

SPECIAL REPORT

MASTER COPY



PERS 179

THE ASSESSMENT OF BANTU SUPERVISORS:
A REVIEW OF STUDIES CONDUCTED TO DATE

NATIONAL INSTITUTE FOR PERSONNEL RESEARCH
COUNCIL FOR SCIENTIFIC AND INDUSTRIAL RESEARCH

CSIR Special Report PERS 179
UDC 658.311.56-052.23:622.342.1] (680=963)
Johannesburg, Republic of South Africa, September 1972

HSRC Library and Information Service

HSRC
Private Bag X41
PRETORIA
0001

Tel.: (012) 202-2903
Fax: (012) 202-2933



RGN
Privaatsak X41
PRETORIA
0001

Tel.: (012) 202-2903
Faks: (012) 202-2933

RGN-Biblioteek en Inligtingsdiens



HSRC Library and Information Service

RGN-Biblioteek en Inligtingsdiens

DATE DUE - VERVALDATUM

--	--

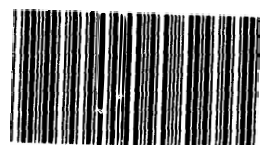
2:

ATE

M.A. Biesheuvel

NATIONAL INSTITUTE FOR PERSONNEL RESEARCH
COUNCIL FOR SCIENTIFIC AND INDUSTRIAL RESEARCH

September 1972.



* P B 9 6 5 0 1 *

ISBN 0 7988 0258 8 CSIR Special Report PERS 179

Published by
National Institute for Personnel Research,
South African Council for Scientific
and Industrial Research,
P.O. Box 10319,
Johannesburg, Republic of South Africa.

September 1972

Printed in the Republic of South Africa by
National Institute for Personnel Research

CONTENTS

	<u>page</u>
1. SYNOPSIS	1
2. THE INTRODUCTION OF APTITUDE TESTING ON THE GOLD MINES .	2
3. THE LEADERLESS GROUP SITUATION AS A MEANS OF ASSESSING SUPERVISORY ABILITY	3
4. EARLY RESEARCH ON THE LEADERLESS GROUP TEST	5
5. FURTHER RESEARCH AIMED AT IMPROVING THE LEADERLESS GROUP TECHNIQUE	9
6. REASONS FOR DISSATISFACTION WITH THE LEADERLESS GROUP TEST	15
7. PHYSICAL APPLICATION OF THE LEADERLESS GROUP TEST	16
7.1 The Administration of the Technique	16
7.2 The Use made of the Results Obtained	17
8. THE DEVELOPMENT OF A CRITERION	20
9. THE RESTANDARDIZATION OF THE LEADERLESS GROUP TEST	23
10. THE ATTEMPT TO DEVELOP A WEIGHTED APPLICATION BLANK	27
11. CONCLUSIONS	30
12. RECOMMENDATIONS	32
12.1 "Short-term" Recommendations	32
12.2 "Long-term" Recommendations	35
BIBLIOGRAPHY	37

1.

SYNOPSIS

It is envisaged that if the economic viability of South Africa continues to grow, as in fact it must, more effective utilization of the non-white labour force will have to be made. It is evident that a start in this direction has already been made through an increasing utilization of Bantu operatives in industry. However, the break-through is most likely to come from the employment of Bantu supervisors. Because of cultural similarity and consequently a common outlook, and mutual understanding, it is clear that Bantu will be more highly suited to supervising work groups of their own people than Europeans. For this reason, the identification of supervisory ability in Bantu is an area requiring particular attention in the future. This report reviews research done in South Africa to date in this field, but because the research has been virtually restricted to the Gold Mining Industry, it will deal almost exclusively with this domain. It is regarded as fortunate that such a large pool of information is available as a source from which hypotheses can be generated. Finally, certain recommendations are made as to approaches for future research into the Bantu supervisory function, with the aim of formulating a strategy for selection and training.

2. THE INTRODUCTION OF APTITUDE TESTING ON THE GOLD MINES

In 1946, the N.I.P.R. was commissioned by the Gold Producers Committee of the Transvaal Chamber of Mines, to introduce a programme of aptitude testing for the Bantu Labour on the Gold Mines. One of the chief purposes of the introduction of testing procedures was to enable a classification of the labour resources into broad occupational categories, such that a more efficient matching of ability with job demands could be achieved. Our interest in this report centres around the highest skill grade, the supervisory category. Particular attention was paid to selection for this category because of the great importance attached to the job of Boss Boy (the colloquial term given to the supervisory level) by mining officials. Firstly, the Boss Boy needs to be familiar with all tasks performed in his section. For this it is assumed that a certain minimal intelligence level is required, hence only those candidates whose scores on the screening battery fall into the upper duodec scale divisions are acceptable material. Secondly, he has to be able to handle a gang of labourers, organizing the team work so as to get the required work done. It was for the purpose of measuring this quality that the Leaderless Group Test (L.G.T.) was constructed.

Up to this stage, the selection of Boss Boys had been haphazard and influenced by many traditionally-held views in the mining industry. In so far as the introduction of the special test battery for the selection of the supervisory category was an improvement on the old system, and that selection had been placed on a more rational scientific basis by use of a standard instrument, the introduction of the L.G.T. was a worthwhile innovation. But although training-wastage rates had been substantially reduced, the test's total capacity for predicting effective Boss Boy performance had not yet been established.

It is the aim of this report to review and integrate the work which has been done from the time of the test's introduction, attempting to see how these problems were tackled, whether they were successfully resolved and in what way improvements could be implemented.

3. THE LEADERLESS GROUP SITUATION AS A MEANS OF ASSESSING SUPERVISORY ABILITY

The leaderless group task is a socio-dramatic situation. It is a simulation of a situation in which leadership characteristics of a supervisory nature are called for. A practical problem is presented to an unstructured group of six people, and its solution depends on the active co-operation and organization of the whole team of men. As the Boss Boy fulfills the function of supervisor of a team of workers, it was felt that this technique would be a suitable one to adopt.

As a technique for the selection of people with leadership potential, the leaderless group situation has been used for military selection by Britain, Germany and America (18) but assessments of performance were largely subjective. The aim of the N.I.P.R. was to give it the status of a psychometric test i.e. a standardised instrument discriminating reliably between individuals by means of a scale measuring the psychological constructs under consideration. Although this type of performance test has appeal, it poses serious problems for the test constructor in terms of its reliability and validity. In order to create a test situation which was an accurate simulation of the on-the-job situation and permitted generalization to the latter, consideration had to be given to including elements of the work situation into the problems set. The tasks were to be of a sufficient difficulty-level to stimulate enough interaction to allow the desired qualities to emerge, but not too difficult to preclude successful solution of the problem.

It is a frequently-expressed criticism of this technique, as applied in the mining industry, that the problems became so well-known that training schools were set up in mine compounds and in the homelands to teach prospective candidates how to solve them. If true, this allegation would be a valid criticism, as knowledge of the solution could make it easier for a candidate to overcome other inhibitory pressures preventing him

from taking the initiative, and place him in a position which would facilitate his ability to control the group. However, this is not a necessary relationship - solution to the problem is the 'apparent' rather than the true aim of the test. While the subjects are led to believe the aim of the exercise to be the correct solution of the problem, the real aim is to observe how members of the group set about solving the problem. The man who knows how to solve it, may not be capable of co-ordinating group action. However, certain other drawbacks to the application of this technique for the purpose of Boss Boy selection come to mind.

1. In each leaderless group situation, the performance of each member of the group is relative to that specific group situation - a product of the interactions existing within the group. An individual placed in a different situation amongst different people may easily display a different behaviour pattern.
2. It is a cultural norm of the Bantu people that age should be deferred to. A younger member of the group is therefore less likely to take command of a situation if the group contains older men. An attempt should therefore be made to keep the groups homogeneous as regards age.
3. In the Leaderless Group situation, all subjects are in mutual competition - a feature which has no part in the actual Boss Boy function.

4.

EARLY RESEARCH ON THE LEADERLESS GROUP TEST (L.G.T.)

Prior to the construction of tests, an analysis was conducted on all tasks performed underground. On the basis of information gathered through observation of the performance of jobs underground, and consultations with mining officials and Bantu mine-workers, an idea of the personality qualities considered desirable in a Boss Boy could be obtained. Once the Leaderless Group Technique had been decided upon as the most suitable method for assessing these qualities, several practical problem situations requiring team co-operation for their solution were constructed. The administration and scoring procedures for each of these situational tests was basically the same. Each group consisted of six testees, and two assessors were required to rate the performance of each group. Test assessors were required firstly to rate the group's performance on a three-point scale, and secondly to make check marks during the test session according to whether certain items of behaviour regarded as indicative of supervisory skill, were evident. The eleven characteristics to be rated were as follows: (a brief operational definition of each was available in the manual).

Dominance
 Activity
 Initiative
 Co-operativeness
 Organising ability
 Perseverance
 Responsibility
 Self-confidence
 Acceptability in the Group
 Planning Ability
 Practical Ability

However, the final leadership rating took the form of a subjective assessment on a 9-point scale of each testee's performance. Thus the actual checking of behaviour items was used only as a guideline for assessors; it was not a quantitative measure in itself.

The great problem facing investigators at this stage, and one which has characterized all research on the L.G.T., was to find a suitable criterion against which to validate it. Two different criteria have been used in validation studies. 1) A measure of performance in the training course and 2) A measure of job efficiency. The initial investigation was conducted at two different mines (2). On one mine, four criteria were decided upon:

- 1) Mine Captain's rating of the Boss Boys in their sections as inferior or superior.
- 2) Mine Captain's examination (a test of mining practices administered to all Boss Boys by two mine captains.)
- 3) An assessment on an N.I.P.R. rating scale by Shift Bosses whose Boss Boys had been tested.
- 4) A rating on the same scale by miners under whom the Boss Boys worked.

On the second mine slightly different criteria were used but they covered the same broad functional areas. Examination of the intercorrelations between criteria indicate that the reliability of these criterion measures was disappointingly low. The highest relationship found between two independent assessments was between Shift Boss's rating and Mine Captain's rating at the first mine ($r = ,588$, S.E. $\pm ,105$). Even when two assessments on the same Boss Boy were made by the same official with an intervening period of one month between assessments, reliability was low: for Shift Bosses $r = ,685$ and for miners, $r = ,654$. There appeared to be no consistent agreement as to what the Boss Boy function involved and the shortcomings of validating a test on these unreliable criteria were obvious. The only avenue open to investigators, was to demonstrate that the test showed as great an agreement with the proficiency criteria, as the latter could agree amongst themselves (2). Therefore multiple correlations between those criteria which showed the highest agreement amongst themselves

and the test scores were calculated. Validity coefficients although not reaching ,6 which was regarded at the time as "good", were nevertheless of acceptable magnitude. The investigators could claim no more for the test than that it provided as good a prediction of Boss Boy efficiency as any group of mining officials, and that, being a standard instrument, was likely to do so more consistently.

Between 1949 and 1952 revalidation studies were conducted on a number of different samples at different mines (14, 15, 16). Slightly different criteria were used. Hudson (5) reports using four criteria: 1) An occupational service criterion. 2) Training course results. 3) A trade test. 4) A rating scale for job performance. Reliability estimates were available for the trade test and the rating scale. For the former $r = ,93$ (obtained from the split half technique and the Spearman Brown formula). For the latter $r = ,70$ was obtained between first and second ratings by one assessor and $r = ,88$ between two independent raters. The various revalidation studies provided evidence that the selection battery was discriminating between different ability levels in the way required by the mining industry. The demonstration of this was provided as follows. When comparing the percentile ranks of the mean test performance of a representative sample of new-to-industry Bantu, with that of a sample of trained Boss Boys, 81% of the former group were found to score less than the average trained Boss Boy on the screening battery alone. In the same manner, 66% of this representative sample scored below the mean of a group of experienced workers nominated by mining officials for Boss Boy training. The addition of the L.G.T. to the initial screening battery produced an added improvement: the test performance of the average trained Boss Boy on the composite battery exceeded the test performance of 75% of those who were nominated for training (14).

The comparison of training wastage rates on a sample of Boss Boy trainees selected by three different methods, gave further

support to the method of test selection. The following table was based on 1 187 cases, representative of new-to-industry workers.

Comparison of Wastage Rates for Boss Boy
Trainees selected by different methods

Type of Selection	Expected	Observed
Chance Selection	=====	36%
Mines Selection	IIIIIIIIIIII	17%
Test Selection	===== IIIIII (15)	3,6% 3%

Thus the L.G.T. had justified its existence as part of the test battery. However, there was still room for improvement as revealed by the thorough investigation into its use and psychometric properties, conducted by Kruger (8).

5. FURTHER RESEARCH AIMED AT IMPROVING THE LEADERLESS GROUP
TECHNIQUE

The aim of Kruger's research was stated as follows:

- (1) The detailed analysis of different sorts of Boss Boy Tasks.
- (2) The leadership qualities in terms of which they should be selected.
- (3) Suitable test situations, test instructions and administrative procedures.
- (4) Determination of reliability and validity of suitable criteria.
- (5) The factors influencing test performance and the criteria.

(Kruger p. 14)

The project consisted of four different stages: Alpha and Beta stages, where largely theoretical considerations were dealt with and the Gamma 56 and 57 stages which concentrated on improving its characteristics as a measuring instrument.

Special consideration was given to developing meaningful instructions, laying down standard administrative procedures and developing a reliable and valid scoring system. To a large extent the success of the test as a psychometric device hinged on the latter. It became clear in this investigation that the L.G.T. is very susceptible to subjective interpretation.

In the final form, five different leadership factors (arrived at through detailed job analysis) were found to discriminate between good and bad Boss Boys. For four of these factors - active leadership, planning of work, improving work methods and personal relationships - assessments were to be anchored to certain critical incidents occurring at different stages of the test. By requiring that the test administrator merely indicate the appearance or lack of appearance of these behaviour items, the subjective assessment characteristic of previous

scoring methods was greatly reduced. To obtain some measure of inter-scorer reliability, product-moment correlation coefficients were calculated between the assessments made by pairs of test administrators for the first 50 groups tested. The obtained $r = ,967$ was an improvement on the $r = ,867 \pm ,012$ obtained with the original form (4) and indicates that an acceptable standard of objectivity had been reached. The 5th factor - acceptability - was rated by obtaining the candidate's nominations of whom in their group would be to them most acceptable as a Boss Boy. Using Kendall's coefficient of agreement (with a maximum value of ,735), $r = ,69$ was obtained, indicating also a satisfactory level of agreement.

Validating the test once again presented particular difficulties centering around the question of an effective and valid criterion. Finally, a weighted criterion was adopted, reflecting both the candidate's performance on the training course and his effectiveness on the job. The training test was a standardized oral test, the questions relating specifically to the various courses comprising the training period. The second criterion was a forced-choice rating scale, related to specific behaviours considered characteristic of a good Boss Boy. It was to be completed by the white mining officials under whom the Boss Boy worked. While both criterion measures showed acceptable reliability coefficients, on what basis was one to assume that they were a valid reflection of an efficient Boss Boy? A factor analysis of 13 sub-criteria, covering as many indices of Boss Boy performance as possible was considered as the best means of revealing a general underlying success factor. Three factors were obtained: the factor on which the greatest number of sub-criteria loaded was termed "general reputation" (the general impression of effectiveness made by the Boss Boy), and was taken as the underlying success factor. The combination of the two criterion measures into a weighted criterion was based on their loadings on this factor. ($r = ,44$ for the "Trade test" and $r = ,66$ for the rating scale). The multiple correlation between this factor and the weighted criterion was $r = ,709$.

In the cross validation, the validity coefficient was low, $r = ,232$ (product moment correlation) - just significant at the 5% level. Kruger argues that this is an underestimate of the relationship between the predictor and the criterion because the sample was homogeneous due to pre-selection. He then proposes that a correction is required in order to yield a more accurate estimate of the validity of the test. Because the restriction in range of the sample was due to both direct and indirect influences, it could not be numerically determined according to Thorndike's correction procedure. Thus Kruger's only alternative was to make an estimate of what the "true" validity coefficient was expected to be. This he placed at $r = ,7$. If one examines Kruger's reasoning in arriving at this figure, it can be concluded that it is a rather optimistic assessment.

As regards his conclusion that a correction is required to account for the pre-selection of his sample, he is partly justified. Those constituting the sample were not representative of the population for which the test is designed i.e. those who obtain the necessary score on the general screening battery, to make them eligible for testing on the L.G.T. The sample was composed of those who were successful on the latter test and therefore did not include all those cases who qualified for testing. However, he is not justified in assuming as he did, that correction should extend even further, so as to account for the elimination of testees in the first two stages of selection on the G.A.B.¹ For the purpose of considering the L.G.T. as a predictor, testees eliminated at these stages are of no interest. We are not concerned with determining the test's predictive capacity for this "total" population of mine intake, we are only interested in the degree to which it will predict success among the candidates for Boss Boy selection.

Kruger arrived at the correlation estimate of $r = ,7$ because when constructing the test the multiple correlation between the rating on the test and an arbitrarily weighted combined

1. General Adaptability Battery.

criterion was $r = ,72$. Because in the cross-validation he was now dealing with a criterion whose validity he had established, there is not sufficient cause to assume that a validity coefficient of similar magnitude would be obtained in a cross validation. (The reason for conducting a cross validation is after all to obtain independent verification of results). Firstly, the criterion differed in the two studies in that for the first it was combined with arbitrary weights and in the second, the weights were determined according to their meaningful relationship to an underlying success factor, as revealed in the validation of the criterion. Thus the very fact that the validity of the criterion had not yet been established in the first study, is an important reason why the validity coefficient from this study cannot be accepted as indicating finally the validity of the test as a predictor. Kruger recognized these facts and it was for these reasons that he conducted a cross validation, it is thus strange that he should put forward this argument.

There was a clear recognition in Kruger's work that indirect influences pertaining to the work situation itself played a significant role in influencing the validity coefficient. These seemed to indicate that the Boss Boy is not a precisely definable entity. The problems encountered in the development of criteria in the initial investigations would also point to this conclusion. The picture appears to be complicated by a number of factors. Not the least of these is the attitude of the white mine-worker. The Boss Boy occupies a role as the mediator between labourers and white supervisor, and his on-the-job performance is intricately bound up with the expectations and quality of management employed by his supervisor. It would probably be an advantage if job descriptions of the Boss Boy positions could be communicated right down the operational line. But even if the European mining officials underground did have access to job descriptions, there are bound to be individual differences in the attitudes of these supervisors to their Boss Boys. This was clearly apparent in the

early work done on the L.G.T. where low agreement was found when miners, shift bosses and mine captains were required to rate the on-the-job performance of their Boss Boys (2). Some may prefer a man who acts strictly according to orders without using any initiative of his own, others may like a Boss Boy to display more initiative in carrying out his duties, etc. In other words, style of leadership is not something which can be controlled through a job description. Traditionally-held beliefs as to the suitability of different tribal groups for these positions influences the attitude of the white supervisor to his Boss Boy. Both these factors have a notable indirect influence on the criterion. In addition to this, not all Boss Boys operate in the same physical conditions, do the same sort of work or have the same number of labourers to control. It would not be out of place to comment that consideration should be given to the total situation in which the incumbent will have to operate if optimal benefit is to be achieved by the introduction of selection tests, e.g. communication practices between labourers and miners, and even more specifically the training of Boss Boys, are two important variables in the work situation.

The problem concerning the predictive validity of this test had still not been altogether successfully resolved. At this stage, the greatest usefulness of the test still appeared to lie in the reduction of training wastage. However, Kruger's investigation had made a valuable contribution to the use of the L.G.T. in the Gold Mining Industry, particularly through the development of a new scoring method. It is therefore strange that this method did not replace the original L.G.T. procedure. The manual for the test still deals with the original scoring method requiring assessors to make a subjective assessment on a nine-point scale, using their observations of the appearance or lack of appearance of the eleven behavioural items. The reasons for this are not known. It is the writer's opinion, that a rating exercise of this nature requires not only intensive training, but also an understanding

of the psychological dimensions underlying the behaviour being observed. Ideally, for the technique to be properly used, assessors should be graduates in psychology. The practical situation on the mines precludes the possibility of employing graduates or even people with any formal training in psychology as assessors. The validity of subjective assessments on observed behaviour patterns made by men who do not have the background knowledge and skills acquired through formal training must be doubted. Given this situation therefore, everything should be done to create a scoring system where subjective judgments are excluded as far as possible and scoring is anchored to critical incidents of behaviour. This Kruger had to a large extent succeeded in doing.

6. REASONS FOR DISSATISFACTION WITH THE LEADERLESS GROUP TEST

The recent series of investigations (7, 9, 10, 11, 12, 19) was prompted by reports that the test had lost its effectiveness. Mauer (9) reports that the reasons for dissatisfaction as revealed by questionnaires sent out to all mining companies and interviews conducted with personnel and training officers of 33 mines not using the test or supplementing it with other predictors, were as follows:

- 1) The test is now too old and the solutions to the problems presented are well-known.
- 2) Selection errors occur.
- 3) The test tends to select bullies.
- 4) Once one leader has emerged, there is no opportunity for the other members of the group to show leadership.
- 5) Problems with test administration.

In general, this survey revealed that the purpose of the test has not clearly been understood by those who are using it. The failure to realize that the solution to the problem is not as important as the behaviour displayed in achieving a solution, underlines the need for careful training of test administrators. Without a proper understanding of the test as a whole, it is doubtful whether the ratings completed will be reliable.

In order to determine the current usefulness of the test and to restore some respectability to it, it had to be investigated from all aspects. Such an investigation could hope to provide a useful explanation for criticism (1) (2) and (6) above. In addition, the whole method of selecting Boss Boys in the industry had to be reviewed.

7. PHYSICAL APPLICATION OF THE LEADERLESS GROUP TEST

The first question to be asked was: How is the L.G.T. being used in the industry at present? This question could be answered in two ways:

- 1) How is the test actually being administered?
- 2) How are the test results being used?

7.1 The Administration of the Technique

It is clear that the reliability and effectiveness of the test depends on vigorous adherence to standard procedure. A survey of the application of the test in industry was called for before considering revising or changing the test itself.

In 1967, Van Heerden (19) reported on a survey conducted at 23 test centres. Attention was focussed on the tests, the European test administrators, Bantu test administrators and the test procedure followed. Results showed that deviations from standardized procedure laid down in the handbook were the rule rather than the exception. In some cases the tests themselves as well as the scoring procedures had been modified and changed. Both European and Bantu testers were often only vaguely informed about the procedure of test administration, and scoring was in many cases haphazard and unenthusiastic. The greatest shortcoming was in the training of test administrators. Both amount and quality of training varied considerably, as did educational level and experience of administrators. This factor was largely responsible for the inconsistencies which had appeared in the scoring of the test. It should be remembered that when the L.G.T. was in its experimental stage, and an interscorer reliability of $r = ,869 \pm ,082$ was reported, the assessors were well trained in the procedure and were probably more highly involved in the test situation and thus more highly motivated to make accurate assessments than many of the test administrators today, for whom it has become merely a routine task. Unfortunately, no procedure for training was laid down in the manual for the

original scoring system, so that no provision was made for standardization in training. Had Kruger's scoring system been adopted, a step by step training guide for test administrators was laid out, which would have been a valuable asset in maintaining a standard method of assessment. One of the reasons for the loss of effectiveness of the test had been clearly outlined by Van Heerden's report. Suggestions were made that reliability could be greatly restored through careful selection of the test administrators, and the development of a systematic training programme for both European and Bantu administrators so that some measure of standardization in scoring be ensured.

This survey indicated the need for careful selection of the test administrators themselves and pointed out an unfortunate lack of systematic training of the test users, both European and Bantu.

7.2 The use made of the results obtained

Mauer's survey (9) revealed that in most cases test scores were supplemented by other selection devices. It was apparent that no uniform policy on the various mines for the selection of the supervisory category existed and there was not a great deal of confidence in the L.G.T. as a means of selection. It appeared that the mining industry as a whole favoured a selection strategy based on a number of different variables, rather than relying solely on test scores. This approach holds some promise. It may well be that the general screening battery and the L.G.T. together do not adequately cover the facets of a Boss Boy's function. However, it is of obvious desirability that a selection strategy should be a standardized programme. Unfortunately, none of the variables being used had been standardized or validated, although many of them held promise. The only criterion of their usefulness was the subjective reports of the mining staff. The frequency with which certain predictors were being used was determined. A brief discussion of the most important ones follows:

Recommendations by Mining Staff were widely used with differing degrees of reliance and in different forms. It is a technique susceptible to the dangers of subjective assessment and the "halo effect". However, if used carefully, subject to regular checks, it could be valuable in identifying the "casualties" of a selection programme, through their on-the-job performance. Because in any test there is a range of probable error of measurement, it is conceivable that certain individuals with potential do not achieve the test scores required to enable them to undergo training. Thus performance in the work situation can be a valuable indication of an individual's level of competence. It is perhaps in this case that a panel interview should receive consideration. It would ensure that a general consensus as to the suitability of a particular candidate recommended, existed; this would prevent the selection of Boss Boys who can only function effectively under the white mining staff who recommend them. Each member of the panel should then be equipped with other background information such as experience, age, test scores and possibly merit ratings.

Experience is accepted as a necessary requirement by most companies, but the time period and the quality of experience specified vary considerably. This factor could be rationalised and organized through the use of promotional routes.

Promotional Routes This system would be of value in making the most effective use of the labour force, but as with "recommendations by mining staff", must be used with care and discretion in order that it have the desired effect. There cannot be any justification for using it alone without test scores as experience at different levels e.g. winch driver or loco driver, requires different skills altogether to the supervisory skills of a Boss Boy. If it is to be used as a supplementary means of recruiting Boss Boys, then a set amount of experience at each level considered necessary to acquire that particular skill would have to be specified.

Merit Ratings are fairly generally used for a number of differing purposes, and with varying degrees of reliability. To be effective as a technique, the attributes to be rated must be meaningful and clearly defined. Raters should be made aware not only of the purpose of the exercise, but also the seriousness of such a task - it was found that very often merit rating systems were used by raters more for their own end, rewarding favourites and not giving proper consideration to all candidates concerned.

Biographical and Physical Variables were often considered to be important contributory factors to Boss Boy success. It was the aim of a later report to attempt to determine the predictive validity of a number of these factors. (The results of these investigations are reported later in the present report).

While no one predictor emerged from this survey which could effectively replace the L.G.T., a number of worthwhile points had been made which deserved investigation. It was also clear that there is increasing awareness of the importance of the organizational climate within which the Boss Boy works. The formulation of a personnel policy specifying promotional routes and a system of performance appraisal would be a valuable aid to the effective use of manpower. These could provide another avenue for the recruitment of Boss Boys.

The next phase of the research project was the development of a new criterion and subsequently the restandardization of the L.G.T. and the investigation into the possibility of developing other predictors.

8. THE DEVELOPMENT OF A CRITERION

8.1 Job Analysis

It was considered that the drop in predictive capacity of the L.G.T. may be due largely to the fact that the Boss Boy job has not remained static. With the introduction of a system of selection and classification of labour, it is highly likely that changes in the structure of the labour force occurred, as well as probable alterations in the quality of management.

Mauer's development of the criterion (10, 11) was an extremely thorough and well-planned project. In order to get an accurate reflection of the job of a Boss Boy, very detailed job descriptions were prepared on the Stopping and Development Boss Boys. These are the most predominant category of Boss Boy in the industry. By this stage, the N.I.P.R. system of job analysis had developed into a comprehensive and systematic model. By this method the jobs are analyzed chiefly in terms of the complexity of the decision-making they demand of the incumbent, and takes into account also the controls and checks under which he operates, the educational level and experience required to perform the job. Careful consideration was paid to ensure that the descriptions were representative of all stopping and development jobs in the industry and were not just a function of the particular jobs described. Descriptions were carried out at two mines, one an old mine, the other a new one. Once completed, they were submitted to the mine management as a check. As very little difference was found between the two jobs, a composite job description was developed, and it was recommended that it be verified at a third mine to provide a further check. Comments on the descriptions by the different mining groups to whom they were circulated, revealed that there were no fundamental differences throughout the industry.

8.2 The Check List as a Criterion

Steyn (17) has pointed out that while many sophisticated predictive tests have been developed, the comparative sophisti-

cation of criterion-development is low. He suggests that the poor relationship between training criterion measures and job performance generally found, may be eliminated if a criterion incorporating the actual task elements is adopted. Mauer therefore decided to adopt Siegel's method of a checklist, based on the detailed analysis of the actual task of the Boss Boy (11). This would more closely satisfy the conditions indicated by Styn as pre-requisite for an effective criterion, namely practicability, relevance, reliability and comprehensiveness.

From the job descriptions, 62 task-anchored items were generated. The response to each item was to be given on a five point scale. The instrument was then given to the shift boss and miner in charge of each of 156 Boss Boys in the industry. A factor analysis of the ratings was performed and four factors with the largest number of high loadings were isolated out of the twelve factors which emerged. These four factors were interpreted as follows:

Subordinate Relations
 Production and Communication
 Accident Prevention
 Fundamental Physical Job Requirements.

An item analysis was then conducted on each of these four factors. The standard error of measurement and the reliability using Kuder-Richardson Formula 20 with Ferguson's extension were calculated. Results are shown in Table I.

Table I

	Subordinate Relations	Production & Communication	Accident Prevention	Physical Job Requirements
S.E.M.	2,72	2,99	2,32	2,46
r	0,95	0,93	0,89	0,83

(Mauer (11))

Reliability of the scales of Accident Prevention and Physical Job Requirements could be further increased if a few more items were generated as Mauer recommends. A performance criterion had thus been developed which showed reliability and the ability to discriminate between good and bad boss boys. This direct performance measure would appear to be the most satisfactory of the criterion measures developed to date.

With a newly-developed criterion measure, the next logical development was the assessment of existing and potential predictor variables. In 1970, two reports were published on the work related to this aspect (7, 12).

9. THE RESTANDARDIZATION OF THE LEADERLESS GROUP TEST

The Anglo-American Research Unit was responsible for restandardizing the L.G.T. (7). Essentially the restandardization produced

- (a) the construction of new practical problems.
- (b) the construction of a new rating scale for scoring the test.

As the Leaderless Group situation is based on the assumption that the leadership function which is to be predicted is of a supervisory nature, it had to be established whether this was in fact still the case. To this end, three questionnaires were constructed. Two of these served the purpose of gauging the opinions of white mining personnel on the importance of duties of a Boss Boy and the third to obtain the opinion of Bantu miners themselves regarding preferred Boss Boy characteristics. It must be noted that this study was conducted on Boss Boys in non-productive areas and not on the Stopping and Development Boss Boys used for Mauer's criterion study. This is perhaps unfortunate although in the Anglo American Corporation it was unavoidable as all initial placements of Boss Boys are in non-productive areas. Results of these questionnaires served two purposes:

- 1) To indicate that the Boss Boy job was still supervisory in nature.
- 2) Important factors in the performance of a Boss Boy job could be divided into two groups. One group related to actual technical knowledge, safety training and supervisory practice, which it was concluded could be dealt with by the mines training course. The other group involved the capacity for organizing, controlling and co-ordinating group activity for which the L.G.T. was felt to be a suitable instrument.

The new test problem constructed, the Rondavel Test, introduced

a number of deviations from the standard Leaderless Group problems. A time limit was imposed and an element of competition was introduced by having two test situations operating concurrently. It is uncertain what the reason for this was as the Boss Boy job is not one involving competition. The rationale underlying these innovations was not given in the report.

In deciding between the four different methods of rating performance which were considered, special consideration was given to the facility with which Bantu test administrators could master the scoring system - a valuable approach considering the difficulties of administration previously encountered. The Systematic Graphic Rating Scale was finally adopted. Four factors are incorporated in the scale; (1) Organizing, (2) Controlling, (3) Co-ordinating, and (4) Problem-solving. Each one of these aspects is rated on a five-point scale, scale points being five statements arranged in descending order e.g. For the factor "co-ordinating" scale points range from "Got the whole group to work together" at the top end to "Does work himself" at the lower end.

A follow-up validation study was conducted on a sample of 54 trained Boss Boys placed on the job for four to six weeks. Validation was done against two ("Subordinate Relations" and "Production and Communication") of the four factors in Mauer's criterion schedule. ("Accident Prevention" and "Physical Job Requirements" were not used as it is the custom on gold mines of the Anglo-American Corporation to make initial Boss Boy appointments in non-production areas, where these two factors do not apply). The validity coefficients obtained were: Subordinate Relations $r = ,313$; Production and Communications $r = ,205$. The coefficient obtained for the former is significant at the 5% level whilst the latter is not significant. These figures were obtained for the Systematic Graphic Rating Scale. Validity coefficients obtained for the other two experimental rating methods were not as high. On the basis

of the following figures, the investigators conclude that the test is discriminating between good and bad Boss Boys as determined by rating of their on-the-job performance. It is felt that these conclusions are somewhat hasty.

	Job Accepts	Job Rejects
Test Accepts	17%	13%
Test Rejects	30%	40%

As the size of the sample is known to be 54 cases these percentages can be converted into number of cases as follows:

	Job Accepts	Job Rejects
Test Accepts	9,18	7,02
Test Rejects	16,2	21,6

Particularly in the "Job Accepts" category, the discriminatory ability of the test appears somewhat doubtful. It would thus seem to be premature to conclude that the new method has successfully demonstrated that it can separate the 'successful' Boss Boys from the 'unsuccessful' ones.

The sample size for this study is felt to be rather small - a factor which must result in caution when drawing conclusions. It is not clear from the report what the basis was for selecting the 54 to make up the validation sample, out of the 90 in the original sample. The reasons why the other 36 were not included could conceivably have significant implications for interpreting the validation results. While Hugo recognized certain limitations in the sample used, it is not felt to be sufficient to indicate that rectifying these could be expected to improve the position. Rather, the test should remain in an experimental stage until a complete revalidation is conducted. Recommendations for such a study will be discussed at the end of this report.

The main positive contributions of the project were (a) its demonstration that the role of a Boss Boy is still a supervisory one, and (b) the development of the Systematic Graphic Rating Scale on which Bantu raters can be trained. This could fruitfully be used in the development of an actual training programme.

10.

THE ATTEMPT TO DEVELOP A WEIGHTED APPLICATION BLANK

The final step in the series of investigations was an attempt to give some scientific status to the predictors used by the industry to supplement test scores (12). The development of a Weighted Application Blank was considered to be the best way of identifying the predictive capacity of a number of biographical or physical variables thought to be relevant. In addition to 19 variables related to biographical or physical information (including those most commonly considered to be important by the industry) and other characteristics of the job, L.G.T. scores and scores on the General Adaptability Battery (G.A.B.) were included in the analysis. The L.G.T. ratings in this case had been obtained before the revision of the technique by the Anglo-American Research Unit.

The predictive validity of other sources of information was also investigated.

Field Dependence/Independence

This was a theory put forward by Witkin to explain how individuals differ with respect to the degree of support they require from their environmental surroundings, in order to function effectively. It was felt to be relevant in this context because the field independent person tends to be characterized by a higher self esteem and capacity for initiating and organizing activity than the field dependent person. The latter in contrast is unable to rely too heavily on his own judgment in making decisions, and does not have the same tendency to initiate and organize. The test used was the Rod and Frame test.

Degree of Urbanization

The term "urbanization" implies a certain familiarity with the objects and mores of a Western Technological Culture. Because the mines form part of this particular culture, it would seem feasible that the rural man who has had no contact with Western technology will find it more difficult to adapt to the

mining situation than the man who has had some exposure to its cultural objects. The Urban-Rural Scale (4), developed as part of the International Biological programme was used to measure this construct.

The Welsh Figure Preference Test

Stopping and Development Boss Boys are required to train their subordinates. The ability to acquire the skill of teaching others is therefore an important requisite of these men. The decision to investigate the predictive capacity of this test - a personality test not dependent on language ability - was taken because Mokoatle reported that the Scale of Conformance showed a positive correlation with a test constructed to measure the facility with which T.W.I. principles were learned. All variables were related to the four criterion dimensions and careful consideration was given to keeping the sample of the same composition as the validation sample. As the χ^2 (chi square) difference between low and high scorers was not significant, the development of a Weighted Application Blank was not warranted. It must also be noted that many of the variables most widely considered as contributing to Boss Boy effectiveness were not found to bear a significant relationship with the criterion scales. This applies to all of the following variables identified by Mauer (12) as the most commonly used supplementary biographical information. Age, educational level, fluency in one of the official languages or Fanakalo, seniority and ethnic affiliation. The fact that the qualities measured by the L.G.T. appear to have less bearing on effective Boss Boy performance as time passes, would seem to indicate that length of experience would be an important factor in determining Boss Boy performance. In this study, however, this variable was not found to relate significantly to the criterion.

Probably the most significant finding in this study was the lack of predictive capacity displayed by the L.G.T. scores. In view of the findings relating to the unstandardized

administration and scoring procedures followed, and the lack of adequate training of assessors, this is hardly surprising. While we can therefore conclude that under the conditions in which it is now in use, it has lost a great deal of effectiveness, we are not yet in the position to indicate whether the technique itself has lost or maintained its suitability for the selection of Boss Boys.

11. CONCLUSIONS

Knowledge concerning the functions of Bantu supervisory positions comes almost exclusively from research on the Gold Mines. The strategy for the selection of this category has now been investigated from almost all angles, but the picture emerging is unfortunately still not altogether clear. It is notable that in its present form (both the "orthodox" version and Hugo's revision), the L.G.T. is not demonstrating effective predictive validity. As an instrument, the L.G.T. has never been able to demonstrate very high long-term predictive validity. But when administered in a standardized form with a scoring method which is designed to eliminate as far as possible the influence of subjective judgment, it can be an effective means of selecting those who are likely to succeed in the training course. Thus we have not yet had a clear demonstration that the technique has no worth at all; it appears that the present position has arisen mainly as a result of poor administrative procedures. What is clear though, is that the technique is not a simple one - it places far greater demands on the ability of the test administrator than does the screening battery. In fact, in terms of the requirements laid down by the South African Psychological Association for registration of test users, it should most logically be considered as a C-level test. In terms of this classification of tests, a C-level test may only be made available to somebody who is registered as a psychologist with one of the recognized professional bodies in the country. Usually this means the possession of a Master's degree in psychology.

Therefore, when considering using this instrument for selection purposes, attention should first be paid to the conditions under which it will be administered, in particular to those who will be assessing test performance. In the Gold Mining situation for instance, where qualified personnel are not available to fill the positions of aptitude testing officers and Bantu test administrators, the possibility of the L.G.T.

fulfilling a truly effective function is doubtful. It is expected that while in certain organizations in secondary industry, such qualified personnel would be available, this is unlikely to be a general rule. The only way in which one can to a degree overcome this problem so as to give the test wider application, is to attempt to gear the scoring of the test to the level of the personnel who are going to have to use it. The graphic rating scale developed by Anglo American holds some promise as such a method.

It is quite clear that the selection of Bantu first line supervisors is one of the major problems facing South African Industry. There appears to be a growing trend to extend the responsibilities of the Bantu supervisor to include many of the tasks traditionally considered the sole responsibility of the European front line supervisor. The selection problem posed is a difficult one. The picture emerging from the investigations into the selection of Bantu supervisors reviewed above, is that their function is not a simple one; methods used to date to identify supervisor potential are not sufficient to account fully for all aspects of his function. Furthermore the demands of secondary industry are very likely different from those of the mining industry on which studies have been conducted. This refers not only to the actual task of the supervisor, but also to the conditions under which selection will take place viz. the number of testees to be tested at a time, the amount of time available for testing and the qualifications of the test administrators.

12. RECOMMENDATIONS

It appears that what is required is a thorough empirical investigation into leadership and supervisory potential amongst Bantu workers, so that the increasing demands of industry can be met through the use of a comprehensive task and relevant selection procedure. This would necessarily be a long-term investigation. However, this is not to say that attempts to improve upon and make the most effective use of the methods currently in use, should be abandoned. It would appear, therefore, that recommendations fall into two categories, "short-term" and "long-term".

12.1 "Short-term Recommendations"

(1) Replication of Hugo's study

A replication of Hugo's study it is felt, would be valuable, but because it is a validation of a procedure designed specifically for the Gold Mining context, it would necessarily have to be done in the mining industry.

(2) Investigation of Assigned Leadership Tasks

Thinking in terms of the actual test situation, one alternative to the Leaderless Group situation has been considered i.e. assessment in an assigned leadership situation. It may be that by placing an individual into a situation in which he is assigned the task of organizing a group of men in solving certain problems, a more accurate idea of his ability in this respect will be obtained than in a situation where he has to emerge amongst a group to assert his leadership. The supervisor in the job situation is placed in a position by virtue of which he carries certain authority, and is expected to carry out a number of duties which require organising skills. Thus it may be that the assigned Leadership Task represents a closer simulation of the true task of the supervisor. It must be noted that it is inevitable that the quality of performance of each individual member of the Leaderless Group, depends to a

large extent on the standard of performance of the group as a whole i.e. individuals of low ability may show up well in a poor group simply because they are the best of a poor group. In a good group, certain individuals may not have a chance to give a performance which is a true reflection of their ability, because one member may dominate the group. The assigned leadership task will to a large extent eliminate this. However, the problem of the method of scoring the test which was one of the greatest difficulties with the L.G.T. has not been avoided. Other problems are also added, e.g. the composition of the group to be led, and the nature of the tasks to be performed.

(3) Establishment of the process whereby supervisory potential can be identified

The process of selecting supervisors can be envisaged as taking the following form.

(a) Recruitment: As Mauer (12) showed, of all variables investigated, G.A.B. score bore one of the highest relationships to the criterion (as could be expected): score on the screening battery should be the major criterion for identifying possible candidates (as has been the case up to now). As all psychological tests, even the most reliable, have a certain margin of error, it must be expected that there will be a certain small proportion of individuals who will show promise in the work situation but who do not qualify in terms of aptitude test scores, to undergo training as supervisors. As in general the mining industry has a shortage of men who are eligible for training as supervisors, every effort needs to be made to utilize all the potential available. Some system of performance appraisal, whereby "casualties" of the selection procedure who have nevertheless proved themselves to have valuable potential on the job, is needed.

While the L.G.T. has shown consistently low long-term predictive validity, as a predictor of success on the training course, it has been satisfactory. This is in line with Bass' findings

that the validity of certain predictors tends to decrease over time. It is therefore felt, that should it be decided to maintain the test, the scores could most profitably be used to predict success on the training course, rather than final effective Boss Boy performance.

(b) Training

With the use of the L.G.T. to predict training success, training must necessarily become an integral part of the selection procedure. Thus the outcome of the training course in the form of an examination will become the predictor of on-the-job performance. The existence of a satisfactory measure of job performance in the form of Mauer's checklist, makes validation of the training course examination possible.

Thus, a considerable degree of importance must be attached to the training period. Training refers not only to equipping the man with technical skills he will be required to perform, but also equipping him with an understanding of the principles of supervision which are relevant to his particular function. Mokoatle (13) has indicated that T.W.I. principles can be learned by the Bantu supervisory level on the mines. In order that the trainees achieve an understanding of the principles being imparted, he suggests that the programmes be adapted to the specific context of the Boss Boy, incorporating problems which are particularly meaningful to the situation in which he will operate and are at a level which he will easily grasp. Furthermore, he emphasizes the need to consider the climate of management within which he will be required to exercise this function and apply the principles he has learned. It was Mokoatle's conclusion that the two T.W.I. programmes which have the most relevance for the training of Boss Boys, are "Job Instruction" and "Job Relations".

In secondary industry, this knowledge, gained from the mining industry, is particularly valuable. From the start, an attempt should be made to encourage the use of the training

period as part of the selection procedure.

These "short-term" suggestions relate mainly to techniques of assessment, and do not question fundamentals e.g. concerning "leadership personality". For this, a long-term investigation is needed.

12.2 "Long-Term Recommendations"

Leadership behaviour can be thought of as a broad range of skills facilitating a person's ability to assume a position of control in a group. Supervision, as it is referred to in the industrial context, must be seen as a more specific category within this broader frame of reference. It is the situation in which he exercises leadership skills which defines the category termed 'supervision'. Thus supervisory ability must be seen as a function of leadership. In the "classical" situation, the natural leader will emerge as controller of a group by virtue of his personality qualities in relation to the character of the group. The supervisor on the other hand is appointed to such a position and is responsible for the performance of a fairly specific set of tasks. For this, he requires the ability to control the co-operative action of the group, but combines this with his function of first-line representative of management. He thus forms an important communication link necessitating the possession of a specific set of skills.

Theory on leadership which has been used to explain behaviour patterns of supervisors in industry stems chiefly from empirical research conducted on Westernized groups. One should therefore be wary in taking these over to explain the behaviour of Bantu supervisors - a totally different cultural group. A number of questions need first to be answered, and it is envisaged that this can only be done by means of an empirical research project into leadership behaviour in the Bantu community.

The following are considered to be the questions which require closest consideration:

- 1) What are the requirements of Secondary Industry regarding the post of Bantu first-line supervisor? What are the demands of the job at present and what are his responsibilities in the future likely to be? The function of a Bantu supervisor needs to be investigated across industry to establish whether one can rightly speak of a "supervisor" whose function is common to all industry, or whether he is specific to the nature of the work performed.
- 2) What are the leadership skills which facilitate effective supervision?
- 3) In the identification of variables which can validly be used as the basis of selection, the following three questions need to be answered:
 - (a) Is the effective leader in the industrial context distinguishable by certain personality qualities and a characteristic need structure?
 - (b) Are training and experience on the job more important considerations than personality variables in determining ability to supervise?
 - (c) What will the effect of education be in determining leadership style?

Only when satisfactory answers can be given to these basic questions, can a strategy for the selection and training of the Bantu supervisor be drawn up.

BIBLIOGRAPHY

1. BASS, B.M. Further Evidence on the Dynamic Character of Criteria. Personnel Psych. 1962, 15 93-97.
2. BIESHEUVEL, S., HUDSON W. Validation of Boss Boy Selection Tests. N.I.P.R. 170 (C.S.I.R.) Project 16 Third Interim Report 1949.
3. BIESHEUVEL, S. Personnel Selection Tests for Africans. S.A. Journal of Science. vol. 49 (1) 1952.
4. GRANT, G.V. The organization of Mental Abilities of an African Ethnic group in cultural transition. Dissertation submitted for the degree of Ph.D., University of the Witwatersrand, 1969.
5. HUDSON, W. The Occupational Classification of Africans. Dissertation submitted for the degree of Ph.D., University of the Witwatersrand, 1953.
6. HUDSON, W. and KRUGER, C.F. The selection of African supervisors by Leaderless Group Tests. Inter-African Labour Institute Bulletin. 1958, 5, 10-27.
7. HUGO, P. Restandardization of the Leaderless Group Test for the selection of Bantu workers in the Gold Mining Industry. Report by the Personnel Research Unit, Anglo American Corporation of South Africa Ltd., submitted to Chamber of Mines of S.A. Nov. 1970.
8. KRUGER, C.F. Die Keuring van Voorjongens in die Goudmyne. Dissertation submitted for the degree of Ph.D., Stellenbosch University 1959.
9. MAUER, K.F. Current Boss Boy Selection Practices in the Mining Industry. Report submitted to the Chamber of Mines of S.A. Nat. Inst. for Pers. Res. (C.S.I.R.) C/Pers 166, 1970.

10. MAUER, K.F. Stopping and Development Boss Boy Jobs. An Interim Report. Report submitted to the Chamber of Mines of S.A. Nat. Inst. for Pers. Res. (C.S.I.R.) C/Pers 170, 1969.
11. MAUER, K.F. A criterion of Boss Boy Performance. Report submitted to the Chamber of Mines of S.A. Nat. Inst. for Pers. Res. (C.S.I.R.) C/Pers 177, 1970.
12. MAUER, K.F. An Attempt to Develop a Weighted Biographical Data Sheet for use in Boss Boy Selection. An Investigation into the Validity of Biographical and Certain Test Variables for Predicting Boss Boy Performance. Report submitted to the Chamber of Mines. Nat. Inst. for Pers. Res. (C.S.I.R.) C/Pers 187, 1970.
13. MOKDITLE, N.B. Effects of Training Bantu Supervisors using Training Within Industry (T.W.I.) Programmes. M.Sc. Thesis for the University of South Africa 1964.
14. S.A.C.S.I.R. Bulletin of the National Institute for Personnel Research. Monograph supplement No. 1. 1950. Aptitude Tests for Native Labour on the Witwatersrand Gold Mines. Validity of Test Procedures.
15. S.A.C.S.I.R. N.I.P.R. Project H61/173. Final Report on the experimental use of Boss Boy Selection Tests at Modderfontein B Gold Mining Company Limited.
16. S.A.C.S.I.R. N.I.P.R. Project H61/175. Final Report on the use of selection and classification tests for Native Mineworkers at East Daggafontein Mines Limited. 1952.
17. STEYN, D.W. The Criterion - Stagnation or Development? Psychologia Africana 1969, 12, 192-211.
18. THE O.S.S. ASSESSMENT STAFF. Assessment of Men. Rinehart Co. Inc. New York. 1948.

19. VAN HEERDEN, C.J. Selection of Boss Boys for the Gold Mines.
Report submitted to the Chamber of Mines
of S.A. Nat. Inst. for Pers. Res.
(C.S.I.R.) C/Pers 150, 1967.

