

Scaling Interpersonal Checklist Items to a Circular Model

Clarence C. McCormick and Jack A. Kavanagh
Loyola University of Chicago

The items of the Interpersonal Checklist (ICL) were scaled to a circular model by two different procedures: (1) the items were sorted by two different samples into categories corresponding to the labels of the eight ICL octants and (2) each item was scaled twice by a third sample, first on the 9-point bipolar scale Hate-Love, and second on the 9-point bipolar scale Dominance-Submissiveness. The two sets of ratings were found to correlate $-.09$, indicating that the two postulated dimensions presumed to underly the circular order are orthogonal. The items were then plotted into the plane formed by using the two scales as axes. The circular scale values calculated for the two sorting procedures correlated $.95$. The scale values obtained from the sorting procedure and those from the two-dimensional procedure correlated $.89$. In general, the plotted items followed a circular order from close synonymy to antonymy and back. Several gaps on the circle were found, indicating an inadequate sampling of items. Scale values were calculated for each of the eight ICL scales using the circular scale values as item weights. When these values were plotted and compared with a factor plot of the eight ICL scales, the plots were remarkably similar. Many items were found to be displaced by the scaling procedures from the placements given by the authors of the ICL. Most of these displacements were found to be related to an intensity dimension postulated by the authors of the ICL. In general, the mild/moderate items were scaled toward the Love and Dominance poles, and the strong/extreme items were scaled toward the Hate and Submissive

poles, thus pulling the items away from the scales they were intended to represent. Some of the major implications of the use of these procedures in the construction of personality instruments are discussed.

A simple scaling technique can be used to scale personality items into a circular structure. One model for this structure has been presented by Guttman (1954). He labeled the circular structure a circumplex and argued that psychological tests and/or scales can be related to each other in a circular order according to the degrees of correlation among them. Table 1 presents a hypothetical intercorrelation matrix for six tests which make up an equally spaced uniform perfect circumplex. The rows and columns of any given empirical matrix of intercorrelations can be permuted by trial and error to examine for possible fit to the model. He also defined for empirical data a quasicircumplex as a circumplex plus deviations.

Empirical circumplexes for psychological variables are, in fact, being reported more and more often, and they seem to be found in almost every area of psychological assessment. Guttman (1957) first reported empirical circumplexes for Thurstone's tests of mental ability and for the Wechsler-Bellevue Intelligence Scales. The order found was similar for both batteries: Verbal, Numerical, Figural, and back to Verbal. Recently, Mukherjee (1975) has confirmed this

Table 1
Intercorrelations for One Possible, Equally-spaced,
Uniform, Perfect Circumplex*

Test	1	2	3	4	5	6
1	1.00	.75	.50	.25	.50	.75
2	.75	1.00	.75	.50	.25	.50
3	.50	.75	1.00	.75	.50	.25
4	.25	.50	.75	1.00	.75	.50
5	.50	.25	.50	.75	1.00	.75
6	.75	.50	.25	.50	.75	1.00
Total	3.75	3.75	3.75	3.75	3.75	3.75

*After Guttman (1954)

order with the Weschler Preschool and Primary Scale. Schaefer (1961) and Slater (1962) have reported a circumplex order for Minnesota Multiphasic Personality Inventory (MMPI) scales. Cole (1973) reported that four major vocational interest inventories—Strong, Kuder, Holland, and American College Testing Program (ACT)—each exhibit a similar circular order among their scales. Schaefer (1959) found a circumplex order among rating scales for maternal behavior, as did Roe and Siegelman (1963) and Slater (1962). Similar circumplexes have been reported for ratings of child behavior by Schaefer and Bayley (1963), Becker and Krug (1964), and Baumrind and Black (1967). Schaefer and Plutchik (1966) studied terms that refer to emotions and to personality traits and found that each of the structures formed similar circumplex orders among the variables.

Freedman, Leary, Ossorio, and Coffey (1951) developed a circular model for interpersonal behavior which formed the basis for the construction of the Interpersonal Checklist (ICL; La Forge & Sucek, 1955). The circular order for the scales of the ICL was confirmed by Rinn (1965). Lorr, Klett, and McNair (1963) developed a set of rating scales for psychotic behavior that followed the circumplex order, and Lorr and McNair (1963) developed an inventory of interpersonal behaviors on the same pattern. Finally, Stern (1970) found a circumplex order among the scales of his Activities Index, which

had been constructed to assess Murray's Need variables. The applicability of the model would seem to be quite general (e.g., aptitude and achievement tests, vocational interest inventories, and personality inventories). Wiggins (1968) summarized the situation most succinctly:

The reader of this considerable body of research is struck by several areas of apparent agreement. Despite substantial differences in universes of content, populations studied, media of observation, and test construction procedures, there are remarkable "convergences" in the conceptions underlying the variables employed and the structural relations found to obtain among them. (p. 320)

The empirical circular structures which have been presented to date have consisted of the intercorrelations among, or the factor loadings of, the scales of the various tests. For the most part the items used in the scales have been selected on the basis of the opinion of one or more experts. Item intercorrelations, or item-scale correlations, are then used as a basis for combining the items into scales. These intercorrelations are obtained from responses of people who use the items to describe themselves or others. The problem then exists of separating instrument variance from person variance. Test construction procedures might be better controlled if the test item characteristics could be operationally separated from use of the items. One possible

procedure would be to scale the items for placement prior to their use as measurement instruments.

The scaling of personality items (especially checklist items) has frequently been recommended (Guilford, 1954). Most frequently the scaling focuses on the question, how much?: How much more favorable is one item relative to another (Anderson, 1967)? or How much more abnormal and how much more intense (Buss & Gerjuoy, 1957)? That is, a linear continuum is assumed and judgments are oriented by this assumption. However, Guttman's circumplex involves a different type of order, an order of kind rather than one of degree. Concepts such as "greater than" or "less than" have no meaning on a circular continuum. The distinction is the same as that made by Stevens (1961) between prothetic and metathetic continua.

Differences of kind have generally been approached as classification problems (nominal scaling) in Stevens' (1951) scheme. The scores from most personality inventories are arranged as profiles of arbitrarily ordered nonoverlapping variables in what Cooley (1971) has called the parallel-stalk model. Guttman's circumplex model makes the assumption that at least some psychological variables can be *ordered* along some latent metathetic continuum.

Woodworth (1938) faced a similar problem in the study of judgments of emotional expression in posed pictures of human faces. He argued that the low agreement among judges found in previous studies was due to the fact that the judgments had been simply classified as right or wrong and that what was required was a scale to measure how far wrong an error of judgment might be. If an underlying continuum is assumed to exist, then much of the error could be due to trouble in discriminating between neighboring points on the continuum. By trial and error permutations of the frequencies of judgments, the order of the judgments was adjusted so that the frequencies ran a regular course both vertically and horizontally, with more overlap between adjacent than nonadjacent cells. In other words, the frequencies were permuted to

present the maximum correlations between pose and category. This adjusted order resulted in a six-step scale of emotional expression, where the judgments were very seldom displaced by more than one step.

Schlosberg (1941) used Woodworth's procedure to study the judgments of 45 subjects who sorted 72 pictures into bins labeled with the names of Woodworth's six scale intervals. In studying the resulting frequency distributions, he noted that by using the criterion of more overlap between adjacent than nonadjacent categories, the sixth (or last step) was more closely related to the first (as well as to the fifth) step than to the other steps of the scale. He argued that this implied that the scale was circular, rather than linear, as Woodworth had supposed. In a further study, Schlosberg (1952) presented an interesting validation of his results. He argued that the circular distributions implied that two orthogonal dimensions (Pleasantness-Unpleasantness and Attention-Rejection) underlay the judgments. In three separate studies, subjects sorted the pictures and rated each picture twice, using the postulated orthogonal dimensions as 9-point rating scales. Schlosberg then used the two dimensions as the axes and the ratings of each as the Cartesian coordinates to plot each of the pictures in the two-dimensional system. Finally, he correlated the angular placements of the pictures with the scale values obtained through the simple sorting procedure. The resulting correlations were gratifyingly high over the three studies (.94, .92, and .96, respectively). The fact that Schlosberg was able to obtain frequency distributions that were circular, while using a simple nominal category sorting procedure, encouraged the authors of this study to believe that similar results could be obtained with personality items.

Method

Subjects and Procedures

The 128 items and the 8 octants of the ICL (La Forge & Suczek, 1955) were selected for use as stimuli in this study because of the evidence that

the scales of this instrument exhibit a reasonably close approximation to a circular order (Rinn, 1965). The hypothesis of the present study was that if subjects were to sort the 128 items of the ICL into the randomly presented octant labels, frequency distributions that would exhibit ordered circularity would be obtained. If so, the items could then be supplied with circular scale values (degrees or radians) that would indicate their placements on the circle.

The items were scaled in three separate studies. In one, the following categories were used as the scale labels into which the items were sorted:

1. Docile-Dependent (JK)
2. Cooperative-Overconventional (LM)
3. Responsible-Overgenerous (NO)
4. Managerial-Autocratic (AP)
5. Competitive-Exploitive (BC)
6. Blunt-Aggressive (DE)
7. Skeptical-Distrustful (FG)
8. Modest-Self-effacing (HI)

Two terms were used for each category, the first term indicating a mild expression of the trait and the second indicating an extreme expression. However, some of the resulting frequency distributions exhibited bimodal distributions, which could have been due to some of the subjects focusing on one of the terms for item placement while others focused on the other term. Therefore, in the second study a second independent scaling was conducted using the following single-label categories:

1. Docile (JK)
2. Conventional (LM)
3. Responsible (NO)
4. Managerial (AP)
5. Competitive (BC)
6. Critical (DE)
7. Skeptical (FG)
8. Self-effacing (HI)

In the third study the ICL items were rated on two separate bipolar dimensions. Love-Hate and Dominance-Submission were set up as two 9-point scales with each of the points anchored by an adverb. The adverbs used at each of the scale divisions were: *Extremely, Strongly, Moderately,*

Mildly, Neutral, Mildly, Moderately, Strongly, Extremely.

The subjects were undergraduate and graduate students. In the first study (sorting the items using the two-term category labels) there were 94 females and 33 males. The second study (sorting items with the one-term category labels) employed a separate sample of 82 females and 27 males. The third study (rating the items on the two bipolar scales) employed another sample of 105 females and 34 males.

In the two sorting studies the subjects were given envelopes containing the eight category labels typed on 3" × 4" slips and the 128 items typed on 1" × 3" slips. The items were given random numbers to replace the ICL numbers. Both the items and the category labels had been well shuffled before placing them in the envelopes, in order to randomize their presentation order. The subjects were asked to take the eight large slips out of the envelopes, to shuffle them, and to place them face up in any order on their desks. They were then asked to take out the 128 items, to shuffle them, to read each item, and then to place it in the category in which it seemed best to belong. After the sorting, they were given a tally sheet to record the numbers of the items they had placed in each of the categories.

In the rating scale study the items were presented in random order down the left side of each page, with the dimension on which the items were to be rated at the top of each page. Subjects were instructed to judge to what extent the item seemed to be related to either of the poles of the bipolar dimension presented at the top of the page and then to place a check in the blank space of the appropriate column opposite that item. If the item did not seem to be related to the dimension, they were to place the checkmark in the Neutral column. The items were rated twice, once for each dimension. The procedure took about 20 minutes.

Analyses

The sorting task. The number of judges who sorted each item into each of the eight categories

was obtained. Items were then grouped into eight subsets, based on their modal assignment to a category (i.e., if more judges assigned Item 1 to the AP category than to any other category, it was included in the AP subset, along with all other items whose modal assignment was AP). The distributions of assignments to the eight categories were then averaged (i.e., an average frequency distribution of assignments to the eight categories was obtained for each of the eight item subtests).

The logic of this procedure may not be immediately obvious. It is assumed that the category labels are ordered along a circular continuum. In the ideal case it would be expected that each interval between adjacent labels was equal. If there are eight categories, it would then be expected that each interval would cover a 45-degree arc on the circle. Further, it would be expected that the items would vary around the central point in each interval. The points representing the items within each interval can be considered as the endpoints of vectors on a unit circle. If their value (in coordinates) are added together, and if the interval is well covered, the sum will represent the central point of the interval. The central vectors are the averages of the item vectors within each interval. In the empirical case, it is not known that the category labels are in fact equidistant, nor is it known that each interval is well covered with items. Some central tendency of the item distributions (say the mode, for convenience) can be considered to be an indicant of the interval into which the item is likely to scale. All the frequency distributions of those items that fall into each interval can then be averaged together. The mean of this average distribution will, by the logic of vector addition, be an estimate of the central point of the interval.

The eight average distributions can then be examined for fit to the circular model. If each of the distributions were exactly the same, the rows and columns (using relative frequencies) would sum to unity, the grand sum would equal the number of categories (k), and each interval or

category (column) sum would equal $1/k$. In the empirical case the column sums will not be exactly equal; however, each sum can be considered to reflect proportionate interval sizes. This matrix can thus be examined for fit to the criteria for a circumplex. The empirical interval sizes can be used to estimate the midpoint of each interval, which can in turn be used to calculate scale values for the items.

To calculate the circular scale values (Ross, 1938), appropriate sines and cosines were used as the coordinates of the midpoint of each of the eight intervals. These were then weighted by the category frequencies for each item, and the polar coordinates were calculated. An example of the procedure is presented in Table 2. Scale values were calculated using two different estimates of the midpoints of the intervals around the circle: (1) the midpoints estimated from the empirically obtained interval sizes, and (2) the midpoints estimated on the assumption that the interval sizes were equal. These two sets of scale values were correlated, as were the scale values obtained from each of the separate scalings. Finally, the structure of the two-dimensional space defined by the categories and the item scale placement was examined.

The rating scale study. The ratings for each of the items on each of the dimensions were averaged across all subjects. The average ratings of the items on each dimension were correlated to examine the degree of relationship between the two dimensions. The dimensions were then set orthogonal to one another and the average ratings on each were used as the coordinates for each of the items in this reference system. Polar coordinates (vector lengths and angles) were then calculated for each of the items for comparison with similar values calculated from the sorting procedure.

Finally, the items were grouped together as scales in accordance with the scale placement of the items as given in the ICL manual (Leary, 1956). The scaling values obtained from the sorting and rating procedures were used to calculate the positions of the scales relative to one another

Table 2
An Example of the Calculations for Obtaining Angular Scale
Values from a Circular Frequency Distribution

Item	Scale	f	cos	sin	x	y
122	NO		.380	.925		
	AP		.918	.397		
	BC	1	.938	-.348	.938	-.348
	DE		.403	-.915		
	FG	1	-.362	-.932	-.362	-.932
	HI	26	-.920	-.391	-23.920	-10.166
	JK	67	-.890	.457	-59.630	30.619
	LM	14	-.321	.947	-4.494	13.258
Total		109			-87.468	32.431

$$r = [(-87.468)^2 + (32.431)^2]^{1/2} = 93.287$$

$$R = 93.287/109 = .856$$

$$\theta = \tan^{-1}(-87.468/32.431) = 290.3^\circ$$

on the circle. These were then compared to the scale positions reported by Rinn (1965) from his factor analysis of the scales. The differences between the individual item placements as given in the ICL manual and as obtained here were also examined.

Results

Circular Scaling

Given that the basic purpose of this study was primarily to demonstrate the feasibility of the procedures rather than to be a definitive normative study, the frequency distributions for the males and females were combined. However, visual inspection of the distributions suggested that about six of the 128 items might exhibit some sex differences in their category assignments.

If the categories are arrayed so that maximum overlap of frequencies can be seen in adjacent columns and rows, the resulting frequency distributions exhibit a circular order. These average frequency distributions are presented in Table 3 (for both sets of scalings) as *relative* frequencies for ease of comparison. The individual item distributions are available in McCormick (1977).

Examination of Table 3 indicates that both matrices meet the criteria listed by Guttman (1954) for a quasicircumplex: (1) the largest values lie along the main diagonal, (2) the values decrease and then increase again along the diagonal running toward the lower left and upper right corners, (3) the column sums tend to be equal, and (4) the value of any one cell tends to be similar to the value of the cell in the column and row prior to and above it, with a carry-around from the last to the first columns. One of the most striking aspects of the matrices, given the random-order presentation of the categories, is that they are approximately symmetrical. Further, the assumption of a latent continuum seems reasonable in that the frequencies show a regular rise and decline around the mode. Notably this technique results in the eight octant labels showing the same circular order as postulated for them by the authors of the ICL. Inspection of the eight distributions shows that some of them (e.g., JK, NO, DE) were quite symmetrical and that others were more skewed (e.g., LM, BC, FG, HI).

Since the order for the categories is demonstrably circular, scale values can be calculated by the procedure illustrated in Table 2. As noted above, in an equally spaced circulant matrix the column sums can be expected to be equal. In

Table 3
Average Item Distributions (in Relative Frequencies)

Item Subsets Whose Modal Assignment Was to This Category	Average Distributions of Assignments to Each of the Eight Categories for the Items Modally Assigned to That Category							
Two Term Categories (N = 127)								
Scale	NO	AP	BC	DE	FG	HI	JK	LM
NO	.52	.04	.02	.01	.01	.14	.09	.18
AP	.14	.49	.11	.14	.02	.04	.01	.05
BC	.02	.13	.50	.17	.11	.04	.01	.02
DE	.03	.15	.19	.44	.15	.02	.01	.01
FG	.01	.05	.10	.13	.63	.03	.04	.01
HI	.10	.01	.02	.02	.02	.54	.11	.07
JK	.05	.01	.01	.01	.03	.16	.63	.10
LM	.21	.02	.02	.01	.01	.11	.15	.47
Total	1.08	.90	.97	.93	.97	1.08	1.16	.91
One Term Categories (N = 109)								
NO	.50	.13	.06	.03	.01	.06	.07	.14
AP	.08	.58	.16	.12	.01	.03	.00	.02
BC	.01	.10	.49	.15	.08	.13	.01	.03
DE	.01	.05	.11	.53	.20	.07	.01	.02
FG	.04	.04	.09	.18	.54	.07	.02	.02
HI	.04	.02	.04	.05	.06	.46	.25	.08
JK	.06	.02	.01	.00	.00	.18	.60	.12
LM	.26	.04	.02	.00	.01	.11	.16	.40
Total	1.00	.98	.98	1.06	.92	1.11	1.12	.83

terms of the proportions presented in Table 3, it would be expected that each of these sums would equal 1.00. As can be seen, the actual proportions do not vary much from the expected value of unity. If the original matrix of frequencies is considered, the column totals should, for equal intervals, each be 1/8 (.125) of the grand total. If these proportions are considered in terms of arcs on the circle in degrees, each interval would be 45 degrees in length. The actual interval lengths, as calculated for the one-term sorting procedure, varied from 37 (LM) to 51 (JK).

Since interval lengths were not very far (at most 6 degrees) from 45 degrees for any of the categories, the midpoints of these intervals were

taken as the scale values for the ordered set of categories. The sines and cosines for these midpoints can be seen in Table 2. These were weighted by the frequency of item placements in each of the categories to calculate the resultant as the scale value for that item, as shown in Table 2. Scale values were also calculated using exactly equal interval midpoints. Correlations between the values calculated using the equal interval midpoints and those calculated from the empirically determined interval size were obtained. For the double-labeled data, the product-moment correlation between the two sets of scale values was .993; for the single-labeled sets, the correlation was .9998. The means and standard deviations of the differences were (1)

-9.2 ± 3.8 for the double-labeled set and (2) -2.8 ± 2.0 for the single-labeled set. Clearly, both the correlations and the average differences indicate that the empirical interval sizes were exceptionally close to the equal interval values. However, the single-labeled set seems to be, just as clearly, much closer to equal intervals than the double-labeled data.

The correlation between the scale values calculated from the double-labeled set of categories and those from the single-labeled set was .95. Clearly, the angular values calculated for the items were exceptionally similar between the two studies. The major effect of the use of the double labels was to increase the variability of the distributions; its effect on the ordering of the items around the circle was, in general, a very small one. This high a correlation between the results of the two studies suggests that this scaling procedure exhibits very high reliability. Given the near identity of the results between the two studies, only the values calculated for the empirically determined midpoints of the single-valued study will be reported here.

Table 4 presents the calculated scale values for each of the 128 ICL items, listed according to their octants as given in the ICL manual (Leary, 1956). The empirically calculated interval size of each of the categories is presented in the row below the category labels. As noted above, the vector length divided by N can be considered an index of the variability of the scale values. These values are presented in Table 5.

Consideration of the details of these tables is facilitated by presentation of the results of the two-dimensional scaling study. As noted above, the 128 items of the ICL were also scaled by a third group of subjects on two bipolar 9-point scales, using the dimensions of Love-Hate and Dominance-Submission. Following Schlosberg's (1952) argument, this study was conceived as a validation study of the sorting procedure. The 9 points of both scales were supplied with numbers ranging from -4 to $+4$ with 0 at the center or neutral point. The mean scale scores were then calculated for each of the 128 items. These

are presented in Tables 6 and 7. Positive values indicate the Dominance and Love poles, respectively. The numbers reflect the rated intensity of each item along each of the dimensions. The last row of each table gives the average intensity rating for the 16 items in each octant.

Examination of Tables 6 and 7 reveals a pattern that suggests the basic circular structure of these data. The pattern is seen most clearly in the row of column averages; they appear to behave approximately like sine and cosine curves. For the Dominance dimension (Table 6) the highest average is seen for the AP (Managerial) scale, as would be expected. The values gradually decline, pass through the neutral point, become most negative for the HI (Self-effacing) scale, and then increase once again; octant HI is presumed to be diametrically opposite AP. The pattern is similar for the Love dimension (Table 8), although the antimode is not exactly where it would be expected (FG rather than DE). Note, however, that the intervals between the categories do not appear to be equal, given the unequal differences among the averages.

Although the orthogonality of the two dimensions has been previously demonstrated by factor analysis (Rinn, 1965), it can also be examined by correlating the two sets of scale values across the 128 items. The correlation between the two sets of values was $-.09$, a value not significantly different from zero; thus, this study supplies independent evidence that the two dimensions are indeed orthogonal, as postulated by the authors of the ICL.

Given the orthogonality of the dimensions, they can be used as the axes in the Cartesian coordinate system. The Dominance-Submission scale was taken to be the ordinate and the Love-Hate scale as the abscissa, and the items were plotted in the plane using the mean scale values as the coordinates. The angular placements of the items were obtained by calculating $\arctan y/x$; and the vector lengths, by taking the square root of $(x^2 + y^2)$. The angles and vector lengths are presented in Tables 8 and 9, respectively. The angular values calculated from the sorting

Table 4
Correspondence Between Item-Scale Assignment as Provided in the
I.C.L. Manual and that Obtained by the Scaling Procedure

ICL "Intensity"	NO (0-45)		AP (45-89)		BC (89-132)		DE (132-180)		FC (180-222)		HI (222-272)		JK (272-323)		LM (323-360)	
	Item	Angle	Item	Angle	Item	Angle	Item	Angle	Item	Angle	Item	Angle	Item	Angle	Item	Angle
Mild	29	6*	1	13	5	13	9	49	13	21	17	32	21	334	25	5
	31	16*	3	59*	7	28	11	38	15	248	19	335	23	354	27	349*
Moderate	30	26*	2	20	6	48	10	40	14	214*	18	271*	22	298*	26	323*
	32	341	4	90	8	146	12	159*	16	217*	20	289	24	273*	28	356*
	61	1*	33	31	37	48	41	67	45	125	49	273	53	297*	57	318*
	62	308	34	28	38	38	42	50	46	199*	50	253*	54	331	58	277
	63	14*	35	46*	39	63	43	175*	47	163	51	296	55	331	59	348*
	64	345	36	42	40	111*	44	46	48	240	52	313	56	306*	60	350*
Strong	93	288	65	109	69	125*	73	152*	77	181*	81	233*	85	298*	89	290
	94	289	66	99	70	62	74	129	78	169	82	286	86	293*	90	280
	95	287	67	95	71	152	75	133*	79	151	83	294	87	293*	91	303
	96	337	68	97	72	121*	76	168*	80	174	84	292	88	287*	92	310
	125	290	97	105	101	132*	105	174*	109	177	113	291	117	295*	121	272
	127	282	99	68*	103	148	107	165*	111	162	115	290	119	290*	123	318
Extreme	126	316	98	97	102	141	106	158*	110	170	114	252*	118	284*	122	290
	128	281	100	94	104	157	108	146*	112	191*	116	279	120	284*	124	301

*Items whose scale values fall into the I.C.L. octants to which they were initially assigned.

Table 5
 Variability Index (R=r/n) for the ICL Items
 (One-term sorting procedure)

ICL "Intensity"	NO		AP		BC		DE		FG		HI		JK		LM	
	Item	R														
Mild	29	.82	1	.80	5	.67	9	.78	13	.44	17	.29	21	.65	25	.80
	31	.84	3	.91	7	.78	11	.61	15	.43	19	.75	23	.66	27	.74
Moderate	30	.73	2	.74	6	.68	10	.83	14	.70	18	.81	22	.66	26	.55
	32	.56	4	.82	8	.44	12	.95	16	.58	20	.90	24	.52	28	.73
	61	.70	33	.76	37	.62	41	.75	45	.51	49	.77	53	.80	57	.75
	62	.75	34	.84	38	.72	42	.75	46	.94	50	.75	54	.71	58	.57
	63	.63	35	.85	39	.92	43	.72	47	.60	51	.88	55	.68	59	.81
	64	.57	36	.87	40	.96	44	.65	48	.66	52	.70	56	.72	60	.74
Strong	93	.84	65	.57	69	.67	73	.77	77	.79	81	.68	85	.86	89	.88
	94	.72	66	.58	70	.44	74	.39	78	.84	82	.88	86	.90	90	.78
	95	.58	67	.74	71	.55	75	.65	79	.56	83	.89	87	.90	91	.75
	96	.29	68	.76	72	.66	76	.73	80	.78	84	.94	88	.86	92	.74
	125	.80	97	.74	101	.47	105	.86	109	.70	113	.93	117	.87	121	.73
	127	.80	99	.93	103	.51	107	.80	111	.64	115	.90	119	.89	123	.69
Extreme	126	.60	98	.53	102	.57	106	.79	110	.77	114	.84	118	.87	122	.86
	128	.74	100	.71	104	.67	108	.60	112	.86	116	.79	120	.89	124	.74

Table 6
Means of ICL Items Rated on the 9-Point Bipolar
Scale Dominance-Submission (+4 to -4)

ICL "Intensity"	NO		AP		BC		DE		FG		HI		JK		LM	
	Item	Mean	Item	Mean	Item	Mean	Item	Mean	Item	Mean	Item	Mean	Item	Mean	Item	Mean
Mild	29	.19	1	.70	5	1.19	9	1.26	13	.55	17	.58	21	-.62	25	.26
	31	.27	3	1.78	7	1.35	11	.94	15	.98	19	-.32	23	.07	27	.32
Moderate	30	.76	2	.91	6	1.70	10	1.38	14	-1.01	18	-1.76	22	-1.46	26	-.87
	32	-.08	4	2.19	8	1.37	12	1.32	16	-1.15	20	-2.21	24	-1.94	28	.24
	61	.35	33	1.24	37	1.78	41	1.62	45	1.83	49	-1.70	53	-1.45	57	-.28
	62	-.47	34	.99	38	1.94	42	1.38	46	.26	50	-2.56	54	-1.06	58	-.99
	63	.27	35	1.86	39	1.70	43	.78	47	1.77	51	-2.32	55	-.70	59	.14
	64	-.29	36	2.09	40	2.05	44	1.70	48	-1.46	52	-.87	56	-.97	60	.27
Strong	93	-2.19	65	1.68	69	1.67	73	1.59	77	.34	81	-1.88	85	-1.72	89	-2.55
	94	-1.72	66	1.63	70	1.78	74	1.25	78	.19	82	-1.86	86	-2.45	90	-2.06
	95	-1.09	67	2.29	71	1.72	75	2.09	79	.80	83	-2.45	87	-2.76	91	-.73
	96	.95	68	2.63	72	2.15	76	.95	80	1.52	84	-1.91	88	-2.12	92	-.53
	125	-1.86	97	1.71	101	1.25	105	1.38	109	.32	113	-2.27	117	-2.01	121	-1.04
	127	-1.86	99	2.27	103	1.46	107	1.14	111	1.94	115	-2.61	119	-1.98	123	-.30
	Mean															
Extreme	126	-.14	98	1.45	102	1.94	106	1.32	110	1.46	114	-2.56	118	-2.80	122	-2.23
	128	-.81	100	2.64	104	1.25	108	1.91	112	.09	116	-2.57	120	-2.30	124	-.66
	Mean	-.48		1.76		1.64		1.38		.53		-1.83		-1.64		-.69

Table 7
Means of ICL Items Rated on the 9-Point Bipolar
Scale Love-Hate (+4 to -4)

ICL "Intensity"	NO		AP		BC		DE		FG		HI		JK		LM	
	Item	Mean	Item	Mean	Item	Mean	Item	Mean	Item	Mean	Item	Mean	Item	Mean	Item	Mean
Mild	29	1.91	1	1.47	5	1.73	9	.96	13	.76	17	1.17	21	1.59	25	1.69
	31	1.95	3	.56	7	1.43	11	1.35	15	.06	19	.50	23	1.79	27	2.06
Moderate	30	1.76	2	1.29	6	1.30	10	.94	14	-1.29	18	.21	22	.22	26	1.24
	32	2.17	4	-1.19	8	-0.96	12	-1.52	16	-0.50	20	-0.04	24	-0.10	28	2.30
	61	2.04	33	1.25	37	1.39	41	.20	45	-0.30	49	-0.32	53	.06	57	1.48
	62	.89	34	1.47	38	1.04	42	1.10	46	-0.62	50	-0.83	54	1.15	58	.21
	63	1.99	35	1.31	39	-1.12	43	-1.43	47	-0.70	51	-0.35	55	.91	59	1.68
	64	2.16	36	.94	40	.17	44	1.21	48	-0.78	52	.70	56	1.51	60	2.04
Strong	93	-.86	65	-.56	69	-1.31	73	-1.40	77	-2.22	81	-1.19	85	.10	89	-.40
	94	.24	66	-.97	70	-.15	74	-.09	78	-1.48	82	.10	86	-.53	90	-.08
	95	.94	67	-1.83	71	-1.99	75	-2.24	79	-1.45	83	-.22	87	-.42	91	1.30
	96	.24	68	-.77	72	-1.42	76	-1.63	80	-1.64	84	.24	88	-.55	92	1.47
	125	.03	97	-.91	101	-1.42	105	-1.95	109	-1.91	113	-.12	117	.18	121	.22
	127	.39	99	-.60	103	-1.95	107	-2.04	111	-1.08	115	-.35	119	-.03	123	1.55
	126	1.47	98	-.63	102	-2.08	106	-2.91	110	-2.12	114	-1.59	118	-.70	122	-.27
128	.66	100	-1.81	104	-2.22	108	-2.00	112	-2.07	116	-1.25	120	.02	124	2.04	
Mean	1.12		.00		-.41		-.59		-1.08		-.21		.32		1.16	

Table 8
Correspondence Between Item-Scale Assignment as Provided in the
I.C.L. Manual and that Obtained by the Scaling Procedure

ICL "Intensity"	NO (0-45)		AP (45-89)		BC (89-132)		DE (132-180)		FG (180-222)		HI (222-272)		JK (272-323)		LM (323-360)	
	Item	Angle	Item	Angle	Item	Angle	Item	Angle	Item	Angle	Item	Angle	Item	Angle	Item	Angle
Mild	29	6*	1	26	5	34	9	53	13	36	17	26	21	339	25	9
	31	8*	3	72*	7	43	11	35	15	86	19	328	23	2	27	9
	30	23*	2	35	6	53	10	56	14	218*	18	277	22	279*	26	325*
	32	358	4	95	8	125*	12	139*	16	248	20	269*	24	267	28	6
Moderate	61	10*	33	45*	37	52	41	83	45	105	49	260*	53	273*	57	349*
	62	332	34	34	38	61	42	51	46	157	50	252*	54	317*	58	282
	63	8*	35	55*	39	94*	43	151*	47	112	51	262*	55	323*	59	5
	64	352	36	66*	40	85	44	55	48	242	52	309	56	327	60	7
	93	292	65	108	69	128*	73	131	77	171	81	238*	85	273*	89	261
	94	278	66	121	70	95*	74	94	78	172	82	273	86	258	90	268
Strong	95	311	67	129	71	139	75	96	79	151	83	265*	87	261	91	331*
	96	76	68	106	72	124*	76	150*	80	137	84	277	88	255	92	340*
	125	271	97	117	101	139	105	145*	109	171	113	267*	117	275*	121	282
	127	282	99	105	103	143	107	151*	111	119	115	262*	119	269	123	349*
	126	354	98	113	102	137	106	156*	110	145	114	238*	118	256	122	263
Extreme	128	309	100	124	104	151	108	136*	112	178	116	244*	120	270	124	342*

*Items whose scale values fall into the I.C.L. octants to which they were initially assigned.

Table 9
 Vector Lengths (Resultants of Intensity Ratings)
 From the Two-Dimensional Study

ICL "Intensity"	NO		AP		BC		DE		FG		HI		JK		LM	
	Item	R														
Mild	29	1.92	1	1.63	5	2.10	9	1.59	13	.94	17	1.31	21	1.71	25	1.64
	31	1.97	3	1.86	7	1.97	11	1.64	15	.98	19	.60	23	1.79	27	2.09
Moderate	30	1.92	2	1.58	6	2.14	10	1.67	14	1.64	18	1.77	22	1.48	26	1.52
	32	2.17	4	2.20	8	1.67	12	2.01	16	1.25	20	2.21	24	1.94	28	2.32
	61	2.07	33	1.76	37	2.26	41	1.63	45	1.90	49	1.73	53	1.45	57	1.52
	62	1.00	34	1.78	38	2.23	42	1.77	46	.67	50	2.69	54	1.56	58	1.02
	63	.93	35	2.28	39	1.70	43	1.63	47	1.90	51	2.34	55	1.15	59	1.69
	64	2.18	36	2.29	40	2.06	44	2.09	48	1.66	52	1.12	56	1.80	60	2.06
Strong	93	2.35	65	1.78	69	2.12	73	2.12	77	2.24	81	2.23	85	1.72	89	2.58
	94	1.74	66	1.89	70	1.79	74	1.25	78	1.49	82	1.87	86	2.50	90	2.06
	95	1.43	67	2.93	71	2.63	75	2.10	79	1.65	83	2.46	87	2.79	91	1.50
	96	.98	68	2.74	72	2.58	76	1.88	80	2.24	84	1.93	88	2.19	92	1.57
	125	1.86	97	2.02	101	1.90	105	2.39	109	1.94	113	2.28	117	2.02	121	1.06
	127	1.90	99	2.35	103	2.44	107	2.33	111	2.22	115	2.64	119	1.98	123	1.58
Extreme	126	1.48	98	1.58	102	2.84	106	3.19	110	2.57	114	3.01	118	2.89	122	2.25
	128	1.04	100	3.20	104	2.55	108	2.76	112	2.07	116	2.86	120	2.30	124	2.15

procedure (Table 4) correlated .89 with those calculated from the two-dimensional procedure (Table 8), in the same range as those reported by Schlosberg (1952) in his comparisons of the two procedures.

The mean difference in the angular placements of items between the two scaling procedures was 3.5 degrees. If two exceptionally disparate items (Items 15 and 96) are dropped from the calculations, the average difference reduces to 1.9 degrees. This indicates that calculating angles from the beginning of the NO category in the sorting procedure is equivalent to calculating angles from the Love pole of the x -axis. If the circle had been started, at say, AP, the average difference between the placements would have been about 45 degrees.

The sizes of the absolute differences in angular placement values are also of interest. Frequency distributions of these differences are presented in Table 10. For the sorting and two-dimensional scaling (column 1) procedures, about two-thirds of the items were within a half-category of one another, and almost 98% of the items were placed within one category-interval.

A similar comparison can be made for the item placements as determined by the two scal-

ing procedures, and those from the original ICL construction procedures. The items, as placed by the authors of the ICL, can be viewed as lying on vectors at the center of each of the eight intervals and can be supplied with the appropriate angular values. Differences between the scaling and the original placements can then be calculated. The frequency distributions of these differences are also presented in Table 10. Only about one-third of the items are placed within half a category-interval of one another, and only about two-thirds lay within one category-interval of one another. About one-third of the items show discrepancies of placement larger than a full category interval. Clearly, item placement agreement is considerably poorer in comparing the scaling placements with the original ICL placements than in comparing the scaling placements with one another.

If the scaling values are examined, the larger discrepancies between the scaling and the original ICL placements are found to be systematically related to one aspect of the original test construction procedures, namely, the intensity dimension. Prior to its publication the ICL went through four revisions. During one of the several tryouts of the ICL, La Forge and Suczek

Table 10
Frequency Distributions of the Absolute Differences
Between the Item Placements Given by the
Sorting and Two-Dimensional Scaling Procedures
and the Original ICL Placements

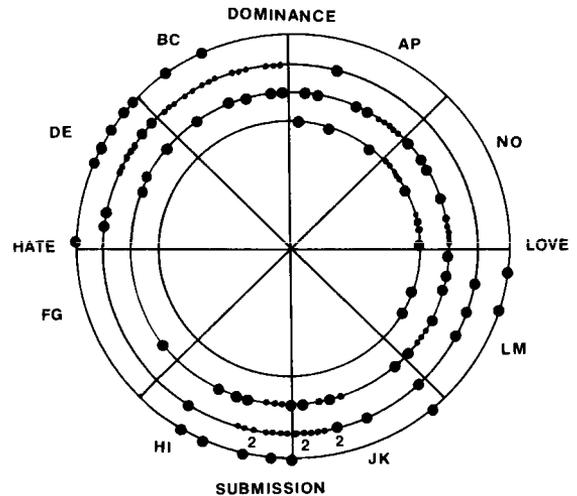
Angular Differences	Sorting-2D		ICL-Sorting		ICL-2D	
	f	%	f	%	f	%
0-22.5	87	67.9	48	37.5	44	34.4
22.5-45.0	38	29.7	43	33.6	42	32.8
45.0-67.5	1	0.8	18	14.0	17	13.3
67.5-90.0	0	0.0	5	3.9	13	10.1
90.0-112.5	1	0.8	10	7.8	8	6.2
112.5-135.0	0	0.0	2	1.6	2	1.6
135.0-157.5	0	0.0	1	0.8	1	0.8
157.5-180.0	1	0.8	1	0.8	1	0.8

(1955) had noted that their subjects were “. . . checking more words on the right hand (friendly) side of the circle than on the left (hostile).” To try to correct for this bias, intensity ratings for each of the items were established on the basis of frequency of response by the subjects. They used a rough rule that items of mild intensity (1) should be answered “yes” by about 90% of the subjects; those rated moderate (2), by about 67%; those rated strong (3) by about 33%; and those rated extreme (4), by about 10%. Intercorrelations among items of the same frequency were studied. The final form was developed in conference (four to six psychologists), where wordings were modified, items were dropped, new items were added, and so forth. The purpose of these procedures was to try to obtain a balance among the categories in the frequency of yes responses.

There would be perfect agreement between the item scale values in Tables 4 and 8 and the original ICL octants if all the angular values in each column clustered around the central vector for that octant. Although this result is clearly not obtained for most of the items, the discrepancies are not random throughout the table. Rather, many of the items seem to be displaced in groups. For example, most of the strong/extreme items in the lower half of the NO column all fall in the JK interval. Conversely, some of the mild/moderate items in the upper half of the BC column fall in the NO interval. The consistencies underlying these displacements are best seen in Figure 1.

Figure 1 displays the item placements from the two-dimensional scaling. Each of the four concentric rings shows the scaling placements of items at one of the intensity levels as determined by the authors of the ICL. The inner ring shows the placements of the mild (high frequency of endorsement) items; the second ring shows the moderate; the third ring, the strong; and the outer ring, the extreme items. Each of the rings represents a unit circle; their distance from the origin of the figure is not related to the vector lengths calculated from the intensity ratings of

Figure 1
Angular Item Placement of the 128 Items
of the ICL from the
Two-Dimensional Scaling Procedures
(See Text for Explanation of the Four Rings)



this study. The circles have been segmented into eight equal intervals labeled to orient the reader to the relationships between the two-dimensional scaling and the categories of the sorting procedure.

On the inner ring, almost all of the mild items have been scaled into the first quadrant; there are none of these items located in octants BC through JK. The sparseness of items in these octants on the second ring shows that the same trend exists for the moderate items. An approximate converse is found for the extreme and strong items on the outer and third rings. There are no strong items in the AP and NO octants; rather, they are located in the DE and HI octants.

Seemingly, then, the mild/moderate items of the ICL have been scaled toward the Love and Dominance poles, and the extreme/strong items have been scaled toward the Hate and Submissive poles. The ratings have thus pulled the items away from the placements given by LaForge and Suczek (1955) on the ICL—often to the opposite side of the circle. Exactly how this

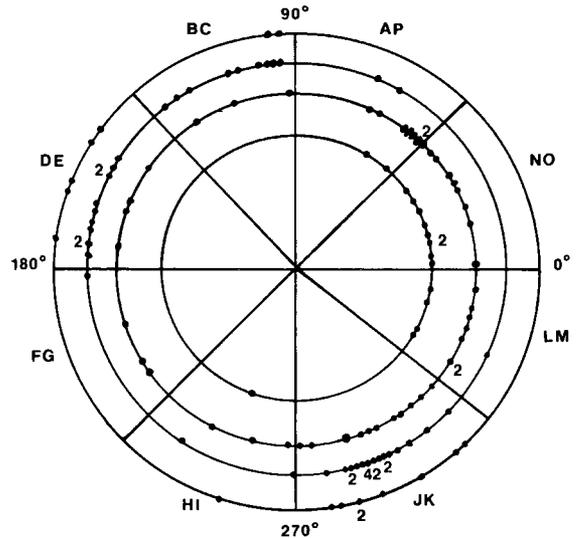
happens can be seen by examining the average ratings in Tables 6 and 7. Item 17 (Able to criticize self) can be considered as an example of a mild item. In order for the scale value to agree with the original ICL placement, the item should have received negative ratings on both dimensions; and, further, the Submission-Dominance rating should have been approximately in the strong interval (a rating of about 3), while the rating on the Hate-Love axis would have had to lie in the mild interval (a rating of about 1). The ratings actually received were the opposite of these (.58 on Dominance, 1.17 on Love); both are positive, and the rating on the Hate-Love axis is more extreme than that on the Submission-Dominance axis. Thus, the scaling placed the item in the NO octant.

As an example of an extreme item, Item 100 (Dictatorial) in the AP scale should be rated around neutral on the Love-Hate dimension and given an extreme rating near the Dominance pole. The actual ratings were 2.64 on Dominance and -1.91 (toward Hate). Although the rating on the Submission-Dominance axis is as expected, subjects tended to move the item toward the Hate pole of the Hate-Love axis, placing the item in the BC octant. Thus, the discrepancies between the original ICL placements of items on the circle and the scaling placements can be traced to differences in the intensity ratings along one or the other, or both, of the dimensions.

The same general trend can be seen in the data from the sorting procedure. These placements are given in Figure 2. As can be seen, the mild items are almost entirely absent from the inner ring in the BC through HI octants and are, instead, located in the AP, NO, and LM octants; the extreme items are absent from AP and NO and, for the most part, are located in DE and JK. Thus, sorting the items into the ICL octants gives much the same structure as that obtained by the two-dimensional rating procedure. In the sorting procedure, the differences cannot be traced to the intensity ratings; it can, however, be noted (from Table 4) that Item 17 is displaced

Figure 2

Angular Placements of the 128 Items of the ICL from the One-Term Scaling Procedure (See Text for Explanation of the Four Rings)



from the HI to the NO octant and that Item 100 is displaced from the AP to the BC octant, just as in the two-dimensional study. Seemingly, the subjects in the present study seemed to be interpreting these items quite differently than was intended by La Forge and Suczek (1955).

Although the item scale values given by the two scaling procedures are generally quite similar and thus correlate very highly, a comparison of Figures 1 and 2 shows some important differences. For example, the items in Figure 1 that are spread out over the HI and JK octants seem to be shifted by the sorting procedure (Figure 2) entirely into octant JK. This difference is probably a function of the selection of category labels used in the sorting procedure, an issue to be developed further below.

Another aspect of interest in Figures 1 and 2 is the difference in item concentrations around both circles. For example, the third quadrant, especially the FG octant, is largely devoid of items. If the circle is conceived of as a sample space, this segment is not well represented by the sample of items used in the present ICL.

There are other, smaller gaps in the circle as well (e.g., the larger segment around the center of JK; the last half of the DE octant in Figure 1). One consequence of an uneven distribution of items in a circular structure is that if the items are combined into scales, the resultant, like any average, will be affected by the unequal densities of the items. If the items in octant DE (Figure 1) were combined into a scale, the mean value would be toward the top of the octant rather than at its center. If this operation were performed for each of the scales, an uneven spacing of the scales would be found around the circle. These unequal intervals between the scales could be perceived as “gaps” in the circular structure. In fact, this is exactly what has been found in factor analytic studies of the ICL scales (Lorr & McNair, 1963; Rinn, 1965; Stern, 1970). These gaps are generally attributed to failure to incorporate some additional dimensions as scales in the ICL. For example, Stern (1970) argued that scales for Introversion and Extraversion would fill a couple of the gaps in the ICL structure.

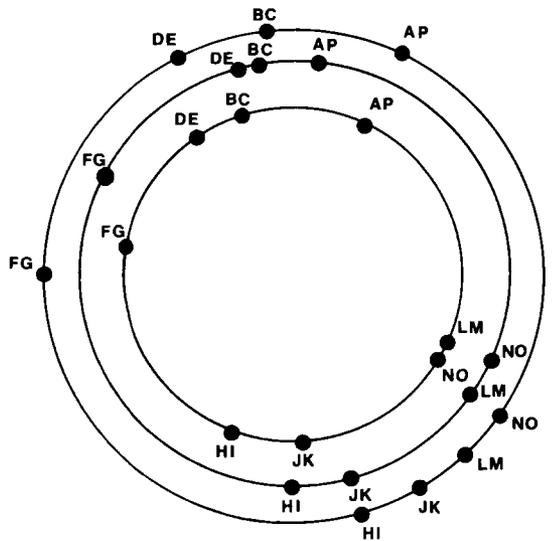
Although the “interpersonal circle” may indeed be inadequately conceptualized by the present ICL scales, the gaps between the scales found in the factor analytic studies might, for the most part, be determined by the relative lack of homogeneity among the items of each of the scales, as seen in Table 4 and 8. That is, when the ICL is used to describe the self and the items are summed into octant scores, including items with quite disparate angular values in each of the scales may function to pull some of them apart and to pull others closer together. To examine this possibility the resultants of the items were calculated for each of the eight octants for both scaling procedures. The results are presented in Figure 3. The outer ring presents the results from the sorting procedure, and the middle ring those from the two-dimensional data. The inner ring presents the findings by Rinn (1965).

The correspondence between the scale placements for the three sets of data seems remark-

able, especially given the differences in samples, data collection, and data analysis between the procedures. As Stern (1970) noted, the basic structure shows two clusters of scales with their centroids approximately 180 degrees apart, so that the impression gained is that there may be only a single dimension. The gaps between the NO and AP scales and between the FG and HI scales are those which Stern felt could be filled by adding scales for Introversion and Extraversion.

However, the data suggests that the primary basis for the gaps lies in the heterogeneity of the items that were placed into the ICL octants, relative to their individual angular weightings as determined by the scaling procedures. To underscore the effects of the item placements, the items have been regrouped into homogeneous scales, based on the scaling results. For the sorting procedure, the items were rank-ordered by their angular scale values and were placed into

Figure 3
 Placement of the ICL Scales on the Circle Inner Circle: Plot of Factor Headings (Rinn, 1965); Middle Circle: Plot of Resultant Angles (2D Scalings, Table 8); Outer Circle: Plot of Resultant Angles (Sorting Scaling, Table 4)



scales on the basis of being included in one of the empirically determined intervals around the circle. The resultants of each of these sets were then calculated to give the scale placements for each of the eight groups of items. A similar procedure was followed for the two-dimensional scaling data, except that the intervals were determined by taking successive 45 degree intervals beginning at the Love pole. The results are presented in Figure 4; the outer circle is based on the sorting study; and the inner circle, on the two-dimensional data.

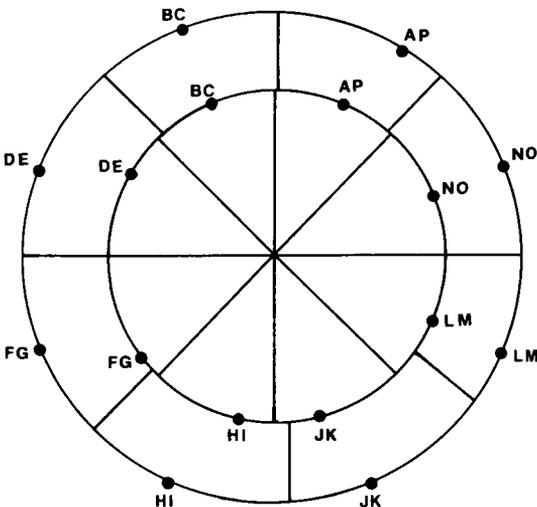
The most obvious result is that the scales now range around the full circle, rather than being polarized into two groups. The two large gaps have disappeared. The basis for the differences between Figures 3 and 4 is the rearrangement of the items in the scales. It seems clear that the gaps were *not* due to lack of adequate categories, as inferred by Lorr and McNair (1963) and Stern (1970); rather, they seem to be a result

of combining items that do not cluster closely enough together to give homogeneous scales. If Rinn's (1965) data were rescored in accordance with these scaling data, the factor plot should then appear much like Figure 4, rather than Figure 3. This inference is, of course, readily testable.

However, examination of the scale placements in Figure 4 in closer detail indicates that some gaps between the scales still exist. For example, the placements of DE and FG from the two-dimensional scaling (inner circle) are not located at the centers of their respective intervals, leaving a rather large gap between them. Moreover, this gap, as well as the others, is largely due to an inadequate sample of items in the ICL. Reference to Figure 1 makes this clear. The FG scale placement is based on only one item, and that one is located far from the center of the octant. If there were more items available in the octant, and if they were distributed evenly, the FG scale would be more centrally located. Examination of the DE interval in Figure 1 demonstrates the importance of an even distribution of items in the octant. Although there are 20 items in this octant, most of them are located in the upper half of the interval, and their average (the resultant) is, of course, located toward the segment of higher density. The same effect can be observed for the HI and JK sectors of the circle. These data lead to the conclusion that the major problem with the ICL in its present form is not an inadequate number of categories but an inadequate sampling of items from the total content domain. Specifically, the ICL lacks enough items covering withdrawn types of behavior (Maccoby & Masters, 1970).

When the items are listed in their rank order, the items strongly reflect an underlying circular *continuum*. The meanings of the items change gradually away (in both directions) from any item which is initially focused upon, just as colors shade gradually one into another around the color circle. Table 11 is presented as an example. The items are those which scaled into the BC (Competitive) interval in the sorting proce-

Figure 4
Resultants of the Scaled Items Grouped into
Successive Intervals around the Circle;
Inner Circle: Two-Dimensional Procedure,
Equal Intervals;
Outer Circle: Sorting Procedure,
Empirical Intervals



ture. The interval ranged from 88 to 132 degrees, with its center thus at 110 degrees. Reading from the top of the list down, items at the beginning of the interval carry connotations clearly related to the adjacent interval (AP-Managerial), gradually changing to pick up narcissistic-competitive meanings toward the center of the list, and then tending to pick up the connotations of Critical, which is the central meaning of the next adjacent category, DE.

The gradual shift in meanings continues around the circle. Indeed, as a point half-way around the circle from the initial focus is approached, the meanings are now the opposite of those at the starting point. That is, antonyms are found approximately 180 degrees opposite one another. For example, Item 66 (Acts important) scales at 99 degrees; and at 279 degrees is Item 116 (Spineless). Item 40 (Likes to compete with others) scales at 111 degrees and Item 122 (Agrees with everyone) scales at 290 degrees. Item 101 (Somewhat snobbish) scales at 132 degrees, whereas Item 92 (Likes everyone) scales at 310 degrees. Item 102 (Egotistical and conceited) scales at 141 degrees, and Item 52 (Modest)

scales at 313 degrees. The accuracy of the antonym placements is impressive when the procedures by which the data were collected and the scale values calculated are considered. These results hold for the two-dimensional data as well (with some relatively minor differences for specific items).

If these scaling procedures are used to construct personality assessment devices, it is important to have a relatively stable estimate of the precision of the item placements. In the sorting procedure, if the original frequency distribution of responses is rectangular, then the item could be placed in any sector of the circle and would therefore be quite useless in making inferences about the personality of the individual, since there could be no certainty of the meaning of the term for that individual. The measure of variability appropriate to the data collected by the sorting procedure is the vector length for the item divided by the number of observations ($R = r/N$), as shown in Table 2. Ross (1938) refers to this quantity as an index of variability—the lower the quantity the more variable the original distribution. Indeed, in the circular

Table 11
Statements Located in the Interval 88–132 Degrees
(BC-Competitive) based on the Sorting Procedure

Item	Angle	Statement
100	94	dictatorial
67	95	bossy
68	97	dominating
98	97	expects everyone to admire him
66	99	acts important
97	105	tries to be too successful
65	109	always giving advice
40	111	likes to compete with others
72	121	shrewd and calculating
45	125	resents being bossed
69	125	boastful
74	130	self-seeking
101	133	somewhat snobbish

normal distribution (von Mises distribution) this quantity estimates the variance. In the equation for the circular normal distribution, it is analogous to the reciprocal of the variance in the more widely known standard normal distribution and can thus be used to calculate a standard error for the placement of items scaled by the sorting procedure (Mardia, 1972; Pearson & Hartley, 1972). As a rough frame of reference for interpretation of the vector length values in Table 5, if all the subjects place the item into any one category, the index will equal unity, for the vector length will equal the number of subjects. The more that the subjects disagree as to the placement of the item, the lower the ratio will be. Examination of Table 5 shows that these values range from .29 to .95. From inspection of the original frequency distributions, values below .65 seem too variable in meaning to be useful. Approximately 23% of the items on the ICL exhibited at least this much variability and thus should probably have been considered for rejection in the construction of this checklist.

The variability problem for the two-dimensional data is much more complex. The basic distributions are obtained from linear unidimensional scaling procedures, and calculation of the variance is quite problematical, given the many truncated distributions that occur when the items are rated at the extremes of the scales (Guilford, 1954). Under these conditions even the interquartile range is not very useful; and therefore this report does not present the standard deviations or the interquartile ranges for the values presented in Tables 6 and 7. Burington and May (1970) present some statistical techniques for two-dimensional data that may help resolve some of these problems.

The vector lengths calculated for the two-dimensional data do not reflect variability; rather, they are the results of the intensity ratings of the items along each of the axes. That words and phrases vary in some intensity aspect of their referents is obvious. Indeed, this aspect of various terms has been the most common subject of most scaling studies. However, it is questionable

whether these resultants have any simple relationship to the intensity level of the words and phrases scaled in a unidimensional manner on the two different scales. As noted earlier, the authors of the ICL conceived of each of the items as falling at one of four different intensity levels. To examine the relationship between the resultants and these four intensity levels, the levels have been dichotomized into mild/moderate and strong/extreme groups and the biserial correlation coefficient has been calculated with the resultant vector lengths. The correlation between the two sets of values was $.49 \pm .09$. This correlation is clearly significant; however, it is not very large. Also, the range of vector lengths has been divided by four to form four levels of intensity, which were compared with the four ICL levels using the gamma coefficient; gamma was .51, approximately the same as the biserial coefficient. Clearly, the relationship is only a moderate one. Whether this is due to the methods used to arrive at the ratings during construction of the ICL, some inadequacy of the resultants as a measure of intensity, or a combination of the two, is as yet unclear.

Discussion

The primary purpose of this study was to demonstrate the feasibility of using the Schlosberg (1941, 1942) scaling techniques to scale personality items to a circular scale. Clearly, the results show that the procedures are feasible. For the sorting procedure, with two separate samples of subjects, when the randomly presented categories were ordered according to the "confusions" or overlap in the frequency distributions, the resulting order confirmed that which was postulated for the category labels of the ICL. Moreover, the circular structure obtained for these frequencies clearly approximates the circumplex model developed by Guttman (1954, 1957) for test intercorrelations. In addition, with a third sample, after scaling the items for magnitude along each of the two bipolar dimensions presumed to underly the circu-

lar order of the categories (Dominance-Submission and Love-Hate) and then using the dimensions as orthogonal axes of a Cartesian coordinate system and the average scale values for each item as coordinates, essentially the same circular values for most of the items were obtained as in the sorting procedure. In addition, when the items were combined into the original ICL scales and the scale values obtained in the study were used as weights, the positions of the scales agreed quite closely with those obtained from Rinn's (1965) plot of the factor loadings of the ICL scales used for self-descriptions.

These results, then, are consistent with a large body of empirical research findings that give similar circumplex organizations for social and emotional behavior (Golden, 1969; Schaefer, 1961). However, the thrust of these studies is toward comparison at the scale, rather than the item, level of analysis. Several recent studies have demonstrated that analysis at the item level is very destructive to the composition of the scales of several multifactorial inventories (Eysenck & Eysenck, 1969; Howarth and Browne, 1971; 1972; Sells, Demaree, & Will, 1970). That is, the items are found to shift from the scales to which they had been allotted by the authors of the inventories to others in the set, or else to cluster into entirely new factors than those postulated by those authors. This instability of item placement, it is argued, is due primarily to the subjective, intuitive judgments involved in originally allocating items to the scales of the inventories. This method of constructing scales is, in turn, seen as stemming from the fact that the inventory authors did not have the capacity, now available, for performing satisfactory factor analyses at the item level.

It has been proposed that since that capacity is now available, the instruments used to measure personality will be updated by reconstructing the tests with more comprehensive item-factoring. Rather, it might be more useful to use the procedures of this study to scale the items to the circular continuum and then select those lying within a few degrees of one another (say, within

5 or 10 degrees) to represent the units of the scales. Scaling these inventories to the circular continuum seems reasonable in light of the fact that all three can be represented as circular structures in a two-dimensional Cartesian coordinate system at the level of their relatively similar highest order factors, as pointed out by Eysenck and Eysenck (1969).¹

The authors of the present study recommend item scaling rather than item factoring for several reasons. Probably the most important of these is that the scaling procedures present considerably more information about the meaning of the item than do the item intercorrelations. For one thing, the scaling procedures give a measure of variability of response to the item, which can be used to study item ambiguity and item discrimination; and this measure can also be used as a basis for statistical inference procedures in the construction of tests, as well as in their use in self- or other-descriptions. For another, the scaling procedures give a simple, direct-numerical, and structural representation of the item placements on the circular continuum. This information is not given directly by the item intercorrelations. One possibility might be to transform the intercorrelations to proximity measures; but this would represent a considerable increment in labor over the scaling procedures. At any rate, although some relationship undoubtedly exists between the scaled values and the item intercorrelations, it is probably a complex one and requires thorough study.

Another advantage afforded by the scaling procedures, as compared with intercorrelations, is the possible use of the vector lengths in the two-dimensional scaling as estimates of the degree of intensity of the items. At present most tests, including the ICL, simply sum the items

¹It can be added here that all the recent reviewers of the ubiquitous two-dimensional frame of reference for psychological variables have concluded that the considerable variety of labels for the basic dimensions found in the literature is due primarily to the different authors each stressing, as their basic reference points, axes which are merely minor rotational differences within the same structure.

(unit weighting) to obtain scale scores (Nunnally, 1978). It is conceivable that use of the vector lengths as weights when calculating scale scores might give more useful scale scores when the scales are used as assessment devices. Study of the issue is, of course, required.

Yet another reason for recommending item scaling rather than item factoring is that the former would seem to be considerably more economical, as well as more efficient. A glance at Figures 1 and 2 is sufficient to indicate that if it is desired to construct a test with several items clustering within some relatively narrow sector (say 10 degrees), several hundred items will have to be examined. In an item-factoring study this would require several thousand subjects to assure stable results (Nunnally, 1978). However, the available evidence indicates that reliable scale values can be obtained with a relatively small number of judges (Edwards, 1957).

The similarity of the results obtained with each of the scaling procedures used in this study has been stressed. However, it is worthwhile to also consider some of the differences between the methods. One advantage of the sorting procedure over the two-dimensional procedure is that it is more economical. The sorting procedure requires only one judgment from each subject for each statement. In addition, that judgment would seem to be a relatively simple one for the subjects, requiring only that they decide which of the standard stimuli (the category labels) is most similar to the statement. Thus, this procedure would seem to have some advantages over other commonly used procedures in psychological scaling, such as paired comparisons and multidimensional scaling (Nunnally, 1978), in that many more stimuli can be scaled. The sorting procedure has other advantages over other commonly used unidimensional scaling techniques. For example, random presentation of the standard stimuli eliminates the "error of central tendency," or the bias against giving extreme judgments (Guilford, 1954). The circularity of the distributions obtained also eliminates concern over another undesirable characteristic

of unidimensional scaling—the marked truncation of many distributions, which often makes it impossible to assign values to the ends of the scales or to estimate the variability of the item distributions (Guilford, 1954).

The sorting procedure has, however, one fairly clear disadvantage: Many of the item distributions exhibit marked skewness, which complicates the development of a simple model for the data. However, this characteristic can be moderated by using more category labels. In this study eight 45 degree intervals were used, but it might have been better to use a dozen 15 degree intervals, with 15 degree "empty" sectors between them. This, in turn, requires a careful selection of category labels for those intervals. For example, the use of the label "Conventional" may have led to many of the distributions being more skewed than they would have been had a term such as "Cooperative" been used. One empirical method for resolving this problem would be to use the two-dimensional procedure to map a sample of terms and then to select from that sample 12 terms that are equally spaced around the circle for use as the category labels in the sorting procedure.

The two-dimensional procedure suffers from many of the defects of unidimensional scaling procedures. For one, it requires two judgments from each subject for each statement and is thus less economical than the sorting procedure. For another, each of the scalings is unidimensional and therefore suffers from the many defects of these procedures such as the error of central tendency and of truncated distributions. In fact, given these problems, it seems quite remarkable that there is such high agreement in item placement between the two procedures, as was demonstrated by Schlosberg (1952) and the present study.

However, the two-dimensional procedure has some advantages over the sorting procedure as well. For example, the two-dimensional procedure gives a potential estimate of the intensity level of the terms scaled. Another advantage of the two-dimensional procedure is that it permits

a test of the hypothesis that more than two orthogonal dimensions are required to represent the relationships among the items. Several of the authors who originally postulated a two-dimensional structure for their data later added a third dimension to their system (e.g., Schaefer, 1971; Schlosberg, 1954). However, the issue remains ambiguous. For example, Abelson and Sermat (1962) used a multidimensional scaling procedure to study the structure found by Schlosberg (1954) in the judgments of facial expressions of emotion. They recovered Schlosberg's three dimensions but concluded that two dimensions served as well in replicating the items' scale values. If a third orthogonal dimension is important, many items should scale around the neutral point on both dimensions, implying a third dimension along which these items would vary (Stevens, 1951).

It would also be possible to examine further the validity of these results with a procedure such as that used by Miller (1969). The sorting procedure used in this study is related to the traditional psychophysical method of comparing stimuli to a set of standard stimuli, whereas the sorting procedure used by Miller is more closely related to a paired comparisons procedure. In that Miller required his subjects to merely sort the stimuli into separate piles based on their judgments of similarity between items, it is a much less restrictive sorting procedure than the one used in this study. However, Miller did restrict his subjects by focusing them on one *given* definition for each item. With this restriction removed, the data collected could be analyzed by a multidimensional scaling procedure to examine for the sufficiency of a two-dimensional solution, as Abelson and Sermat (1962) did with Schlosberg's (1954) data. It would also be of interest to examine the data thus collected, as Miller (1969) did his, for any possible hierarchical relation among the items. This type of analysis considers the items as contrasts, rather than as bipolar, continuous dimensions, and an examination comparing these two approaches on the same data could prove illuminating.

The two scaling procedures imply different interpretations of the meaning of the terms in this semantic domain. For the sorting procedure, it is assumed that most terms in a language can be found to have a number of different but more or less related senses; and, further, some of these senses are shared by other terms so that their meanings overlap. The result is, as Deese (1965) has noted, that the meaning of any given term is not given by a single response but rather by the potential distribution of responses to that term. When using a term, any given individual will, in his/her communications, tend to emphasize one of these senses, depending on his/her background and the context of the communication. These shared senses between terms might be considered to be Guttman's (1957) "elementary components," which are basic to the circumplex ordering.

The two-dimensional scaling operations lend themselves to a different interpretation of the meanings of the terms. In this system, it seems natural to consider any given term's primary meaning as made up of varying proportions of two orthogonal features. The two interpretations are not, of course, mutually exclusive. In analogy with the development of secondary color terms, a small set of "focal terms" (Rosch, 1975) or a "minimal contrastive set" (Miller & Johnson-Laird, 1976), which are used to cover a wide range of shades of color (or meaning), can be conceived. For example, "blue" and "green" are often combined and some intermediate shade of color is referred to as bluish green. If the distinction becomes widespread and useful to the speakers of the language, the shade might be given some special term of its own, such as "turquoise." In the same way "loving-dominance" might be distinguished as "nurturance."

Use of these techniques would clearly have some value in helping to overcome some problems that plague the users of the Semantic Differential. As the authors of this technique (Osgood, Suci, & Tannebaum, 1957) have noted, "One of the difficult methodological problems faced . . . unsuccessfully . . . is to demonstrate

that the polar terms we now use are true psychological opposites, i.e., fall at equal distances from the origin . . . and in opposite directions . . ." (p. 327). Since the two-dimensional scaling procedure permits an estimate of distance from the origin as well as of circular placement, the polar terms now in use on Semantic Differential scales could be tested for their psychological opposition in terms of vector lengths as well as angular placement.

The Semantic Differential has been widely used in cross-cultural research (Osgood, 1962). One recurring criticism of this work focuses on the adequacy of the translation of the polar terms used in the scales. This criticism has, of course, been applied to all translations of psychological instruments. These scaling techniques could be of value in mitigating any possible effects of translation. For example, the two-dimensional scaling procedure requires the translation of only the four "focal" terms used to anchor the two dimensions. These dimensions could be used to scale a sample of items from the language's lexicon, and appropriate items could then be selected to form scales. The orthogonality of the dimensions could, of course, be assessed; and the resultant structure could be compared with the English original.

Equivalence of meaning is not just a problem for translation across different languages; as is well known, stimuli may have different meanings for different subjects (e.g., Fiske & Pearson, 1970). Yet most test-scoring procedures are based on the assumption that any given item has comparable meanings for all of the respondents (Scott, 1968). The results of the present study confirm the hypothesis that many of the ICL items did not scale into the octants on the circle intended by the ICL's authors. Furthermore, comparison of Rinn's (1965) factor plot and the scale resultants in this study indicates that subjects tend to use the items to describe themselves in much the same way as indicated by the scaling values.

References

- Abelson, R. P., & Sermat, V. Multidimensional scaling of facial expressions. *Journal of Experimental Psychology*, 1962, 63, 546-554.
- Anderson, W. H. Likeableness ratings of 555 personality-trait words. *Journal of Personality and Social Psychology*, 1967, 9, 272-279.
- Baumrind, D., & Black, A. E. Socialization practices associated with dimensions of competence in pre-school boys and girls. *Child Development*, 1967, 38, 291-328.
- Becker, W., & Krug, R. S. A circumplex model for social behavior in children. *Child Development*, 1964, 35, 371-396.
- Burlington, R. S., & May, Jr., D. C. *Handbook of probability and statistics with tables* (2nd ed.). New York: McGraw-Hill, 1970.
- Buss, H., & Gerjuoy, H. The scaling of items used to describe personality. *Journal of Consulting Psychology*, 1957, 21, 361-369.
- Cole, N. On measuring the vocational interests of women. *Journal of Counseling Psychology*, 1973, 20, 105-112.
- Cooley, W. W. Techniques for considering multiple measurements. In R. L. Thorndike (Ed.), *Educational measurement* (2nd ed.). Washington DC: American Council on Education, 1971.
- Deese, J. *The structure of associations in language and thought*. Baltimore: Johns Hopkins Press, 1965.
- Edwards, A. L. *Techniques of attitude scale construction*. New York: Appleton-Century-Crofts, 1957.
- Eysenck, H. J., & Eysenck, S. G. *Personality structure and measurement*. San Diego CA: Knapp, 1969.
- Fiske, D. W., & Pearson, P. H. Theory and techniques of personality measurement. *Annual Review of Psychology*, 1970, 20, 49-86.
- Freedman, M., Leary, R., Ossorio, A., & Coffey, H. The interpersonal dimension of personality. *Journal of Personality*, 1951, 20, 143-161.
- Golden, P. C. A review of children's reports of parent behavior. *Psychological Bulletin*, 1969, 71, 222-236.
- Guilford, J. P. *Psychometric methods* (2nd ed.) New York: McGraw-Hill, 1954.
- Guttman, L. A new approach to factor analysis: The radex. In P. Lazarsfeld (Ed.), *Mathematical thinking in the social sciences*. Glencoe IL: Free Press, 1954.
- Guttman, L. Empirical verification of the radex structure of mental abilities and personality traits. *Educational and Psychological Measurement*, 1957, 17, 391-407.

- Howarth, E., & Browne, J. A. An item factor-analysis of the 16 PF. *Personality*, 1971, 2, 117-139.
- Howarth, E., & Browne, J. A. An item factor-analysis of the Eysenck Personality Inventory. *British Journal of Social and Clinical Psychology*, 1972, 11, 162-174.
- La Forge, R., & Sucek, R. The interpersonal dimension of personality. III. An interpersonal checklist. *Journal of Personality*, 1955, 24, 94-112.
- Leary, T. *Multilevel measurement of interpersonal behavior*. Berkeley, CA: Psychological Consultation Service, 1956.
- Lorr, M., Klett, C. J., & McNair, D. M. *Syndromes of psychosis*. New York: Macmillan, 1963.
- Lorr, M., & McNair, D. M. An interpersonal behavior circle. *Journal of Abnormal and Social Psychology*, 1963, 67, 68-75.
- Maccoby, E., & Masters, J. C. Attachment and dependency. In P. H. Mussen (Ed.), *Carmichael's manual of child psychology* (Vol. 2, 3rd ed.). New York: John Wiley, 1970.
- Mardia, K. V. *Statistics of directional data*. London: Academic Press, 1972.
- McCormick, C. C. *A study of a procedure for the scaling and calibration of personality test items applicable to the circumplex model*. Unpublished doctoral dissertation, Loyola University of Chicago, 1977.
- Miller, G. A. A psychological method to investigate verbal concepts. *Journal of Mathematical Psychology*, 1969, 6, 169-191.
- Miller, G. A., & Johnson-Laird, P. N. *Language and perception*. Cambridge MA: Belknap Press, 1976.
- Mukherjee, B. N. The factorial structure of Wechsler's Preschool and Primary Scale of Intelligence at successive age levels. *British Journal of Educational Psychology*, 1975, 45, 214-226.
- Nunnally, J. C. *Psychometric theory* (2nd ed.). New York: McGraw-Hill, 1978.
- Osgood, C. E. Studies on the generality of affective meaning systems. *American Psychologist*, 1962, 17, 10-28.
- Osgood, C. E., Suci, G. K., & Tannebaum, R. H. *The measurement of meaning*. Urbana IL: University of Illinois Press, 1957.
- Pearson, E. S., & Hartley, H. O. *Biometrika tables for statisticians*. Cambridge, England: Cambridge University Press, 1972.
- Rinn, J. L. Structures of interpersonal domains. *Psychological Review*, 1965, 72, 445-466.
- Roe, A., & Siegelman, M. A parent-child questionnaire. *Child Development*, 1963, 34, 355-369.
- Rosch, E. Cognitive reference points. *Cognitive psychology*, 1975, 7, 532-547.
- Ross, R. T. A statistic for circular scales. *Journal of Educational Psychology*, 1938, 29, 384-389.
- Schaefer, E. S. A circumplex model for maternal behavior. *Journal of Abnormal and Social Psychology*, 1959, 59, 226-235.
- Schaefer, E. S. Converging conceptual models for maternal behavior and for child behavior. In J. Glidewell (Ed.), *Parental attitudes and child behavior*. Springfield IL: Thomas, 1961.
- Schaefer, E. S. From circular to spherical conceptual models for parent behavior. In J. P. Hill (Ed.), *Minnesota symposium on child psychology* (Vol. 4). Minneapolis: University of Minnesota Press, 1971.
- Schaefer, E. S., & Bayley, N. Maternal behavior, child behavior, and their intercorrelations from infancy to adolescence. *Monographs of the Society for Research in Child Development*, 1963, 28, (3, Serial No. 87).
- Schaefer, E. S., & Plutchik, R. Interrelationships of emotions, traits, and diagnostic constructs. *Psychological Reports*, 1966, 18, 399-410.
- Schlosberg, H. A scale for the judgment of facial expressions. *Journal of Experimental Psychology*, 1941, 29, 497-510.
- Schlosberg, H. The description of facial expression in terms of two dimensions. *Journal of Experimental Psychology*, 1952, 44, 229-237.
- Schlosberg, H. Three dimensions of emotion. *Psychological Review*, 1954, 61, 81-88.
- Scott, W. A. Attitude measurement. In G. Lindzey & E. Aronson (Eds.), *Handbook of social psychology*, Vol. II: *Research methods* (2nd ed.). Reading MA: Addison-Wesley, 1968.
- Sells, S. B., Demaree, R. C., & Will, D. P., Jr. Dimensions of personality: 1. Conjoint factor structure of Guilford and Cattell trait markers. *Multivariate Behavioral Research*, 1970, 5, 391-422.
- Slater, P. E. Parent behavior and the personality of the child. *Journal of Genetic Psychology*, 1962, 101, 53-68.
- Stern, G. *People in context*. New York: John Wiley, 1970.
- Stevens, S. S. Mathematics, measurement and psychophysics. In S. S. Stevens (Ed.), *Handbook of experimental psychology*. New York: John Wiley, 1951.
- Stevens, S. S. The psychophysics of sensory function. In W. A. Rosenblith (Ed.), *Sensory communication*. Cambridge MA: MIT Press, 1961.
- Wiggins, J. S. Personality structure. *Annual Review of Psychology*, 1968, 19, 293-350.
- Woodworth, R. S. *Experimental psychology*. New York: Holt, 1938.

Acknowledgments

This article is based on a dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy at Loyola University of Chicago.

Author's Address

Send requests for reprints or further information to Jack Kavanagh, Chairman, Department of Foundations of Education, Loyola University of Chicago, 820 N. Michigan Ave., Chicago IL 60611.