

Item Format and the Structure of the Personal Orientation Inventory

Wayne F. Velicer, Carlo C. DiClemente, and Donald P. Corriveau
University of Rhode Island

Two versions of the Personal Orientation Inventory were administered to 317 subjects. One version employed the standard two-choice response format. The other version used a six-choice response format. The purpose of this study was (1) to determine if a multiple-response format resulted in improved psychometric properties, (2) to compare the component structure of the two versions, and (3) to compare the empirically derived scales with the theoretically defined scales. The results showed a slight improvement for the multiple-response format, but with poorly defined component patterns. The change in format resulted in a change in component structure. The components derived from both versions did not correspond to the theoretical scales. An analysis indicated that the only well-defined component from either response format could be interpreted as measuring social desirability responding rather than measuring content. A follow-up questionnaire indicated greater subject acceptance of the six-choice version.

Most structured personality inventories employ a two-choice response format. The selection of this response format seems to have been dictated primarily by practical considerations, such as ease of scoring. However, the current widespread use of computerized scoring services reduces the importance of this factor. Several recent instruments, such as the Comrey Personality Scales (Comrey, 1970) and the Tennessee Self-Concept Scale (Fitts,

1965), have employed five- or seven-choice Likert-type response formats. The potential advantages of the multiple-response format include (1) improved reliability, (2) improved validity, (3) greater subject acceptance, and (4) more clearly defined structure.

Of all these issues, the reliability issue has been the most widely researched. The findings, however, are inconsistent. Jahoda, Deutsch, and Cook (1951) and Ferguson (1941) reported increased reliability with increases in the number of response categories. Bendig (1954), Komorita (1963), Peabody (1962), and Matell and Jacoby (1971) reported no improvement with multiple-response formats. Komorita and Graham (1965) found an increase in reliability with multiple-response formats only for scales with relatively homogeneous items. Masters (1974) reported an improvement only for a scale that had a low initial total score variation.

The use of different test instruments and different research methodologies makes direct comparisons difficult. As a potential resolution of these contradictory findings, Velicer and Stevenson (1978) suggested that the change in format may sometimes result in a change in structure. The finding of Komorita and Graham (1965) with respect to "heterogeneous" scales supports this argument. In the only study to control for structure, Lissitz and Green (1975), employing a monte carlo approach, found that reliability increased as the number of response

APPLIED PSYCHOLOGICAL MEASUREMENT
Vol. 8, No. 4, Fall 1984, pp. 409-419
© Copyright 1984 Applied Psychological Measurement Inc.
0146-6216/84/040409-11\$1.80

categories increased from two to five, with no further improvement as the number of response categories was increased beyond five. Thus, structure seems to mediate the effect of format on reliability.

The area of test validity is more difficult to study empirically. With respect to internal validity, it can be argued that a multiple-response format permits the subject to provide more meaningful and finer distinctions than a two-choice format, resulting in scales that are better defined. This would result in higher loadings for items on the components and more of the total variation accounted for. A more efficient and valid sampling of items could be assembled and possibly fewer items used to measure each construct. In one of the few studies to investigate external validity, Govia (1983) reported higher validity coefficients for the multiple-response version compared to the two-choice version of the same set of items when identical scoring was employed.

The issue of subject reaction to the type of response format has not been extensively investigated. Jones (1968) reported that subjects generally preferred a multiple-response form to a two-choice form. The present study also investigated subject preference in regard to response format.

One aspect that has been investigated extensively is the effect of response format on factorial validity. Studies have employed the following personality inventories: (1) Rotter I-E Scale (Joe & John, 1973); (2) Eysenck Personality Inventory (Oswald & Velicer, 1980; Velicer & Stevenson, 1978); (3) Comrey Personality Scales (Comrey & Montag, 1982); and (4) Buss-Durkee Hostility Inventories (Velicer, Govia, Cherico, & Corriveau, in press). The studies all reported higher loadings and a larger amount of the total variance accounted for with the multiple-response format. In addition, some of the studies found differences in the number of components retained and the nature of the components.

The purpose of the present study was to investigate the effects of format change on factorial validity employing the Personal Orientation Inventory (POI; Shostrom, 1963). This inventory differs in a number of ways from the previously studied inventories. The POI is a long (150-item) instru-

ment, employs a forced-choice response format, and involves scales that are theoretically rather than empirically defined. Possibly the most unique aspect of the POI is the content area. The inventory was developed to measure the values and behaviors important in the development of the self-actualizing individual. Two methods of scoring are available: two broad scales involving all the items or 10 subscales. In addition to assessing the effects of a change in response format, the present study provided information about the extent to which the theoretical scales can be reproduced empirically. Previous studies (Lorr & Knapp, 1974; Silverstein & Fisher, 1974) were limited and inconclusive.

The POI has been criticized for a number of flaws: (1) high scale overlap, (2) low reliability for some scales, and (3) lack of control for social desirability (Bloxom, 1972; Coan, 1972). Nevertheless, the POI has proven to be one of the most popular personality assessment instruments available, particularly as an outcome measure employed in clinical and counseling settings; Knapp (1976) provided a summary of POI use. A partial reason for this popularity may stem from the fact that the POI attempts to measure a set of constructs that are relatively unique when compared with more traditional inventories.

This study was designed to serve several purposes: (1) to compare the empirically derived scales with the theoretically derived scales of the POI, (2) to determine if the change from the two-choice to the six-choice format will affect the component structure, and (3) to assess the improvement in psychometric properties as a result of format change. The extensive popularity of the POI makes the limited information on the psychometric properties of the instrument a serious concern. The availability of computer scoring for personality inventories permits researchers to easily employ Likert-type response formats if the advocated advantages can be established.

Method

Two versions of the POI were administered to student volunteers from two large introductory-level psychology classes at two sessions one week apart.

Approximately half the students took Version I in the first session and Version II in the second session, with the order reversed for the remaining subjects. A follow-up questionnaire was administered at the end of the second session and a short debriefing was provided. Students earned extra credit in their course by participating in both sessions. A total of 317 subjects completed both versions and was included in the present study. Approximately 30 additional students were eliminated from the analysis because of incomplete information.

Version I of the POI consisted of the traditional two-choice response format and hereafter is referred to as the binary version. Version II employed a six-choice response format ranging from 1 = "I agree a great deal more with A than with B" to 6 = "I agree a great deal more with B than with A." This format was originally employed by Joe and John (1973). This version hereafter is referred to as the Likert version.

For each version of the inventory, a principal components analysis was performed on the 150×150 matrix of item intercorrelations. The Minimum Average Partial (MAP) correlation method (Velicer, 1976) was used to determine the number of components to extract. A varimax rotation was performed on the component pattern.

A critical issue in this study was the method of determining the number of components to retain. The MAP procedure (Velicer, 1976) possesses a number of advantages. The procedure is based on the concept of partialling out components until the average squared partial correlation reaches a minimum. Components retained by this procedure are clearly "common" components. Zwick and Velicer (1982, 1984) reported, on the basis of two large simulation studies, that MAP is a highly accurate method. In contrast, Silverstein and Fisher (1974) attempted a principal components analysis of the POI based on the eigenvalue greater than unity rule. They found that 54 components would be retained. Using this rule, they concluded that the solution did not appear to make sense, and turned to cluster analysis as an alternative. Lorr and Knapp (1974) employed the scree test (Cattell, 1966; Cattell & Vogelmann, 1977) and retained 15 components in their analysis. However, six or seven

of the retained components were so poorly defined that they were uninterpretable. In the present study, the eigenvalue greater than unity rule would have resulted in 52 and 50 components retained for the two versions, and the scree test provided no clear stopping point for either version. The MAP procedure would be expected to retain only a limited number of components that are well defined.

Results

POI Binary Format

The analysis of the traditional binary format version of the POI resulted in eight components, which together accounted for 22.7% of the variance. Typically, the interitem correlations were very low with an average squared value of .0087. Only items with a loading of .30 or higher on at least one component were considered to be salient for that component.¹

Cross-classification tables were formed in order to compare the results of the components analysis with the theoretically defined subscales. In effect, the varimax rotated matrix of component loadings is replaced by a matrix of 1s and 0s, where 1 was employed if the loading is above the cutoff, that is, salient, and 0 if the loading is below the cutoff. The concept of interpreting salience has been extensively justified by many authors (Cattell, 1949; Cattell, Balcar, Horn, & Nesselroade, 1969; Horn, 1969; Revelle & Rocklin, 1979; Velicer, Peacock, & Jackson, 1982). This approach is particularly appropriate for this problem since salience information is the only information available for the theoretical scales. If the two matrices consisting of salience indicators are labeled A and A*, the cross classification matrix can be expressed as $A'A^*$.

The cross classification matrix comparing the components from the binary format with the 10 theoretically defined subscales is presented in Table 1. Items that occurred on more than one component or were scored on more than one subscale were counted each time. An inspection of the re-

¹A table showing the items that were salient for each of the components under the two formats is available from the first author.

Table 1
Cross Classification of Items* between the Components of the Two-Choice
and Six-Choice Versions and the Theoretical Scales

Theoretical Scale	Two-Choice Component								Six-Choice Component								Possible	
	1	2	3	4	5	6	7	8	1	2	3	4	5	6	7	8		9
Self-Actualizing Value	5	2	1	2	2	1	2	2	4	2	5	2	0	2	3	2	3	26
Existentiality	4	3	5	2	1	7	1	2	2	9	4	6	2	1	2	0	3	32
Feeling Reactivity	7	2	3	1	0	0	0	1	3	1	3	2	2	6	0	0	1	23
Spontaneity	8	1	1	1	0	0	0	1	3	1	3	1	1	2	0	0	1	18
Self-Regard	9	0	0	1	1	1	0	1	10	1	6	0	0	1	0	1	0	16
Self-Acceptance	11	1	1	0	2	2	0	1	12	5	1	2	1	0	0	0	1	26
Nature of Man	1	0	0	0	0	0	6	2	1	0	3	2	0	0	4	0	2	16
Synergy	1	0	0	0	1	0	1	2	0	0	3	0	0	0	2	0	0	9
Acceptance of Aggression	7	1	6	0	0	1	2	4	5	2	0	1	1	8	4	0	0	25
Cap. for Intimate Contact	7	3	2	3	1	0	0	1	3	1	3	6	3	4	1	1	2	28

*Includes all items with loadings > .30 on any component.

Table 2
Cross Classification of Items* between the Components of the Two-Choice Version and
the Silverman and Fisher (1974) Clusters

Cluster	Two-Choice Component								Six-Choice Component										
	None	1	2	3	4	5	6	7	8	None	1	2	3	4	5	6	7	8	9
I. Self-Acceptance	8	19	4	0	2	2	0	1	1	8	15	0	7	4	2	0	0	1	2
II. Existentiality (A)	4	1	1	3	2	1	4	0	2	3	0	6	1	3	1	0	1	0	1
III. Nature of Man	3	3	0	0	1	1	0	4	0	1	0	1	2	0	1	2	0	3	
IV. Existentiality (C)	3	0	0	3	0	0	1	0	0	1	3	4	0	0	0	0	1	0	0
V. Spontaneity	1	5	0	1	0	0	0	0	0	2	3	0	1	1	0	0	0	0	1

*Includes all items with loadings > .30 on any component.

sults suggests a very poor correspondence between the components and the 10 theoretical subscales. The two general scales were not included in the cross classification, since neither corresponded to any of the components. Component 1 seems to be a very general component with at least five items from 7 of the original 10 scales. Forty-seven of the total 150 items have a loading of .30 or higher on this component. The items with the highest loading tend to come from the Self-Acceptance (Sa) and Self-Regard (Sr) scales, permitting a tentative identification of this component as Self-Perception. An alternative interpretation of this component is that it is a stylistic factor, representing the social desirability content of the items. There is no attempt to control for response distortion within the POI. The content of the initial marker items deals with fear of failure and making errors, uncertainty, insecurities, combined with self-confidence and general optimism.

The second component cannot be matched with any of the original POI scales. The content of the initial marker variables suggests a person who views his/her moral values as self-determined but who is also a conformist in those values. Component 3 draws most of its items from the Existentiality (Ex) and Acceptance of Aggression (A) scales. However, these items represent less than 25% of the items involved in these two scales. The content of the items composing this scale deals with a person's acceptance of anger and dishonesty in his/her own self.

Components 4 and 5 cannot be matched with any of the original POI scales. Both are poorly defined. The highest loading for an item on either component was .409. The content of Component 4 deals with isolation and rigid self-control, and the content of Component 5 deals with present orientation and dependence on others.

Component 6 has seven of the nine items in common with the Existentiality scale. The content of the items deals with compulsivity and perfectionism. Component 7 has all six items from the Nature of Man scale and the content deals with the belief that people in general are good, trustworthy, and cooperative. Component 8 cannot be matched with any of the original POI scales. The content is a

vague moralistic, judgmental approach to relationships.

The components derived from the binary format were also compared to the results of the Silverstein and Fisher (1974) cluster analysis. Using a prison population, Silverstein and Fisher found five basic item clusters. The items were cross classified between the five clusters and the eight components (see Table 2). The first cluster (Self-Acceptance) can be matched with the first component. The fifth cluster (Spontaneity) can also be matched with the first component. It should be noted that Clusters 1 and 5 were positively correlated. The third cluster (Nature of Man) can be matched with Component 7. The other two clusters do not permit unambiguous matching. The Lorr and Knapp (1974) study did not present adequate information for comparisons to be made.

POI Likert Format

The analysis of the Likert version of the POI resulted in nine components, which together accounted for 28.6% of the variance. For comparison purposes, if nine components had been extracted for the binary format POI, 24.3% of the variance would have been accounted for. Correlations between items were again very low with an average squared correlation between the items of .0123. Again, an item with a loading of .30 or above on at least one component was considered meaningful. A cross classification of items by the nine components and the 10 original subscales was developed to assist in interpretation (see Table 1).

Once again the general correspondence between the component and the original POI subscales is very poor. The first component includes slightly fewer items than the first component of the binary version but still involves 35 of the 150 items. The component consists largely of items from the Self-Regard and Self-Acceptance scales and is tentatively labeled Self-Esteem. An alternative interpretation of this component as a measure of social desirability is also possible.

Component 2 overlaps most highly with the Existentiality scale. The content of the items describes a person who is diligent, motivated, and will bend

Table 3
Cross Classification of Items* Between the Components of
the Two-Choice Versions and the Six-Choice Version

Two-Choice Component	Six-Choice Component								
	1	2	3	4	5	6	7	8	9
1	30	1	10	3	1	4	0	0	2
2	2	0	0	7	0	1	0	0	0
3	2	3	0	1	0	2	2	0	1
4	1	2	0	1	0	0	0	2	0
5	3	0	2	1	0	0	0	0	3
6	1	8	0	0	1	0	0	0	0
7	0	0	1	0	0	0	3	0	2
8	0	1	1	1	1	0	2	0	0

*Includes all items with loadings > .30 on any two-Choice Component and any Six-Choice Component. Items that did not load > .30 in both forms were deleted.

rules a little to achieve an end. Component 3 cannot be clearly matched with any of the original POI scales. The content includes the concepts of a present orientation and a social orientation. Component 4 does not correspond clearly to any of the POI scales. The content of the items deals with a person who views his/her moral values as self-determined but who is also largely a conformist. Component 5 cannot be matched with any of the original POI scales and is poorly defined. The highest loading for any item was .350. The content is vague and generally deals with social commitments and emotions.

Component 6 draws most of its items from the Acceptance of Aggression and Feeling Reactivity scales. The content of the items deals with both the ability to accept and to express feelings in interpersonal settings. Component 7 cannot be clearly matched with any of the original POI scales. One aspect of this component deals with the nature of man as good and cooperative but now combined with a rigid morality that sees the world in dichotomies. Component 8 consists of only four items that deal with privacy, withdrawal, and independence. Component 9 does not correspond with any of the original POI scales. The content of the items deals with self-interest and distrust of others.

The components derived from the Likert version were also compared to the results of the Silverstein and Fisher (1974) cluster analysis (see Table 2). The first cluster can be matched with the first component. The second cluster can be matched with the second component. The other three clusters do not permit unambiguous matching.

Comparison of the Binary and Likert Versions

The items were cross classified by the eight components of the binary version and the nine components of the Likert version to assist in comparing the two formats (see Table 3). Several clear correspondences emerged. The first components from both versions have 30 items in common. This component can be labeled Self-Esteem. Ten items from the first component represent the majority of the items in the third component of the Likert version. This component involves a present and social orientation. The change to the Likert format could be viewed as resulting in a split of the large general component from the binary format into two more interpretable components.

Component 6 from the binary version corresponds to Component 2 of the Likert version. These

items deal with compulsivity, perfectionism, and a tendency to bend the rules slightly to achieve goals. Component 2 of the binary version is identified with Component 4 of the Likert version. These items are described as dealing with the concept of moral values being self-determined but still largely in conformity with society. A tentative match is also made between Component 7 of both versions. Both contain items dealing with the nature of man as good and trustworthy, but the components also include additional aspects that are not shared. Four components from each version cannot be matched with any corresponding component from the other version.

For the comparison between the binary and Likert versions, an alternative analysis is possible. Since this comparison involves two different sets of items measured on the same subjects, scale scores were formed on both versions and the correlations between the scale scores were calculated. In this case, the scale scores were calculated as the sum of the five items with the highest loading on each component. Items with negative loadings were reflected before summation. Table 4 presents the 8×9 matrix of correlations between the two sets of scales. An inspection of the table supports the match between Component 1 (binary) and Component 1 (Likert), between Component 6 (binary) and Component 2 (Likert), and between Component 2 (binary) and Component 4 (Likert), all of which involve correlations of approximately .60. The next

highest correlations in the mid .30s are not considered to represent a clear match. These results are consistent with the analysis based on the cross-classification table, with the exception of supporting the tentative match between Component 7 of both versions.

Social Desirability Ratings

The interpretation of the first component from both the binary and the multiple-response formats was ambiguous. Should it be interpreted as a content scale or a social desirability scale? Within the context of the MMPI, the first component has been interpreted as a measure of social desirability rather than content (Jackson & Messick, 1961; Reddon, Marceau, & Jackson, 1982). In order to determine the extent that social desirability can account for responding on the first component, Social Desirability Scale Values (SDSV) were determined for each of the 150 items of the POI. A total of 25 subjects rated each item on a six-point scale, from 1 = "A is a *great deal* more desirable than B" to 6 = "B is a *great deal* more desirable than A." The mean rating was employed as the SDSV for the item, and the correlation was calculated between the loading on the first (unrotated) component and the SDSV. For the binary response format the correlation was .76, and for the multiple-response format the correlation was .83. These correlations support the interpretation that the first

Table 4
Correlation between Scale Scores based on the Eight Components of the Two-Choice Version and The Nine Components of the Six-Choice Version

Score	1	2	3	4	5	6	7	8	9
1	.61	.07	-.23	.24	.03	-.36	-.13	-.05	-.17
2	.21	.04	-.10	.59	.08	-.20	-.05	-.22	-.19
3	.12	-.43	.01	.15	-.16	.00	-.26	-.05	-.20
4	.13	.12	-.07	.14	.25	.05	.08	-.33	-.07
5	.16	.07	.10	.04	-.06	.09	.08	.03	.30
6	.15	.58	.00	.03	.23	-.03	.07	-.05	-.03
7	-.10	-.03	.04	-.02	.00	.02	.29	.04	.33
8	.04	.21	-.03	.19	.21	-.05	-.01	-.11	-.09

component from both formats measures primarily the social desirability of the items and, therefore, a content interpretation is not appropriate.

Results of the Follow-up Questionnaire

The follow-up questionnaire consisted of 10 items answered in a five-choice Likert response format (from "Strongly Agree" to "Strongly Disagree"). The three items of greatest interest dealt with preference of response format (see Table 5 for the items, means, and standard deviations). The reactions were mixed with respect to which format made it easier to respond. The subjects slightly preferred six choices and strongly felt that six choices permitted them to accurately describe themselves. In some of the general questions, the subjects described the directions as clear, their responses as honest, and the items as interesting. They also described themselves as glad they participated in the study and interested in finding out more about inventories of this type.

Discussion

The study involved three major questions. The

first was an evaluative question, Is a multiple-response format superior to a two-choice format? Three aspects of this question are involved: reliability, internal validity, and subject acceptance. The second question concerned comparability: Are the two forms of the instrument comparable or does the change in format result in a change in structure? The third focused on the extent to which the theoretical scales of the POI could be supported empirically.

With respect to the first question, the multiple-response format seemed to be slightly superior. Part of the difficulty with this question stems from the heterogeneous nature of the item pool. However, the multiple-response version resulted in more variance accounted for and more clearly defined components. For comparison purposes, consider the first component of the binary version. In the Likert version this general component apparently split into Component 1 and Component 3. However, Component 1 of the Likert version was still better defined than Component 1 of the binary version. The eigenvalues were 10.99 for the binary version and 13.146 for the Likert version. The loadings for the Likert version were generally higher. This occurred despite the fact that the Likert ver-

Table 5
Means and Standard Deviations for Items on the
Followup Questionnaire (N=332)

Item	Mean	Standard Deviation
1. I found the items interesting	2.01	0.73
2. I found it easier to answer with two choices than six choices	2.70	1.47
3. I could describe myself more accurately with two choices than six choices	3.91	1.14
4. I tried to remember how I had answered last week	2.96	1.18
5. I generally answered all the items honestly	1.55	0.68
6. I would like to know more about inventories like this	2.17	0.90
7. I prefer two choices to six choices	3.23	1.38
8. I am glad I participated in this study	2.09	0.79
9. I found many items confusing	3.03	1.16
10. I found the directions to be clear	1.46	0.65

sion included only a subset of the items found in Component 1 of the binary version. Each eigenvalue of the Likert version was higher than the corresponding eigenvalue of the binary version, indicating more reliable components. However, the multiple-response version was only slightly superior to the binary version, which was a marked contrast to the findings of previous studies (Comrey & Montag, 1982; Oswald & Velicer, 1980; Velicer et al., in press; Velicer & Stevenson, 1978).

The results with respect to subject acceptance were more clear cut. The subjects found it easier to respond to two choices, but felt that they could describe themselves more accurately with six choices and preferred the six-choice format. Since the quality of subject responses and the cooperation of the subject are critical aspects in the use of structured personality inventories, these results alone should motivate research in the area of multiple-response formats.

With respect to the comparability of the two versions, it was quite clear that the change in format affected the structure. Component 1 from the binary format apparently split to form Components 1 and 3 of the Likert version. Three other components could be cross identified. Four components from each version could not be matched. This lack of comparability of structure is consistent with previous findings on other instruments (Joe & John, 1973; Oswald & Velicer, 1980; Velicer et al., in press; Velicer & Stevenson, 1978). Given that even slight changes in the wording of an item can affect its allocation to a scale and its performance, it should be hardly surprising that a change in response format has an equally dramatic effect.

The results of this study did not support the present scoring of the POI. For the binary version, support was provided for a single scale combining the Self-Acceptance and Self-Regard scales. However, this scale seemed to be more appropriately characterized as measuring primarily social desirability rather than content. The Nature of Man scale was also supported. The Existentiality scale seemed to be partially represented by two components. No evidence was provided for using either of the two major scales or the remaining subscales. This result is consistent with the previous finding of Silverstein

and Fisher (1974) and Lorr and Knapp (1974). The percentage of variance accounted for was also very low.

The lack of empirical support for the theoretical scoring seems to stem from two problems: (1) the lack of precision in the definition of the original constructs and (2) the allocation of the same items to different scales. This lack of differential validity calls into question the use of the POI as a criterion measure in outcome studies. For the only well-defined component, the first, it seems that the components measure stylistic aspects of responding rather than self-acceptance. Therefore, it is concluded that the employment of this scale as an outcome measure is seriously compromised.

The content area covered by the POI is considered to be both important and uniquely different from traditional personality inventories, and the development of a new, psychometrically sound and theoretically well-defined instrument is urged for this area. The type of procedure recommended by Jackson (1970, 1971) would be both consistent with these goals and is likely to result in a much improved inventory. The popularity of the POI indicates the need for an instrument involving the content covered by the POI. The results of the present study, however, make any research relying primarily on the POI of questionable utility.

References

- Bendig, A. W. (1954). Reliability and the number of rating scale categories. *Journal of Applied Psychology*, 38, 38-40.
- Bloxom, B. (1972). Personal Orientation Inventory [Review]. In O. K. Buros (Ed.), *The seventh mental measurements yearbook* (pp. 290-292). Highland Park NJ: Gryphon Press.
- Cattell, R. B. (1949). A note on factor invariance and the identification of factors. *British Journal of Psychology*, 2, 134-138.
- Cattell, R. B. (1966). The scree test for the number of factors. *Multivariate Behavioral Research*, 1, 245-276.
- Cattell, R. B., Balcar, K. R., Horn, J. L., & Nesselrode, J. R. (1969). Factor matching procedures: An improvement of the *s* index; with tables. *Educational and Psychological Measurement*, 29, 781-792.
- Cattell, R. B., & Vogelmann, S. (1977). A comprehensive trial of the scree and K G criteria for determining

- the number of factors. *Multivariate Behavioral Research*, 12, 289–325.
- Coan, R. W. (1972). Personal Orientation Inventory [Review]. In O. K. Buros (Ed.), *The seventh mental measurements yearbook* (pp. 292–294). Highland Park NJ: Gryphon Press.
- Comrey, A. L. (1970). *Manual for the Comrey Personality Scales*. San Diego: Educational and Industrial Testing Service.
- Comrey, A. L., & Montag, I. (1982). Comparison of factor analytic results with two-choice and seven-choice personality item formats. *Applied Psychological Measurement*, 6, 285–289.
- Ferguson, L. W. (1941). A study of the Likert technique of attitude scale construction. *Journal of Social Psychology*, 13, 51–57.
- Fitts, W. (1965). *Manual: Tennessee Self-Concept Scale*. Nashville TN: Counselor Recordings and Tests.
- Govia, J. M. (1983). *A comparison of multi-dimensional measures of aggression*. Unpublished master's thesis, University of Rhode Island, Kingston.
- Horn, J. L. (1969). On the internal consistency of reliability factors. *Multivariate Behavioral Research*, 4, 115–125.
- Jackson, D. N. (1970). A sequential system for personality scale development. In C. D. Spielberger (Ed.), *Current topics in clinical and community psychology* (Vol. 2, pp. 61–96). New York: Academic Press.
- Jackson, D. N. (1971). The dynamics of structured personality tests: 1971. *Psychological Review*, 78, 229–248.
- Jackson, D. N., & Messick, S. (1961). Acquiescence and desirability as response determinants of the MMPI. *Educational and Psychological Measurement*, 21, 771–790.
- Jahoda, M., Deutsch, M., & Cook, S. W. (Eds.). (1951). *Research methods in social relations*. New York: Dryden Press.
- Joe, V. C., & John, J. C. (1973). Factor structure of the Rotter I-E scale. *Journal of Clinical Psychology*, 29, 66–68.
- Jones, R. R. (1968). Differences in response consistency and subjects' preferences for three personality inventory response formats. *Proceedings of the 76th Annual Convention of the American Psychological Association*, 3, 247–248.
- Knapp, R. R. (1976). *Handbook for the Personal Orientation Inventory*. San Diego: EdITS.
- Komorita, S. S. (1963). Attitude content, intensity and the neutral point on a Likert scale. *Journal of Social Psychology*, 61, 327–334.
- Komorita, S. S., & Graham, W. K. (1965). Number of scale points and the reliability of scales. *Educational and Psychological Measurement*, 4, 987–995.
- Lissitz, R. W., & Green, S. B. (1975). Effect of the number of scale points on reliability; A monte carlo approach. *Journal of Applied Psychology*, 60, 10–13.
- Lorr, M., & Knapp, R. R. (1974). Analysis of a self-actualization scale: The POI. *The Journal of Clinical Psychology*, 355–357.
- Masters, J. R. (1974). The relationship between number of response categories and reliability of Likert-type questionnaires. *Journal of Educational Measurement*, 11, 49–53.
- Matell, M. S., & Jacoby, J. (1971). Is there an optimal number of alternatives for Likert scale items? Study I: Reliability and validity. *Educational and Psychological Measurement*, 31, 657–674.
- Oswald, W. T., & Velicer, W. F. (1980). Item format and the structure of the Eysenck Personality Inventory: A replication. *Journal of Personality Assessment*, 44, 283–288.
- Peabody, D. (1962). Two components in bipolar scales: Direction and extremeness. *Psychological Review*, 69, 65–73.
- Reddon, J. F., Marceau, R., & Jackson, D. N. (1982). An application of singular value decomposition to factor analysis of MMPI items. *Applied Psychological Measurement*, 6, 275–283.
- Revelle, W., & Rocklin, T. (1979). Very simple structure: An alternative procedure for estimating the optimal number of interpretable factors. *Multivariate Behavioral Research*, 14, 403–414.
- Silverstein, A. B., & Fisher, G. (1974). Cluster analysis of Personal Orientation Inventory items in a prison sample. *Multivariate Behavioral Research*, 9, 325–330.
- Shostrom, E. L. (1963). *Personal Orientation Inventory*. San Diego: EdITS.
- Velicer, W. F. (1976). Determining the number of components from the matrix of partial correlations. *Psychometrika*, 41, 321–327.
- Velicer, W. F., Govia, J. M., Cherico, N., & Corribeau, D. P. (in press). Item format and the structure of the Buss-Durkee Hostility Inventory. *Aggressive Behavior*.
- Velicer, W. F., Peacock, A. C., & Jackson, D. N. (1982). A comparison of component and factor patterns: A monte carlo approach. *Multivariate Behavioral Research*, 17, 371–388.
- Velicer, W. F., & Stevenson, J. F. (1978). The relation between item format and the structure of the Eysenck Personality Inventory. *Applied Psychological Measurement*, 2, 293–304.
- Zwick, W. R., & Velicer, W. F. (1982). Factors influencing four rules for determining the number of components to retain. *Multivariate Behavioral Research*, 17, 253–269.
- Zwick, W. R., & Velicer, W. F. (1984). *A comparison of five rules for determining the number of components*

in complex data sets. Unpublished manuscript, University of Rhode Island, Department of Psychology, Kingston.

Acknowledgment

An earlier version of this paper was presented at the Eastern Psychological Association Meeting, Philadelphia, April 1979. The authors thank J. Michael Govia for his assistance in this study.

Author's Addresses

Send requests for reprints or further information to Wayne F. Velicer, Psychology Department, University of Rhode Island, Kingston RI 02881, U.S.A. C. DiClemente is currently at the Texas Research Institute of Mental Sciences, Houston TX, U.S.A. D. Corriveau is at Southeastern Massachusetts University.