ARMORED MEDICAL RESEARCH LABORATORY

FORT KNOX, KENTUCKY

First Supplementary Report

On

PROJECT NO. 30 - FIELD TEST OF ACCEPTABILITY AND ADEQUACY OF U. S. ARMY C, K, 10-IN-1 AND CANADIAN ARMY MESS TIN RATIONS

Subject: Numerical Requirements for Statistically Valid Results in Field Test of Acceptability of Rations

Project No. 30

24 March 1945

ARMORED MEDICAL RESEARCH LABORATORY
Fort Knox, Kentucky

Project No. 30
SPMEA 430.2

24 March 1945

NUMERICAL REQUIREMENTS FOR STATISTICALLY VALID RESULTS IN FIELD TEST OF ACCEPTABILITY OF RATIONS

1. PROJECT NO. 30 - Field Test of Acceptability and Adequacy of U. S. Army C, K, 10-in-1 and Canadian Army Mess Tin Rations. First Supplementary Report.

a. Authority: Letter SPQRD 400.112 (Ration Acceptability, C and K) and 4th Ind. A.S.F., Office of the Surgeon General, SPMDO, 4 April 1944.

b. Purpose: To determine with regard to field testing of acceptability of rations:

(1) Number of men to be used as test subjects
(2) Number of meals on which data should be obtained
(3) Duration of test

2. DISCUSSION:

a. Since field tests must be based on some form of sampling it is essential that test subjects be representative of the population to which the results are to be applied, and conditions in the field must resemble as closely as possible the expected conditions for which the test material is designed. It is understood that no test will be satisfactory unless the observers are trained and have absolute control of the subjects to enforce isolation from extraneous sources of food and to control the activity of the subjects.

b. Moreover, in setting up a field test of acceptability of rations it is important to know the number of men, the quantity of data and the length of time required to obtain results which have statistical validity. If such information is not available in advance the test may be carried out on too few men, insufficient data may be gathered and the test may last too short a time; thus, even the most carefully collected data may give results which are not significant. On the other hand, if more than enough subjects are used for longer periods than required and the data obtained are more than sufficient, waste occurs in man hours of test subjects and observers.

c. Study of the data obtained in The Field Test of Acceptability and Adequacy of Rations* carried out by the Armored Medical Research Laboratory at Topside, Colorado in the summer of 1944 permits the definition of size, time and quantity of data necessary for this type of field test.

* A.M.R.L. Report on Project No. 30: Acceptability and Adequacy of U. S. Army Rations C, K, 10-in-1 and Canadian Army Mess Tin Rations, 22 November 1944.

d. It is emphasized that acceptability of rations will be affected by a number of variables such as ambient temperature; season, rainfall, stress and level of work. Accordingly any inference as to adequacy or techniques of testing which may be drawn from the Topside study must be interpreted with caution when applied to tests of other rations under different conditions. With these limitations in mind, however, certain useful generalization may be made.

3. CONCLUSIONS:

a. In short tests (2-3 weeks) little accuracy is sacrificed by collecting opinion data only every third or fourth day.

b. In longer tests (2 months) no significant accuracy is lost if data are obtained only once a week.

c. Data from the first few days may show more variations than those obtained later on in a 3-week test.

d. Where there are only relatively few items to test, a period long enough for each menu or item to appear only 3 or 4 times will apparently give results little different from those obtained when the test components are used several times as often.

e. If acceptability of rations for weeks or months is desired, it is necessary to test the rations for the entire period but data need be collected no more frequently than every 6 or 7 days.

f. From the statistical point of view, groups of at least 2 platoons (60-80 men) are needed to give reliable information on the relative acceptability of items in a single ration. For reasons of organization, supply, command, morale and losses from normal attrition, groups of company size (100-150 men) are more desirable.

4. RECOMMENDATIONS:

a. The information contained in this report be made available to agencies engaged in ration tests.

Report submitted by:
Charles R. Henderson, Captain, SnC
William B. Bean, Major, MC
William F. Ashe, Major, MC

APPROVED Willard Machle
WILLARD MACHLE
Colonel, Medical Corps
Commanding

2 Incls.
#1 - Appendix
#2 - Figure 1

APPENDIX A

MATERIALS AND PROCEDURE

1. For purposes of this study, the following were obtained from the original acceptability data for E Company on K Ration for 19 consecutive days:

a. Total of good, fair, and poor ratings on each item for each day of the test.

b. Total during 19 days of good, fair, and poor ratings on each item for each man of the company.

c. From each of these totals was calculated the percentages of good ratings. Working with these percentages, analysis of variance methods were used to calculate "date X item discrepance" and "men X item discrepance", each within food groups.

d. With these two statistics available, it is possible to estimate the reliability of results obtained from different numbers of men for varying periods of time.

2. RESULTS

a. It should be made clear that the discussion to follow is concerned primarily with comparison of different items of the same general type when eaten concurrently by the same men. The comparison between items to be used is the average percentage of "good" ratings by which one item exceeds another. If only a portion of the available data is used, the advantage one item has over another will not be exactly the same as when all data are used. The probable differences between results using the data for all men of E Company for nineteen (19) days on the K Ration and results using data from fewer days and data from fewer men are tabulated in Tables 1 and 2.

Incl. #1

TABLE 1

Number of Days Data Used	Maximum Probable Error in Difference Between Means of any Two Items.* (Per Cent Rated Good)
1	15.3
2	10.5
3	8.3
4	7.0
5	6.1
6	5.3
7	4.7
8	4.2
9	3.8
10	3.4
11	3.1
12	2.8
13	2.5
14	2.2
15	1.9

* Mean Item A (19 days)—Mean Item B (19 days)—Mean Item A (n days)—Mean Item B (n days); $P > .95$

TABLE 2

Number of Men's Data Used	Maximum Probable Error in Difference Between Means of any Two Items. (Per Cent Rated Good)
10	33.7
20	23.8
30	19.4
40	16.8
50	15.1
60	13.7
70	12.7
80	11.9

Incl. #1

b. Illustrating the use of the above tables, coffee and bouillon were rated good 67% and 47% of the time respectively by E Company during the period studied, a difference of 20%. Had data from only four days been picked at random for Company E for these two items, the difference between the two items would probably (P > 95%) fall in the range 20% ±7% or 13% to 27%. The odds are better than even that it would fall between 17% and 23%. If data had been selected from only one day, the difference would probably (P > 95%) fall in the range 5% to 35%. From Table 2 it may be seen that using the 19 days' data from 50 of the 110 men would be likely to give results as far from the actual results observed as would be the case when only 1 day's data on all men were used. Similarly the use of 2 days' data from all men would give results somewhat more accurate than would use of 80 men's data for the entire 19 days.

c. Figure 1, on which are plotted the day by day acceptability of 3 beverages, illustrates why little accuracy is sacrificed when only a relatively small proportion of the different day's data is used. After the first 3 days, the rank of the items remained the same. Minor fluctuations in absolute ratings are not considered to be important.

d. The wide differences among men (even though in the same company) in their opinions of the relative merits of different items are illustrated in Tables 3 and 4--Table 3 shows the ratings given by three different men to three different beverages. The discrepancies among them are almost identical with those for all men of the company.

TABLE 3

Percentages of Time Different Men Rated Different Beverages Good

Subject	Lemon Juice	Bouillon	Coffee
E101	50	0	59
E103	20	97	79
E112	7	0	36

Table 4 illustrates the wide difference of opinion among Company mates as to the rank of each of 3 beverages.

TABLE 4

Percentage of Men Ranking Each of Three Beverages
First, Second and Third Choice

Beverage	First	Second	Third
Lemon Juice	33	29	38
Bouillon	24	22	54
Coffee	43	49	8

To illustrate further the differences which may complicate interpretation of acceptability questionnaires, some of the differences among platoons in the same company and among companies are presented in Tables 5 and 6:

TABLE 5

Average Percentage of Good Ratings of Confections by Different Platoons in the Same Company

Platoon	Sugar	Gum	Fruit Bar	Caramels	Sweet Chocolate Bar
1	91	92	55	89	82
2	82	76	59	75	81
3	94	93	72	87	83
4	89	90	75	89	88

The platoon by item discrepance is statistically significant.

TABLE 6

Differences Among Companies in Acceptability Scores* Given Experimental C Ration Item Groups

Item Group	G Company	X Company	Y Company
Meats	2.74	2.87	2.59
Cereals	2.91	2.94	2.80
Beverages	2.78	2.89	2.74
Confections	2.97	2.97	2.93
Biscuits	2.91	2.87	2.64

* Score calculated by the equation 3 (No. of good) +2 (No. of Fair) +1(No. of Poor) divided by the total number of ratings.

e. The results of this study indicate that the most promising way in which to reduce the labor of obtaining ration acceptability information is through reducing the number of days of data collection rather than through reducing the number of test subjects.

f. If information regarding the relative acceptability of items for a short period of time (say 3 weeks) is desired, there appears to be little value in obtaining data more often than every 3 or 4 days. If one suspects that there will be a trend in acceptability upwards or downwards during a period as short as 3 weeks men should be kept on the ration for that length of time. Collection of data only on each third or fourth day would give results very little different from data collected daily.

g. If information is desired on long time acceptability, the only apparent way to obtain it is to feed the ration for a long time. Extrapolation of results from a short test is a dangerous procedure. If a long test is planned, even longer

intervals between data collection days would be possible. (Five to 7 days should give valid results).

h. In contrast to the small variability from day to day in the relative acceptability of ration items, different men in the same company differ widely in their opinions of the relative desirability of different items. For example, the odds are even that any 2 men of a company picked at random will differ in their opinions as to how much better one item is than another by nearly 30%. This being the case, the use of too small numbers of men in evaluating a ration is likely to result in data which are not representative of the population in which one is interested.

i. Furthermore, the fact that platoons differ within companies and that companies themselves differ in their opinions of the relative merits of rations and ration items indicates that great care must be exercised in selecting test personnel representative of the larger group for whose feeding the ration is proposed.

j. If it is desired to compare different rations, the test plan should be set up with at least 2 platoons or better a company for each ration; and in which groups are systematically rotated from one ration to another to smooth the differences among groups. On the basis of this study, it would appear that an entire company, or at the least two platoons, are needed to test ration items adequately for acceptability.

FIG. I

DAILY FLUCTUATION IN ACCEPTABILITY OF K RATION BEVERAGES

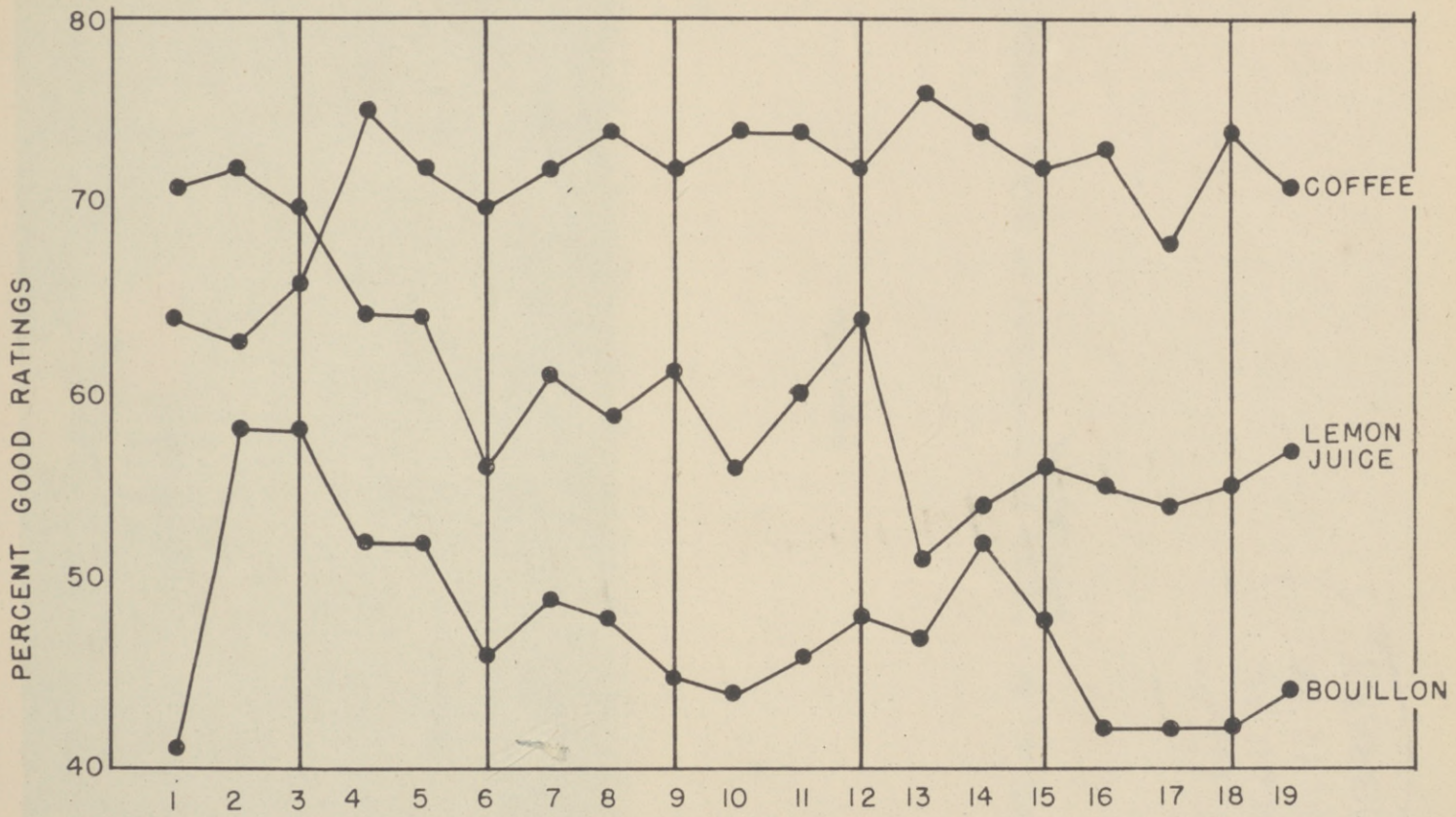

FIG. I

Incl. #2