AD-766 887

AN ANALYSIS OF THE EFFECT OF AIR CREW FACTORS ON B-52 BOMBING RESULTS

Kenneth H. Biehle

Air Force Institute of Technology
Wright-Patterson Air Force Base, Ohio

June 1973

AD 766887

D D C

SEP 28 [illegible]

Reproduced by
NATIONAL TECHNICAL
INFORMATION SERVICE
US Department of Commerce
Springfield, VA 22151

UNITED STATES AIR FORCE

AIR UNIVERSITY

AIR FORCE INSTITUTE OF TECHNOLOGY

Wright-Patterson Air Force Base, Ohio

58

AN ANALYSIS OF THE EFFECT OF AIR CREW FACTORS ON B-52 BOMBING RESULTS

THESIS

GSA/SM/73-4 Kenneth H. Biehle
Major USAF

Approved for public release; distribution unlimited

Unclassified
Security Classification

DOCUMENT CONTROL DATA - R & D

(Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified)

1. ORIGINATING ACTIVITY (Corporate author)
Air Force Institute of Technology (AFIT-EN)
Wright-Patterson AFB, Ohio 45433

2a. REPORT SECURITY CLASSIFICATION
Unclassified

2b. GROUP

3. REPORT TITLE
AN ANALYSIS OF THE EFFECT OF AIR CREW FACTORS ON B-52 BOMBING RESULTS

4. DESCRIPTIVE NOTES (Type of report and inclusive dates)
AFIT Thesis

5. AUTHOR(S) (First name, middle initial, last name)
Kenneth H. Biehle
Major USAF

6. REPORT DATE
June 1973

7a. TOTAL NO. OF PAGES
55

7b. NO. OF REFS
16

8a. CONTRACT OR GRANT NO.

b. PROJECT NO.

c.

d.

9a. ORIGINATOR'S REPORT NUMBER(S)
GSA/SM/73-4

9b. OTHER REPORT NO(S) (Any other numbers that may be assigned this report)

10. DISTRIBUTION STATEMENT
Approved for public release; distribution unlimited

11. SUPPLEMENTARY NOTES
Approved for public release; IAW AFR 190-17
JERRY C. HIX, Captain, USAF
Director of Information

12. SPONSORING MILITARY ACTIVITY

13. ABSTRACT

This study examines the effect of aircrew and equipment factors on bomb scores achieved by electronically simulated bomb releases in B-52 aircraft. A total of 260 low level, synchronous releases performed by one SAC squadron over a six-month period were analyzed. Multiple linear regression analysis was used to analyze the effect of 16 crew factors and two equipment factors on bomb scores. At a .05 level of significance, no single crew factor was found to consistently affect bomb scores. It was also found that certain of the aircraft have a greater effect (either positive or negative) on bomb scores than any of the crew factors which were studied.

DD FORM 1 NOV 65 1473

Unclassified
Security Classification

14 KEY WORDS	LINK A		LINK B		LINK C	
	ROLE	WT	ROLE	WT	ROLE	WT
Air Crew Effectiveness factors.						
Bomber air crew effectiveness.						
B -52 bombing process; aircrew inputs.						

AN ANALYSIS OF THE EFFECT OF AIR CREW FACTORS ON B-52 BOMBING RESULTS

THESIS

Presented to the Faculty of the School of Engineering of
the Air Force Institute of Technology
Air University
in Partial Fulfillment of the
Requirements for the Degree of
Master of Science

by

Kenneth H. Biehle, B.S.
Major USAF
Graduate Systems Analysis

June 1973

Approved for public release; distribution unlimited.

ic

Preface

This paper attempts to determine whether the bomb scores achieved by B-52 crews on routine training missions are primarily determined by the crew, and if so, what factors of crew experience and background are most important in the bombing process. The study is believed to be timely in that there has been a large turnover of flying personnel in the Strategic Air Command (SAC) in recent years, and SAC commanders at all levels are faced with the problem of trying to maintain SAC's traditional high level of performance with a crew force that is relatively inexperienced.

I wish to express my gratitude to my advisor, Lieutenant Colonel Charles J. Doryland for the assistance he gave me throughout the formulation and preparation of this thesis. His insistence that I determine early my objectives, scope, and methodology resulted in a framework to work in, without which I would have had great difficulty turning the literally thousands of numbers in the data into a readable paper in the time available.

I also wish to thank Brigadier General Thomas F. Rew, and Brigadier General Thomas P. Conlin, former and present commander of the 17th Bomb Wing respectively, for their support, the many personnel of the 17th Bomb Wing who assisted me in gathering data, and Major Jon Hobbs of the AFIT Systems Management Department for his help with the mathematical formulation.

Table of Contents

Abstract

This study examines the effect of aircrew and equipment factors on bomb scores achieved by electronically simulated bomb releases in B-52 aircraft. A total of 260 low level, synchronous releases performed by one SAC squadron over a six-month period were analyzed. Multiple linear regression analysis was used to analyze the effect of 16 crew factors and two equipment factors on bomb scores. At a .05 level of significance, no single crew factor was found to consistently affect bomb scores. It was also found that certain of the aircraft have a greater effect (either positive or negative) on bomb scores than any of the crew factors which were studied.

I. Introduction

Systems analysts are trained, as the name implies, to analyze systems. The type and size of system can vary considerably, but it will include three parts or stages: (1) resources (or "inputs") which are (2) going through some process to (3) produce an output. The systems analyst's job is to determine, if possible, the optimum way to use the available resources to produce the maximum output within the given constraints. The classic examples have generally come from the field of weapons development and acquisition, where dozens of firms were involved, and millions of dollars were at stake. (Ref 4:243)

This study attempts to analyze a system which consists of: (1) B-52 aircrew members who are (2) planning and flying training missions to (3) produce bombing results. The aircrew members consist of aircraft commanders, radar navigators, and navigators of the 34th Bombardment Squadron, 17th Bombardment Wing, Wright-Patterson Air Force Base, Ohio. The training missions are routine missions of the type flown in the U.S. by SAC crews. The bombing results are electronically simulated bomb releases, with the circular error of probability (CEP) of impact computed in azimuth and distance from the target point by a ground station. (Azimuth is not considered in this study.) The result achieved on a bomb release is commonly called a "bomb score," and that term will be used hereafter in this paper. Bomb scores, then, are the output measure for the process under study.

It may be argued that from a B-52 squadron or wing commander's point of view, bomb scores are not a proper measure of crew productivity. Certainly there are others, such as navigation, electronic countermeasures, and knowledge of a crew's assigned Emergency War Order (EWO) mission. But bomb scores are a valid output measure for bomber crews for two reasons. The first and foremost reason is that the ultimate SAC mission is to drop bombs on enemy targets, and it follows that it will do little good to do everything perfectly all the way to the target area if the crew then can not hit the target. The second reason, and one which may be of more immediate interest to commanders, is that the success or failure of a SAC bombardment wing on its semi-annual operational readiness inspections (ORI) is determined almost entirely by its bombing performance. However, whether or not bomb scores are a realistic measure of output may be a moot point. The more important point is that if bombing results can be improved to some extent by the way available crew members are formed into crews, then it is of interest to commanders.

Research Objective

This study attempts to determine what effect, if any, various factors of aircrew member background and experience have on bombing results. Moreover, it attempts to determine if, given a squadron in which there is a wide range of experience, there are ways to form crews which will result in the best overall bomb scores. Thus, the research objective can be stated as follows: To determine if the bombing results of a B-52 squadron, with current manning, might be improved by proper selection of aircrew combinations. Note that there is no attempt to determine a mathematical "model" for predicting bomb scores for a given crew or all the crews.

A Previous Study

At this point it is well to point out that a similar study was conducted in 1967 by Colonel Herman L. Gilster, then of the United States Air Force Academy. (Ref 16) That study differs from this one in two important respects. The stated objective of Colonel Gilster's study was to determine whether there is a significant relationship between a crew member's background and how well he performs at bombing on a given ORI. This research also sought to determine relationships between crew member backgrounds and bombing performance. But the overriding objective is to then use these findings to determine whether combinations of crew members, rather than individual capability, may be of more interest to commanders.

But the really major difference in the two studies is in the crew members. Since 1967, throughout SAC, there has been a large turnover of personnel in the three crew positions which are used for input variables in this study. A comparison of the experience of the aircraft commanders in the two studies serves to illustrate the point. On the 387 crews which were studied in 1967, the aircraft commanders averaged 3,946 total flying hours, 1,586 B-52 hours, and 12.1 years commissioned service; in this study the figures are 1,573 total flying hours, 818 B-52 hours, and 4.8 years commissioned service. Similar comparisons could be shown for the other two positions, although the differences are most pronounced in the aircraft commander position.

The results of Colonel Gilster's study, as presented in the abstract accompanying the study, are given below. This thesis does not attempt to either confirm or refute those findings.

This report analyzes five qualitative and six quantitative input variables to determine their effect on SAC B-52 combat crew output as measured by simulated wartime exercise bombing accuracy. Of the qualitative variables, if we disregard the inherent differences in accuracy caused by equipment variations in different models of the B-52, we find that crew output can be predicted most reliably by crew designation. Select crews deliver the most accurate bombs and Ready crews the least accurate bombs. Of the quantitative variables, the aircraft commander's experience level, expressed either by total flying hours or total commissioned service, plus his educational level provide the most reliable predicators of crew output. A crew's bombing accuracy improves as the aircraft commander's experience level increases to approximately 4,500 flying hours or 11 years commissioned service. Beyond this point, accuracy begins to decrease at an increasing rate until at approximately 6,700 hours/16 years, it equals that of a new crew. Also, accuracy improves as the aircraft commander's educational level rises from high school graduate to the master's degree, but the incremental increase is not as great at higher levels of education as it is at lower levels. Finally, the report discusses possible changes to personnel policies as a result of these findings and recommends further studies of crew output in fighter and transport aircraft units.

Qualifications

A qualification must be put on the results as presented in this paper. Actual bomb scores are classified information and of course cannot be revealed in an unclassified paper. Therefore, when numerical results are presented in the paper, only relationships are shown and not absolute figures.

This study, and the system, are both obviously limited in scope. The system is only a part of a larger system. The latter statement can, of course, be made about any system, but it is more immediately evident in this case because directly involved in a bomb run also are: (1) the other members of the aircrew, (2) the aircraft and its integral equipment, and (3) the ground station which scores the simulated bomb release. In succeeding chapters, the impact of, and assumptions regarding these other factors are treated.

II. The Variables in the System

There are essentially three ways for a B-52 to drop bombs: from high altitude, at low level, and from a stand-off position using airborne missiles. The tactic most likely to be used in a nuclear engagement, and therefore the type most often practiced on training missions, is the low level drop (or "release" as it is more commonly called.) Consequently, low altitude releases are the only type considered in this study.

There are also two types of low altitude releases: (1) synchronous releases, using radar bombing techniques, and (2) alternate releases, which simulate failure of some component of the radar bombing system. Since the synchronous releases involve a greater degree of crew coordination, and because the crew has greater control over the accuracy of a synchronous release, only synchronous releases are considered in this study. In short then, only low level, synchronous releases are considered. The bomb scores on those releases are the dependent variable in the system model.

The number of synchronous releases per sortie on the missions studied in this paper varied from one to six, with the average being approximately three. Releases on training missions may be made in succession on several targets during one "run," or in a series of passes over a target. Based on information gained in interviews with crew members, in this study all releases on a mission are considered as separate (independent) events. Each release, then, is considered as a data point, or observation.

Other Crew Members

The aircraft commander, radar navigator, and navigator are involved in a closely coordinated team effort during a synchronous bombing run. The functions of the three other crew members (copilot, electronic warfare officer, and defensive systems operator) during a bomb run are considered, for purposes of this study, to have a negligible effect on the resulting bomb scores. It is true that any one, or all three, of these other crew members may play some role in the bomb run. But the degree to which they will enter into the bomb run, either in planning or performing it, is generally determined by the aircraft commander or radar navigator. Since this is so, the contribution of these other three positions during the bomb run should essentially be reflected in the experience of the aircraft commander or radar navigator. The experience of each of the latter two is included among the variables investigated for effect on bomb scores.

Aircraft

Inevitably there are differences in the bombing performance of different aircraft, or of the same aircraft on different days. The accuracy of the radar, which is highly dependent on calibration and "stability" in the aircraft's electrical system, is the most obvious of the differences which can exist. There are those who believe that differences in aircraft are more responsible for variations in bomb scores than any other factor. If that is so, then it is possible that this one factor could dominate all others to the extent that it may be impossible to determine other significant variables by analytical means. That possibility was considered in this study and is addressed in later chapters.

RBS Sites

The bomb scores which are recorded on training missions are computed by radar bomb scoring (RBS) sites which track the aircraft on radar during its bomb run. This introduces another variable into the system in that several sites are utilized, and again there may be variations in the performance of the equipment at different sites, or at the same site at different times. This possibility is taken into account in the study and is again addressed in later chapters.

Integral Crew Experience

Since a bomb run obviously involves close coordination between the aircraft commander, radar navigator, and navigator, it is of course conceivable that experience together as a crew would be the most significant factor in determining the bomb scores which a crew achieves. The procedures to be used by a crew are spelled out in detail in various directives, and checklists are used for all phases of a mission, but different individuals will still have their own techniques. If the other crew members are familiar with these techniques, the crew effort, it seems, should become better coordinated and more productive. Whether experience together as a crew contributes more significantly to a crew's bomb scores than individual factors, therefore, had to be investigated in the study.

Individual Factors

The variables of primary interest in this study are, of course, the individual crew members, or more specifically, factors of experience, capability and background which the crew members possess. With regards to experience, the following are considered for each crew member studied:

(1) total flying time, (2) B-52 flying time, (3) years of rated service, and (4) length of time in present crew position. The factors considered in the area of capability are: (1) standing (by percentile) in undergraduate pilot training or navigator training class upon completion of training, (2) performance on B-52 combat crew training school (CCTS) standardization evaluation (stanboard) mission upon completion of CCTS, (3) performance on 17th Bomb Wing initial qualification stanboard mission in present crew position, and (4) performance on three most recent routine stanboard missions when (2) and (3) above were not available.

Also considered is one factor which might best be called a motivation factor, that being whether an individual had been flying another type of aircraft operationally prior to his present assignment in B-52 aircraft. It is conceivable that young pilots being transferred to bombers might be somewhat disgruntled, and consequently might not be performing to the best of their ability.

Lastly, one factor which might be listed under either "capability" or "motivation" was considered, that being an individual's college (undergraduate) grade point average. No consideration was given to either the field of study or the school attended.

Assumptions Regarding Other Variables

One can readily think of numerous other variables which might affect bomb scores. Some which come to mind immediately are such things as weather in the target area, day or night mission, and mission length. In performing the analysis, factors of that type are assumed to balance out over the period of time and the number of sample points (releases) studied.

There are also intangibles such as how well the crew members get along with each other, and whether the crew as a whole is motivated toward the mission. Factors of that nature quite possibly affect crew and individual performance more than any other. But they are difficult, at best, to measure and quantify; moreover, during the time period which this study covers, there were many crew changes. As a result of those crew changes, on only 19 percent of the releases studied in this paper, had the pilot, radar navigator and navigator flown this type of mission together as a crew more than four times previously. Furthermore, with the exception of the crews which spent a majority of the time period on temporary duty in Southeast Asia, not one crew remained together for the entire time period covered by the study. For these reasons then, factors of crew personality, morale, and motivation were not considered in this study.

The Data Base

The period of time covered by this study is 1 July 1972 through 31 December 1972. The majority of the low level synchronous releases performed by 34th Bomb Squadron combat ready integral crews, and scored by a RBS site, during that six-month period, are included as data points in the study. Only missions which were performed by line, combat ready crews were included. This excluded: (1) missions flown by stanboard crews, (2) missions where members of the squadron or wing staff were performing duties in one of the three primary crew positions, and (3) missions on which instructors were conducting upgrade training in one or more of the three primary crew positions. In short, the missions included in the study were those which were performed by the type of crews and crew members which a commander could normally expect to have assigned to his B-52 squadron.

Data on crew members' flying experience was obtained from individual flight records available in the 17th Bomb Wing during January 1973. As a result, missions on which one of the three primary crew positions had been occupied by someone who then retired, transferred or separated prior to 1 January 1973 could not be included in the study. Crew members who became qualified after 1 July 1972 were included in the study.

The net result of the inclusions and exclusions of the above paragraphs was a total of 260 releases on 89 missions, performed by 23 different aircraft commanders, 22 radar navigators, and 24 navigators, comprising 49 different crew combinations. "Unreliable" releases which were a result of crew error are included in the total. ORI missions were not included.

Stanboard Data Determination

The factors of crew member experience, capability, and background as outlined in Chapter II are self-explanatory with the exception of performance on stanboard missions. Data for stanboard flight inputs was taken from SAC Form 817 on which the stanboard evaluator records his evaluation of the individual's performance. All phases of the mission, from mission planning, through the flight, to forms completion, are graded in detail. A grade is assigned for each phase of the mission, and each item within that phase, as follows: "H" for "highly qualified," "Q" for "qualified," "C" for "conditionally qualified," and "U" for "unqualified." For this study, an ordinal scale of 3 for H, 2 for Q, 1 for C, and 0 for U, was arbitrarily applied. The average of all grades on the form was then used as the input.

Since crew member evaluation data is not required to be kept indefinitely in individual flight records, data on CCTS and 17th Bomb Wing initial qualification stanboard missions was not available on all individuals. The absence of this information, however, is not considered a serious loss of data. The reason the data was missing invariably was because it had been several years since the individual had taken the stanboard flight. Since standards, procedures, and individuals inevitably change with time, the value of the data would be questionable at best. (In performing the analysis, a stanboard input factor was applied for each crew member from either a combination of the CCTS and the initial qualification stanboards, or recent annual proficiency stanboards.)

RBS Site Identification

Of the 89 missions included in this study, all but ten were flown against one of three primary RBS sites used by the 17th Bomb Wing during that period. Those three sites will be identified only as A, B, and C in the paper. The total missions and releases flown against each site are as follows: site A--26 missions, 79 releases; site B--29 missions, 73 releases; site C--24 missions, 82 releases. As can be seen, 24 releases, or less than ten percent of the total, were at sites other than A, B, and C.

Integral Crew Experience Factors

To obtain inputs for crew experience together as a crew, three crew member combinations and three experience criteria were used. The crew member combinations are: (1) aircraft commander, radar navigator and navigator together, (2) aircraft commander and radar navigator together, and (3) radar navigator and navigator together. The crew experience criteria are: (1) first or second low level mission together, (2) third or fourth low level mission together, and (3) fifth or greater low level mission together. (Two missions together are roughly equivalent to one month together. Obviously, to get this experience data, crew composition on missions flown prior to July had to be looked at.)

Pilot Training, Navigator Training and GPA

Data on individuals' standing in undergraduate pilot training or navigation training, were obtained by questionnairre. Data on college grade point average (GPA) was obtained from the Air Force Institute of Technology admissions branch. Again, data on these two items was not available on all individuals.

The Independent Variables

The basic independent variables which were included in the analysis are the following:

1. Aircraft Commander (AC) total flying time

2. AC B-52 flying time

3. AC years of rated service

4. AC composite stanboard input

5. AC previous assignment in other operational aircraft.

6. AC years (or fractions thereof) as B-52 AC

7 through 12. For radar navigators (RN), corresponding to 1 through 6 for aircraft commanders.

13 through 16. For navigators (N), corresponding to 1 through 6 for aircraft commanders with the exception of previous assignment and years as a B-52 navigator. The sample was too small on previous assignment, and years as a B-52 navigator is highly correlated with years of rated service.

17. RBS Site A

18. RBS Site B

19. RBS Site C

20. RBS Site other than A, B, or C

21. Combination of AC-RN-N flying either their first or second low level mission together as a crew.

22. AC-RN first or second mission together

23. RN-N first or second mission together
24. AC-RN-N third or fourth mission together
25. AC-RN third or fourth mission together
26. RN-N third or fourth mission together
27. AC-RN-N more than four missions together
28. AC-RN more than four missions together
29. RN-N more than four missions together

The variables listed above were available and included in the analysis for all 260 data points. A variety of data combinations had to be chosen to include those variables for which data was not available on all individuals. Those variables were: CCTS stanboard results, 17th Bomb Wing initial qualification stanboard results, standing in undergraduate pilot training or navigator training, and college GPA.

The next chapter outlines the method used to analyze the data.

III. Methodology

The method used in this study to try to determine significant crew inputs in the low level synchronous bombing process is multiple linear regression. In simplest terms, this means that the process of releasing a bomb is modeled by using an equation of the form,

$$Y = B_0 + B_1X_1 + B_2X_2 + . . . + B_kX_k + e$$

where Y represents the bomb score (the dependent variable) which results from combining the aircraft commander's B-52 flying time, the radar navigator's B-52 flying time, the navigator's total flying time, etc. etc., which are represented by the X's (the independent variables). The B's (the coefficients) indicate the contribution of each independent variable to the score, and they can be either positive or negative, depending on whether the associated variable has the effect of making the scores larger or smaller. The e is a random error term which represents the variability of Y due to errors in making the observations.

If all of the 260 bomb scores being studied, and the variables which went into producing each score, are put together as one system of linear equations, then by multiple linear regression (the method of least squares) the coefficients can be determined. One then has an equation which might be used for predicting a bomb score within some range with a known level of confidence.

Several features of multiple linear regression formulation made it desirable for this study, even though a model for predicting bomb scores was not an objective. The first is that it allows the separate influences of the independent variables on the bomb scores to be observed and, more importantly, simple statistical tests will show whether these influences are actually significant or whether they might best be disregarded. Also, the regression method, through the use of qualitative variables assigned a value of 1 or 0, allows variables to be included which are not otherwise quantifiable. An example of the use of the latter in this study is the RBS sites. If a release was made at site A, the value 1 was given to that variable for that release, and the value 0 was given to sites B and C. In the resulting regression equation, the value of the coefficient of the variable corresponding to site A would indicate the number of units of distance to be added to, or subtracted from, the score if a release were made at site A, relative to those sites not then included in the equation. Statistical tests will also show whether the coefficient of a qualitative variable is significant in the equation.

Assumptions

In order for the use of multiple linear regression analysis to be valid, several assumptions have to be made. The first of these is the assumption of linearity. This assumption means that, for the range of values being studied, the relationship between Y and the X's is a straight-line relationship, the X's being either a recorded value or some transformation of that value. For example, if the X's in the equation were aircraft commander B-52 flying time, and navigator years as a B-52 navigator, it would be assumed that the relationship between bomb scores and aircraft commander B-52 flying time is a linear one, and that the relationship between bomb scores and navigator years as a B-52 navigator is also a linear one. If there were some reason to believe that bomb scores are directly related to, say, the logarithm of the aircraft commander's B-52 flying time, then it might be included as a variable and the assumption of linearity would apply to it.

The other assumptions apply to the error terms which exist in the observations. The error terms are assumed to be normally distributed with a mean of zero and a constant variance, and to be mutually independent in the statistical sense. An example of what this means in this study is that of the RBS sites' scoring of the releases. The recorded score may vary from what it actually was, due to errors in observation. But the errors are assumed to have a normal (bell-shaped) distribution with a mean of zero. The constant variance assumption in this case means that the variance of this error is assumed to be the same for a fixed setting of the X's in the equation. The errors are mutually independent if the error in measuring one release does not contribute to the error in measuring another.

Although there can never be complete assurance that these assumptions will hold, familiarity with the process being studied will lead one to believe that they are reasonable. The assumptions were accepted for this study.

Analytical Method

The regression analysis was performed on a GE 600 computer, utilizing the Biomedical Computer Program (BMD) stepwise regression routine, one of the many routines in the BMD package developed by the University of California at Los Angeles. A brief description of what the routine does is presented in Appendix A.

Several features of the BMD stepwise regression routine made it attractive for this analysis. As with most routines, the mean and standard deviation of each variable, and a correlation matrix of all variables, is displayed. At each step, as a variable is entered into or removed from the regression equation, the routine prints out (along with other statistics) the following: (1) an R^2 value, (2) the analysis of variance F-statistic, and (3) the coefficients of the variables in the equation at that step and the standard error of those coefficients, from which a t-statistic can be computed.

The table of means and standard deviations of the variables was particularly useful for determining the relative experience of the crew members in the various groupings of data which were used. The correlation matrix is useful for determining when multi-collinearity--correlation between values of two or more variables, which can lead to invalid statistical tests on the B's--exists. Also, the correlation between the dependent variable and the individual independent variables provides some indication of which independent variables might prove to be significant.

The R^2 statistic provides a general indication of how well the regression line "fits" the data. In other words, R^2 is a measure of the percentage of variance of the dependent variable about its mean which can be explained by the independent variables in the equation. By observing the magnitude of the increase in R^2 at each step of the routine, one can determine whether the fit of the line is improving at a continuous rate, and hence gain some idea of whether the entering variables are meaningful.

The BMD stepwise regression routine has provisions for specifying the minimum level of significance that a variable must have in order to enter the regression equation, and the level each variable must maintain in order to remain in the equation. In order to allow a large number of variables to enter the equation so that the contribution of each could be observed, the level of significance for both entry and removal was purposely kept low (approximately the 50 percent confidence level) in this study.

The following decision rule was established to determine when a variable was significant in the statistical sense: If at each step in the routine at which the F-ratio for the equation at that step was large enough that one could reject the null hypothesis that all coefficients in the equation were zero, the t-ratio for a coefficient was large enough that one could reject the null hypothesis that the coefficient was equal to zero, then the variable (factor) associated with that coefficient was to be considered significant. All tests were to be made at the 95 percent significance level; i.e., the probability of rejecting the null hypothesis when it is true is .05.[1] The variable might have either a positive or negative effect on the score, depending on whether the coefficient in the equation was positive or negative.

A general method of analysis was established which could utilize the statistics described above and the writer's empirical knowledge of the data base. First a computer "run" was to be made using all available data points and variables. Then the data would be divided into groups, each of which had a common factor which corresponded to a like factor in the other groups, and each of these groups would be analyzed. For example, the data might be divided (or "blocked") into three groups according to whether the score came from RBS site A, B, or C, and a run would be made on each group. The results of these runs would be compared to see what similarities existed. The correlation matrix, and the table of means and standard deviations would be examined for possible causal relations which may have been responsible for dissimilar results. The data blocks might then be sub-divided again, provided the resulting number of data points would be at least 30. At each step also, the correlation matrix would be examined

---

[1] This is often referred to as a .05 level of significance, but .95 is thought to have greater appeal for the average reader and thus will be used throughout this paper.

for closely correlated independent variables, each of which had approximately the same correlation with the dependent variable. For example, navigator total flying time and navigator B-52 flying time could be expected to be closely correlated, and to each have approximately the same correlation with bomb scores. When possible, one of the variables would be dropped in the following run.

The objective of the method just described was to identify crew member factors which would consistently prove to be significant under like conditions. If either individual or crew factors, which would enable commanders to put together aircrew combinations which would improve their squadrons' overall bombing results, could be thus conclusively identified, the research objective would be realized.

IV. Results

This chapter presents the results of the analysis of the data. Starting with the first computer run which was made, each "run" or series of runs is described, and the significant variables (at the .95 level as described in Chapter III) which were identified in each run are listed. The R^2 value of the equation containing only the significant variables is also shown. The rationale for performing each successive run or series of runs is then explained.

When significant variables are listed, the following notation is used: a variable described as "positive" is one which improves the bomb score; i.e., reduces the CEP, while a "negative" factor is one which increases the CEP.

Comment on conclusions in this chapter is limited to that which is necessary to explain the rationale behind performing the various analysis.

Initial Results

The first run included all 260 data points. The 29 variables detailed in Chapter II were used. The only significant variable which was identified was, "Aircraft Commander previous assignment in other operational aircraft," a positive factor. The R^2 value for the equation containing only the constant term and the one significant variable was .024, a totally unacceptable result

Aircraft Included As Qualitative Variables

Since none of the RBS sites had emerged as a significant variable in the initial run, it was decided to check the other factor external to the crew in the bombing process, the aircraft, to find if they were perhaps dominating the analysis. A run was made which included qualitative variables for the 14 aircraft. The results were:

Aircraft 008 -- Positive

Aircraft 009 -- Positive

Aircraft 023 -- Negative

AC previous operational assignment -- Positive

$R^2 = .14$

Observation of the scores of the individual aircraft revealed that in some cases there was a considerable difference in scores from month to month, or from one point in time to another point several months later. These differences could probably be attributed to gradual failure, or replacement, of some component of the bombing system. In view of this, it was decided to further investigate the influence of the aircraft on the bombing process by analyzing the data for July through September, and October through

December, separately. These two periods were chosen only on the basis of having convenient reference points while also providing enough data points in each group to permit further analysis. Had there been sufficient data points, one-month or two-month time periods would have been preferable.

The July-September data included 154 data points. The variables were the same as for the preceding run except that two of the aircraft were not included since they had no releases during the period. The results for the July-September period:

Aircraft 008 -- Positive

Aircraft 009 -- Positive

Aircraft 023 -- Negative

$R^2 = .13$

The October-December data included 106 points, the same variables as previously, and 10 aircraft, the remaining aircraft having had three or less releases during the period. The results:

Aircraft 008 -- Positive

Aircraft 012 -- Positive

Aircraft 023 -- Negative

RN-N more than four missions together -- Positive

$R^2 = .36$

From the results of these runs which included the aircraft as variables, it appeared that at least some of the aircraft consistently had a greater effect on bomb scores than any crew factor. The effect could be positive, as in the case of aircraft 008, or negative as for aircraft 023.

If the bomb scores were being determined primarily by the aircraft regardless of a crew's capability, the influence of the aircraft would have to be overcome before significant crew inputs could be identified. The problem can be visualized as in Figure 1.

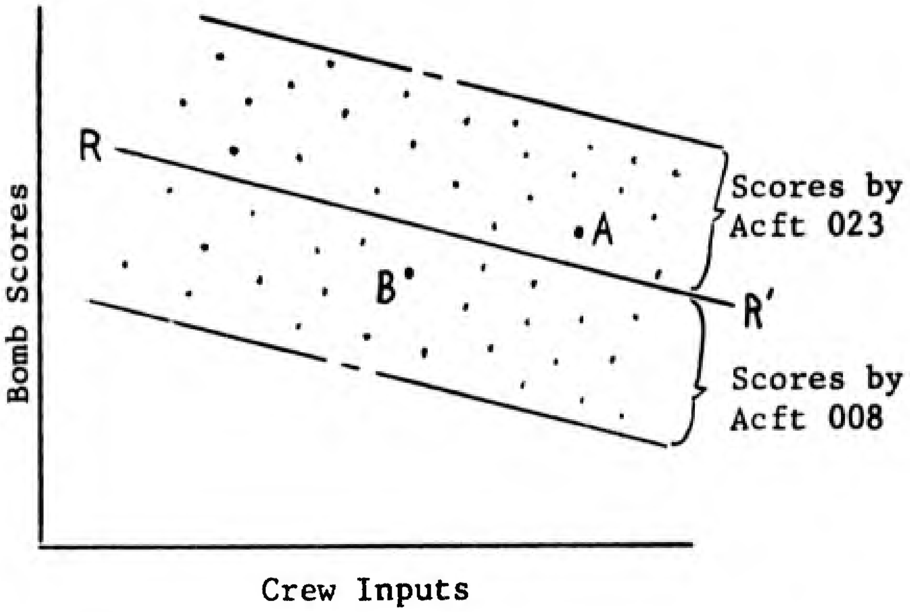

Figure 1. Effect of Aircraft Bias

Referring to Figure 1, with all data points generated by aircraft 023 and 008 included in one analysis, the regression method might fit line RR' to the data. If one of the crew inputs were, say, RN B-52 flying time, and the RNs who generated points B and A had 800 and 1,200 hours of B-52 time, respectively, an increase of B-52 flying time would appear to have an adverse effect on bomb scores; i.e., make the scores larger. If the analysis were to be done only on those scores made with either aircraft 023 or 008 however, an approximation of the relationship of flying time to bomb scores could be determined.

The large number of variables being investigated, and the relatively small number of scores available per aircraft, precluded using just one aircraft for the analysis. But if a group of aircraft were all observed to be getting approximately the same scores, then analysis of that group should effectively remove the influence of the aircraft and allow the influence of crew factors to be determined.

Blocking By Aircraft

A method of blocking the aircraft was devised as follows:

1. The average scores for each aircraft for the July through September, and October through December time periods were computed. All reliable scores for all aircraft were used; i.e., the scores from missions which had been excluded

from the analysis for the reasons outlined in Chapter II were also included (with the exception of ORI missions).

2. The average scores for each aircraft/period were then arranged in ascending order and examined for natural groupings by which the aircraft might be categorized. In Table I on the following page, the results are shown.

3. The aircraft/period were arbitrarily divided into the three groups indicated simply by observing that the scores were relatively clustered in those groups, with "breaks" between groups. Hereafter, these groups will be referred to as group I (good), group II (medium) and group III (poor). There were 66 data points in group I, 104 in group II, and 90 in group III.

Grouping aircraft of similar quality in this manner, and analyzing within groups, should have the effect of starting each crew out on an equal basis for achieving bomb scores. Within a group of like aircraft then, it should be possible to determine whether crew factors of experience and capability have a measurable effect on bomb scores.

Results of Aircraft Blocking

An analysis of each of the three aircraft groups, using all the original variables, was made. The significant variables and R^2 for each group are as follows:

Table I

AVERAGE SCORES BY AIRCRAFT/PERIOD IN ASCENDING ORDER

	Average Score	Aircraft	Period
Group I	7.2	012	Oct - Dec
	11.4	009	Jul - Sep
	12.3	008	Jul - Sep
	12.7	008	Oct - Dec
	13.6	025	Jul - Sep
	13.6	025	Oct - Dec
Group II	15.2	1008	Oct - Dec
	16.7	009	Oct - Dec
	16.8	039	Oct - Dec
	17.0	1008	Jul - Sep
	17.3	043	Oct - Dec
	17.9	051	Jul - Sep
	19.5	011	Jul - Sep
	19.7	043	Jul - Sep
	20.0	012	Jul - Sep
	20.3	014	Jul - Sep
	20.9	040	Oct - Dec
Group III	23.1	040	Jul - Sep
	23.5	039	Jul - Sep
	24.1	026	Jul - Sep
	26.1	021	Jul - Sep
	28.3	023	Jul - Sep
	34.8	023	Oct - Dec

Group I:

AC previous operational assignment -- Positive

RN total flying time -- Negative

Navigator stanboard input -- Positive

$R^2 = .21$

Group II:

Navigator B-52 flying time -- Negative

$R^2 = .05$

Group III:

RN B-52 flying time -- Positive

RN-N more than four missions together -- Positive

$R^2 = .12$

It had been expected that, with the external influence of the aircraft on the bomb scores effectively removed, the crew factors which affect the scores could be determined, and that the same factors would prove to be significant in the analysis of each group. One can see from the results that no single crew factor proved to be significant in more than one group. Perhaps the bias caused by aircraft was not the only factor which was preventing significant crew inputs from being identified.

Based on examination of data and a knowledge of the crew members' backgrounds, and in line with the research objective of improving selection of aircrew combinations, it was decided to further divide the data.

Blocking By Crew Members

By cursory examination of the data on individual crew members, it was obvious that aircraft commanders and radar navigators could be grouped roughly into categories. It is conceivable that if there were a great deal of difference in the individual contribution of experienced vs. inexperienced ACs or RNs on a crew, unless separate analysis were performed for crews with experienced personnel at either of those two positions, and crews with relatively inexperienced personnel at either of those two positions, it would be impossible to determine the significance of any crew factors in producing bomb scores. Therefore, in order to try to determine factors which would be applicable to a particular range of AC or RN experience, groups of ACs and RNs were identified, then analyzed separately within the same aircraft groups that were used previously. The aircraft commander groups are hereafter identified as group 1 and group 2; the radar navigator groups are hereafter identified as group A and group B.

Group 1--This group consisted of those ACs who, after pilot training, had flown in Southeast Asia for one year, then had been assigned to B-52s. They trained at CCTS as

ACs, then flew as ACs in the 17th Bomb Wing without having flown as copilots. There were eight ACs in this group. They had an average of 3.1 years of rated service, 371 hours of B-52 flying time, 1,242 hours total flying time, and less than one year as a B-52 AC.

Group 2--This group consisted of those ACs who had been assigned to B-52s upon completion of pilot training, trained at CCTS as copilots, then flew as copilots in the 17th Bomb Wing for approximately one and one-half years before upgrading to AC. There were seven ACs in this group. They had an average of 3.1 years of rated service, 916 hours of B-52 flying time and total time, and less than one year as AC.

The remaining eight ACs were generally older and more experienced as ACs than those in groups 1 and 2. They had an average of 7.4 years of rated service, 1,164 hours of B-52 flying time, 3,610 hours total flying time, and 2.1 years as a B-52 AC. However, this eight accounted for only 63 of the 260 total data points, so separate analysis of their performance within aircraft groups was not possible.

Group A--This group consisted of nine RNs, all of whom had recently been assigned to B-52s for the first time and had no experience as a B-52 navigator. They had an average of 14.8 years of rated service, 341 hours of B-52 flying time, 3,015 hours total flying time, and slightly more than one-half year as a RN. (Several of this group had had B-47 experience, but in all cases that had been at least eight years previous.)

Group B--This group consisted of the remaining 13 RNs, all of whom had spent most of their careers in B-52s and had also spent considerable time as B-52 navigators prior to upgrading to RN. They had an average of 9.8 years of rated service, 2,340 hours of B-52 flying time, 3,240 hours total flying time, and slightly more than two and one-half years as a RN.

While there was also a wide range of flying time experience among navigators, with just two exceptions all had flown in B-52s exclusively, and distinct groupings could not be identified. For informational purposes only, the following data on the navigators in the analysis is presented: the average years of rated service was 3.3 with a range from 1 to 9; average B-52 flying hours was 951 with a range from 90 to 2,540.

Results of Blocking by Crew Members

The results of this series of runs were generally mixed and inconclusive. The results are shown below.

AC Group 1, Aircraft Group I; 30 data points:

No significant variables.

AC Group 1, Aircraft Group II; 39 data points:

RN B-52 flying time -- Positive

N B-52 flying time -- Negative

$R^2 = .27$

AC Group 1, Aircraft Group III; 32 data points:

AC B-52 flying time -- Positive

RN-N more than four missions together -- Positive

$R^2 = .23$

AC Group 2, Aircraft Group II; 39 data points:[1]

No significant variables.

AC Group 2, Aircraft Group III; 41 data points:

RN B-52 flying time -- Positive

$R^2 = .13$

RN Group A, Aircraft Group I; 34 data points:

N B-52 flying time -- Positive

$R^2 = .16$

---

[1] AC Group 2, Aircraft Group I, had only 16 data points, so that grouping could not be analyzed.

RN Group A, Aircraft Group II; 46 data points:

No significant variables.

RN Group A, Aircraft Group III; 40 data points:

N B-52 flying time -- Positive

$R^2 = .15$

RN Group B, Aircraft Group I; 32 data points:

No significant variables.

RN Group B, Aircraft Group II; 58 data points:

P-RN first or second mission together -- Positive

RBS Site B -- Positive

$R^2 = .28$

RN Group B, Aircraft Group III; 50 data points:

RN B-52 flying time -- Positive

$R^2 = .06$

With this series of runs, the analysis using variables for which data was available on all crew members was complete. The conclusions are presented in Chapter V.

Other Runs

A variety of runs was also made to investigate the relationship to bomb scores of those variables for which data was not available on all crew members. To reiterate, those variables are: CCTS stanboard check, 17th Bomb Wing initial qualification check, standing in pilot or navigator training class, and college grade point average. None of these variables proved to be significant at the 95 percent level.

V. Conclusions and Recommendations

Aircrew Factors

Based on the results of this study, the one conclusion that can be drawn about the effect of aircrew factors of experience and capability on B-52 bombing results is this: with current manning, such as existed in the 17th Bomb Wing during the last half of 1972, the experience and/or capability of the individual crew members has little effect on bomb scores. At the 95 percent significance level, no single factor was found to consistently effect the scores either positively or negatively.

Separate analysis of scores achieved in aircraft categorized as "good," "medium," or "poor" (as described in Chapter IV) failed to identify crew member factors which consistently had a significant effect on bomb scores. Separate analysis of scores achieved by crews with experienced radar navigators, and by crews with relatively inexperienced radar navigators, failed to identify crew member factors which consistently had a significant effect on bomb scores for any range of experience. Likewise, separate analysis of scores achieved by crews with aircraft commanders who were relatively new to the B-52, and by crews with aircraft commanders with B-52 experience exclusively, produced only negative results.

Although in a majority of the computer runs some crew member factors were found to be statistically significant, the R^2 values of the equations containing the factors indicated that only a very small percentage of the variance of the bomb scores--less than 29 percent in all cases--could be explained by those factors. Because of this, combined with the fact that no factors were found to consistently be significant under similar conditions, based on this study, no recommendations can be made for selecting aircrew combinations to improve bomb scores.

The Aircraft

A conclusion regarding factors other than crew member factors is this: individual aircraft can have a greater effect on bomb scores than any singly identifiable crew factor. When all available data points were used in an analysis which included all variables, three of the four variables which proved to be significant were aircraft. Two of these aircraft, one with a positive and one with a negative effect, also proved to be significant in separate analysis of scores generated in the July through September, and October through December time periods. In the latter analysis the largest R^2 value attained during the entire study was achieved. While the majority of the aircraft did not prove to affect scores significantly, it is interesting to note that a least one aircraft (aircraft 008) had a great enough positive effect, and another (aircraft 023) a great enough negative effect to consistently have a greater effect on bomb scores than any crew member factors. Based on that information one can conclude that, given those factors considered in the analysis, the best way to improve overall bomb scores would be to improve the bombing performance of the aircraft.

Recommendations for Further Study

It is clear that before the effect of crew member inputs in the bombing process can be analyzed, the inherent differences in the performance of the various aircraft have to be overcome. It is recommended that any future studies of this type address that problem first. One possibility which was not explored in this study would be a weighting factor applied to the various aircraft based on the relative merits of the aircraft as determined by maintenance personnel or the aircrews.

As pointed out in Chapter II, numerous variables which could effect bomb scores were not examined in this study.

Most obvious among the crew variables is crew morale and/or motivation. It is recommended that any future studies of this type include some measure of those factors.

If air crews do not have a great deal of control over the bomb scores they achieve, perhaps bombing accuracy should carry little weight in determining officer effectiveness. Similarly, at the wing level, bombing reliability should not be the (effectively) only criteria for determining success on an ORI. It is recommended that SAC give consideration to both of these points.

The fact that crew member stanboard performance did not prove to be a significant factor in determining bomb scores may be an indication that stanboard evaluations do not provide a good measure of crew members' ability to perform the assigned mission. It is recommended that stanboard evaluation procedures and criteria be examined to determine whether crew member performance is, in fact, being measured.

The process of a crew dropping bombs from an aircraft onto a target is a system wherein inputs determine an output. As such, the system is subject to analysis to determine which combinations of inputs produce the optimum output. The inputs will be changing over time; as the inputs change, the process changes, so new analysis should be performed on a continuing basis.

Bibliography

1. Action, Forman S. Analysis of Straight-Line Data. New York: John Wiley and Sons, Inc., 1959.

2. Anderson, T. W. An Introduction to Multivariate Statistical Analysis. New York: John Wiley and Sons, Inc., 1958.

3. Blackman, R. B. Linear Data-Smoothing and Prediction in Theory and Practice. Reading, Massachusetts: Addison-Wesley Publishing Company, Inc., 1965.

4. Enthoven, Alain C. and Smith, K. Wayne. How Much is Enough? New York: Harper and Row, Publishers, 1971.

5. Freund, John E. Mathematical Statistics. Englewood Cliffs, New Jersey: Prentice-Hall, Inc., 1971.

6. Goldberger, Arthur I. Topics in Regression Analysis. New York: The MacMillan Company, 1968.

7. Hajek, Jaroslav. A Course in Nonparametric Statistics. San Francisco: Holden-Day, 1969.

8. Huang, David S. Regression and Econometric Methods. New York: John Wiley and Sons, Inc., 1964.

9. Linnik, Yu. V. Method of Least Squares and Principles of the Theory of Observations. New York: Pergamon Press, 1967.

10. Mendenhall, William. Introduction to Linear Models and The Design and Analysis of Experiments. Belmont, California: Wadsworth Publishing Co., 1968.

11. Meyer, Paul L. Introductory Probability and Statistical Application. (Second Edition) Reading, Massachusetts: Addison-Wesley Publishing Company, 1970.

12. Revesz, Pal. The Laws of Large Numbers. New York: Academic Press, 1968.

13. Snyder, Lt. Col. William P. National Security Management: Case Studies in Military Systems Analysis. Washington, D.C.: Industrial College of the Armed Forces, 1967.

14. Sprent, Peter. Models in Regression. London: Methuen and Co. Ltd., 1969.

15. _______. "Basic Concepts and Analysis of Measurement Data." Army Materiel Command Pamphlet 706-110. (1969).

16. Technical Report 69-1. United States Air Force Academy, Colorado, 1969.

Appendix A

BMD Stepwise Regression Routine

This program computes a sequence of multiple linear regression equations in a stepwise manner. At each step one variable is added to the regression equation. The variable added is the one which makes the greatest reduction in the error sum of squares. Equivalently it is the variable which has the highest partial correlation with the dependent variable partialed on the variables which have already been added; and equivalently it is the variable which, if it were added, would have the highest F value. In addition, variables can be forced into the regression equation. Non-forced variables are automatically removed when their F values become too low. Regression equations with or without the regression intercept may be selected.

Output from this program includes:

(1) At each step:

  a. Multiple R

  b. Standard error of estimate

  c. Analysis-of-variance table

  d. For variables in the equation:

    1. Regression coefficient

        2. Standard error

        3. F to remove

    e. For variables not in the equation:

        1. Tolerance

        2. Partial correlation coefficient

        3. F to enter

(2) Optional output prior to performing regression:

    f. Means and standard deviations

    g. Covariance matrix

    h. Correlation matrix

(3) Optional output after performing regression:

    i. List of residuals

    j. Plots of residuals vs. input variables

    k. Summary table

The limitations per problem are:

(1) Maximum number of variables: 80

(2) Maximum number of cases: 9,999

(3) Maximum number of variables to be plotted: 30

Vita

Kenneth H. Biehle was born in [redacted]. He graduated from high school in Perryville, Missouri, attended Southeast Missouri State College for two years, then entered the United States Air Force Academy. He received his Air Force Commission and his Bachelor of Science degree from the Air Force Academy in 1960, then entered pilot training. After receiving his pilot's wings, he was with the Military Airlift Command for five years, flying C-124 aircraft. He was next assigned to the Defense Atomic Support Agency for one year, then flew KC-135 aircraft in SAC for three years prior to entering the Air Force Institute of Technology in August, 1971.

Permanent Address: [redacted]

This thesis was typed by Mrs. Dorothy K. Mote.