Cranfield

College of Aeronautics Report No 9108
April 1991

An investigation into the potential use of Psychometric Instruments for the selection of Firearms Officers

Ms.Susan Gleave, Dr.John Harris and Dr.Andrew Guppy

Applied Psychology Unit
College of Aeronautics
Cranfield Institute of Technology
Cranfield, Bedford MK43 0AL. England

Cranfield

College of Aeronautics Report No 9108
April 1991

An investigation into the potential use of Psychometric Instruments for the selection of Firearms Officers

Ms.Susan Gleave, Dr.John Harris and Dr.Andrew Guppy

Applied Psychology Unit
College of Aeronautics
Cranfield Institute of Technology
Cranfield, Bedford MK43 0AL. England

ISBN 1 871564 29 8

£8.00

LIST OF CONTENTS

LIST OF APPENDICES

BACKGROUND

During the past few years there has been a substantial increase in the number of incidents involving armed police officers. This trend is a cause of concern to police and public alike. Whenever such incidents occur they invariably attract a great deal of media attention which sometimes portrays the police action in an unfavourable light.

The firearms officer is expected to perform a distasteful, dangerous and highly stressful duty for the protection of society. These officers, in carrying out such duties, are also subjected to unreasonable pressures and criticisms by many sections of the public. Research has also demonstrated that the effect of a shooting incident on the officer involved can be traumatic and occasionally results in permanent psychological damage. For such a sensitive and demanding role, it is important that measures are adopted to select appropriate individuals.

There are a great array of individual differences between people, they vary in a great many physical and psychological characteristics. It is in an employer's interest to select the best person(s) for the job, from a usual wide array of applicants. Psychology has been able to assist in the selection process by offering techniques and methodologies to improve the selection of personnel. Current procedures for the selection of firearms officers have involved the use of psychometric techniques, although there has been little standardisation between forces on the methods employed. Partly as a consequence of this, there has been some concern about the value of such methods in this selection process. It was for these reasons that Bedfordshire Police Headquarters supported an investigation into the potential use of psychometric instruments for the selection of firearms officers.

The aim of this study was to identify a way of assisting the development of the psychological aspects of the current selection procedure for firearms applicants. It was decided that this would be best satisfied with the development of a job performance criterion measure that would reliably differentiate between officers.

INTRODUCTION

There are two options for developing suitable job performance criteria; objective measurements and subjective measurements. As objective measurements were unavailable, a subjective method of eliciting performance criteria was needed. In this instance, judgemental indices were employed defined in terms of personal characteristics and observed job-related behaviour. The two rating formats (trait and behavioural) were compared for reliability to decide which would provide the most valid criterion.

Before any improvements to the selection procedure can begin, a clear understanding of the job is essential. By establishing a concrete picture of the job, the selector gains an impression of the characteristics/abilities a successful candidate should possess. This is achieved by means of a job analysis.

Job Analysis is a means of eliciting job related information. This study employed the Critical Incident Technique (CIT, Flanagan, 1954) as its method of job analysis. The CIT utilizes a collection of anecdotes (i.e. incidents) of job-related behaviour that are critical to performance. These behaviours reflect particularly outstanding or poor performance of a job. These incidents are elicited from "subject matter expert" (SME), these are usually the job incumbents, and/or supervisors and trainers. These individuals are used to generate a great number of relevant incidents. This is done through individual and/or group interviews. The incidents are then edited and classified into job dimensions. These job dimensions represent a composite picture of job characteristics.

The CIT is used to evaluate observable behaviour that is seen to be critical to job success or failure, it involves the recording of behaviour that has been observed. The critical incidents provide a direct measure of what the individual actually does that leads to unsuccessfully or successfully accomplishing a specific task. One major advantage when employing this technique is that there are not a rigid set of rules which predispose the collection of data. There are however a flexible set of principles which can be modified so that the technique can be used in different situations.

There has been some debate as to the relative merits of the CIT. Empirical work has shown the information produced to be of a high quality (Levine et al 1980). The work of Andersson and Nilsson (1964) has shown CIT to be both reliable and valid with reference to its content validity (ie relevance and the reliability of the data collection). Schwatz and Baskett (1976), in their research of the taxonomies developed with the CIT, have shown the categories formed to be robust. Because of the delicacy of this research and the possible repercussions of employing a method of selection that is not wholly explicit, a method of job analysis was needed that has been proven to be both reliable and valid, it was for these reasons that the critical incident technique was employed.

METHOD

The study comprised of three phases:-
a) The collection of critical incidents and traits through interviews
b) The development of a behavioural questionnaire (BOS) and a trait questionnaire
c) The administration of behavioural and trait questionnaires to job experts in order to provide ratings of incumbents.
d) The analysis of results using cluster analysis and a number of techniques to test for reliability.

a) Interviews to identify incidents and traits
The first step of this phase was to identify the general aim of the job under investigation. This was achieved by means of an initial interview with job experts. The second step was to develop a standard protocol that was to be presented to the interviewees that would elicit relevant information.

Data were collected during individual in-depth interviews. The researcher prepared a briefing paper to be read at the beginning of the interview to explain what was required from the interviewees. The instructions were identical for each in-depth interview to ensure uniformity. The data were collected from personnel at two different locations: Bedfordshire police headquarters, and Luton airport police station. Ten interviews were conducted and interviewees comprised of ex-unit firearms officers, team leaders, trainers, Zulu officers and Containment officers. Data were therefore collected from a diversity of sources as recommended by previous research (Dissanayake Eltes & Harris, 1990). Before the interviews commenced, the researcher inquired as to whether the respondents objected to the interview being tape-recorded.

b) Development of Behavioural and Trait Questionnaires
Once the interview information was transcribed, the behaviours and traits were edited. The behaviours and traits were then subjectively clustered. Behavioural Incidents that described the same behaviour were grouped together and given a descriptive label.

Traits that essentially described the same characteristic were clustered together and the most descriptive trait was selected. Five point (Likert) scales were attached to the behavioral items and the traits. The anchor-points of the rating scales used to evaluate the behaviour or traits were gained from the job inventory literature.

The incidents were edited by three psychologists with experience in the field of job analysis. Latham et al (1979) suggested that the use of experts has no significant disadvantages over using job incumbents for the editing of incidents and has the benefit of maximising the subject pool.

The Behavioural Observation Scale (BOS, Ronan and Latham 1974; Latham, Wexley and Rand, 1975 Latham and Wexley 1977) and a trait scale were therefore developed from the incidents and traits collected from the initial interviews.

c) Job Experts Rating of Incumbents
The BOS and the Trait Scale were developed to identify a criterion that would assist in the selection of firearms officers. Both these rating scales were given to two trainers to rate current firearms officers.

The raters had to rate the same 19 firearms officers on both the BOS and the Trait rating scale. Both raters were asked to consider each behaviour and the range of settings in which its occurrence is possible, the setting may include training sessions and periods on duty. They were then asked to rate the frequency that each officer exhibits a behaviour.

This procedure was repeated for the traits, only they were asked to indicate the degree to which the ratee exhibits the characteristic and not the frequency.

d) Cluster and Reliability Analysis
When the completed questionnaires were returned, the responses were coded, and the data analyzed. The extraction of important clusters was achieved through cluster analysis. The reliability of the derived criterion measures was assessed using tests of reliability and inter-judge agreement.

RESULTS

The first step was to find homogeneous clusters of behaviours and traits. This was achieved using Cluster Analysis, which identifies relatively similar groups of cases based on selected attributes.

The Wards method of Cluster Analysis produced a five cluster solution for the behaviours and a nine cluster solution for the traits. For an explanation of the trait variables went into the nine trait cluster solution see Appendix A. Listed below is a brief description of each trait cluster, preceded by a descriptive label attached to each trait cluster.

CLUSTER NAME	DESCRIPTION OF CLUSTER
POLICE OFFICER	Represents the traits a police officer is expected to possess.
16PF	This is a collection of traits that represent dimensions on the 16PF.
CAUTIOUS	This is a collection of traits pertaining to a cautious nature, at the scene of a firearms incident.
PUBLIC	A collection of traits pertaining to dealing with the public at or outside of a firearms incident.
UNDESIRABLE	A collection of traits that are undesirable in a firearms officer that if displayed would result in ill feeling within the group.
STRESS	Is a collection of traits that could be exhibited when the individual is under stress.
INCIDENT	Represents traits that would have an effect on the result of an incident.
CONTAINMENT	Includes traits desired in an individual to be effective in the role of containment officer.
ZULU	This cluster represents traits desired in an individual to be effective in the role of Zulu officer.

For an explanation of the behaviour variables that went into the five behaviour cluster solution see Appendix B. Listed below is a brief description of each behaviour cluster, preceded by a descriptive label.

CLUSTER NAME	DESCRIPTION OF CLUSTER
GUN HANDLING	This cluster represents behaviours concerning the handling of firearms.
ZULU	Represents behaviours likely to be exhibited by Zulu officers at a firearms incident.
GOOD OFFICER	Includes behaviours likely to be exhibited by conscientious officers concerning any police duty.
PROFESSIONALISM	Cluster four represents behaviours that would not have an effect on a firearms incident, but makes the difference between a professional and non professional officers.
NATURAL	This represents behaviours that would be exhibited by a "natural" firearms officer, with an education in tactics.

In terms of assessing the reliability of these clusters, two approaches were adopted. Firstly, the internal consistency of the clusters was assessed using Cronbach's alpha statistical index (see Appendix C). Secondly, a measure of inter-judge reliability (ie the extent that the two raters independently provide similar ratings on given aspects of the same ratee's behaviours or traits) was obtained from ratings by the two trainers (see Appendix D).

If a scale is reliable then there should be high internal consistency and high inter-judge agreement. Although almost all clusters were found to be reliable in terms of internal consistency, there were only three clusters which also showed significant inter-judge agreement. These were two trait clusters and one behaviour cluster and they are described in more detail below.

Cluster five of the traits was found to be reliable. This consisted of traits undesirable in a firearms officer and likely to cause ill feeling within the other firearms officers. Examples of these traits include if the officer is boastful or complaining. Also if the officer is flippant then this could lead to the other officers feeling that s/he was not taking their job seriously, resulting in a lost in confidence in that person at a firearms incident.

Cluster seven of the traits was also reliable, and included traits that were seen to be relevant to performance at firearms incidents. Quick thinking, forgetful, easily confused, nervous and switched on can all be seen to effect the result of a firearms incident. For example, if the officer becomes easily confused and nervous then there is a high possibility that there would not be an easy resolution of the incident, in addition to this there they would run the risk of placing their and fellow officers lives in danger. Quick thinking, switched on, adaptable, capable officers can be relied upon to diffuse an armed situation with the minimum of hinderance. Also, officers that remain switched on can be relied upon and inspire confidence for the other officers. If an officer is adaptable and quick thinking, then, should an incident change from the plan, then he can be relied upon to take this in his stride and formulate an alternative plan of action to accommodate this new sequence of events.

Cluster four of the behaviours was also reliable and consisted of behaviours seen to refer to professionalism. For example, an officer that argues with a fellow officer at a non-armed incident is not behaving with a professional attitude to his work, as one role of a police officer is to set an example to the public. An officer that is flippant in firearms training sessions gives the impression that he is not taking his job seriously.

The 16PF cluster was omitted from the analysis, as the traits were not generated from the critical incident elicitation and are therefore not relevant to this study.

SUMMARY

In general, it must be noted that the study was of an exploratory nature and the interpretation of the results should be tempered accordingly. In particular, the small number of job incumbents restricted the study considerably and the reliability of the findings would be improved by extending the study to include other forces.

As previously stated, the study identified nine trait and five behaviour clusters. With the exception of two trait clusters, all the clusters formed from the behaviour and trait ratings had high internal consistency, that is to say each cluster had an acceptable Alpha value. For the behaviour clusters no comparisons could be made with rater two as a result of missing values (see Appendix C).

The inter-judge reliabilities of the two rating formats produced disappointing results. There appeared to be little evidence of both raters independently providing similar ratings, that is, there was little or no consistency of responses (see Appendix D).

There were only three clusters that yielded acceptable correlation coefficients, those were cluster five and cluster seven for the traits and cluster four of the behaviours. One possible explanation for this is that ineffective behaviours and undesirable traits were employed within the scales. Behaviours that are perceived to be potentially dangerous, for instance, are unlikely to be rated as occurring within officers currently in the firearms unit.

Therefore, it would be fair to say that any behaviours or undesirable traits elicited from the critical interviews could have been exhibited by officers no longer employed within the unit. The repercussions of this are that the ineffective items do not differentiate between existing officers. That is, no officer has ever displayed an ineffective behaviour or trait that the trainers or supervisors perceived to be potentially harmful and consequently all officers are in fact 'good officers' and not that there is no difference between the variables. Alternatively, the high inter-item correlations could be the result of rater bias, or halo effect, that is the rater rated the scales the same way oblivious to whom he was rating. Both these alternatives would explain the high Alpha values, as all the effective variables would be rated positively and the ineffective variables negatively, resulting in high internal consistency, and consequently poor inter-judge reliabilities.

The aim of the study was to enhance the psychometric aspects of selection by developing a reliable criterion measure. This could not be achieved at the macro level, as no one method (Behaviour vs Traits) was seen to be more reliable than the other. The aim may be satisfied in that the clusters of the behaviours and traits yielded three very respectable inter-judge reliabilities, therefore a combination of these three clusters may satisfy the aim and offer a reliable criterion on which to improve selection.

However this would not yield a comprehensive method of appraisal as important aspects of the job would be overlooked because they did not discriminate between individuals. For example, a firearms officer needs to be an accurate shot, it is therefore assumed that officers serving on the unit are in fact accurate shots. When the officers were rated on this item there would be a restriction of range, that is, there would be little or no spread in the ratings between the raters. This results in a phenomenon called 'truncation' where a correlation cannot be accurately calculated as there is insufficient difference between the rating of the raters. This could lead to an item being omitted from the criterion, as it would appear not to differentiate between individuals, even though this is still an important factor pertaining to firearms officers that needs to be addressed.

One interesting factor that arose as a result of the cluster formations is that there are two types of firearms officers, that is Containment officers and Zulu officer. Although there is a main core of activities that both officers share (eg shooting accuracy and remaining switched on), both these types need to posses different traits and, as a result of their activities within the unit, display different behaviours. For example, a Zulu officer would need to be proficient in the entry of a house, and adept with the use of a ballistic shield, whereas a containment officer needs to be vigilant. Concentrating on these different requirements for the selection of these types of officers has not yet been addressed. It is recommended that the differences in these officers be scientifically identified, and rating scales developed prior to further examination of the utility of psychometric procedures.

REFERENCES

ANASTASI, A. Psychological testing, 5th edition, 1982, Macmillan publishing Co.

CRONBACH, L.J. Essentials of Psychological Testing. 4th Edition, 1984, Harper and Row.

FLANAGAN, J.C. The critical incident technique. Psychological Bulletin, Vol. 51, No. 4, July 1954, pp.327-358.

HAYS, W.L. Statistics, 4th Edition, 1988, Holt, Rhinehart and Winston, Inc.

LATHAM, G.P.
FAY, C.
SARRI, L. The development of behavioral observation scales for appraising the performance of foremen. Personnel Psychology, Vol. 32, 1979, pp.299-311.

LATHAM, G.P.
WEXLEY, K.N. Behavioral observation scales for performance appraisal purposes. Personnel Psychology, Vol. 30, 1977, pp.225-268.

LATHAM, G.P.
WEXLEY, K.N.
RAND, T.M. The relevance of behavioral criteria developed from the critical incident technique. Canadian Journal of Behavioral Sciences, Vol. 7, No. 4, 1975, pp.349-358.

LEVINE, W.L.
ASH, R.A.
BENNETT, N. Exploratory comparative study of four job analysis methods. Journal of Applied Psychology, Vol. 65, 1980, pp.524-535.

NORUSIS, M.J. SPSS-X Advanced Statistics Guide, 2nd Edition, 1988, SPSS Inc.

RONAN,W.W.
LATHAM, G.P. The reliability and validity of the critical incident technique. Studies in Personnel Psychology, Vol. 6, 1974, pp.55-64.

APPENDIX A

Explanation of variables in the 9 trait clusters

TRAIT VARIABLES IN THE 9 CLUSTERS

CLUSTER NUMBER	VARIABLES
1	Suspicious, Risk Taking
2	Tender minded, Venturesome,Happy-go-lucky Outgoing, Eager, Experimenting, Imaginative, Reasoning, Self-sufficient, Talkative
3	Apprehensive, Reserved, Aggressive
4	Quick tempered, Noisy, Abrasive, Civil
5	Boastful, Complaining, Flippant, Boisterous
6	Tense, Immature, Controlled, Emotionally stable, Over reacting, Easily frustrated, Calm
7	Nervous, Quick thinking, Shrewd, Intelligent, Leadership potential, Forgetful, easily confused,capable, Adaptable, Coordinated, Switched on
8	Conforming, Smartly dressed, Safety Conscious, experienced,Tactically aware, Professional, Dependable,Conscientious, Dedicated
9	Positive attitude to killing, Streetwise, Assertive, Confident, Self reliant

APPENDIX B

Explanation of variables in the 5 behaviour clusters

BEHAVIOUR VARIABLES IN THE 5 CLUSTERS

CLUSTER	VARIABLES
1	Points weapon without due care. Intentionally fires at unthreatening target. Fires weapon by accident. Flinches at gun recoil. Freezes and needs prompting to initiate action. Maintains constant awareness of firearms status. Follows exact details of tactical/orthodox procedure when intelligence is accurate. Maintains awareness of lines of fire. Touches rig unnecessarily.
2	Withholds fire when real threat present. Is an accurate shot. Effective in forcing entry to property for search. Uses Ballistic shield effectively. Establishes good team work with partner.
3	Displays concern for safety of bystanders at incident. Delays action until planning is finalised. Displays firearms unnecessarily. Is abrasive dealing with bystanders during incident Uses excessive forces when arresting armed suspects Questions orders assigning routine police duties. Advertises firearm status when unnecessary. Remains discrete when carrying firearms Communicates clearly and accurately with other officers at incidents. Contributes sensibly to rapid development of tactical plan. Uses minimal necessary force when arresting armed suspects.
4	Argues with fellow officers at incident. Obeys orders of officers without question. Is abrasive after resolution of incident. Behaves flippantly during training sessions. Diplomatic when dealing with superior officers.

5 Remains switched on after length periods of inactivity.
Overreacts to minor irritants during classification shoots.
Becomes restless when awaiting incident to commence
Requests unnecessary detail at briefing.
Communicates excessively at incidents.
Makes assumptions on incomplete information.
Relinquishes appropriate cover prematurely.
Liaises effectively with unarmed officers at incidents concerning safety.
Provides effective fire cover for other officers.
Effective and systematic when conducting a house search.
Observant of fine detail relevant to planning.
Improvises appropriately when tactical plan/intelligence is inaccurate.
Maintains a safe distance from the public
Adopts correct strategic positioning.
Effective in choosing suitable cover at incidents.
Improvises to cover errors by colleagues
Communicates clearly with fellow officers, conveying intentions for immediate action.

APPENDIX C

Cronbach's Alpha for all clusters for both raters

CRONBACH'S ALPHA FOR ALL CLUSTERS FOR BOTH RATERS

	RATER 1	RATER 2	RATER 1	RATER 2
CLUSTER	TRAITS		BEHAVIOURS	
1	.7273 N=8	-	.8797 N=10	-
2	-	-	.9359 N=10	-
3	.3670 N=8	.1376 N=16	.9100 N=10	-
4	.6082 N=8	.9087 N=16	.8755 N=10	-
5	.7545 N=8	.8343 N=16	.8659 N=10	-
6	.7853 N=8	.9026 N=16		
7	.8942 N=8	.9019 N=16		
8	.9260 N=8	.8115 N=16		
9	-.1048 N-8	.7398 N=16		

APPENDIX D

Interjudge correlations and standard deviations
for trait and behaviour clusters

INTER-JUDGE CORRELATIONS AND STANDARD DEVIATIONS FOR BOTH TRAIT AND BEHAVIOUR CLUSTERS

CLUSTER	SD RATER1	SD RATER2	TRAIT 'r'	SD RATER1	SD RATER2	BEHAVIOUR 'r'
1	1.231 N=17	0.713 N=19	0.3900 N=17	3.562 N=14	2.629 N=19	0.1845 N=5
2	-	-	-	4.340 N=13	2.053 N=8	0.2094 N=14
3	2.062 N=19	1.677 N=19	0.0524 N=19	6.700 N=11	3.536 N=2	-
4	4.245 N=19	2.789 N=18	0.2557 N=18	3.113 N=19	3.394 N=18	0.5856 N=18
5	4.126 N=19	3.194 N=19	0.7328 N=19	8.428 N=13	-	-
6	4.004 N=19	3.687 N=19	0.2775 N=19			
7	7.535 N=17	5.242 N=19	0.6432 N=17			
8	6.317 N=18	3.894 N=19	0.2347 N=18			
9	2.761 N=18	1.887 N=16	0.4685 N=15			