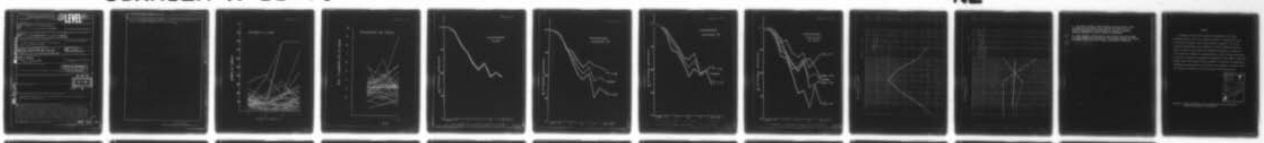
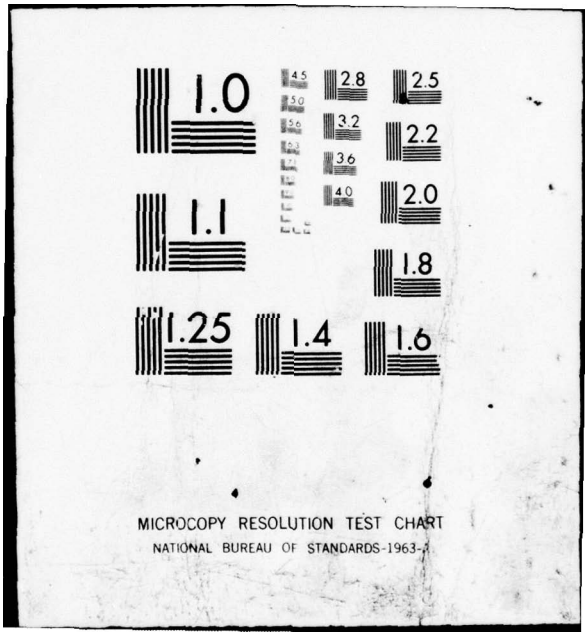
AD-A069 850 ARMY RESEARCH INST OF ENVIRONMENTAL MEDICINE NATICK MA F/G 5/10
THE FALLACY OF THE AVERAGE PERSON,(U)
MAY 79 B J FINE, J L KOBRICK

UNCLASSIFIED USARIEM-M-11/79 NL

| OF |
AD
A069850



END
DATE
FILMED
7-79
DDC


1.0
1.1
1.25
1.4
1.6
1.8
2.0
2.2
2.5
2.8
3.2
3.6
4.0
4.5
5.0
5.6
MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

UNCLASSIFIED
SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

DDC

(1) LEVEL II

# REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

| | |
|---|---|
| 1. REPORT NUMBER: M 11/79 | 2. GOVT ACCESSION NO. / 3. RECIPIENT'S CATALOG NUMBER |
| 4. TITLE (and Subtitle): (6) The Fallacy of the Average Person | 5. TYPE OF REPORT & PERIOD COVERED |
| | 6. PERFORMING ORG. REPORT NUMBER |
| 7. AUTHOR(s): (10) Bernard J. Fine, Ph.D. and John L. Kobrick, Ph.D. | 8. CONTRACT OR GRANT NUMBER(s) |
| 9. PERFORMING ORGANIZATION NAME AND ADDRESS: US Army Research Institute of Environmental Medicine, Natick, MA 01760 | 10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS |
| 11. CONTROLLING OFFICE NAME AND ADDRESS: (14) USARIEM-M-11/79 | 12. REPORT DATE: (11) 21 May 1979<br>13. NUMBER OF PAGES |
| 14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office): (12) 24p. | 15. SECURITY CLASS. (of this report)<br>15a. DECLASSIFICATION/DOWNGRADING SCHEDULE |

16. DISTRIBUTION STATEMENT (of this Report)

Distribution of this document is unlimited.

**DISTRIBUTION STATEMENT A**
**Approved for public release; Distribution Unlimited**

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

DDC RECEIVED JUN 14 1979 B

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

average performance; sets; stereotyping; group performance; individual performance; individuality

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

An analysis is made of the origins and transmission of so-called scientific "truth", and of the influence of "sets", or predispositions for certain ways of thinking, upon current concepts of behavior. One of the principal prevailing sets is the tendency to interpret behavioral research results in terms of group rather than individual performance, in which use of the concept of "average" is transformed from an index of group performance to one which represents the individual performances of each member of the group. Examples taken from

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

040 850

Unclassified
SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

79 06 01 052

research data are presented to illustrate the dangerously misleading conclusions which can be drawn by such procedures, and their potentially dangerous consequences for public believes about human behavior.

Figure 1

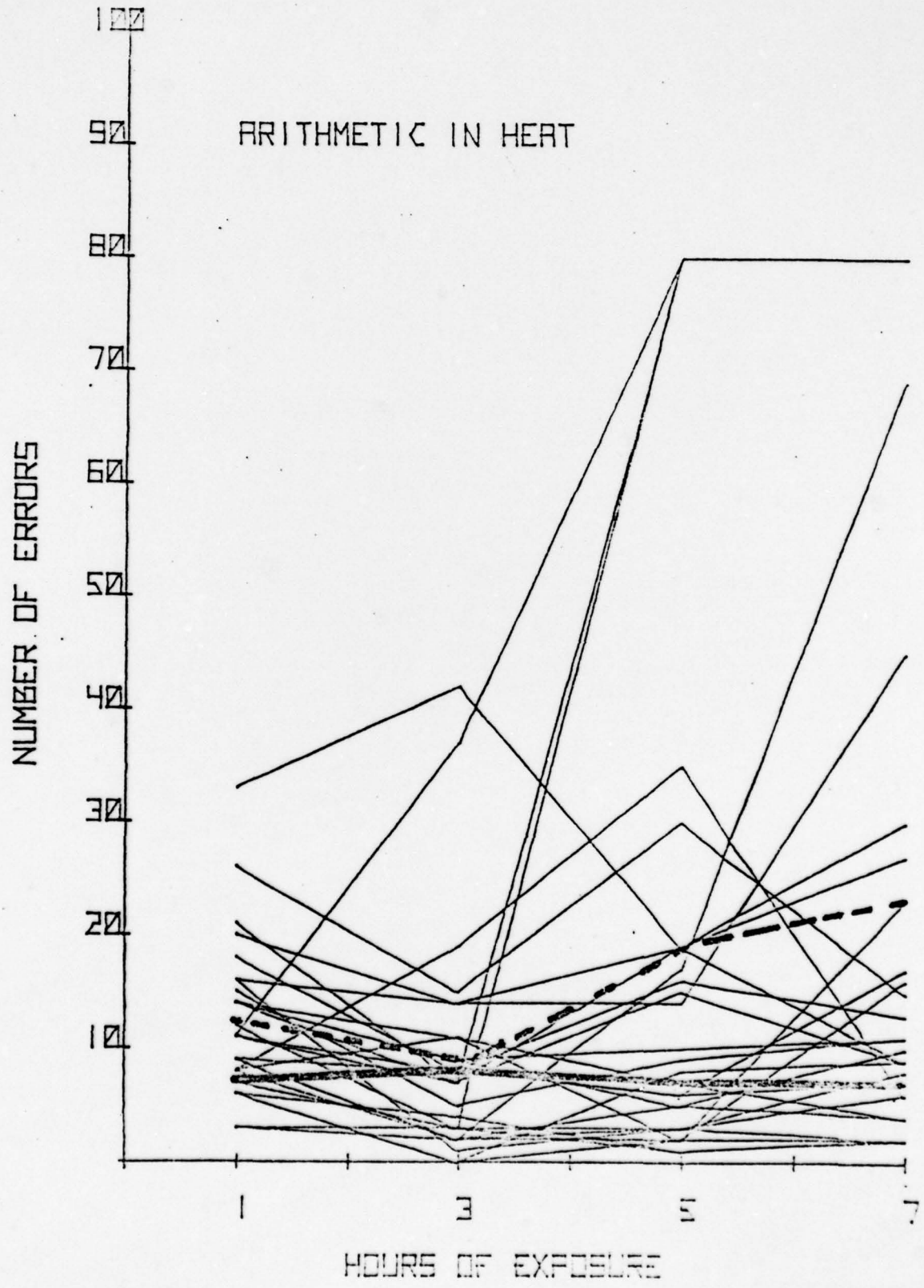

ARITHMETIC IN HEART
NUMBER OF ERRORS
100
90
80
70
60
50
40
30
20
10
1
3
5
7
HOURS OF EXPOSURE

Figure 2

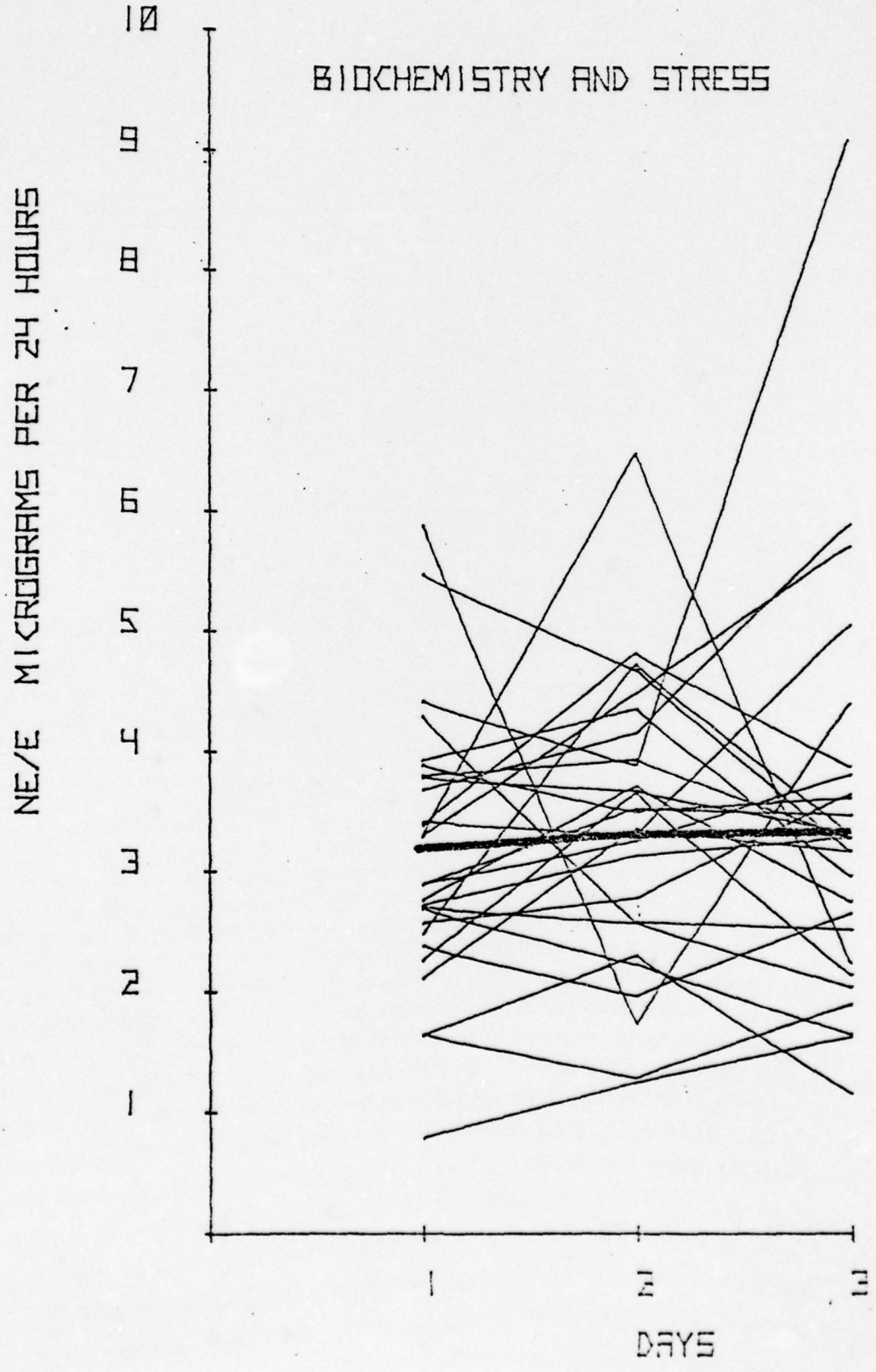

BIOCHEMISTRY AND STRESS
NE/E MICROGRAMS PER 24 HOURS
10
9
8
7
6
5
4
3
2
1
1
2
3
DAYS

Figure 3

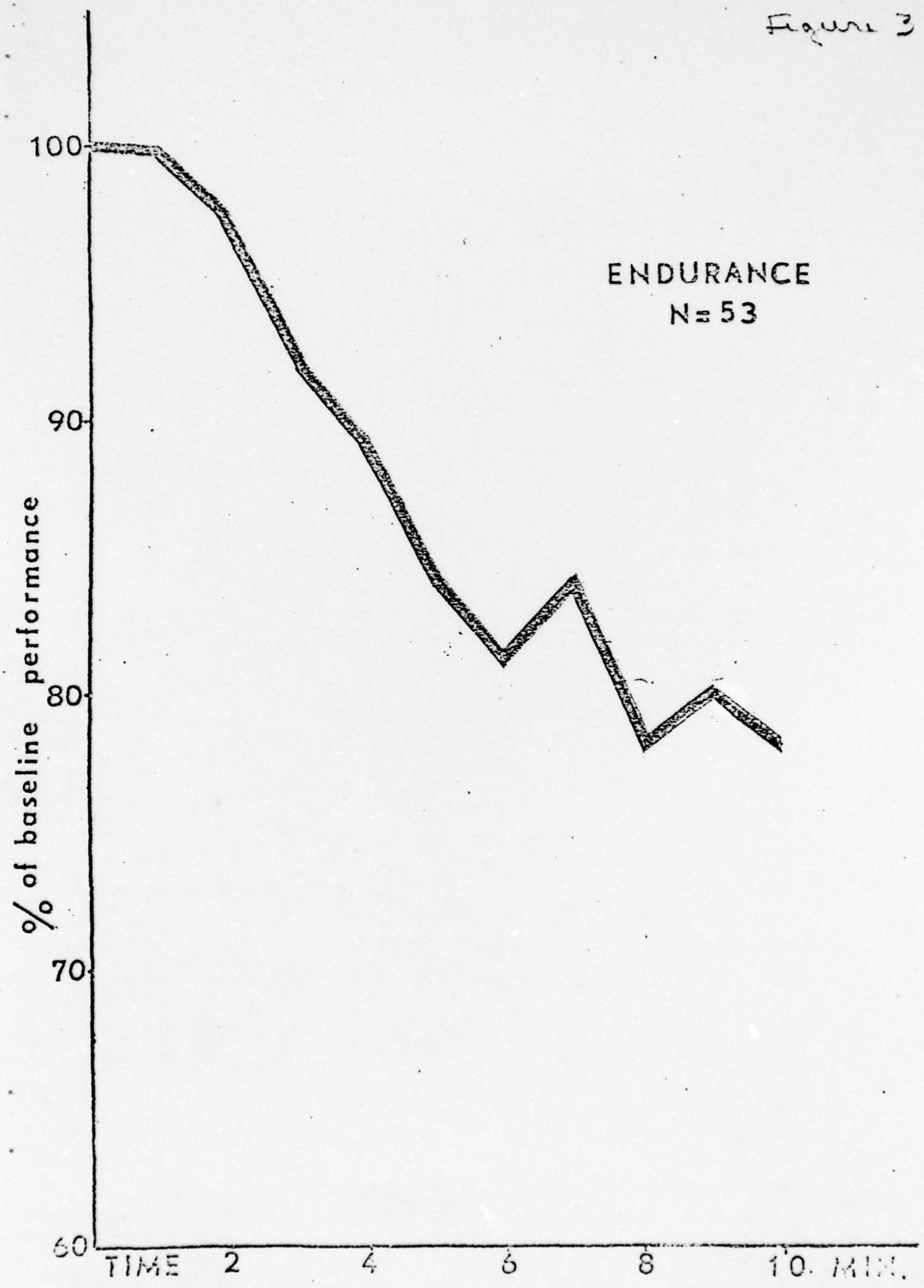

ENDURANCE
N=53
% of baseline performance
100
90
80
70
60
TIME
2
4
6
8
10. MIN.

Figure 4

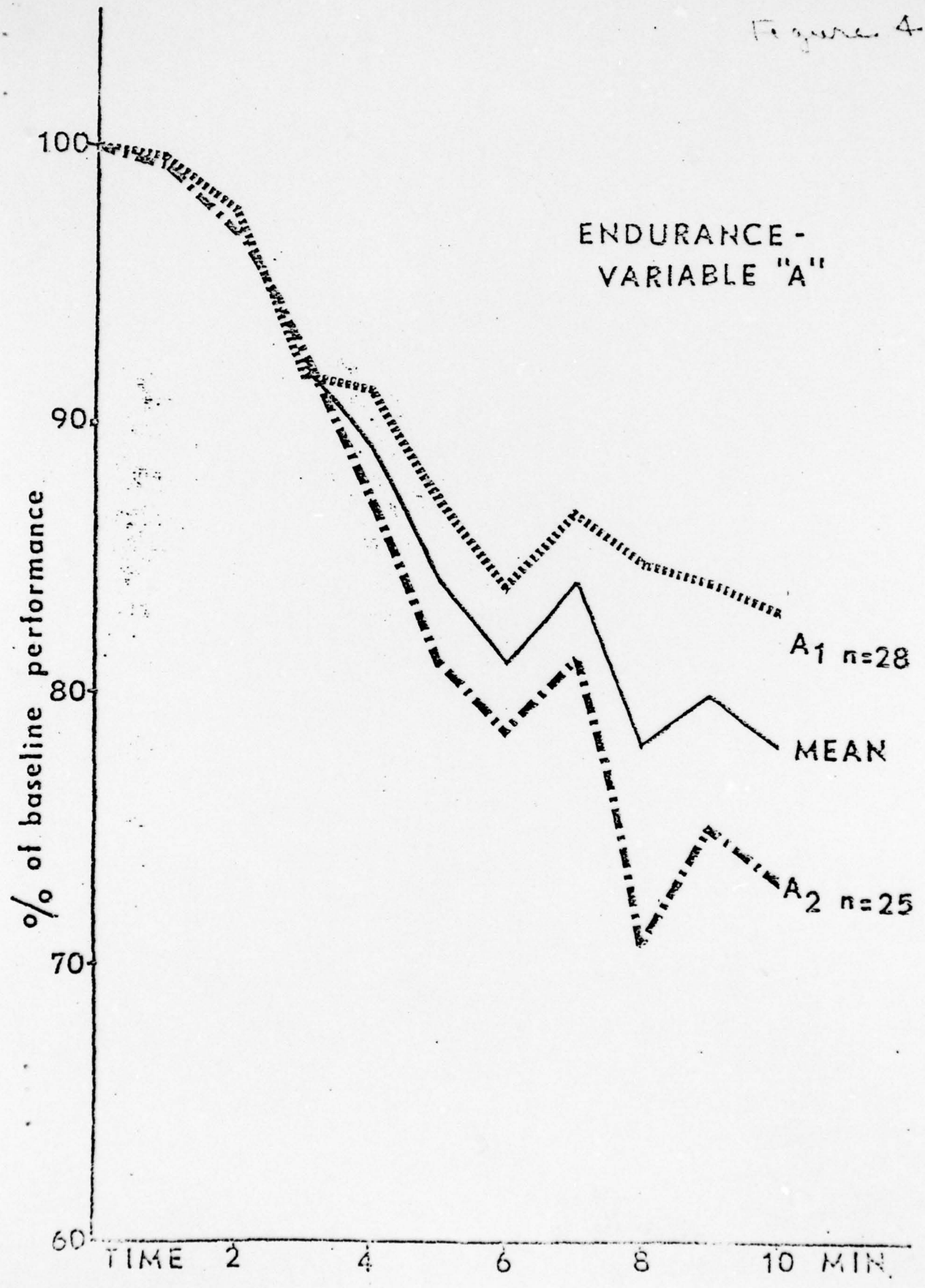

ENDURANCE -
VARIABLE "A"
100
90
80
70
60
% of baseline performance
TIME 2
4
6
8
10 MIN.
A1 n=28
MEAN
A2 n=25

Figure 5

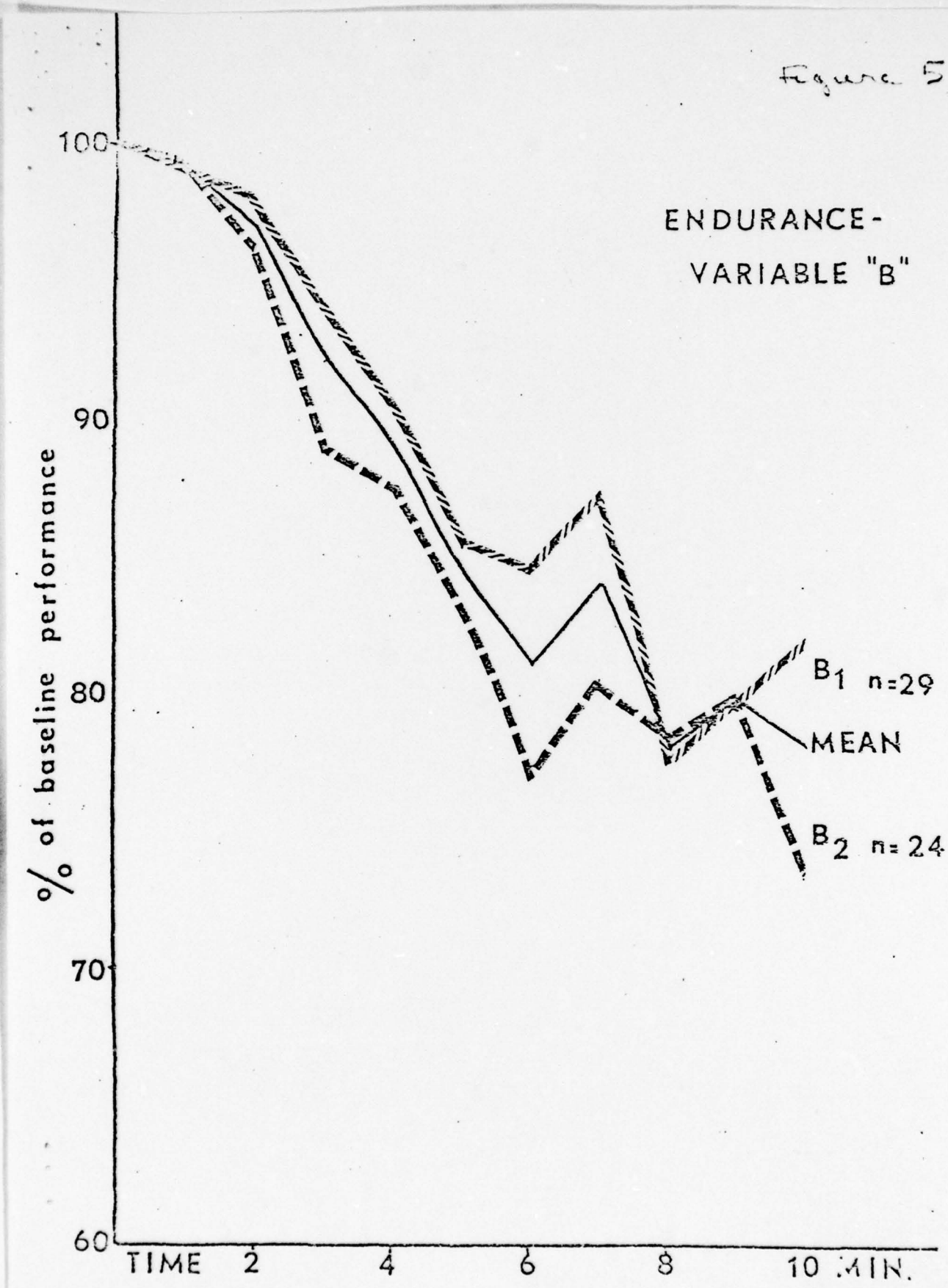

ENDURANCE-
VARIABLE "B"
100
90
80
70
60
% of baseline performance
TIME
2
4
6
8
10 MIN.
B1 n=29
MEAN
B2 n=24

Figure 6

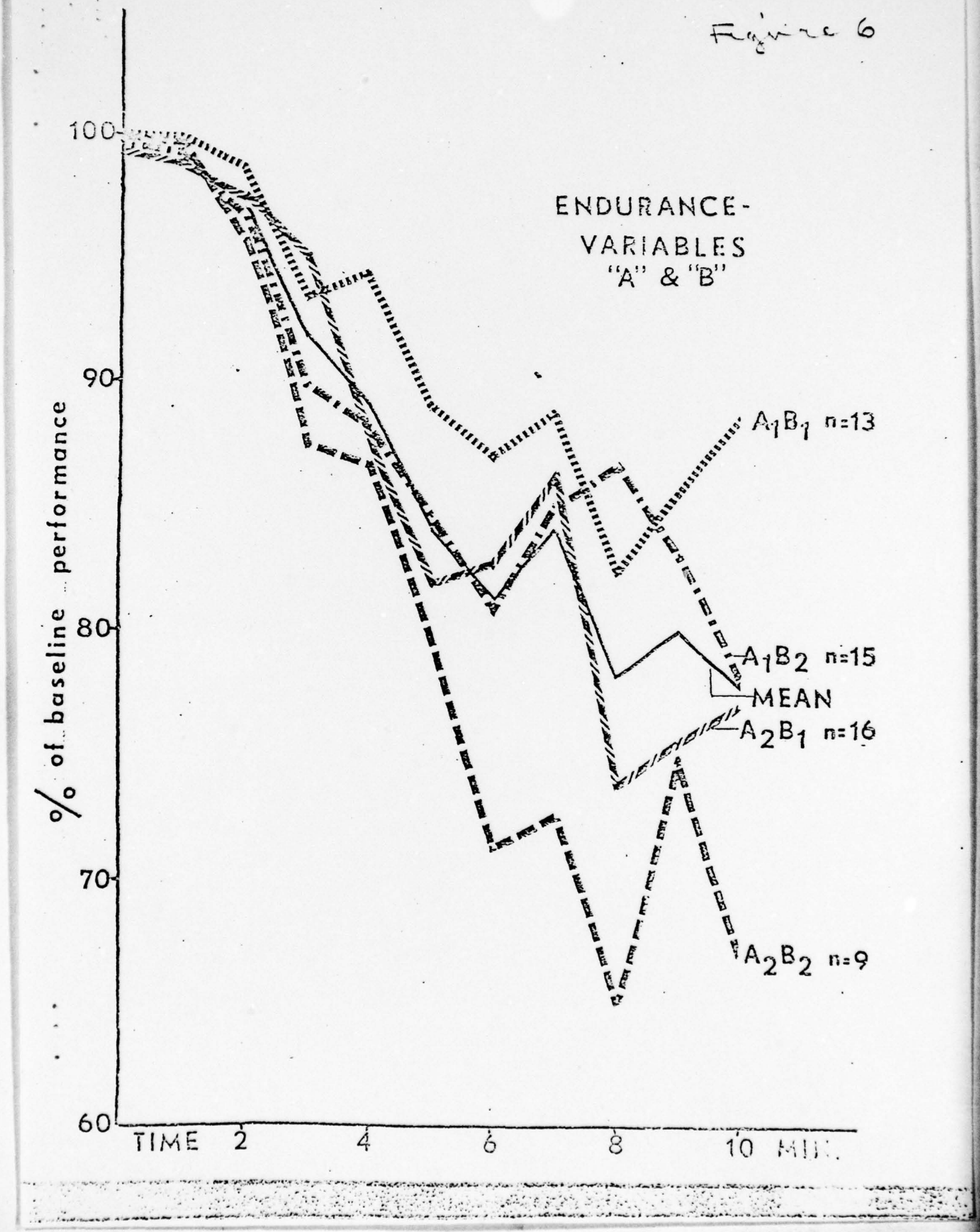

ENDURANCE-
VARIABLES
"A" & "B"
% of baseline performance
100
90
80
70
60
TIME 2
4
6
8
10 MIN.
$A_1B_1$ n=13
$A_1B_2$ n=15
MEAN
$A_2B_1$ n=16
$A_2B_2$ n=9

Figure 7a

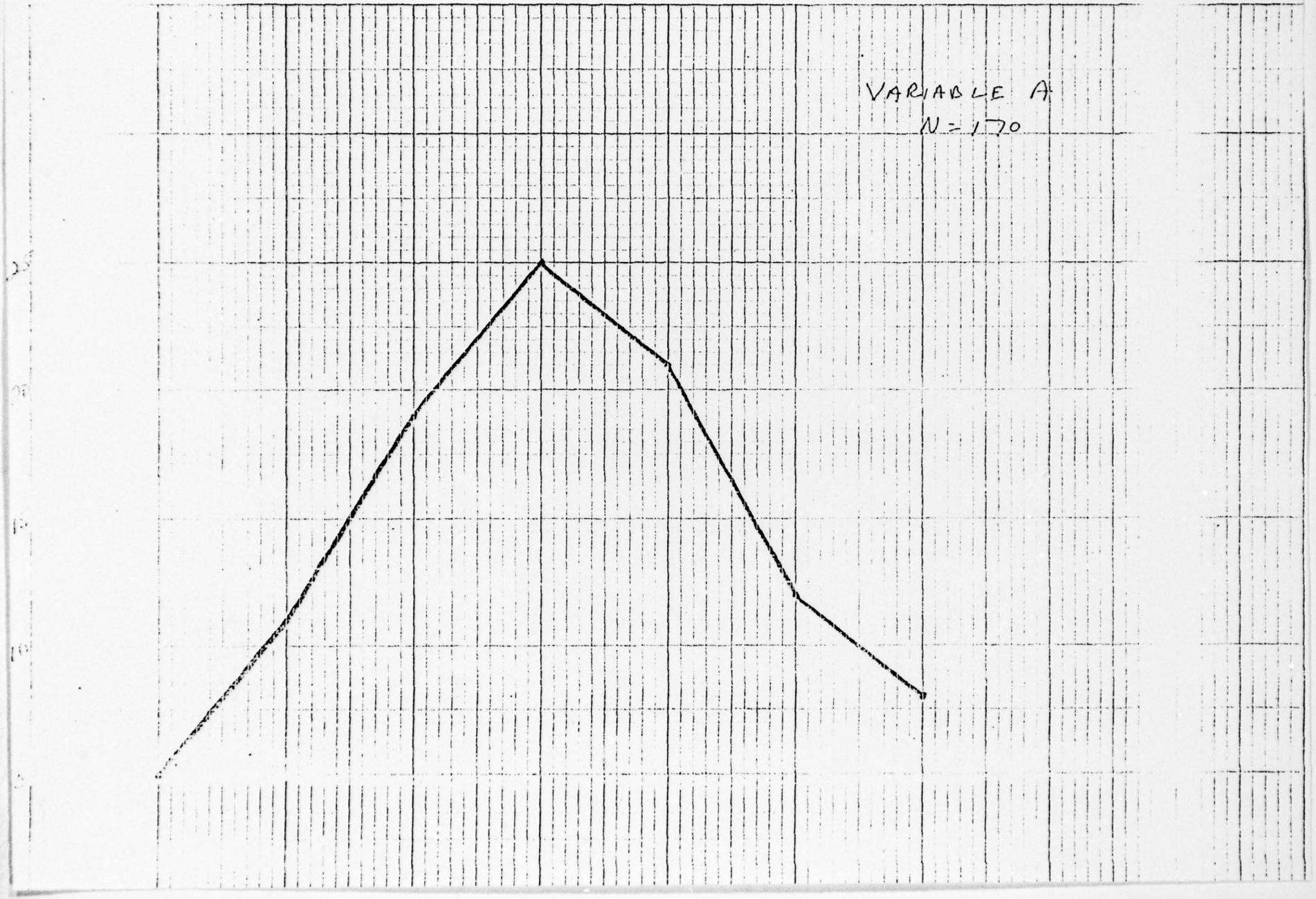
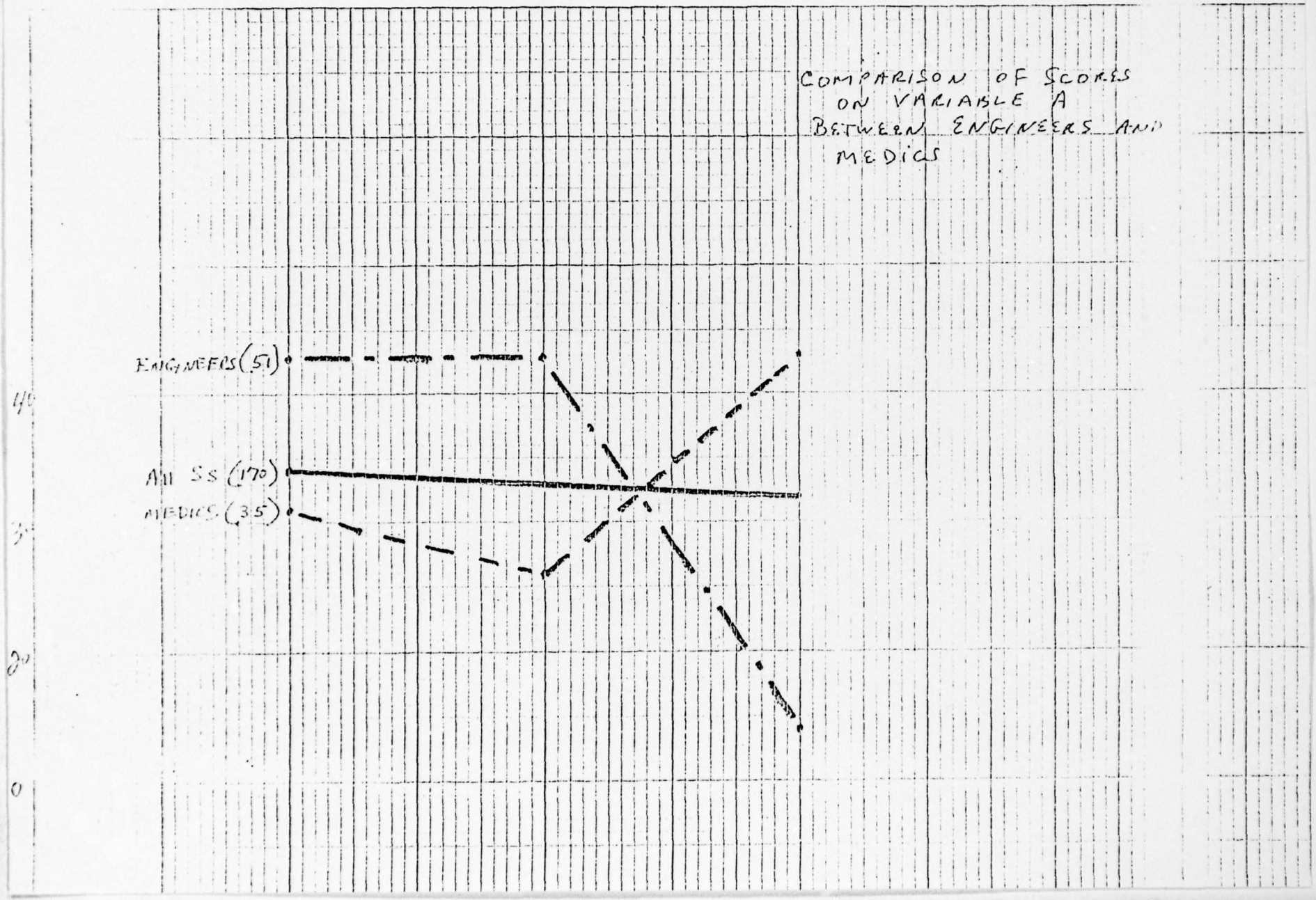


Figure 1.8



1. The views, opinions, and/or findings contained in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy, or decision, unless so designated by other official documentation.

2. Human subjects participated in these studies after giving their free and informed voluntary consent. Investigators adhered to AR 70-25 and USAMRDC Regulation 70-25 on Use of Volunteers in Research.

ABSTRACT

An analysis is made of the origins and transmission of so-called scientific "truth", and of the influence of "sets", or predispositions for certain ways of thinking, upon current concepts of behavior. One of the principal prevailing sets is the tendency to interpret behavioral research results in terms of group rather than individual performance, in which use of the concept of "average" is transformed from an indiex of group performance to one which represents the individual performances of each member of the group. Examples taken from research data are presented to illustrate the dangerously misleading conclusions which can be drawn by such procedures, and their potentially dangerous consequences for public believes about human behavior.



Keywords: average performance; sets; stereotyping; group performance; individual performance; individuality

# The Fallacy of the Average Person

Bernard J. Fine and John L. Kobrick

In recent years there has been considerable growth of interest in and concern with human behavior. The main focus of this increased attention has been on certain aspects of individuality, particularly self-awareness. Terms such as "finding one's self" or "doing one's thing" have become part of our everyday vocabulary, and a spate of do-it-yourself "insight" books has saturated the marketplace. This remarkable quest for self-knowledge seems to have no end in sight.

We are not sure what all of this means, but it has caused us to re-examine what we know about the origins and transmission of scientific knowledge. Reflecting upon the scientific journal articles we read daily in the course of our work, upon the scores of scientific presentations we have attended, and upon our own combined research experience of many years, we have begun to develop strong reservations about the kind of information we see reaching the public or, for that matter, reaching other scientists, in the guise of "truth" or "fact." Our concern is not only with scientific truth as related to self-enhancement, but with truth in all areas of the behavioral sciences.

The origins of material passed on to the public as "truth" are very difficult to trace. Even when the material is referenced to particular scientific journals, one frequently finds that such "truths" really are amalgamations of the research results of many articles. It is our impression that even though some of this information is based on the unsubstantiated opinions of authors of dubious competence, most of

these "scientific truths" come from reputable investigators well trained in their specialties. Such research results are usually published first in scientific journals, and later become consolidated and summarized periodically by textbook authors who presumably possess enough training to insure competence. This material eventually filters down to the general public as behavioral science "truths" in the form of popularized articles written by the scientists themselves, or by journalists skilled in writing about science.

A great deal can be lost in the transition from a scientific journal article to the popularized account of it in a magazine or newspaper. It might appear at first that such losses occur during the translation from the scientific language into everyday, ordinary language. However, it is our impression that only some misinterpretations occur in this way. By far the the most important information loss, or loss of "truth," occurs within the original scientific articles, when the authors themselves interpret the research results and draw conclusions about what was accomplished or discovered.

Our main purpose here is to try to clarify some aspects of this process for the general reader who we believe has developed an overly trusting attitude toward scientists, and is inclined to believe that scientific interpretations are unquestionable.

In the course of their training and in subsequent work experience, scientists develop certain outlooks which may be termed "sets," in the psychological sense of the word. Sets are predispositions to think in certain ways, and, in may cases, can be very strong and enduring. The medieval belief that the world was flat is but one of many examples. At that time, evidence to the contrary was rejected and even ridiculed because it was not consistent with the set held by contemporary scholars.

We do not mean to imply that sets are unique to scientists. Based on their own experiences, nearly everyone develops sets or biases about many aspects of the world in which they live. Many sets become integral parts of a person's "character", to the extent that external attempts to change those particular sets may be perceived as threatening and to be resisted.

The sets that scientists have are of particular importance in this regard, because science is explicitly and necessarily based on objectivity and on unbiased consideration of all points of view. Yet, because they have deep personal significance to the holder, sets in science, like other highly valued sets, are rarely changed by those who have them. Scientists are taught that science, as represented by the scientific method, is "objective," and that if they adhere rigorously to the "method," objectivity is assured. Values are rarely discussed; they merely are considered to be outside the realm of science. As a result, relatively few scientists are able to perceive that their way of conceptualizing scientific problems may be wrong. Rather, they assume that phenomena contradictory to their sets are due to flaws in their methods or in the methods of others, but not to anything inherently wrong with the thought processes of the set itself. Thus, science becomes faith, and the scientist becomes a pursuer of truth, but only within the framework of his sets about the world. This tends to substantially impede scientific progress.

Behavioral science in this country is particularly oriented around a set in which research results are interpreted in terms of groups rather than individuals. That is, in the typical research study, the performance scores of the individual members of a group are averaged, and this average is then used to represent the performance of the group. We do not question this procedure. However, we are

concerned about another practice of many authors in which the concept of "average" is unwittingly transformed from an index of group performance to one which represents the individual performance of each member of the group.

It is difficult to pin-point just how this transformation takes place. It seems that once the group averages and their associated variances have been calculated and pertinent statistical tests have been performed, most investigators, because of their training, orientation, and sets, feel that they have completed their data analysis. The possibility of examining the individual scores that were used in the calculation of the group averages and variances is rarely entertained. Thus, group averages and variances become the only points from which the interpretation and discussion of the results can be developed.

Furthermore, as they proceed, most authors manifest their set for groups by ignoring the variances, the only remaining indicators of individual differences in behavior. They focus their attention solely on the group means. Through a curious semantic circumlocution which is difficult to describe but easy to recognize, the group mean comes to be thought of as though it was the actual score of every member of the group. Statements such as "exposure to temperatures of 95°F for two hours resulted in a decrement of motor performance," or "daily physical exercise leads to significant improvement in cognitive performance," typically are made. They are based on group mean changes in the face of large variances which clearly indicate that many individuals were not only unaffected by the variables in question, but possibly changed in directions opposite to that noted in the author's conclusion. It is this type of empirical "foundation" upon which much of behavioral science has been built.

This entire procedure is dangerously misleading, because of what we term

"reverse stereotyping," a process by which each individual is improperly labelled by the average characteristics of the group. We call this abuse of logic the "fallacy of the average person." This fallacy has profoundly influenced the focus and implications of most American research in the behavioral sciences and leads us to question much of what has been published as scientific truth.

We believe that the "fallacy of the average person" stems from a gross misconception of the degree to which people actually differ from one another in virtually every aspect of human functioning. The influence of this fallacy pervades the American behavioral sciences, and, in our judgment, has led both scientists and public alike to overestimate the degree to which behavior can be predicted, understood and explained by currently available knowledge.

The issue is crystallized by questions such as: "Who is the 'average' person?" "What is the meaning of an 'average' response?" It can be very instructive to consider some examples which speak for themselves, and which show where some of the problems lie.

The examples we cite are from our own research. We do this for two reasons. First, the problems we write about are widespread, and we do not wish to single out particular individuals. Second, individual scores upon which group means are based are rarely published in journal articles or textbooks, and, therefore, are unavailable for use as examples. However, we know that the problems do exist outside of our own research efforts because, in our experience, we have seen raw individual data from many research studies, and the ways in which authors handled the individual data and formed conclusions.

Figure 1 shows the errors of 28 soldiers who participated in a study of the effects of a hot climate on performance. The errors occurred on a task involving

arithmetic calculations. The heavy solid line represents the average performance of the men over a period of seven hours under normal temperature conditions. The heavy broken line represents the average performance of the same group of men under severe heat stress. The difference between the two curves was shown statistically to be highly significant. On the basis of the average curves alone, one would certainly conclude that exposure to heat caused a substantial increase in errors. In Figure 1, separate curves for the performance of each man under heat stress (narrow lines) are also shown. These individual curves indicate clearly that the overall effect of heat was due to the reactions of only about 1/3 of the group, and that the rest of the men were affected either slightly or not at all. Within the traditional set to which we have referred above, only the two average curves would have been reported, and the conclusion would have been drawn that heat strongly affected performance on this task. The implication would have been left with the reader that everyone in the study was adversely affected by the heat, and another behavioral science "truth" would have come into being. This exemplifies the reverse stereotyping process described above, since an adverse reaction to heat would now be attributed to each of the men, but which actually occurred in only a small number of them.

The situation depicted by Figure 1 is one in which a highly significant statistical finding can be shown to be quite limited in scope and generalizability. The next example illustrates the opposite kind of situation, one in which a statistically non-significant result masks considerable behavioral change among the study participants. The information represented in Figure 2 is taken from a study of the biochemical responses of 27 soldiers to a variety of stress-producing experiences on three successive days. The values shown are the ratios of two

stress-related hormones, adrenaline and noradrenaline, expressed as micrograms secreted per 24 hours.

The group average curve (heavy line) in Figure 2, clearly indicates that the group as a whole showed little change from day to day. We are quite certain that the group set of which we speak would have led to a conclusion that the daily conditions had no differential effect on the individual biochemical responses of the soldiers. Through the reverse stereotyping process, the reader would have been left with the impression that the individual soldiers were quite consistent and unchanging from day to day in their responses. However, the separate performance curves (narrow lines) show large differences within and between individuals in patterns of change over the three days. It is quite obvious that there is much more to these data than group averages alone would show. In the actual published account of this research, different clusters of personality measures were shown statistically to be highly related to performance on each day, indicating a strong correspondence between type of individual and type of situation.

The group set to which we refer emphasizes similarities among people regarding the structure and function of mind and body. In substance, this approach assumes that "a body is a body is a body," and, therefore, that "on the average," men are interchangeable. While obvious inter-individual differences such as age and sex may be taken into consideration in the design and analysis of experiments, the group set implicitly assumes that the basic underlying processes which "govern" behavior, and, thus, the resulting behavior itself, are the same from person to person. The experiment now becomes a device to determine <u>that</u> performance or behavior which, once known, is assumed to be standard under the given circumstances for all individuals.

An alternative to the group set is one in which the differences among people are emphasized. This set assumes that while people in general are similar, they still may differ quantitatively and qualitatively from one another in many ways which are manifested in their behavior. There may be many sources for these behavioral differences, ranging from genetically determined variations in nervous system strength and sensitivity to environmentally or culturally induced behavioral patterns.

It is assumed that these kinds of differences can be categorized and that people can be sorted into "types" on the basis of the categorizations. For example, individuals may be classified on the basis of such things as personality traits (anxiety, aggressiveness), psychodiagnostic categories (schizophrenic, neurotic), disease proneness (A vs. B coronary types) or brain hemisphere dominance (left or right). Experiments related to this point of view are designed specifically to study the differences between people, and an awareness of the characteristics or "types" of people is critical.

Let us compare the "average person" or group set with the "individual differences" set by an example which illustrates the differences between the two approaches.

Figure 3 shows the average performance curve for 53 soldiers on a severely difficult test of physical endurance. The curve shows the average minute-by-minute changes (decrements) in performance for the group. A set oriented around an average performance concept would lead to the conclusion that there was a characteristic 20% decrement in performance of this task over a 10-minute period.

The same performance is illustrated in Figure 4 except that the data have been plotted for two sub-groups of the 53 soldiers. The groupings are based on the

soldiers' scores on a personality test, referred to here as Variable A, for illustrative purposes. Group A1 includes those soldiers ranking above the median of the distribution of scores of Variable A, and Group A2 those ranking below the median. The large differences in performance between Groups A1 and A2 are readily observable.

In Figure 5, the same index of performance of the 53 soldiers is plotted for a second variable, "B." Sub-group B1 includes those ranking above the median of the distribution of scores of this personality test, and sub-group B2 includes those ranking below the median. Again, differences between the two groups can be seen.

The potential interpretive power of an individual differences approach can be seen in Figure 6. Here, the 53 soldiers are divided into four sub-groups, representing the possible combinations of the two "A" categories with the two "B" categories. The dramatic differences in performance between sub-groups A1B1 and A2B2 illustrate that the original curve, based on the total group average has little meaning.

There is another level of criticism of the group set as it relates to the quest for scientific "truth." We have, heretofore, observed that group means can be extremely deceptive when used to describe the behavior of the individual members of the group, and that logically it is erroneous to use a group statistic to represent the behavior of an individual. It also can be quite misleading to make generalizations about behavior from group to group without considering the kinds of people which make up each group.

To illustrate: we recently tested a group of 170 soldiers at an Army base. The members of the group came from units representing four different military specialties, engineers, medics, headquarters personnel and supply. Among the tests

administered was that for Variable A, the same one shown in Figure 4 to be related to a physical endurance task. Figure 7a shows the frequency distribution of scores on Variable A for the 170 men. It can be seen that the distribution conforms almost perfectly to the ideal normal probability curve.

In Figure 7b, this normal curve is expressed differently and is shown as the heavy black straight line. Here, the distribution of the 170 scores has been divided into approximately equal thirds. By this device, 34% of the 170 people scored between 0 and 15 on Variable A, 33% between 16 and 22, and 32% between 23 and 36.

Two of the occupational specialty sub-groups, medical and engineering, were then selected from the 170 total. Their scores are represented separately in Figure 7B, medical personnel by the red line and engineering personnel by the green line. It can be seen in Figure 7B that the two occupational sub-groups differ dramatically in the distribution of scores on Variable A. Whereas 43% of the engineers had scores ranging from 0-15, 31% of the medical personnel had scores in that category. Even more impressive is the upper third of the distribution which shows that 45% of the medics scored in that category as opposed to only 14% of the engineers.

Remembering the relationship shown in Figure 4 between physical endurance and Variable A, it is obvious that quite different physical endurance performance curves would be obtained between engineers and medics; the engineer performance would be more like the performance of the A2 group shown in Figure 4, whereas the performance of the medical personnel would be more like the curve shown for the A1 group in the same figure.

Since the group-oriented researcher tends to believe and act on the premise

that all men are basically the same, the source of his test subjects is apt to make little difference to him. In practice, any concern he may still have is reduced further by simple lack of availability of subjects. Thus, many of the conclusions reached by American behavioral science, in addition to being contaminated by "the fallacy of the average person," have been derived predominantly from research on college freshmen and sophomores. The fact that the student participants are from a school of engineering or from a medical school is nearly always stated quite properly in the method section of published articles, but there is usually no mention of the possibility that students in different academic areas may be quite different types of people with different bases for their behavior. In the final analysis, generalizations are frequently made about human behavior based on conclusions from single studies, with no regard for the non-comparability of groups of people.

The conclusions from these studies always seemed to us to be conservatively stated and appropriately tempered with caution, in keeping with a scientific tradition which frowns upon broad generalizations or rash claims about one's research findings. However, in writing about the "fallacy of the average person," we came to realize that the conservativeness and caution expressed by many behavioral scientists is misplaced and illusory.

Broad, uncautious generalizations are continually being made in the behavioral sciences, tempered by phraseology which gives the illusion of conservatism and caution, but which accomplishes this by lulling the reader with conservative-sounding cliches. Phrases such as "in our opinion," "despite the small numbers of subjects in this study," "it appears that," "the data suggest that," etc., are among the more common ones used. A properly cautious statement might read "While the results show a significant effect of heat on group performance after

seven hours of exposure, it should be noted that the significant result primarily was due to the influence of the exposure on only one-third of the subjects; the rest of the subjects either were unaffected or actually improved their performance." Instead of this type of statement, a sentence with a broad generalization of the results is more likely to be used, preceded by a qualifying phrase designed simply to conform to the investigator's set for caution and conservatism in a superficial way. Thus, we find statements such as: "It appears that exposure to 95°F for seven hours produces a decrement in performance," or "the data suggest that exposure to 95°F for seven hours produces...." The need for caution seemingly has been served, the investigator is reassured, and another fallacious generalization enters the body of "scientific knowledge," destined to become "truth" simply by its publication.

These are strong words and are sure to be unpopular with some of our colleagues who will probably accuse us of "putting down" behavioral science. That is not our intent. We have attempted in this article to address what we perceive to be a major problem. At the same time, we hasten to acknowledge that some of our own research is subject to the criticisms we have made.

It is a matter of great concern to us that the behavioral science literature has become so vast and is expanding so rapidly that it is difficult even for specialists, let alone lay people, to keep abreast of it. While many of our colleagues seem to get quite enthused about all of the "knowledge" being generated, we are hardpressed to find very many scientifically acceptable behavioral science truths in the literature. To us, much of the published research has only narrow meaning because it is based on "the fallacy of the average person." Many of the so-called "facts" which have been generated simply cannot be applied to the solution of real problems because these "facts" are oriented toward a

mythical "everyperson" who does not exist. Much of behavioral science as it is now practiced does not accurately reflect the full gamut of human behavior. To the extent that individuals or organizations pursue courses of action based on over-generalized "truths," behavioral science may even be dangerously misleading the public.

In this article, we have tried to inform potential users or consumers of behavioral science information about certain pitfalls which exist in the field as it is now practiced by most scientists in this country. Perhaps an informed public, primed to ask the right questions, can do much to change these current sets.