of some skills from the set. A factor analysis of the nine skills revealed that this is not the case. Although there is a degree of communality, the nine skill ratings of the MAC essentially provide unique information on the subjects' performance.

Intercorrelations of the Criterion

The intercorrelations of the criterion refers to the correlations between the dimension ratings within each criterion instrument. As with the predictor, the presence of uniformly high intercorrelations between the dimensions of a criterion instrument may denote communality. A more likely interpretation, however, is that it denotes the existence of a halo effect. A halo effect occurs when a high rating on one dimension influences the rater to give the subject comparable ratings on the other dimensions. Since the criterion raters (unlike the MAC assessors) were not trained in observing and rating a subject's performance, they could conceivably make such a mistake. Therefore, this interpretation of uniformly high coefficients is logical. Yet it can be problematic, for the subject under scrutiny may indeed be superior in all dimensions.

The MAC evaluation revealed that the dimension ratings provided by the subjects exhibit comparably low intercorrelation coefficients on both the Job Performance Scale and the Effectiveness Scale.[9] This indicates that, compared to the other rater groups, the subjects generally saw greater differences in their own performances on the various dimensions. This result, incidently, is consistent with the findings of the Schmitt et al. (1983) study. Secondly, the dimension ratings provided by the other three rater groups exhibit intercorrelation coefficients that range from moderate to moderate-to-high, but they are never so uniformly high as to preclude the possibility of separate interpretations of the various dimensions. Therefore, there is no evidence of undue influence exerted on the dimension ratings by a halo effect. This result is also consistent with the findings of the Schmitt et al. (183) study.

Inter-Rater Reliability of the Predictor

The inter-rater reliability of the predictor is an index of the degree of agreement in the skill ratings provided by the MAC assessors. This index is important, because in an assessment center high inter-rater reliability is considered a prerequisite to validity. The inter-rater reliability index of the MAC was computed using two methods, the Pearson r and the coefficient alpha. Table 6 displays the Pearson r correlations of the skill ratings provided by the three MAC assessors. Each assessor was arbitrarily designated Assessor 1, 2 or 3; and

[9] The intercorrelations of the dimension ratings can not be computed for the Overall Rating Item, since this criterion instrument consists of a single dimension.

correlations were computed for each of the nine skills assessed at the MAC. An examination of Table 6 reveals that the correlations are high with 62.9% of the coefficients falling in the 0.80's. These results are indicative of a high degree of agreement among the MAC assessors.

Table 6

Inter-Rater Reliability of the MAC Skill Ratings Based on the Pearson r

1. Leadership

Assessor			
1			
2	.90		
3	.88	.89	
	1	2	3

2. Organizing and Planning

Assessor			
1			
2	.88		
3	.87	.87	
	1	2	3

3. Perception

Assessor			
1			
2	.87		
3	.81	.86	
	1	2	3

4. Decision Making

Assessor			
1			
2	.88		
3	.82	.86	
	1	2	3

5. Decisiveness

Assessor			
1			
2	.85		
3	.85	.87	
	1	2	3

6. Interpersonal

Assessor			
1			
2	.77		
3	.71	.73	
	1	2	3

7. Adaptability

Assessor			
1			
2	.73		
3	.74	.74	
	1	2	3

8. Oral Communication

Assessor			
1			
2	.72		
3	.74	.76	
	1	2	3

9. Written Communication

Assessor			
1			
2	.86		
3	.89	.86	
	1	2	3

Note. All coefficients are significant at $p \leq 0.05$ for $N = 106$.

Table 7

Inter-Rater Reliability of the MAC Skill Ratings Based on Coefficient Alpha

1.	Leadership	.96
2.	Organizing and Planning	.95
3.	Perception	.94
4.	Decision Making	.94
5.	Decisiveness	.94
6.	Interpersonal	.89
7.	Adaptability	.89
8.	Oral Communication	.89
9.	Written Communication	.95

Table 7 displays the coefficient alpha of the skill ratings provided by the MAC assessors. Coefficient alpha, which ranges in value from zero to one, is generally utilized as a measure of agreement among items in a written test. In applying coefficient alpha to this situation, each MAC skill was regarded as a single test and each assessor as a single item within the test. An examination of Table 7 reveals that all the MAC skills exhibit a coefficient alpha of 0.89 or higher. As in Table 6, these results are indicative of a high degree of agreement among the MAC assessors.

Since the Schmitt et al. (1983) study used the same two methods to compute the inter-rater reliability index, it facilitated a comparison of the results of the study and the MAC evaluation. Before the comparison is made, however, two differences should be noted. First, since the NASSP centers routinely collect the data necessary to compute the inter-rater reliability index, Schmitt was able to draw from the centers' records and base his computation on 340 subjects. The MAC, however, does not routinely collect this data. The inter-rater reliability index of the MAC is based on only 106 subjects, because the data needed to compute the index was collected solely for this evaluation. Secondly, the NASSP assessment centers employ six assessors per assessment, while the MAC employs three.

The comparison of the Schmitt study and the MAC evaluation reveal that the inter-rater reliability index computed by means of the Pearson r is higher for the MAC. Schmitt reported that "most of the correlations were in excess of 0.60", while most of the MAC correlations were in excess of 0.80. In terms of the coefficient alpha, it should be acknowledged that the inter-rater reliabil-

ity index computed by this method is enhanced by an increase in the number of assessors (i.e., the items in the test). Nevertheless, despite the disparity in this regard, the results were quite similar. Schmitt reported a minimum coefficient alpha of 0.90 for all skills, which is essentially equal to the 0.89 of the MAC.

Consistency of the Criterion

Table 8 displays the correlations between the composite ratings of the criterion instruments (i.e., the sum of all the dimension ratings within an instrument). These correlations provide information on the consistency of the ratings across the instruments. Thus, they reflect to a degree the reliability of the criterion.

Table 8

Correlations Between the Composite Ratings of the Criterion Instruments

A. Supervisor

	J	E	O
J			
E	.76 (110)		
O	.75 (109)	.89 (109)	

C. Subordinates

	J	E	O
J			
E	.78 (99)		
O	.09[a] (97)	.30 (98)	

B. Support Persons

	J	E	O
J			
E	.77 (76)		
O	.69 (76)	.87 (76)	

D. Subjects

	J	E	O
J			
E	.37 (115)		
O	.42 (114)	.81 (115)	

Note. All coefficients are significant at $p \leq 0.05$ unless designated otherwise. The n for each cell appears in parentheses. J = Job Performance Scale; E = Effectiveness Scale; O = Overall Rating Item.

[a]Not significant at $p \leq 0.05$.

A review of Table 8 reveals that the consistency correlations of the composite ratings provided by both the supervisors and the support persons are generally high. By contrast, the correlations of the composite ratings provided by both the subjects and the subordinates exhibit some low correlations. These low correlations, however, are not unexpected. In the case of the subjects, the low consistency correlations are generally a reflection of the previously noted low intercorrelations of the ratings provided by this group. In the case of the subordinates, the low consistency correlations are attributable to this group's propensity for opting not to respond to certain items in the instruments. Of the four rater groups, it should be noted that the subordinates invoked this option most often. This is understandable, given that the subordinates probably had the least direct contact with the subjects. While invoking this option does not affect the correlations involving individual dimension ratings, a single "missing" dimension rating can adversely affect the correlations involving the composite ratings. For example, the composite ratings of the Job Performance Scale and the Overall Rating Item correlated only 0.09 (see Table 8, C). Yet, the correlations between the individual dimension ratings of these respective instruments are moderate with 80.0% of the coefficients falling in the 0.40's and 0.50's.

In brief, the consistency correlations of the criterion instruments displayed in Table 8 exhibit some low coefficients. These coefficients, however, are attributable at least in part to a procedural factor in the computation. Generally the correlations in the table are high with 58.3% of the coefficients falling in the 0.70's and 0.80's. This represents positive evidence of the consistency of the ratings across the criterion instruments.

Validity Correlations

The correlations between the predictor and the criterion represent the most crucial statistical relationship in this phase of the evaluation. These correlations, known as validity correlations, are essentially a measure of how accurately the MAC predicts the subjects' subsequent job performance. In computing the validity correlations, the MAC overall skill ratings of the subjects function as the predictor, and the subjects' subsequent job performance ratings function as the criterion. Of the 121 subjects assessed, 47 were hired as school-level administrators during the 1982-83 school year, and 54 during the 1983-84. In addition, the group included 20 subjects who, although unsuccessful at the MAC, were incumbent school-level administrators. The job performance ratings were obtained in June of 1983 for the original 47 subjects, and again in June of 1984 for the entire group of 121 subjects.

The validity correlations based on the June of 1984 job performance ratings of all 121 subjects are displayed in Tables 9 through 12 and Tables 14 through 18. Specifically, Tables 9 through 12 display the correlations between the overall skill ratings of the MAC and the dimension ratings on the Job Performance Scale provided respectively by the supervisors, the support persons, the subordinates and the subjects. Tables 14 through 17 display equivalent data for the Effectiveness Scale; and Table 18 displays equivalent data for the Overall Rating Item.

A review of Tables 9 through 12 reveals certain differences in the validity correlations based on the Job Performance Scale across rater groups. Focusing first on the supervisors, the validity correlations in Table 9 exhibit a positive statistical relationship between the MAC overall skill ratings and the dimension ratings provided by this group. Almost all the coefficients in this table are positive, and 37.0% of them are statistically significant.[10]

[10] A statistically significant relationship is defined in this report as one that would occur only 5 times or less in 100 by pure chance, i.e., $p \leq 0.05$.

Table 9

Correlations Between the MAC Overall Skill Ratings and the Dimension Ratings Provided by the Supervisors on the Job Performance Scale

Skills/Dimensions:	1. Curriculum Objectives	2. Curriculum: Individual Progress	3. Student Activities: Supervision	4. Student Activities: Participation	5. Support Services	6. Directing Student Behavior	7. Staff Evaluation	8. Developmental Activities	9. Community Relations	10. Interpersonal Effectiveness	11. Community Relations: Parents	12. Coordination with District	13. Fiscal Management	14. Maintenance of School Plant	15. Structures Communication
1. Leadership	.10 (102)	.13 (96)	.17 (91)	.18* (96)	.11 (99)	.04 (101)	.00 (106)	.00 (103)	.10 (101)	.10 (106)	.09 (102)	.11 (98)	.05 (85)	.04 (99)	.09 (104)
2. Organizing and Planning	.15 (102)	.22* (96)	.25* (91)	.22* (96)	.15 (99)	.14 (101)	.13 (106)	.09 (103)	.17* (101)	.19* (106)	.21* (102)	.16* (98)	.14 (85)	.07 (99)	.17* (104)
3. Perception	.12 (102)	.20* (96)	.18* (91)	.30* (96)	.20* (99)	.20* (101)	.14 (106)	.10 (103)	.20* (101)	.19* (106)	.17* (102)	.16 (98)	.12 (85)	.12 (99)	.17* (104)
4. Decision Making	.14 (102)	.21* (96)	.24* (91)	.25* (96)	.15 (99)	.22* (101)	.16* (106)	.10 (103)	.17* (101)	.20* (106)	.15 (102)	.21* (98)	.11 (85)	.11 (99)	.18* (104)
5. Decisiveness	.08 (102)	.16 (96)	.06 (91)	.15 (96)	.07 (99)	.06 (101)	.04 (106)	.07 (103)	.09 (101)	.10 (106)	.10 (102)	.06 (98)	.08 (85)	.08 (99)	.04 (104)
6. Interpersonal	.22* (102)	.15 (96)	.32* (91)	.20* (96)	.23* (99)	.25* (101)	.15 (106)	.06 (103)	.21* (101)	.13 (106)	.09 (102)	.19* (98)	.09 (85)	.06 (99)	.21* (104)
7. Adaptability	.16* (102)	.20* (96)	.22* (91)	.26* (96)	.24* (99)	.26* (101)	.16* (106)	.11 (103)	.27* (101)	.23* (106)	.20* (102)	.20* (98)	.06 (85)	.12 (99)	.21* (104)
8. Oral Communication	.17* (102)	.22* (96)	.17 (91)	.16 (96)	.15 (99)	.13 (101)	.05 (106)	.06 (103)	.14 (101)	.07 (106)	.09 (102)	.10 (98)	.01 (85)	.04 (99)	.13 (104)
9. Written Communication	.08 (102)	.18* (96)	.13 (91)	.09 (96)	.02 (99)	.04 (101)	.01 (106)	.07 (103)	.06 (101)	.00 (106)	.09 (102)	.10 (98)	-.02 (85)	.01 (99)	.09 (104)

Note. The n for each cell appears in parentheses.

*p ≤ 0.05.

Table 10

Correlations Between the MAC Overall Skill Ratings and the Dimension Ratings Provided by the Support Persons on the Job Performance Scale

Skills/Dimensions:	1. Curriculum Objectives	2. Curriculum: Individual Progress	3. Student Activities: Supervision	4. Student Activities: Participation	5. Support Services	6. Directing Student Behavior	7. Staff Evaluation	8. Developmental Activities	9. Community Relations	10. Interpersonal Effectiveness	11. Community Relations: Parents	12. Coordination with District	13. Fiscal Management	14. Maintemance of School Plant	15. Structures Communication
1. Leadership	.34* (67)	.35* (59)	.22* (63)	.17 (69)	.13 (61)	.26* (70)	.31* (72)	.20* (66)	.14 (64)	.18 (72)	.22* (65)	.25* (62)	.25* (52)	.19 (63)	.27* (71)
2. Organizing and Planning	.34* (67)	.18 (59)	.13 (63)	.11 (69)	.10 (61)	.16 (70)	.23* (72)	.22* (66)	.17 (64)	.17 (72)	.17 (65)	.18 (62)	.28* (52)	.14 (63)	.21* (71)
3. Perception	.31* (67)	.20 (59)	.18 (63)	.18 (69)	.04 (61)	.23* (70)	.12 (72)	.13 (66)	.15 (64)	.17 (72)	.16 (65)	.18 (62)	.16 (52)	.06 (63)	.25* (71)
4. Decision Making	.35* (67)	.20 (59)	.22* (63)	.16 (69)	.10 (61)	.15 (70)	.18 (72)	.13 (66)	.05 (64)	.22* (72)	.12 (65)	.26* (62)	.22 (52)	.07 (63)	.25* (71)
5. Decisiveness	.29* (67)	.10 (59)	.12 (63)	.15 (69)	.00 (61)	.13 (70)	.14 (72)	.13 (66)	.20* (64)	.14 (72)	.10 (65)	.19 (62)	.18 (52)	.18 (63)	.19 (71)
6. Interpersonal	.27* (67)	.22* (59)	.07 (63)	.02 (69)	-.12 (61)	.10 (70)	.14 (72)	.04 (66)	-.03 (64)	.09 (72)	.06 (65)	.05 (62)	.04 (52)	.04 (63)	.28* (71)
7. Adaptability	.27* (67)	.23* (59)	.11 (63)	.13 (69)	-.02 (61)	.23* (70)	.24* (72)	.13 (66)	.08 (64)	.27* (72)	.12 (65)	.11 (62)	.02 (52)	.08 (63)	.30* (71)
8. Oral Communication	.31* (67)	.28* (59)	.11 (63)	.16 (69)	.06 (61)	.23* (70)	.17 (72)	.21* (66)	.07 (64)	.15 (72)	.04 (65)	.15 (62)	.13 (52)	.08 (63)	.23* (71)
9. Written Communication	.20* (67)	.28* (59)	.03 (63)	.05 (69)	.01 (61)	.25* (70)	.19 (72)	.21* (66)	.07 (64)	.08 (72)	.09 (65)	.14 (62)	.15 (52)	.15 (63)	.24* (71)

Note. The n for each cell appears in parentheses.

*$p \leq 0.05$.

Table 11

Correlations Between the MAC Overall Skill Ratings and the Dimension Ratings Provided by the Subordinates on the Job Performance Scale

Skills/Dimensions:	1. Curriculum Objectives	2. Curriculum: Individual Progress	3. Student Activities: Supervision	4. Student Activities: Participation	5. Support Services	6. Directing Student Behavior	7. Staff Evaluation	8. Developmental Activities	9. Community Relations	10. Interpersonal Effectiveness	11. Community Relations: Parents	12. Coordination with District	13. Fiscal Management	14. Maintenance of School Plant	15. Structures Communication
1. Leadership	.01 (75)	.07 (41)	.07 (40)	.03 (55)	.07 (41)	.02 (75)	-.07 (73)	.03 (66)	-.05 (29)	-.08 (86)	-.14 (48)	-.30 (16)	-.17 (24)	.03 (56)	-.22* (69)
2. Organizing and Planning	.15 (75)	.07 (41)	.25 (40)	.21 (55)	.10 (41)	.07 (75)	-.01 (73)	.12 (66)	.15 (29)	-.04 (86)	-.03 (48)	.03 (16)	.05 (24)	.24* (56)	-.18 (69)
3. Perception	.11 (75)	.06 (41)	.12 (40)	.12 (55)	-.09 (41)	-.03 (75)	-.08 (73)	.09 (66)	.00 (29)	-.09 (86)	-.14 (48)	.00 (16)	-.30 (24)	.06 (56)	-.18 (69)
4. Decision Making	.07 (75)	.16 (41)	.00 (40)	.11 (55)	.08 (41)	.10 (75)	-.01 (73)	.19 (66)	.05 (29)	.00 (86)	-.02 (48)	.06 (16)	.03 (24)	.10 (56)	-.09 (69)
5. Decisiveness	.17 (75)	.25 (41)	.02 (40)	.14 (55)	.04 (41)	.08 (75)	-.02 (73)	.06 (66)	.09 (29)	.00 (86)	-.02 (48)	-.12 (16)	.17 (24)	.10 (56)	-.17 (69)
6. Interpersonal	.20* (75)	.11 (41)	.06 (40)	.09 (55)	.10 (41)	.09 (75)	.03 (73)	.06 (66)	.24 (29)	-.02 (86)	-.08 (48)	.04 (16)	.05 (24)	.09 (56)	-.07 (69)
7. Adaptability	.05 (75)	.11 (41)	-.20 (40)	-.07 (55)	.00 (41)	.09 (75)	-.08 (73)	-.05 (66)	.02 (29)	-.02 (86)	-.15 (48)	-.06 (16)	-.11 (24)	-.01 (56)	-.21* (69)
8. Oral Communication	.28* (75)	.22 (41)	.00 (40)	.09 (55)	.13 (41)	.26* (75)	.07 (73)	.17 (66)	.11 (29)	.00 (86)	-.05 (48)	.09 (16)	-.06 (24)	.11 (56)	-.10 (69)
9. Written Communication	.14 (75)	.22 (41)	.09 (40)	.05 (55)	.10 (41)	.05 (75)	.00 (73)	.08 (66)	.11 (29)	-.01 (86)	-.05 (48)	-.14 (16)	.00 (24)	.00 (56)	-.16 (69)

Note. The n for each cell appears in parentheses.

*$p \leq 0.05$.

Table 12

Correlations Between the MAC Overall Skill Ratings and the Dimension Ratings Provided by the Subjects on the Job Performance Scale

Skills/Dimensions:	1. Curriculum Objectives	2. Curriculum: Individual Progress	3. Student Activities: Supervision	4. Student Activities: Participation	5. Support Services	6. Directing Student Behavior	7. Staff Evaluation	8. Developmental Activities	9. Community Relations	10. Interpersonal Effectiveness	11. Community Relations: Parents	12. Coordination with District	13. Fiscal Management	14. Maintenance of School Plant	15. Structures Communication
1. Leadership	.10 (107)	-.07 (98)	.15 (97)	.10 (107)	.06 (110)	.07 (114)	.02 (113)	-.06 (111)	-.03 (111)	.04 (114)	-.08 (108)	.04 (110)	-.10 (93)	.10 (110)	-.13 (113)
2. Organizing and Planning	.13 (107)	.06 (98)	.19* (97)	.14 (107)	.11 (110)	.04 (114)	.21* (113)	.00 (111)	.08 (111)	.13 (114)	.00 (108)	.06 (110)	-.10 (93)	.03 (110)	-.08 (113)
3. Perception	.08 (107)	.04 (98)	.10 (97)	.13 (107)	.00 (110)	.05 (114)	.16* (113)	-.08 (111)	.03 (111)	.06 (114)	-.03 (108)	-.05 (110)	-.10 (93)	.07 (110)	-.08 (113)
4. Decision Making	.14 (107)	-.10 (98)	.15 (97)	.03 (107)	.04 (110)	.11 (114)	.12 (113)	-.01 (111)	.03 (111)	.11 (114)	-.07 (108)	.02 (110)	-.13 (93)	.06 (110)	-.04 (113)
5. Decisiveness	.16* (107)	.01 (98)	.10 (97)	.04 (107)	.07 (110)	.08 (114)	.18* (113)	-.05 (111)	.02 (111)	.10 (114)	.00 (108)	.03 (110)	-.16 (93)	.07 (110)	-.06 (113)
6. Interpersonal	.11 (107)	.03 (98)	.04 (97)	.03 (107)	.02 (110)	-.04 (114)	.11 (113)	-.09 (111)	-.05 (111)	-.02 (114)	-.13 (108)	-.08 (110)	-.19* (93)	-.03 (110)	-.06 (113)
7. Adaptability	.14 (107)	-.02 (98)	-.04 (97)	.00 (107)	-.03 (110)	.00 (114)	.08 (113)	-.07 (111)	-.04 (111)	-.01 (114)	-.18* (108)	.00 (110)	-.16 (93)	-.09 (110)	-.07 (113)
8. Oral Communication	.05 (107)	.04 (98)	.03 (97)	-.02 (107)	-.06 (110)	.00 (114)	.00 (113)	-.13 (111)	-.09 (111)	.08 (114)	-.10 (108)	.02 (110)	-.17* (93)	-.10 (110)	-.10 (113)
9. Written Communication	.03 (107)	-.02 (98)	.13 (97)	-.03 (107)	-.05 (110)	.00 (114)	.01 (113)	-.08 (111)	-.02 (111)	.11 (114)	-.01 (108)	.06 (110)	-.25* (93)	-.07 (110)	-.09 (113)

Note. The n for each cell appears in parentheses.

*$p \leq 0.05$.

Likewise the validity correlations in Table 10 exhibit a positive statistical relationship between the MAC overall skill ratings and the dimension ratings provided by the support persons. Almost all the coefficients in this table are positive, and 32.5% of them are statistically significant.

Unlike Tables 9 and 10, the validity correlations in Table 11 exhibit no statistical relationship between the MAC overall skill ratings and the dimension ratings provided by the subordinates on the Job Performance Scale. Although 62.2% of all the coefficients in this table are positive, very few of them are statistically significant. Nevertheless, given the tendency of the values of validity correlations coefficients to rise over time, the preponderance of positive coefficients in Table 11 suggests that these validity correlations may in time exhibit a positive statistical relationship.

The validity correlations in Table 12 exhibit no statistical relationship between the MAC overall skill ratings and the dimension ratings provided by the subjects on the Job Performance Scale. The number of positive coefficients in this table is approximately equal to the number of negative coefficients. Therefore, unlike Table 11, there is no evidence in Table 12 to suggest that these validity correlations may eventually exhibit a positive statistical relationship.

Since the validity correlations of the Schmitt et al. (1983) study are based on the Job Performance Scale, this affords an opportunity to compare the results depicted in Tables 9 through 12 with the results of the Schmitt study. Before this comparison is made, however, certain differences between the Schmitt study and the MAC evaluation should be noted. First, the rater groups consisting of support persons are not comparable. In the Schmitt study this group consisted of "secretaries, janitorial staff, etc.", while in the MAC evaluation it consists of assistant principals. Likewise, the rater groups consisting of subordinates are not comparable. In the Schmitt study this group consisted of designated "senior level [teachers] who were likely to observe and be familiar with the [subject's] work". Although in the MAC evaluation this group also consists of teachers, they were selected randomly. Finally, the job tenure of the subjects hired after a favorable assessment is not comparable. In the Schmitt study the tenure of these subjects ranged up to four years, while in the MAC evaluation it only ranged up to two years.

Despite this difference in tenure, an examination of Table 13 reveals that the proportion of positive and statistically significant validity correlations in the MAC evaluation compares favorably with the Schmitt study. Table 13 also reveals that the MAC evaluation and the Schmitt study concur on the statistical relationship derived from the ratings provided by the two comparable rater groups, i.e., the supervisors and the subjects. In the MAC evaluation, the validity correlations based on the supervisors' ratings exhibit a positive statistical relationship, while the validity correlations based on the subjects' ratings exhibit no statistical relationship; the Schmitt study yielded the same results.

By contrast, the MAC evaluation and the Schmitt study differ on the statistical relationship derived from the ratings provided by the rater groups which were not comparable, i.e., the support persons and the subordinates. In the case of the support persons, the validity correlations based on this group's ratings exhibit a positive statistical relationship in the MAC evaluation, but in the Schmitt study they exhibit no statistical relationship. These different results are understandable, since the support persons in the MAC evaluation (i.e., assistant principals) generally have a greater knowledge of the subject's job and more direct contact with the subject than the support persons in the Schmitt

Table 13

Comparison of the Validity Correlations Based on the Job Performance Scale in the MAC Evaluation[a] and the Schmitt Study[b]

Rater Group/Study:	MAC	Schmitt	MAC	Schmitt
	% of Negative r		% of Positive r	
Supervisors[c]	0.7	19.4	99.2	80.5
Support Persons	2.2	38.8	97.7	61.1
Subordinates	37.7	21.1	62.2	78.8
Subjects[c]	46.6	35.0	53.3	65.0
	% of Negative r Significant at $p \leq 0.05$		% of Positive r Significant at $p \leq 0.05$	
Supervisors[c]	0	0.5	37.0	46.6
Support Persons	0	2.[illegible]	32.5	7.2
Subordinates	1.4	2.2	2.9	20.0
Subjects[c]	2.9	7.2	3.7	13.3

[a]The validity correlations of the MAC evaluation consist of 135 coefficients per rater group.

[b]The validity correlations of the Schmitt et al. (1983) study consist of 180 coefficients per rater group.

[c]Comparable rater group.

study (i.e., secretaries, janitors, etc.). In addition, the different results exhibited by the validity correlations based on the ratings provided the subordinates (i.e., teachers) may also be attributable to a disparity in the degree of job knowledge and subject contact. The validity correlations based on the ratings provided by the designated senior level teachers in the Schmitt study exhibit a positive statistical relationship; however, those based on the ratings provided by the randomly selected teachers in the MAC evaluation exhibit no statistical relationship.

Tables 14 through 17 are the next set of tables to be reviewed. These tables display the validity correlations based on the Effectiveness Scale, which is the second criterion instrument utilized in the MAC evaluation. Focusing first on the supervisors, the validity correlations in Table 14 exhibit a positive statistical relationship between the MAC overall skill ratings and the dimensions ratings provided by this group. Almost all the coefficients in this table are positive, and 38.1% of them are statistically significant. Likewise, the validity correlations in Table 15 exhibit a positive statistical relationship between the MAC overall skill ratings and the dimension ratings provided by the support persons. Indeed, the validity correlations in this table exhibit the strongest positive statistical relationship in the MAC evaluation. All the coefficients in the table are positive, and 73.0% of them are statistically significant.

Unlike Tables 14 and 15, the validity correlations in Table 16 exhibit no statistical relationship between the MAC overall skill ratings and the dimension ratings provided by the subordinates on the Effectiveness Scale. Although 70.6% of the coefficients in this table are positive, very few are statistically significant. This preponderance of positive coefficients, which is similar to the results on the Job Performance Scale (see Table 11), suggests that in time these validity correlations may exhibit a positive statistical relationship.

The validity correlations in Table 17 exhibit no statistical relationship between the MAC overall skill ratings and the dimension ratings provided by the subjects on the Effectiveness Scale. The number of positive coefficients in this table is approximately equal to the number of negative coefficients. Therefore, unlike Table 16, there is no evidence in Table 17 to suggest that these validity correlations may eventually exhibit a positive statistical relationship.

Table 18 is the last table to be reviewed in this section. It displays the validity correlations based on the Overall Rating Item, which is the third criterion instrument utilized in this study. Since this instrument consists of a single dimension, it is possible to display the validity correlations derived from all four rater groups in this single table. A review of Table 18 reveals that the validity correlations in the first and second columns, which are based on the dimension ratings provided by the supervisors and support persons respectively, exhibit positive statistical relationships. All the coefficients in these two columns are positive, and 66.6% of the coefficients in each column are statistically significant. By contrast, the validity correlations in the third and fourth columns, which are based on the dimension ratings provided by the subordinates and subjects respectively, exhibit no statistical relationships.

In summary, the results of the MAC evaluation reveal that the validity correlations based on the supervisors' ratings exhibit a positive statistical relationship in all three criterion instruments (see Tables 9, 14 and 18). Furthermore, the validity correlations based on the support persons' ratings exhibit comparable results (see Tables 10, 15 and 18). By contrast, the validity correlations

Table 14

Correlations Between the MAC Overall Skill Ratings and the Dimension Ratings Provided by the Supervisors on the Effectiveness Scale

Skills/Dimensions:	1. Adaptability	2. Conferring with Parents	3. Decision Making	4. Decisiveness	5. Interpersonal	6. Management Skill	7. Observing Teachers	8. Oral Communication	9. Paper Work	10. Planning and Organizing	11. Perception	12. Leadership	13. Technical Know-How	14. Written Communication
1. Leadership	-.01 (106)	.15 (106)	.00 (105)	.08 (105)	.07 (105)	.07 (104)	.10 (104)	.16* (104)	.08 (104)	.02 (105)	.07 (105)	.11 (106)	.01 (105)	.15 (106)
2. Organizing and Planning	.11 (106)	.19* (106)	.06 (105)	.11 (105)	.18* (105)	.15 (104)	.22* (104)	.29* (104)	.19* (104)	.09 (105)	.13 (105)	.20* (106)	.06 (105)	.33* (106)
3. Perception	.12 (106)	.18* (106)	.09 (105)	.18* (105)	.13 (105)	.15 (104)	.17* (104)	.16* (104)	.13 (104)	.12 (105)	.20* (105)	.22* (106)	.11 (105)	.25* (106)
4. Decision Making	.13 (106)	.20* (106)	.12 (105)	.22* (105)	.15 (105)	.19* (104)	.27* (104)	.17* (104)	.18* (104)	.14 (105)	.17* (105)	.22* (106)	.13 (105)	.29* (106)
5. Decisiveness	.01 (106)	.13 (106)	.00 (105)	.10 (105)	.09 (105)	.09 (104)	.15 (104)	.13 (104)	.09 (104)	.02 (105)	.08 (105)	.12 (106)	.03 (105)	.22* (106)
6. Interpersonal	.14 (106)	.13 (106)	.16* (105)	.19* (105)	.17* (105)	.17* (104)	.24* (104)	.19* (104)	.14 (104)	.17* (105)	.14 (105)	.14 (106)	.12 (105)	.24* (106)
7. Adaptability	.20* (106)	.20* (106)	.13 (105)	.18* (105)	.27* (105)	.22* (104)	.25* (104)	.24* (104)	.15 (104)	.14 (105)	.20* (105)	.24* (106)	.14 (105)	.26* (106)
8. Oral Communication	.10 (106)	.13 (106)	.03 (105)	.11 (105)	.11 (105)	.16* (104)	.13 (104)	.27* (104)	.10 (104)	.09 (105)	.14 (105)	.15 (106)	.09 (105)	.23* (106)
9. Written Communication	.04 (106)	.06 (106)	-.03 (105)	.04 (105)	-.01 (105)	.06 (104)	.12 (104)	.19* (104)	.06 (104)	.04 (105)	.08 (105)	.09 (106)	.05 (105)	.26* (106)

Note. The n for each cell appears in parentheses.
*$p \leq 0.05$.

Table 15

Correlations Between the MAC Overall Skill Ratings and the Dimension Ratings Provided by the Support Persons on the Effectiveness Scale

Skills/Dimensions:	1. Adaptability	2. Conferring with Parents	3. Decision Making	4. Decisiveness	5. Interpersonal	6. Management Skill	7. Observing Teachers	8. Oral Communication	9. Paper Work	10. Planning and Organizing	11. Perception	12. Leadership	13. Technical Knowledge	14. Written Communication
1. Leadership	.32* (73)	.24* (72)	.25* (73)	.30* (72)	.24* (73)	.34* (72)	.36* (70)	.41* (72)	.34* (73)	.32* (73)	.27* (72)	.31* (71)	.31* (72)	.37* (72)
2. Organizing and Planning	.30* (73)	.24* (72)	.17 (73)	.20* (72)	.26* (73)	.30* (72)	.33* (70)	.37* (72)	.24* (73)	.26* (73)	.16 (72)	.25* (71)	.20* (72)	.36* (72)
3. Perception	.30* (73)	.26* (72)	.22* (73)	.21* (72)	.23* (73)	.24* (72)	.25* (70)	.33* (72)	.22* (73)	.24* (73)	.24* (72)	.22* (71)	.24* (72)	.31* (72)
4. Decision Making	.31* (73)	.15 (72)	.18 (73)	.19 (72)	.27* (73)	.24* (72)	.25* (70)	.27* (72)	.21* (73)	.24* (73)	.13 (72)	.19* (71)	.16 (72)	.26* (72)
5. Decisiveness	.29* (73)	.19 (72)	.19 (73)	.19 (72)	.22* (73)	.25* (72)	.18 (70)	.25* (72)	.24* (73)	.23* (73)	.13 (72)	.24* (71)	.15 (72)	.28* (72)
6. Interpersonal	.18 (73)	.11 (72)	.13 (73)	.13 (72)	.14 (73)	.22* (72)	.24* (70)	.15 (72)	.14 (73)	.23* (73)	.04 (72)	.08 (71)	.09 (72)	.16 (72)
7. Adaptability	.34* (73)	.20* (72)	.26* (73)	.21* (72)	.30* (73)	.31* (72)	.25* (70)	.26* (72)	.21* (73)	.26* (73)	.20* (72)	.23* (71)	.22* (72)	.27* (72)
8. Oral Communication	.26* (73)	.27* (72)	.18 (73)	.19 (72)	.26* (73)	.27* (72)	.26* (70)	.32* (72)	.24* (73)	.33* (73)	.10 (72)	.22* (71)	.18 (72)	.34* (72)
9. Written Communication	.19* (73)	.21* (72)	.12 (73)	.26* (72)	.15 (73)	.26* (72)	.15 (70)	.28* (72)	.25* (73)	.21* (73)	.04 (72)	.19 (71)	.17 (72)	.31* (72)

Note. The n for each cell appears in parentheses.

*$p \leq 0.05$.

Table 16

Correlations Between the MAC Overall Skill Ratings and the Dimension Ratings Provided by the Subordinates on the Effectiveness Scale

Skills/Dimensions:	1. Adaptability	2. Conferring with Parents	3. Decision Making	4. Decisiveness	5. Interpersonal	6. Management Skill	7. Observing Teachers	8. Oral Communication	9. Paper Work	10. Planning and Organizing	11. Perception	12. Leadership	13. Technical Know-How	14. Written Communication
1. Leadership	.00 (92)	-.09 (57)	.03 (80)	-.09 (84)	.05 (90)	-.10 (73)	-.14 (58)	.08 (93)	-.19 (55)	-.05 (52)	-.04 (73)	.03 (78)	.09 (56)	.00 (78)
2. Organizing and Planning	.09 (92)	-.03 (57)	.12 (80)	.03 (84)	.07 (90)	.03 (73)	.00 (58)	.21* (93)	-.08 (55)	.08 (52)	.03 (73)	.13 (78)	.20 (56)	.16 (78)
3. Perception	.01 (92)	-.08 (57)	.05 (80)	-.01 (84)	.02 (90)	-.06 (73)	-.11 (58)	.03 (93)	-.17 (55)	.01 (52)	-.03 (73)	.05 (78)	.12 (56)	-.01 (78)
4. Decision Making	.14 (92)	.00 (57)	.07 (80)	.03 (84)	.10 (90)	-.01 (73)	.00 (58)	.17* (93)	-.03 (55)	.04 (52)	.05 (73)	.11 (78)	.15 (56)	.10 (78)
5. Decisiveness	.14 (92)	.00 (57)	.15 (80)	.04 (84)	.09 (90)	.01 (73)	.00 (58)	.14 (93)	-.03 (55)	.07 (52)	.05 (73)	.09 (78)	.10 (56)	.12 (78)
6. Interpersonal	.12 (92)	.00 (57)	.14 (80)	.04 (84)	.04 (90)	-.07 (73)	.03 (58)	.07 (93)	-.17 (55)	.07 (52)	.12 (73)	.10 (78)	.15 (56)	.01 (78)
7. Adaptability	.01 (92)	-.08 (57)	.06 (80)	-.10 (84)	.04 (90)	-.19* (73)	-.10 (58)	.11 (93)	-.20 (55)	-.05 (52)	.09 (73)	-.05 (78)	-.05 (56)	.03 (78)
8. Oral Communication	.15 (92)	.05 (57)	.32* (80)	.13 (84)	.14 (90)	.10 (73)	.07 (58)	.17* (93)	.09 (55)	.22 (52)	.19* (73)	.18 (78)	.24* (56)	.17 (78)
9. Written Communication	.10 (92)	.04 (57)	.10 (80)	-.03 (84)	.07 (90)	.00 (73)	-.02 (58)	.14 (93)	.01 (55)	.08 (52)	.03 (73)	.06 (78)	.10 (56)	.12 (78)

Note. The n for each cell appears in parentheses.

*p ≤ 0.05.

Table 17

Correlations Between the MAC Overall Skill Ratings and the Dimension Ratings Provided by the Subjects on the Effectiveness Scale

Skills/Dimensions:	1. Adaptability	2. Conferring with Parents	3. Decision Making	4. Decisiveness	5. Interpersonal	6. Management Skill	7. Observing Teachers	8. Oral Communication	9. Paper Work	10. Planning and Organizing	11. Perception	12. Leadership	13. Technical Know-How	14. Written Communication
1. Leadership	-.09 (115)	.03 (115)	.00 (115)	.00 (115)	-.10 (115)	-.03 (115)	-.03 (114)	.06 (114)	.06 (116)	-.07 (116)	.08 (116)	-.03 (116)	-.06 (115)	.09 (116)
2. Organizing and Planning	-.01 (115)	.10 (115)	.02 (115)	.01 (115)	-.08 (115)	-.04 (115)	.00 (114)	.10 (114)	.10 (116)	.00 (116)	.07 (116)	.01 (116)	-.02 (115)	.13 (116)
3. Perception	.01 (115)	.16* (115)	.02 (115)	.01 (115)	-.04 (115)	-.01 (115)	.01 (114)	.12 (114)	.07 (116)	-.03 (116)	.14 (116)	.04 (116)	.04 (115)	.11 (116)
4. Decision Making	.00 (115)	.05 (115)	-.06 (115)	.00 (115)	-.07 (115)	-.05 (115)	.00 (114)	.05 (114)	.00 (116)	-.05 (116)	.03 (116)	-.05 (116)	.02 (115)	.11 (116)
5. Decisiveness	-.03 (115)	.07 (115)	.01 (115)	-.01 (115)	-.08 (115)	.03 (115)	-.01 (114)	.12 (114)	.04 (116)	-.06 (116)	.06 (116)	.02 (116)	.00 (115)	.15* (116)
6. Interpersonal	-.15* (115)	.03 (115)	-.12 (115)	-.03 (115)	-.13 (115)	-.14 (115)	-.04 (114)	.15 (114)	.00 (116)	-.12 (116)	-.02 (116)	-.07 (116)	.00 (115)	.15* (116)
7. Adaptability	.00 (115)	.00 (115)	-.08 (115)	-.09 (115)	.00 (115)	-.12 (115)	-.08 (114)	.08 (114)	-.01 (116)	-.11 (116)	-.07 (116)	.01 (116)	.00 (115)	.10 (116)
8. Oral Communication	-.06 (115)	-.01 (115)	-.08 (115)	-.06 (115)	-.11 (115)	-.09 (115)	-.12 (114)	.15 (114)	.04 (116)	-.05 (116)	-.07 (116)	-.08 (116)	.01 (115)	.19* (116)
9. Written Communication	-.13 (115)	-.02 (115)	-.12 (115)	-.11 (115)	-.12 (115)	-.04 (115)	-.16* (114)	.13 (114)	.09 (116)	-.07 (116)	-.08 (116)	-.11 (116)	.00 (115)	.24* (116)

Note. The n for each cell appears in parentheses.

*p ≤ 0.05.

Table 18

Correlations Between the MAC Overall Skill Ratings and the Dimension Rating Provided by Each Rater Group on the Overall Rating Item

Skills/Rater Groups:	Supervisors	Support Persons	Subordinates	Subjects
1. Leadership	.08 (105)	.36* (73)	-.19* (95)	-.02 (115)
2. Organizing and Planning	.23* (105)	.29* (73)	.02 (95)	.07 (115)
3. Perception	.23* (105)	.23* (73)	.10 (95)	.09 (115)
4. Decision Making	.26* (105)	.16 (73)	.02 (95)	.10 (115)
5. Decisiveness	.14 (105)	.19 (73)	.02 (95)	.12 (115)
6. Interpersonal	.22* (105)	.14 (73)	-.06 (95)	.00 (115)
7. Adaptability	.24* (105)	.19* (73)	-.05 (95)	.03 (115)
8. Oral Communication	.22* (105)	.24* (73)	.02 (95)	.00 (115)
9. Written Communication	.11 (105)	.20* (73)	.01 (95)	-.01 (115)

Note. The n for each cell appears in parentheses.

*$p \leq 0.05$

based on the subordinates ratings exhibit only a preponderance of positive coefficients in the Job Performance Scale and the Effectiveness Scale. While this situation suggests that a positive statistical relationship may in time be exhibited, at the present the validity correlations based on the subordinates' ratings exhibit no statistical relationship in all three instruments (see Table 11, 16 and 18). In the case of the subjects, the validity correlations based on this group's ratings exhibit likewise no statistical relationship in all three instruments (see Table 12, 17 and 18). Finally, a comparison between the results of the MAC evaluation based on the Job Performance Scale and the results of the Schmitt et al. (1983) study revealed that the two studies concur on the statistical relationship derived from the ratings provided by the two comparable rater groups, i.e., the supervisors and the subjects. Furthermore, despite the greater job tenure of the subjects in the Schmitt study, the proportion of positive and statistically significant validity correlations in the MAC evaluation compares favorably with the Schmitt study (see Table 13).

Longitudinal View of the Validity Correlations

Since the job performance ratings of the original 47 subjects were obtained both in June of 1983 and June of 1984, it afforded an opportunity to view a portion of the validity correlations of the MAC longitudinally. The specific focus was on the previously noted tendency of validity correlations to rise over time. To determine if the MAC exhibits this tendency, the composite validity correlations for 1983 and 1984 were computed. The composite validity correlations are the correlations between the sum of the MAC overall skill ratings and the sum of the dimension ratings on each criterion instrument provided by each rater group. The increases in these correlations from 1983 to 1984 were tested for statistical significance utilizing a test of significance of the difference between correlation coefficients for two correlated samples. The test yielded no statistical significance.

Despite the lack of significance in the increases of the composite validity correlations, there was a discernible rise in the validity correlations of the MAC between 1983 and 1984. Tables 19 through 21 display the 1983 and 1984 correlations between the MAC overall skill ratings and the composite dimension ratings on the three criterion instruments. An examination of these tables reveals that the positive coefficients across the tables rose from 55.5% in 1983 to 68.5% in 1984. Consequently, there is some evidence that the validity correlations of the MAC have risen over time.

Check for Criterion Contamination

Since the MAC is an operational assessment center, the subjects' skill ratings were not kept completely confidential. Criterion contamination could have contributed to the positive statistical relationships exhibited by the validity correlations. To rule out this possibility, the evaluation incorporated a procedure suggested by the Huck and Bray (1976) study as a check for criterion contamination. This procedure is essentially a comparison of the criterion ratings provided by the raters who were aware of the assessment center ratings with the criterion ratings provided by the raters who were unaware. The purpose of the comparison was to ascertain if the knowledge of the assessment center ratings biased the criterion ratings and produced inflated validity correlations.

Of the four rater groups in the MAC study, only the subjects and the supervisors had access to the MAC skill ratings. The subjects were routinely informed of their skill ratings as part of the assessment process. This knowledge, however, did not result in inflated validity correlations. Indeed, the validity correlations based on the criterion ratings provided by the subjects exhibited no

Table 19

Correlations Between the MAC Overall Skill Ratings and the 1983 and 1984 Composite Ratings on the Job Performance Scale for the Original 47 Subjects

Skills/Year and Rater Groups:	1983 Supervisors	1983 Support Persons	1983 Subordinates	1983 Subjects	1984 Supervisors	1984 Support Persons	1984 Subordinates	1984 Subjects
1. Leadership	.25* (43)	.19 (28)	.21 (35)	.11 (46)	.06 (42)	.36* (28)	.00 (38)	-.05 (45)
2. Organizing and Planning	.04 (43)	-.22 (28)	.05 (35)	.07 (46)	-.01 (42)	.16 (28)	-.04 (38)	-.01 (45
3. Perception	.18 (43)	-.06 (28)	.16 (35)	-.01 (46)	.09 (42)	.22 (28)	.11 (38)	.09 (45)
4. Decision Making	.12 (43)	-.12 (28)	.17 (35)	-.08 (46)	-.05 (42)	.24 (28)	-.04 (38)	.03 (45)
5. Decisiveness	.06 (43)	-.11 (28)	.10 (35)	-.05 (46)	-.06 (42)	.15 (28)	-.11 (38)	.03 (45)
6. Interpersonal	.13 (43)	.04 (28)	-.03 (35)	-.04 (46)	.01 (42)	.26 (28)	.06 (38)	-.05 (45)
7. Adaptability	.06 (43)	-.06 (28)	-.01 (35)	-.16 (46)	.11 (42)	.30 (28)	-.04 (38)	-.09 (45)
8. Oral Communication	.04 (43)	-.01 (28)	.11 (35)	-.32* (46)	-.01 (42)	.47* (28)	.08 (38)	-.04 (45)
9. Written Communication	.07 (43)	.01 (28)	-.04 (35)	-.17 (46)	-.05 (42)	.36* (28)	-.02 (38)	.16 (45)

Note. The n for each cell appears in parentheses.

*$p \leq 0.05$.

Table 20

Correlations Between the MAC Overall Skill Ratings and the 1983 and 1984 Composite Ratings on the Effectiveness Scale for the Original 47 Subjects

Skills/Year and Rater Groups:	1983 Supervisors	1983 Support Persons	1983 Subordinates	1983 Subjects	1984 Supervisors	1984 Support Persons	1984 Subordinates	1984 Subjects
1. Leadership	.30* (43)	.22 (28)	-.05 (35)	.03 (47)	.16 (42)	.38* (28)	-.06 (38)	-.11 (45)
2. Organizing and Planning	.13 (43)	-.12 (28)	-.09 (35)	.02 (47)	.02 (42)	.21 (28)	-.04 (38)	-.10 (45)
3. Perception	.13 (43)	-.01 (28)	.15 (35)	.12 (47)	.10 (42)	.16 (28)	.00 (38)	-.06 (45)
4. Decision Making	.02 (43)	-.05 (28)	.22 (35)	.04 (47)	.01 (42)	.21 (28)	.00 (38)	-.08 (45)
5. Decisiveness	-.01 (43)	-.14 (28)	.19 (35)	.02 (47)	.00 (42)	.00 (28)	-.06 (38)	.00 (45)
6. Interpersonal	.23 (43)	.00 (28)	.03 (35)	.00 (47)	.09 (42)	.13 (28)	.12 (38)	-.12 (45)
7. Adaptability	.13 (43)	-.03 (28)	-.02 (35)	-.06 (47)	.24 (42)	.17 (28)	-.07 (38)	-.03 (45)
8. Oral Communication	.17 (43)	.02 (28)	.01 (35)	.00 (47)	.18 (42)	.32* (28)	.13 (38)	.00 (45)
9. Written Communication	.06 (43)	-.02 (28)	-.03 (35)	.13 (47)	.02 (42)	.28 (28)	-.07 (38)	.20 (45)

Note. The n for each cell appears in parentheses.

*$p \leq 0.05$.

Table 21

Correlations Between the MAC Overall Skill Ratings and the 1983 and 1984 Composite Ratings on the Overall Rating Item for the Original 47 Subjects

Skills/Year and Rater Groups:	1983 Supervisors	1983 Support Persons	1983 Subordinates	1983 Subjects	1984 Supervisors	1984 Support Persons	1984 Subordinates	1984 Subjects
1. Leadership	.30* (43)	.19 (28)	.11 (30)	-.08 (46)	.13 (41)	.27 (28)	-.18 (36)	-.12 (44)
2. Organizing and Planning	.08 (43)	-.14 (28)	.08 (30)	-.11 (46)	.12 (41)	.19 (28)	.00 (36)	.00 (44)
3. Perception	.02 (43)	-.27 (28)	.15 (30)	-.10 (46)	.15 (41)	.15 (28)	.02 (36)	.01 (44)
4. Decision Making	.00 (43)	-.22 (28)	.33* (30)	-.21 (46)	.07 (41)	.12 (28)	.06 (36)	.09 (44)
5. Decisiveness	.02 (43)	-.31 (28)	.29 (30)	-.06 (46)	.00 (41)	-.14 (28)	-.03 (36)	.21 (44)
6. Interpersonal	.16 (43)	-.27 (28)	.04 (30)	-.23 (46)	.05 (41)	.14 (28)	.00 (36)	.07 (44)
7. Adaptability	.11 (43)	-.20 (28)	.18 (30)	-.27* (46)	.30* (41)	.08 (28)	.05 (36)	.08 (44)
8. Oral Communication	.17 (43)	-.20 (28)	.29 (30)	-.28* (46)	.19 (41)	.28 (28)	-.04 (36)	.04 (44)
9. Written Communication	.00 (43)	-.17 (28)	.15 (30)	-.07 (46)	.08 (41)	.24 (28)	-.01 (36)	.24 (44)

Note. The n for each cell appears in parentheses.

*$p \leq 0.05$.

statistical relationship (see Tables 12, 17 and 18). Thus, the check for criterion contamination in this situation is not warranted. Furthermore, it is not possible, since the process of routinely informing all these subjects of their skill ratings precludes the application of the procedure.

The check for criterion contamination, therefore, was confined to the supervisors. This group of raters was not routinely informed of the subjects' skill ratings, but they did have access to them by virtue of their positions within the school system. Consequently, prior to rating the subjects' job performance, the supervisors were asked: "Are you aware of the skill ratings the [subject] received from the Management Assessment Center?" Of the 108 supervisors providing criterion ratings, 42 responded "yes" to this question and 66 responded "no". Based on these responses, the subjects were divided into two groups. One group consisted of the subjects rated by the 42 "aware" supervisors, and the other group consisted of the subjects rated by the 66 "unaware" supervisors.

The comparison of the two groups was based on the data displayed in Table 22. They include the means and standard deviations of the composite overall skill ratings of the MAC, the means and standard deviations of the composite dimension ratings of each criterion instrument, and the validity correlations based on these ratings. The differences between the corresponding values of the predictor and the criterion ratings were tested for significance utilizing analysis of variance. The test revealed that the differences are not statistically significant. This indicates that based on the predictor ratings these two groups of subjects performed similarly at the MAC, and this similarity in performance was reflected in the subsequent criterion ratings. The fact that the criterion ratings provided by the aware supervisors did not differ significantly from those provided by the unaware is evidence of a lack of criterion contamination.

Additional evidence is provided by the coefficient values of the validity correlations, which are the reverse of what would be expected had there been criterion contamination. The validity correlations based on the dimension ratings provided by the unaware supervisors are actually higher than those based on the aware supervisors. Criterion contamination, therefore, could not have inflated the validity correlations. To rule out the unlikely possibility that it deflated the validity correlations, the differences in the corresponding coefficient values were tested for significance utilizing Fisher's Z transformation in a test of significance between the correlation coefficients for two independent samples. The test revealed that the differences are not statistically significant. Thus, the application of the procedure suggested by the Huck and Bray (1976) study reveals no evidence of criterion contamination in the ratings provided by the supervisors.

Comparison of the Selection Processes: Results

The final phase of the MAC evaluation involved a comparison of the district's present selection process for school-level administrators with the former selection process. The former selection process essentially consisted of a series of interviews for the qualified candidates. The present selection process differs in the use of the MAC to screen the qualified candidates prior to the interviews. The subjects in the comparison consisted of 109 incumbent school-level administrators. The interview-only process was used to select 62 of them. These 62 subjects were hired during the 1981-82 school year, the last year the interview-only selection process was operational. The remaining 47 subjects were hired the following school year using the MAC-interview selection process. To equate the difference in job tenure between the two groups, the comparison was staggered in time. The June of 1983 job performance ratings of the interview-only subjects were compared to the June of 1984 job performance ratings of the MAC-interview subjects.

Table 22

Check for Criterion Contamination

Ratings and Correlations	Subjects Rated by: Aware Supervisors n = 42	Unaware Supervisors n = 66
Composite Overall Skill Rating of MAC		
Mean	42.4	44.8
Standard Deviation	11.7	10.9
Job Performance Scale Composite Rating		
Mean	63.2	66.2
Standard Deviation	14.6	16.0
Effectiveness Scale Composite Rating		
Mean	53.4	55.3
Standard Deviation	10.9	11.6
Overall Rating Item		
Mean	7.2	7.2
Standard Deviation	1.6	1.8
Correlation of the Composite Overall Skill Rating of MAC and the Composite Rating of:		
Job Performance Scale	.05	.11
Effectiveness Scale	.08	.24
Overall Rating Item	.19	.24

Table 23 displays the job performance ratings for the two groups of subjects, as well as the results of the statistical comparison of these values utilizing analysis of variance (ANOVA). Specifically, the second column in Table 23 displays the means of the composite job performance ratings of the 62 interview-only subjects provided by the four rater groups on the three criterion instruments. The third column displays equivalent data for the 47 MAC-interview subjects. The fourth column displays the F value resulting from the statistical comparison of each set of means using ANOVA. The final column in the table displays the significance of each F value.[11]

A review of the second and third column in Table 23 reveals a pattern of consistently higher job performance ratings given to the MAC-interview subjects by the supervisors and support persons. This pattern is of interest, because, as noted in the second phase of the evaluation, the job performance ratings provided by the supervisors and the support persons exhibit criterion-related validity. Nevertheless, the job performance ratings of the MAC-interview subjects are not of sufficient magnitude to preclude the possibility that their advantage over the ratings of the interview-only subjects resulted from pure chance. Indeed, an examination of the final column in Table 23 reveals that none of the differences in the means are statistically signficant at a probability value of $p \leq 0.05$. This indicates that based on the job performance ratings neither group of subjects exhibits a clear advantage over the other.

[11] The significance of the F value refers to the probability that the difference between the two means was due to pure chance.

Table 23

Comparison of the Job Performance Ratings of the Interview-Only Subjects and the MAC-Interview Subjects

	Mean Composite Ratings		ANOVA	
Supervisors' Ratings on the:	Interview-Only Subjects	MAC-Interview Subjects	F	Significance of F
Job Performance Scale	7.52 (56)	7.81 (43)	1.097	.297
Effectiveness Scale	65.80 (56)	70.33 (43)	2.204	.141
Overall Rating Item	56.73 (56)	58.95 (43)	1.269	.263
Support Persons' Ratings on the:				
Job Performance Scale	7.48 (42)	7.48 (29)	.000	.986
Effectiveness Scale	65.90 (42)	67.83 (29)	.190	.664
Overall Rating Item	54.81 (42)	57.38 (29)	.680	.412
Subordinates' Ratings on the:				
Job Performance Scale	8.31 (55)	8.25 (44)	.143	.706
Effectiveness Scale	76.02 (55)	75.39 (44)	.054	.816
Overall Rating Item	62.73 (55)	63.27 (44)	.165	.685
Subjects' Ratings on the:				
Job Performance Scale	8.31 (55)	8.25 (44)	.143	.706
Effectiveness Scale	76.02 (55)	75.39 (44)	.054	.816
Overall Rating Item	62.73 (55)	63.27 (44)	.165	.685

Note: The n for the cell appears in parentheses.

CONCLUSIONS

The conclusions of the evaluations of the Management Assessment Center (MAC) will be addressed in this section of the report. It should be noted that this section adheres to the same format employed in the Design of the Evaluation and the Results of the Evaluation, in that the conclusions of the three phases of the evaluation are addressed separately.

Interview and Survey of the MAC Assessors: Conclusions

Since the interview and the survey of the MAC assessors were conducted after the completion of assessment cycle 1 in 1981-82, the conclusions of this phase of the evaluation were presented in the preliminary report which was published in March of 1984. Basically the preliminary report noted that during the first year of operation, the MAC had experienced some start-up problems. The MAC staff, however, had been very responsive in addressing these problems, and thus had facilitated the subsequent development of the MAC. Consequently, the MAC assessors, who were in a unique position to observe the operation of the center, were very supportive of both the MAC staff and the MAC process. Beyond these general conclusions, the preliminary report presented a number of ancillary conclusions which will not be replicated in this report. A reader who wishes to review all the conclusions of this phase of the evaluation should contact the Office of Educational Accountability and request a copy of Preliminary Report on the Evaluation of the Management Assessment Center.

Criterion-Related Validity of the MAC: Conclusions

The second phase of the evaluation involved the criterion-related validity of the MAC. The design of this phase encompassed several data analysis procedures; they included: (a) the intercorrelations of the predictor, (b) the intercorrelations of the criterion, (c) the inter-rater reliability of the predictor, (d) the consistency of the criterion, (e) the validity correlations, (f) the longitudinal view of the validity correlations, and (g) the check for criterion contamination. The results of these data analysis procedures yielded the following conclusions:

- A degree of communality is revealed by the intercorrelation of the nine MAC skill ratings. This communality implies that some of the skills are tapping a common factor in the subject's performance. This situation, however, is not unusual given the similarity of some of the skills. Indeed, a moderate level of intercorrelation under the circumstances is to be expected. The overriding consideration is whether the intercorrelations are so uniformly high as to suggest that some of the skill ratings are redundant. A factor analysis revealed that this is not the case with the MAC skill ratings. Although they exhibit a degree of communality, each skill is essentially providing unique information on the subject's performance.

- As with the MAC skill ratings, uniformly high intercorrelations of the job performance dimension ratings may denote a communality across the dimensions. A more likely interpretation, however, is that it denotes the existence of a halo effect. A halo effect occurs when a high rating on one dimension influences the rater to give the subject comparable ratings on the other dimensions. Since the job performance raters (unlike the MAC assessors) were not trained in observing and rating a subject's performance, they could conceivably make such a mistake. An examination of the

Intercorrelations of the dimensions reveals that coefficients range from low to moderate-to-high, but they are never so high as to suggest that each dimension rating is not providing unique information on the subject's job performance.

- The most important measure of the reliability of an assessment center is the inter-rater reliability, which is an index of the degree of agreement among the assessors on the ratings of the subjects. This index is important, because in an assessment center high inter-rater reliability is considered a prerequisite to validity. The inter-rater reliability of the MAC was computed using two methods, the Pearson r and coefficient alpha. Both methods yielded a high index of inter-rater reliability, indicating a high degree of agreement among the MAC assessors. Depending on the method used to compute the index, the inter-rater reliability of the MAC assessors is either equal to or higher than the inter-rater reliability of the NASSP assessment center assessors reported in the Schmitt et al. (1983) study.

- The correlation of the composite job performance ratings on each criterion instrument provides a measure of the consistency of the ratings. These correlations of consistency yielded comparably higher coefficients for the job performance ratings provided by the supervisors and the support persons than the subordinates and the subjects. Nevertheless, on the whole the correlations of consistency were high, which reflects favorably on the reliability of the three criterion instruments.

- The positive statistical relationships exhibited by the validity correlations indicates that the MAC did to a degree predict the subsequent job performance of the subjects. Furthermore, a comparison of the results of the MAC evaluation with the Schmitt et al. (1983) study reveals that the proportion of positive and statistically significant validity correlations are comparable. This outcome is impressive, because the tendency of the values of validity correlations coefficents to rise over time and the shorter job tenure of the MAC subjects should have combined to yield comparably lower validity correlations for the MAC.

- Two sets of validity correlations were computed for the original 47 subjects using the job performance ratings obtained in June of 1983 and June of 1984. A comparison of these two sets of validity correlations reveals a discernible rise in the values of the coefficients over the one year span. While the difference in values is not statistically significant, it does concur with several studies that have noted this tendency of the values of the validity correlations coefficients to rise over time.

- Since the MAC is an operational assessment center, the subjects' skill ratings were not kept completely confidential. If the criterion raters were aware of the subjects' MAC skill ratings, it could have contributed to the positive statistical relationship exhibited by the validity correlations. Of the four groups of criterion raters, only the supervisors and the subjects themselves had access to the MAC skill ratings. The subjects were routinely informed of their ratings during the assessment process. However, the validity correlations based on the job performance ratings provided by this group exhibited no statistical relationship, so criterion contamination was not an issue. The supervisors, although not routinely informed of the MAC skill ratings, had access to them by virtue of their positions within the school system. Consequently, the check for criterion contamination suggested by the Huck and Bray (1976) study was

applied to this group. It yielded no evidence of criterion contamination. Thus, having knowledge of the subjects' MAC skill ratings did not bias the supervisors when they rated the job performance of the subjects.

Comparison of the Selection Processes: Conclusions

The final phase of the MAC evaluation involved the comparison of the district's present selection process for school-level administrators with the former selection process. The former selection process essentially consisted of a series of interviews for the qualified candidates. The present selection process differs in the use of the MAC to screen the qualified candidates prior to the interviews. The results of the comparison of the two selection processes yielded the following conclusions:

- The comparison of the job performance ratings of the MAC-interview subjects and the interview-only subjects revealed that the MAC-interview process is not superior to the interview-only process in the selection of better school-level administrators. Despite the validity of the MAC, its incorporation into the selection process is not justified by the selection of better candidates. Thus, it must be concluded that under the existing operating procedures the MAC has no utility.[12]
- The unsatisfactory performance of the MAC-interview selection process can be attributed to two possibilities. The first possibility is that the MAC and the district's interview procedure have comparable validity. Consequently, nothing was gained by incorporating the MAC into the selection process. This possibility is highly unlikely, since research has shown that the interview method only exhibits validity correlations coefficients in the 0.10's (Reilly & Chao, 1982). The MAC's current validity correlation coefficients based on the ratings of the supervisors and support persons are in the 0.20's, and there is evidence that they are still rising. A more likely possibility is that, with the existing high passing rate, the MAC results are not really being used. In other words, the minimum passing score of the MAC is such that the few candidates who are eliminated from consideration would probably have been eliminated anyway by the interviews. Under the circumstances, the interviews in effect became the overriding factor in both selection processes. Thus, there was no advantage in incorporating the MAC into the selection process, not because of a deficiency in its validity, but because its validity was essentially not used.

In brief, the evaluation of the MAC revealed that after a job tenure of two years or less, the correlations between the subjects' MAC skill ratings and the job performance ratings exhibited criterion-related validity. Furthermore, the longitudinal trend of the validity correlations, as well as the high inter-rater reliability of the MAC assessors, are indications that the values of validity correlations coefficients will continue to rise. How high they will go can only be speculated at this time. However, if the research in the field is an indication, the validity correlation coefficients should be in the low 0.40's when the subjects achieve a job tenure of approximately five years. Nevertheless, the MAC's level of validity in the future or in the present are moot issues, if the

[12] Utility refers to "the determination of expected institutional gain or loss (outcomes) anticipated from various courses of actions "(Cascio, 1982). The fundamental question of utility analysis is: Are the costs of a course of action justified by the returns?

passing rate remains so high as to preclude the center's utility. The importance of this issue cannot be overstated. Perhaps Dr.[13] Neal Schmitt phrased it best when he made the following comment about the MAC:

> Your pass rate, as well as those of other centers, is really arbitrary and can be manipulated. Actually, the optimal use of the assessment center data ... is to pick the highest ranking individuals. Use of [a high] success rate is tantamount to ignoring the results of the [MAC] (N. Schmitt, personal communication, March 23, 1984).

[13] Dr. Neal Schmitt, who headed the validity study of the NASSP assessment centers, served as a consultant in the evaluation of the MAC.

RECOMMENDATIONS

The conclusions drawn from the evaluation of the MAC serve as a basis for the following recommendations:

- The MAC will not have utility until the minimum passing score is raised. It is, therefore, recommended that a qualified consultant be retained to advise the district on this matter. The employment of a consultant is recommended, because a higher minimum passing score will likely increase the adverse impact of the MAC. Consequently, it would be prudent to insure that all actions taken regarding this matter are legally defensible.

The raising of the minimum passing score of the MAC is the overriding recommendation of this evaluation. The remaining recommendations are all contingent on the implementation of this recommendation.

- Assuming an appropriate upward adjustment of the minimum passing score, it is recommended that the district retain the MAC as a part of its selection process of school-level administrators. This recommendation is based on the established criterion-related validity of the MAC, as well as the demonstrated competence of the MAC staff. The MAC clearly offers the district a method of improving the effectiveness of the existing selection process.

- The existing procedure of the MAC does not provide the data needed to compute the inter-rater reliability of the assessors. As previously noted, the index of inter-rater reliability is the most important measure of the reliability of an assessment center. This index is useful in monitoring the training and performance of the assessors. For this reason, the MAC procedure was temporarily altered to provide the evaluation with this index. And, despite the fact that the evaluation revealed the index to be high at this time, this situation could change. Therefore, it is suggested that provisions be made to compute the inter-rater reliability index of the MAC whenever there is a substantive change in the training procedure or in the cadre of assessors.

BIBLIOGRAPHY

Boehm, V.R. (1982). Establishing the validity of assessment centers. Pittsburgh: Development Dimensions Press.

Bray, D.W., & Grant, D.L. (1966). The assessment center in measurement of potential for business management. Psychological Monographs, 80, (17, Whole No. 625).

Bureau of Staff Development. (1985). Management assessment in the Dade County Public Schools the fourth cycle. Miami: Dade County Public Schools.

Cascio, W.F. (1982). Costing human resources: The financial impact of behavior in organizations. Boston: Kent Publishing.

Division of Industrial - Organizational Psychology. (1980). Principals for the validation and use of personnel selection procedures. (2nd ed.). Dayton: American Psychological Association.

Equal Employment Opportunity Commission. (1978). Uniform guidelines on employee selection procedures. Federal register. Washington, D.C.: Civil Service Commission, Department of Labor, & Department of Justice.

Hersey, P.W. (1977). NASSP's assessment center - from concept to practice. NASSP Bulletin, 61, 74-76.

Hinrichs, J.R. (1978). An eight-year follow-up of a management assessment center. Journal of Applied Psychology, 63, 596-601.

Huck, J.R., & Bray, D.W. (1976). Management assessment center evaluations and subsequent job performance of black and white females. Personnel Psychology, 29, 13-30.

Mitchel, J.O. (1975). Assessment center validity: A longitudinal study. Journal of Applied Psychology, 60, 573-579.

Office of Educational Accountability. (1984). Preliminary report on the evaluation of the Management Assessment Center. Miami: Dade County Public Schools.

Ogawa, R.T. (1984, April). NASSP assessment centers: Opportunities for research. Paper presented at the Annual Meeting of the American Educational Research Association, New Orleans, LA.

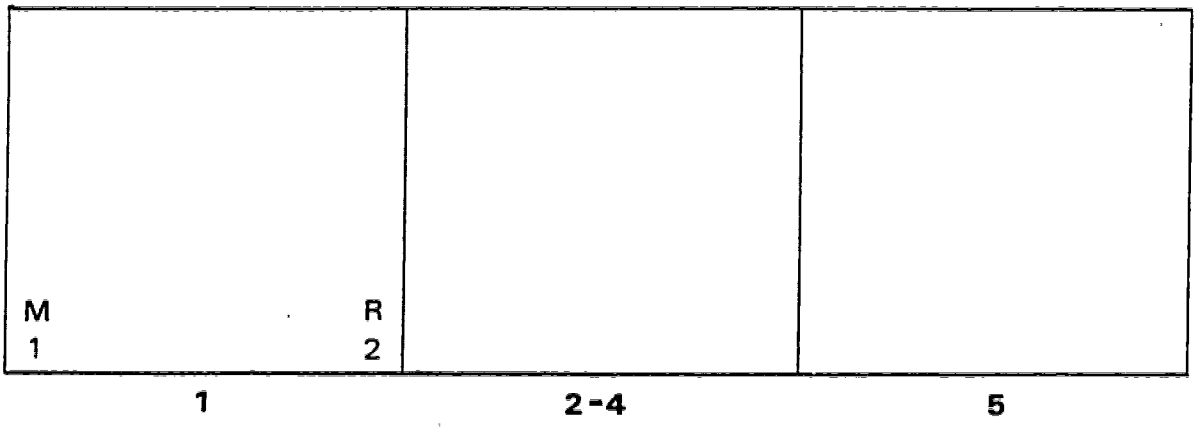
Reilly, R.R., & Chao, G.T. (1982). Validity and fairness of some alternative employee selection procedures. Personnel Psychology, 35, 1-62.

Schmitt, N., Noe, R., Meritt, R., Fitzgerald, M., & Jorgensen, C. (1983). Criterion-related and content validity of the NASSP assessment center. Reston, VA.: National Association of Secondary School Principals.

Task Force on Assessment Center Standards. (1980). Standards and ethical considerations for assessment center operations. The Personnel Administrator, 25, 35-38.

APPENDIX

M/R RATING INSTRUMENT



INCUMBENT/RATER CODE

DADE COUNTY PUBLIC SCHOOLS

OFFICE OF EDUCATIONAL ACCOUNTABILITY

SECTION A: RATER DATA

RATER DATA

Instructions: All raters should respond to items 1 through 7.

DO NOT WRITE IN THIS AREA

1. Rater's age: _____ years

2. Sex: female ____ (1) male ____ (2) — 6 - 7

3. Ethnicity: — 8

White, not of Hispanic origin	____ (1)
Black, not of Hispanic origin	____ (2)
Hispanic	____ (3)
Asian or Pacific Islander	____ (4)
American Indian or Alaskan native	____ (5)

9

4. Current Work Location:

Elementary School	____ (1)
Junior High School	____ (2)
Senior High School	____ (3)
Area Office	____ (4)
Central Office	____ (5)

10

5. Current position:

Teacher	____ (1)
Asst. Principal	____ (2)
Principal	____ (3)
Director	____ (4)
Other: (Please specify.) _____	____ (5)

11

6. Number of years with the Dade County Public Schools: _____ years — 12-13

7. Are you rating yourself (i.e., the rater and the incumbent are the same person)? ____ (1) yes ____ (2) no — 14

Instructions: Item 8 is only applicable to incumbents coded M (see box 1 on the cover). Only a rater who is currently a principal or an area director should respond to item 8. A principal who is rating him/herself, however, should not respond to this item.

DO NOT WRITE IN THIS AREA

8. Are you aware of skill ratings the incumbent received from the Management Assessment Center? yes___ no___
   1 2

15

Instructions: If you are not rating yourself, skip items 9 through 12 and go directly to page 5. If you are rating yourself, please respond to items 9 through 12.

9. Were you evaluated at the Management Assessment Center as part of the selection procedure for your current job? yes___ no___
   1 2

16

10. How many months have you held your current job? (Include summer months regardless of employment status. Include time served on an "interim" basis, if applicable.) ____month/s

11. What was the last job you held prior to assuming your current job?

17-18

Teacher	1
Department Head	2
C.S.I. Director	3
Activities Director	4
Counselor	5
Assistant to the Principal	6
Assistant Principal	7
Assistant Principal at a different school level	8
Principal at a different school level	9
Other: (Please specify.)_____	10

19-20

21 - 22

12. How many years did you hold your previous job identified in item 11? ____year/s

SECTION B: JOB PERFORMANCE SCALE

JOB PERFORMANCE SCALE

Instructions: This rating scale is designed to measure the important dimensions of a principal's or assistant principal's job. You are to use the scale to evaluate the job performance of a designated incumbent. Each dimension of the scale should be considered separately and carefully, as some incumbents may do well on one dimension but not another. These ratings may include some areas of the job which you do not see the incumbent doing or for which you do not have enough information to base a decision. In this situation, feel free to indicate you are unable to rate the dimension in question.

Each dimension has a rating scale which is numbered from 1 (the lowest rating) to 6 (the highest rating). Examples of job performance behavior associated with various numerical ratings are provided. These are actual examples obtained from teachers, students, principals, support staff, and parents in a national study of school principals and assistant principals. To make a particular rating does not imply that the incumbent actually does or has done the things described by that rating. The examples are used to illustrate the kinds of behavior that can be expected from an incumbent who is typically rated high, average or low on each dimension. Thus, the examples represent a point of reference to help you decide how to rate the incumbent.

The actual rating procedure will be to read each dimension and the examples of job performance behaviors. If you feel that you have insufficient information to rate the dimension, check the box at the top, right-hand side. Otherwise decide which numerical rating best reflects the incumbent's job performance, and check the appropriate box on the right. Note that there is only one dimension per page, so CHECK ONLY ONE BOX PER PAGE.

1. CURRICULUM AND INSTRUCTIONAL LEADERSHIP: MONITORING CURRICULUM OBJECTIVES

CHECK ONLY ONE RESPONSE

If you have insufficient information to rate this dimension, check this box. ☐

RATING

6 ☐ High

Examples of Behavior of Individuals who are Usually Rated "High" on this Dimension:

- Takes a stand on issues even when they are controversial; supports and stands up for educational values.
- Constructs a schedule which maximizes student options, yet minimizes course conflicts.
- Visits classrooms to monitor the curriculum actually being taught in school.

5 ☐

4 ☐

Examples of Behavior of Individuals who are Usually Rated "Average" on this Dimension:

- Reads necessary materials to have a thorough knowledge of state or district curriculum standards and requirements.
- Looks at new texts so that he/she can influence the selection of materials and textbooks.
- Reviews and makes comments on weekly lesson plans in terms of overall curriculum objectives.

Averag

3 ☐

Examples of Behavior of Individuals who are Usually Rated "Low" on this Dimension:

2 ☐

- Does not utilize special curriculum programs provided by the district.
- Adopts curriculum materials without the input of teachers or in spite of their recommendations.
- Does not allow any curricular experimentation by faculty members.

1 ☐ Low

DO NOT WRITE IN THIS AREA

23

2. CURRICULUM AND INSTRUCTIONAL LEADERSHIP: MONITORING INDIVIDUAL PROGRESS

CHECK ONLY ONE RESPONSE

If you have insufficient information to rate this dimension, check this box. ☐

RATING

6 ☐ High

Examples of Behavior of Individuals who are Usually Rated "High" on this Dimension:

- Initiates a program with the help of student-parent associations to allow students to take advanced classes offered at nearby schools.
- Organizes student help sessions which meet after school hours for students who are failing.

5 ☐

Examples of Behavior of Individuals who are Usually Rated "Average" on this Dimension:

- Reviews records of individual student progress including standardized scores and basic skills program.
- Institutes a classroom where several things are taught (e.g., ceramics, painting) because of a lack of enough interest in any one of them to justify individual classes.
- Sponsors a biweekly reading rally in which all students stop what they are doing and read for 20 minutes.

4 ☐

Averag

3 ☐

Examples of Behavior of Individuals who are Usually Rated "Low" on this Dimension:

- Allows students to hand carry scores of certain tests home; students are able to open these unsealed envelopes and compare scores.
- Schedules all advanced courses in the morning so that academic students can take only 2 out of 5.

2 ☐

1 ☐ Low

DO NOT WRITE IN THIS AREA

3. COORDINATION OF STUDENT ACTIVITIES: SUPERVISION

CHECK ONLY ONE RESPONSE

If you have insufficient information to rate this dimension, check this box. ☐

RATING

Examples of Behavior of Individuals who are Usually Rated "High" on this Dimension:

- Initiates an awards banquet and establishes a letter for superior academic performance.
- Evaluates all activities and student needs with input from faculty, student council, and student surveys, then acts on results.
- Meets regularly with student leaders to coordinate activities and take suggestions.
- Reschedules a homecoming event which had been scheduled for the night before the SAT.

Examples of Behavior of Individuals who are Usually Rated "Average" on this Dimension:

- Schedules a sporting event during the academic day for the benefit of bused and/or poor students who cannot come to night activities.
- Organizes parent-student teams to work on fund raising activities.
- Organizes faculty teams to play against varsity teams.

Examples of Behavior of Individuals who are Usually Rated "Low on this Dimension:

- Organizes a student fund raising activity without talking with students.
- Cancels a pep meeting because of a drinking party by a few students.
- Allows extra-curricular activity and school classes to conflict so that students have to be pulled out of class to participate in band or sports.

6 ☐ High

5 ☐

4 ☐ Average

3 ☐

2 ☐

1 ☐ Low

DO NOT WRITE IN THIS AREA

25

4. STUDENT ACTIVITIES: PARTICIPATION

CHECK ONLY ONE RESPONSE

If you have insufficient information to rate this dimension, check this box. ☐

RATING

6 ☐ High

Examples of Behavior of Individuals who are Usually Rated "High" on this Dimension:

- Initiates a program so that handicapped students are incorporated into various activities, such as plays or sports.
- Participates in extracurricular school activity by actually working at the function, such as fun fair, school dinners.
- Provides student government officers with the opportunity to solicit student participation.

5 ☐

Examples of Behavior of Individuals who are Usually Rated "Average" on this Dimension:

4 ☐

- Talks to students who are not participating in extracurricular activities to get them involved.
- Attends extracurricular activities.
- Encourages staff to participate in extracurricular activities but does not demand such participation.

Averag

3 ☐

Examples of Behavior of Individuals who are Usually Rated "Low" on this Dimension:

- Tells the students that certain activities are important and then doesn't support them by his/her presence or allowing time in the schedule for them.
- Selectively participates in only certain school activities (e.g., foot ball).

2 ☐

1 ☐ Low

DO NOT WRITE IN THIS AREA

5. DIRECTION OF SUPPORT SERVICES

CHECK ONLY ONE RESPONSE

If you have insufficient information to rate this dimension, check this box. ☐

RATING

6 ☐ High

Examples of Behavior of Individuals who are Usually Rated "High" on this Dimension:

- Devises a system whereby librarian is able to coordinate library resources (e.g., books, magazines) with needs of teachers as indicated by lesson plans.
- Includes clerical/custodial staff in all school parties and meetings.

5 ☐

- Verbally acknowledges the completion of tasks by school maintenance and food services personnel.

Examples of Behavior of Individuals who are Usually Rated "Average" on this Dimension:

4 ☐

- Initiates a program to have parents receive first aid training so there can be personnel in the clinic during the lunch hour.
- Organizes groups of support services people so that they have a more unified voice.
- Asks cafeteria and custodial staff for input in school decisions.
- Asks teachers to periodically complete a form concerning the quality of support services as an aid in monitoring these activities.

Averag

3 ☐

Examples of Behavior of Individuals who are Usually Rated "Low" on this Dimension:

2 ☐

- Assigns hall and cafeteria duty without staff input.
- Allows a community group to use a school projector without the permission of the media specialists.
- Does not consult with custodians when making requisitions for supplies.
- Neglects to inform secretary of the school programs and events.

1 ☐ Low

DO NOT WRITE IN THIS AREA

6. SUPPORT SERVICES: DIRECTING THE BEHAVIOR OF STUDENTS

CHECK ONLY ONE RESPONSE

If you have insufficient information to rate this dimension, check this box. ☐

RATING

Examples of Behavior of Individuals who are Usually Rated "High" on this Dimension:

- Maintains up-to-date staff manuals including statements on discipline policies, which communicate all procedural matters.
- Works with teachers and students to handle discipline problems.
- Helps counselor develop programs for incoming students, such as orientation.
- Sets guidelines for student behavior which are not threatening and enforces them objectively.
- Conducts a cafeteria survey to determine student needs in that area.

6 ☐ High

5 ☐

4 ☐

Examples of Behavior of Individuals who are Usually Rated "Average" on this Dimension:

- Develops a staggered release schedule on rainy days to alleviate traffic jams in front of school caused by parents picking up their children.
- Develops special instructional programs for lunch period so that students do not just play in gym or hang around.
- Starts a program in which highest attending classroom is recognized with a free period and a party.

Avera

3 ☐

Examples of Behavior of Individuals who are Usually Rated "Low" on this Dimension:

- Discontinues an arts program because a small group of students ruined a kiln.
- Suspends a student with whom the counseling staff is working closely; does not check with staff first.
- Does not discipline students because he/she does not know what to do or does not have time.
- Places a student in a special education program for reading problems without proper evaluation.

2 ☐

1 ☐ Low

DO NOT WRITE IN THIS AREA

7. STAFF EVALUATION

CHECK ONLY ONE RESPONSE

If you have insufficient information to rate this dimension, check this box. ☐

RATING

Examples of Behavior of Individuals who are Usually Rated "High" on this Dimension:

- Provides constructive feedback after observations and calls in specialists if additional help is needed; then observes again and gives more feedback.
- Encourages evaluation of him/herself by staff.
- Consults with individual staff members on a periodic basis to develop individual standards of performance (goals and objectives) and reviews subsequent accomplishment of goals.
- Writes a note to each teacher who has led a particular event or done something special to show appreciation and give praise.

6 ☐ High

5 ☐

4 ☐

Examples of Behavior of Individuals who are Usually Rated "Average" on this Dimension:

- Disciplines employee who is not functioning properly.
- Asks permission to come to observe classroom.

Averag

3 ☐

Examples of Behavior of Individuals who are Usually Rated "Low" on this Dimension:

- Doesn't afford a teacher adequate opportunity to refute accusations made by students.
- Circulates a staff bulletin criticizing teachers when only 3 or 4 teachers were guilty and should be reprimanded individually.
- Criticizes or belittles staff members in front of other staff members, students or parents.

2 ☐

1 ☐ Low

DO NOT WRITE IN THIS AREA

8. DEVELOPMENTAL ACTIVITIES

CHECK ONLY ONE RESPONSE

If you have insufficient information to rate this dimension, check this box. ☐

RATING

Examples of Behavior of Individuals who are Usually Rated "High" on this Dimension:

- Provides in-service programs for staff which include dealing with student behavior problems and interactions with parents.
- Attends seminars, workshops, etc. for professional growth.
- Provides time for teachers who have attended developmental activities to communicate information to the rest of the staff.

Examples of Behavior of Individuals who are Usually Rated "Average" on this Dimension:

- Devises class schedule to allow for staff development periods.
- Uses staff meetings to encourage and publicize in-service training.
- Allows teachers time off to attend meetings and classes.

Examples of Behavior of Individuals who are Usually Rated "Low" on this Dimension:

- Tells staff to get involved in staff development problems when he/she does not get involved.
- Plans staff development activities with no staff input.
- Does not help teachers who have problems but works toward documenting problems and removing teacher.
- Does not allow staff to go to professional meetings even at their own expense.

6 ☐ High

5 ☐

4 ☐

Average

3 ☐

2 ☐

1 ☐ Low

DO NOT WRITE IN THIS AREA

30

9. COMMUNITY RELATIONS

CHECK ONLY ONE RESPONSE

If you have insufficient information to rate this dimension, check this box. ☐

RATING

6 ☐ High

Examples of Behavior of Individuals who are Usually Rated "High" on this Dimension:

- Informs community about the basics of operating the school and educational objectives; asks for the ideas, opinions and support of these people and listens to their suggestions.
- Works with various community and local groups to develop cooperation with the school.
- Institutes a course which is aimed at developing cultural awareness in an integrated community.

5 ☐

4 ☐

Examples of Behavior of Individuals who are Usually Rated "Average" on this Dimension:

- Sends stories to the community newspaper and/or invites newspaper correspondents to school events.
- Organizes senior citizen breakfasts to explain school programs.
- Coordinates school activities with church and community events.

Average

3 ☐

- Takes two students to the Rotary Club every month as a reward and to promote good relations.

Examples of Behavior of Individuals who are Usually Rated "Low" on this Dimension:

- Stays away from community organizations such as Jaycees, Lions, Kiwanis, etc.
- Defends his/her position to the community rather than listening to their requests for certain programs.
- Speaks in educational jargon to explain school programs to community members.
- Speaks badly of the school system to the news media.

2 ☐

1 ☐ Low

DO NOT WRITE IN THIS AREA

31

10. INTERPERSONAL EFFECTIVENESS

CHECK ONLY ONE RESPONSE

If you have insufficient information to rate this dimension, check this box. ☐

RATING

Examples of Behavior of Individuals who are Usually Rated "High" on this Dimension:

- Shows a sense of humor in times of conflict.
- Interacts with students during the lunch hour in the cafeteria.
- Encourages parents to go directly to a teacher with a compliment instead of relaying the message him/herself to the teacher.
- Displays students' art work in main office as encouragement.
- Comes from behind desk to discuss problems, takes notes, and refers to this file later when another discussion takes place.

6 ☐ High

5 ☐

4 ☐

Examples of Behavior of Individuals who are Usually Rated "Average" on this Dimension:

- Leaves a note of encouragement to someone who has made a mistake, to carry on and start over.
- Informs staff of all contributions to school activities by faculty members and communicates appreciation.
- Allows staff member to go home during the day to solve a family problem.

Average

3 ☐

Examples of Behavior of Individuals who are Usually Rated "Low" on this Dimension:

- Delegates work but then takes work away or interferes because he/she thinks it is not getting done correctly.
- Uses language and words which students do not understand when communicating with them.
- Establishes a policy without giving staff or students a rationale.
- Gets into "shout-outs" with students in the hallways.

2 ☐

1 ☐ Low

DO NOT WRITE IN THIS AREA

32

11. COMMUNITY RELATIONS: PARENTS

CHECK ONLY ONE RESPONSE

If you have insufficient information to rate this dimension, check this box. ☐

RATING

Examples of Behavior of Individuals who are Usually Rated "High" on this Dimension:

- Starts a procedure whereby parents come to meet teachers individually to discuss child's progress.
- Organizes coffees in parents' homes to interact with parents in a nonschool environment.
- Writes a letter to all parents inviting them to school, spends an evening talking with them and answering questions.

6 ☐ High

5 ☐

Examples of Behavior of Individuals who are Usually Rated "Average" on this Dimension:

- Implements a program in which parents follow class schedule of their children for a shortened day.
- Organizes a system for buses to pick up parents and bring them to school events in order to promote community involvement.
- Calls parents to remind them of parent-teacher conference so as to ensure attendance.

4 ☐ Average

3 ☐

Examples of Behavior of Individuals who are Usually Rated "Low" on this Dimension:

- Refuses to develop or initiate a parent advisory group.
- Controls PTA by dominating officers, promoting participation of only one segment of school population, or failing to follow through on their requests or suggestions.
- Does not have a system of reporting students' discipline problems to parents.
- Tells parents they do not really know their children when problems arise.

2 ☐

1 ☐ Low

DO NOT WRITE IN THIS AREA

33

12. COORDINATION WITH DISTRICT AND OTHER SCHOOLS

CHECK ONLY ONE RESPONSE

If you have insufficient information to rate this dimension, check this box. ☐

RATING

6 ☐ High

Examples of Behavior of Individuals who are Usually Rated "High" on this Dimension:

- Participates in professional organization problem solving projects aimed at improving the functioning of central administration services which impact directly on the school.
- Balances district activities against building priorities.

5 ☐

- Encourages use of district resource personnel to develop needed programs.
- Informs central office of possible complaints or problems at school or in community so they will be aware of it ahead of time.

4 ☐

Examples of Behavior of Individuals who are Usually Rated "Average" on this Dimension:

- Arranges meetings with central office administration to speak to staff at school on various issues.
- Implements decisions at the building which were made at the district level even if not involved in making decisions.
- Participates in district level committees when called upon.

Average

3 ☐

Examples of Behavior of Individuals who are Usually Rated "Low" on this Dimension:

- Makes all decisions without the advice of any immediate superior.

2 ☐

- Is so involved in district committees that he/she is always out of the building.
- Does not communicate with district personnel unless absolutely necessary.
- Never volunteers for district committees.

1 ☐ Low

DO NOT WRITE IN THIS AREA

13. FISCAL OR MONETARY MANAGEMENT

CHECK ONLY ONE RESPONSE

If you have insufficient information to rate this dimension, check this box. ☐

RATING

Examples of Behavior of Individuals who are Usually Rated "High" on this Dimension:

- Involves all staff in establishing priorities for the allocation of resources and materials.
- Periodically reviews and shifts allocations based on current needs; obtains agreement from all department heads.
- Maintains strict requirements on the procedures to follow with purchase orders and budgets of departments.
- Acquaints the staff with the daily operations budget and how it relates to their curriculum.
- Calls several other schools and sources to get materials when they are not available at the local school.

6 ☐ High

5 ☐

4 ☐

Examples of Behavior of Individuals who are Usually Rated "Average" on this Dimension:

- Monitors expenditures of students' activity funds.
- Provides funds for special art projects so that students would not need to bring materials from home.
- Keeps staff informed of budget.

Average

3 ☐

Examples of Behavior of Individuals who are Usually Rated "Low" on this Dimension:

- Does not comply with standard accounting procedures.
- Misspends school money in one area and compensates by taking money out of budget in another area.
- Spends an inordinate amount of school funds on athletics.
- Requires staff to fill in requisition forms for small items, such as toilet paper and light bulbs.

2 ☐

1 ☐ Low

DO NOT WRITE IN THIS AREA

35

14. MAINTENANCE OF SCHOOL PLANT

CHECK ONLY ONE RESPONSE

If you have insufficient information to rate this dimension, check this box. ☐

RATING

Examples of Behavior of Individuals who are Usually Rated "High" on this Dimension:

- Works with new students to develop school pride and reduce vandalism. Older students are given responsibility for taking new students "under their wing".
- Initiates a program to clean up graffiti in school; provides students with cleaning materials and develops a contest for cleanest area, thereby unifying students and staff.
- Establishes a process which allows staff and community members to provide input for the orderly improvement of school plant facilities and equipment.

Examples of Behavior of Individuals who are Usually Rated "Average" on this Dimension:

- Takes an active part in keeping the building clean (e.g., picking up the paper in hallways, repairing bulletin boards).
- Obtains necessary materials when things need to be fixed for ongoing daily functions in the building.
- Sets up a graffiti board in the restrooms to minimize vandalism.

Examples of Behavior of Individuals who are Usually Rated "Low" on this Dimension:

- Unilaterally determines which school facilities are to be improved.
- Does not provide a safe storage place for expensive material.
- Never foresees equipment problems; reacts to peoples' requests and emergencies.
- Ignores a report of unsafe playground equipment until something happens to a child.

6 ☐ High

5 ☐

4 ☐

Average

3 ☐

2 ☐

1 ☐ Low

DO NOT WRITE IN THIS AREA

15. STRUCTURES COMMUNICATION WHICH PROVIDE FOR COOPERATION AMONG VARIOUS GROUPS IN SCHOOL

CHECK ONLY ONE RESPONSE

If you have insufficient information to rate this dimension, check this box. ☐

RATING

Examples of Behavior of Individuals who are Usually Rated "High" on this Dimension:

- Plans meetings of staff, supervisors, and parents to air concerns regarding school programs or problems.
- Communicates through meetings, memos, or personal communications to tell people about school or things that impact on school's functioning.
- When dealing with irate or concerned parents, consults with the personnel involved.
- Sets up communication devices so that teachers as a group can convey needs to support service personnel (e.g., librarian, cafeteria staff, media specialists).

6 ☐ High

5 ☐

4 ☐

Examples of Behavior of Individuals who are Usually Rated "Average" on this Dimension:

- Distributes a weekly newsletter to inform staff of school events.
- Keeps appointment calendar on desk so anyone can make an appointment at any time.
- Trains students to make announcements of events each day.
- Meets with whole staff once a month with no agenda, to discuss problems that have developed.

Average

3 ☐

2 ☐

Examples of Behavior of Individuals who are Usually Rated "Low" on this Dimension:

- Sees people only by appointment.
- Makes decisions without involving those whom it directly affects.
- Organizes an Advisory Committee but never goes to a meeting.
- Communicates only by mail, P. A., or memo with staff and students.
- Stays in office all day.

1 ☐ Low

DO NOT WRITE IN THIS AREA

SECTION C: EFFECTIVENESS SCALE

EFFECTIVENESS SCALE

DO NOT WRITE IN THIS AREA

Instructions: Various dimensions of school administrators are listed below. Read each one carefully and rate the incumbent by making a check mark at the scale value which represents your judgment of what the person is really like. If you are not able to give a reliable rating on a particular dimension because you do not have enough information, mark the "DO NOT KNOW" box.

1. Adaptability:

How well does this person adapt his or her behavior and approach in dealing with different situations and different people?

38

|__| do not know

|__| 5 extremely well

|__| 4 above average

|__| 3 satisfactory

|__| 2 needs improvement

|__| 1 falls short of requirements

2. Conferring with Parents:

How effectively does this person deal with parent conferences where he or she needs to review material relating to a student experiencing a problem at school, to meet with the parent(s) to discuss the student's problem in an attempt to remedy the situation, and perhaps to prepare a brief written summary of the outcome of the meeting for his or her immediate supervisor?

39

|__| do not know

|__| 5 extremely well

|__| 4 above average

|__| 3 satisfactory

|__| 2 needs improvement

|__| 1 falls short of requirements

DO NOT WRITE IN THIS AREA

3. Decision Making:

How logical and sound are the judgments this person makes in using resources, determining courses of action and defining solutions to problems?

| | do not know 40

| | 5 extremely well

| | 4 above average

| | 3 satisfactory

| | 2 needs improvement

| | 1 falls short of requirements

4. Decisiveness:

How firm and conclusive is this person in making decisions, rendering judgments, taking action and defending decisions, judgments and actions when challenged?

| | do not know 41

| | 5 extremely well

| | 4 above average

| | 3 satisfactory

| | 2 needs improvement

| | 1 falls short of requirements

5. Interpersonal:

To what extent is this person sensitive to others and behaves in a manner which reflects an appreciation of the needs, feelings and capabilities of others? How effectively does he or she show courtesy, tact and understanding, regardless of a person's status or position, accept interpersonal differences and develop rapport with others?

42

| | do not know

| | 5 extremely well

| | 4 above average

| | 3 satisfactory

| | 2 needs improvement

| | 1 falls short of requirements

DO NOT WRITE IN THIS AREA

6. Management Skill:

How much leadership does this person provide through directing and coordinating the activities of others, by delegating authority and responsibility and in providing the means to follow-up?

| | do not know 43

| | 5 extremely well

| | 4 above average

| | 3 satisfactory

| | 2 needs improvement

| | 1 falls short of requirements

7. Observing Teachers:

How effectively does this person use teacher observation skills by reviewing the teacher's background, observing the teacher instructing a class, completing an Observation Form and subsequently meeting with the teacher to give feedback on the performance observed?

| | do not know 44

| | 5 extremely well

| | 4 above average

| | 3 satisfactory

| | 2 needs improvement

| | 1 falls short of requirements

8. Oral Communications:

How clearly and effectively does this person speak or express information orally, making proper use of techniques, such as voice inflection, eye contact, and good grammar and vocabulary?

| | do not know 45

| | 5 extremely well

| | 4 above average

| | 3 satisfactory

| | 2 needs improvement

| | 1 falls short of requirements

DO NOT WRITE IN THIS AREA

9. Paper Work:

How effectively does this person structure, organize and deal with memos, letters, correspondence and other reports which vary in their importance and urgency; where he or she must write letters, memos, give instructions and, from time to time, clarify or explain action taken?

46

| | do not know

| | 5 extremely well

| | 4 above average

| | 3 satisfactory

| | 2 needs improvement

| | 1 falls short of requirements

10. Planning and Organizing:

How well does this person establish objectives, schedules, priorities and strategies for himself or herself, as well as others, to accomplish specific results?

47

| | do not know

| | 5 extremely well

| | 4 above average

| | 3 satisfactory

| | 2 needs improvement

| | 1 falls short of requirements

11. Perception:

How well does this person identify the critical elements of a situation, interpret implications of alternative courses of action and evaluate factors essential to the solution of a problem?

48

| | do not know

| | 5 extremely well

| | 4 above average

| | 3 satisfactory

| | 2 needs improvement

| | 1 falls short of requirements

DO NOT WRITE IN THIS AREA

12. Leadership:

How effectively does this person lead and motivate students, parents, teachers, school officials and others? Does he or she receive loyalty from them and inspire in them a sense of purpose and desire to do what they are asked?

| | do not know 49

| | 5 extremely well

| | 4 above average

| | 3 satisfactory

| | 2 needs improvement

| | 1 falls short of requirements

13. Technical Know-how:

How effective is this person in acquiring, utilizing, and applying the technical knowledge required by his or her current position?

| | do not know 50

| | 5 extremely well

| | 4 above average

| | 3 satisfactory

| | 2 needs improvement

| | 1 falls short of requirements

14. Written Communication:

How well does this person express information effectively and clearly in writing, making proper use of grammar and vocabulary?

| | do not know 51

| | 5 extremely well

| | 4 above average

| | 3 satisfactory

| | 2 needs improvement

| | 1 falls short of requirements

15. Overall Rating:

In comparison with all others presently holding the same job, how would you rate this person in overall effectiveness in his or her position?

DO NOT WRITE IN THIS AREA

52

Rating	Stanine	%ile
|__|	9	96-100
|__|	8	89-95
|__|	7	80-88
|__|	6	61-79
|__|	5	40-60
|__|	4	21-39
|__|	3	12-20
|__|	2	5-11
|__|	1	1-4

16. Remarks:

If you wish to comment on the Effectiveness Scale or offer any suggestions, please do so.

The information you have provided on this rating instrument will be used to evaluate the selection procedure of school level administrators. Hopefully, through our mutual efforts the selection process can be improved. Thank you for your cooperation.

Please return the completed rating instrument to:

Mr. Joe Gomez
Board Administration Bldg.
Mail Code 9999
Room 800

Use the enclosed return envelope.

The School Board of Dade County, Florida adheres to a policy of nondiscrimination in educational programs/activities and employment and strives affirmatively to provide equal opportunity for all as required by:

Title VI of the Civil Rights Act of 1964 - prohibits discrimination on the basis of race, color, religion, or national origin.

Title VII of the Civil Rights Act of 1964, as amended - prohibits discrimination in employment on the basis of race, color, religion, sex, or national origin.

Title IX of the Education Amendments of 1972 - prohibits discrimination on the basis of sex.

Age Discrimination Act of 1967, as amended - prohibits discrimination on the basis of age between 40 and 70.

Section 504 of the Rehabilitation Act of 1973 - prohibits discrimination against the handicapped.

Florida Educational Equity Act - prohibits discrimination on the basis of race, sex, national origin, marital status or handicap against a student or employee.

Veterans are provided re-employment rights in accordance with P.L. 93-508 (Federal) and Section 295.07, Florida Statutes, which also stipulates categorical preferences for employment.

"END OF DOCUMENT"