

Assessment centres seem to offer a credible alternative to interviews in personnel selection.

Selection Methods for Educational Administrators in the USA

Petros Pashiardis

International Journal of Educational Management, Vol. 7 No. 1, 1993, pp. 27-35.
© MCB University Press, 0951-354X

Purposes

The selection of administrators (Principals and Assistant Principals) for America's schools has always been given a degree of importance. However, this process is being given the utmost importance because recent findings from the body of research, called *Effective Schools Research*, indicate that a school's leadership is its heart and soul. The effective schools research has indicated that when a school has competent leadership (combined with other factors), it is on the right track of becoming an effective school and of achieving excellence, where students and teachers get the most out of their interactions and experiences[1].

This article has three purposes. To describe:

- (1) different methods of principal selection currently employed,
- (2) the Assessment Centre (AC) Method of principal selection, and
- (3) the University of Texas at Austin Assessment Center (UT-AC) in particular.

A literature review will follow presenting some current practices in order to acquaint the reader with ACs and the UT-AC in particular.

Review of the Literature

As mentioned earlier, the principal selection process is currently gaining more and more importance both in America and in Europe[2].

However, even today, the vacancy stage of a principalship does not correspond well with the weights placed on the importance of the position. As Baltzell and Dentler[3-5] cite, "Instead, this stage tends to send a message about a rather lower-level, middle-management or supervisory post". The job descriptions do not really reflect the importance of the position. Therefore, it is important that the qualifications required for such an important leadership position be made explicit and overt. Until now, instructional leadership was not a well-specified or widely-applied criterion for selecting principals. Now, it is becoming the main factor used in examining the suitability of a candidate for the principalship.

The most pervasive method (process) for selection has been (and still is) the interview technique (after some initial screening takes place). However, usually screeners express great concern about the limitations of the screening interview in terms of its ability to predict success on the job[3;4;5, p. 11]. Because of this method's subjectivity, a broader participation is usually required in the interview committees in order to give various constituencies the opportunity to examine the candidate and thus legitimize the method. The final decision always rests upon the superintendent for recommendation, and to the school board for final hiring. Usually the superintendent controls the process from the beginning but does not become overtly involved until the end. In addition to the above, the process is more often political and chance-ridden. Baltzell[3,4,5] present some cases of principal selection and conclude that:

Our cases make the point, however that selection pathways are often chance-ridden, that even when merit and equity triumph they may do so for reasons that have to do with the maneuvers of transfer, the consolidation of a superintendent's authority, or the turning of the wheel of political fortune[4, p. 29].

Owing to the above problems and inconsistencies with the interview process, school districts and educational research centres have been experimenting with different methods/approaches of principal selection. Four of those approaches are going to be presented briefly and a fifth (ACs) is to be examined in more detail.

The Administrative Intern Programme (AIP) in Hayward Unified School District, California, began in 1979. After vacancies are announced, applicants submit letters of

intent and complete problem exercises which are evaluated. Candidates, who pass this first screening, are then interviewed for 30 minutes, then given another set of problems and, if they pass, are admitted to the AIP. The AIP includes a variety of opportunities for training, such as visits, observations, conferences, workshops, interviews, substituting, attendance at board and council meetings, and formal training sessions. Its cost has been slight when contrasted with the extensive benefits in increased competence, organizational unity and instructional services of many Hayward public schools[3-5].

The Administrative Training Programme (ATP) of the Montgomery County Public School system in Maryland is a three-part programme which involves: (1) a career development phase; (2) an administrative internship; (3) a district-run AC to assess interns as they complete the training programme. The administrative internship is an intensive, on-the-job training programme. Interns are systematically evaluated by a district team and their principal. Based on their performance, interns are either offered a position or not.

Broward County's (Florida) Principal Selection Process has three main elements: (1) An "eligibility list", which describes certain selection criteria and career ladders; (2) a "vacancy screening", which rates eligible candidates against the specific requirements of each vacancy; and (3) the "vacancy interview", which selects a finalist from among the most highly rated eligible candidates.

In addition, Baltzell and Dentler[3-5] describe certain clues to essential principal selection procedure elements such as: (a) *Merit* (Vacancy, Selection criteria, Applicant pool, Screening, and Appointment decisions); (b) *Sex and racial equity* (keep those ratios in mind); (c) *Legitimacy* (credibility).

Teachers and parents will increase their confidence in the system and will follow the lead of principals to the extent that they interpret the principal selection process as operating with integrity and on the basis of merit[4, p. 59].

The Integrated Appraisal Approach proposed by Brush and Schoenfeldt[6] uses data gathered on the job. Appropriately collected job-performance data would have the advantage of being (1) available without the use of costly and contrived experiences, (2) clearly job-related, (3) defensible on the basis of content-oriented validity, and (4) more meaningful for employee development. The authors say further that in order to achieve the above, it is necessary to (a) perform a *job-analysis* to establish the critical tasks and abilities required for each management position, (b) provide assessor *training*, and (c) *establish procedures* (they use multiple raters and a variety of evaluation methods just like ACs). Furthermore, employee behaviour is observed, recorded, and evaluated through multiple lines over a longer period of time.

However, the authors do acknowledge that the cost of this method could be the same as that of the ACs and the yield in predicting performance could be the same as well.

Overall, the various internship models can be seen as the main alternative to ACs. In some cases internships give better indications about job performance because the intern is actually on the job. However, they are time-consuming and expensive and not flexible when there is an immediate vacancy which needs to be filled right away.

The other alternative, advocating the continuous presence of an evaluator, proposed by Brush and Schoenfeldt[6], does not seem to be feasible, mainly because one would need full-time managers on the job doing nothing else but evaluating. It is next to impossible for managers to devote the time and energy necessary to gather reliable information for evaluation purposes.

Thus, ACs seem to be the most viable alternative for personnel selection. However, ACs have some disadvantages as has been explained in the literature review. Besides costs, coaching and other disadvantages that have been explained there is a more serious problem. What do ACs measure and how well do they predict? As Schmitt *et al.*[11] and Gomez[7] report, the NASSP skills do measure uniquely what they are supposed to be measuring. In other words, there is no multicollinearity among the variables (dimensions). However, those results contradict the Ogawa and Oxaal[12] results which indicated a large degree of multicollinearity among the skills/dimensions of the NASSP AC. Therefore, the skills/dimensions area of the AC are still questionable as to what they measure and if they measure it uniquely. Furthermore, the statistical analyses and interpretations performed in some of the studies[11,7], are, at first glance, excellent. However, the researchers subjectively explain discrepancies when the results do not conform to their expectations.

Moreover, in Gomez's evaluation[7], Dr Schmitt was used as a consultant and several of his instruments were used in obtaining the data. Thus, one would be able to conclude that Gomez's evaluation of the Dade County Public Schools' AC is not all that objective and independent. Therefore, one is more inclined to accept Ogawa's and Oxaal's results[12], or demand that more research is done in order to establish the validity of ACs.

Also, a close examination of the literature suggests that significantly higher correlations exist between AC predictions and job potential ratings (median $r = 0.63$) than between AC ratings and actual job performance (median $r = 0.33$)[6]. In relation to the above, it is interesting to note that Gomez[7] reports that when comparing the interview process with the management AC of Dade County Public Schools process, none proves to be superior to the other. His explanation for this

outcome is that, given the minimum passing score of the management AC, few candidates are eliminated, probably the same ones who would have been eliminated by using the interview process. Thus, he recommends that the minimum passing score of the management AC be raised, therefore, increasing its screening capability. The author of this article agrees with Gomez that the power of an AC lies in being able to screen candidates by using more rigorous techniques and higher passing standards. However, even with the above disadvantages, ACs seem to offer a credible alternative to more questionable hiring practices such as interviews.

Districts which emphasize internships, do so in large part in order to obtain performance-based criteria, because they feel that the weight ascribed to interviews probably exceeds the power of prediction that can be gained under the best circumstances. However, districts which do not have internship positions and still want performance-based criteria for their candidates are turning increasingly towards assessment centres which are the real focus of this paper.

Assessment Centres (ACs)

The ACs have been used for years (since the 1950s) as a personnel selection method in industry. Current estimates of the number of ACs in operation in the United States range from 300 to 1,000[7]. A brief description of ACs will be presented as well as some criticisms of this method. The very first use of ACs was to evaluate candidates for irregular warfare, under the auspices of the Office of Strategic Services (OSS) during World War II. ACs have several distinctions:

- (1) The use of multiple, contrived situations (for example, business simulations) to observe behaviour.
- (2) The presence of several trained assessors who pool their evaluations along a variety of specified dimensions (for example, the assessee's leadership, risk-taking, and administrative abilities).
- (3) The evaluation of several candidates at one time.
- (4) Extensive feedback, written or verbal, to either the candidate or management, or both[6].

Moreover, a true AC conforms to the "Standards and Ethical Considerations for AC Operations." This set of standards was endorsed by the Third International Congress on the AC method in 1975 and revised in 1980. It defines an AC as follows:

An AC 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[8, p. 35].

Assessment Centres in Education

During the 1970s, the AC approach to personnel identification and selection began moving into education, and since 1975 the National Association of Secondary School Principals (NASSP) of Reston, Virginia, in conjunction with the American Psychological Association, developed an AC model. The purpose of this pilot project was to demonstrate another approach to selecting potentially successful administrators which might set a standard for others to emulate[9]. By 1984, there were 19 ACs organized according to the NASSP model. These ACs served more than 200 school districts in 15 states[10]. Thus, this project is the most extensively used model of its kind in public education so far.

According to the NASSP the centre is an event not a place and is conducted as frequently as the adopting district desires. Each centre is conducted by six well-trained assessors whose charge is to observe, measure and evaluate the candidates as they complete various exercises and simulations[3-5].

At each centre the assessors evaluate each participant on 12 skills/dimensions:

- (1) *Problem analysis* (which measures the candidate's ability to seek out relevant data and analyse complex information for problem solving);
- (2) *Judgement* (ability to reach logical conclusions and make high quality decisions);
- (3) *Organizational ability* (ability to plan, schedule, and deal with pressure);
- (4) *Decisiveness* (ability to recognize when a decision is required and to act quickly);
- (5) *Leadership* (ability to provide the necessary style of leadership needed for a group to accomplish a task);
- (6) *Sensitivity* (ability to perceive the needs, concerns, and personal problems of others and deal effectively with others);
- (7) *Stress tolerance* (ability to perform under pressure and during opposition);
- (8) *Oral communication* (ability to make a clear oral presentation of facts or ideas);
- (9) *Written communication* (ability to express ideas clearly in writing);
- (10) *Range of interests* (competence to discuss a variety of subjects);
- (11) *Personal motivation* (need to achieve in all activities attempted);
- (12) *Educational values* (possession of a well-reasoned educational philosophy; receptiveness to new ideas and change)[9].

When the assessors' discussions are complete and final ratings have been assigned, a detailed report is prepared on each candidate. The reports cover areas of strength, needs for improvement and, training and development recommendations. The assessment report is then given to the candidate and discussed in a confidential feedback interview.

After the first ACs had taken place, researchers began to wonder about their ability to predict performance. In 1979, a three-year validity study of the NASSP ACs was undertaken by a team of Michigan State University psychologists headed by Dr Neal Schmitt. The results of the study included a number of significant, positive validity correlations between AC skill ratings and subsequent job performance[11]. Similar results were revealed by the evaluation of the Management Assessment Centre (MAC) of the Dade County Public Schools in Florida, undertaken by Dr Joseph Gomez[7]. Moreover, the evaluation showed that:

- (1) The inter-rater reliability (which is considered very important to validity in an AC) was high;
- (2) The validity correlations were substantially higher than those generally produced by the interview method;
- (3) The validity correlations compared favourably with those of other ACs; and
- (4) There is evidence that the validity correlations are still rising[7, pp. 44-46].

ACs' Weaknesses

However, others undertook similar validation studies and revealed different results. Ogawa and Oxaal[12] wanted to examine the (1) validity of the NASSP's 12 dimensions using data from the Intermountain-NASSP AC Project (IACP) of the University of Utah. An initial problem which they encountered in their analysis was that the 12 dimensions/skills were highly correlated, thus multicollinearity was a problem for the analysis and therefore regression was problematic. Subsequently, they thought factor analysis and therefore regression was problematic. Subsequently, they thought factor analysis revealed that scores of all 12 skills shared one factor[12]. Surprised by these results, they took the original findings and reports by Schmitt and his associates and they factor-analysed their research. This time, instead of one, they identified four factors: analytical skills, self-presentation skills, sensitivity and written communication.

In reviewing a larger number of studies, Thornton and Byham[13] discovered that many skill dimensions have been found to correlate significantly with the final overall ratings.

Sackett and Hakel[14], using data collected at AT&T over three years, found that five dimensions contributed

significantly to the overall rating: leadership, organizing and planning, decision making, behaviour flexibility, and energy.

The above descriptions indicate conflicting views about the validity of what ACs actually measure and how and are regarded as disadvantages of ACs.

Other weaknesses include ACs' costs. As Brush and Schoenfeldt[6] report, Joseph Steger, former Dean of the School of Management at Rensselaer Polytechnic Institute in Troy, New York, estimated that the testing (paper and pencil) cost was \$68.82, whereas the AC cost was \$540.25. It generally cost an additional \$24,996 to go from a validity of 0.45 (with tests) to 0.52 with ACs.

Another weakness is coaching — future participants will know about the materials being used after the previous participants tell them (it would be very expensive to change constantly the materials being utilized). Also heterogeneity of assesseees could be a problem, when assesseees realize that others among the group have higher status than they do. Probably their performance will be adversely affected.

A final problem relates to job relatedness — are all exercises relevant? Is watching someone performing on-the-job the same as performing under simulated situations? Some researchers believe that there is lack of job relatedness among the exercises involved[6]. As the same authors purport:

Some studies suggest the possibility that AC staffs are appraising candidates not so much on an objective evaluation of observed performance as on their own familiarity with their preferences of management (p. 70).

ACs' Strengths

Nevertheless, there are several strengths associated with ACs:

- (1) Assessor training — going through an AC is a valuable experience because assessors learn how to observe, report, and evaluate, therefore increasing their own ability to handle personnel management functions.
- (2) Reliability — generally, inter-rater reliability is reported in the 70s and 80s range[6].
- (3) Career development — feedback to the employee in the form of effective career development counselling is important.

A detailed description of the University of Texas AC will be presented in order to familiarize the reader with current uses of the AC technique in the university setting. A comparison between the UT-AC and other existing centres will also be attempted.

The University of Texas Assessment Center

The Executive Leadership Assessment Center as it is called was initiated in the spring of 1987 with no specific model in mind. The purpose of the centre is to assist the Department of Educational Administration in making more informed decisions regarding the admission of graduate students into the master's programme, mid-management and supervision certification programmes and eventual doctoral candidacy. (The information concerning the UT-AC was gathered through an interview and materials given by Dr Carole Veir, a Professor in the Department of Educational Administration at UT, to the author.)

The areas of emphasis for the centre were developed by a team of professors from the department, through a series of meetings and revisions. The development team consisted of Drs Harris, McIntyre, Rippey, Baker, Veir, and Valverde. After some initial decisions regarding the content and process were made and some ideas were formulated, a group of 25 prominent and well-respected practitioners from the state of Texas were gathered at UT to discuss and review the need, concept, content, direction, and field involvement in the centre. The reaction of all participants towards an AC of this nature was extremely favourable and therefore, the idea was further pursued.

The initial assessment activities were determined through a synthesis of the literature on assessment centres in operation and the research on them; the NASSP activity areas; and others that the team thought would enhance the admissions process and/or were seen as critical information for decision making and/or programme planning. The committee felt that decisions based solely on the biographical data sheets, GRE scores, and rating forms were no longer yielding enough data or the types of data the faculty thought were needed to assure the stature of students the university wanted to draw. In addition, the department was in the process of raising its admission standards to a level that warranted the closer scrutiny of incoming students, to ensure that the student was doctoral material when entering the programme. Therefore, it was felt that the AC process would allow the department to screen in more qualified students and not just screen out potentially good administrators based on data known to be of questionable validity, e.g. letters of reference and grades.

After initial decisions were made regarding the areas to be assessed and the possible means of gathering the data, each professor was responsible for developing a specific activity or instrument for use in the centre. As with all such undertakings, the process was revised continuously to upgrade and modify the current instrumentation. Several changes have been made since the first centre was held, and it is anticipated that these changes will continue until satisfactory instrumentation yields the appropriate results.

Some of the instruments are administered at home before the candidates come to the AC. The rest are administered at the university. The areas currently being assessed are:

- *Writing samples* — one of the most critical pieces of the centre for both graduate student skill-knowledge and for a skill needed as an administrator. The writing samples are given as in-basket activities. However, it is the most difficult to grade and process information; it needs to be simplified or an answer key needs to be developed.
- *Watson-Glaser* — diagnostic in nature, good screening instrument to look at critical thinking skills, used in conjunction with the verbal scores on the GRE (especially if GRE score is below 450); it has been used as part of the testing process in the department since 1952 in FEA.
- *Self-assessment* — also diagnostic in nature, tests knowledge of good teaching practices (pre-TTAS), has an answer key, can be scored by a secretary, teaching assistant, or aide.
- *Interviews* — seen by the students as the most useful learning activity for them during the assessment process; it has yielded very stable data, the comments are useful to students, and is seen by the field personnel as a very fruitful activity that allows them to stay in contact with a potential pool of candidates and stay in tune with the university training programme and its emphasis areas.
- *Oral presentation* — (it measures organizational ability, oral communication, sensitivity, decisiveness, stress tolerance) — seen by most students as the most intimidating activity, one of the most useful and fun in the final analysis; seen by field personnel as revealing and yielding very useful information about interpersonal skills, ability to function on their feet, and pull from a bank of knowledge about a variety of subjects that may surface during the process; it is also perceived by field personnel as informational and enjoyable as well.
- *KAT-knowledge assessment test about curriculum/teaching* — it is diagnostic in nature, tests knowledge of leadership, has answer key, and can be scored by a secretary, teaching assistant or aide.
- *Myers-Briggs* — (personal attributes, leadership, sensitivity) — students seem to like this instrument and have expressed an interest in knowing more about the results.
- *EAEP* — (leadership behaviour) — this has only been administered once so far, and as Dr Veir said, they do not yet have enough data to determine its overall usefulness; However, it is perceived that this is a good instrument to use in making better informed decisions about entrance into the doctoral programme and student professional growth from entrance into the programme to the point of doctoral studies. If it is feasible it could be used as a pre/post programme exam.

As Dr Veir said, "Most of the instruments have been used three or four times and can begin to establish some baseline data for cutoff scores. We have data now and can follow through the programme and into the students' performance in the field. However, we now need to take a close look at the assessment data and scoring based upon this information, and develop a follow-up instrument for determining (a) if the assessment process is making a difference in the students we are entering into the programme, (b) that the good performance of these students is not only observed in the university setting but also once they are on the job, (c) what we will modify and change and what will be left in the process" (personal interview with Dr Veir).

Another characteristic of the UT-AC is that it is labour intensive. To date, there has been no problem in pulling people in from the field to assist (in fact most of them seem to enjoy it, as Dr Veir mentioned), but it is something that needs to be addressed.

Furthermore, a clear distinction should be made between those assessment centres (that currently exist in the field) and the NASSP centre on one hand that are designed for the purpose of screening personnel into and/or out of administrative positions, and the UT centre on the other. The UT centre has not been designed to be diagnostic or prescriptive in itself but was designed to assist decision makers in making informed decisions about academic potential of students who are entering a training programme with the notion of developing the necessary skills to become an effective administrator at some time in the future. There are areas that would pertain to both types of centre (i.e. stress tolerance, interpersonal communication skills). However, there are no other centres which could be found in the literature who explore these skills prior to entrance into a training programme for administrators. All the existing centres examine the finished product, in the hope of putting that product to work in the near future. This is the vacuum which UT wanted to fill by creating its own AC centre.

Critical Analysis of the Reviewed Literature

It is interesting to see how many methods/processes professionals and researchers have developed in their effort to choose the most suitable persons for administrative positions in public education. The interview process is not the most suitable as it has often been reported. As Baltzell *et al.* [3-5] report, "The weight screeners tend to ascribe to their interviews probably exceeds the power of prediction that can be gained under the best of circumstances" (p. 62). Therefore, the interview method will not be the subject of examination of this section. This portion of the article will rather consist of a critical review of the other methods presented and the assessment centres in particular. The critical reviews will be presented in the same order as the programmes were presented previously.

The Administrative Intern Programme in California has been quite successful, but not all districts can afford to include the difference activities required by the programme and in addition to provide internships. The same applies for the Administrative Training Programme in Maryland which involves internships. What if, at the end of the internship period, the intern proves not to be the suitable candidate? By that time, the administration at the intern's school has been negatively affected. In addition, the district needs to start searching from the beginning to find a better candidate and thus, valuable time has been lost.

On the other hand, the Florida (Broward County) process is nothing but an elaborate form of interviewing. However, the clues provided about merit, sex, racial equity and legitimacy are valid considerations which any method could follow.

Overall, the various internship models can be seen as the main alternative to ACs. In some cases internships give better indications about job performance because the intern is actually on the job. However, they are time-consuming and expensive and not flexible when there is an immediate vacancy which needs to be filled right away.

The other alternative, advocating the continuous presence of an evaluator, proposed by Brush and Schoenfeldt [6], does not seem to be feasible, mainly because one would need full-time managers on the job doing nothing else but evaluating. It is next to impossible for managers to devote the time and energy necessary to gather reliable information for evaluation purposes.

Thus, ACs seem to be the most viable alternative for personnel selection. However, ACs have some disadvantages as has been explained in the literature review. Besides costs, coaching and other disadvantages that have been explained there is a more serious problem. What do ACs measure and how well do they predict? As Schmitt *et al.* [11] and Gomez [7] report, the NASSP skills do measure uniquely what they are supposed to be measuring. In other words, there is no multicollinearity among the variables (dimensions). However, those results contradict the Ogawa and Oxaal [12] results which indicated a large degree of multicollinearity among the skills/dimensions of the NASSP AC. Therefore, the skills/dimensions area of the AC are still questionable as to what they measure and if they measure it uniquely. Furthermore, the statistical analyses and interpretations performed in some of the studies [7,11] are, at first glance, excellent. However, the researchers subjectively explain discrepancies when the results do not conform to their expectations.

Moreover, in Gomez's evaluation [7], Dr Schmitt was used as a consultant and several of his instruments were used in obtaining the data. Thus, one would be able to conclude that Gomez's evaluation of the Dade County Public

Schools' AC is not all that objective and independent. Therefore, one is more inclined to accept Ogawa's and Oxaal's results[12], or demand that more research is done in order to establish the validity of ACs.

Also, a close examination of the literature suggests that significantly higher correlations exist between AC predictions and job potential ratings (median $r = 0.63$) than between AC ratings and actual job performance (median $r = 0.33$)[6]. In relation to the above, it is interesting to note that Gomez[7] reports that when comparing the interview process with the management AC of Dade County Public Schools process, none proves to be superior to the other. His explanation for this outcome is that, given the minimum passing score of the management AC, few candidates are eliminated, probably the same ones who would have been eliminated by using the interview process. Thus, he recommends that the minimum passing score of the management AC be raised, therefore, increasing its screening capability. The author of this article agrees with Gomez that the power of an AC lies in being able to screen candidates by using more rigorous techniques and higher passing standards. However, even with the above disadvantages, ACs seem to offer a credible alternative to more questionable hiring practices such as interviews.

In evaluating the UT-AC one should be aware of two factors: first, it is one of its kind at a university setting and fairly new and second, as Dr Veir said, the ordinary ACs are designed to screen people who already possess the necessary skills which are required by the position they are seeking, whereas the UT-AC is designed mainly to search for potential.

Bearing in mind the above, the UT's AC effort is very commendable. However, certain questions are raised in order to increase its validity:

- (1) How valid is this screening method? (Dr Veir said that they are in the midst of gathering data which will be analysed.)
- (2) How sound is the practice of involving people from the field to function as assessors? Dr Veir said they want people from the field in order to get a broad field view perspective. However, it seems that the danger of perpetuating bad practices from the field is there.
- (3) Is one day of assessment (this is how long the UT-AC is) enough? Is it not threatening to the candidates? Dr Veir responded that "If you cannot tell in one day, then how long would it take?" Well, this is the point of view which advocates of the interview method have as well.

These questions are being raised in an effort to make the process more researched-based and hopefully more valid when the data are examined.

Conclusions — Discussion

The three purposes stated at the beginning of this article were fulfilled: a brief description of different methods of principal selection was given; a description of ACs in general and the UT's AC in particular followed. The different methods described at the beginning were presented as current alternative practices, but the main scope was to look at ACs and at the UT-AC in particular as a method of "preventing medicine". It is probably cheaper and less time-consuming to establish more university-based ACs than wait until the candidate is applying for a job.

Nevertheless, the AC method seems to be a very promising method for personnel selection. There are certain disadvantages and limitations as described in previous sections of this article, but, this method is more comprehensive, rigorous and complete than other current methods of personnel selection. However, ACs should be further investigated especially in the area of multicollinearity among the skills being measured. Research will probably provide us with fewer than the 12 skills of a good principal currently being measured at the NASSP AC.

UT's AC is an excellent effort (one of its kind) to screen people before they even enter the field of public school administration. It is a proactive method. Since it is a fairly new event, it should be further researched and probably altered. However, the AC process is a more rigorous way to screen people because it is not just a two-hour paper and pencil test or a typical 30-minute interview; rather it is a whole array of activities where every candidate is given the chance to shine in one area or the other.

Recommendations

- (1) One recommendation accrues from the fact that in all the literature reviewed, no study was found about on-the-job counselling and training as a follow-up of the selection process (except measuring effectiveness or non-effectiveness on job performance). However, it is important to find out how many of those hired need or express the need for further counselling and guidance on their jobs. This could serve as a comparison measure between different selection processes.
- (2) Also counselling of those who were not selected for the position in a district is very important, especially if the process of selection is very competitive and the candidates are employed by the district. It is important to convey the message to the unsuccessful candidates that their non-selection has nothing to do with their abilities or qualifications, but that it is mainly a match with the specific position's requirements[3,4,5].
- (3) Furthermore, more research should be undertaken to examine what is being measured and to what

degree. However, even if fewer (than the NASSP 12) skills are found important for predicting performance or the existence of other factors is uncovered, Thornton and Byham[13] insist that the information from all 12 skills of the NASSP AC should be used because:

- Non-linear and configured judgement patterns which are not accounted for in a statistical analysis may exist,
- Behavioural information reported and discussed on less important skill dimensions may contribute to the predictive accuracy of the overall rating,
- Unique points of view may be uncovered, owing to the individual differences in weighted dimensions which may, in fact, be a strength and,
- Dimension ratings are important in further development and diagnostic progress.

However, the fact that Ogawa and Oxaal[12] found one factor subsuming all other dimensions is very disturbing. As the authors pose the question,

Could it be that assessors make overall intuitive judgments about candidates and bring scores on the specific skill dimensions into line with the initial general assessment? If such were the case, then it is possible that ACs simply legitimize personnel decisions based on such traditional criteria as social similarity to those individuals who make personnel decisions (p. 16).

- It is possible that trainers could be better trained to rate more independently (without a halo effect in mind). This possibility needs to be further researched.
- In addition, think-aloud protocols could be employed to trace the decisions and the information used in making those decisions as assessors rate candidates on skills and overall performance. Think-aloud protocols have been proved to be a very objective source of data[15-17].
- Another recommendation for making the AC process more sound and to guard against previous participants coaching candidates, would be for ACs around the country (or even the world) to develop a pool of materials and exchange them so that there would be some kind of rotation and interchange, thus safeguarding the validity of the process.

Further research is definitely needed about ACs. However, they do promise to create personnel selection practices that are based on more sound scientific evidence.

References

1. Duttweiler, P.C. and Hord, S.M., *Dimensions of Effective Leadership*, SEDL, Austin, TX, 1987.
2. Van der Perre, C. and Vandenberghe, R., *Training Programs for School Leaders in Europe*, Catholic University, Leuven, Belgium, 1984.
3. Baltzell, D.C., and Dentler, R.A., *Selecting American School Principals: A Sourcebook for Educators*, Abt Associates Inc., Cambridge, MA, 1983.
4. Baltzell, D.C., and Dentler, R.A., *Selecting American School Principals Research Report*, 31 January 1983.
5. Baltzell, D.C. and Dentler, R.A., *Selecting American School Principals: Executive Summary*, 31 January, 1983.
6. Brush, D.H. and Schoenfeldt, L.F., "Identifying Managerial Potential: An Alternative to Assessment Centers", *Personnel*, May-June 1980, pp. 68-76.
7. Gomez, J.J., *Final Report on the Evaluation of the Management Assessment Center*, Dade County Public Schools, Miami, FL, 1985.
8. Task Force on Assessment Center Standards, "Standards and Ethical Considerations for Assessment Center Operations", *The Personnel Administrator*, Vol. 25, 1980, pp. 35-8.
9. Hersey, P.W., "NASSP's Assessment Center — From Concept to Practice", *NASSP Bulletin*, Vol. 61, 1977, pp. 74-6.
10. Ogawa, R.T., *NASSP Assessment Centers: Opportunities for Research*, paper presented at the Annual Meeting of the American Educational Research Association, New Orleans, LA, April, 1984.
11. Schmitt, N., Noe, R., Meritt, R., Fitzgerald, M. and Jorgensen, C., *Criterion-related and Content Validity of the NASSP Assessment Center*, National Association of Secondary School Principals, Reston, VA, 1983.
12. Ogawa, R.T. and Oxaal, I., *A Reexamination of NASSP Assessment Center Ratings: Research Notes*, a paper presented at the annual meeting of the American Education Research Association, San Francisco, CA, 1986.
13. Thornton, G.C. and Byham, W.C., *Assessment Centers and Managerial Performance*, Academic Press, New York, 1982.
14. Sackett, P.R. and Hakel, M.D., "Temporal Stability and Individual Differences in Using Assessment Information to Form Overall Ratings", *Organizational Behavior and Human Performance*, Vol. 23, 1979, pp. 120-37.
15. Payne, J.W., Braunstein, M.L. and Carroll, J.S., "Exploring Predecisional Behavior: An Alternative Approach to Decision Research", *Organizational Behavior and Human Performance*, Vol. 22, 1978, pp. 17-44.
16. Einhorn, H.J., Kleinmuntz, D.N. and Kleinmuntz, B., "Linear Regression and Process-tracing Models of Judgment", *Psychological Review*, Vol. 86, 1979, pp. 465-85.
17. Ericsson, K.A. and Simon, H.A., "Verbal Reports as Data", *Psychological Review*, Vol. 87, 1980, pp. 215-51.

Further Reading

- Balmores, N.R., *The Use of Field-based Experience to Train Educational Leaders: A Historical Analysis*, a paper presented at the annual meeting of the American Educational Research Association, New Orleans, LA, 1988.
- Baltzell, D.C. and Dentler, R.A., *Local Variations in the Selection of School Principals*, a paper presented at the annual meeting of the American Educational Research Association, New York, 1982.
- Baltzell, D.C. and Dentler, R.A., *School Principal Selection Practices: Five Case Studies*, 1982.
- Boehm, V.R., *Establishing the Validity of Assessment Centers*, Development Dimensions Press, Pittsburgh, PA, 1982.
- Bray, D.W. and Grant, D.L., "The Assessment Center in Measurement of Potential for Business Management", *Psychological Monographs*, Vol. 80 No. 17, 1966.
- Geismar, T.J., Morris, J.D. and Weppner, D.B., *Selecting Mentors for Principalship Interns*, a paper presented at the annual meeting of the American Educational Research Association, New Orleans, 1988.
- Hersey, P.W., "Selecting and Developing Educational Leaders: A Search for Excellence", *NASSP Bulletin*, Vol. 69, 1986, pp. 1-2.
- Hinrichs, J.R., "An Eight-year Follow-up of a Management Assessment Center", *Journal of Applied Psychology*, Vol. 63, 1978, pp. 596-601.
- Huck, J.R. and Bray, D.W., "Management Assessment Center Evaluations and Subsequent Job Performance of Black and White Females", *Personnel Psychology*, Vol. 29, 1976, pp. 13-30.
- Janis, S.L., *Victims of Groupthink*, Houghton Mifflin, Boston, MA, 1972.
- Job Descriptions for Principals and Assistant Principals: The Current Trends*, NASSP, Reston, VA, 1979.
- Jones, R. and Godfrey, G.E., "Increasing One's Potential to Become an Administrator", *NASSP Bulletin*, Vol. 69, 1986, pp. 20-2.
- Landholm, D.S., "Center Helps One's Monitoring of Strengths, Weaknesses", *NASSP Bulletin*, Vol. 69, 1986, pp. 24-5.
- McCall, D.S., "Twelve Skill Dimensions — Professional Benefits", *NASSP Bulletin*, Vol. 69, 1986, pp. 32-3.
- Meehan, M.L., *Training Principals to be Classroom Management Trainees: Yes, You Can Teach Old Dogs New Tricks*, paper presented at the annual meeting of the American Educational Research Association, Washington, DC, 1987.
- Mitchel, J.O., "Assessment Center Validity: A Longitudinal Study", *Journal of Applied Psychology*, Vol. 60, 1975, pp. 573-79.
- Nisbett, R. and Ross, L., *Human Inference: Strategies and Shortcomings of Social Judgment*, Prentice-Hall, Englewood Cliffs, NJ, 1980.
- Office of Educational Accountability, *Preliminary Report on the Evaluation of the Management Assessment Center*, Dade County Public Schools, Miami, FL, 1984.
- Payne, J.W., "Task Complexity and Contingent Processing in Decision Making: An Information Search and Protocol Analysis", *Organizational Behavior and Human Performance*, Vol. 16, 1976, pp. 366-87.
- Payne, J.W., "Contingent Decision Behavior", *Psychological Bulletin*, Vol. 92, 1982, pp. 382-402.
- Personal interview with Dr Carole Veir, 20 October 1988.
- Pritchett, J.H., "Process Can Enrich Management Techniques", *NASSP Bulletin*, Vol. 69, 1986, pp. 34-5.
- Reilly, R.R. and Chao, G.T., "Validity and Fairness of Some Alternative Employee Selection Procedures", *Personnel Psychology*, Vol. 35, 1982, pp. 1-62.
- Russo, J.E. and Rosen, L.D., "An Eye Fixation Analysis of Multialternative Choice", *Memory and Cognition*, Vol. 3, 1975, pp. 267-76.
- Rutherford, W.L., Hord, S.M. and Thurber, J.C., "Preparing Principals for Leadership Roles in School Improvement", *Education and Urban Society*, Vol. 17 No. 1, 1984, pp. 29-48.
- Wendell, F.C., "Individuals Benefit from Exercises, Feedback", *NASSP Bulletin*, Vol. 69, 1986, pp. 23-4.