A FIELD STUDY OF THE EFFECT OF EXTREME WINTER CONDITIONS ON A TEST OF MOTOR PERFORMANCE

A Thesis
Presented to
the Department of Psychology,
The University of Manitoba.

In Partial Fulfillment
of the Requirements for the Degree
Master of Arts

by
Richard Noel Frizell

August 1949.

TABLE OF CONTENTS

CHAPTER PAGE

LIST OF TABLES

LIST OF FIGURES

CHAPTER I

THE PROBLEM AND THE EXPERIMENTAL APPROACH

I. THE PROBLEM

The general problem to which we seek an answer is whether exposure to severe winter conditions affects motor performance, and if so to what extent. The general problem has a more practical side in that it is desirable to know to what extent the individual is incapacitated in carrying out a manual task under cold weather conditions, equipped with the appropriate protective clothing and hand-gear. On a realistic motor task, to what extent is the individual handicapped by the weather conditions, and to what extent by the clothing he must wear? What advantage if any has the acclimatized individual over those not used to working under such extreme conditions? Is the manually skilled person more, or less, affected than the unskilled? These are some of the questions which we shall attempt to answer. They are important questions from a practical point of view, and their answers will have a bearing on the methods and value of selecting and training men who cope with these extreme conditions.

II. THE EXPERIMENTAL APPROACH

Nature and Description of the Test. An experimental task was sought which would approach as nearly as possible a realistic manipulation, such as might be required of one who must operate and make minor repairs and adjustments on machines. At the same time the task must admit of good experimental control. It was judged that some operation employing nuts and bolts would have as wide a practical application as any, and that the relatively simple nature of such an operation would allow the control necessary to the experimental test.

The apparatus devised consists of three series of bolts mounted with their heads sunk into a horizontal wooden base. The three series of bolts are graded as to size, the heaviest (series H) being of 1-1/4" diameter, the medium size (series M) of 5/8" diameter, and the lightest (series L) of 5/16" diameter. The nuts fitted to these bolts have the dimensions (diameter and thickness) 2-1/4" x 1-1/4", 1" x 3/4", and 9/16" x 1/4". The heavy nuts are square, the smaller sizes hexagonal. Nuts and bolts are of the ordinary commercial grade in mild steel. Nuts were of a rather slack fit, turning easily with the hand and at no point requiring forcing.

For convenience in handling and transportation the H series was mounted on one frame and the M and L on a second. In the smaller frame, the two series were separated by five inches. Working clearance, i.e. the distance between the nuts when these are placed on the bolts, was for the H series 2½", for the M 2¾", and for the L 3-1/4".

Starting with the nuts placed on the bolts so that the top of the nut was flush with the top of the bolt[1], the subject was required to turn them down to the end of the thread as quickly as possible, and his time was recorded for the completion of each of the three series. The test thus constitutes a crude measure of reaction time and manual dexterity. A similar test (described by Tiffin[2]) was constructed by C.A. Drake for the measurement of manual speed.

Experimental Design. In all, three test groups were employed, designated as follows: i) Ft. Churchill, ii) University, iii) D.I.D. The Ft. Churchill and D.I.D. experiments were carried out at or near Ft. Churchill, Manitoba during the winter months (January to March, 1949). The remaining part of the experiment was conducted at the University of Manitoba under mild summer conditions (May 1949.

---

1 With the nuts placed thus, the number of turns remaining on the five H bolts is 65, on the four M bolts 60, and on the three L bolts 48.

2 Tiffin, J., Industrial Psychology. (New York: Prentice-Hall Inc., 1946) p.304.

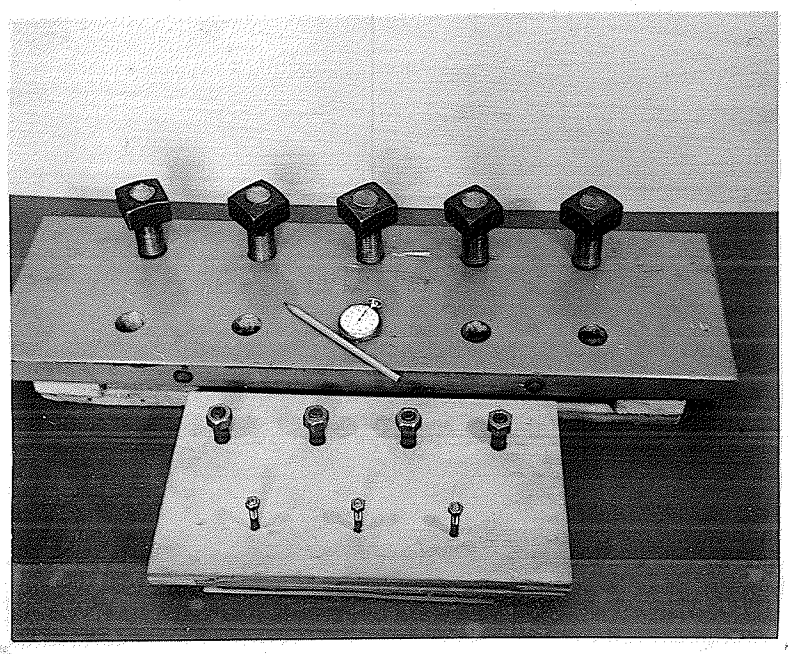

Fig. 1 - Test Apparatus

The basic experimental design involved testing under the following three conditions:

A - Outside, with normal winter clothing including the Canadian Army parka and fleece-backed issue mitts with felt liners.

B - Inside, with the same clothing as above.

C - Inside, with bare hand, and the parka removed.

The order in which the conditions were encountered by the subject was varied in three ways according to the following scheme:

		order of conditions		
subjects	1.	A	B	C
	2.	B	C	A
	3.	C	A	B

This pattern was repeated on successive groups of three subjects.

The effect of such a rotation of conditions is to counterbalance factors which may affect test results, and which operate with respect to position in the series[3]. Practice, since it is a cumulative effect, is one such factor. In the case of the first subject, the test score under condition C receives the greatest benefit of practice. In the second subject it receives only the average benefit, and in the third subject it receives only the minimum.

---

3 T.G. Andrews, Methods of Psychology (New York: John Wiley & Sons, 1948) pp. 13-15.

The second series of tests, administered at the University of Manitoba provided a control on the effect of weather, and supplied reliability data on the tests which could not be derived from the original material collected at Ft. Churchill.

In order to control the same 'series' factors, the rotation pattern described above was retained in the University experiment. Due to the limitations of time and the number of subjects available, only eighteen individuals were tested here. This change altered the original pattern only with regard to the numbers performing the three different bolt-orders. Three groups of six were tested in the orders HML, LMH, and MLH respectively, whereas in the Ft. Churchill experiment the first named order was administered to twelve subjects.

In addition to the original schedule of tests, retests were administered to the University subjects under two of the three conditions, viz. B and C. These are the two basic test conditions; the A situation merely adds a combination of variable weather factors to B. The purpose of the retests was of course to derive reliability correlations for the tests under the two conditions which were identical in the winter and summer situations. It also gives us added information on the extent of practice effects.

The third series of tests was conducted on a group of military personnel engaged in clothing trials carried out by the Directorate of Inter-Service Development in the vicinity of Ft. Churchill. It was originally planned to administer tests at intervals of one-half hour over a three-hour period, this repeated for several days. The nature of the subjects' duties, however, permitted only two tests on any given day. The first, administered in the morning, was designed to give a measure of their performance under normal conditions of minimum stress. The second, administered after two to three hours exposure to the weather was designed to measure the effect of the outdoor conditions. In fact, however, the first trial of the day proved to be the "stress" situation, as the subjects generally reported a more comfortable and warm condition after the outdoor exercise than they experienced on emerging from their tents in the morning. This did not however, essentially change the design of the experiment, but only the treatment of the date.

Only the heavy bolt series was used with these subjects, and one trial on the series constituted the test. As all equipment had to be hand-hauled on toboggans, it was not feasible to carry more.

Administration of the Test. The tests were administered in the same way to the Ft. Churchill and the University subjects.

Before beginning, the subject was permitted a practice turn on one bolt in each series with the mitt only in order to give him the "feel" of the test, and as an aid in communicating to him what was required. No practice was given with the bare hand as this is a familiar enough operation, and lacking the complications which the mitt involves. It was pointed out for example that when the mitt is not cleanly withdrawn at the end of a turn, the nut has a tendency to turn backwards with the return twist of the hand. He was instructed to employ as far as possible a deliberate turning action, using the thumb and fingers, and to avoid rolling the nut down with the end or side of the mitt. He was warned against trying to make the nut spin freely with a flip of the hand, but to keep it under control of the fingers at all times. It was suggested that he try to maintain a constant turning technique. Little difficulty was experienced with the subjects in this regard.

Each subject was to start from the side opposite to his preferred hand, and as all tested were right-handed, all proceeded from left to right in each series.

i) Ft. Churchill Group

Preceding the outdoor test, in the winter experiment, a minimum exposure period of fifteen minutes was required. During this period the subject was kept in a relatively inactive standing position. He was not permitted to pace about or to chafe his hands or body. He was allowed to smoke, and was generally engaged in conversation which effectively discouraged more vigorous physical activity. It was contrived to have the subject stand clear of the shelter of buildings so that he should have the full effect of the wind. If the subject had been engaged in some outdoor activity prior to the test (driving, walking, etc.), even if such activity occupied considerably more than the minimum exposure time, he was required to stand another five minutes before beginning. This was considered necessary because, from our own experience, even a relatively slow walk in heavy arctic clothing produces considerable body heat.

In those cases where the subject's inside test followed that outside, a warming period was given before the inside test was started. This continued until the subject reported comfortably warm, usually about ten minutes. The test apparatus was also warmed and dried by a stove so that no frost or moisture remained on the metal parts. This was easily accomplished within the period required for the subjects' warming by placing the apparatus on top of a stove.

Between the bolt series i.e., between H and M, and M and L, a rest of thirty seconds was permitted, and between adjacent tests indoors, with mitt and with bare hand, a rest of one minute. During these periods subjects used any means they pleased to limber or relax their arms.

ii) University Group

The test was administered in precisely the same way to the University group on the two inside tests. Preceding the outdoor test, however, there was no exposure period required. While this control could be criticized for not reduplicating the exposure period, it was felt that the outdoor conditions were not sufficiently different from indoors to affect the subjects' comfort. It did not seem sufficiently important to warrant detaining the subjects, particularly as all were on jobs, and had a limited time to give to the test.

The time between indoor and outdoor trials was the interval of three or four minutes necessary for transporting the test apparatus from one situation to the other. All other rest periods were the same as those given to the original group.

iii) D.I.D. Group

Only the heavy bolt test was administered to the D.I.D. subjects, as indicated in the previous section. Instructions given the subjects were the same, although no practice turns were allowed on the initial test. For the first test of the day, no preliminary exposure period was required, but the test was given as soon as the subject went outside. The second test was given immediately on return to the camp after the morning's exercises.

Conditions under which Tests were Administered.

i) Ft. Churchill Group

On the indoor trials at Ft. Churchill, all tests were conducted with the test apparatus placed on a firm table of standard height. Outdoors, conditions were not quite so uniform. At the location where seventeen of the subjects were tested, a bench was used which was within an inch of table height, and firmly set. The remainder of the subjects worked on a packing case outside which was four inches lower than the table used indoors. This might conceivably have increased the outside scores of these subjects, although the lower support did not seem to offer any additional difficulty, and there were no complaints that such was the case.

It might be observed here also that the larger frame (for the H series) had a base $1\frac{1}{2}$" thicker than the smaller one. This was necessary to accommodate the large heads on the heavy bolts, and in any case should not have any bearing on relative performance, since the added height was a constant factor on the heavy task.

It was remarked that in the outside situations the bumpy trampled-snow underfooting might also detract from performance. Our own experience convinced us that such was not the case to any noticeable extent. The tasks are light, involving little body movement, so that a good footing is not difficult to maintain even on soft or slippery ground.

It was the practice to place the subject during the test in such a position that the wind was either broken by a building or taken on the subject's back, so that he had clear vision, and the danger of frostbite was removed. The deep parka hood provided a good sun-shade, and since the test apparatus is of a relatively dark color, the effect of sunlight reflected from the snow was minimized.

It was feared at first that blowing snow might increase the difficulty of the task by lodging in the threads of the bolts, and stiffening the action of the parts. In the two instances where blowing snow was present to any considerable extent, however, we could observe no such effect when we reset the nuts for the next trial, and subjects remarked that they could detect no difference.

The only difference that existed between the inside test with mitts and parka and the outside one, so far as clothing is concerned, is that for the inside trial the subject was permitted to put down his parka hood. From our own experience with the test there is no notable difference in the degree of restriction affected by wearing the hood up. In the Canadian Army design, all of the weight of the parka is taken on the shoulders, not on the head as is the case in some patterns. The important effect of the parka for our purposes, in addition to the extra weight on the body, is the constricting and binding effect on the arm, particularly the lower forearm which is also covered by the gauntlet of the mitt. Moreover, it was felt that the additional warmth produced by the hood on the indoor trial would be a source of irritation and discomfort to the subject.

ii) <u>University Group</u>

As in the case of the original subjects, the University Group was also tested indoors on a standard table. A support of the same vertical dimension was used outdoors, except in the case of the first five subjects who used a workbench four inches lower than the table. Four of the subjects reported that they discerned no difference in the effect of this condition on the difficulty of the test. One reported that the lower bench made the task somewhat less fatiguing.

The same parka and mitts were used at the University as were used in Ft. Churchill. In the majority of cases at Ft. Churchill however, the subjects used their own clothing.

It seems unlikely that this fact in itself is of any practical importance, since all used the same model parka and mitts, and since the latter are made in only one size.

As in the case of the indoor trial at Ft. Churchill, the University subjects were permitted to wear the parka hood thrown back in both the indoot and outdoor tests. The reasons are the same as those which apply to the indoor trial at Ft. Churchill.

iii) D.I.D. Group

Test conditions were least well controlled in the case of the D.I.D. trials. It was the practice to place the frame on a small wooden chest lodged firmly on a shelf of snow blocks or a snow enbankment. As the test team moved camp frequently it was possible only once to use the same location for testing on two successive days. Under these circumstances it was difficult to maintain uniformly comfortable conditions. A standard mitt was used by all the subjects; a plain leather mitt with an index finger compartment. The finger compartment was not used, as it seemed to no advantage, and even a hinderance on the smaller nuts.

Subjects.

i) Ft. Churchill Group

The twenty-four subjects used in the Fort Churchill experiment were all male civilian personnel employed at D.R.N.L. or visitors to the camp. In terms of occupation they included the administrative and office staff of the Laboratories, maintenance people (mechanics and general duties), laboratory technicians, scientific observers and research workers. In terms of northern experience they include those with a few days residence, and some of several years. The average daily exposure to weather which each experiences ranges from a few minutes in the case of some of the office workers, to several hours in the case of mechanic-drivers. They are a heterogeneous group, but with few exceptions have had little specialized training in mechanical skills.

ii) University Group

The eighteen subjects tested at the University included seven academic people, seven skilled workers; mechanics, machinists, and radio technicians, three dairy employees, and one unskilled laborer. On the whole they constitute a more mechanically skilled group than the Ft. Churchill subjects. Like the Ft. Churchill people, they were all volunteer subjects.

iii) D.I.D. Group

The D.I.D. test subjects were all military personnel except two civilian observers, and without exception they possessed no technical or mechanical training. At the time of testing, all had had about four weeks indoctrination in the Army School at Ft. Churchill, a considerable amount of such time being spent in outdoor exercises and in living outdoors in tents.

Kinds of Data Collected. In addition to time scores, records were kept of the weather conditions at Ft. Churchill which prevailed at the time of testing. These included temperature, wind velocity, and wind-chill factor: also such special conditions as sun-shine, and the presence of blowing snow. Subjective feelings of cold were also noted from reports taken before and after testing. The subject was encouraged to give verbal reports also during the test on his feelings of fatigue, and a running record was kept. A short personal history was taken, which included information about the subject's occupation, the length of his experience in the north, and that part of the day normally spent out of doors.

At the University, the same kind of meteorological data was noted. There were no reports of subjective feelings of cold taken of course, and in all cases the subjects found the outdoor situation comfortable. Reports on fatigue were taken as in the case of the Ft. Churchill subjects.

During the D.I.D. trials, wind velocity and temperature were read at the time of testing, and a single fatigue report was taken at the end of the test. A cold report was also recorded for each testing.

CHAPTER II

BACKGROUND LITERATURE

The effects of exposure to extreme cold conditions on human functions and behavior have merited a good deal of scientific attention in recent years. The need for information of this kind was made urgent during the 1939 - 1945 World War when the United Nations were faced with the necessity of maintaining operational fronts in arctic waters and territories. The governments of the United States and Great Britain have been particularly active in this kind of research, and since the end of the war have continued an accelerated program of investigation. A great deal of this effort has been directed towards problems of a purely physiological nature, and the results of these, in addition to their value to medicine, have had an important bearing on the requirements of rations and on the design and construction of shelters, clothing, machines and equipment.[1]

Psychological problems have received a share of attention in these investigations also. Two broad areas of inquiry can be distinguished: that which is concerned with changes effected in the whole personality by continued exposure to severe cold; and temporary disturbances of a psychophysiological nature effected by relatively short periods of exposure. It is to this latter category, of course, to which the present study belongs.

---

[1] See Bibliography, articles by Gagge, Pierce and Rees, and Roth and Gabrielson.

It is difficult to evaluate the work which has been done on these problems, as the available literature on them is so scant. There are two reasons why this should be the case. In the first place, the bulk of recent cold research has been done under the direction of, or in cooperation with the Defense Departments of the countries concerned, and the resulting information is generally guarded to some degree for security reasons. The second reason is that many of these studies are of a long-range order and continue over a period of years, so that their results are slow in reaching the stage of publication. This fact was confirmed in recent conversation with the writer by Dr. N.H. Mackworth of the Cambridge University Laboratories, where the greater part of such investigation is currently being carried out for the British government.

An excellent survey of published results in this field (up to 1947) has been provided for us by Macdonald Critchley.[2] Discussing the effects of short exposures to cold, Critchley remarks that these have rarely been the object of psychometric study. A noteworthy exception which he cites are the important investigations which T. Bedford has conducted on the effect of lowered temperatures on efficiency, as measured by the rate at which subjects are able to assemble the links of a bicycle chain. In this case, extremes of cold were not included for study, and differences from room temperature were relatively small. Nevertheless, he found significant changes to occur in dexterity.

---

2 M. Critchley, "Effects of Climatic Extremes", Brit. J. Industr. Med., 1947, 4, p.164.

A drop in temperature from 75°F to 68°F for example, was accompanied by a 12% increase in time required for the task.

Armstrong and Grow have described disturbances of judgment and intellect which occur in airmen flying at high altitudes and which are ascribed to cold.

As a part of a larger study, Dr. Mackworth has recently[3] conducted field investigations at Ft. Churchill on the effect of short exposures to cold on tactile and kinesthetic functions. The results of these are not yet published.

Very nearly all of the psychometric studies of cold which have been undertaken to the present time are the product of the cold chamber. The fact that so little has been attempted in the field makes this investigation what is essentially a pilot study. The ground we are breaking has remained hitherto almost untouched. Lacking the precise facilities for control which the laboratory enjoys, we can scarcely expect definitive answers to the questions we pose of the same order which those controls would permit. The value of a field study lies, quite apart from the precision of its results, in the issues it opens, in the suggestions which its results make for future investigation, and in the new problems which it may chance to uncover in the practical situation. The cure for its crudeness is repetition and further investigation.

---

[3] January to March, 1949.

CHAPTER III

EXPERIMENTAL RESULTS

I. RAW DATA

i) Ft. Churchill Group

All the raw data obtained on the twenty-four subjects tested at Ft. Churchill is included in Appendix A. Here the raw scores appear in the order in which the corresponding sub-tests were administered. The order of conditions under which the tests were given is also contained in the legend. In addition to scores, data is included on the occurence and degree of fatigue during the test, and the subjective cold report; also a rating of each subject for acclimatization. (Weather is described in terms of temperature, wind velocity, and wind chill.) The data on fatigue, subjective cold and acclimatization are described, and the methods by which they were quantified are explained in the appropriate places in the text.

ii) University Group

Appendix B contains the same kind of data for the eighteen subjects tested at the University. In this case of course, cold reports do not apply, nor the allied measures of exposure time and degree of acclimatization.

iii) D.I.D. Group

Raw data for the D.I.D. test group is summarized in Appendix C. Scores are given for each of the two tests given each day, and also scores on the indoors tests which were administered under both bare-hand and mitted conditions after the completion of the main set of tests.

The graphs in Figs. 2 -4 (pp.22- 25) illustrate the distribution of scores obtained by the Ft. Churchill subjects. Figs. 5 - 9 are based on the comparable data for the University group.

Table I (below) is a summary of the ranges means and standard deviations from the corresponding graphs for Ft. Churchill. Table II contains the same data for the University group. Data from the retest scores is included with the original measures. Means and mean deviations for the D.I.D. data follow in Table III.

All correlational data will be presented in Part III of this Chapter.

TABLE I

SUMMARY OF DISTRIBUTIONS

FT. CHURCHILL

	Outside-mitt			Inside-mitt			Bare-hand		
	H	M	L	H	M	L	H	M	L
Range	140	165	83	154	137	131	44	38	16
Mean	112.7	143.1	102.9	99.5	126.2	82.6	58.5	48.4	26.7
S.D.	30.7	39.8	23.6	27.8	38.6	28.9	13.7	9.0	4.2

TABLE II

SUMMARY OF DISTRIBUTIONS

UNIVERSITY OF MANITOBA

		Outside-mitt			Inside-mitt			Bare-hand		
		H	M	L	H	M	L	H	M	L
Range	1.	79	117	71	68	107	67	61	31	13
	2.+				61	80	72	36	24	18
Mean	1.	76.3	96.2	55.7	73.4	94.4	56.4	51.8	40.3	22.6
	2.+				64.8	76.9	44.4	42.6	37.9	20.0
S.D.	1.	21.7	31.1	20.1	21.1	26.1	11.7	13.5	8.1	3.2
	2.+				17.3	19.0	17.1	9.7	6.9	4.8

+ Retest results

TABLE III

D.I.D. GROUP - MEAN SCORES AND MEAN DEVIATIONS

	Day 1		Day 2		Day 3		Day 4		Inside	
	1	2	3	4	5	6	7	8	mitt	bare
M	85.0	66.6	60.9	55.3	46.5	45.7	42.4	39.8	49.7	30.1
M.D.	17.8	11.4	15.1	11.4	7.1	11.0	3.8	4.2	10.2	4.8

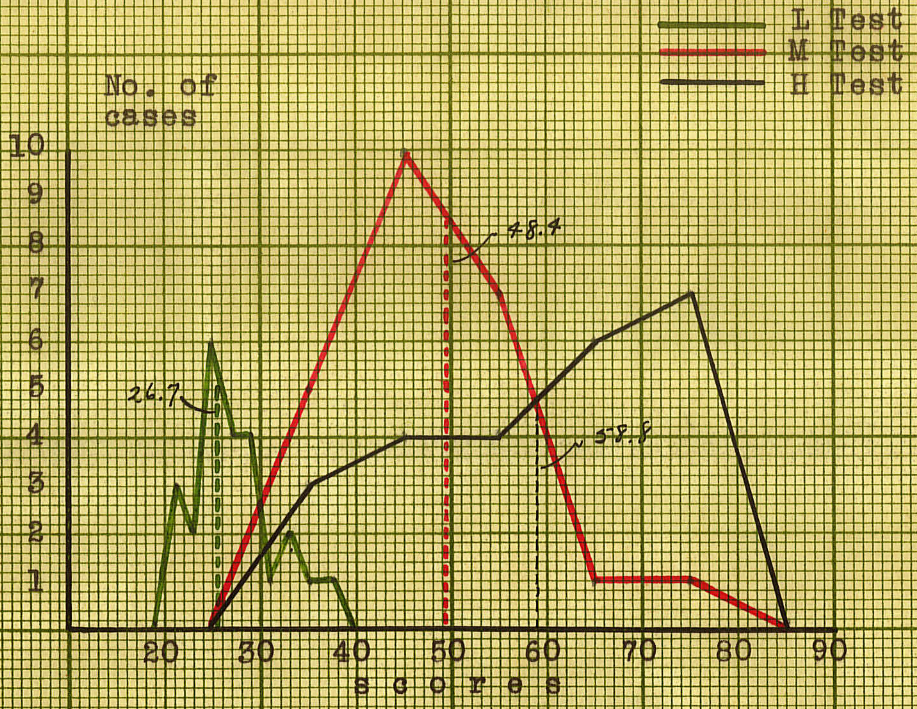

Fig. 2 - Distribution of scores on "Bare-hand" Test, Ft. Churchill

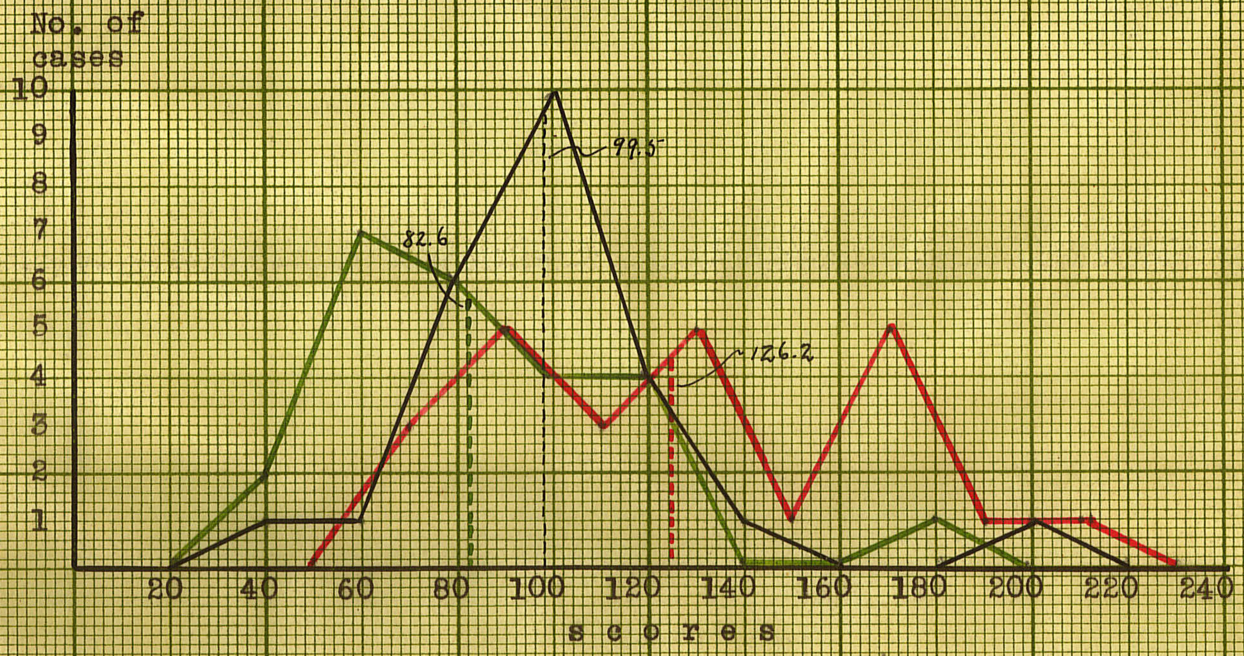

Fig. 3 - Distribution of scores on "Inside-mitt" Test, Ft. Churchill

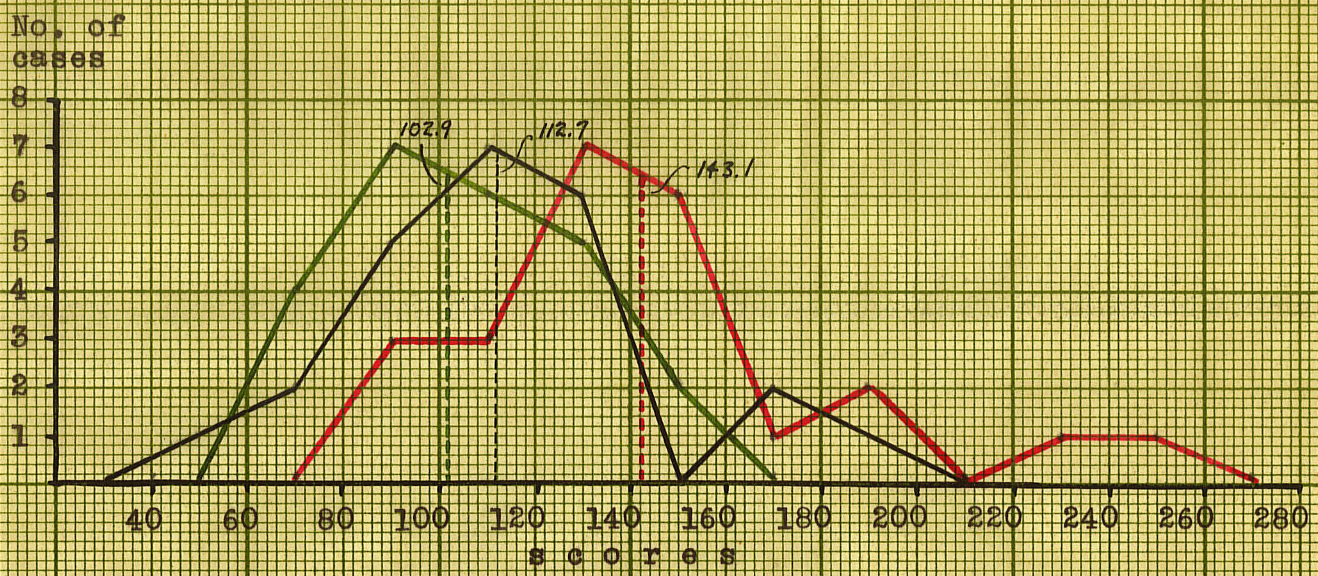

Fig. 4 - Distribution of scores on "Outside-mitt" Test, Ft. Churchill

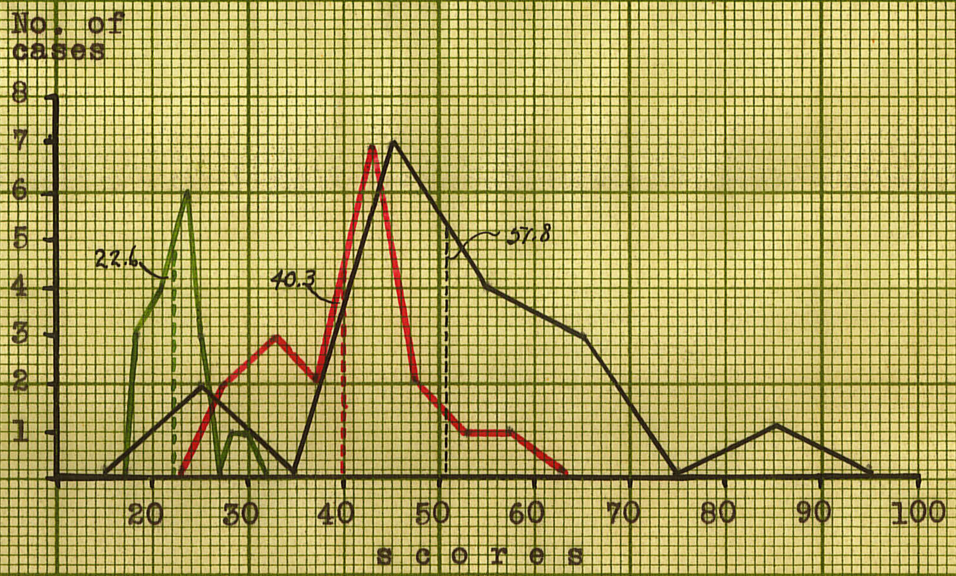

Fig. 5 - Distribution of scores on "Bare-hand" Test, University

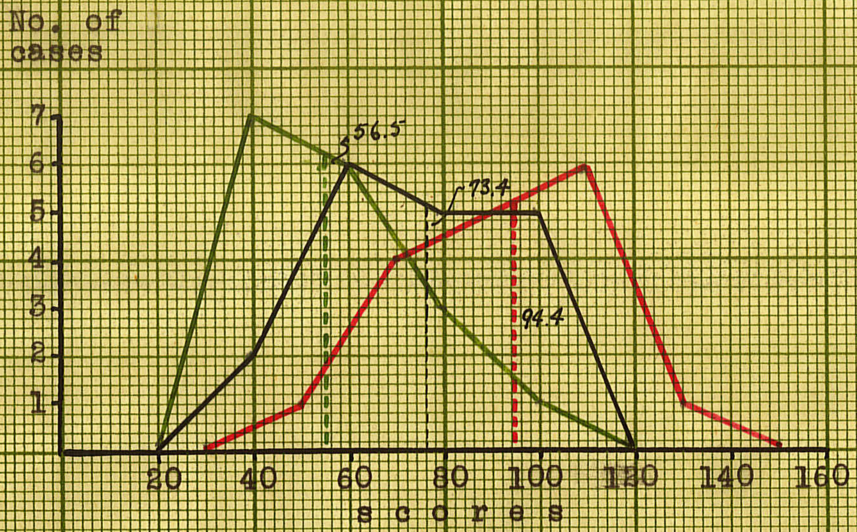

Fig. 6 - Distribution of scores on "Inside-mitt" Test, University

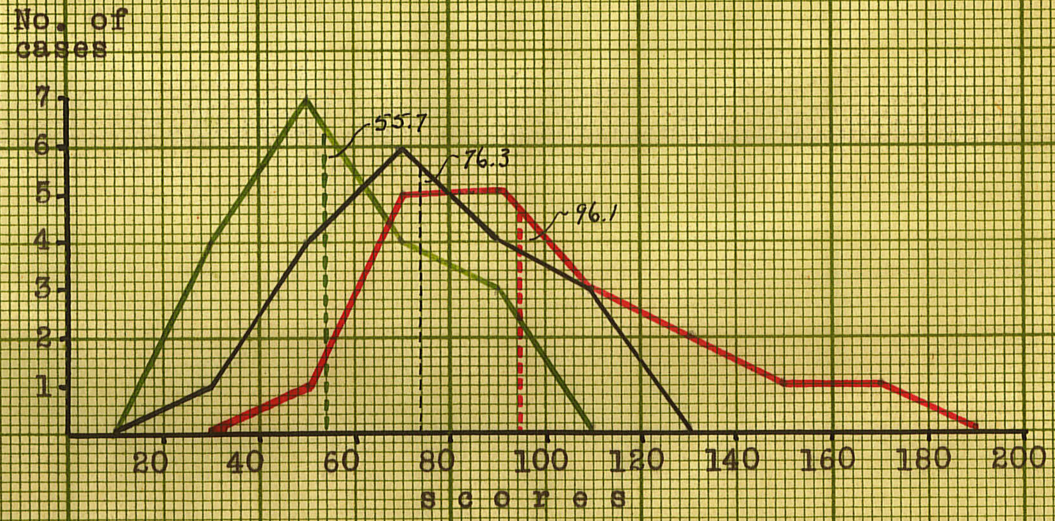

Fig. 7 - Distribution of scores on "Outside-mitt" Test, University

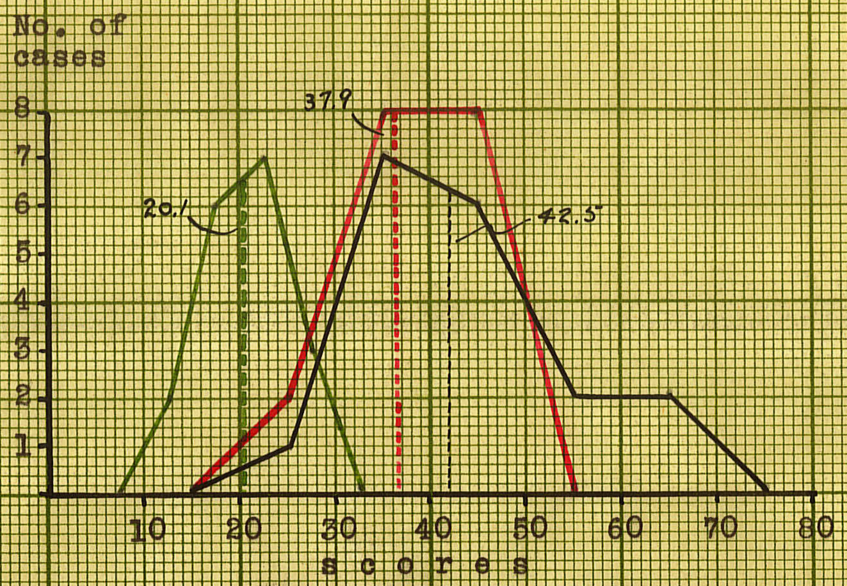

Fig. 8 - Distribution of scores on "Bare-hand" Test (Second Trial), University

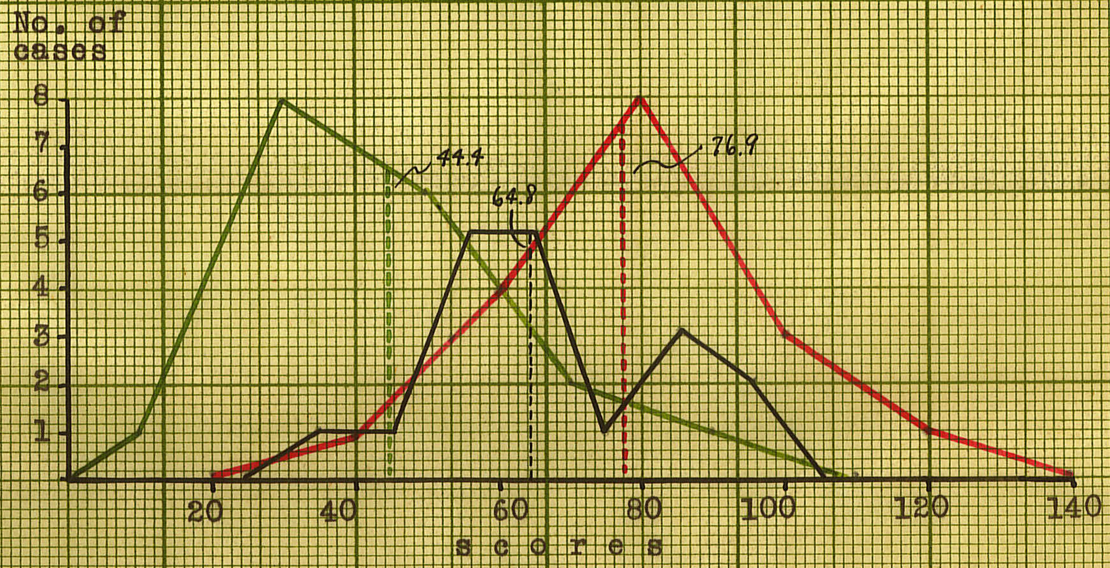

Fig. 9 - Distributions of scores on "Inside-mitt" Test (Second Trial), University

II. TREATMENT OF DATA ON THE BASIS OF GROUP DIFFERENCES

Group Differences in Test Scores Due to Outdoor Clothing. A measure of the effect on group scores of outdoor clothing, including the mitt, may be calculated on the basis of the difference in means between Bare-hand and Inside-mitt scores. The differences between the means which are observed in the Ft. Churchill data in the case of the H, M and L tasks are 41.04, 77.79 and 55.83 respectively (Table I). All are increases in the mitted as against the Bare-hand tasks. They are all substantial, representing increases in the Bare-hand mean scores of 70.1%, 160.7% and 208.6%.

The significance of the differences of the means was calculated with the following results[1]:

i) H Test

M_1(Mitt-inside)	= 99.54	SE M_1= 5.93
M_2(Bare-hand)	= 58.50	SE M_2= 2.92
Difference	= 41.04	SE D = 6.60

CR= 6.85, significant above .01 level of probability

---

[1] The SE D was calculated on the formula for uncorrelated means (SE D= $\sqrt{SE\,M_1^2 + SE\,M_2^2}$) in each of these three cases, instead of that for correlated means (SE D=$\sqrt{SE\,M_1^2 + SE\,M_2^2 - 2r_{12}\,SE\,M_1\,SE\,M_2}$) which, of course, these are. This would not have meaning fully changed the picture, however, since the differences already obtained are outside the limits of the tables published.

ii) M Test

M_1(Mitt-inside) = 126.20 SE M_1= 8.23

M_2(Bare-hand) = 48.41 SE M_2= 1.91

D = 77.79 SE D = 8.44

CR = 9.21, significant above .01 level.

iii) L Test

M_1(Mitt-inside) = 82.58 SE M_1= 6.15

M_2(Bare-hand) = 26.75 SE M_2= .88

CR = 9.16, significant above .01 level.

In all three cases, then, we have significant differences beyond any question of doubt: all are well within the .01 level of probability.

Group Differences in Test Scores Due to Exposure.

i) Ft. Churchill Group

The difference in mean scores effected by exposure to weather can be shown by the same procedure as was employed with clothing. Making use of the differences of the means obtained in the Mitt-inside and Mitt-outside conditions at Ft. Churchill, we have the following differences which are attributable to the effect of exposure (from Table I).

H - - 13.17, a mean increase over the inside score of 13.2 per cent.

M - - 16.87, a mean increase over the inside score of 13.3 per cent.

L - - 20.38, a mean increase over the inside score of 24.7 per cent.

The differences here are not as substantial in terms of per cent increase on inside scores as they were in the comparison of Inside-mitt and Bare-hand. In calculating their significance, the formula for correlated means was used.

i) H Test

M_1(Mitt-outside) = 112.71   SE M_1 = 6.49 r_{12} = .69

M_2(Mitt-inside) = 99.54   SE M_2 = 5.79

D = 16.9   SE D = 4.85

CR = 2.71, significant at less than .02 level.

ii) M Test

M_1(Mitt-outside) = 143.1   SE M_1 = 8.29 r_{12} = .64

M_2(Mitt-inside) = 126.2   SE M_2 = 8.05

D = 16.9   SE D = 8.83

CR = 1.80, significant at .10 level.

iii) L Test

M_1(Mitt-outside) = 102.9   SE M_1 = 4.91 r_{12} = .67

M_2(Mitt-inside) = 82.6   SE M_2 = 6.02

D = 20.3   SE D = 4.29

CR = 4.73, significant at .01 level.

The L test, then, shows significant difference in mean scores between the outside and inside trials to a very high level of probability. On the H test, the group is also affected by the weather conditions to a significant degree, though we cannot make this statement with the same degree of assurance as in the case of the L test. Taken by itself, the difference shown on the M test is not a significant one, but its direction reaffirms the two significant results.

ii) University Group

This group provides a control for the exposure factor. Only very small differences obtained between inside and outside situations in terms of "weather" conditions, consequently no significant differences between means should be expected. Below are means and standard errors for inside and outside scores for each of the three tests.

i) H Test

M_1(Outside-Mitt) - 76.32 SE M_1 - 21.70

M_2(Inside-mitt) - 73.40 SE M_2 - 21.07

ii) M Test

M_1(Outside-mitt) - 96.16 SE M_1 - 31.09

M_2(Inside-mitt) - 94.44 SE M_2 - 26.06

iii) L Test

M_1(Outside-mitt) - 55.73 SE M_1 - 20.10

M_2(Inside-mitt) - 56.47 SE M_2 - 11.70

Differences here are so obviously insignificant that a formal computation seems superfluous, except possibly in the case of H, where r between Inside-mitt and Outside-mitt is very high. Here the significance was computed:

D - 2.92 SE D - 2.14

CR - 1.36 P - .30

proving this difference insignificant also.

The Relation of Sunjective Cold Report to Performance

An American Army report has suggested that behavior observed in the "Fox-Hole" Study has more relation to subjective reports of cold than to objective measures such as skin temperature. We shall determine whether such is the case in the nut and bolt test.

i) Ft. Churchill Group

The subjects were divided into three groups according to their reports of cold. The "cold" group included subjects who reported cold to the extent of a total of three points or more on the rating scale in the two ratings given. The "Slightly cold" group reported cold to the extent of a total of two points, and the "not cold" group reported no appreciable cold; a zero rating. (See rating scale in App.A.)

The differences between mean scores made on the Outside-mitt and Inside-mitt tests for the three groups are shown below. Each difference is also shown as a per cent increase on the Inside-mitt mean score.

"Cold" (7 subjects)	H	M	L
Difference Out/In	12.2	27.0	20.0
% of mean In score	14.1	19.0	24.6

"Slightly Cold" (9 subjects)	H	M	L
Difference Out/In	17.9	20.9	15.9
% of mean In score	20.0	17.6	17.6

"Not Cold" (8 subjects)	H	M	L
Difference Out/In	8.6	20.0	29.8
% of mean In score	7.8	16.5	21.1

On the M test, there appears a direct relationship between subjective report and performance. The "cold" subjects deteriorate in the outside test to the extent of 19.0% of their mean inside score. The "slightly cold" deteriorate by only 17.6%, and "not cold" by 16.5%. In the two other tests, however, this relation does not obtain. On the H test, the "cold" deteriorate less than the "slightly cold", although both show a much more severe effect than the "not cold". On the L test, the performance of the "slightly cold" suffers less than that of the "cold", but also less than that of the "not cold".

Although some relationship between performance and subjective cold is indicated, it is not borne out with sufficient consistency in these tests to enable us to make a conclusive statement.

ii) D.I.D. Group

It must be pointed out here that the cold reports taken on the D.I.D. Subjects received a different treatment than those taken on the Ft. Churchill group. In the latter case, ratings were ascribed,on an absolute scale. In the case of the D.I.D. group, on the other hand, the ratings represent a comparison between the subject's feelings at one time with his feelings at another, with no reference to an absolute judgment. An equivalent rating might have been given the Ft. Churchill subjects by putting the question in this form: "Were you warmer or colder on the inside test as compared to the outside test?".

While it is not possible to group the D.I.D. subjects according to the degree of cold which they experienced, it is still possible to compare the differences between their two scores on any day with their comparative feelings on the second of the two trials. Test scores were divided into five groups, according to whether subjects reported "much warmer" (++), "a little warmer" (+), "slightly colder" (-), "much colder" (--), or "about the same" (0) on the second trial of the day as compared to the first. If there is any relation between these reports of cold and the subjects' performance, it should be revealed in the differences between the mean scores obtained on the two tests.

Differences in mean scores on the two trials are shown below, together with the number of subjects in each group.

	++	+	0	-	--
D (two trials)	-10.0	-7.5	-9.1	+8.0	+6.0
N	5	10	17	2	1

The differences in mean scores are expressed positively and negatively to indicate the direction of the change. A positive gain means an increase in score on the second trial; a negative gain means a decrease.

The general direction of the results indicates some corrsspondence between subjective feelings of cold and performanece. Those who felt colder on the second trial made poorer scores than they did initially. Those who felt warmer improved their initial scores. The over-all picture is not completely consistent, however. Those subjects who reported "much colder" did not increase in score as much as those who reported "slightly colder". Also, there is a substantial improvement in the scores of those subjects who experienced no change in their subjective state. The fact that practice operates to reduce scores may mask the effect of subjective cold, since practice operates to increase those differences which are evidenced by the "warmer", and to decrease the differences evidenced by the "colder" subjects.

<u>The Effect of Acclimatization on Subjects' Differences Between Inside and Outside Scores</u>. It is commonly supposed that individuals who have been acclimatized or conditioned to a severe climate suffer less from its psychological and physiological effects than those unused to such extremes. We shall see whether there are significant differences between the outside/inside changes in score for the "acclimatized" subjects in the Ft. Churchill group as compared with those of the relatively unacclimatized. The group was divided according to ratings based on the length of time which the subject had resided in the Churchill area immediately prior to testing, and the amount of time he habitually spent out of doors. The resulting classification is strictly one of degree. The "more acclimatized" are those who on the average spent more than one hour per day outdoors, and had been residing in Ft. Churchill for more than one month.. The ensuing division resulted in fourteen subjects

being assigned to the "more acclimatized" group, and ten to the "less acclimatized".

Mean scores obtained by each group on the three sub-tests under the Inside-mitt and Outside conditions are reproduced below, with the difference which obtain in each case, and the percentage of the inside score which each difference represents.

"More acclimatized"

	H	M	L
Out	109	132	95
In	103	119	80
d	6	13	15
% Incr.	5.8	10.9	18.7

"Less acclimatized"

	H	M	L
Out	117	157	115
In	95	136	95
d	22	21	20
% Incr.	23.1	15.4	21.0

The significance of the differences between the differences (d's above) was calculated. The differences which occur in the H test are the ones most nearly approaching:[2]

---

2 Garrett's method for small independent samples might have been used here with justification. This would raise P to about .05. See H. E. Garrett, Statistics in Psychology and Education (New York; Longmans, Green & Co., 1947) p. 204-5.

H Test

M_1 - 6.4 SE M_1 - 5.46

M_2 - 22.6 SE M_2 - 7.81

d - 15.6 SE D - 9.52

CR - 1.69

At 22 degrees freedom, P - .10

Although we are not justified in placing confidence in the difference of the differences between the two groups, (the fact that the same tendency shows up in all the tests) increases the chance that such differences are real.

The Relation of Objective Cold to Performance.

The effect of objective cold was investigated by the same technique as was employed with subjective cold. The Ft. Churchill group was divided in two on the basis of the temperatures which obtained during the subjects' outdoor tests. Those tested at -11°F and below constitute the "colder conditions" group, and those tested at -10°F and above constitute the "milder conditions" group. The first group contains thirteen subjects, the second eleven. The unequal division resulted from tied conditions at the critical temperature. The particular figure of -11°F was chosen simply to give approximately equal groups.

Below are shown the differences between mean scores on the inside and outside tests for the two groups; also the per cent deterioration which each difference represents, on the basis of the inside scores.

"Colder conditions" (13 subjects)

	H	M	L
Difference Out/In	2.7	10.0	8.7
% of Mean in score	2.5	7.6	9.7

"Milder conditions" (11 subjects)

	H	M	L
Difference Out/In	25.4	25.0	26.0
% of Mean In score	27.0	20.8	34.9

In each case, the subjects tested under the colder conditions showed less deterioration outside than did those tested under the milder conditions. The difference between the differences for the two groups on the L test was tested for significance. This is the difference which taken in conjunction with the S.D.'s of the distributions appeared most likely to be significant.

L Test

M_1(d's of "cold" group) - 8.7

M_2(d's of "mild" group) - 26.0

$S.D._{12}$ -23.0[3]

CR - 1.83 P - .10

We conclude that the differences are not significant.

3 By Garrett's method for small independent samples.

The Relation of Subjects' Proficiency in the Tests to Changes in Performance Due to Exposure.

i) Ft. Churchill

Is deterioration of performance on these tasks related to the subject's proficiency in them? Since the tasks are not highly related, different subjects may be proficient in different tasks, therefore, subjects were divided equally into "proficient" and "non-proficient" groups on the basis of their inside scores as judged by three criteria: their inside scores on the H, M, and L tests. Differences between mean scores on the inside and outside trials were then computed for the "proficient" and "non-profient" groups.

Judged by the criterion of inside scores on H, the mean differences in scores for the two groups are presented below.

Criterion of Proficiency - H (critical score 95)

(Proficient Subjects -No's. 1,2,3,5,6,8,9,12,14,17,18,24)

		Prof.	Non-Prof.
Mean d Out - in	H	17.83	8.50
	M	34.50	1.66
	L	21.41	20.16

Criterion of Proficiency - M (critical score 125)

(Proficient subjects- No's. 2,4,5,6,9,12,13,14,15,17,18,22.)

		Prof.	Non-Prof.
	H	13.00	13.33
Mean d Out-In	M	28.58	5.58
	L	26.41	15.16

Criterion of Proficiency - L (critical score 76)

(Proficient subjects - No's.3,4,5,6,7,9,12,13,14,17,21,22)[4]

		Prof.	Non-Prof.
Mean d Out-In	H	8.91	17.41
	M	18.41	15.75
	L	31.41	10.16

Summary of Differences

Prof. - Non-Prof.	Criterion of Prof. H	M	L
H	9.33	- .33	- 8.50
M	30.83	23.00	2.66
L	1.25	11.25	21.25

With only two exceptions, viz. the differences between proficient and non-proficient in the H test when the criteria of proficiency are the inside M and L tests respectively, the fact is that the proficient subjects deteriorate to a greater degree than do the less proficient. It will be noted from the summary of differences above, that when the criterion of proficiency is score on the L test, the greatest deterioration is also on that test. Similarly in the case of the M test. The only exception is in the H test. When the criterion of proficiency is the H test scores, the greatest deterioration is exhibited on the M test.

4 Membership in the proficient group appears to be a matter of chance. For example, those proficient in H are not necessarily proficient in L. The proficient groups are significantly related in only one instance, viz. between M and L, where Chi-square = 6.

ii) D.I.D.

A similar enquiry of the effect of a stress situation on proficiency was conducted on the results of the D.I.D. testing. In this case, proficiency was judged on the basis of scores made in the non-stress situation.

We have previously indicated that these subjects were given two trials on each testing day. The first, or morning trial, was administered when the subjects were first out of their tents, and were comparatively stiff and cold. The second trial was called the "non-stress", since immediately before it the subjects had been engaged in vigorous physical exercise, and were comparatively warm and comfortable. The subjects were classified as proficient or non-proficient on the basis of their performance in the non-stress situation which occurred in the second trial on each of three days. Differences between the stress and non-stress scores are based on the average differences shown on three days of testing.

i) Proficiency based on "Non-Stress" Trial, Day 1
(Proficient Subjects, Nos. 4,5,6,7)

Prof. group (M diff. in scores)	9.1
Non-Prof. group	7.4

ii) Proficiency based on "Non-Stress" Trial, Day 2
(Proficient subjects No's. 3,4,5,9)

Prof. group (M diff. in scores)	10.4
Non-Prof. group	6.4

iii) Proficiency based on "Non-Stress" Trial, Day 3
(Proficient Subjects, No's. 3,4,5,9)
(Read in Day 2 grouping, as subjects
were the same)

The above results contain practice effects, and the possibility remains that the proficient group obtained the most benefit from practice. As a check on this, the practice effect was removed in the following way. The difference between the stress and non-stress scores is attributable to practice and the stress effect. If we compute the differences shown on successive trials between stress and non-stress situations, and compare the differences with those obtained in going from a non-stress to a stress situation, we can remove practice in much the same way as we did with the Ft. Churchill data. We have defined the stress trial as the first one of each day. In the case of any given subject, the transition from the second trial on one day to the first trial on the following day represents a change in the non-stress to stress direction. The second test of the pair will, in effect, be increased (i.e., made poorer) by the stress factor and reduced (improved) by practice. We can represent the difference between the two scores then as $y - x$, where y is practice and x i stress.

The difference between subjects' scores on the first trial of each day as compared with their scores on the second represents a change from stress to non-stress. This difference can be symbolized as $y + x$ where practice and stress operate in the same direction, toward a reduction of the subject's first score.

Subjects were first divided into proficient and non-proficient classes on the basis of their total scores on three non-stress trials (i.e., 2,4, and 6 in Table III). The difference in means between their second and third trials, and their fourth and fifth trials were computed and equated to y - x. The difference in means between trials 1 and 2, and between 3 and 4 were equated to y + x. The equations were then solved as before, for both the proficient and non-proficient subjects. The results are shown below.

Proficient (Subjects 1,4,5, and 9)

Practice	-	11.67
Stress	-	5.67

Non-Proficient (Subjects 2,3,6,7 and 8)

Practice	-	9.37
Stress	-	1.37

The findings here confirm those on the Ft. Churchill group; the proficient suffer more from the introduction of stress than do those who are less proficient.

Practice Effects Inherent in the Tests. One reason for selecting a simple motor task was to rule out practice effects. In addition, the rotation method of presenting conditions was employed with the University and Ft. Churchill subjects for the purpose of balancing out what practice effects did exist so that the mean scores would be directly comparable on the basis of the variables of clothing and exposure. While differences between mean scores seem to be well protected against the variables of practice (by the nature of the test and by rotation), nevertheless they may still affect correlations, and consequently had to be computed.

i) Ft. Churchill Data

Let us consider again the rotation scheme which was applied to the order of the presentation of conditions.

		order of conditions		
	1.	A	B	C
subjects	2.	B	C	A
	3.	C	A	B

The difference in scores between the Inside-Mitt and Bare-hand conditions (B and C), is the result of two factors: a) the effect of the mitt, and b) the effect of learning. The effect of the mitt is constant. The effect of the learning factor is varied in three ways. In the case of the first subject, practice increases the difference to the extent of one repetition; in the case of the second subject also it increases the difference to the extent of one repetition; in the case of the third subject it decreases the difference to the extent of two repetitions.

Formalizing the above argument we can show the difference between scores in B and C for the first and second cases = $x + y$. The average difference for case three = $x - 2y$.

We now have two solvable equations:

$$x - y = d_1$$

$$x - 2y = d_2$$

$$3y = d_1 - d_2$$

$$y = \frac{d_1 - d_2}{3}$$

$$x = d_1 - y$$

The obtained values for x and y are the mean differences in the scores attributable to the mitt and to practice respectively.

This calculation was performed on the difference between the Inside-mitt and Bare-hand scores for each of the three subtests with the following results.

<u>H Test:</u>

Practice = 6.32 (10.8% of Bare-hand score)

Mitt = 41.04 (70.0% of Bare-hand score)

<u>M Test:</u>

Practice = 7.71 (16.0% of Bare-hand score)

Mitt = 77.79 (160.7% of Bare-hand score)

<u>L Test:</u>

Practice = 4.79 (17.0% of Bare-hand score)

Mitt = 55.83 (208.7% of Bare-hand score)

The effect which the mitt exerts on performance is clearly much greater than that of practice, as is seen in a comparison of the difference which each contributes in terms of percent increase in the mean bare-hand scores for each of the subtests. The mitt has its greatest impeding effect on the L test, and least on the H, as may be expected. Practice, on the other hand, betters performance on the L test more than it does on the M, and to a greater extent on the M than on the H. It seems reasonable that such should be the case. Since the mitt presents more of a problem on light nuts, there is more opportunity there for learning a better technique.

The same method as was described above was used to draw out the effects on mean scores of practice and exposure between tests under conditions A and B.

Results of the calculations for each of the three subtests are shown below.

H Test: Practice = 3.39 (3.4% of inside score)
Exposure = 13.17 (13.2% of inside score)
M Test: Practice = 3.17 (2.5% of inside score)
Exposure = 17.08 (13.5% of inside score)
L Test: Practice = .10 (.1% of inside score)
Exposure = 20.79 (25.1% of inside score)

There is a considerable discrepancy between the values obtained for the effect of practice in this case, and that which we found to apply between conditions B and C. Apparently practice is greater between B and C than it is between B and A. This effect might be explainable on the basis of the relative pleasantness or unpleasantness of the change involved in going from A to B or from the B to C tasks. The mitted task is a trying one from the subject's point of view, and he normally goes from it to the bare-hand task with considerable alacrity.

The bare-hand task is much less fatiguing, on the basis of the subjects' reports, and it comes almost as a reward. When we are concerned with the B - C and C - B transitions, as we were in our first calculations for the effect of practice, a "pleasant" transition occurs in two cases out of three (i.e., C follows B in two of the three possible orders). When we are concerned with A and B, as we were in our second set of calculations, no such "pleasant" combination occurs. If our assumption is correct that such "pleasant" transitions enhance the learning effect, or alter motivation to give that appearance, we have accounted at least in part for the discrepancy.

ii) University Data

Since exposure, as judged by differences in mean scores, has no significant effect on performance under the summer conditions which obtained at the University, we can ignore it as a factor. The three trials with the mitt (one Outside-mitt and two Indoors-mitt, test and retest) can, then, be considered simply as first, second, and third trials, disregarding the in-out difference. To get an uncomplicated picture of the effect of learning, we must also disregard what practice may be contributed by the one Bare-hand trial which is in each case interposed between two of the mitted trials.

The mean scores obtained on the three trials with the mitt are given below for each of the three subtests.

		H	M	L
	1.	76.61	98.76	59.41
Trials	2.	72.48	86.24	52.27
	3.	64.83	71.33	44.44

In the case of the H Test, one repetition reduces the initial score 5.4%, and another repetition reduces the new score 10.5%. In the M test, learning proceeds at the rate of 12.7% in the first period, and 17.3% in the second; and in the L test at the rate of 12.0% in the first practice period, and 15.0% in the second. Three trials is scarcely enough on which to base any prediction on the form of the learning curve. It appears, however, that initially at least the curve is positively accelerated. The fact that the drop in score is more pronounced between the second and third trials than it is between the first and second, may bear a relation to the time interval between the tests. The first two trials were necessarily given to the subject during the one test period; that is, they are part of the same series of three trials under A, B and C. The third trial is always a retest, administered an hour or two later, and in two cases out of three it is the first measure of the retest period, preceding the Bare-hand trial. This being the case, the third measure frequently suffers less from the effect of fatigue than does the second, which tends to augment the apparent learning effect.

In any event, there is no question that practice has a marked effect on performance on these tests, and is evidenced on the M and L tests to a greater degree than it is on the H. The two lighter tasks are the more difficult, and it is on these that the subject is most radically affected by the mitt. It seems reasonable to expect him to improve to a greater degree on them with successive trials as he becomes used to them and develops new ways of coping with them.

iii) D.I.D. Data

In the repeated trials given the D.I.D. subjects on the H test, the variable of altered physical stress was the only one introduced into the pure learning picture. Since the effects of this stress were very slight, we shall disregard them, and consider the changes in mean scores which are otherwise attributable to practice.

The following summary shows the mean scores for the group on each trial, and their mean deviations, (repeated from Table III).

	Trials							
	1	2	3	4	5	6	7	8
M.	85.0	66.6	60.9	55.4	46.5	45.7	42.4	39.8
M.D.	17.8	11.4	15.1	11.4	7.1	11.0	3.8	4.2

It will be noted in Appendix C that the fifth and sixth trials include only nine of the eleven subjects, and the seventh and eighth only five. Mean scores and mean deviations are plotted in the following graph.

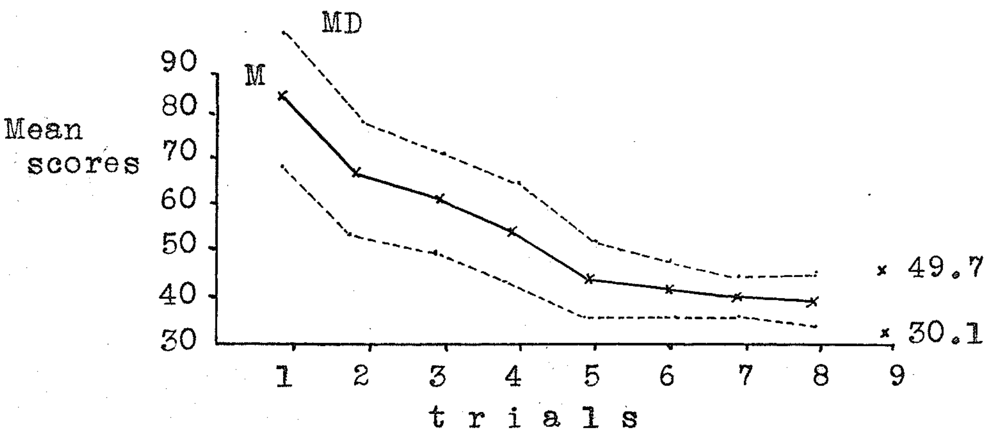

Fig. 10 - Means and Mean Deviations, D.I.D.

Mean scores on the inside tests with mitt and with bare-hand, which were administered only once at the end of the experiment are also plotted on the graph.

The practice effect is well defined, and follows a typical learning curve pattern, negatively accelerated. Mean deviations also tend to follow the same form: practice reduces the group's variability of performance.

The Inside-mitt test which followed the other trials after an interval of five days shows a return to about the fourth or fifth trial's level, indicating an apparent loss in performance over the interval. The Bare-hand score gives an indication of how closely the outdoor trials approached this level of performance in seven practice periods.

III TREATMENT OF DATA FOR THE PURPOSE OF DETERMINING DIFFERENTIAL EFFECTS OF VARIOUS FACTORS

Description of Data. Table IV contains all the correlation coefficients (Pearson r's) calculated between scores on the same test, including

(a) self-correlations on test-retest carried out under the same conditions,[5]

(b) correlations for scores on the same tests under different conditions

1. Bare-hand vs. Inside-mitt
2. Bare-hand vs. Outside-mitt
3. Inside-mitt vs. Outside-mitt

Values derived from the retest correlations are the reliabilities of these tests. With them we shall be able to evaluate the other correlations: e.g. the H test yields a very high reliability and therefore, when the correlation calculated on this test under different conditions differ, the difference may be ascribed to the effect of the conditions with a considerable degree of assurance. The reliabilities of the M and L tests are not nearly so high, and this is reflected in the inconsistent picture which the data on these tests yield. Where reliabilities are so low, differences in correlations between test scores obtained under different conditions may be due either to the conditions or the unreliability of the tests.

---

[5] On the University subjects only.

TABLE IV

CORRELATION DATA

Conditions	Test Combinations HH	MM	LL
Bare-hand vs. Bare hand	.84	.55	.74
Bare-hand vs. Mitt-inside	.76	.17	.45
	(.44)	(.31)	(.28)
Bare-hand vs. Outside-mitt	.70	.34	.35
	(.22)	(.06	(.07
Inside-mitt vs. Inside-mitt	.95	.15	.70
Inside-mitt vs.Outside-mitt	.93	.34	.64
	.93+	.10+	.69+
	(.69)	(.64)	(.67)

Bracketed figures are Ft. Churchill data.

+Outside-mitt vs. Inside-mitt 2nd trial.

Values derived from correlations described under (b) make it possible to ascertain whether or not the introduction of a new factor affects all subjects in the same way, or whether it affects them in different ways. For example, as we have shown in the previous section, the introduction of the mitt causes a drop in efficiency as is measured by group means. The question remains whether this change affects the subjects equally, or whether the effect is differential to the extent of significantly altering the relative position of each man in the group.

A similar question may be put with regard to exposure and its effects.

Table V contains correlations between the subtests H, M an L. A great number of combinations are possible here, viz.

(a) those which occur under the same conditions (Bare, Inside-mitt, Outside-mitt)

(b) those which occur under different conditions

1. Bare hand vs. Inside-mitt
2. Bare hand vs. Outside-mitt
3. Inside-mitt vs. Outside-mitt

Correlations under Condition (a)

Those combinations which occur under the same conditions were all calculated for both the Ft. Churchill and University groups, using the first trial in each case on the University group.where two trials were administered.

TABLE V

CORRELATION DATA

Conditions	Test Combinations		
	HM	HL	ML
Bare-hand vs. Bare-hand	.86 (.65)	.51 (.69)	.59 (.77)
Inside-mitt vs. Inside-mitt	.66 (.40)	.77 (.65)	.61 (.74)
Inside-mitt vs. Outside-mitt (H out - M in H out - L in M out - L in)	.61 (.43)	.67 (.32)	.65 (.73)

The values derived here indicate to what extent the tests are alike (in that they test the same thing), and to what extent consistent results may be expected between the three subtests. Obviously, however, these figures are not to be interpreted independently of their reliabilities. For example, the correlation of .86 between H and M on the Bare-hand task is slightly higher than the reliability coefficient for the H and significantly higher than the reliability coefficient for the M test. Under these conditions we cannot place a great deal of confidence in the .86.

It may be well to note here, and this observation applies to the whole treatment of correlations, that reliability coefficients may be high for two tests, and yet their intercorrelations may be low. There is no reason, e.g., that the H test results with the bare-hand and those with the mitt should inter-correlate significantly even though the H test has proved highly reliable under both of these conditions.

Correlations under Condition (b)

The first two combinations under (b) were omitted. There are a great number of possible combinations contained in these two. It was felt that no advantage would be gained in calculating the relations in this chain of combinations, since it contains so many variables that chance would play a predominant role.

The inter-correlations were calculated between Inside-mitt and Outside-mitt conditions because the differences here (between Ft. Churchill and University groups), if significant, would corroborate the findings on the effects of exposure which were given by the first set of figures in Table V. For example, H in the outside condition and M in the inside condition (University group) $r = .61$. The same combination at Ft. Churchill yields an r of .43. If we can show a significant difference between these two, that difference can be attributed to the effect of the rigorous conditions, to which the Ft. Churchill subjects were exposed in their outdoor tests.

With these general observations we shall turn now to a consideration of the results.

Reliabilities. Under both conditions (Bare-hand and Inside-mitt) in which retests were conducted, the H task is clearly the most reliable of the three, correlating with itself .84 with the bare hand, and .95 with the mitt. It is a highly reliable test, and the addition of the mitt apparently increases its reliability substantially. That such is the case has been borne out to some extent by our observations of the actual performance of the test subjects. The nuts on the H series bolts are heavy, and they are quite easy-running, facts which make it a temptation for the subject to allow the nut to "coast" between turns. This lack of control, however, which undoubtedly contributes to some unreliability in the bare hand condition, seems to be checked by the mitt. In this case the soft, bulky palm of the mitt tends to arrest the nut at the end of the twist: it is not as easily or quickly taken clear of the nut as the bare fingers are. These facts may account, in part at least, for the increased reliability which the mitt contributes to the test.

The light task is also a fairly reliable one, though not to as great a degree as the heavy. It correlates with itself to .74 with the bare hand, and drops slightly to .70 with the mitt.

The M bolts constitute the least reliable of the subtests, having the low self-correlation of .55 in the bare hand condition and the almost negligible correlation of .15 with the mitt. The essential unreliability of the M task was suggested to us in several ways. In the first place it was observed from our notes on fatigue which were kept in some detail on the original record sheets, that the concensus indicates the M subtest to be the most fatiguing of the three, in the bare-hand trial. This seems to be an effect of using a purely finger manipulation on the same order as that employed with the L nuts. The M nuts are almost too large for this kind of manipulation, but not large enough to use the method required by the H nuts, viz. a "whole-hand" turning action. Another disturbing effect which applies to the mitted task, concerns the fact that the M bolts were threaded right down to the level of the surface of the frame in which they were mounted. This was also the case with the H bolts of course, but here it presented no special difficulty because the H nuts are thick enough to allow an easy grip even when the nut is down flush to the frame. The M nuts on the other hand are only about one half as thick as the H, and the last few turns are particularly difficult when the mitt is worn. The bulk at the tip of the mitt makes it difficult to get the fingers close enough to the frame to grip the nut in the usual manner. It was frequently found necessary to modify the grip by turning the hand over in a forward direction, the backs of the knuckles towards the frame, or by retracting the hand so that the palm is parallel to the frame, and applying the turn with the thumb and lateral side of the index finger only. Some subjects adjusted easily to this difficulty, some slowly, some not at all.

The same difficulty was obviated in the light task by the fact that the L bolts were threaded only to within 5/8" of the surface of the frame.

Differential Effects. We have already shown on the basis of group results, that changing conditions led to significant changes in test performance. What we must now determine is whether or not these changes affect all the subjects more or less equally or whether there is a real difference in the way that individuals react to those conditions. In testing for these differential effects, the general method adopted here was to calculate the difference between two observed correlations by R.A. Fisher's method.[6]

---

[6] R.A. Fisher, Statistical Methods for Research Workers, (London: Oliver and Boyd, 1938), p. 208.

a) Differential Effects of Mitt

On the basis of mean differences, the effect of the mitt is to increase time scores on the tasks. Are all subjects equally affected? If we obtain correlations 1) by test retest under condition A, and 2) by use of the same tests under conditions A and B, then if the two r's are significantly different, the difference may be attributed to the difference between conditions A and B.

The difference between r's for Bare-hand vs. Bare-hand and Bare-hand vs. Inside-mitt were tested for significance by Fisher's method with the following results (University group:

i) H Test

Bare-hand vs. Bare-hand -- .84
Bare-hand vs. Inside-mitt -- .76
n' - 18

	r	z	N - 3	recip.
Bare-hand vs. Bare-hand	.84	1.2221	15	.0666
Bare-hand vs. Inside-mitt	.76	.998	15	.0666
		.233 ± .365		

CR - .6109 P - .54

P at .54 indicates that differences would exceed this one by chance 54 per cent of the time. They cannot therefore be regarded as coming from differently correlated populations.

ii) M Test

Bare vs. Bare	.55
Bare vs. In.-mitt	.17
CR - 1.334 P -	.18

iii) L Test

Bare vs. Bare	.74
Bare vs. In.-mitt	.45
CR - 1.275 P -	.205

We must conclude that, as compared to the bare hand task, there is no significant differential effect created by the mitt.

The correlation between Bare-hand and Inside-mitt may also be compared to the reliability coefficient of the mitted task. A glance at Table V will show that there is no significant difference in the case of the M and L tests. The difference does appear to be significant in the H test, and this is proven by our calculation:

Inside-mitt vs. Inside-mitt	.95
Bare-hand vs. Inside-mitt	.76

Cr - 2.289 P - .02

Compared with the highly reliable mitted task, the change to bare hand has a significant differential effect.

(b) Differential Effect of Exposure

It was not possible to use the same procedure as above to test the significance of the differential effect of exposure, since reliability coefficients are lacking for the

Ft. Churchill group. We may, however, use the same basic statistical technique employing the University group as a control. At the University tests were conducted under conditions inside and outside which were very much the same, while for the Ft. Churchill group these conditions were markedly different. If the correlation between Inside-mitt and Outside-mitt for the University group is, significantly higher than the corresponding figure for Ft. Churchill, then we can assume that the exposure affected the Ft. Churchill group differentially. The calculation involving the H test is shown below.

<u>H Test</u>

Inside-mitt vs. Outside-mitt (U.) .93

Inside-mitt vs. Outside-mitt (C.) .69

CR - 2.426 P - .015

The difference is significant to a high degree of probability. With the M and L tests, however, a similar treatment reveals no significant differences.

<u>M Test</u>

Inside-mitt vs. Outside-mitt (U.) .44

Inside-mitt vs. Outside-mitt (C.) .64

CR - .828 P - .41 .

<u>L Test</u>

Inside-mitt vs. Outside-mitt (U.) .64

Inside-mitt vs. Outside-mitt (C.) .67

CR - .017 P - .98

The results on the M and L tests are not surprising considering the initial unreliability which these tests showed.

A second line of argument, based on a series of relations each of which are open to chance variations, yields some corrobatory evidence. For the University group, the Bare-hand vs. Inside-mitt correlation is .76. The same correlation for the Ft. Churchill group is .44. The test conditions were equivalent for the two groups, and we should expect to find that they are not samples from differently correlated populations. Such indeed is the case:

Bare-hand vs. Inside-mitt (U.) .76 N - 18

Bare-hand vs. Inside-mitt (C.) .44 N - 24

CR - 1.556 P - .12

and the coefficients cannot be regarded as coming from differently correlated populations.

If we now consider the correlations between Bare-hand and Outside-mitt, the conditions are different in one respect for the two groups. In the Ft. Churchill situation outside conditions were rigorous; in the University situation they were mild. If a significant difference is found here we may assume they are from differently correlated populations, and that the reason for this difference is to be found in the weather conditions, since we have shown that where conditions are the same they are not from differently correlated populations.

Bare-hand vs. Outside-mitt (U.) .70

Bare-hand vs. Outside-mitt (C.) .22

CR - 1.906 P - .054

This may be counted as barely significant, but in the light of the low P (.12) shown above, the conclusion is something less than decisive. The whole argument rests upon the interrelations between six sets of data under three sets of conditions, and so chance variance has an opportunity to play an important role. The trend of the evidence here, however, corroborates the more conclusive results shown in the first part of the discussion.

Intercorrelations Between the Subtests (Table VI). The arguments presented above could be further corroborated by considering the correlations between the different tests. Such arguments, however, are based even more tenuously than the last presented, since we are now dealing with a new variable, i.e., the difference between the tests.

First of all, the argument that the two groups, University and Ft. Churchill, do not come from differently correlated populations may be checked by comparing the r's for the two groups between the different subtests carried out under the same conditions. The two pairs of r's showing the greatest difference are indicated below, together with the calculation of P.

Bare-hand(Heavy)	vs.	Bare-hand(Medium)	.86	(U.)
" " "	vs.	" " "	.65	(Ft.C.)

CR - 1.551 P - .12

Inside-Mitt (H)	vs.	Inside-mitt (M)	.66	(U.)
" " "	vs.	" " "	.40	(Ft.C.)

CR - 1.105 P - .27

Neither difference is significant. These data reaffirm the fact that we are not dealing with differently correlated populations.

The differences which exist between the r's for the Ft. Churchill and University groups under the two conditions, viz. Inside-mitt and Outside-mitt are likewise not significant. The greatest difference which obtains here (that between H outside and L inside) was calculated, and its non-significance demonstrated.

H (Outside-mitt)	vs. L (Inside-mitt)	.67	(U.)
" "	vs. " "	.32	(Ft.C.)
CR - 1.438	P - .15		

If a significant difference had been found here, it would have corroborated our findings on the effect of exposure which was demonstrated in our discussion under section (b). This argument depends upon a chain of relationships, each link being open to chance variations. It is not therefore surprising that the results are indecisive.

CHAPTER IV

DISCUSSION OF RESULTS AND CONCLUSIONS

Up to this point, the manner in which this investigation was carried out has been described, and the results have been presented under a number of separate headings. In this section, certain inadequacies of the experimental set-up will be noted, and the results already stated will be discussed with a view to showing (a) how they must be qualified, (b) how the various results are interrelated, and (c) to show where possible what implications they may have for the Army.

The Tests. Results which are satisfactorily consistent have been obtained with the use of the heavy bolts, and fair consistency with the light. The medium, however, at no point yield a significant result except in the comparison of Bare-hand vs. mitt performance. It was proven to be a highly unreliable test, and some explanation of this fact has been attempted.

Each of the tests can be improved. All would benefit from the spacing between bolts being increased. Machine cut threads are a necessity to control "spinning" (see p.6). The M series would be more satisfactory if the bolts were longer, and an increase of number of bolts in the H should make it even more reliable as a test.

Most, if not all of these difficulties might have been overcome had the tests involved turning bolts into a threaded plate, although the added weight would have increased transportation difficulties on the field.

A wider range of motor tests, particularly work-samples of Army tasks involving the larger muscle groups which could have been put into standardized form, would have been desirable.

General Conditions. Length of exposure period was brief, and nothing is shown of the effects of varying periods of exposure, and how subjects might adapt to them.

To maximize exposure effects subjects kept inactive. The question of how exercise would affect motor performance is not known, although D.I.D. results indicate that exposure plus inaction yield a deficit in performance over that of exposure plus action. The term "exposure" as used in the text, should then be interpreted as meaning a combination of exposure plus inaction.

The effect of winter clothing on these tasks is shown to be very considerable, but their differences are based on initial performance without any adequate opportunity for practice. Results of D.I.D. group suggest that the Ft. Churchill results do not represent what would happen under conditions of continual practice.

Effect of Clothing. There is a very marked and significant difference between the bare-hand performance and performance while wearing winter clothing under inside conditions. As already indicated, these results must be interpreted in relation to the D.I.D. data which suggest very strongly that the effect of clothing on performance may be radically reduced by training.

The evidence concerning the differential effect of clothing on performance is not decisive. There is not sufficient evidence to show that the mitt affects some individuals more than others.

A high differential effect with the mitt would point up the need of redesigning the mitt to meet requirements of individual fit. That there is no suggestion of a need in this respect does not rule out the possibility that other designs might show a less hampering effect on manual tasks of this sort.

<u>Effect of Exposure.</u> A number of factors must be considered here. First of all, the exposure period was a very brief one, nor is the amount of exposure in any sense equated for all subjects, since tests were carried out over a period of 11 days, and conditions of temperature,wind, etc. were changed from day to day. Effects of exposure were intensified by keeping subjects inactive. In addition, there was a quick change from warm to cold and vice versa, and it may be that the rapidity of change is the challenge to adaptability rather than the cold as such.

In any case the subjects were placed in a situation of some discomfort characterized by cold, and showed a significant deterioration of performance in two of the tests (viz.

L and H), the most reliable ones. A difference in the same direction is shown in the M test, but at a probability level of only .10 - .05. The whole picture leaves small room to doubt the effect of exposure on the test performance of this group.

These results are made more secure by virtue of the results obtained by repeating the experiment under summer conditions, and finding no suggestion of deterioration due to the outside conditions in this case.

The results are confirmed again by the D.I.D. data, which, while leaving no doubt as to the validity of the above conclusions, do at the same time modify the over-all picture with respect to the effect of exposure.

As already noted, the cold and warm situations on the trail had to be defined in terms of the subjects' statements concerning his feelings of cold. Sinces these were overwhelmingly on the side of being warmer on the second trial of the day, this was treated as equivalent to the inside condition for the Ft. Churchill group. The first test of the day followed a period of exposure and inaction (sleeping and rising in the cold). The second test followed a period of greater exposure, and fairly vigorous action. Inaction seems to be the decisive factor, since test results show poorest performance under the first condition. Whether the situation here may be called exposure effects is debatable, but at least the inactive period in the cold preceding the first trial of the day corresponds reasonably well to the exposure period in Ft. Churchill, and the warming up by exercise has a certain equivalence to the warming up indoors. Accepting this parallel between the two experimental groups, it is found that there is deterioration for the D.I.D. group of 8.33 seconds. Whether or not this difference is statistically significant, it does supply another bit of confirmatory evidence.

Superimposed on this picture of observations due to cold is another pronounced result from the D.I.D. tests. That is that living in the open under severe winter conditions as these subjects were, a very pronounced practice effect was apparent. On the whole, performance under conditions of exposure and inaction was poorer than under conditions of greater exposure and action, but the practice effect was greater than the effect of the first set of conditions, so that the cold performance of e.g., Day 3 was better than the warm performance of Day 2. Indeed, the outdoor performance, after two practices per day for four days was markedly superior to the indoor performance obtained at the end of the expedition, five days after returning to Ft. Churchill.

With respect to the differential effect of exposure, the argument had to rest on the one reliable test of the series. The evidence fits the reasonable a priori assumption that individuals are not equally affected by the cold. What evidence we have points to the fact that selection may play a role in obtaining men who will not show a great variability in performance under varying degrees of exposure, and that, as will be developed in more detail later, training and indoctrination offset the effect of cold on motor performance.

Acclimatization. The data on the Ft. Churchill group indicate that the ten men who were acclimatized, i.e, that had spent more than one month in Ft. Churchill, and who spent at least one hour per day outside, showed less effect from the cold than the fourteen who were not acclimatized.

One interesting aspect here lies in the fact that the acclimatized group reported subjective cold more frequently than those who,were acclimatized. The suggestion is that the factors of adjustment are not so much physiological as psychological. If this is so, and if the acclimatized person has simply learned to work with his discomfort, it may mean that the importance of training and indoctrination will overshadow that of selection. Such a conclusion is strengthened by the evidence already noted from the D.I.D. group; training under continuous exposure to cold brough about a radical improvement even on such a simple task as we presented, and at the same time brought about a similarly marked reduction in individual differences. This latter effect is probably a function of the task, but then many of the Army tasks involve such elementary dexterity.

Practice Effects. The initial concern over practice effects with the Ft. Churchill group was due to our awareness that they might have been affecting the proper interpretation of the data. If practice effects had been marked, they would have reduced the correlations obtained. In addition to this, on the assumption that the Bare-hand test provided no practice for the mitted trials, the difference in means between inside and outside performance would have been larger than they should to the extent of 1/3 the practice effect. However, as well as could be ascertained the bare hand provided practice, and consequently the corrections for practice on the above noted difference appeared to be unwarranted. Similarly the practice effects were not of great enough magnitude to affect the correlations, so that no use was made of them by way of the original purpose for ascertaining them.

The fact that practice does affect performance, even on these simple tasks, is of importance for the purpose of experimental design for further research. It would be highly desirable that proficiency in any test situation be based on considerable practice before attempting to measure the effects of exposure on it.

Other implications of the practice effect, especially that shown by the D.I.D. group have been discussed earlier and need not be repeated here.

<u>Proficiency and its Relation to Exposure Effects.</u> It has been demonstrated that the subjects who are the most proficient in a given test are the most susceptible to exposure effects. This fact is not amenable to a simple explanation.

Proficiency was based on scores made under warm conditions. In the small D.I.D. group the membership of the proficient group was fairly consistent regardless of which trial proficiency was based on. In the Ft. Churchill group, however, membership in the proficiency group varied as the criterion test shifted from H to M to L. Only between the membership of the M and L groups was there sufficient identity to give a Chi-square of significance.

If membership in the proficient group was consistent, one might argue from the basis of the individuals involved. That is, that the proficient group would be made up of those who on the whole were putting forth maximum effort, and consequently any factor tending to alter performance would be revealed since there would be little room for compensatory effort.

On the basis of the Ft. Churchill group, however, such an argument fails, for the proficient group has a shifting personnel. Here it is the task that is the key, rather than the individual.

Disregarding the M test, which has given inconsistent results throughout, there does appear to be a pattern. If the H test is the criterion, then compared to the non-proficient, the proficient show the greatest deterioration in the same task, none in the M and again in the L. If L is the criterion, then compared to the proficient the non-proficient show the greatest deterioration in the same task, less in the M, and a negligible difference in the H. Here deterioration seems to be directly related to the task, and not a function of the individual as a whole. This cannot be checked in the D.I.D. data for only the H test was used by them.

It should be kept in mind that while the proficient group deteriorate more than the non-proficient, their outside performance is still better than that of the non-proficient. It would seem that cold tends to reduce individual differences in performance, but does not, of course, eliminate them.

Subjective Cold. Only in the experiment on the Ft. Churchill group was there any attempt to compare subjects one with another with respect to their feelings of cold. The results showed no relation between cold report and performance on the test.

Quite a different result was obtained in the U.S.Army "Fox-Hole Study" conducted in Alaska. Here, in a manner comparable to that employed in our own experiment, subjective cold reports were quantified and compared with subjects' performance. It was found that a high degree of correspondence existed between the two: those who felt the cold most had the poorest scores on the performance test. An objective measure of cold, skin temperature, was found to bear no such relation to the scores.

Objective Cold. Our treatment of the relation of objectively cold conditions to performance revealed no reliable difference between the deterioration suffered by those tested under comparatively mild conditions and the deterioration suffered by those tested under more severe conditions.

POINT SUMMARY

The Experiment

1. An experiment was conducted to determine the effect of exposure to sub-arctic winter conditions on performance on a manual dexterity test.
2. The test was a timed one, and involved turning down nuts on three series of bolts, each series of a different size.
3. On the field, two groups of subjects were employed under different conditions:
   - (a) A group of 11 subjects, mainly military personnel, tested on the trail in the environs of Ft. Churchill before and after periods of exposure with activity.
   - (b) A group of 24 civilian personnel, tested at Ft. Churchill under more rigidly controlled conditions,
     - i) Indoors, with indoor clothing and bare hand,
     - ii) Indoors, with outdoor clothing including mitt,
     - iii) Outdoors with outdoor clothing including mitt.
4. A third group of subjects were tested at the University of Manitoba in mild summer weather under the same conditions as in (b) above.

Conclusions

1. The H series bolts provided a highly reliable test.
2. The fact that the M and L tests were essentially unreliable made the results from these tests unreliable and difficult to interpret.
3. On the basis of group performance, there is a significant deterioration of performance effected by the wearing of outdoor clothing.
4. This effect is general in nature, and is not individually differential to any significant degree.
5. A marked group deterioration is evinced by exposure in the case of the Ft. Churchill subjects.
6. This deterioration also takes a differential form, individual subjects being effected in different ways.
7. The effects of exposure were confirmed by results on the D.I.D. group.
8. The D.I.D. data also revealed the very large practice effects inherent in the tests.
9. The practice effects tended to overcome the effects of cold, and to reduce individual differences.
10. Ft. Churchill results indicated that the more acclimatized subjects suffered less deterioration in performance than did the less acclimatized.
11. The fact that the acclimatized reported cold as often as the less acclimatized suggested that the adjustment the former made was a psychological rather than a physiological one.

12. Results of both the Ft. Churchill and D.I.D. investigations showed that the proficient suffered more from exposure than did the non-proficient, although the former still maintained the higher level of performance.
13. Some positive relation between performance and subjective feelings of cold was found in the D.I.D. results. In this case, the subjects' feelings in one situation were compared with their feelings in another.
14. In the Ft. Churchill tests, where subjects' feelings were compared with those of other subjects, no such relation could be found.
15. No relation was observed between objective measures of cold and performance.

BIBLIOGRAPHY

Andrews, T.G., editor, *Methods of Psychology*. New York: John Wiley and Sons, Inc., 1948. 716 pp.

Bedford, T., *Trans. Inst. Min. Eng.*, 1937, 94, p. 76.

Critchley, M., "Effects of Climatic Extremes", *Brit. J. Industr. Med.*, 1947, 4, p.164.

Critchley, M., *Brit. Med. J.*, 1945, 2, p.145.

Fisher, R.A., *Statistical Methods for Research Workers*. Seventh edition; London: Oliver and Boyd, 1938. 356 pp.

Gagge, A.P., A.C. Burton, and H.C. Bazett, *Science*, 1941, 94, p. 428.

Garrett, H.E., *Statistics in Psychology and Education*, Third Edition; New York: Longmans, Green and Co., 1947. 487 pp.

Pierce, F.T., and W.H. Rees, *Shirley Institute Memoirs*, 1944, XIX, 5, p.341.

Roth, G.M., and M.M. Gabrielson, *Amer. J. Physiol.*, 1940, 140, p.168.

Tiffin, J., *Industrial Psychology*. New York: Prentice-Hall Inc., 1946. 386 pp.

APPENDIX "A"

RAW DATA -- FT. CHURCHILL

Subject	Date	Temp. °F.	Wind m.p.h.	Wind Chill	Res. in Ft. C.	Acclimatization	Exposure Time	Order of Tests	1st Test Scores			Test on which fatigue began	Degree of fatigue	Cold Report 1.	Cold Report 2.	2nd Test Scores			Test on which fatigue began	Degree of fatigue	Cold Report 1.	Cold Report 2.	3rd Test Scores			Test on which fatigue began	Degree of fatigue	Cold Report 1.	Cold Report 2.
1	Feb. 28	-11	25	2000 E	2 mos	++	20 min	HML	82	220	*125	2	2	3	2	82	160	110	1	3	-	-	42	38	24	2	1	-	-
2	Mar. 1	- 1	12	1398 E	2 wks	-	15 min	HML	90	102	82	2	1	-	-	64	48	25	0	0	-	-	116	110	*80	1	2	0	0
3	Mar. 1	- 1	12	1390 E	2 wks	+	15 min	HML	70	46	29	0	0	-	-	108	134	*84	1	1	2	1	93	132	75	0	0	-	-
4	Feb. 28	- 7	18	1750 E	2 mos	-	15 min	HML	132	140	*100	1	3	2	3	98	125	65	1	2	-	-	35	35	21	2	1	-	-
5	Mar. 2	-15	10	1710 E	3 mos	+	15 min	HML	82	84	59	0	0	-	-	46	44	27	0	0	-	-	72	90	*81	2	1	2	1
6	Mar. 1	- 1	12	1390 E	2 wks	+	20 min	HML	69	54	26	0	0	-	-	138	106	*71	1	2	2	0	86	91	56	1	2	-	-
7	Mar. 1	- 3	15	1450 E	2 mos	+	15 min	HML	125	195	*126	1	3	3	3	98	165	69	1	2	-	-	75	57	32	1	1	-	-
8	Mar. 8	- 25	10	1750	3 mos	-	15 min	HML	89	198	170	2	2	-	-	55	38	24	0	0	-	-	92	250	*154	1	1	1	1
9	Mar. 4	-12	18	1780 E	2 mos	++	1 hr	HML	44	35	22	0	0	-	-	55	85	*72	2	1	1	1	44	66	39	2	1	-	-
10	Mar. 6	-25	12	1820	4 mos	+	15 min	HML	109	135	*104	1	3	1	2	110	125	82	1	3	-	-	57	45	25	3	1	-	-
11	Mar. 7	-25	15	1910	2 mos	+	15 min	HML	141	203	110	1	2	-	-	62	54	34	0	0	-	-	162	137	*89	1	2	3	4
12	Mar. 8	-25	10	1750	4 mos	+	15 min	HML	39	44	27	0	0	-	-	79	101	*75	2	2	1	1	67	107	52	2	3	-	-
13	Mar. 8	-14	16	1780	3 mos	+	15 min	LMH	82	89	*97	2	2	1	1	76	77	125	1	2	-	-	28	45	64	0	0	-	-
14	Mar. 8	0	4	1110	3 mos	+	15 min	LMH	52	72	95	2	2	-	-	21	48	74	0	0	-	-	104	142	*95	1	1	0	0
15	Mar. 8	-17	12	1730	2 mos	+	15 min	LMH	23	44	41	2	1	-	-	98	132	*112	1	3	1	1	89	99	96	1	2	-	-
16	Mar. 8	-12	14	1670	1 mo.	-	15 min	LMH	133	156	*113	1	2	0	0	101	129	101	1	2	-	-	24	39	51	3	1	-	-
17	Mar. 10	+ 4	Calm	- -	1 mo.	-	15 min	LMH	45	96	92	1	2	-	-	21	49	55	2	2	-	-	127	141	*81	1	2	0	0
18	Mar. 8	+ 10	14	1305	3 mos	-	15 min	LMH	26	52	33	0	0	-	-	108	198	*164	3	2	1	1	103	97	88	-	-	-	-
19	Mar. 9	0	10	1360	2 mos	-	15 min	MLH	156	117	*107	1	2	1	1	165	111	113	1	2	-	-	59	30	74	0	0	1	1
20	Mar. 8	- 14	15	1720	4 mos	+	15 min	MLH	177	101	198	1	3	-	-	73	37	71	3	1	-	-	172	123	*195	1	3	0	0
21	Mar. 9	- 6	14	1580	2 wks	-	15 min	MLH	54	29	67	0	0	-	-	140	78	*124	1	2	0	0	152	72	97	1	2	-	-
22	Mar. 8	- 11	12	1590	1 yr +	++	15 min	MLH	130	85	*101	1	1	0	0	113	58	123	2	1	-	-	63	29	77	3	1	-	-
23	Mar. 8	- 12	14	1670	2 days	-	15 min	MLH	167	115	97	1	3	-	-	55	33	74	1	1	-	-	138	145	*122	1	2	1	1
24	Mar. 9	- 6	14	1580	2 wks	-	15 min	MLH	43	25	65	0	0	-	-	137	111	*124	2	2	0	0	127	91	84	1	1	-	-

E - Estimated

Cold Report: 0 - not cold
1 - slightly cold

Cold Report (contd)
2 - moderately cold
3 - uncomfortably cold
4 - extremely cold

Fatigue — 0 - not tired
1 - slightly tired
2 - uncomfortably tired
3 - extremely tired

Test Scores
- underlining denotes - Inside - mitt
- asterisk " - Outside - mitt
- plain " - Bare - hand

APPENDIX "B"

RAW DATA -- UNIVERSITY

Subject	Date	Temp. °F.	Wind m.p.h.	Order of Tests	1st Test Scores	Test on which fatigue began	Degree of fatigue	2nd Test Scores	Test on which fatigue began	Degree of fatigue	3rd Test Scores	Test on which fatigue began	Degree of fatigue
A	MAY 13	65	10	HML	118 172 *82	2	1	99 118 99	2	2	60 44 29	0	0
			(2ND	TRIAL)				92 105 77	1	2	37 38 24	0	0
B	MAY 13	65	10	HML	105 156 67	1	2	53 33 21	0	0	106 91 *34	1	3
					91 37 15	1	3	38 46 11	2	2			
C	MAY 13	65	10	HML	49 32 18	0	0	71 85 *61	1	3	61 84 57	1	3
					41 34 18	0	0				52 75 55	1	2
D	MAY 13	65	10	HML	109 129 *63	1	1	104 120 70	2	1	88 57 23	0	0
								86 82 36	2	2	62 36 19	3	1
E	MAY 16	68	8	HML	55 95 50	1	3	27 26 18	0	0	51 55 *24	1	2
					46 62 39	1	3	35 31 13	2	1			
F	MAY 16	68	8	HML	60 43 21	0	0	74 110 *95	1	2	70 108 57	1	2
					43 33 20	0	0				53 72 40	1	2
G	MAY 18	55	15	LMH	58 72 *39	1	1	41 49 44	2	2	23 41 49	0	0
								32 62 51	2	2	19 29 39	0	0
H	MAY 16	68	8	LMH	38 76 74	2	1	23 39 49	0	0	41 67 *65	1	1
					37 60 67	1	2	21 39 42	0	0			
I	MAY 16	68	8	LMH	21 34 42	0	0	60 91 *55	2	2	44 75 54	2	2
					22 31 34	3	1				38 77 51	2	3
J	MAY 18	55	15	LMH	45 92 *72	2	3	46 80 62	1	2	22 44 50	0	0
								39 75 62	1	2	21 44 45	0	0
K	MAY 18	55	15	LMH	58 112 62	2	2	19 35 47	0	0	51 110 *55	2	1
					48 80 54	2	2	18 35 43	0	0			
L	MAY 18	55	15	LMH	25 43 57	3	1	48 83 *79	3	1	48 87 77	2	2
					25 45 48	0	0				42 75 65	2	2
M	MAY 20	66	9	MLH	69 34 *57	1	1	64 32 37	2	1	26 20 28	0	0
								61 31 31	2	2	24 12 28	0	0
N	MAY 20	66	9	MLH	60 53 87	1	2	43 24 56	0	0	66 50 *87	1	2
					70 44 68	1	2	41 23 51	0	0			
O	MAY 18	55	15	MLH	45 23 43	2	1	142 49 *83	1	3	116 51 77	1	2
					40 17 35	0	0				117 45 71	2	3
P	MAY 20	70	12	MLH	107 77 *97	1	2	96 72 95	1	2	48 31 69	1	1
								91 64 80	1	3	48 29 64	1	1
Q	MAY 20	70	12	MLH	101 49 64	1	3	41 24 42	0	0	67 39 *68	1	2
					81 31 60	1	2	40 27 31	2	1			
R	MAY 18	55	15	MLH	51 22 63	0	0	124 92 *89	1	2	104 85 92	1	3
					48 22 50	0	0				102 87 87	1	3

Fatigue
- 0 - not tired
- 1 - slightly tired
- 2 - uncomfortably tired
- 3 - extremely tired

Test scores
- = underlining denotes - Inside - mitt
- asterisk " - Outside - mitt
- plain " - Bare - hand

APPENDIX "C"

RAW DATA -- D. I. D.

Subjects	Trial	Test Scores Day 1	Test Scores Day 2	Test Scores Day 3	Test Scores Day 4	Inside - mitt score	Inside - bare score
1	1	85	69	44	38	41	29
	2	66 +	55 +	37 ++	34 ++		
2	1	140	119	70	50	72	28
	2	113 +	96 ++	60 ++	40 +		
3	1	78	61	31	44	42	20
	2	68 ±	51 ++	42 -	40 ±		
4	1	60	45	32	37	36	25
	2	53 ±	33 ±	25 ±	35 ±		
5	1			81	66	40	31
	2			66 ±	36 +		
6	1		113	34	41	50	27
	2		45 ±	35 +	40 +		
7	1		91	46	53		
	2		60 ±	60 ±	54 ±		
8	1		69	52	47	60	32
	2		52 ±	54 +	58 ±		
9	1	80	63	52	43	44	38
	2	71 ±	68 -	60 +	50 ±		
10	1		63	68	49		
	2		69 --	50 ±	35 +		
11	1			71	52	63	41
	2			69 ±	51 ±		

METEOROLOGICAL DATA:

	TEMP:	WIND
DAY 1	-40^{0} F	14 MPH
DAY 2	- 28	4
DAY 3	- 8	2
DAY 4	- 15	5