QMFCIAF REPORT NO. 11-62

THE LENGTH AND DIRECTION OF RATING SCALES

Interim Report
June 1962

QUARTERMASTER FOOD AND CONTAINER INSTITUTE FOR THE ARMED FORCES
QUARTERMASTER RESEARCH AND ENGINEERING COMMAND, U.S. ARMY
CHICAGO 9, ILLINOIS

AD ______ Accession No. ______
QM Food & Container Institute for the Armed Forces,
QM Research & Engineering Command, U. S. Army,
Chicago 9, QMFCIAF Rpt. No. 11-62
Date June 1962 Proj. No –
pp 20 tbl 7 fig. 1

The Length and Direction of Rating Scales by R. W. Seaton

Reliability of preference differences among homogeneous and heterogeneous food samples was studied as a function of two rating-scale variables, physical length and vertical vs. horizontal printing.

Primary Field: **Human Factors Research**
Secondary Field(s): **Psychometrics**

UNCLASSIFIED

I. Psychometrics
II. Scaling (Psych.)
1. Seaton, R.W.

AD ______ Accession No. ______
QM Food & Container Institute for the Armed Forces,
QM Research & Engineering Command, U. S. Army,
Chicago 9, QMFCIAF Rpt. No. 11-62
Date June 1962 Proj. No. –
pp 20 tbl 7 fig. 1

The Length and Direction of Rating Scales by R. W. Seaton

Reliability of preference differences among homogeneous and heterogeneous food samples was studied as a function of two rating scale variables, physical length and vertical vs. horizontal printing.

Primary Field: Human Factors Research
Secondary Field(s): Psychometrics

UNCLASSIFIED

I. Psychometrics
II. Scaling (Psych.)
1. Seaton, R.W.

AD ______ Accession No. ______
QM Food & Container Institute for the Armed Forces,
QM Research & Engineering Command, U. S. Army,
Chicago 9, QMFCIAF Rpt. No. 11-62
Date June 1962 Proj. No. –
pp 20 tbl 7 fig. 1

The Length and Direction of Rating Scales by R. W. Seaton

Reliability of preference differences among homogeneous and heterogeneous food samples was studied as a function of two rating-scale variables, physical length and vertical vs. horizontal printing.

Primary Field: Human Factors Research
Secondary Field(s): Psychometrics

UNCLASSIFIED

I. Psychometrics
II. Scaling (Psych.)
1. Seaton, R.W.

AD ______ Accession No. ______
QM Food & Container Institute for the Armed Forces,
QM Research & Egineering Command, U. S. Army,
Chicago 9, QMFCIAF Rpt. No. 11-62
Date June 1962 Proj. No. –
pp 20 tbl 7 fig. 1

The Length and Direction of Rating Scales by R.W. Seaton

Reliability of preference differences among homogeneous and heterogeneous food samples was studied as a function of two rating-scale variables, physical length and vertical vs. horizontal printing.

Primary Field: Human Factors Research
Secondary Field(s): Psychometrics

UNCLASSIFIED

I. Psychometrics
II. Scaling (Psych.)
1. Seaton, R.W.

QMASO Form 12-437
Rev 3 Jul 61

AD Accession No.
QM Food & Container Institute for the Armed Forces,
QM Research & Engineering Command, U. S. Army,
Chicago 9, QMFCIAF Rpt. No. 11-62
Date June 1962 Proj. No -
pp 20 tbl 7 fig. 1

The Length and Direction of Rating Scales by R. W. Seaton

Reliability of preference differences among homogeneous and heterogeneous food samples was studied as a function of two rating-scale variables, physical length and vertical vs. horizontal printing.

Primary Field: **Human Factors Research**
Secondary Field(s): **Psychometrics**

UNCLASSIFIED

I. Psychometrics
II. Scaling
(Psych.)
1. Seaton, R.W.

AD Accession No.
QM Food & Container Institute for the Armed Forces,
QM Research & Engineering Command, U. S. Army,
Chicago 9, QMFCIAF Rpt. No. 11-62
Date June 1962 Proj. No. -
pp 20 tbl 7 fig. 1

The Length and Direction of Rating Scales by R. W. Seaton

Reliability of Preference differences among homogeneous and heterogeneous food samples was studied as a function of two rating scale variables, physical length and vertical vs. horizontal printing.

Primary Field: **Human Factors Research**
Secondary Field(s): **Psychometrics**

UNCLASSIFIED

I. Psychometrics
II. Scaling
(Psych.)
1. Seaton, R.W.

AD Accession No.
QM Food & Container Institute for the Armed Forces,
QM Research & Engineering Command, U. S. Army,
Chicago 9, QMFCIAF Rpt. No. 11-62
Date June 1962 Proj. No. -
pp 20 tbl 7 fig. 1

The Length and Direction of Rating Scales by R. W. Seaton

Reliability of preference differences among homogeneous and heterogeneous food samples was studied as a function of two rating-scale variables, physical length and vertical vs. horizontal printing.

Primary Field: **Human Factors Research**
Secondary Field(s): **Psychometrics**

UNCLASSIFIED

I. Psychometrics
II. Scaling
(Psych.)
1. Seaton, R.W.

AD Accession No.
QM Food & Container Institute for the Armed Forces,
QM Research & Egineering Command, U. S. Army,
Chicago 9, QMFCIAF Rpt. No. 11-62
Date June 1962 Proj. No. -
pp 20 tbl 7 fig. 1

The Length and Direction of Rating Scales by R.W. Seaton

Reliability of preference differences among homogeneous and heterogeneous food samples was studied as a function of two rating-scale variables, physical length and vertical vs. horizontal printing.

Primary Field: **Human Factors Research**
Secondary Field(s): **Psychometrics**

UNCLASSIFIED

I. Psychometrics
II. Scaling
(Psych.)
1. Seaton, R.W.

QMFCIAF REPORT NO. 11-62

PROJECT: Human Factors in Quartermaster Corps Operations

TASK: Attitude toward and acceptance of Quartermaster materiel

PHASE: Methods of measuring and predicting acceptance

THE LENGTH AND DIRECTION OF RATING SCALES

Interim Report

by

Richard W. Seaton

Food Acceptance Branch, Food Division

June 1962

Quartermaster Food and Container Institute for the Armed Forces

THE LENGTH AND DIRECTION OF RATING SCALES

Concern with the problem of obtaining sensitive and reliable measures of consumer judgment has stimulated efforts toward improving rating scales so that judges may use them with greater ease and precision. It has been found, e.g., that having labelled points on a rating scale improves rating performance, and that scales with five or six points yield more information than those with only two of three points.

Among the problems remaining for the scale designer is that of the optimum physical length of the scale. Short scales may lack legibility due to crowding, but long scales may be so spread out as to be confusing. Generally, scales are from three to six inches long, but longer or shorter scales might be more productive. A second practical problem is scale direction, i.e., vertical or horizontal orientation. Does direction make any difference in rater performance? A review of the literature on scaling showed no research on these two problems despite their practical and possible theoretical implications. This study was undertaken to provide systematic evidence on performance effects of the length and direction of rating scales.

Theoretical Considerations

In performing the rating task, the rater is provided with an index, the rating scale, which represents the degrees of an attribute in a stimulus class or the degrees of some characteristic in the rater's responses to the stimulus class[1]. Categories on the rating scale can be viewed as representing ideal conceptual degrees of an attribute to which each stimulus event is compared (Coombs, 1953). The rater's task is thus seen as one of matching the perceived level of an attribute in a stimulus event with the several categories of a scale. "The problem for the rater in rating each stimulus is one of saying that the large number of events (represented by the stimulus) is the same as, or different from the events represented by each category on the rating scale" (Bendig, 1954).

This process has been likened to encoding, transmission and decoding operations in information theory (Garner & Hake, 1951). The stimuli are conceived as representing (encoding) a range of levels of an attribute in the stimulus class, the rater is conceived of as the transmission channel, and the rating scale is the decoding device. There are four parameters to the decoding operation. <u>Maximum response information</u> (bit potential) is set by the number of available categories used and by their frequency of use; it approximates maximum response information on moderately structured scales (Bendig, 1954), since raters tend to use most of the available categories when there are no more than 6-10.

---

[1] This differentiation between stimulus attributes and attributes of response to a stimulus is somewhat artificial. Potentiality of eliciting a response (e.g., "like slightly") can be considered an attribute of the stimulus and, contrariwise, estimation of an attribute level in a stimulus can be considered a response; also empirical work on scaling has revealed little or no qualitative difference between measurements of the two kinds of attributes. (In this connection see discussion in Volkman, Hunt, and McGourty, 1940).

Response equivocation is equivalent to the average uncertainty or error of response to a given stimulus due to intra- and inter-rater disagreement; it is subtracted from gross obtained information to give a measure of residual transmitted information, i.e., variation in ratings attributable solely to inter-stimulus differences (Erikson & Hake, 1955).

Key problems in designing reliable rating scales are maximizing obtained information and minimizing concomitant response equivocation error. Unfortunately, these two data attributes are not always independent. Transmitted information, for example, can be shown to augment regularly with increasing number of scale categories (Bendig & Hughes, 1953; Bendig, 1954; Jones, Peryam, and Thurstone, 1955). However, response equivocation also increases with increased number of categories. Accordingly, the reliability of increased transmitted information does not notably increase with increasing number of scale categories (Bendig, 1953; Bendig, 1954; Eng, 1948; Erickson and Hake, 1955; Jones, 1959).

Thus for a scale with a given number of categories, the path to increasing reliability lies in decreasing error relative to transmitted information. Increasing the heterogeneity of stimuli appears to have this effect. Not only is variation in ratings reflecting inter-stimulus differences enhanced, but error declines slightly; reliability is accordingly increased (Engen, Levy, and Schlosberg, 1958; Hunt and Flanery, 1938; Erikson and Hake, 1955; Bendig, 1955). Bendig, noting this error reduction and basing his analysis on Volkman (1951), comments: "The rater tends to adjust the psychological length of the scale to fit the range of the stimuli," (Bendig, 1954). Accordingly, each scale category spans a greater psychological interval, and the midpoints of adjacent categories have greater psychological difference in meaning. With greater contrast of meaning between adjacent scale categories, the stimuli-into-categories matching task of the rater may be eased and he may perform more reliably. A similar interpretation may apply to the study of Bendig (1955), in which increases in reliabilities resulted from extending the range of meaning encompassed by verbal anchors at the scale extremes. One would expect, too, that reliability would improve if inter-category contrast is enhanced by substitution of verbal for numerical scale anchors, as was found by Bendig (1953).

Generally, then, findings on reliability in the rating task suggest that matching a stimulus series to a series of rating scale categories is relatively more error-free as members of either series tend to be more distinct. Thus, increasing the density (i.e., the number of stimuli) in a stimulus series of given range diminishes contrast between members of that series; such more dense series were directly associated with decreasing intra-judge reliability of ratings (Volkman, Hunt and McGourty, 1940).

The physical spacing and layout of rating scales may be a factor affecting contrast of scale categories and accordingly the reliability of ratings. A review of the literature reveals almost no research on geometric dimensions as they relate either to distinctiveness of categories or to reliability of ratings. Guilford (1936) has recommended that rating scales be held to an over-all length close to five inches for reasons of clarity and continuity, but evidence for this conclusion is not given.

Discrimination of deviations of the positions of dots in patterns has been found to be an inverse function of the dispersion of the dots within the patterns (French, 1953). On the other hand, wide separation between stimuli appears to be related directly to distinctiveness in discrimination learning experiments (Wodinsky, Varley, and Bitterman, 1954). These findings may relate to the question of clarity of spacing of scale categories. Presumably the closer one brings distinctly different categories together, the greater the contrast between them; but if this step leads to the merging or overlapping of the meanings of the two categories, distinctiveness between them will be reduced.

The experiment described below was aimed at testing possible effects of varying the rating scale category spacing to see if these may affect the reliability with which subjects rate a set of relatively homogeneous or a set of relatively heterogeneous food samples on horizontal and vertical scales.

Method

Six different scales were tested. All included nine categories and required subjects to check their level of liking of different samples of milk. The two extremes and the middle of the nine categories were labelled as shown in Figure 1. A nine-point vertical labelled rating scale (Peryam and Pilgrim, 1957) usually is used in these tests. The taste-test laboratory includes a bank of semi-enclosed booths, each having a stool, table and opaque window through which food samples are successively passed for rating. Milk is often tested, and having to taste three variant formulations also was usual for the subjects.

The fact that only three, rather than all nine, of the categories of the experimental scales were labelled represented a deviation from usual practice. However, instructions to subjects to check scale categories reflecting their liking for individual successive samples were typical of usual practice.

To avoid confounding of experimental conditions (heterogeneity of samples, and length and direction of scales) with session effects all 12 treatment combinations were presented to 12 three-person sub-groups on each of six successive working days. Each daily session included 36 subjects. No subject participated in more than one session, so variance estimates between sessions or subsets of sessions could be considered independent. A total of 18 subjects were included in each of the 12 experimental conditions.

Because session effects on preparation of samples and on subjects' rating performance might be expected to contribute to error, the ratings of the 18 subjects taking part in each experimental condition were divided into three subsets of six; each subset included ratings of the six subjects testing under a given condition during two successive daily test sessions. Since there were 12 experimental conditions and three pairs of successive daily sessions, there were a total of 36 subsets. Each six-person subset of ratings was treated by analysis of variance with partitioning of the 17 degrees of freedom as shown in Table 1.

Of particular interest in the present study were the mean square variances associated with differences between ratings of samples ("obtained information"), and associated with error ("response equivocation").

Ratios between these two variance estimates in each of the 36 analyses were expressed as Fisher's z (Fisher, R. A. and F. Yates, Statistical Tables for Biological, Agricultural, and Medical Research. New York: Hafner Publishing Co., Inc., 1957, p. 2 and Table 5), (Fisher and Yates, 1957) to which unity was added so that ratios lower than one would be expressed positively. The 36 z values so obtained were classified by sample heterogeneity, scale direction and scale length, and analysis of variance of the transformed variance ratios was performed. With error df = 8 in each of the 36 separate analyses of subsets, the distribution of z values may have deviated moderately from normality; however, such moderate deviation would not be expected to affect much the applicability of the analysis of variance. (Anderson, 1961).

Results

Error variance (response equivocation) estimates for each of the 36 subjects were derived as shown in Table 1. The average values obtained from the three subsets in each of the 12 experimental conditions appear in Table 2. These values represent the averages of denominators of variance ratios in the 36 analyses of variance forming the central data in this report. Apparently scale direction made little over-all difference in error, but there is a trend toward less error when more heterogeneous samples were judged. This trend is in agreement with previous evidence (Engen, Levy, and Schlosberg, 1958; Erikson and Hake, 1955) showing increased reliability when stimuli are more diverse. The effect of scale length appears non-linear; however, lowest error was observed for the longest scale.

Table 1

Partitioning of variance in subsets

Source	df
Milk Samples (M)	2
Days (D)	1
M X D	2
Subjects (S) w/in Days	4
S X M w/in Days (error)	8
Total	17

Table 2

Averages of three subset estimates of error variance (subject-sample interaction within test days) in each of 12 experimental conditions

Samples:	**Homogeneous**			**Heterogeneous**			**Homogeneous & Heterogeneous**		
Scale Directions:	**Horiz.**	**Vert.**	**Both**	**Horiz.**	**Vert.**	**Both**	**Horiz.**	**Vert.**	**Both**
Scale Length									
2-3/4"	1.98	2.37	2.18	2.84	1.69	2.27	2.41	2.03	2.22
5½"	3.43	2.00	2.72	1.82	3.19	2.51	2.63	2.60	2.61
8¼"	1.79	2.34	2.06	1.32	1.55	1.44	1.56	1.94	1.75
Total	2.40	2.24	2.32	1.99	2.14	2.07	2.20	2.19	2.19

Averages of estimates of inter-sample variances (obtained information) in each of the experimental conditions appear in Table 3. The longest scales had the largest inter-sample variance, the medium length scales ($5\frac{1}{2}$") produced the lowest inter-sample variance. However, differences between scales in terms of transmitted information were dependent on scale direction; the medium-length vertical scale and the long horizontal scale produced less inter-sample variation than their opposites of the same lengths. This interaction makes unfeasible generalization about the relative amounts of information obtained by the scales of different lengths.

The main aim of the study was to ascertain if scale length might produce changes in the information/error ratio -- i.e., in net transmitted information -- either by increasing obtained information or decreasing error. Fisher's z values for variance ratios obtained from the three analyses of variance computed within each of the 12 experimental conditions are shown in Table 4. The analysis of variance of these normalized ratios (Table 5) confirms the expected increase in net transmitted information obtained when samples were more variable and therefore more distinguishable. The analysis also suggests that scales of different lengths produced different amounts of information. However, the concomitant interaction of scale length with scale direction, coupled with data in the table of averages, restricts the generality of the main effects of scale length. Variance ratios among the horizontal scales remained relatively constant regardless of scale length, whereas among vertical scales the middle of the three lengths produced less information than did the longer and shorter lengths.

Table 3

Averages of three subset estimates of inter-sample variance in each of 12 experimental conditions

Samples:	Homogeneous			Heterogeneous			Homogeneous & Heterogeneous		
Scale Direction:	Horiz.	Vert.	Both	Horiz.	Vert.	Both	Horiz.	Vert.	Both
Scale Length:									
2-3/4"	2.65	4.72	3.68	12.65	9.07	10.86	7.65	6.89	7.27
5½"	2.87	.93	1.90	17.87	4.04	10.95	10.37	2.48	6.42
8¼"	2.59	4.98	3.79	8.46	14.98	11.72	5.53	9.98	7.75
Total	2.70	3.54	3.12	12.99	9.36	11.18	7.85	6.45	7.15

Table 4

Averages of three subset estimates of normalized variance ratios expressing net transmitted information in each of 12 experimental conditions

Samples:	Homogeneous			Heterogeneous			Homogeneous & Heterogeneous		
Scale Direction:	Horiz.	Vert.	Both	Horiz.	Vert.	Both	Horiz.	Vert.	Both
Scale Length:									
2-3/4"	1.16	1.21	1.19	1.73	1.80	1.76	1.45	1.50	1.47
5½"	.96	.43	.69	2.03	1.16	1.58	1.49	.78	1.14
8¼"	1.03	1.43	1.23	1.90	2.16	2.03	1.47	1.80	1.61
Total	1.05	1.02	1.04	1.89	1.70	1.79	1.47	1.36	1.41

Table 5

Analysis of variance of z values

Source	df	Sum of sq. dev.	Mean sq.	F.	Signif.(p)
Scale length (L)	2	1.5226	.7613	2.79	.10
Scale direction (D)	1	.1064	.1064	<1	
Sample variability (V)	1	5.1234	5.1234	18.79	.01
L X D	2	1.7487	.8742	3.21	.10
L X V	2	.1528	.0764	<1	
D-V	1	.0621	.0621	<1	
L X D X V	2	.2701	.1351	<1	
Replicate	2	.9136	.4568	1.68	n.s.
Error*	22	5.9979	.2726		
Total	35	15.8976			

*Pooled error term consisted of all interactions involving replicates.

Of subsidiary interest in the experiment was the question of how scale form and sample heterogeneity might affect ratings of a constant sample. Since one sample (milk with 40 percent dilute evaporated milk) was used in all experimental conditions, mean ratings of this sample obtained under the 12 different experimental conditions on each of the six days of testing (see Table 6) were subjected to analysis of variance as shown in Table 7. Generally the mean ratings show little variability except for session variation. There is a suggestion, however, that scale direction affected the level of mean ratings.

Table 6

Mean ratings, averaged over six test days, for fresh milk mixed with 40% of evaporated milk and evaluated under each of 12 experimental conditions

Samples:	Homogeneous			Heterogeneous			Homogeneous & Heterogeneous		
Scale Direction:	Horiz.	Vert.	Both	Horiz.	Vert.	Both	Horiz.	Vert.	Both
Scale Length:									
2-3/4"	5.30	5.90	5.60	5.45	6.18	5.82	5.37	6.04	5.71
5½"	5.60	5.88	5.74	5.00	5.65	5.32	5.30	5.77	5.53
8¼"	5.71	6.15	5.93	6.11	5.45	5.78	5.92	5.80	5.86
Total	5.54	5.98	5.76	5.52	5.76	5.64	5.53	5.87	5.70

Table 7

Analysis of variance of ratings of constant sample

Source	df	Sum of sq. dev.	Mean sq.	F.	Signif.(p)
Heterogeneity (H)	1	.2450	.2450	1	
Test days (T)	5	14.0917	2.8183	3.92	.01
Scales (S)	(5)	(5.3250)	(1.0650)	(1.48)	
Direction (D)	1	2.0672	2.0672	2.87	.10
Length (L)	2	1.2700	.6350	1	
L X D	2	1.9878	.9939	1.38	
H X S	5	3.2567	.6513	1	
H X T	5	6.7833	1.3567	1.86	
S X T	25	28.7333	1.1493	1.60	
H X S X T	25	17.9850	.7194		
Total	71				

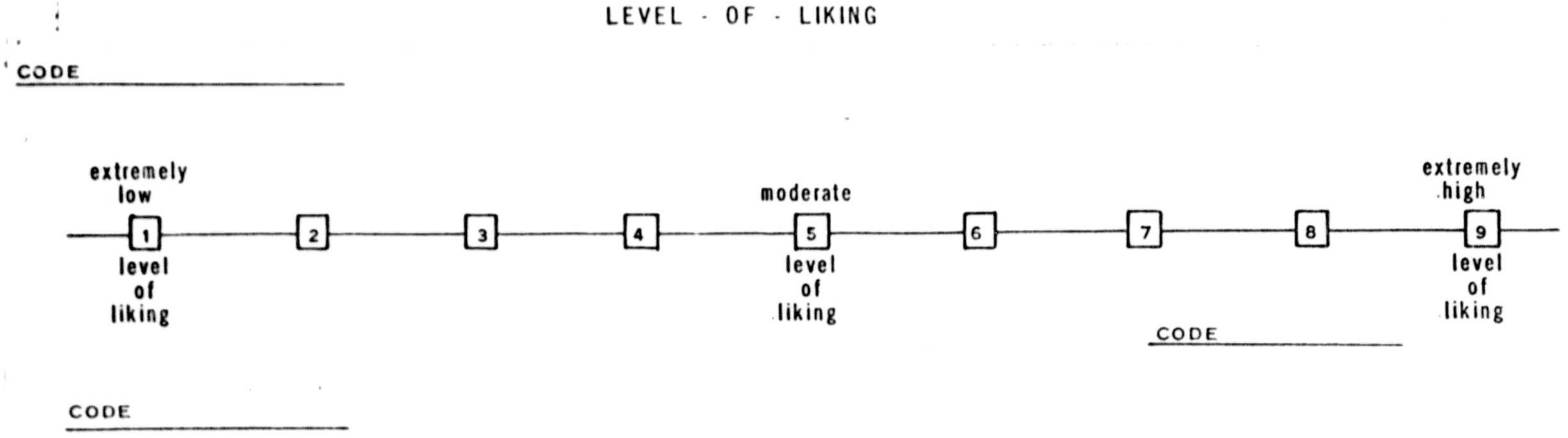
Discussion and Conclusion

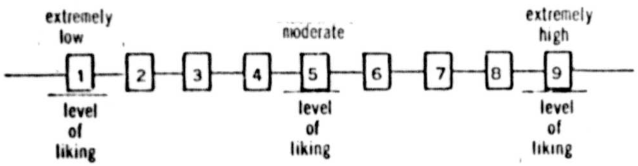
Indications of a non-linear effect (at the 10 percent level) of scale length on net transmitted information only among the vertical test scales were not expected. An accidental variation in scale reproduction may account for this. Through a printing error, the bottom end of the vertical 5½" scale touched the foot of the page so that there was no margin between the bottom of the scale and the bottom of the page. Since all nine anchor points on the scale appeared on the page, it was not thought necessary to reprint the scale. Nevertheless, the fact that the bottom end of the scale ran off the page may have given the impression of poorer psychological anchoring. If this were so, the variation in sensitivity of vertical scales might be an artifact. Since it was at the 10 percent level, one might conclude that the experiment gave no indication that the length or direction of a scale affects its sensitivity. Thus, space-saving short scales would appear to be as effective as more extended scales, regardless of whether they are represented vertically or horizontally and regardless of the variability among stimuli.

A difference in the mean ratings of a constant sample was suggested (again, only at the 10 percent level) between the horizontal and vertical scales. This difference was not unexpected. Peryam *et al* (1960, pp. 12-13) have shown that subjects using horizontal labelled scales rate stimuli more favorably when the favorable extreme is on the left (beginning) end of the scale. In the present experiment, the horizontal scales had the favorable extreme at the right while the vertical scales had the favorable extreme at the top. If one assumes the subjects read scales from left to right or top to bottom, then the cited findings suggest that there is a tendency for subjects to stay closer to the beginning point of a scale. If this were so, then in the present experiment one would expect that ratings on the favorable-unfavorable vertical scale would be more favorable than on the unfavorable-favorable horizontal scale -- and this was observed.

The theoretical basis for this study was the suggestion of others that increasing the psychological length and category density of fixed-interval scales increases the reliability with which stimuli are classified into scale categories. This is supposed to occur because scale intervals are more distinct. The hypothesis tested was that extending the physical length of fixed-interval scales might make categories more distinct and thereby enhance reliability of ratings. The data provided no reliable support of this hypothesis. Apparently, short scales having a given number of intervals and anchors have as much discriminatory power as do longer scales with the same number of intervals and anchors.

This "Level - of - Liking" food rating scale differs from the Preference scale usually used in the taste - test laboratory. Therefore please look this scale over carefully before you begin to taste the samples. For each sample, check the numbered box which best shows your "level - of - liking" for that sample.





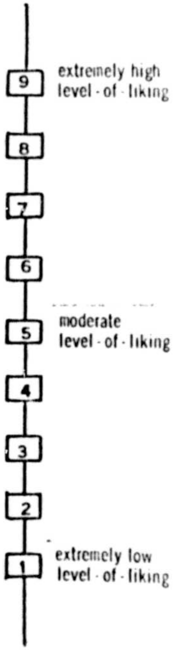


Figure 1. Examples of instructions, long horizontal scale, short horizontal scale, and short vertical scale.

References

1. Anderson, N. H. Scales and statistics: parametric and non-parametric. Psychol. Bull., 58: 305-316, 1961.

2. Bendig, A. W. The reliability of self-ratings as a function of the amount of verbal anchoring and of the number of categories on the scale. J. Appl. Psychol., 38: 38-41, 1954.

3. Bendig, A. W. Transmitted information and the length of rating scales. J. Exp. Psychol., 47: 303-308, 1954.

4. Bendig, A.W. Rater reliability and the heterogeneity of scale anchors. J. Appl. Psychol., 39: 37-39, 1955.

5. Bendig, A. W. and J. B. Hughes, II. The effect of the amount of verbal anchoring and number of rating-scale categories upon transmitted information. J. Exp. Psychol., 46: 87-90, 1953.

6. Coombs, C. H. Theory and methods of social measurement. In Festinger, L. and D. Katz (eds.), Research Methods in the Behavioral Sciences. Chapt. 2, New York, Dryden, 1953.

7. Eng, W. E. An experimental study of the reliability of rating scales for food preference discrimination. Unpublished M.A. thesis, Northwestern University, 1948.

8. Engen, T., N. Levy, and H. Schlosberg. The dimensional analysis of a new series of facial expressions. J. Exp. Psychol., 55: 454-58, 1958.

9. Erikson, C. W. and H. W. Hake. Absolute judgment as a function of stimulus range and number of stimulus and response categories. J. Exp. Psychol., 49: 323-32, 1955.

10. Fisher, R. A. and F. Yates. Statistical Tables for Biological, Agricultural, and Medical Research. Hafner Publ. Co., Inc., p. 2 and Table 5, New York, 1957.

11. French, R. S. The discrimination of dot patterns as a function of number and average separation. J. Exp. Psychol., 46: 1-9, 1953.

12. Garner, W. R. and H. W. Hake. The amount of information in absolute judgments. Psychol. Rev., 58: 446-59, 1951.

13. Guilford, J. P. Psychometric methods. McGraw Hill, N. Y., 1936.

14. Hunt, W. A. and J. Flanery. Variability in the affective judgment. Amer. J. Psychol., 51: 507-13, 1938.

15. Jones, L. V. Some invariant findings under the method of successive intervals. Amer. J. Psychol., 72: 210-20, 1959.

16. Jones, L. V., D. R. Peryam, and L. L. Thurstone. Development of a scale for measuring soldiers' food preferences. Food Res., 20: 512-20, 1955.

17. Peryam, D. R. and F. J. Pilgrim. Hedonic scale method of measuring food preferences. Food Technol., 11: (9), Supplement, 9-14, 1957.

18. Peryam, D. R., B. W. Polemis, J. M. Kamen, J. Eindhoven, and F. J. Pilgrim. Food Preferences of Men in the U.S. Armed Forces. Chicago: Quartermaster Food and Container Institute for the Armed Forces, 1960.

19. Volkman, J. Scales of judgment and their implications for social psychology. In Rohrer, J. H. & Sherf, M. (eds.). Social psychology at the crossroads. New York: Harper, Chapt. 11, 1951.

20. Volkman, J., W. A. Hunt, and M. McGourty. Variability of judgment as a function of stimulus density. Amer. J. Psychol., 53: 227-84, 1940.

21. Wodinsky, J. M. A. Varley, and M. E. Bitterman. Situational determinants of the relative difficulty of simultaneous and successive discrimination. J. Comp. Physiol., 47: 337-40, 1954.

DISTRIBUTION

Chief, Research and Engineering (1)
Department of the Army
The Pentagon, Washington 25, D.C.

Chief, Research Analysis Division (1)
Army Research Officer
Office, Chief, Research and Dev.
Department of the Army
Washington 25, D.C.

Research and Engineering Division (6)
Office of The Quartermaster General
ATTN: Services Office
Department of the Army
Washington 25, D.C.

The Quartermaster General (1)
ATTN: Charles N. Gardner
Department of the Army
Washington 25, D.C.

Commanding General (40)
ATTN: Tech. Information Branch
Tech. Services Division
QM R&E Command, U. S. Army
QM R&E Center
Natick, Mass.

Commanding Officer (1)
Field Evaluation Agency
QM R&E Command, U. S. Army
Ft. Lee, Va.

Commanding Officer (1)
U. S. Army Medical Research and
Nutrition Lab., Fitzsimons Gen. Hosp.
Denver Colorado (Food Reports)

The Library (1)
U. S. Army Leadership HRU
P. O. Box 787 Presidio of
Monterey, Calif. (Food Reports)

Defense Research Member (4)
Canadian Joint Staff
2450 Massachusetts Ave, N.W.
Washington, D.C.

Chief, Bureau of Supplies and Accounts (W/2)(1)
Department of the Navy
Washington 25, D.C.

Commanding Officer (1)
U.S. Navy Subsistence Office
Naval Weapons Plant
Washington 25, D.C.

Chief, Bureau of Supplies and Accounts (O/W)(1)
Department of the Navy
Washington 25, D.C.

Officer-in-Charge (1)
U. S. Naval Supply Center
Naval Supply Depot
Bayonne, New Jersey

Chief, Signal Corps Packaging Standards Ofc. (1)
U. S. Army Signal Supply Agency
Tobyhanna Signal Depot
Tobyhanna, Penn. (Container Reports)

Librarian (1)
QM Technical Library
Ft. Lee, Va.

Commandant (1)
Hq., U. S. Marine Corps
Washington 25, D.C. (O.I.C. Supply Bn.)

(3)
British Joint Services Mission (Army Staff)
British Embassy Annexe
Washington 8, D.C.

Major L. G. Clark (3)
Australian Military Mission
2001 Connecticut Ave., N.W. Box 4837
Washington 8, D.C.

QMFCIAF

Commandant, Assistant Commandant, Scientific Director, Deputy Scientific Director, Library, all office, division, and branch chief, Navy Liaison Officer - 1 each.
Air Force Liaison Officer (6)