The development of behaviorally-anchored rating scales (BARS) will permit researchers measuring the effectiveness of school administrators to utilize sound, ethnographic methods instead of relying on traditional, theory-based rating systems. In this document the author describes his own modified process for developing BARS and explains the strengths and weaknesses of the BARS approach. The author asked five school principals to develop lists of essential administrative tasks and to describe in their own terms practical behaviors that were effective and ineffective in performing those tasks. The author compiled the results and is currently testing their validity with a second group of principals. When his project is concluded, the author will have a list of 14 tasks, with related behaviors ranging from the effective to the ineffective. The BARS approach is cited as encouraging descriptive validity, providing for adequate descriptive language, and taking into account the social meaning of behaviors, three factors that empirical behavior measures tend to overlook. Shortcomings of BARS, according to the author, include its dependence on multiple criteria, the complexity and time-consuming nature of the scale-developing process, and the subjectivity involved in interpreting measurement results. (PGD)
INTRODUCTION

The author has conducted research with school principals in which he has attempted to develop measures of administrative effectiveness. The methodology employed in this research follows the approach generally referred to as ethnography.* There are a number of conceptual and methodological problems associated with any attempt to develop measures of a construct, such as effectiveness, no matter what the methodological approach used. Such problems may be easy to understand and appreciate at the conceptual level but they are difficult, if not impossible, to resolve at the operational level. These problems seem to persist whether one uses a logical empirical or a phenomenological (in this research ethnographic) approach.

Many of the traditional ways of investigating administrative behaviour (e.g. the questionnaire method investigating pre-conceived categories of behaviour) have not been able to reveal much of what Sargent and Belisle (1957, p.3), a quarter of a century ago, referred to as the "real stuff" of educational administration. They stated:

All... efforts to discover, organise, and relate abstract knowledge about administrative

* For an explanation of the Ethnography Approach see Duignan (1981)
behaviour rests ultimately in the depth and accuracy with which they tap into the stuff of administrative behaviour, which consists of a particular administrator behaving in a particular situational sequence of interactions.

There has been a marked reluctance by researchers in educational administration to allow the practising administrator to tell his own story in his own words. Shulman (1970, p. 377) stressed that the best way to study job-related behaviour is to do so in situations that are as natural and close as possible to real life. He supported the notion that the researcher should focus on “active life situations” so as to “maximize the similarity between the conditions in which they study behaviour and those other conditions to which researchers may ultimately wish to make inferences.” The Behaviourally Anchored methodological procedures used in this research, in the opinion of this researcher, tap into the “real stuff” of administrative behaviour and allows the practising administrator “tell his own story in his own words”.

DEVELOPING BEHAVIOURALLY ANCHORED MEASURES OF EFFECTIVENESS (BARS)

Instead of taking a theory of administrative behaviour and its underlying causes, this approach starts with a description of actual behaviour. These measures are based on a naturalistic description of behaviour. The author will first describe the steps involved in carrying out the generally accepted procedures of BARS as reported by Bennett and Langford (1979, p. 64-66). However, in its original form, this approach is time consuming and costly. Because of these problems, many modifications of the approach have been used in Management. This author has modified the approach substantially in the study of the effectiveness of school principals without, in his opinion, compromising its key strengths.

The technique involves asking a small group of practising administrators in this study principals — to state in their own terms the qualities, traits or criteria which are important in carrying out their work effectively. From these descriptions the researcher can identify the essential tasks or qualities of the job as seen through the eyes of the practitioner. The next step is to ask the same group to provide examples of behaviour, drawn from their experience and observation, of “effective” and “ineffective” behaviour in relation to each task or quality. The researcher may at this stage add the words “could be expected to” in preparation for the next stage. However, some researchers prefer to leave these words out as they tend to shift the focus of the practitioner’s responses from what he actually does to what he might or should do.

The next step is to ask another group of principals to “retranslate” these statements of behaviour by allocating them to the original criteria. Some statements may be dropped at this stage if there is no clear agreement on which criteria they belong to, and criteria may be similarly ignored if behaviour examples are not consistently reassigned to them.

The final step requires yet another independent group of principals to rank the behavioural statements on a five- seven- or nine-point scale according to their perceived degree of effectiveness or ineffectiveness. Those which command little agreement are again dropped, leaving a rating scale to which points of assessment are “behaviourally anchored”.

A Modified Approach to BARS

The following steps were used by this research in developing behaviourally anchored measures for school principals.

1. A group of five principals was asked to identify the tasks they saw as “essential” in carrying out their jobs effectively. The researcher then developed a list of the most commonly

* Some Behaviourally Anchored Studies of Effectiveness in management have also included “average behaviour”, but as Bennett and Langford (1979, p. 65) pointed out, it is difficult to get adequate statements of average behaviour since it is a very human tendency, when perceiving and describing behaviour, to emphasis the extremes rather than the middle range.” The category “average behaviour” was excluded in this study of principals.
mentioned tasks.* These were: problem solving; decision making, communicating with staff; planning, delegating to others, introducing innovation; evaluating; educational leadership; managing conflict; managing time, managing stress; relating to others; supervising staff; and allocating resources.

2. These same principals were asked to provide in their own words, (in written statements and in an interview situation) examples of "effective" and "ineffective" behaviour in relation to these tasks.

3. The researcher condensed their behavioural statements into a composite profile of effective and ineffective behaviours in relation to each task. Example of such statements of behaviour in terms of the task COMMUNICATING WITH STAFF were:

   Effective behaviour. "Balancing face-to-face communication with other methods — memos, and formal informal communication."

   Ineffective behaviour. "Using largely impersonal modes, e.g. using written notices and the P.A. system as almost the sole means of communicating with staff."

4. The composite profile of behaviour was presented to the same group of principals for comment. They were told that this profile was the "researcher's construction" using the statements and language of the principals.

5. The modified profile was given to an independent group of five principals for their reaction. They were told that the behaviours included in the profile were examples of effective and ineffective behaviour provided by a group of colleagues. These principals were asked to comment on the "appropriateness" of the behavioural statements, from their own experience and observations. They were encouraged to provide behavioural examples of their own if they disagreed with those on the profile. The research is at present at stage 5. The remaining steps will be described as they will be carried out.

6. The profile will again be modified to include any "common themes" expressed by this group of principals. The modified profile will then consist of fourteen administrative tasks with related "effective" and "ineffective" behaviour.

* These categories of behaviour are referred to here as tasks. Some, no doubt, would question the use of this label. While some of these behaviours could be regarded as processes the term task best describes them for the purpose of this study.
7. A further refinement of this instrument will be to place the behaviours related to each task at either end of a five-point rating scale. An example is given in Figure 2. The principals from the two groups involved in the study (ten principals, five in each group) will be asked to provide examples of (a) good performance (b) average performance, and (c) poor performance. The researcher feels that examples of such behaviour will be easier to provide once the extremes of “extremely good” and “extremely poor” performance are established. It would, in the researcher’s opinion, be too demanding, if not extremely confusing, to obtain examples of all these categories of performance in Step 2 as outlined earlier.

8. In the final step, the researcher will compile “representative” behaviour statements for each of the points on the scale (the extreme points will already have been established). The final instrument will contain fourteen such scales (one for each task identified earlier) with their behaviourally anchored statements of various degrees of effective and ineffective behaviour. Such scales could be used by principals themselves as self-analysis instruments.

ADVANTAGES OF BARS APPROACH

In several respects, the behaviourally anchored approach, although still requiring considerable development, is an attractive alternative to conventional rating methods. Because practising administrators are heavily involved in the construction of this instrument, the tasks and the behavioural examples should have a high degree of meaning for the users — the practitioners themselves. The approach ensures that only descriptively valid items are included in the instrument because generalisations are carefully made from specific illustrations of effective and ineffective behaviour.

The Importance of Descriptive Validity in Qualitative Research

Descriptive validity should be a prime concern of a researcher using an ethnographic approach. Erickson (1978, p. 4) stated that descriptive validity is essentially “an issue of functional relevance, from the actor’s point of view.” Descriptive validity has frequently been neglected, ignored or taken for granted in research studies of administrative behaviour. In the tradition of the logical empirical or positivist approach, the actor’s point of view was, frequently, not sought and the interpretation of behaviours and events was made by the researcher within the framework of an established theory. The findings of such research, devoid of any input from respondents with regard to the meaning of events, have, generally, not been received with enthusiasm by practitioners. In such research endeavours, practitioners are often suspicious of the researcher’s intentions and are sceptical of his ability to understand what life is really like on the “firing line”. These suspicions and doubts contribute to the problem that is generally referred to as the theory-practice gap.

The problems of the so-called theory-practice gap and of validity are further compounded by the fact that the researcher and the practitioner often use different languages of description. The researcher often imposes his own world view on his observations by the very language he chooses to describe them.

The need for an adequate language of description. Over thirty years ago Cassirer (1946, p. 28) noted that it is only through the medium of language that we can look meaningfully at our world. He stated:

All theoretical cognition takes its departure from a world a ready pre-formed by language; the scientist, the historian, even the philosopher, lives with his objects only as language presents them to him.

The language chosen by the researcher to describe his observations can “colour” their real meanings. Karl Popper (1959, p. 59) reminded us that there is no purely “phenomenological language”, no language to describe something sensed and not interpreted. Fearing [commenting on the ideas of
the anthropologist Benjamin Whorf (cited in Von Bertalanffy, 1973, p. 235) supported this point of view when he stated:

The commonly held belief that the cognitive processes of all human beings possess a logical structure which operates prior to and independently of communication through language is erroneous. It is Whorf’s view that the linguistic patterns themselves determine what the individual perceives in the world and how he thinks about it.

How can the researcher resolve this language problem and at the same time increase the validity of his descriptions? Erickson (1978, p. 1) offered some advice. He suggested that we need to develop an appropriate data language for the description of real-life events, because our present vocabulary and syntax of description is based on an inadequate analysis of actual behaviour. He emphasised that for narrative description to be valid it must be derived from an adequate functional theory of (administrative) behaviour. In order to develop such a theory, we must, first of all, develop a language of description which will enable us to establish the underlying social meanings involved in the face-to-face interactions at the firing line. The need to understand the social meaning of behaviour. A number of authors (Vidich, 1955; Wilson, 1977; McCutcheon, 1978; Erickson, 1978; Anguera, 1979) have emphasised that human behaviour is significantly influenced by the settings in which it occurs. Many aspects of human behaviour, to be understood, must be interpreted in terms of their social importance. Wilson made this point rather forcefully when he suggest ...

Sociologists studying in organisations assert the importance of the traditions, roles, values, and norms that are part of life in organisations. Much behaviour in organisations is influenced by the participants' awareness of these mental states and by pressures generated by others who are influenced by these states.

The researcher who is investigating and trying to interpret the meaning of administrative behaviour must endeavour to understand what Erickson (1978) referred to as the social organisation of the behaviour he is observing. He can describe a principal's behaviour in terms of physical acts at a low level of inference, e.g., “Who does what, with whom, when, and where?” However, such mere description is an inadequate foundation for generating theories of action with regard to administrative behaviour. Trow (1977, p. 35) pointed out that “The data he (researcher) collects are not a substitute for the interpretive influence.” The researcher, then, is obliged to explain his observations within the framework of the social organisation from which they are taken. Erickson (1978) exhorted social science researchers to develop an adequate theory of the nature of social relations in everyday life because he believes that human behaviour is always mediated in social relationships. He (1978, p. 4-5) propounded that the theory entailed in the description of a connected sequence of events across time “is in essence a theory of its social organisation.” By social organisation he meant the meanings participants ascribe to events based on their roles in the organisation.

The problems of developing an adequate language of description and of capturing the social meaning and organisation of behaviours and events are, therefore, important considerations for the researcher concerned with descriptive validity and with making conceptual sense of the world he observes. These problems are also closely related to each other at the operational level. The problem of developing an adequate language of description is closely linked to the problem of understanding the social meaning and organisation of events. Erickson (1978, p. 5) suggested that the theory of social organisation contained in any description is inextricably caught up with the language used. He stated:

The theory of social organisation entailed in description is embedded in the key terms and relations contained in the description; in the very nouns and adjectives one chooses as labels for the cast of characters ... and in the verbs and adverbs one chooses as labels for the kinds of actions those characters perform together ... Such theory is also embedded in the descriptive syntax accounting for sequence relationships among the actions, in the points of functional climaxes or crises identified in those sequences, and in the terms indicating standards for judgment of the social appropriateness of those actions.
The researcher's task in overcoming these problems — developing a language of description which reflects the social meaning and organisation of observed behaviour — is a difficult one. The effort, however, is worthwhile because it is one way in which he can hope to allay the practitioner's fears and suspicions and, thereby, try to bridge the theory-practice gap. Using an ethnographic approach, e.g., BARS, the researcher can begin developing his language of description by allowing practitioners to tell it as they see it, describing their understanding of Being-in-the-World.

In addition to increasing the descriptive validity of the effectiveness measures, the BARS approach has a number of other advantages.

The specificity of the behavioural examples allows the user to match his own behaviour against effective behaviour as described by a number of his colleagues. In this way, well designed BARS can highlight weaknesses in performance and help identify training needs. Such behavioural statements prefixed by the words 'could be expected to' can provide clear directions for improvement to the user.

More importantly, the Behaviourally Anchored approach helps to bridge the gap between theory and practice. The close relationship that develops between the researcher and the practitioner; the high level of involvement of the practitioner in the construction of the instrument; the use of the practitioner's "working language" in the behavioural examples; the opportunity for the researcher to check his interpretations and generalisations (and thus the social meaning of the practitioner's descriptions) with different groups; and the multiple validation checks (feedback from initial group and later another, independent group) help to break down the barriers of suspicion and scepticism that often exist between the researcher and practitioner.

The development of the composite profile and the provision of feedback from those whose behavioural statements are included in the profile and are key validation techniques. The use of such profiles and feedback mechanisms is similar to an approach which has received wide acceptance as a method for developing descriptively valid documentary films. The technique was used with great effectiveness by the National Film Board (NFB) of Canada in its Challenge for Change Programme. The approach had its greatest impact in influencing social change on the island of Fogo off the Northeast coast of Newfoundland, and so became known throughout the world of documentary film as "The Fogo Process." The key to the success of this technique was the use of vertical films (vertical studies) which portrayed "linear chunks of the reality" of life on the island. A brief examination of The Fogo Process follows.

The Fogo Process: Emphasis on Process not Product

In the 1960's a majority of the people on Fogo Island — a small isolated island divided by religion, occupation and geography — were unemployed and conditions were reaching a crisis point. The Government offered the residents an option — resettlement on the mainland.

The National Film board, in conjunction with Memorial University, under the direction of a Mr. Colin Low decided to make a film of the resettlement using the conventional documentary approach — focusing on key issues. Before long, however, Low changed his methods drastically. Low, [reported in Gwyn, (1972, p. 5)] explains why:

When I went to Fogo, I thought that I would make one, or perhaps two or three films. But as the project developed, I found that people were much freer when I made short vertical films: each one the record of a single interview, or a single occasion.

The final result of Low's effort has twenty-eight short films, each centred around a personality or an event rather than an issue, each expressing an aspect of life on Fogo island.

Low's next steps were fundamental to the success of his documentary venture and were to be copied in ensuing years by documentary filmmakers all over North America. He screened each film for the participants and allowed them to suggest

* A programme designed to "improve communications, create greater understanding, promote new ideas and provoke social change."
changes, additions, and deletions. Later he screened the edited versions to wider audiences using their suggestions to create a "true picture" of life on the Island. In this way, opposing factions within the communities were given a chance to share each others views. More importantly, the participants in the film were able to view and modify their own responses and the filmmaker's interpretation of their world. The filmmaker in turn was able to check his interpretation of the social meaning and organisation of the behaviour and events portrayed on the film.

Far more important than the films themselves, therefore, was the process of making them (prior to this, documentary filmmakers were more concerned with the product, a fault of educational researchers as well) and more important still, according to Low, was the process of validating them through trial screenings with the participants and others. The end result was that through looking at each other and themselves, Fogo Islanders began to recognise the commonality of their problems as important. New economic initiatives were taken by the Islanders and resettlement became unnecessary.

The lessons learned by the filmmakers on Fogo are lessons that most educational researchers would do well to note. They learned that:

1. it is a necessity that the filmmaker be familiar with the setting and its actors. Low used a native of Fogo to familiarise him with "the mystique" of the Island.
2. the participants in the film must be involved in the selection of the material which purports to represent their world. It is not the filmmaker's view of their world that must be depicted on film, but their description and interpretation of Being-in-their-World.
3. vertical films were effective instruments for portraying various aspects of life on the Island. Because of their narrow focus and short duration, they were found to be ideal for generating discussion and feedback from the participants. In this way, the participants in the film make the judgment on whether or not the image captured on the screen is an accurate representation of reality.

The use of different groups by the filmmaker for providing feedback provides a test of what McCutcheon (1978) referred to as Referential Adequacy. Referential adequacy is a test of how well an account of an event accords with the real world. Criticism of films, books and music can be subjected to this validation test because the object being reviewed is accessible to a wide audience of critics who can judge for themselves the accuracy of the critic's account. The greater the referential adequacy, the higher the validity of the description. Descriptive validity is undoubtedly enhanced through the process of "playback for feedback" of vertical films. The importance of this playback technique as a validation check should not be underestimated. This is how the cricket announcer can be so much more certain than the umpire of the "reality" in a close call for stumping. The umpire, unlike the announcer, doesn't have the luxury of instant replay to check the validity of his decision.

As educational researchers, we cannot often afford the luxury of recording observed behaviour on film or tape for more detailed examination at a later time. Neither can we afford the time to spend lengthy periods of time observing, interacting with, practitioners in their natural surroundings. The methodology used by this researcher in the development of BARS for school principals may help the educational researcher overcome these problems. The process used is similar to that used by Low in the Fogo Process. The initial group of principals generate the behavioural examples of effective and ineffective behaviour which form the basis of the composite profile; which is akin to Low's vertical film. In fact, this approach to the development of effectiveness measures can be justly called a vertical study. The composite profile is "played back" to the principals who participated in its construction and modifications are made—just as Low screened his vertical studies to the participants for feedback. Later an independent group of principals are asked to provide further feedback on, and clarification of, the descriptions presented to them. Low did likewise. And so the process goes on (as described earlier) until the final product is produced—a product subjected to rigorous tests of referential adequacy and descriptive validity.
There is, however, much developmental work left to be done in refining the Behaviourally Anchored procedures. The approach has a number of shortcomings.

**SHORTCOMINGS OF BARS APPROACH**

Some of the shortcomings are common to all measures based on multiple criteria. Most approaches to assessing effectiveness rely on a series of relatively discrete criteria (for example, ability to communicate, relationships with others, success rate of students). This use of multiple criteria for assessing effectiveness can be somewhat problematic. It is possible that in some situations these criteria may be in conflict. For instance, a task oriented principal may be "effective" because he is able to obtain good results on matriculation. However, these results may have been obtained because of the pressure he placed on staff to perform at a high level. Such pressures could lead to a straining of relations. Within a limited time perspective; a principal may be high on one criterion and low on another and still be classified, at least by some audiences, as effective. It is unrealistic to expect that a principal must score positively on all effectiveness criteria at any one time in order to be effective.

The development of BARS is a time-consuming and involved process. Many demands are placed on the cooperating practitioners and on the researcher. There is still a large element of subjectivity involved in the use of the instrument, i.e., someone still has to rate somebody else (unless it is used as a self-evaluation tool) on a scale of 1 to 5. Collapsing down complex behaviour to conform to a discrete point on any scale is a very subjective exercise. However, the fact that in BARS each point of the scale is accompanied by a "typical" specific behavioural example, helps, at least partially, overcome this problem.

There is a danger that BARS as an instrument might consist of no more than a series of loosely knit behavioural statements about general effectiveness. In order to avoid this, the researcher has to play a crucial role. He has to act as the "photographer", "technician", and "editor". He must make interpretive decisions when developing the composite profile. He must be able to listen to and select from a wide variety of narrative descriptions, the "representative" behaviours for each task.

Wilson (1977, p. 281) highlighted the difficult task faced by the qualitative researcher when he stated:

> The qualitative research enterprise depends on the ability of the researcher to make himself a sensitive research instrument by transcending his own perspective and becoming acquainted with the perspectives he is studying.

The BARS approach, then, has a number of rough edges. In the words of Bennett and Langford (1979, p. 122):

> This is by no means a perfect technique. Subjectivity on the part of the designers and users cannot be ruled out, altogether. . . . There will never be universal panaceas, of course, for there can be no such thing. But these measures do represent a step forward toward establishing clear, concise and practical guides (for measuring performance.)

**CONCLUDING REMARKS**

The BARS approach shows promise as a means of developing effectiveness measures for school administrators. Using this ethnographic approach, the researcher should be able to generate descriptively valid behavioural statements of effective and ineffective behaviour. In order to achieve descriptive validity there is a need to develop a "language of description" which reflects the social meaning and organization of the behaviours and events being described by the researcher. The absence of such a language has, in the past, contributed to the theory-practice gap.

The researcher should allow the practitioner to "create" the language of description by telling his story as he sees it. The BARS approach uses this method in developing its behaviourally stated examples for the various administrative tasks.

The BARS approach is in need of further refinement. Even the modified version presented by this researcher is time consuming and places a great burden on both the practitioner and the researcher. However, the approach has potential. Perhaps its greatest contribution toward the development of the
practice of educational administration, could be in its use by practitioners as a self-analysis tool. The use of clear, concise and practical behavioural examples of effective and ineffective performance in relation to specified administration tasks, could be of great benefit to the practitioner who is motivated to improve his performance.

REFERENCES


INDEX TO STUDIES IN EDUCATIONAL ADMINISTRATION

6. A.R. Crane, The Preparation of Principals — The Role of Tertiary Institutions —


**STUDIES IN EDUCATIONAL ADMINISTRATION**

This series is published at Armidale, N.S.W., Australia by the Commonwealth Council for Educational Administration as a service to its members throughout the Commonwealth of Nations and beyond.

It normally appears four times a year.

Papers may be submitted for consideration to the Editor, c/o CCEA, Faculty of Education, University of New England, Australia, 2351. Each paper should be typed, double spaced on one side of the paper only. References should in general appear at the end of the work.

A limited number of back issues is available. These will be sent to CCEA members on payment of postage. Non-members should add $AUS 1.50 per copy.